https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Extracting Interpretable Representation of Catchment Hydrological Processes with Deep Learning

Che-You Liu[1], Jun-Hao Chen[2*], Sung-Ching Lee[3], Shijie Jiang[3,4], Hsing-Jui Wang[5], Shao-Yiu Hsu[1*]

[1]Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan (ROC)

[2]Department of Computer Science & Information Engineering, National Taiwan University, Taipei, Taiwan (ROC)

[3]Department of Biogeochemical Integration, Max Planck Institute For Biogeochemistry, Jena, Germany

[4]ELLIS Unit Jena, Jena, Germany

[5]Graduate Institute of Environmental Engineering, National Taiwan University, Taipei, Taiwan (ROC)

Correspondence to: Shao-Yiu Hsu (syhsu@ntu.edu.tw) and Jun-Hao Chen (chenjunhao15@ntnu.edu.tw)

Abstract. Traditional approaches in hydrological modeling typically rely on parameter calibration within prior empirical functional forms and an integrated subsurface bucket. While effective, these fixed mathematical representations can sometimes limit flexibility in capturing complex and localized behaviors. To complement these approaches, we proposed a temporal difference loss regularization (TDLR) framework. Using meteorological and streamflow observations, this framework enables a Long Short-Term Memory (LSTM) network to extract interpretable functional representations of distinct hydrological processes in parallel within a mass-conserving, conceptual bucket architecture. Applications of TDLR across three hillslope catchments (Shihmen, Deji, and Jiji) yielded stable re-extracted function representations under the given structural constraints, offering strong internal consistency between extracted representations of processes. For the integrated subsurface bucket, the LSTM extracted patterns consistent with the transition from saturation-excess to infiltration-excess runoff. Furthermore, results from the Shihmen catchment captured localized variations that plausibly suggest preferential flow dynamics during heavy rainfall—patterns that are often challenging to represent using fixed functional forms. The framework also presents daily variations consistent with transitions from water- to energy-limited regimes in evapotranspiration. Canopy interception exhibited the expected asymptotic plateau, showing high sensitivity to precipitation and its covariance with wind speed. Additionally, the approach captured dynamic routing relationships that align with local morphological and meteorological characteristics. In our study area, the reconstruction accuracy of the extracted functional representations compares favorably with or exceeds both baseline hybrid models and pure black-box LSTMs. By demonstrating that catchment-scale behaviors can be effectively represented as emergent combinations of stable, data-driven relationships, this study provides a complementary perspective for exploring functional representations to advance hydrological modeling.

1 Introduction

Hydrological processes are fundamental components for describing how water is transported and stored in catchments (Beven, 2006b; Gnann and Wagener, 2026). To infer hydrological phenomena quantitatively, we define these 'functional

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

representations' as constitutive mappings that perceptually relate meteorological forcing sequences to moisture fluxes. Combining the functional representations with the water balance constraints is a long-standing, practical approach for simulating complex water movement phenomena in catchments (Clark et al., 2017; Clark et al., 2015). Therefore, functional representation identification serves as the cornerstone of catchment hydrology in hydrological modeling for water resources research and management (Beven, 2006b).

Traditionally, researchers have pursued generalized functional forms for each process across varied catchments, often without explicit verification of their validity across different environments. Consequently, persistent simulation errors leave the verification of universally applicable and fixed process functions unconfirmed. Recently, deep learning has emerged as a general function approximator capable of drastically reducing simulation errors across diverse catchments (Kratzert et al., 2019). This superior capacity suggests the existence of a complex but learnable overall catchment-scale response, achieved by carefully selecting input features for deep learning. However, the hydrologist cannot determine the exact constitutive relationship for each hydrological process by inspecting the deep learning model itself, even with the aid of Explainable AI (XAI) tools (Dwivedi et al., 2023; Mosca et al., 2022). It indicates that, when mapping meteorological forcing to streamflow, deep learning is an efficient surrogate for the overall functional relationship between the two, but lacks interpretability.

Markedly, the interpretability of pure data-driven deep learning models is limited because they lack linkage to prior knowledge in hydrology and field-scale experimental observations (Jiang et al., 2020). To reconcile physical interpretability with simulation flexibility, hybrid modeling that advances parameter calibration has emerged, integrating process-based prior functional forms with deep learning architectures. For instance, the long short-term memory (LSTM) networks embedded with hydrological dynamics can achieve accuracy comparable to purely data-driven models while preserving partial process interpretability (Acuña Espinoza et al., 2024; Feng et al., 2022). Such hybrid frameworks explicitly account for the seasonality of dynamic parameters, combining the explanatory strengths of conventional modeling with the adaptability of deep learning. Another hybrid approach directly encodes the physical relationship into the deep learning architecture (Hoedt et al., 2021; Sahu et al., 2026; Wang et al., 2025) or adds penalty terms to the loss function to limit the deep learning model's behavior. Both hybrid approaches show comparable accuracy to pure data-driven models. Hence, hybrid approaches with advanced parameterization successfully balance the deep learning capacity and maintain interpretability, which relies on the prior functional forms.

However, even with the advances in deep learning in parameterization, there is a lack of clarity regarding the identification of catchment-scale prior functional forms for each hydrological process (Beven, 2006b; Blöschl and Sivapalan, 2006; Peters-Lidard et al., 2017; Kirchner, 2006). The descriptions of hydrological processes, including snow melting, infiltration, evapotranspiration, riverine routing, canopy interception, and baseflow estimation, are based on several strong assumptions. In addition, the closed boundary assumption might fail due to interbasin groundwater flow (Frisbee et al., 2016; Han et al., 2024; Le Mesnil et al., 2020; Peng et al., 2014); the functional forms solved by partial differential equations are not well-suited for the catchment scale with varied initial and boundary conditions, which are weakly constrained in catchment-scale observation. (Beven, 2006b; McDonnell et al., 2007; Peters-Lidard et al., 2017). Therefore, even if the hybrid models achieve

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

competitive accuracy with the prior equations by improving parameter calibration, the representations of the prior functional forms remain uncertain, and the calibrated parameter sets for those functional forms are equifinal (Beven, 2006a).

To be more specific, the simulation residual of the closure problem with conventional models can be decomposed into four layers (Gharari et al., 2021; Gupta et al., 2012). (i) moisture flux and status dynamic equation based on the water balance, (ii) the generic buckets design and the related process interaction, (iii) the prior functional form of the given process, and (iv) the parameter calibration of the given functional forms. During the parameter calibration process, misrepresentation of prior functional forms is not only an unavoidable source of structural error in hydrological modeling but also creates additional residual, which might amplify the equifinality phenomenon. Hence, to address the closure problem, searching for the functional representations directly from data is a potential solution to identify a stable representation and prevent additional equifinality due to possible misrepresentation of prior functional forms. Nevertheless, there was no way to search for an interpretable, stable functional representation.

To directly identify interpretable functional representations from data, we propose temporal difference loss regularization (TDLR) to train the LSTM for data-driven extraction of functional representations. The trained LSTM by TDLR is a constrained function approximator that approaches the hydrological process function in the catchment scale directly from meteorological forcing series, streamflow, and the designed catchment's generic buckets. By embedding a generic bucket structure and enforcing mass conservation in the loss function, we constrain the search space to physically plausible solutions, thereby addressing the challenge of directly identifying stable, interpretable representations from data. The proposed TDLR enables key processes, such as canopy interception, infiltration, routing, and subsurface and groundwater flow, to emerge as time-invariant and interpretable functional representations of the data. These representations provide a structured and physically constrained way to explore dominant processes–response relationships for addressing the closure problem in hydrological modeling. It supports data-driven induction, maintaining the physical meaning and interpretability to advance the understanding of the complex hydrological system at the catchment scale.

2 Study area

2.1 Catchment Characteristics: Climatology and Morphology

The hillslope region exhibits a highly complex, non-linear, and threshold-based process-response function (Bauser et al., 2022; Singh et al., 2021; Zehe and Sivapalan, 2009). The fix of prior functional forms is suspected of losing the representation in the hillslope. To demonstrate that extracted functional representations at the catchment scale may diverge from prior functional forms, we selected three catchments in Taiwan as case studies: Shihmen (764 km^2), Deji (600 km^2), and Jiji (1,718 km^2). Located in a subtropical climate zone, these sites exhibit pronounced dry and wet seasons driven primarily by the Mei-Yu monsoon and typhoons from April to October. Mean annual precipitation for Shihmen, Deji, and Jiji is 2,573 mm, 2,299 mm, and 1,825 mm, respectively. Notably, local convective storms frequently occur in the mountainous regions of these catchments. Because these events are characterized by short duration, localized spatial extent, and extreme intensity, the spatial

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

heterogeneity of precipitation is expected to play a critical role in accurately simulating catchment-scale hydrological responses (Andrea Rinaldo, 1996; Loritz et al., 2021; Wang et al., 2023).

Furthermore, these catchments exhibit dramatic topographic gradients, characterized by minimal transitional zones between steep mountainous hillslopes and near-planar areas (Fig. 1(a)). All catchments have a clear hillslope-plain transition. This highly heterogeneous morphology interacts closely with precipitation spatial distribution, underscoring the potential for functional deformation when applying prior equations to catchment-scale models. Subsurface characteristics exert an equally significant control. In Taiwan, geological faults are predominantly oriented northeast-southwest. The Shihmen catchment is unique among the three study sites not only because its primary stream network aligns parallel to these fault lines, but also because its contributing area is similar across the hillslope and plain regions. The first alignment suggests that subsurface stormflow occurrence in Shihmen introduces a strong potential for preferential fracture flow (Weiler et al., 2005). The second one indicates that, due to similar area contributions from hillslopes and plains, the Shihmen catchment may exhibit a distinct surface routing pattern, as Saint-Venant's assumption is violated. In summary, these morphological characteristics are likely to affect catchment-scale total infiltration, subsurface routing, groundwater dynamics, and baseflow responses (Türk et al., 2026; Weiler et al., 2005). Additionally, the long-term groundwater depletion driven by agricultural pumping is expected to be evident, particularly in the Jiji catchment, reflecting broader regional trends in Central Taiwan (Chen et al., 2024a).

2.2 Hydrometeorological Data, the Heterogeneity and Covariance of Driving Forces

Meteorological driving forces, including precipitation, average wind speed, average, maximum, and minimum temperatures, net radiation, relative humidity, and air pressure, are provided at a daily 1km^2 grid resolution. Daily streamflow data is collected at the respective reservoir gauges and the dam. Because station-measured streamflow represents an integrated catchment response, in Sect. 3.2, we apply a spatial averaging operator to upscale the gridded meteorological forcing data as input to the LSTM. To quantify this spatial heterogeneity and its covariance with daily average precipitation, we use the spatial concentration index (CI) (Martin-Vide, 2004) and spatial moments (Zoccatelli et al., 2011), as illustrated in Fig. 1(b) and 1(c). The formulations for these metrics are defined as:

$$CI = 1 - 2 \times \frac{A_{50\%\,water}}{A} \tag{1}$$

$$\delta_1(t) = \frac{A^{-1}\int_{\Lambda} r(x,y,t)\times d(x,y)da}{P(t)\times A^{-1}\int_{\Lambda} d(x,y)da} \tag{2}$$

$$\delta_2(t) = \frac{A^{-1}\int_{\Lambda} r(x,y,t)\times[d(x,y)-\delta_1(t)A^{-1}\int_{\Lambda} d(x,y)da]^2da}{P(t)\times A^{-1}\times(A^{-1}\int_{\Lambda}[d(x,y)-A^{-1}\int_{\Lambda} d(x,y)da]^2da} \tag{3}$$

where $A_{50\%\,water}$ denotes the raining area contributing 50% of the total daily cumulative precipitation, A and Λ are the area and specific region of the entire catchment, respectively, $r(x,y,t)$ is rainfall area density, and $P(t)$ represents the daily cumulative precipitation. $d(x,y)$ denotes the distance from outlet to each grid. As the CI approaches 1, precipitation becomes highly concentrated, whereas a value closer to 0 indicates a uniform distribution. The variable δ_1 represents the first-order

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

spatial moment. A value of $\delta_1 = 1$ indicates that the precipitation center of mass aligns with the geometric centroid of the catchment. Values greater than 1 indicate the storm center is located toward the headwaters, while values approaching 0 indicate it is concentrated near the outlet. The second-order spatial moment, δ_2, describes precipitation dispersion. A value of $\delta_2 = 1$ implies the dispersion of the rainfall field mirrors the inherent dispersion of the catchment shape. In contrast, values approaching 0 indicate the rainfall is more highly dispersed than the catchment geometry itself.

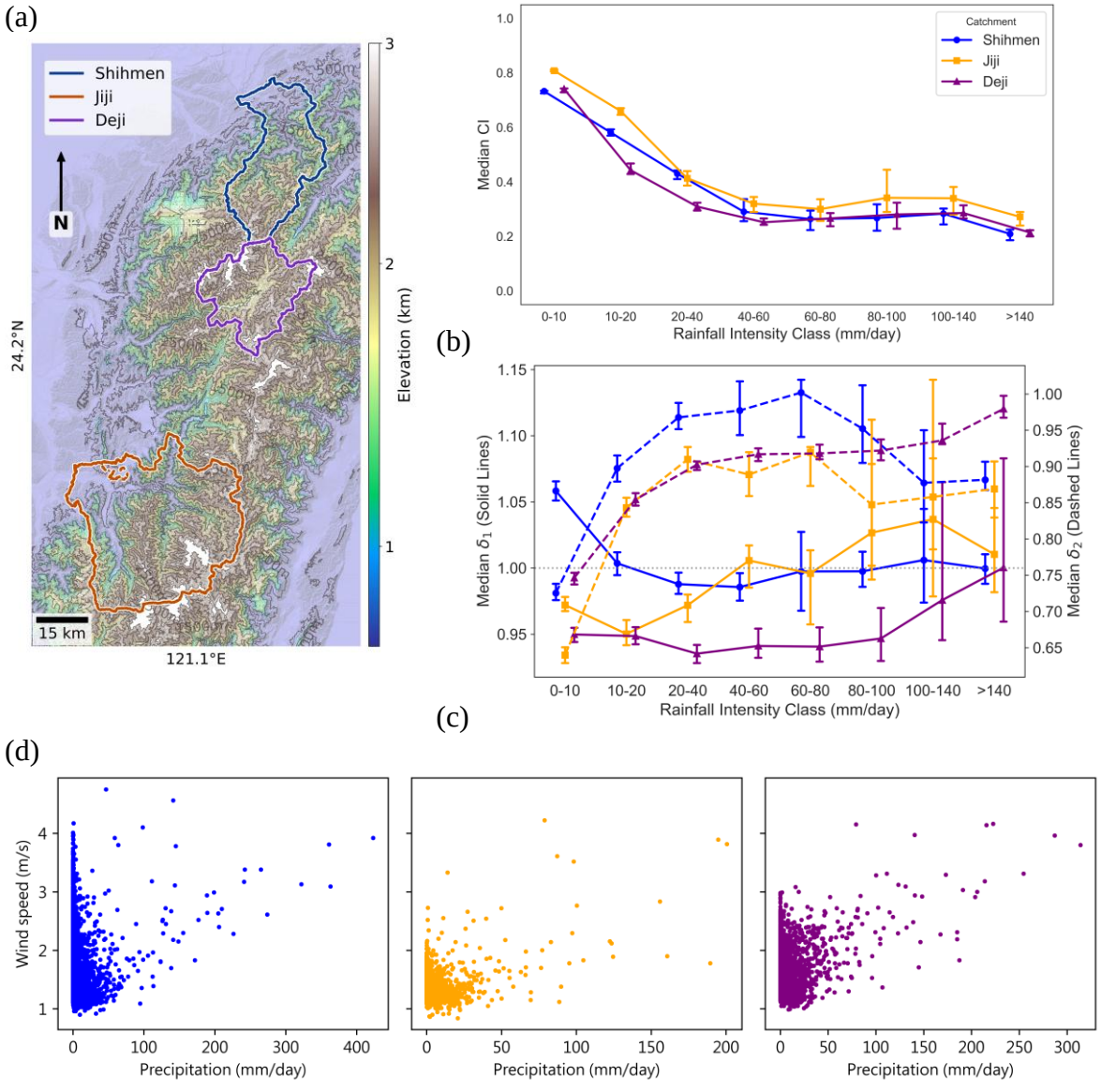


Figure 1. (a) represents the geography information for three catchments. (b) represent the relationship median CI index of the rainfall intensity class change alongside the catchment-scale average precipitation. (c) represents the median spatial moments (δ_1 and δ_2) of the rainfall intensity class. (d) presents the covariance between precipitation and wind speed in three catchments, by using the meteorological record on a daily scale. Blue, orange, and purple denote the cases in Shihmen reservoir, Jiji dam, and Deji reservoir, respectively.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

In Fig. 1(b), the CI patterns reveal a precipitation threshold of approximately 20~40 mm. Below this threshold, spatial non-uniformity is not only highly significant but also concentrated in mountain areas; above it, the rainfall distribution becomes more uniform. Thus, the rainfall area expanded along with the catchment precipitation, from the mountain to the dry areas. As shown in Fig. 1(c), the consistency between the daily average precipitation and the changes in spatial heterogeneity, as well as the first- and second-order spatial moments, indicates the location of the rainfall distribution's center and its dispersion. δ_1 reaches the stable center distance as precipitation increases, in both Shihmen and Jiji cases. However, the storm center is closer to the headwater as the precipitation increases. Hence, the Deji might exhibit a different travel-time process response. Moreover, the dispersion pattern (δ_2), Jiji and Shihmen shows that dispersion decreases with increasing precipitation, but after 80mm, dispersion increases, especially in Shihmen. Differences in rainfall spatial dispersion might lead to differences in infiltration, subsurface, and groundwater responses between Jiji and Shihmen. Hence, the interactions between the spatial heterogeneity and precipitation might play a role in catchment-scale responses, and we show and discuss how these relationships could be captured through the interpretation of the functional representations learned by the proposed framework in Sect. 4.2.

The correlation between daily precipitation and wind speed is robust across all three sites; their specific dynamics vary. The Jiji and Deji catchments demonstrate a clear monotonic relationship, indicating that elevated precipitation is consistently associated with higher wind speeds. Conversely, the Shihmen catchment exhibits a non-linear, "hockey-stick" pattern, as shown in Fig. 1(d), characterized by strong winds even during low-magnitude precipitation events. Given this anomalous wind-precipitation coupling, canopy interception and actual evapotranspiration dynamics in the Shihmen catchment are expected to diverge significantly from those at the other two sites, further complicating the extraction of generalized catchment-scale functional representations.

3 Method

Since the functional representations have far more degrees of freedom than the available constraints, the closure problem in catchment hydrology–that is, the problem of mapping meteorological driving force sequences into the catchment-scale moisture fluxes- is decomposed into two parts: specifying constraints based on prior assumptions for each catchment, and identifying the functional representation that fits the data. We first present the constraints and conceptualize interactions between each functional representation within given conceptual buckets in Sect. 3.1. Within the given buckets, we proposed temporal difference loss regularization (TDLR), using an LSTM as the function approximator to identify the functional representation in Sect. 3.2. The proposed TDLR serves as a new paradigm that searches for functional representations of each process under constraints, thereby narrowing the feasible solution space for the closure problem in hydrological modeling without relying on prior functional forms while retaining interpretability. The benchmark paradigms, pure-data driven and hybrid models with advanced parameterization, are presented in Sect. 3.3

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

3.1 Process functions under given buckets: overview and definition

Interpretation of hydrological processes often relies on perceptual concepts and field-scale observation. In contrast to the XAI approach for opening the black box, hydrological modeling relies on a generic bucket structure as a foundation for interpretability, bridging process understanding with catchment-scale observations.

In this section, we discuss the consideration of buckets in Taiwan, a subtropical climate zone, where canopy interception, transpiration, subsurface and groundwater, and explicit riverine routing should be accounted for as changes in the control volume for the whole catchment. This implies that precipitation can be stored in three buckets: canopy, surface, and integrated subsurface layers. Precipitation, evapotranspiration, streamflow, and interbasin groundwater flow are fluxes in the catchment control volume boundary. Infiltration, baseflow, and direct runoff are the fluxes inside the control volume. In Fig. 2, we describe the interaction between the unsolved functional representations, well-known driving forces, and streamflow. From precipitation to streamflow, it first passes through the canopy layer and then infiltrates into the integrated subsurface bucket. A portion of the precipitation after interception is routed to the catchment outlet. Another portion of the effective precipitation would enter the integrated subsurface bucket, the source of baseflow, through infiltration.

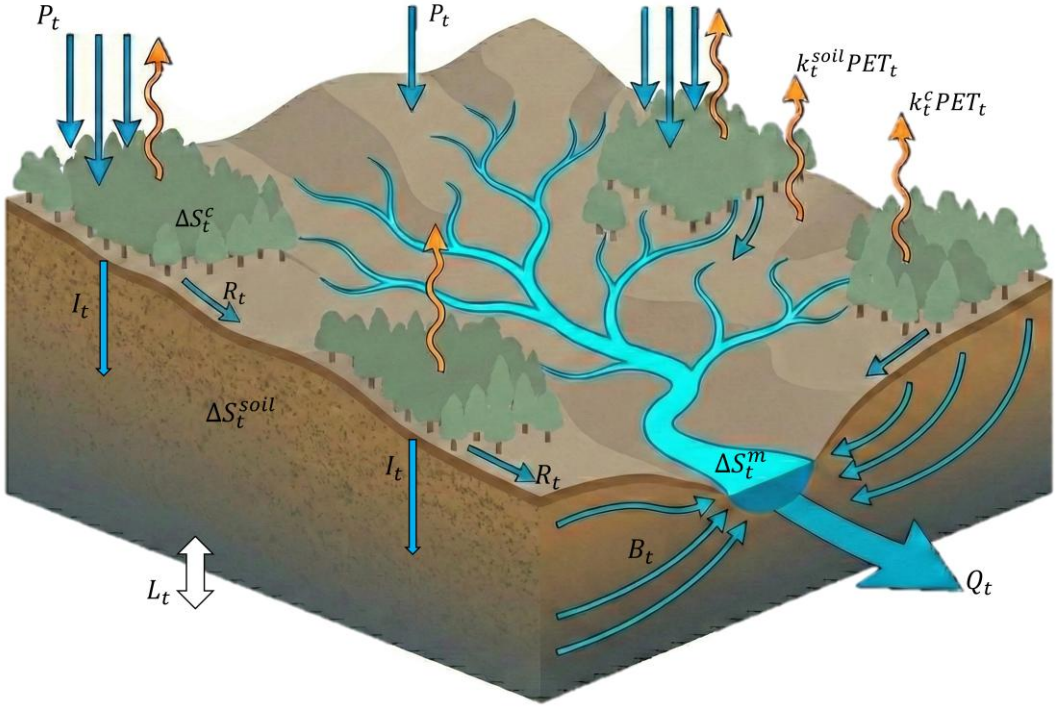


Figure 2. Conceptual structure and related hydrological functional representations of generic buckets. Three main buckets, canopy storage, connected water storage in integrated subsurface layers, and routing storage, denote ΔS_t^c, ΔS_t^{soil}, and ΔS_t^m. The moisture flux process, precipitation, infiltration, direct runoff, baseflow, fraction of evaporation from soil, fraction of transpiration from canopy, and streamflow,

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

denote P_t, I_t, R_t, B_t, k_t^{soil}, k_t^c, and Q_t. L_t denote the residual of the mass conservation constraint. The subscript t indicates the quantities at time t.

Equation (4) provides the catchment-scale moisture fluxes and the dynamics of the connected water storage serving as the generic buckets design, as well as the mass conservation constraint:

$$\Delta S_t^c + \Delta S_t^{soil} + \Delta S_t^m + L_t = P_t - \left(k_t^c + k_t^{soil}\right) PET_t - (R_t + B_t) \quad (4)$$

where ΔS_t^c, ΔS_t^{soil}, ΔS_t^m, and L_t, denote the change of canopy water storage, change of discharge connected water stored in integrated subsurface layers, the storage change of the river routing, and the residual of the mass conservation constraint, respectively. P_t denotes precipitation. k_t^c and k_t^{soil} denote the fractions of evapotranspiration in the buckets of canopy and soil, respectively. R_t, PET_t, and B_t denote the direct runoff, potential evapotranspiration, and baseflow. The subscript, t, represents the process and quantities change at time t on a daily scale. The sum of R_t and B_t is the streamflow, as a well-recognized constraint. In this study, the unsolved functional representations refer to the connected storage changes (ΔS_t^c, ΔS_t^{soil}, and ΔS_t^m), the moisture flux control (k_t^c and k_t^{soil}), direct runoff, and baseflow.

L_t is the residual from mass conservation constraints, not only reflecting the water exchange with the neighboring basins, which might occur through groundwater and subsurface pathways, but also residuals arising from measurement errors and structural mismatches, as shown in Eq. (5):

$$L_t = IGF_t + \epsilon_t \quad (5)$$

where IGF_t and ϵ_t denote the inter basin groundwater flow (Elrashedy et al., 2025; Le Mesnil et al., 2020) and the residual arising from measurement errors and structural mismatches, respectively. If the unresolved functional representations are identified, L_t could be a function of meteorological driving forces, analogous to other functional representations learned by the LSTM. Given that inter-basin groundwater exchange can occur at daily time scales, the correlation between L_t and the driving force could be inferred from the data. A positive IGF_t indicates net lateral groundwater outflow to downstream basins or percolation into deeper aquifers. In contrast, a negative IGF_t suggests groundwater recharge from adjacent basins or deeper subsurface layers to the integrated subsurface bucket.

To simulate the dynamics of the processes and flux in Eq. (4), we first present the process relationship of the baseflow, infiltration, and the dynamics of the integrated subsurface bucket, shown in Eq. (6):

$$B_t = -\Delta S_t^{soil} + (P_t - \Delta S_t^c - k_t^c PET_t)\left(1 - F_{pt}\right) - k_t^{soil} PET_t \quad (6)$$

F_{pt} represents the ratio of water on the soil excluding the infiltration at the current time, t, and $\left(1 - F_{pt}\right)$ denotes the ratio of the water infiltrating into the soil. $(P_t - \Delta S_t^c - k_t^c PET_t)$ and B_t should be the positive value, as the extra penalty in loss. F_{pt}, k_t^c, and k_t^{soil} are the partitioning fractions ranging from zero to one. Notably, the memory effects and threshold behavior in moisture flux modeling are conventionally captured either through prior functions of antecedent conditions or through water-status variables in generic storage buckets. Here, we depart from this practice: F_{pt}, k_t^c, k_t^{soil}, ΔS_t^{soil}, and ΔS_t^c are directly searched without explicitly using water storage states as inputs. Sequences of meteorological driving serve as inputs to the

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

partitioning functions, e.g., F_p, and storage change dynamics (ΔS_t^{soil}). We argued that the expanded input domain implicitly encodes information about the antecedent conditions.

For direct runoff, R_t, effective precipitation would not equal to instantaneous measured direct runoff due to travel time effects, which are represented by riverine routing storage and manifested as the peak delay phenomenon. In this study, we used the Muskingum form to describe the surface riverine routing water storage. Using Eq. (7), the change of the connected water storage of the riverine routing is solved based on the effective travel time (K_t) and wedge storage parameter (x_t). The travel time and the wedge parameter should be highly correlated with the current and previous time-varying direct runoff, both of which are time-variant (Guang-Te and Singh, 1992; Yadav et al., 2015).

$$S_t^m = K_t[(1 - x_t)F_{pt}(P_t - \Delta S_t^c - k_t^c PET_t) + x_t R_t] \quad (7)$$

The current water entering the riverine routing process equals the water excluded from infiltration, canopy interception, and canopy evapotranspiration. With this riverine routing storage assumption and the mass conservation of the surface water, the direct runoff R_t could be calculated with K_t and x_t, as shown in Equations (8) to (11):

$$R_t = C_{0t}F_{pt}(P_t - \Delta S_t^c - k_t^c PET_t) + C_{1t}F_{pt-1}(P_{t-1} - \Delta S_{t-1}^c - k_{t-1}^c PET_{t-1}) + C_{2t}(Q_{t-1} - B_{t-1}) \quad (8)$$

C_{0t}, C_{1t}, and C_{2t} denote time-variant routing coefficients of the Muskingum method (Guang-Te and Singh, 1992; Ponce et al., 1996; Yadav et al., 2015).

$$C_{0t} = \frac{-K_t x_t + 0.5\Delta t}{K_t(1-x_t) + 0.5\Delta t} \quad (9)$$

$$C_{1t} = \frac{K_t x_t + 0.5\Delta t}{K_t(1-x_t) + 0.5\Delta t} \quad (10)$$

$$C_{2t} = \frac{K_t(1-x_t) - 0.5\Delta t}{K_t(1-x_t) + 0.5\Delta t} \quad (11)$$

Δt in our study area is set to be 86400 sec. Additionally, the range of K_t is a positive number, and the range of x_t is from -1 to 0.5. Notably, the negative x_t could be observed in the natural riverine system (Szilagyi, 1992), when the timing of the outflow peak is inside the inflow hydrograph.

From Equations (5) to (11), we describe the domain and range of constitutive relationships between the meteorological driving force sequence and the specific hydrological process response, both of which are anchored in the conceptual buckets of the catchment and in the mass conservation constraint. The extracted functional representations are therefore interpretable because their identification is heavily grounded in prior knowledge of catchment hydrology. The proposed formulation differs fundamentally from conventional state-dependent process-based equations. Rather than carrying memory in the system state, the process response functions learned from the proposed framework, which encodes information directly in the meteorological forcing sequence, thereby each hydrological process can be expressed as a functional representation of that sequence. Consequently, the standard calibration paradigm, in which the state serves as a memory of the sequence-equivariant hydrological response, does not transfer directly to the problem of identifying a time-invariant function. Moreover, calibration must run recurrently over the whole data, decreasing the computational efficiency.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

We formulate the overall interactions between the unsolved functional representations, the force sequences, and stream response, showing that these functional representations are the functions of sequences of driving forces. The formulation rests on two assumptions:

1. A sufficiently long forcing sequence carries enough information for each functional representation.
2. Over the period of record, the functional representations remain time-invariant.

The first assumption indicates that the initial water status is not sensitive to the streamflow construction if the simulation time is long enough, echoing the concept of a warm-up period in conventional hydrological modeling. The second one is the key axiom for constructing the optimization hypersurface in deep learning using gradient descent (Ruder, 2016). It is adopted here as a practical assumption to facilitate model training under approximately stationary conditions. With these two assumptions, the processes, including infiltration (F_{pt}), surface routing (K_t, and x_t), the change of canopy storage (ΔS_t^c), dynamics of the subsurface and groundwater layers (ΔS_t^{soil}), and evapotranspiration (k_t^c and k_t^{soil}), are well represented by meteorological forcing sequences that contain the complete information to describe the antecedent moisture conditions in all buckets. With this formulation, the proposed LSTM training framework can address the closure problem while satisfying mass balance constraints and the prescribed generic bucket design by searching the functional representations.

3.2 Temporal Difference Loss Regularization

The temporal difference loss regularization (TDLR) focuses not only on the temporal difference of the output but also on bounding the output range of the functional representation to remain consistent with the prescribed generic buckets. The workflow for single-LSTM training for functional representation is shown in Fig. 3(c), and the generic buckets with the functional representation rule are shown in Eq. (12).

$$Q_t = M(X_t, X_{t-1}, Q_{t-1}^* | \omega_t, \omega_{t-1}) \quad (12)$$

where Q_t, Q_{t-1}^*, and M denote simulation streamflow, previous time step streamflow, and the structure of generic buckets holding the temporal difference of the streamflow, respectively. X_t is the force series which is upscaled from grid data with an averaging operator. ω_t is the functional representation for each hydrological process learned by single LSTM, including, k_t^c, k_t^{soil}, F_{pt}, ΔS_t^c, ΔS_t^{soil}, K_t, and x_t; the input of the functional representations are past and current meteorological driving force from t-64 days to t. Notably, during the training phase, Q_{t-1}^* is from observed data. But in a streamflow reconstruction task, it could be the simulated streamflow at the previous time step. In Fig. 3(b), a single LSTM model learns the functional representations, but the current loss propagation for the LSTM model includes the previous hydrological response simulated from the same LSTM. This kind of regulatory design for loss propagation encodes the LSTM's focus on the constitutive relationship between the sequence of meteorological driving forces and the dynamics of hydrological responses.

Equation (13) proposed the target LSTM that the model wants to learn for the loss function and data flow design. ω_t is only a function of the meteorological driving force, as shown from Equations (13) to (15).

$$\omega_t = \phi(X_t) \quad (13)$$

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

$$\phi^*(X_t|\theta) = argmin_\theta |Q_t^* - M(X_t, X_{t-1}, Q_{t-1}^*|\omega_t, \omega_{t-1}) + \delta_t|\omega_t| \quad (14)$$

$$\delta_t|\omega_t = a_1 RELU(-B_t|\omega_t) + a_2 RELU(-R_{t-1}|\omega_{t-1}) + a_3 RELU(-(P_t - \Delta S_t^c - k_t^c PET_t)|\omega_t) \quad (15)$$

where $\phi^*(X_t|\theta)$ denotes the LSTM model. $\delta_t|\omega_t$ is the penalty loss when baseflow, direct runoff, and effective precipitation take negative values because the water-balance constraints alone do not enforce the lower bound of zero on direct runoff and baseflow. a_1, a_2, and a_3 control the strength of the penalty and are all set to 1.5. This study aims to extract the functional relationships between the process responses and the meteorological forcing sequence. Notably, the extracted functions, ω_t, for each hydrological process, are interpretable owing to the definition of the functional representation and the generic bucket design.

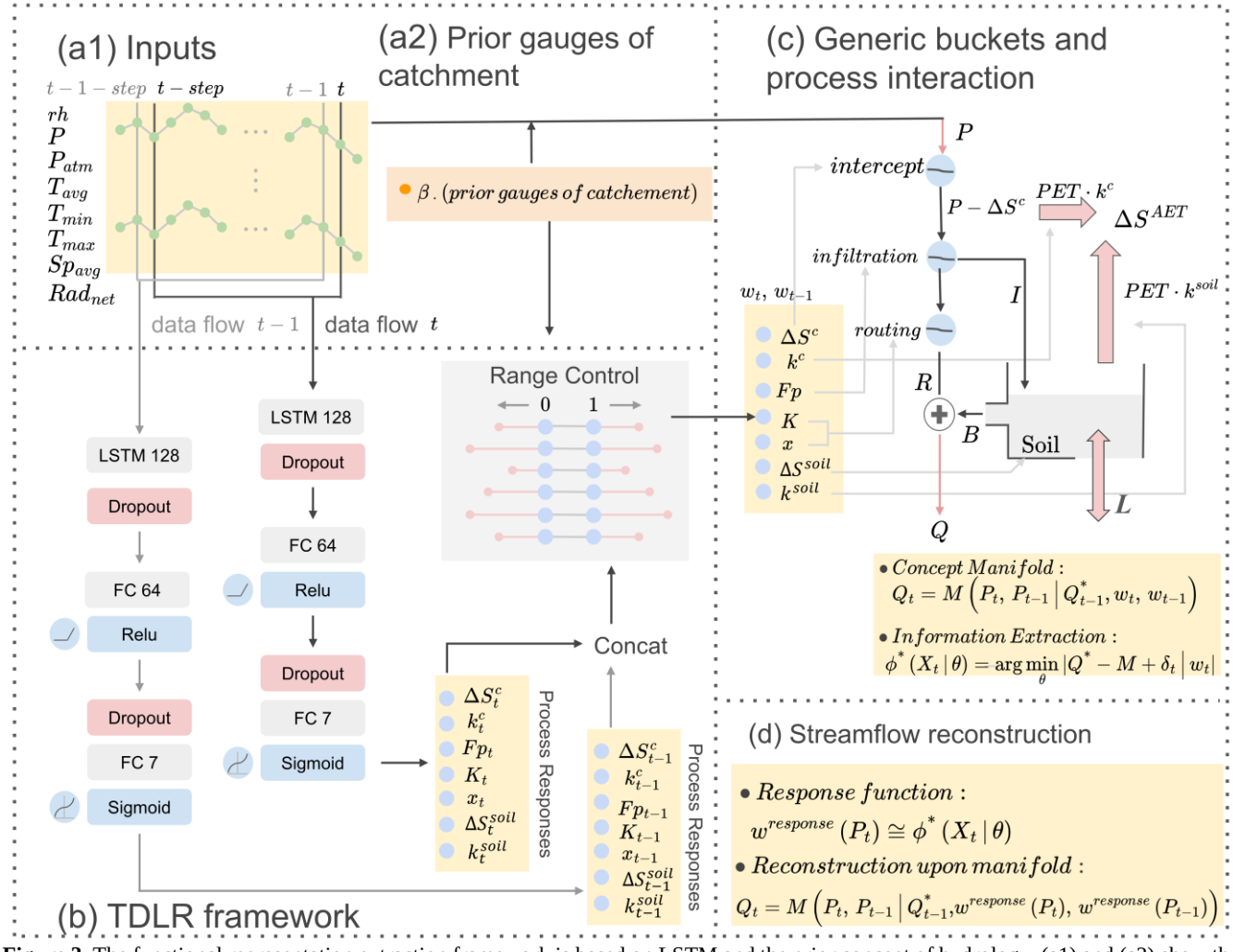


Figure 3. The functional representation extraction framework is based on LSTM and the prior concept of hydrology. (a1) and (a2) show the data input and the prior gauges of the catchment. The meteorological forcing includes rh, P, P_{atm}, T_{avg}, T_{min}, T_{max}, Sp_{avg}, and Rad_{net} are relative humidity(%), precipitation(mm), atmosphere pressure(hpa), daily average temperature(°C), daily minimum temperature(°C), daily maximum temperature(°C), average wind speed(m/s), and net radiation(W/m^2), respectively. (a2) is the prior gauges showing the feasible

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

range of hydrological response. (b) is the LSTM layer with 128 hidden units, with 64-day time steps. FC 64/7 represents a fully connected layer with 64/7 output dimensions. (c) represents the generic buckets and process interactions. Black and red arrows are the fluxes in the boundary and on the boundary, respectively. (d) represents the streamflow reconstruction with the projected relationships ($\omega^{response}(P_t)$) from extracted functional representations (ω_t).

During LSTM training, the chronological sequence is strictly maintained within the input and output windows, preserving the network's ability to capture the critical memory effects of the hydrological system. We employed RMSprop with mini-batching, randomly sampling segmented data sequences to update the network parameters. The mini-batching resampling strategy relies on a fundamental hypothesis: that the underlying hydrological functional representations remain time-invariant throughout the historical record. This assumption of functional stationarity justifies the random batching approach, allowing the LSTM to induce time-invariant response mechanisms. Combined with our proposed loss regularization, this training scheme ensures that optimization traverses a loss hypersurface specifically structured to describe how current and antecedent driving forces map onto current streamflow. Hence, the LSTM is trained to simulate the functional representations in parallel for solving the closure problem in hydrological modeling. Beyond solving the closure problem in catchment hydrology, we want to emphasize that once we know the forcing sequence and process responses at the current and previous times, combined with previous-time streamflow estimates, the current-time streamflow can be estimated. Therefore, we could use the previous time-estimation result and the precipitation sequence to iteratively generate the streamflow sequence without LSTM.

3.3 Benchmark Paradigms: Hybrid and Pure Deep Learning Models

To compare the different paradigms, (i) hybrid model advancing parameterization, (ii) pure-data driven model, and (iii) the TDLR with generic buckets, we implemented the hybrid model based on the Hydrologiska Byråns Vattenbalansavdelning (dPL-HBV) model (Feng et al., 2022; Tsai et al., 2021) and the LSTM model (Kratzert et al., 2019). First, dPL-HBV is a differentiable physics-informed model that hybridizes an LSTM and an HBV model, leveraging the LSTM to learn HBV parameters. It incorporates the LSTM architecture with dynamic or static parameterization of conceptual process equations, as our first benchmark hybrid model. The comparison between dPL-HBV and TDLR should highlight the functional response differences between parameter calibration for the given process equation and that of our proposed framework. To strictly isolate the influence of meteorological forcing in our single-basin setup, we excluded static catchment attributes. Static and dynamic parameterization schemes were applied. In the static scheme, the model uses a single, time-invariant parameter set over the simulation period, serving as a differentiable calibration. For the dynamic scheme, a subset of parameters is allowed to vary daily based on meteorological history, while the remainder remains constant. Second, we introduced an LSTM for daily rainfall-runoff simulation (Kratzert et al., 2019) as the second benchmark model to know the capacity of the data-driven model in our task, for each catchment. Operating in a sequence-to-one configuration, the model processes rolling windows of meteorological forcing sequences to simulate current streamflow. We tested the pure LSTM model with and without a previous timing streamflow input, and with two types of moving windows: 64-day and 365-day. Within this architecture, the LSTM's

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

internal cell states serve as an implicit, unconstrained representation of catchment storage. We presented the comparison in Sect. 4.4.

4 Results and Discussions

In this section, we show and discuss (i) reconstruction accuracy as a measure of how effectively the extracted functional representations reproduce observations. (Sect. 4.1), (ii) the dominant functional dependence of each inferred process on precipitation and related drivers (Sect. 4.2), (iii) the water budget and residuals under the mass conservation constraint (Sect. 4.3), (iv) a paradigm comparison between advancing parameterization with fixed prior functional form, pure-data end-to-end LSTM, and the TDLR framework (Section 4.4), and (v) strong re-extracting stability of the functional representation (Sect. 4.5).

4.1 Hydrograph Reconstruction with Extracted Functional Representations

One of the goals of the hydrological model implementation is to reconstruct hydrographs. With the extracted functional representation and designed generic buckets, the performance and results are shown in Fig. 4.

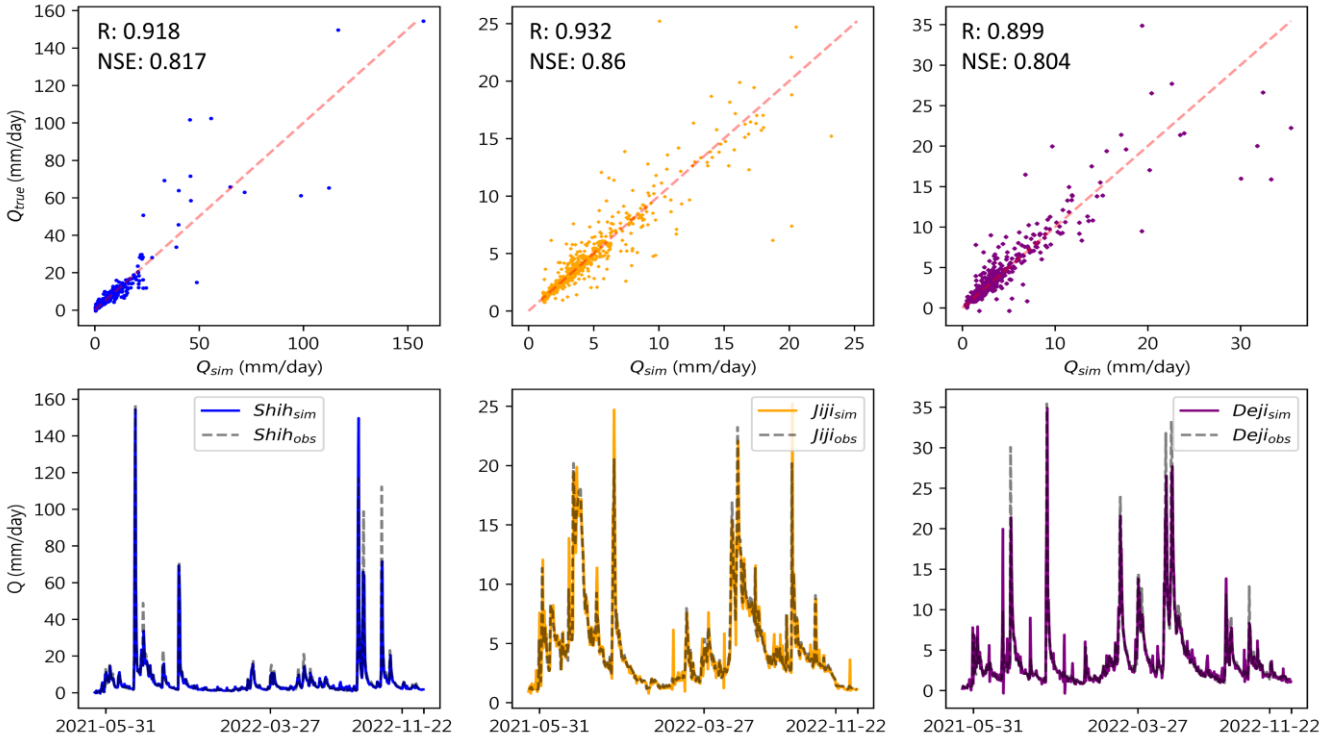


Figure 4. Reconstruction accuracy with the extracted functional representations learned by a single LSTM and the precipitation series from late May 2021 to December 2022, on a daily scale. The blue, orange, and purple colors in the scatter and line plots represent the case in the Shihmen, Jiji, and Deji catchment areas.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Competitive reconstruction accuracy serves as a baseline for testing the representation of the result. All catchments demonstrate competitive reconstruction accuracy for streamflow sequences using precipitation series from May 2021 to December 2022, initial streamflow, and the extracted functional representations. Remarkably, competitive reconstruction accuracy not only serves as an essential requirement for verifying the representation of the overall response in the closure problem but also provides a reliability check of the extracted function's interpretability. Moreover, the high reconstruction accuracy implies that the LSTM, as a functional approximator, can effectively address the memory effect, suggesting that different meteorological sequence combinations with the same current precipitation may yield different hydrological responses.

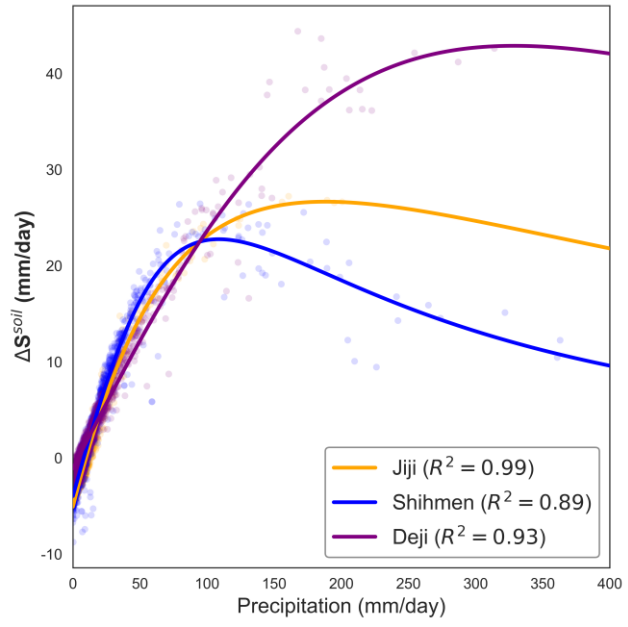
Despite the LSTM being able to deal with the memory effect, the extracted functional representation shows strong covariance between the meteorological driving forces and the rainfall event life cycle. The covariance among current precipitation, other forcing, and spatial heterogeneity is significant. In Appendix A, the current precipitation provides the dominant information to describe the current hydrological response functions, indicating that the residual of the response estimation when we only use precipitation is small enough. It supports those three catchments that exhibit strong covariance among the meteorological driving forces, echoing the characteristic of the meteorological-hydrological-morphological-geological interaction in hillslope regions; therefore, in the following section, we present the projected relationships between current precipitation and current process responses from the extracted functional representations. To be more specific, although we cannot completely verify the identifiability of the extracted representation through the accuracy, the accuracy checking and stability of retraining the LSTM to extract the functional representation (discussed in Sect. 4.5), and the interpretability of the extracted representations (Sect. 4.2), at least, provide a complementary perspective and additional information for addressing the closure problem.

4.2 Interpretable and Projected Relationships in Catchment Scale

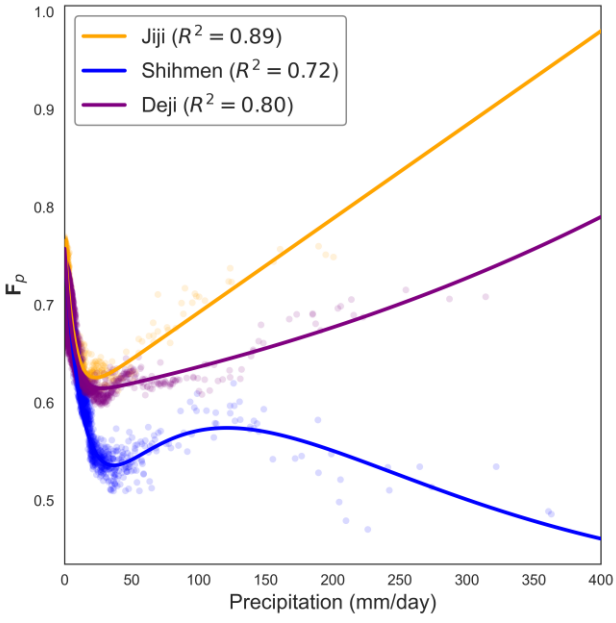
The strength of TDLR lies in its direct search and distinction of functional representations and characteristics across different catchments, guided by the mass conservation constraint and the design of generic buckets. Given the competitive reconstruction accuracy of the extracted functional representations, their interpretability is another contribution of TDLR. Figure 5 summarizes six projected relationships (points, $\omega^{response}(P_t)$) extracted from functional representations and fitted curves (lines), which are used solely to visualize dominant trends. The details of the fitted curves are presented in the Appendix C. Here, ΔS^{soil} is the daily change in the integrated subsurface bucket, ΔS^{c} is the daily change in canopy interception storage, F_p is the fraction of effective precipitation on the soil that become surface runoff (direct runoff), K is the effective travel time parameter, and x is the Muskingum "wedge" parameter controlling translation versus attenuation in routing. In Fig. 5(f), the AET/PET relationship is shown against the ratio in x axis ($(P - \Delta S^{c} - \Delta S^{soil} - \Delta S^{m})$/PET) where ΔS^{m} denotes the change in the subsurface storage term as defined in Eq. (6). To demonstrate that the interpretability of the functional representations could be aligned with the local morphological-meteorological characteristics. In the following subsections, we interpret and discuss these projected relationships with the classic theory and the mechanisms.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

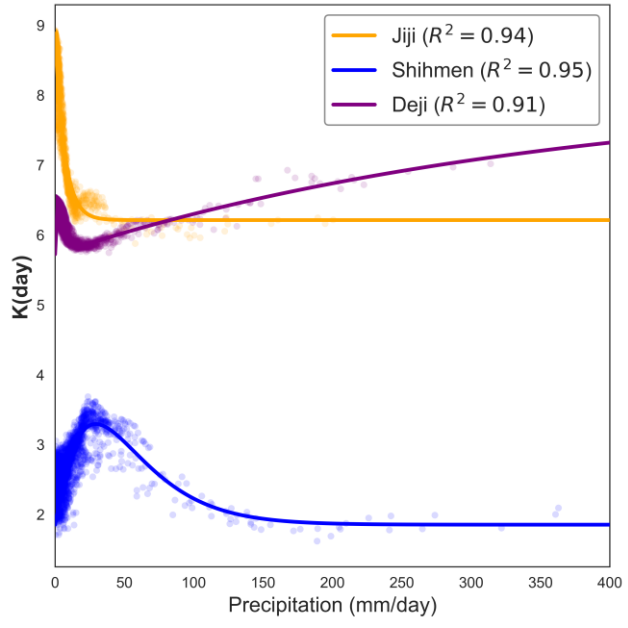
(a)



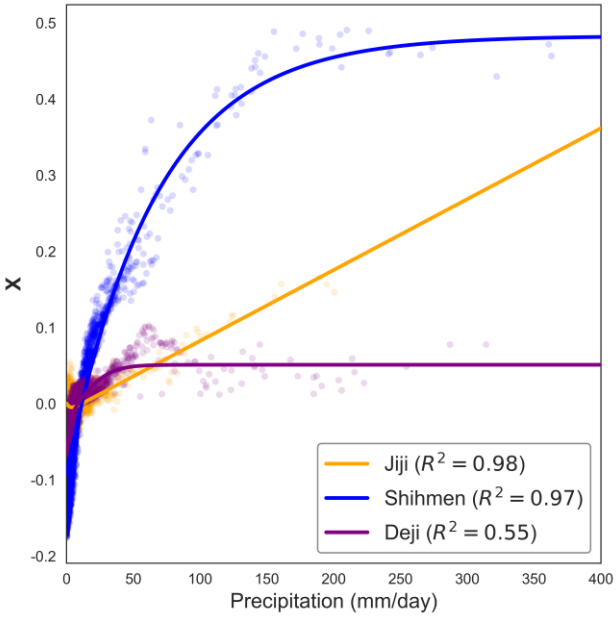
(b)



(c)



(d)



https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

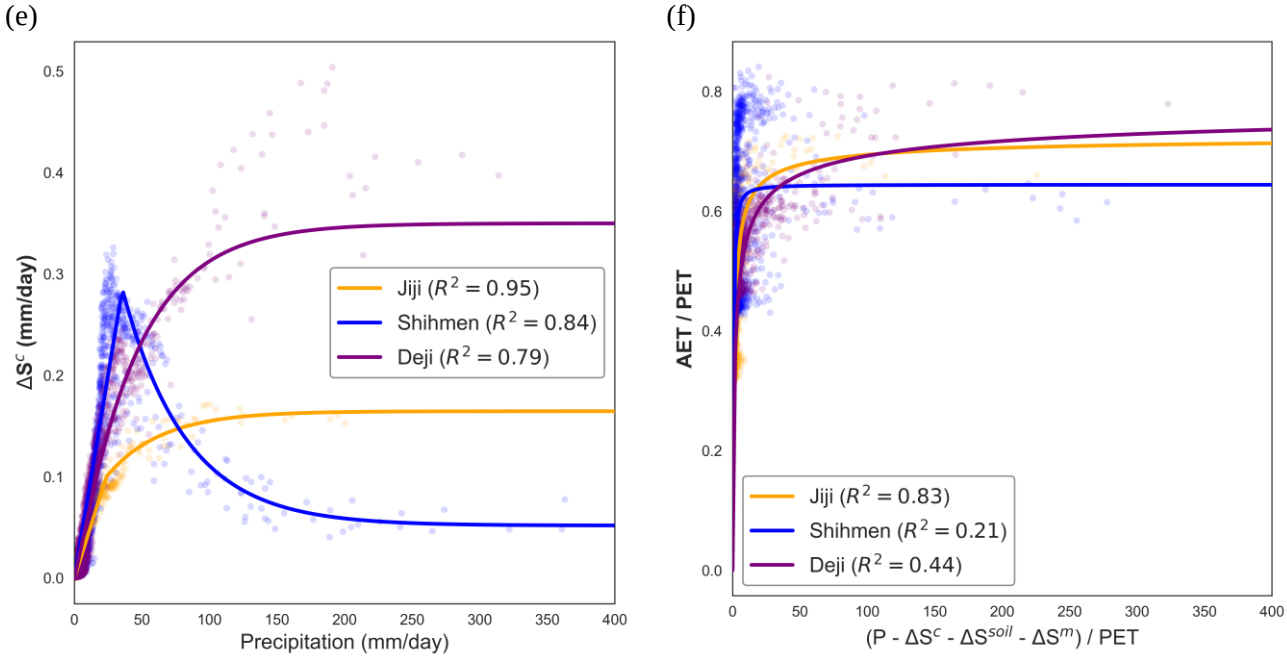


Figure 5. Cross-scale similarity of the projected relationship. ΔS^{soil}, ΔS^{c}, AET, PET, K, x, and F_p, represent the change of water content in the integrated subsurface buckets, canopy bucket, actual evapotranspiration, and potential evapotranspiration, the effective travel time, wedge parameter, and the ratio between the water that becomes the surface runoff and effective precipitation on the soil. The points and lines denote the extracted functional representations by LSTM, and the descriptive continuous function by the points. Blue, orange, and purple colors represent the case in Shihmen, Jiji, and Deji, respectively. The R-squared value indicates the fitness between the points (outputs of LSTM) and the guessed continuous functions.

4.2.1 Runoff Generation and Subsurface Storage Groundwater Dynamics

We first focus on the small precipitation case, (before 20-30 mm/day), as shown in Fig. 5(a) and (b), all catchments show the consistency patterns in $\Delta S^{soil}(P)$ and $F_p(P)$. The decreasing F_p and increasing ΔS^{soil}, are likely dominated by the saturation-excess mechanism (Bernier, 1985; Beven, 1989; Beven and Kirkby, 1979), reflecting the dynamic expansion of the contributing area. The wetting area expansion, along with the rain area expansion, results in water entering the bucket. It echoes the CI patterns shown in Fig.1(b), with consistent wet-area expansion from the mountain region to the dry region. After the 20-40 mm/day, the entire catchment area is wet, then all catchments enter infiltration-excess control (Clapp and Hornberger, 2010; Garen and Moore, 2005), acting like a sponge, F_p increasing, especially in the cases of Jiji and Deji. The covariance between the wet area and the precipitation is also supported by the concept of the spatial probability curve (Fang et al., 2017; Cheng et al., 2006). The inferred subsurface response $\Delta S^{soil}(P)$ (Fig. 5(a)) complements in the saturation-excess mechanism which triggers subsurface stormflow (Beven et al., 2021; Weiler et al., 2005). All catchments show rapid increases in $\Delta S^{soil}(P)$ at low-to-moderate precipitation, consistent with net wetting of the wet area expansion along with the raining area. Also, in

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

both catchments, the $\Delta S^{soil}(P)$ reach the plateau or even a decreasing pattern during the high precipitation range. This pattern may reflect the influence of the subsurface storm flow or preferential fractal flow processes. The subsurface stormflow mechanism usually occurs in a catchment with a clear hillslope or a dramatic change in the water table (Chen et al., 2024b). Hence, the subsurface storage response reaches a plateau or even weakly decreases. Jiji peaks at higher precipitation and then declines more gently, indicating larger effective storage and a slower transition from storage filling to runoff. Deji shows the largest $\Delta S^{soil}(P)$ and a slow approach to saturation, consistent with a larger effective bucket and stronger capacity for precipitation retention at the daily scale. Shihmen reaches an early peak and, however, then shows a decline in $\Delta S^{soil}(P)$ with increasing P. The decline in $\Delta S^{soil}(P)$ suggests a comparatively small effective subsurface storage and/or rapid drainage/overflow, such that additional precipitation is preferentially routed to other pathways rather than retained as subsurface storage when rainfall intensity exceeds a threshold.

In Shihmen, the combination of (i) reduced ΔS^{soil} at very high precipitation and (ii) the decline in F_p at extremely high P (Fig. 5(b)) is consistent with a shift toward rapid subsurface stormflow (Chen et al., 2024b) rather than additional surface runoff during extreme events. This interpretation aligns with the transmissivity feedback or preferential fracture flow concept, whereby hillslope transmissivity and lateral subsurface connectivity increase disproportionately as the subsurface approaches saturation during extreme precipitation (Türk et al., 2026; Xiong et al., 2022). Site-specific geology and structural controls (e.g., fault-aligned pathways) may further enhance rapid lateral drainage during extreme storms. The interpretation of the preferential fracture flow is also consistent with the distinct residual pattern of the mass conservation, as discussed in Sect. 4.3

4.2.2 Surface Routing and Network Connectivity

Fig. 5(c), $K(P)$ shows the clear three patterns in three distinct catchments, (i) monotonic decreasing then reach plateau, (ii) smooth increasing, (iii) increasing first, then decreasing, last reach the plateau. We first discuss the case in Jiji (i) and Deji (ii). The first decrease, then reaching an asymptotic plateau pattern in Jiji, shows the clear celerity increasing more significantly than the effective travel length increasing. We suggest that, in Fig 1(c), although the storm center moves away from the outlet, the precipitation concentration remains constant, and the celerity increases due to the dominant increase in water depth. On the other hand, Deji presents a case of effective travel-length domination. The low-precipitation range still maintains control, but after 20 mm/day, the storm center not only moves away from the outlet but also becomes more concentrated in its center, as shown in Fig. 1(c). As a result, in the Deji case, the effect of a longer travel distance outweighs the celerity increase with streamflow.

Additionally, in Fig. 5(d), from the perspective of riverine connectivity dynamics (Mahoney et al., 2023; Shrestha et al., 2025; Tull and Passalacqua, 2025), both saturated and unsaturated responses indicated that rainfall distribution affects the water stored in surface water buckets within the riverine system. Deji increases rapidly at low precipitation but saturates at a low x value, a behavior recognized as storage-dominated (small x). Jiji shows a near-linear increase without an evident plateau

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

over the observed precipitation range. The unshown plateau pattern in Jiji is due to a greater absence of saturation, which limits effective riverine network connectivity, supported by a larger basin area and broader spatial variability.

Notably, in the case of Shihmen K and x patterns, after the increasing phase, the travel time is similar to Jiji, becoming celerity-dominant, then decreasing, and finally reaching a plateau, with a stable storm center and rain concentration. However, in the low-to-moderate precipitation region, K increases likely due to the negative x value, rather than the prolonged travel length. As shown in Fig. 1(a), all catchments have a significant hillslope-plain transition region, but only the Shihmen catchment has a comparable area contribution between hillslope and plain. The conditions occurring in Shihmen would violate the celerity-discharge relationship derived by kinetic wave theory. When the water routed from headwater to outlet encounters a significant discontinuity in the channel slope, the backwater effect becomes pronounced. Thereby, to maintain mass conservation, the x value would become negative. The headwater would tend to flood the plain first rather than directly reach the outlet, prolonging the effective travel time. After the 20-30 mm/day, the connectivity of the whole catchment increases, both K and x behave as in the framework of the kinetic wave theory and the Saint-Venant Equation.

4.2.3 Canopy Storage Change as a Practical Catchment-Scale Function

Both the Jiji and Deji catchments exhibit an initial increase that approaches a stable asymptotic plateau, as shown in Fig. 5(e). This behavior is consistent with classic interception theory, in which dynamic canopy storage increases with cumulative rainfall until a finite structural capacity is reached, after which excess rainfall is primarily partitioned into throughfall and stemflow (Klaassen et al., 1998; Li and Tian, 2025). During low-to-moderate precipitation events, the daily interception-storage change (ΔS^c) reflects the wet-canopy evaporation; as long as the volume of intercepted water is less than the atmospheric evaporative demand, it is rapidly evaporated back to the atmosphere. In Fig. B1, as precipitation intensity increases and the relative evaporative fraction decreases, the canopy approaches full saturation. Beyond this threshold, the learned curve suggests that rainfall partitioning becomes dominated by stemflow and throughfall, while the maximum retained volume is fundamentally constrained by vegetation geometry and physical characteristics. Consequently, the distinct plateau observed during high-intensity precipitation implies a maximum canopy storage capacity, which could be determined by the fractional vegetation cover and forest stand architecture within the catchment. Deji attains the highest asymptotic limit, suggesting a greater maximum canopy storage, typically associated with denser vegetation structure or higher fractional forest cover.

In contrast, Shihmen exhibits a qualitatively distinct, non-monotonic response. It is plausible, driven by a unique covariance between precipitation and wind speed, as discussed in Sect. 2.1. In this catchment, ΔS^c rises steeply at low precipitation, peaks during moderate events, and subsequently declines toward a lower asymptote under heavy rainfall. As illustrated in Fig. 1(d), low-to-moderate precipitation events at Shihmen frequently coincide with the highest recorded wind speeds. These strong winds could significantly enhance aerodynamic conductance, driving exceptionally high rates of wet-canopy evaporation during the storm and thereby elevating the apparent interception loss, as shown in Fig. B1. However, during larger, more intense precipitation, the low asymptote suggests a reduced canopy storage, with rainfall primarily partitioned into stemflow

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

and throughfall. The effective retention capacity of the foliage could be physically reduced by the combined kinetic energy of heavy raindrop impacts and ambient mechanical shaking of the branches. This dynamic forcing promotes the premature release of intercepted water as a throughfall. This process aligns with the learned decline in ΔS^c, potentially establishing a dynamic capacity limit under extreme storm conditions. Our interpretation of these extracted responses of interception to rainfall in all catchments echoes the premise that canopy interception is a highly meteorologically sensitive and dynamically forced ecohydrological variable (Gao et al., 2024; Véliz-Chávez et al., 2014).

4.2.4 Evapotranspiration, Water Limit, Energy Limit, and Summary

Based on Eq. (4), the actual evapotranspiration (AET) is the product of PET and the sum of the fractions learned by LSTM, $k_t^c + k_t^{soil}$, the ranges of both are from zero to one, reflecting how much water is evaporated. Hence, the ratio of AET and PET is the sum of the fractions. Figure 5(f) shows AET/PET as a function of the ratio calculated by daily water availability (Zhang et al., 2008) and PET. All three basins exhibit a steep rise at the low ratio value, followed by saturation, consistent with the transition from water to energy limitation. However, the inferred plateau values are substantially below 1 (approximately 0.7 for Jiji and Deji, and 0.65 for Shihmen), indicating that even on high "available water" days, AET remains below PET at the daily scale. This sub-unity plateau can reflect both physical and methodological factors. Physically, canopy resistance, stomatal regulation, and soil moisture limitations can keep AET below PET even when precipitation is high. Methodologically, PET is computed here using the FAO-56 Penman–Monteith formulation (Cai et al., 2007), which represents reference-surface conditions; without site-specific roughness and surface resistance parameters for the dominant land covers, PET might be biased relative to actual canopy conditions, directly affecting the inferred AET/PET ceiling (Zhao et al., 2025). The comparatively low R^2 values for this relationship (notably in Deji) indicate that precipitation-derived water availability alone does not explain all day-to-day variability in AET/PET; radiation, humidity, and temperature (which were secondary drivers in Fig. A1) likely contribute to residual variability. Moreover, compared to the canopy interception pattern of precipitation, during low nondimensional water availability, the fraction of evapotranspiration is highest, echoing the local climate characteristic that strong winds can occur not only during high precipitation due to typhoons but also during low precipitation or even on sunny days.

In summary, Section 4.2 presents the projected relationships from extracted functional representations (LSTM outputs) as interpretable mechanisms that reflect the distinct interactions between the morphology and climatology of each catchment. Because these relationships exhibit non-linear, threshold-based behavior without requiring rigid prior functional forms, the projected relationships represent purely data-driven induction of catchment-scale responses. The distinct time-invariant projections across the three catchments imply that a general process-function form for each specific process might not exist, suggesting the limitations of the conventional paradigm with a fixed prior functional form in conceptual hydrological modeling. Additionally, without giving the prior functional form for functional representation extraction, the water budget serves as the

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

axiom of the hydrological model constrained in TDLR. In the next section, we discussed the residual associated with the mass conservation constraint under the closed boundary assumption.

4.3 Residual of the Mass Conservation Constraint and Water Budget

In this study, the L_t represents the sum of the IGF, the structural uncertainty, and measurement error as shown in Eq. (5). If ϵ_t is assumed to act as white noise, the systematic variation in L_t may reflect a proxy signal of IGF when evaluated alongside the extracted moisture fluxes and generic bucket design. Hence, we check two main scopes, the daily scale and the long-term L_t, to speculate the relationships between inter-basin groundwater flow and other residuals in Eq. (5). Therefore, we examine both the daily-scale dynamics and the long-term L_t to explore the relationships between groundwater exchanges and other mass residuals. As shown in Fig. 6 and Table 1, all three catchments demonstrate a clear relationship between precipitation, dynamics of integrated subsurface bucket, and a non-ignorable long-term L_t. It indicates that while the IGF cannot be definitively separated from measurement errors and unrepresented partitioning components, the systematic behavior of L_t suggests that IGF plays a critical role in the overall mass balance.

Table 1. Water Balance, including the residuals of the mass conservation constraint (L_t), water storage change of three buckets, precipitation, streamflow, and AET from 2014 to 2021.

	ΔS^{soil} (mm/day)	ΔS^{c} (mm/day)	ΔS^{m} (mm/day)	P (mm/day)	Q (mm/day)	L_t (mm/day)	AET (mm/day)
Shihmen	-0.430	0.036	0.00082	7.51	4.93	0.86	2.11
Jiji	-1.541	0.0163	-0.013	5	4.72	-0.18	2.02
Deji	-0.663	0.03	-0.0013	6.3	3.93	0.84	2.18

Figure 6 illustrates the relationship between daily precipitation and the mass conservation residual, color-coded by ΔS^{soil}. The magnitude of this residual flux exhibits distinct threshold behavior proportional to precipitation intensity across three catchments. For instance, the Jiji catchment displays a classic threshold response: significant outward flux ($L_t > 0$) occurs primarily when precipitation exceeds 100 mm/day, above which the flux intensity increases sharply. In the Shihmen catchment, the outward flux also increases with precipitation but stabilizes into a plateau beyond 150 mm/day. Conversely, the Deji catchment displays higher dispersion in the precipitation-residual relationship without a single distinct plateau.

We hypothesize that these distinct residual patterns reflect catchment-specific subsurface dynamics, particularly IGF. The limited variation in active subsurface storage relative to precipitation suggests that IGF could heavily influence baseflow sources, consistent with findings by Yang et al. (2025). Mechanistically, intense precipitation may drive rapid infiltration that displaces existing subsurface storage, initiating outward IGF that subsequently recharges adjacent or downstream layers. Under this interpretation, the plateau observed in Shihmen aligns with our earlier discussion (Sect 4.2.1) regarding the activation of preferential fracture flow paths or subsurface stormflow networks, which efficiently route excess water laterally rather than contributing to continuous deep recharge. For Deji, the higher dispersion might be attributed to a larger subsurface storage capacity provided by dense mountainous forest cover, which buffers hydrological connectivity.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

From a long-term water budget perspective (Table 1), all three catchments exhibit a decreasing trend in ΔS^{soil}, with the Jiji catchment experiencing the most severe depletion. Even when considering a potential 10% measurement uncertainty in precipitation and streamflow, the inferred negative trends strongly corroborate local studies and government reports indicating groundwater-level declines and land subsidence in the downstream regions of the Jiji catchment between 2010 and 2020 (Chen et al., 2024a). Furthermore, the actual evapotranspiration (AET) estimates, derived using a dynamic water-availability penalty rather than relying solely on potential evapotranspiration, align well with satellite observations in Taiwan. The minimal storage changes observed in the change of canopy storage (ΔS^{c}) and subsurface routing (ΔS^{m}) are minor and consistent with the fundamental assumptions of long-term catchment surface-water balance and numerical closure.

Overall, Sections 4.1 to 4.3 demonstrate the reconstruction accuracy, the interpretation of the extracted functional representations via projected relationships, and the robustness of the water budget. In the following section, we present a paradigm comparison between traditional parameterization with fixed prior functional forms and the proposed TDLR process-function extraction framework.

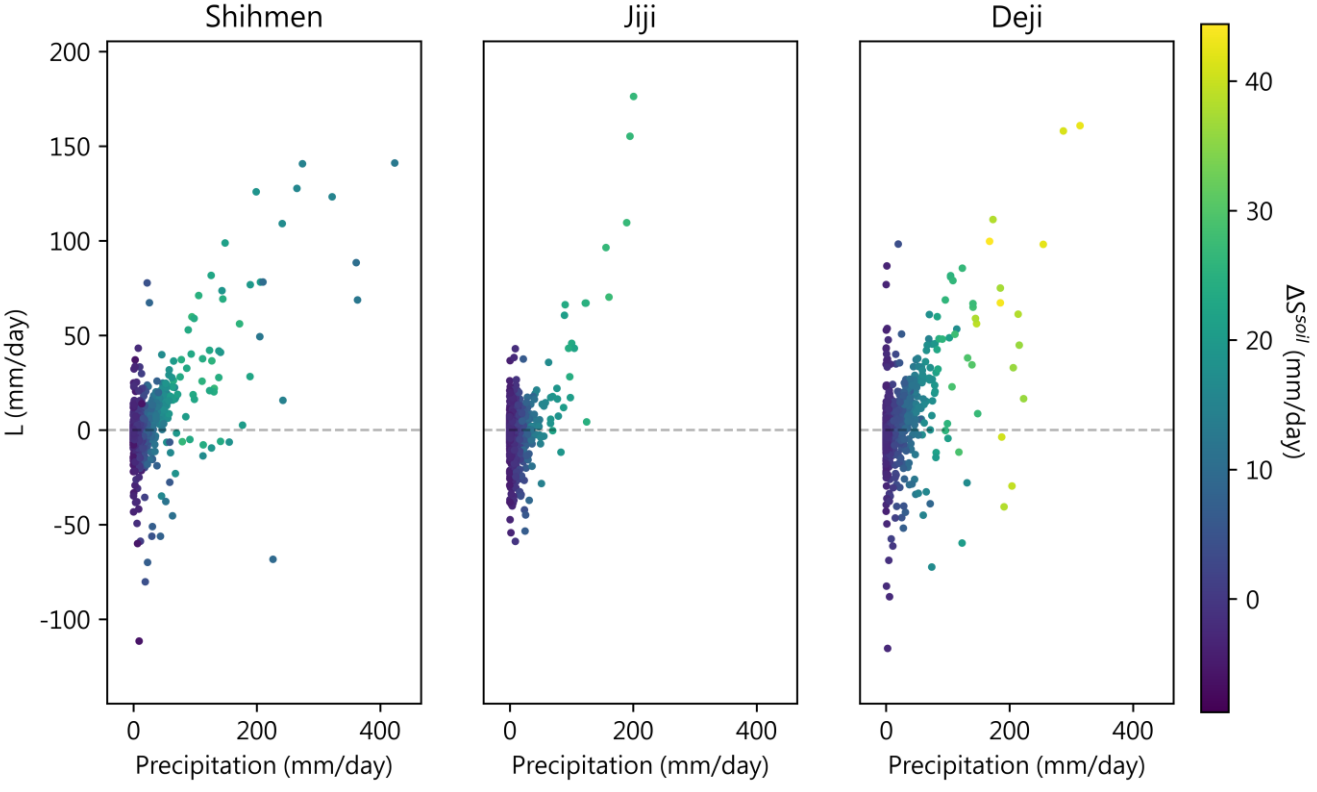


Figure 6. The x-axis for each catchment is daily cumulative precipitation. The y-axis represents the residual of mass conservation constraint (L). The color bar of the points is the water change in the integrated subsurface bucket. The points above the horizontal dashed lines indicate subsurface storage that flows out to recharge downstream or other catchments; the points below the dashed line represent subsurface storage that receives additional supplementation from the catchment.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

4.4 Paradigm Comparison between Parameterization and Functional Searching

The proposed TDLR framework enables LSTM models to directly learn functional representations, offering a new paradigm relative to existing hydrological modeling approaches that rely either on purely data-driven methods or on conceptual models with predefined functional forms. To elucidate the sources of performance differences, in this section, we compared three modeling paradigms: (i) the hybrid modeling approach (dPL-HBV), (ii) purely data-driven LSTM models without prior hydrological constraints (Kratzert et al., 2019), and (iii) the proposed TDLR-based process-function learning framework.

In Table 2, we first evaluated simulation accuracy against the proposed TDLR framework, dPL-HBV with static and dynamic parameters, and four pure LSTMs on different input sequences (64 and 365 days), including or excluding the previous timing streamflow. The results demonstrated that the TDLR framework performs competitively. As expected, the purely data-driven autoregressive LSTMs (using previous-day Q with a 365-day lookback) achieve marginally higher NSE values in the Jiji and Deji catchments (0.87 and 0.84, respectively). Because autoregressive models inherently benefit from continuous streamflow state updating, these scores effectively represent the prediction ceiling of the theoretical deep learning models for these catchments. Notably, when evaluated under strict continuous rainfall-runoff simulation conditions in these hillslope catchments, the similar accuracy between pure data-driven and TDLR indicates that the ability of the LSTM with TDLR makes the best use of the resource, even if we add the generic bucket and mass conservation constraints.

Moreover, the proposed TDLR is competitive with pure LSTMs and outperforms dPL-HBV across all studied catchments (e.g., reaching an NSE of 0.863 in Shihmen). Unlike pure data-driven approaches, both dPL-HBV and TDLR relied on mass conservation and the given generic bucket design, thereby retaining the interpretability of the learning results through the discussion of the structure hierarchy (Gharari et al., 2021; Gupta et al., 2012). The strength of the dPL-HBV relies on parameter calibration within predefined process equations that ensure consistency with mass conservation, generic bucket structures, and assumed functional forms. Nevertheless, when these formulations break down at the catchment scale, the rigid prior functional forms may constrain the flexibility of parameterized hybrid models when representing highly complex catchment-scale responses.

Table 2. The simulation accuracy comparison between the proposed framework, dPL-HBV, and LSTM. We use Nash-Sutcliffe efficiency (NSE) (Gupta et al., 2009) as the metric to assess simulation accuracy during validation periods.

	Extracted functions with TDLR	*dPL-HBV (static/dynamic)*	*Pure LSTM added Q_{t-1} as input (64/365 days)*	*Pure LSTM (64/365 days)*
Shihmen	0.863	0.833 / 0.802	0.86/0.85	0.84/0.86
Jiji	0.860	0.305 / 0.528	0.86/0.87	0.81/0.81
Deji	0.823	0.709 / 0.656	0.82/0.84	0.74/0.72

To highlight process differences from prior forms parameterization and functional representational search, we compared the responses of two key hydrologic projected relationships $\Delta S^{soil}(P)$and $F_p(P)$ from dPL-HBV and TDLR as shown in Fig. 7. The dynamics of integrated subsurface bucket described by TDLR, inferred from TDLR (represented with lines), increases

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

with precipitation but approaches a saturated limit during high-intensity storm events. In contrast, dPL-HBV exhibited a continuously increasing trend with precipitation (represented by points), resulting in lower baseflow during heavy storms. In fact, streamflow often receives a substantial, rapid contribution from subsurface or groundwater sources at the onset of storms (Klaus and McDonnell, 2013; Partington et al., 2012; Yang et al., 2025). The dPL-HBV, based on a linear reservoir assumption, might not capture these rapid subsurface contributions in the hillslope. It could be one of the reasons for the low accuracy of dPL-HBV for these hill-slope catchments. This limitation reflects the broader understanding that catchment-scale hydrological responses are inherently nonlinear, driven by interactions between storage dynamics and hydroclimatic conditions (Wang et al., 2025).

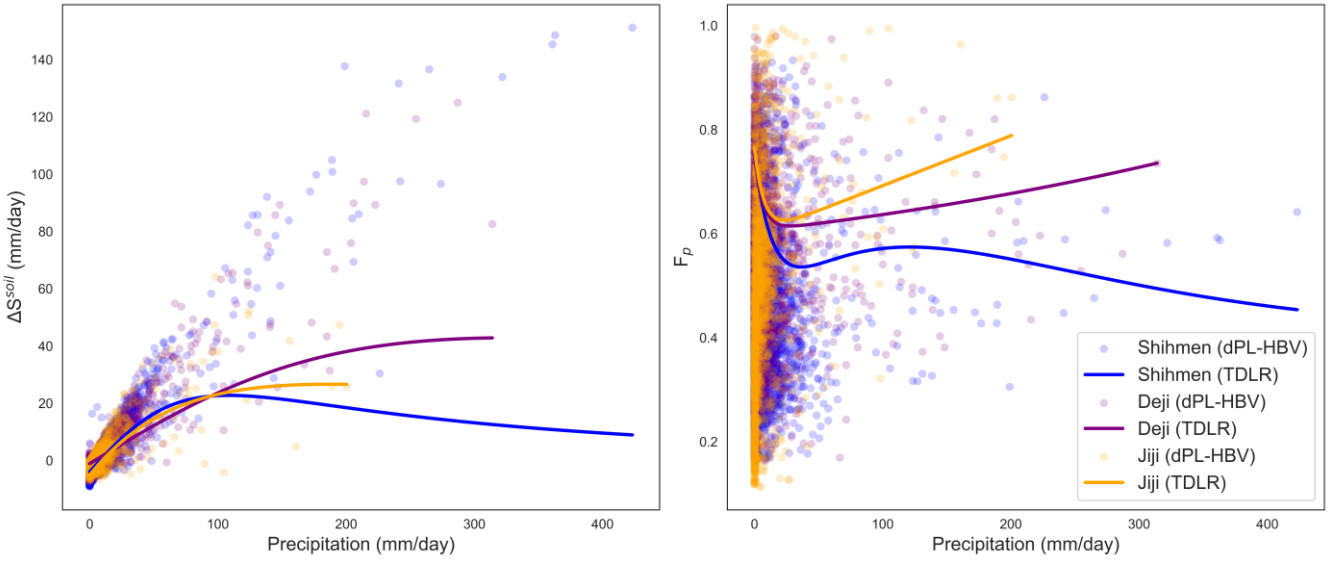


Figure 7. Infiltration-related and the dynamic of integrated subsurface bucket comparison in the scope of projected relationships (TDLR vs dPL-HBV): solid lines show the non-linear regression results based on the extracted functional representations; points show the functional representations from the dpL HBV-dynamic. The method for identifying solid lines is presented in the Appendix C and in Fig. 5.

Similarly, examining the runoff generation mechanism, conceptual models like dPL-HBV utilize established soil-water partitioning equations. While these empirical functional forms are highly effective and interpretable, especially at localized and field scales, extending them to complex catchment-scale phenomena (spatial heterogeneity in driving forces, morphologic, and hydraulic conductivity) can sometimes restrict their flexibility.

We argue that the scattered points from dPL-HBV indicate that similar precipitation events can yield different runoff generation tendencies depending on antecedent soil moisture. While this relationship holds well at the flattened field scale with uniform precipitation intensity and hydraulic conductivity, its direct application to a larger catchment scale (>500 km²) is problematic. This spatial heterogeneity is precisely the motivation behind distributed modeling approaches. In Fig. 7, the wide scatter of points during small to moderate precipitation events suggests a physical unreality: it is highly unlikely that an entire catchment of this size would uniformly exhibit either near-zero or over 80% infiltration simultaneously. Therefore, such rigid field-scale formulations may not adequately represent the spatially integrated hydrological dynamics in our study area.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

As shown in Sect. 4.2, our TDLR framework yields distinct, coherent, and catchment-specific partitioning curves without relying on a prescriptive prior form. These distinct relationships could reflect local morphological-meteorological interactions. This flexibility suggests that TDLR can serve as a complementary approach to extracting functional representations from data at the catchment scale. Notably, while true functional responses at such large catchment scales remain difficult to verify, conventional models often rely on strong assumptions to simulate hydrographs while maintaining interpretability. In this context, TDLR offers a complementary perspective, providing alternative, interpretable functional representations with competitive accuracy at the catchment scale. Furthermore, a key strength of the TDLR lies in the stability of the re-extracted representations. The following section will demonstrate the capacity of TDLR to yield stable, time-invariant functional representations with high re-extraction stability.

4.5 Stability of the Representations Re-extraction with TDLR

Equifinality remains a fundamental and unresolved issue in hydrological modeling, as multiple parameter sets can yield similar simulations within generic bucket structures and empirical formulations. Beven (2006a) highlights that this problem is unavoidable and cautions against attributing physical meaning to individual parameter values (Beven, 2006a). Equifinality limits the ability of models to establish a deterministic mapping between parameter sets and catchment characteristics. As a result, even when a model achieves high accuracy, calibrated parameter sets may lack representative stability and struggle to robustly verify underlying catchment behaviors.

In such a case, checking the stability of the re-extracted functional representations in different training random seeds is necessary to compare the strengths of the functional search paradigm and parameterization. We retrained the LSTM with TDLR 30 times by selecting random seeds and recorded the extracted functional representations in Fig. 8(a). The performance of the training results is shown in Fig. 8(b). Within the scope of function shape stability, canopy storage change, and the AET/PET ratio have the largest ranges among the other four extracted functions. Because the LSTM loss propagation relies on the streamflow simulation error, the order of canopy interception and evapotranspiration-related processes is much smaller than the streamflow directly impacted. The functional searches for these two patterns are heavily dependent on the penalty to offer the extra loss signal for gradient descent. Hence, the representation uncertainty of these two functional representations is most significant.

Remarkably, for each functional representation, shown in Fig. 8(a), the robust stability of the re-extracted patterns provides strong support for the interpretability of the projected relationships from functional representation. The re-extracted patterns show strong conformity across 30 re-extraction runs with different random seeds. It indicates that the functional search with the given generic buckets and mass conservation yields a stable LSTM output. Hence, in these three catchments, TDLR is an effective approach for hydrological modeling at the catchment scale as a potential closure problem solver.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

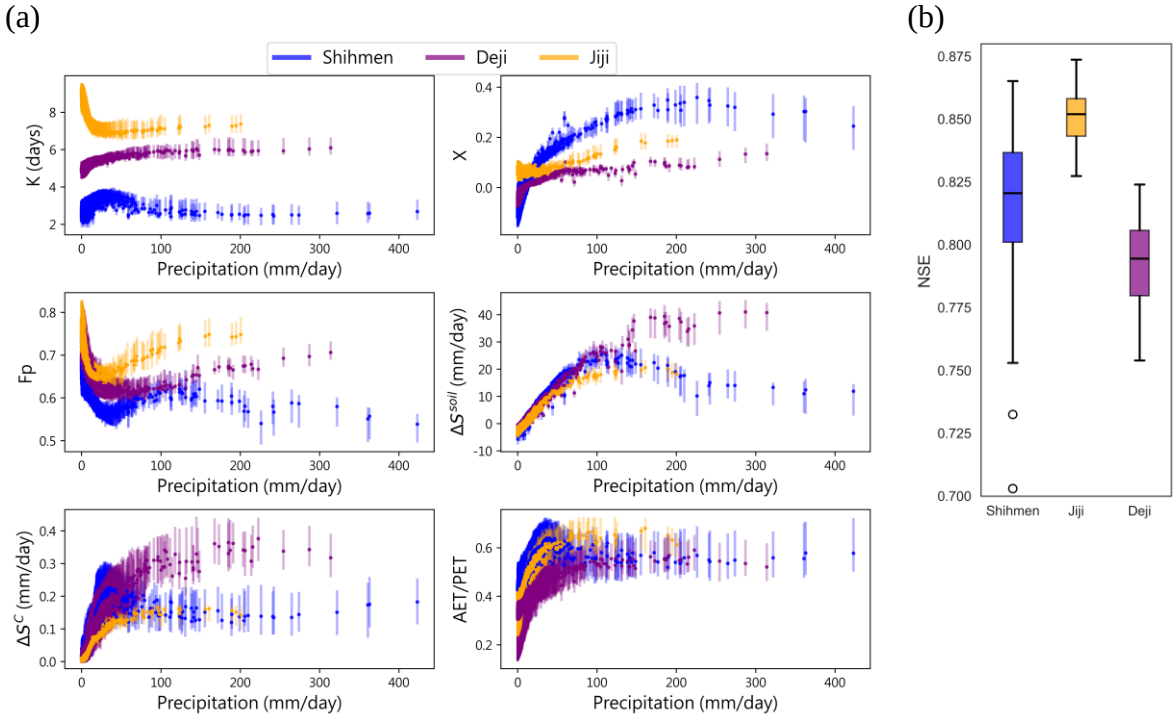


Figure 8. (a) Each functional representation as a function of current precipitation, P. We re-extracted the 30-times functional representations. The Blue, orange, and purple colors represent the case in Shihmen, Jiji, and Deji. We present the bar in the range from 75% to 25% re-extracted functions. The points indicate the median of the 30 times re-extracted representations. (b) provides the 30 re-extracted accuracy results across random seeds 0-29.

Additionally, the unconformity of the parameterized paradigm in hydrological model construction arises because the prior functional form is compensated for parameter optimization, even when it is inappropriate. For instance, when searching for feasible parameter sets, the prior functional form of the specific process does not represent that of the target catchment. The inappropriate form would lead to greater compensation from other functional representations, thereby intensifying the unconformity of the calibrated parameter sets. Therefore, stability of re-extracted representations could be enhanced not only by adding more constraints and data, but also by providing a better representation of prior functional forms, thereby reducing the structural error caused by an inappropriate prior functional form. Particularly, with the proposed functional search to address the closure problem, equifinality persists, translating parameter-set compensation into functional compensation with higher stability of the re-extracted representations and interpretable patterns; in so doing, it advances the credibility of data-driven induction.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

5 Conclusion

We introduced the TDLR framework to train an LSTM to extract functional representations. The TDLR enables LSTM, as a general function approximator, to learn interpretable response functions directly, without imposing prior functional forms. We applied the TDLR to three hillslope-dominance catchments with varying meteorological, hydrological, and morphological conditions, demonstrating that the extracted functional representations can either converge or diverge substantially, providing insight into the similarity and variability of underlying hydrological process controls. Additionally, the results extracted from these three catchments show the strong stability of re-extracted functional representations, enhancing the identifiability of the data-driven method.

Compared with hybrid approaches that boost parameterization within predefined model structures, the functional representations extracted with TDLR capture hydrological responses more effectively in our study area. The results challenge the assumption of transferable, fixed functional forms across catchments, as morphological, geological, meteorological, and hydrological heterogeneity, and their covariation, strongly influence process behavior, particularly in runoff generation, subsurface water movement, and riverine routing. On the other hand, the extracted evapotranspiration and canopy interception functions show greater consistency across the three catchments. Still, the distinct covariance between precipitation and wind speed plays a critical role in breaking the pattern consistency during low-to-moderate precipitation.

The TDLR establishes a hybrid framework that directly extracts the functional representations, revealing a new paradigm of data-driven functional induction for catchment hydrological modeling. Combining the TDLR with conceptual generic buckets achieves highly competitive performance, extracts interpretable and time-invariant response functions that vary across catchments. The proposed framework provides an alternative pathway with mass conservation constraints and conceptual generic buckets to reveal how the meteorological-hydrological-morphological interactions and their heterogeneity shape the hydrological responses.

Code and data availability

The codes and data used to conduct all the analyses in this paper are publicly available at 10.5281/zenodo.19842549

Author contributions

CL conducted the investigation, data curation, formal analysis, validation, visualization, and manuscript preparation, including the original draft. CL, JC, and SH conceived and designed the study. CL and JC developed the machine-learning methodology. SH was responsible for project administration, and funding acquisition. JC and SH share the supervision of the study. SL, SJ, and HW contributed to manuscript revision. All authors participated in discussions throughout the study, provided critical feedback, and approved the final version of the manuscript.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Disclaimer

Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

The authors are grateful for the financial support provided by the National Science and Technology Council (NSTC). Also, thank Qun-Zhan Hung for providing the technical support in writing and figure production.

Financial support

This work was supported by the National Science and Technology Council (NSTC), Taiwan (grant numbers 114-2116-M-002-020-MY3, 114-2222-E-002-014, and 112-2621-M-002-008-MY2).

References

Acuña Espinoza, E., Loritz, R., Álvarez Chaves, M., Bäuerle, N. and Ehret, U. 2024. To bucket or not to bucket? Analyzing the performance and interpretability of hybrid hydrological models with dynamic parameterization. Hydrology and Earth System Sciences 28(12), 2705–2719. doi:10.5194/hess-28-2705-2024

Andrea Rinaldo, I.R.-I. 1996. Geomorphological Theory of the Hydrological Response. Hydrological Processes 10, 803 – 829. doi:10.1002/(SICI)1099-1085(199606)10:6%3C803::AID-HYP373%3E3.0.CO;2-N

Bauser, H.H., Kim, M., Ng, W.R., Bugaj, A. and Troch, P.A. 2022. Richards Equation at the Hillslope Scale: Can We Resolve the Heterogeneity of Soil Hydraulic Material Properties? Water Resources Research 58(12). doi:10.1029/2022wr032294

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Bernier, P.Y. 1985. Variable source areas and storm-flow generation: An update of the concept and a simulation effort. Journal of Hydrology 79(3-4), 195–213. doi:10.1016/0022-1694(85)90055-1

Beven, K. 1989. Changing ideas in hydrology — The case of physically-based models. Journal of Hydrology 105(1-2), 157–172. doi:10.1016/0022-1694(89)90101-7

Beven, K. 2006a. A manifesto for the equifinality thesis. Journal of Hydrology 320(1-2), 18–36. doi:10.1016/j.jhydrol.2005.07.007

Beven, K. 2006b. Searching for the Holy Grail of scientific hydrology: Qt=(S, R, Δt)A as closure. Hydrology and Earth System Sciences 10(5), 609–618. doi:10.5194/hess-10-609-2006

Beven, K.J. and Kirkby, M.J. 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Bulletin 24(1), 43–69. doi:10.1080/02626667909491834

Beven, K.J., Kirkby, M.J., Freer, J.E. and Lamb, R. 2021. A history of TOPMODEL. Hydrology and Earth System Sciences 25(2), 527–549. doi:10.5194/hess-25-527-2021

Blöschl, G. and Sivapalan, M. 2006. Scale issues in hydrological modelling: A review. Hydrological Processes 9(3-4), 251–290. doi:10.1002/hyp.3360090305

Cai, J., Liu, Y., Lei, T. and Pereira, L.S. 2007. Estimating reference evapotranspiration with the FAO Penman–Monteith equation using daily weather forecast messages. Agricultural and Forest Meteorology 145(1-2), 22–35.

Chen, H.-Y., Yeh, H.-F., Ke, C.-C., Yang, Y.-S. and Huang, C.-C. 2024a. Deciphering inter-catchment groundwater flow: A water balance perspective in the Choshui River Basin, Taiwan. Journal of Hydrology: Regional Studies 53. doi:10.1016/j.ejrh.2024.101742

Chen, X., Yu, Z., Yi, P., Chen, P., Hwang, H.-T., Sudicky, E.A. and Simonovic, S.P. 2024b. Water sources and threshold behaviors of streamflow generation in a mountain headwater catchment. Journal of Hydrology 644. doi:10.1016/j.jhydrol.2024.132117

Cheng, C. T., Zhao, M. Y., Chau, K. W., & Wu, X. Y. (2006). Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure. Journal of Hydrology, 316, 129–140. doi:10.1016/j.jhydrol.2005.04.022

Clapp, R.B. and Hornberger, G.M. 2010. Empirical equations for some soil hydraulic properties. Water Resources Research 14(4), 601–604. doi:10.1029/WR014i004p00601

Clark, M.P., Bierkens, M.F.P., Samaniego, L., Woods, R.A., Uijlenhoet, R., Bennett, K.E., Pauwels, V.R.N., Cai, X., Wood, A.W. and Peters-Lidard, C.D. 2017. The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism. Hydrol Earth Syst Sci 21(7), 3427–3440. doi:10.5194/hess-21-3427-2017

Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D., Arnold, J.R., Gochis, D.J. and Rasmussen, R.M. 2015. A unified approach for process-based

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

hydrologic modeling: 1. Modeling concept. Water Resources Research 51(4), 2498–2514. doi:10.1002/2015wr017198

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T. and Morgan, G. 2023. Explainable AI (XAI): Core ideas, techniques, and solutions. ACM computing surveys 55(9), 1–33.

Elrashedy, S., Zhang, Y., Geurink, J.S., Parajuli, K., Wang, H. and Wang, D. 2025. Inter-Basin groundwater flow in West-Central Florida. Journal of Hydrology 660. doi:10.1016/j.jhydrol.2025.133423

Fang, Y. H., Zhang, X., Corbari, C., Mancini, M., Niu, G. Y., & Zeng, W. (2017). Improving the Xin'anjiang hydrological model based on mass–energy balance. Hydrology and Earth System Sciences, 21, 3359–3375. doi:10.5194/hess-21-3359-2017

Feng, D., Liu, J., Lawson, K. and Shen, C. 2022. Differentiable, Learnable, Regionalized Process-Based Models With Multiphysical Outputs can Approach State-Of-The-Art Hydrologic Prediction Accuracy. Water Resources Research 58(10). doi:10.1029/2022wr032404

Frisbee, M.D., Tysor, E.H., Stewart-Maddox, N.S., Tsinnajinnie, L.M., Wilson, J.L., Granger, D.E. and Newman, B.D. 2016. Is there a geomorphic expression of interbasin groundwater flow in watersheds? Interactions between interbasin groundwater flow, springs, streams, and geomorphology. Geophysical Research Letters 43(3), 1158–1165. doi:10.1002/2015gl067082

Gao, X.G., Wang, J.P., Ge, S., Su, S.K., Bai, M.H. and Francois, B. 2024. Investigation of canopy interception characteristics in slope protection grasses: A laboratory experiment. Sci Total Environ 948, 174731. doi:10.1016/j.scitotenv.2024.174731

Garen, D.C. and Moore, D.S. 2005. Curve Number Hydrology in Water Quality Modeling: Uses, Abuses, and Future Directions. Journal of the American Water Resources Association 41(2), 377–388. doi:10.1111/j.1752-1688.2005.tb03742.x

Gharari, S., Gupta, H.V., Clark, M.P., Hrachowitz, M., Fenicia, F., Matgen, P. and Savenije, H.H.G. 2021. Understanding the Information Content in the Hierarchy of Model Development Decisions: Learning From Data. Water Resources Research 57(6). doi:10.1029/2020wr027948

Gnann, S. and Wagener, T. 2026. On Utilizing Spatial Gradients to Discover Functional Relationships in Hydrology. Water Resources Research 62(3). doi:10.1029/2025wr040101

Guang-Te, W. and Singh, V.P. 1992. Muskingum method with variable parameters for flood routing in channels. Journal of Hydrology 134(1-4), 57–76. doi:10.1016/0022-1694(92)90028-t

Gupta, H.V., Clark, M.P., Vrugt, J.A., Abramowitz, G. and Ye, M. 2012. Towards a comprehensive assessment of model structural adequacy. Water Resources Research 48(8). doi:10.1029/2011wr011044

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

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Han, P.-F., Wang, X.-S., Zhou, Y., Yang, Z., Wan, L., Chen, J. and Jiang, X.-W. 2024. Three-dimensional inter-basin groundwater flow toward a groundwater-fed stream: Identification, partition, and quantification. Journal of Hydrology 629. doi:10.1016/j.jhydrol.2023.130524

Hoedt, P.-J., Kratzert, F., Klotz, D., Halmich, C., Holzleitner, M., Nearing, G.S., Hochreiter, S. and Klambauer, G. 2021 MC-LSTM: Mass-Conserving LSTM. Marina, M. and Tong, Z. (eds), pp. 4275—4286, PMLR, Proceedings of Machine Learning Research.

Jiang, S., Zheng, Y. and Solomatine, D. 2020. Improving AI System Awareness of Geoscience Knowledge: Symbiotic Integration of Physical Approaches and Deep Learning. Geophysical Research Letters 47(13). doi:10.1029/2020gl088229

Kirchner, J.W. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research 42(3). doi:10.1029/2005wr004362

Klaassen, W., Bosveld, F. and de Water, E. 1998. Water storage and evaporation as constituents of rainfall interception. Journal of Hydrology 212-213, 36–50. doi:10.1016/s0022-1694(98)00200-5

Klaus, J. and McDonnell, J.J. 2013. Hydrograph separation using stable isotopes: Review and evaluation. Journal of Hydrology 505, 47–64. doi:10.1016/j.jhydrol.2013.09.006

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S. and Nearing, G. 2019. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrology and Earth System Sciences 23(12), 5089–5110. doi:10.5194/hess-23-5089-2019

Le Mesnil, M., Charlier, J.-B., Moussa, R., Caballero, Y. and Dörfliger, N. 2020. Interbasin groundwater flow: Characterization, role of karst areas, impact on annual water balance and flood processes. Journal of Hydrology 585. doi:10.1016/j.jhydrol.2020.124583

Li, Z. and Tian, F. 2025. Derivation and Validation of a Theoretical Canopy Interception Model Based on Raindrop Microphysical Processes. Water Resources Research 61(6). doi:10.1029/2024wr038296

Loritz, R., Hrachowitz, M., Neuper, M. and Zehe, E. 2021. The role and value of distributed precipitation data in hydrological models. Hydrology and Earth System Sciences 25(1), 147–167. doi:10.5194/hess-25-147-2021

Mahoney, D.T., Christensen, J.R., Golden, H.E., Lane, C.R., Evenson, G.R., White, E., Fritz, K., D'Amico, E., Barton, C., Williamson, T., Sena, K. and Agouridis, C. 2023. Dynamics of streamflow permanence in a headwater network: Insights from catchment-scale model simulations. J Hydrol (Amst) 620(A). doi:10.1016/j.jhydrol.2023.129422

Martin-Vide, J. 2004. Spatial distribution of a daily precipitation concentration index in peninsular Spain. International Journal of Climatology 24(8), 959–971. doi:10.1002/joc.1030

McDonnell, J.J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M.L., Selker, J. and Weiler, M. 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resources Research 43(7). doi:10.1029/2006wr005467

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Mosca, E., Szigeti, F., Tragianni, S., Gallagher, D. and Groh, G. 2022 SHAP-based explanation methods: a review for NLP interpretability, pp. 4593–4603.

Partington, D., Brunner, P., Simmons, C.T., Werner, A.D., Therrien, R., Maier, H.R. and Dandy, G.C. 2012. Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model. Journal of Hydrology 458-459, 28–39. doi:10.1016/j.jhydrol.2012.06.029

Peng, T.-R., Lu, W.-C., Chen, K.-Y., Zhan, W.-J. and Liu, T.-K. 2014. Groundwater-recharge connectivity between a hills-and-plains' area of western Taiwan using water isotopes and electrical conductivity. Journal of Hydrology 517, 226–235. doi:10.1016/j.jhydrol.2014.05.010

Peters-Lidard, C.D., Clark, M., Samaniego, L., Verhoest, N.E.C., van Emmerik, T., Uijlenhoet, R., Achieng, K., Franz, T.E. and Woods, R. 2017. Scaling, similarity, and the fourth paradigm for hydrology. Hydrology and Earth System Sciences 21(7), 3701–3713. doi:10.5194/hess-21-3701-2017

Ponce, V.M., Lohani, A.K. and Scheyhing, C. 1996. Analytical verification of Muskingum-Cunge routing. Journal of Hydrology 174(3-4), 235–241. doi:10.1016/0022-1694(95)02765-3

Ruder, S. 2016. An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.

Sahu, G., Mangukiya, N.K. and Sharma, A. 2026. Does MC-LSTM model improve the reliability of streamflow prediction in human-influenced watersheds? Journal of Hydrology 665. doi:10.1016/j.jhydrol.2025.134711

Shrestha, P.K., Samaniego, L., Rakovec, O., Kumar, R. and Thober, S. 2025. A Novel Stream Network Upscaling Scheme for Accurate Local Streamflow Simulations in Gridded Global Hydrological Models. Water Resources Research 61(6). doi:10.1029/2024wr038183

Singh, N.K., Emanuel, R.E., McGlynn, B.L. and Miniat, C.F. 2021. Soil Moisture Responses to Rainfall: Implications for Runoff Generation. Water Resources Research 57(9). doi:10.1029/2020wr028827

Szilagyi, J. 1992. Why can the weighting parameter of the Muskingum channel routing method be negative? Journal of Hydrology 138(1-2), 145–151. doi:10.1016/0022-1694(92)90161-n

Tsai, W.P., Feng, D., Pan, M., Beck, H., Lawson, K., Yang, Y., Liu, J. and Shen, C. 2021. From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling. Nat Commun 12(1), 5988. doi:10.1038/s41467-021-26107-z

Tull, N. and Passalacqua, P. 2025. Flood Wave Attenuation as a Function of Floodplain Storage, Secondary Channel Conveyance, and Discharge. Water Resources Research 61(5). doi:10.1029/2024wr038582

Türk, G., Gey, C.J., Schöne, B.R., Floriancic, M.G., Kirchner, J.W., Leonard, L., Gourdol, L., Keim, R.F. and Pfister, L. 2026. Bedrock geology controls on new water fractions and catchment functioning in contrasted nested catchments. Hydrology and Earth System Sciences 30(2), 343–369. doi:10.5194/hess-30-343-2026

Véliz-Chávez, C., Mastachi-Loza, C.A., González-Sosa, E., Becerril-Piña, R. and Ramos-Salinas, N.M. 2014. Canopy Storage Implications on Interception Loss Modeling. American Journal of Plant Sciences 05(20), 3032–3048. doi:10.4236/ajps.2014.520320

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Wang, H.-J., Merz, R., & Basso, S. 2025. Constructing a geography of heavy-tailed flood distributions: insights from common streamflow dynamics. Hydrol. Earth Syst. Sci. 29(6), 1525-1548. doi.org/10.5194/hess-29-1525-2025

Wang, H.-J., Merz, R., Yang, S., Tarasova, L., & Basso, S. 2023. Emergence of heavy tails in streamflow distributions: the role of spatial rainfall variability. Advances in Water Resources Journal, 171(104359). doi.org/10.1016/j.advwatres.2022.104359

Wang, Y., Zhang, L., Erichson, N.B. and Yang, T. 2025. A Mass Conservation Relaxed (MCR) LSTM Model for Streamflow Simulation Across CONUS. Water Resources Research 61(8). doi:10.1029/2024wr039131

Weiler, M., McDonnell, J.J., Tromp-van Meerveld, I. and Uchida, T. (2005) Encyclopedia of Hydrological Sciences.

Xiong, F., Sun, H., Zhang, Q., Wang, Y. and Jiang, Q. 2022. Preferential flow in three-dimensional stochastic fracture networks: The effect of topological structure. Engineering Geology 309. doi:10.1016/j.enggeo.2022.106856

Yadav, B., Perumal, M. and Bardossy, A. 2015. Variable parameter McCarthy–Muskingum routing method considering lateral flow. Journal of Hydrology 523, 489–499. doi:10.1016/j.jhydrol.2015.01.068

Yang, C., Condon, L.E. and Maxwell, R.M. 2025. Unravelling groundwater–stream connections over the continental United States. Nature Water 3(1), 70–79. doi:10.1038/s44221-024-00366-8

Zehe, E. and Sivapalan, M. 2009. Threshold behaviour in hydrological systems as (human) geo-ecosystems: manifestations, controls, implications. Hydrology and Earth System Sciences 13(7), 1273–1297. doi:10.5194/hess-13-1273-2009

Zhang, L., Potter, N., Hickel, K., Zhang, Y. and Shao, Q. 2008. Water balance modeling over variable time scales based on the Budyko framework – Model development and testing. Journal of Hydrology 360(1-4), 117–131. doi:10.1016/j.jhydrol.2008.07.021

Zhao, J., Sun, S., Yin, Y., Tang, Y., Li, C., Liang, Y., Wang, Y., Winkler, A. and Jiang, S. 2025. Advancing evapotranspiration modeling with optimized soil and canopy resistance combinations. Water Resources Research 61(6), e2024WR039252.

Zoccatelli, D., Borga, M., Viglione, A., Chirico, G.B. and Blöschl, G. 2011. Spatial moments of catchment rainfall: rainfall spatial organisation, basin morphology, and flood response. Hydrology and Earth System Sciences 15(12), 3767–3783. doi:10.5194/hess-15-3767-2011

Appendix A: The Current Precipitation Dominance in Current Hydrological Responses

We provide results from deep learning extraction with competitive simulation accuracy. Figure A1 (a) shows the case in the Shihmen reservoir catchment. The constitutive relationship between the processes related to evapotranspiration ($k_t^c + k_t^{soil}$), runoff generation (F_{pt}), canopy interception (ΔS_t^c), change of soil water bucket (ΔS_t^{soil}), and riverine routing (K_t and x_t) and sequence of driving forces. The results for the other two catchments, Jiji dam and Deji reservoir, are shown in Fig. A1(b) and A1(c), respectively. The y-axis of each subplot in Fig. 4 represents different hydrological functional representations. The x-

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

axis of each subplot serves as the meteorological driving force. Remarkably, daily cumulative precipitation, as the only input value (highlighted in red in Figure A1), provides the dominant information for constructing all functional representations; the second and third are relative humidity and net radiation.

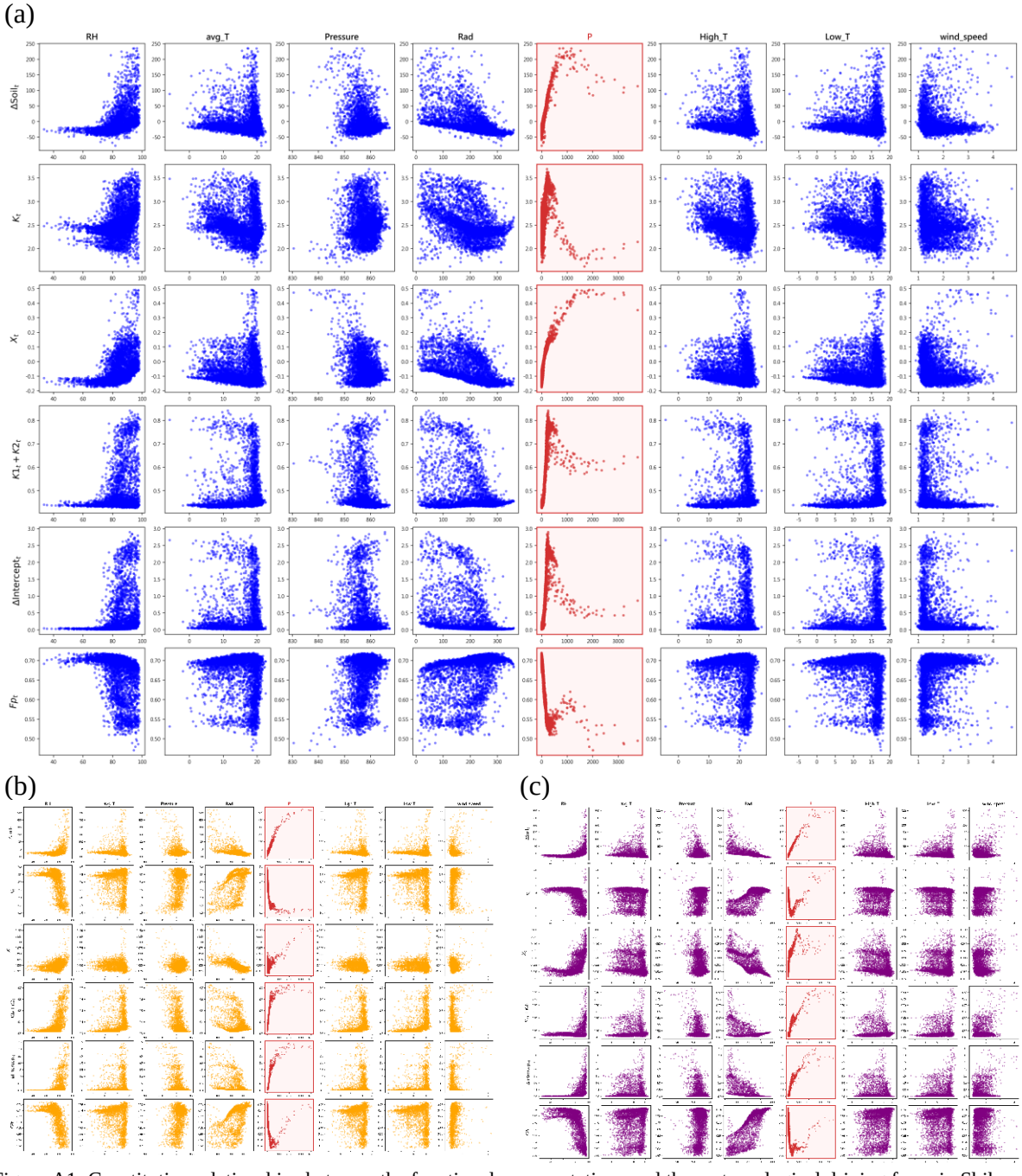


Figure A1: Constitutive relationships between the functional representations and the meteorological driving force in Shihmen reservoir catchment (blue points), Jiji dam catchment (yellow points), and Deji reservoir catchment (purple points). The y-axis of each subplot in

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

Figure A1 represents different hydrological functional representations. The x-axis of each subplot serves as the meteorological driving force. The meteorological driving forces include RH (relative humidity), avg_T (average temperature), Pressure (atmospheric pressure), Rad (net radiation), P (precipitation), High_T (highest temperature), Low_T (lowest temperature), and wind_speed (average wind speed). The functional representations include change of soil water bucket (ΔS_t^{soil}), riverine routing (K_t and x_t), evapotranspiration ($k_t^c + k_t^{soil}$), canopy interception (ΔS_t^c), and infiltration (F_{pt}).

Appendix B: Extracted functional representation: evaporation from the canopy layer

In Shihmen, in the low to moderate precipitation, the evaporation from AET increases along with the covariance of the precipitation-wind speed shown in Figure 1 (d). After 50 mm/day, precipitation and wind speed decrease; as humidity decreases, PET decreases as well. The case of Jiji shows a similar pattern in low-to-moderate precipitation: the AET is slightly decreasing, while the PET decreases in high precipitation. Lastly, the Deji case indicates an increasing trend rather than a plateau in evaporation. Hence, during low to moderate precipitation, wet evaporation dominated canopy interception, but as precipitation increased, throughflow and stemflow took over.

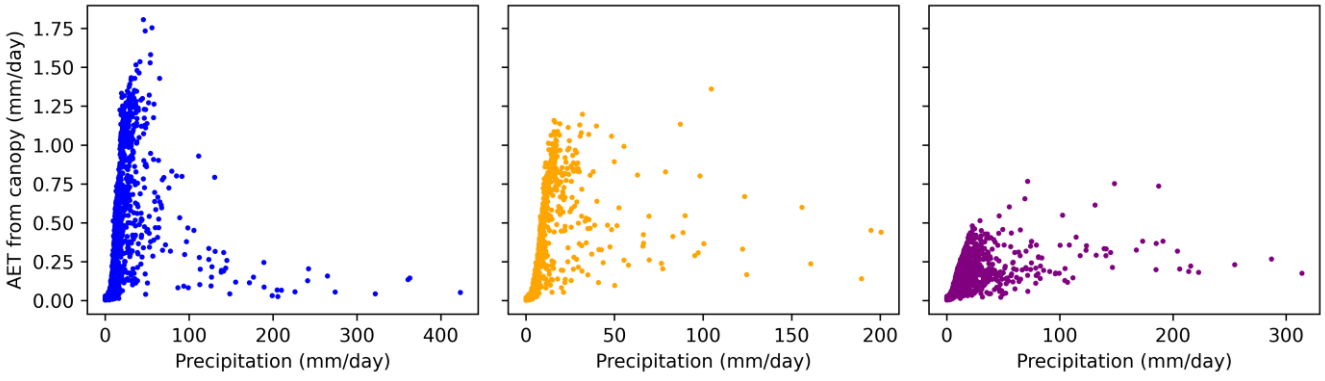


Figure B1: With the extracted functional representation, k_t^c, and the potential evapotranspiration. Blue, orange, and purple are the extracted cases in Shihmen, Jiji, and Deji, respectively.

Appendix C: Curve fitting method for projected relationships in Figure 5

To effectively visualize the underlying physical trends in highly non-uniform scatter plots, characterized by dense, low-variance clusters (low to moderate precipitation) and sparse observation tails (heavy precipitation), a piecewise data aggregation technique was implemented. For each projected relationship in Figure 5, the independent variable was partitioned using a predefined, variable-specific threshold. Within the high-density region below this threshold, the data was discretized into fixed-width bins. To extract the central tendency and mitigate the computational and visual bias of overplotting, the mean of the dependent variable was calculated for each bin containing a minimum of two valid observations. These aggregated bin centers and means were then concatenated with the raw, unbinned observations from the sparse region above the threshold, forming a reformed dataset designed specifically to stabilize the optimization process.

https://doi.org/10.5194/egusphere-2026-2483
Preprint. Discussion started: 20 May 2026
© Author(s) 2026. CC BY 4.0 License.

To capture the complex non-linear behaviors of the catchment response, non-linear least-squares curve fitting was applied to the reformed dataset. This approach allowed for the fitting of highly specific, non-linear analytical forms, such as exponential decay, gamma distributions, and piecewise overshoot functions, bounded by predefined parameter constraints. By anchoring the optimization to the binned means in the dense region and the individual raw points in the sparse tail, the algorithm successfully maps the trajectory of the physical processes without being disproportionately anchored by the sheer volume of data in the low-variance clusters.