https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

Diagnosing Dissolved Organic Carbon Simulation of SWAT-C model Using Machine Learning Approaches

Zehong Huang[1], Shouzhi Chen[1], Yufeng Gong[1], Zheng Wang[2], Zheng Duan[3], Yongshuo H. Fu[1]

[1]College of Water Sciences, Beijing Normal University, Beijing, China
[2]College of Urban and Environmental Sciences, Central China Normal University, Wuhan, China
[3]Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, SE-223 62 Lund, Sweden

Correspondence to: Yongshuo H. Fu (Email: yfu@bnu.edu.cn);

Abstract. Dissolved organic carbon (DOC) plays a critical role in the terrestrial carbon cycle, and accurate simulation of its dynamics is essential for understanding carbon balance and climate change mitigation. However, DOC simulations still involve large uncertainty under complex environmental conditions. To address this challenge, we proposed a Module Diagnosis Framework (MDF) that quantitatively identifies the module-level sources of uncertainty in DOC modeling. The SWAT-MDF integrates the physically based SWAT-Carbon (SWAT-C) model with a data-driven module that employs machine learning algorithms and applies Shapley additive explanations (SHAP) and residual analysis to diagnose the uncertain source of DOC simulation in the Yalong River Basin.We found that the data-driven module based on bidirectional long short-term memory (Bi-LSTM) networks achieved good performance for daily DOC predictions with an average *NSE* = 0.62, R^2 = 0.67, *KGE* = 0.74 while the original SWAT-C model yielded average *NSE* = 0.51 and R^2 = 0.61, *KGE* = 0.55. Despite this improvement, the testing performance remains limited, suggesting that the main uncertainty arises from the structural limitations of SWAT-C and highlighting the need for further structural improvement and module-level diagnosis. The MDF results revealed that the carbon cycle module and pollutant transport module mainly regulated the magnitude and variation of DOC predictions in the original SWAT-C model, and the vegetation growth module and the carbon cycle module were major sources of DOC prediction uncertainties. We therefore proposed that further improvements in DOC prediction in the SWAT-C model should focus on the vegetation growth, carbon cycle modules. Our proposed SWAT-MDF framework significantly enhances the reliability of DOC simulations, and provides a

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

quantitative basis for improving the SWAT-C model and offers a generalizable approach to module optimization in similar coupled modeling frameworks.

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

1 Introduction

Dissolved organic carbon (DOC) is a critical component of the terrestrial carbon-water cycle and constitutes the primary form of carbon exported from watersheds to stream and river system (Evans et al., 2005). It accounts for over 50% of the total terrestrial organic carbon flux entering the oceans (Cole et al., 2007), underscoring its importance in the global carbon cycle. The dynamic behavior of DOC is closely linked to substantial changes in terrestrial carbon storage. Although DOC is not considered a pollutant itself, its excessive presence can facilitate the co-transport of toxic substances such as heavy metals through complexation processes, thereby adversely affecting surface water quality (Harkort and Duan, 2023; Lawlor and Tipping, 2003). Consequently, accurately quantifying DOC dynamics is essential for understanding carbon input and output processes at the watershed scale and provides crucial data support for pollution policy-making and climate change adaptation (Du et al., 2023).

Currently, watershed-scale DOC monitoring predominantly relies on water-quality sampling. However, the limited spatial distribution of fixed monitoring stations and the lack of long-term observational data hinder comprehensive characterization of DOC's spatiotemporal variability (Wang et al., 2025). As a result, physically-based modeling approaches have become indispensable tools in DOC studies. Among these, the Soil and Water Assessment Tool (SWAT) has been widely utilized to simulate hydrological and water quality dynamics under various land-use and climate change scenarios (Srinivasan et al., 1998). Yang and Zhang, (2016) developed a carbon cycling module for SWAT (Zhang et al., 2013), while Du et al., (2019) extended this framework by incorporating updated algorithms to simulate DOC transport and transformation processes (Qi et al., 2020b). The enhanced SWAT-C model has demonstrated promising potential for watershed-scale DOC simulation, yielding relatively accurate results in diverse basins, such as forested watersheds (Lee et al., 2025) and agricultural watersheds (Qi et al., 2020a). However, when the model is applied to new regions, it still faces considerable challenges. The pronounced heterogeneity in land use, topography, and climatic conditions across watersheds necessitates parameter regionalization. Nevertheless, the characterization of watershed features during regionalization remains insufficiently defined (Liu et al., 2024; Razavi and Coulibaly, 2013) which can lead to reduced model performance. This is particularly evident under

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

complex environmental conditions where SWAT-C model tends to underestimate high flows and DOC peak fluxes (Qi et al., 2020a). In addition, incomplete module designs and vague parameter representations exacerbate prediction uncertainties. Existing studies have rarely conducted quantitative, module-level diagnosis; instead, they typically rely on subjective model adjustments, which may lead to inconsistent calibration and unreliable results..

Therefore, developing quantitative module diagnosis frameworks is essential to identify and improve critical sub-modules affecting DOC prediction, thereby enhancing the robustness and generalizability of the coupled modeling approach.

In recent years, machine learning (ML) has demonstrated substantial potential in hydrological modeling, particularly for capturing complex nonlinear interactions among variables (Fan et al., 2020). As data-driven tools, ML models can autonomously learn latent relationships without requiring extensive prior knowledge, and they often outperform traditional process-based models in terms of simulation accuracy, generalizability, and robustness (Lee et al., 2023a; Yao et al., 2023). However, the "black-box" nature of ML models limits their interpretability, raising concerns in environmental decision-making and policy implementation contexts (Carvalho et al., 2019). To address the performance limitations of physically-based models under complex environmental conditions and the limited interpretability of ML approaches, hybrid modeling frameworks have been proposed. Recent studies have integrated hydrological and machine learning models to explore the dominant factors influencing streamflow variability, demonstrating the advantages of combining physical-based and data-driven approaches in hydrological modeling (Ding et al., 2025). In addition, a SWAT-ELM-SHAP hybrid modeling framework has been proposed to improve both predictive accuracy and interpretability of blue and green water simulations in data-scarce watersheds by incorporating ensemble learning and model explainability techniques (Guo et al., 2024). In this framework, the Extreme Learning Machine (ELM), a single-layer feedforward neural network known for its fast training speed, is used to predict blue and green water components. Shapley additive explanations (SHAP) are applied to interpret feature contributions, thereby enhancing model transparency.

Despite recent advancements, accurate simulation of DOC remains challenging due to the complexity of watershed environmental conditions, limited long-term monitoring data, and

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

inadequately developed model modules. The SWAT-MDF quantitatively diagnoses module-level sources of uncertainty in SWAT-C and evaluates their relative contributions under complex environmental conditions. To support this diagnostic process, the data-driven module compares eight algorithms to identify the optimal approach for reproducing the nonlinear dynamics of DOC. Subsequently, Shapley Additive Explanations (SHAP) and residual analysis are applied to interpret this representation, quantify feature and module contributions, and pinpoint sub-modules that dominate simulation uncertainty (Lundberg et al., 2020). The main objective of this study is to establish a quantitative diagnosis of module-level uncertainties in DOC simulations using the SWAT-MDF framework and to provide a scientific basis for targeted model refinement.

2. Materials and methods

2.1. Study area

The Yalong River Basin, located in Southwest China (Fig. 1), is the largest tributary of the Jinsha River, which constitutes the upper reaches of the Yangtze River. Originating from the southern slopes of the Bayan Har Mountains on the Qinghai Tibet Plateau, the river traverses Qinghai Province, Sichuan Province, and the Tibet Autonomous Region. Geographically, the basin extends from 96º52' to 102º48'E and 26º32'to 33º58'N, spanning approximately 1,571 km in length and covering an area of about 136,000 km^2. The basin features highly variable terrain, with an average elevation exceeding 3,000 meters (Liu et al., 2019).

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

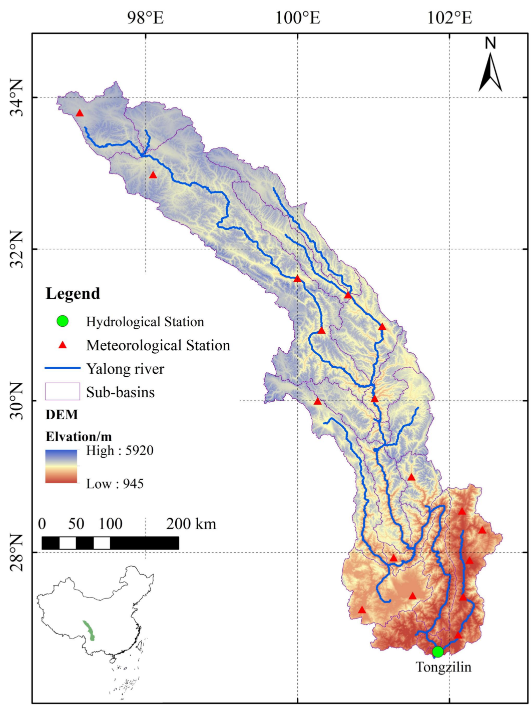

Figure 1. Overview of the Yalong River Basin. The inset at the lower left shows the basin's location within China.

The average annual runoff of the basin is 1,914 m^3/s, and the average annual export of suspended particulate matter reaches approximately $2.55*10^{10}$ kg. Precipitation varies significantly across regions—ranging from 500-800 mm in the headwaters to 900-1,300 mm in the lower reaches, with certain midstream areas receiving as much as 1,000-1,800 mm annually (Li et al., 2014). Influenced by the monsoon climate, the basin experiences distinct seasonal hydrological variation: the wet season spans from May to October, characterized by concentrated rainfall and runoff with pronounced river flow fluctuations; the dry season occurs from November to April, with markedly reduced runoff. Additionally, snowmelt contributes to runoff between March and June each year, supplementing rainfall inputs (Zhao et al., 2021). In this study, the Tongzilin Hydrological Station, located in the lower reaches of the Yalong River, was selected for analysis. Its runoff, sediment, and DOC monitoring data are highly representative and reliable, providing robust support for the investigation of hydrological processes, DOC dynamics, and model optimization in the basin.

2.2. Data source

The SWAT-C model developed in this study was constructed using a suite of input datasets, including meteorological, topographic, soil, and land use information, as summarized in Tab. 1. Meteorological data spanning from 1970 to 2020, including precipitation, relative humidity, maximum and minimum air temperature, solar radiation and wind speed, were obtained from eleven stations (Tab. 1) archived by the China Meteorological Data Service Center. Model calibration and validation were performed using observed runoff, sediment, and DOC data. Daily runoff and sediment observations at the Tongzilin Hydrological Station in the Yalong River Basin for the period 2013-2020 were sourced from the Hydrological Yearbook of the People's Republic of China. DOC observations for 2013-2014 and 2019-2020 were derived from field measurements conducted by Xu et al., (2024) at the same station. Description of DEM, soil data and land use map.

Table 1 .Data Sources for the study.

Data Type	Description	Source
Meteorological	Daily station-based	China Meteorological Data Service Center (https://data.cma.cn/）
DEM	ASTER GDEM 30M	Geospatial Data Cloud (https://www.gscloud.cn/)
Soil Data	HWSD Global Soil Database	Institute of Soil Science, Chinese Academy of Sciences (http://english.issas.cas.cn/)
Land Use Map	2014 land use distribution	CLCD dataset, Wuhan University （http://doi.org/10.5281/zenodo.4417809）
Runoff Data	Daily observations	Hydrological Yearbook of the People's Republic of China
Sediment Data	Daily observations	Hydrological Yearbook of the People's Republic of China
DOC Data	Daily observations	Escalating Carbon Export from High-Elevation Rivers in a Warming Climate (Xu et al., 2024)

2.3. The SWAT-MDF Coupled Modeling Framework

The SWAT-MDF diagnostic framework developed in this study quantitatively identifies and interprets the module-level sources of uncertainty in DOC simulations. It integrates three main components (Fig. 2): the SWAT-C model, a data-driven module, and a module diagnosis framework based on SHAP and residual analysis. These components are connected in a sequential structure, in which outputs from each stage serve as inputs for the subsequent stage, enabling a complete workflow encompassing initial simulation, data-driven module, and comprehensive module diagnosis. Specifically, multiple calibrated SWAT-C outputs (feature parameters) are treated as diagnostic variables that help trace uncertainty propagation through the coupled system. Subsequently, SHAP interpretation and residual analysis are applied to the data-driven results to quantify the contributions of different SWAT-C modules and identify the dominant sources of structural uncertainty within the model.

The first component (Fig. 2a), SWAT-C model, provides the physical and mechanistic foundation of the diagnostic framework (Section 2.3.1). Its simulations are driven by spatial datasets such as digital elevation model (DEM), soil type, land use and slope, as well as meteorological variables including precipitation, relative humidity, maximum and minimum temperature, wind speed and solar radiation. Upon model calibration (More details are provided in Text S1 in the Supplementary Materials.), a subset of key characteristic parameters that govern DOC dynamics within the watershed was identified. These parameters, each linked to specific physical modules in the SWAT-C structure, are used as feature inputs for the second component of the framework. Tab. 2 provides detailed definitions of these parameters and their associated modules within SWAT-C model.

The second component (Fig. 2b) is a data-driven module that establishes nonlinear relationships between SWAT-C simulations and observed DOC data to support subsequent diagnostic analysis. In this module, several ML algorithms are employed to learn process-dependent relationships and generate data representations for interpretability analysis in the next stage. Eight ML algorithms were implemented (Section 2.3.2): convolutional neural network (CNN), support vector regression (SVR), extreme gradient boosting (XGBOOST), multilayer perceptron (MLP), and four variants of long short-term memory (LSTM)—vanilla LSTM (V-LSTM), bidirectional LSTM (Bi-LSTM), stacked LSTM (St-LSTM), and convolutional LSTM (Cv-LSTM)(More details are provided in Texts S2 and S3 in the Supplementary Materials.). A transfer learning strategy was implemented, wherein the models were initially pretrained using long-term simulation outputs from SWAT-C model and subsequently fine-tuned with observed DOC data. Bayesian optimization was applied to ensure model robustness and stability for the subsequent SHAP-based and residual analyses.

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

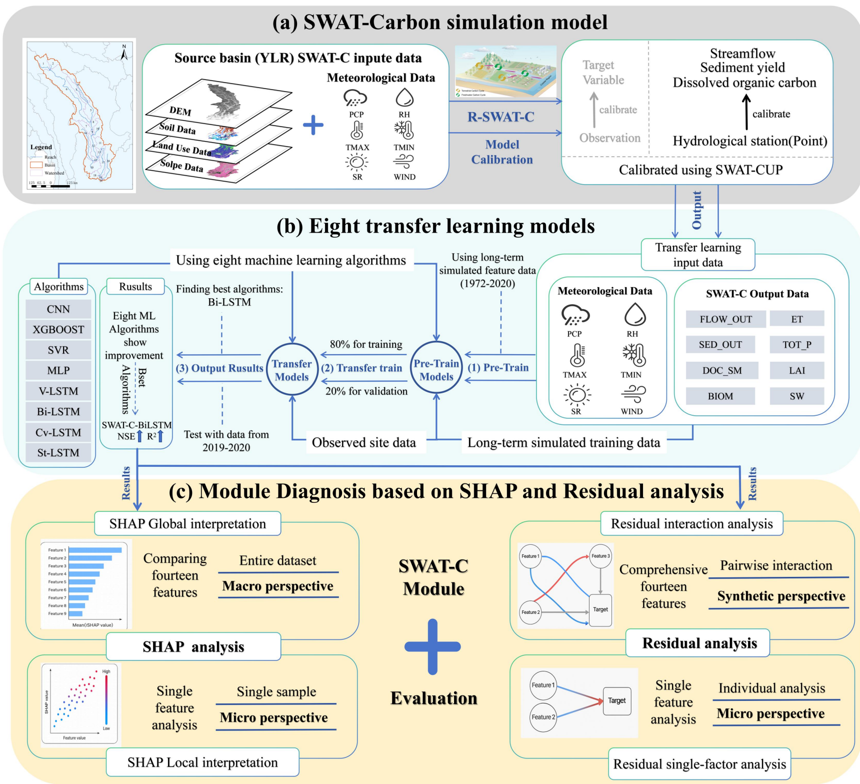

Figure 2. The SWAT-MDF Coupled Model Technical Framework. The computational and diagnosis procedures are illustrated from top to bottom: (a) SWAT-Carbon simulation module; (b) Data-drivem model; (c) Module diagnosis Using SHAP and Residual Analysis.

The third component (Fig. 2c) diagnosis SWAT-C modules using SHAP interpretation and residual analysis (Section 2.3.3). SHAP quantifies both global and local contributions of input features, improving model interpretability. The residual contribution rate is employed to measure each module's relative share of total prediction error, providing an objective comparison of module reliability. By combining SHAP-based feature attribution with residual decomposition, the SWAT-MDF framework pinpoints modules that dominate structural uncertainty and provides a quantitative basis for targeted model refinement. This diagnostic approach shifts the analytical focus from empirical calibration toward mechanistic understanding and structural evaluation of process-based DOC simulations.

Table 2. SWAT-C model Feature Parameters

Data Name	Definition	SWAT-C Module

FLOW_OUT	Streamflow output	Runoff module
SED_OUT	Sediment output	Sediment module
DOC_Simulate	Simulated total DOC	Carbon cycle module
BIOM	Biomass	Biomass module
ET	Evapotranspiration	Evapotranspiration module
TOT_P	Total phosphorus output	Pollutant transport module
LAI	Leaf area index	Vegetation growth module
SW	Soil water content	Soil moisture module
PCP. RH, TMAX, TMIN, SR, WIND	Precipitation, Relative humidity, Max/Min Temperature, Solar radiation, Wind speed	Meteorological forcing module

Note: The specific locations of the SWAT-C modules in the source code are provided in Table S4 in the Supplementary Materials. All feature parameters were derived from the calibrated SWAT-C model.

2.3.1. SWAT-C Model

The SWAT is a physically based, semi-distributed hydrological model capable of continuous simulation. It has been extensively applied in watershed-scale hydrology and water quality studies for replicating the rainfall–runoff process (Rathjens et al., 2015; Srinivasan et al., 1998). By incorporating multi-source inputs such as climate, soil, and land use data, SWAT provides a robust platform for quantitatively assessing the impacts of climate change and anthropogenic activities on watershed hydrology and environmental conditions (Tan et al., 2019).

In the SWAT-MDF coupled modeling framework, we adopted an enhanced version of SWAT known as SWAT-C, which integrates the CENTURY biogeochemical model to enable dynamic simulation of soil organic matter and carbon processes (Zhang et al., 2013). Yang and Zhang (2016) demonstrated the model's reliability for simulating key forest ecosystem variables, such as evapotranspiration and net primary productivity. Building on this, Du et al., (2019) expanded the model's functionality to couple terrestrial and aquatic carbon cycling, facilitating watershed-scale simulation of DOC transport. The current SWAT-C model, built upon the SWAT2012 framework, has demonstrated strong simulation performance in snowmelt-dominated basins (Grusson et al., 2015).

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

Through the innovative integration of core algorithms from water quality models such as QUAL2K and CE-QUAL-W2 (Pelletier et al., 2006), SWAT-C has established a comprehensive carbon cycle simulation system. This system includes modules for soil-derived dissolved organic carbon (DOC) generation and transport, as well as the biogeochemical cycling of DOC within river channels (Lee et al., 2025). A full description of the spatial setup and parameterization is provided in Text S1 in the Supplementary Materials.

2.3.2. Data-driven Models

The data-driven component of the SWAT-MDF framework is implemented through a set of machine learning (ML) models. ML methods, characterized by data-driven modeling strategies, circumvent the dependence of traditional hydrological models on subjective prior knowledge. Their strong capacity for nonlinear fitting has introduced a new paradigm for hydrological simulations (Lee et al., 2023b; Todini, 2007). To address DOC simulation requirements in the Yalong River Basin, this study reviewed and compared commonly used ML approaches in hydrological contexts, see Text S2 in the Supplementary Materials. Eight ML methods were ultimately selected for the data-driven module: CNN, SVR, XGBOOST, MLP, and four LSTM variants, including V-LSTM, St-LSTM, Bi-LSTM, and Cv-LSTM.

DOC observations from 2013-2014 were used as training data, with 20% randomly withheld for validation, while observations from 2019-2020 were used for testing. Given the limited availability of DOC measurements, a transfer learning strategy was adopted. First, models were pretrained on long-term simulated feature parameters from the SWAT-C model (1972-2020) to establish fundamental relationships between inputs and DOC outputs. Second, the pretrained models were fine-tuned using observed data to better match the true data distribution. This process mitigated overfitting caused by data scarcity by leveraging generalized patterns learned from simulations. Text S3 in the Supplementary Materials details the principles of the selected ML methods and the procedures for Bayesian-based hyperparameter optimization.

2.3.3. Calibration and Module Diagnosis of the Coupled Model

This study conducted preliminary simulations of runoff, sediment, and DOC in the Yalong River

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

Basin using the SWAT-C model. Daily outputs were calibrated and validated using DOC observations from 2013 to 2020, together with runoff and sediment data from the Tongzilin Station. The DOC calibration (2013-2014) and validation (2019-2020) periods included 34 and 38 samples, respectively, with consistent periods applied to runoff sediment. Model performance was evaluated using the Nash-Sutcliffe efficiency (*NSE*), coefficient of determination (R^2) and Kling-Gupta efficiency (*KGE*). Model parameters were automatically calibrated using the SUFI-2 algorithm in SWAT-CUP (Abbaspour et al., 2017). Detailed parameter values and sensitivity analysis results are presented in Text S1 in the Supplementary Materials.

To interpret DOC predictions from the coupled model, SHAP was employed. Rooted in cooperative game theory, SHAP considers each input feature as a contributor, with the SHAP value quantifying the marginal contribution of a feature to a specific prediction (Lundberg et al., 2020). Two strategies were adopted: (1) global interpretation, using the mean SHAP values across all samples to rank feature importance; (2) local interpretation, evaluating the SHAP values for individual predictions to determine the direction and magnitude of each feature's influence.

Residuals, which are defined as the differences between predicted and observed values, were further analyzed to assess the model bias and predictive uncertainty. To quantify the influence of input features on prediction errors, residuals were computed for each sample, and polynomial feature expansion was applied to generate nonlinear interaction terms. Ridge regression was subsequently used to explore the relationship between these terms and the residuals (Santos Nobre and Da Motta Singer, 2007; Tyagi et al., 2022). The absolute values of the regression coefficients served as indicators of each feature's contribution to the residuals (Eq. 1). This allowed for the quantification of both individual feature effects and interaction-driven contributions to model errors. To evaluate the robustness of these estimates, a bootstrapping approach was used to repeatedly resample the training dataset (Hongyi Li and Maddala, 1996), and the mean and standard deviation of each feature's contribution rate were computed to characterize uncertainty. Two residual analysis strategies were adopted: (1) Interaction-based: capturing nonlinear feature interactions via polynomial expansion and ridge regression; (2) Single-feature: assessing individual feature contributions without interactions.

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

$$C_j = \frac{|\beta_j|}{\sum_{i=1}^{n}|\beta_i|} \times 100\% \tag{1}$$

where C_j denotes the contribution rate of the j-th feature to the residuals, expressed as a percentage; β_j denotes the ridge regression coefficient corresponding to the j-th feature; n denotes the total number of input features.

3. Results

3.1. Model performance of SWAT-C Model in DOC simulations

The first component of the coupled model, the SWAT-C model, achieved excellent performance in simulating daily runoff and sediment for the Yalong River Basin (Fig. 3). During the calibration period, the SWAT-C model achieved high accuracy for daily runoff (calibration: NSE = 0.91, R^2 = 0.92, KGE = 0.94; validation: NSE = 0.90, R^2 = 0.93, KGE = 0.85). For sediment, both calibration and validation phases yielded excellent results (calibration: NSE = 0.92, R^2 = 0.92, KGE = 0.95; validation: NSE = 0.89, R^2 = 0.93, KGE = 0.79). In contrast, the model's ability to simulate daily DOC was relatively limited, with lower performance (calibration: NSE = 0.59, R^2 = 0.59, KGE = 0.70; validation: NSE = 0.42, R^2 = 0.63, KGE = 0.40). Specifically, the model consistently overestimated DOC concentrations during both high-flow and low-flow periods, resulting in an overall positive bias. Additionally, the simulated seasonal pattern showed a slight temporal mismatch, with peak DOC concentrations in some times occurring later than observed.

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

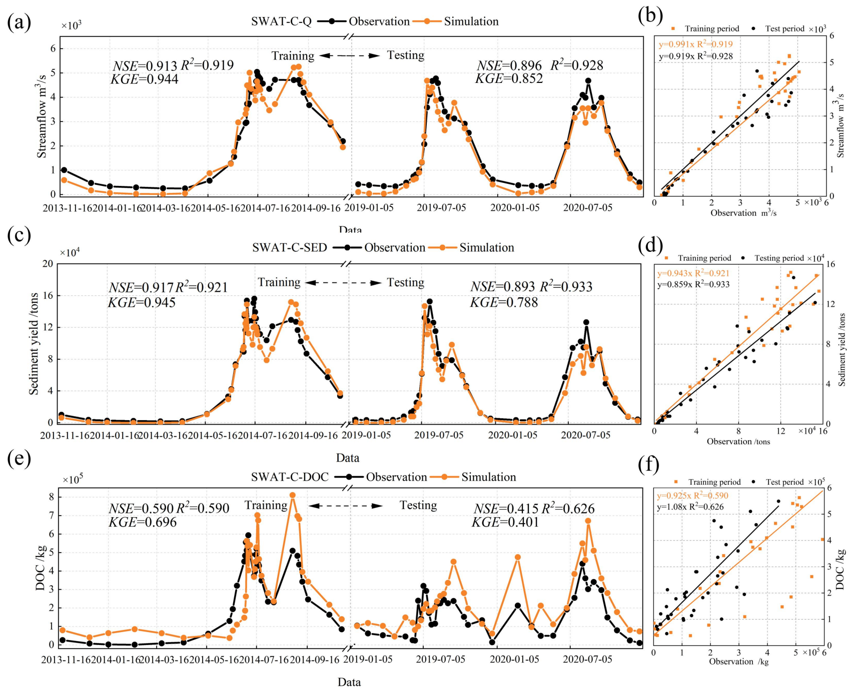

Figure 3. Simulation performance of the SWAT-C model. Line plots (a) streamflow (Q), (c) sediment yield (SED), and (e) dissolved organic carbon (DOC). Scatter plots (b), (d), and (f) show the correlations between observed and simulated values.

3.2. Comparison of Data-driven Module Performance

The performance of the data-driven module in the SWAT-MDF framework was evaluated using eight algorithms to simulate daily DOC dynamics(Fig. 4d). Among these, the Bi-LSTM based data-driven model achieved the best performance (training: NSE = 0.76, R^2 = 0.78, KGE = 0.80, testing: NSE = 0.48, R^2 = 0.55, KGE = 0.67). With the exception of MLP and CNN, most algorithms reproduced DOC variations more effectively than the calibrated SWAT-C model (training: NSE = 0.59, R^2 = 0.59, KGE = 0.70, testing: NSE = 0.42, R^2 = 0.63, KGE = 0.40), though the degree of improvement varied among methods. Violin plots of DOC predictions, together with NSE analysis, indicate that St-LSTM and Bi-LSTM exhibit minimal differences between the training and testing phases, demonstrating relatively strong generalization, robustness, and model stability. In contrast, although XGBOOST, MLP, and SVR performed well during training, they exhibited significant biases

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

and notable drops in performance on the test set. Notably, the Bi-LSTM model effectively captured both short- and long-term dependencies in the data, enabling better simulation of the complex nonlinear dynamics of DOC.

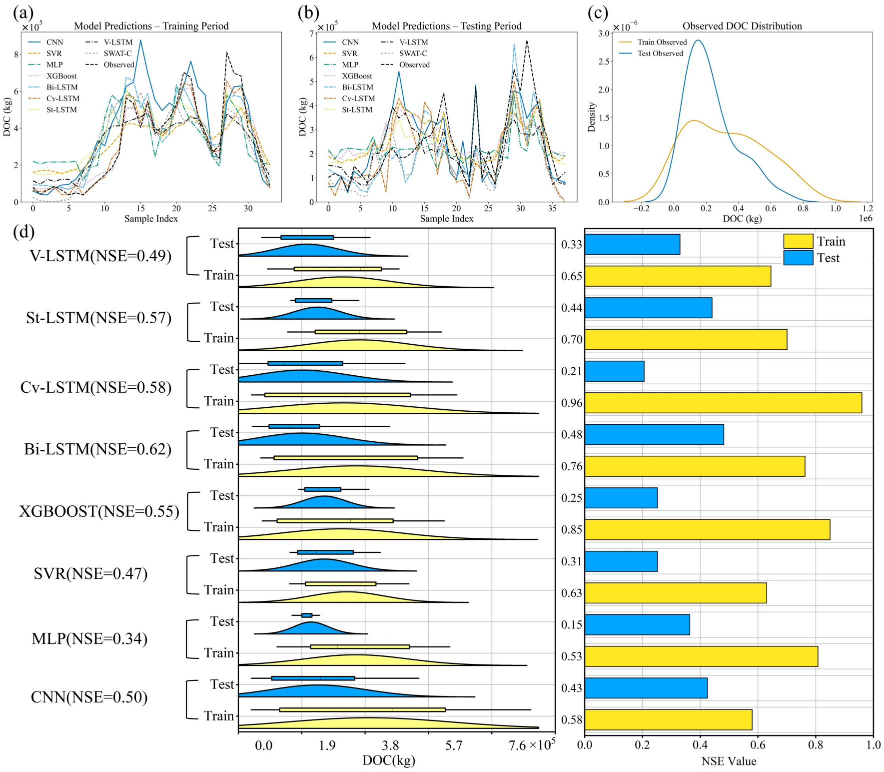

Figure 4. Performance comparison of eight algorithms in DOC prediction. (a) and (b) present time series plots of simulated versus observed DOC values during the training and testing periods, respectively. (c) shows the kernel density estimation of observed DOC distributions for both datasets.(d) compares model-specific distributions of predicted DOC using violin plots (left), alongside their corresponding NSE values (right) for both training and testing periods.

The results of the SWAT-MDF-BiLSTM model highlights the advantage of using physically based features derived from SWAT-C. These features helped the Bi-LSTM model capture key watershed processes and learn complex relationships between input variables and DOC output. Although the integration improved the model's predictive ability, the testing results were still limited, suggesting the model struggles to fully capture DOC dynamics. This finding underscores the necessity of subsequent module-level diagnosis within the SWAT-MDF framework to identify and address these structural

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

limitations. Overall, the integration of data-driven modeling with process-based information provides a reliable foundation for uncertainty attribution and module diagnosis in complex watershed systems.

3.3. SHAP and Residual-Based Model Diagnosis

3.3.1. Global Interpretation of Input Features

Based on the comparison of Data-driven Module Performance in Section 3.2, the SWAT-MDF-BiLSTM model, which demonstrated the highest performance, was selected for further analysis. In this coupled framework, global SHAP interpretation was employed to identify the most influential input features contributing to DOC prediction. Fig. 5 illustrates the global feature importance ranked by the mean absolute SHAP values. This analysis facilitated the identification of dominant hydrological and ecological modules driving DOC simulation and provided insights into how the coupled model captures DOC dynamics within the watershed.

Features with a global importance greater than 5% were defined as key contributors. According to the results (Fig. 5), DOC_Simulate, TOT_P, and PRE were identified as the top three features, accounting for 59.34%, 8.78%, and 6.56% of the total importance, respectively, while other features exhibited relatively low contributions. DOC_Simulate corresponds to the Carbon cycle module, which governs carbon inputs, transformations, cycling, and export-directly shaping DOC behavior. As such, it exerted a substantially stronger influence than other modules. TOT_P reflects the pollutant transport module, which represents the effects of pollution sources on water quality and influences DOC dynamics through pollutant loading and mixing processes. PRE denotes precipitation, which regulates runoff generation, soil water content, and evapotranspiration, thereby modifying DOC transport and concentration across the watershed.

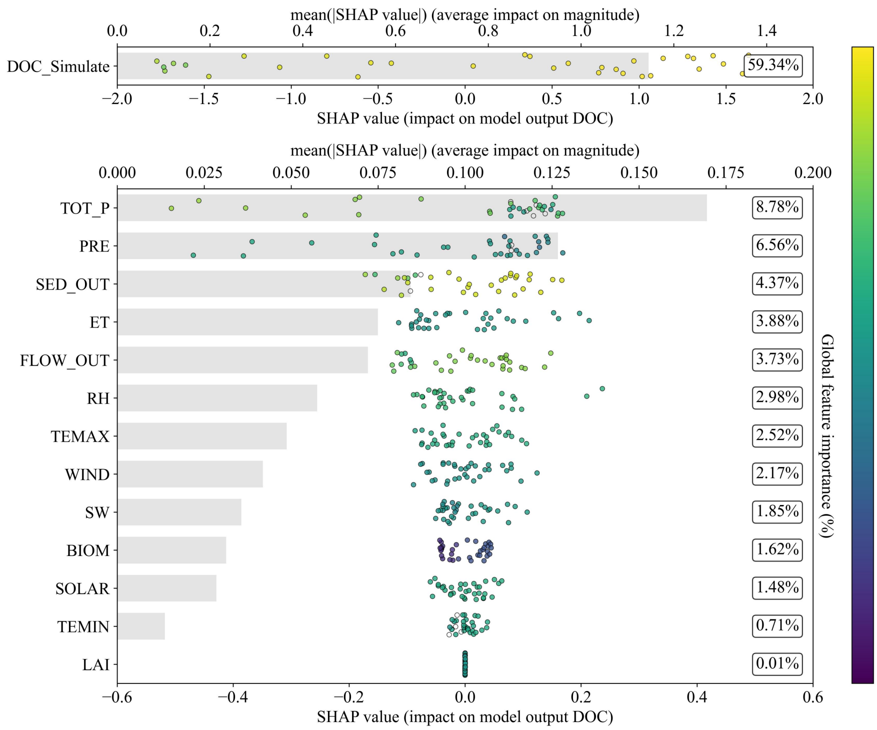

Figure 5. Global SHAP analysis of feature parameters. The upper plot shows the distribution of SHAP values for simulated DOC and its impact on the model's DOC predictions. The lower plot illustrates the distribution of SHAP values for input features along with their average impact on DOC prediction.

3.3.2. Residual Analysis of Input Features

Residual analysis was further conducted to quantify the contribution of input features to DOC prediction residuals within the coupled model. The five interaction terms with the highest contributions were identified as key factors. As illustrated in the residual interaction analysis (Fig. 6a), the dominant contributors were DOC_Simulate-TMIN, DOC_Simulate-SOLAR, LAI, DOC_Simulate-WIND, and ET-SOLAR, with contribution rates of 2.94%, 2.72%, 1.99%, 1.96%, and 1.95%, respectively. Notably, three of these interactions involved DOC_Simulate, implying that other environmental variables influence DOC output primarily through their interaction with the carbon cycle simulation module. Additionally, LAI emerged as a major individual contributor, highlighting the significant influence of the vegetation growth module on model residuals.

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

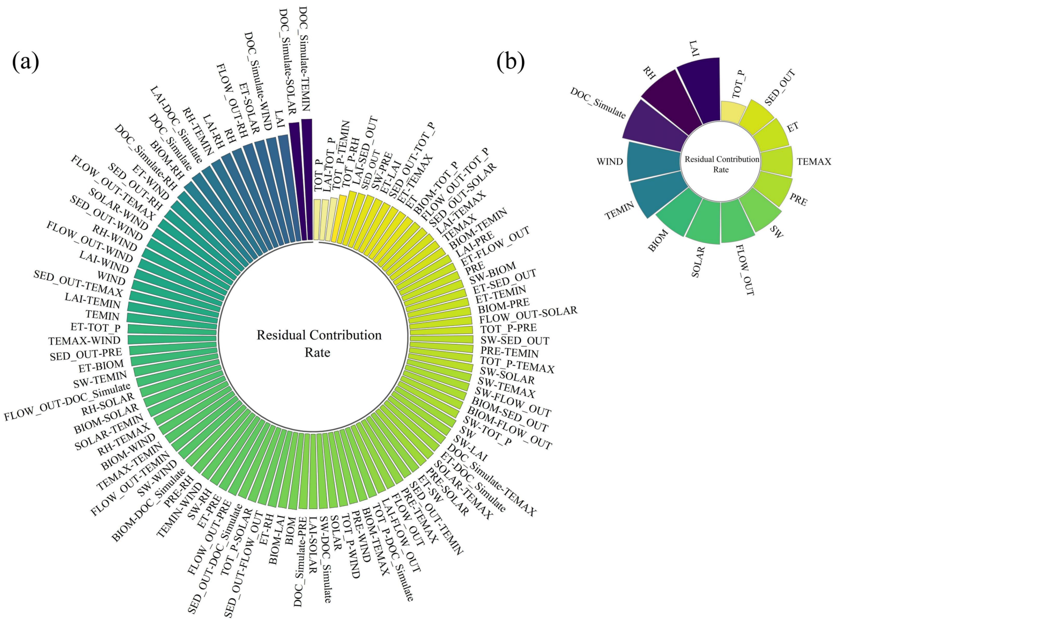

Figure 6. Residual analysis of characteristic parameters. (a) presents the detailed contribution rates for all parameter interactions, while (b) summarizes the primary individual contributions of key parameters. Circular bar plots illustrating the residual contribution rates of input parameters and their pairwise interactions influencing model DOC predictions.

For individual features, the top three contributors to residual variance were identified as key single-factor drivers. As shown in the single-feature residual analysis (Fig. 6b), LAI, RH, and DOC_Simulate contributed 14.65%, 14.01%, and 13.05%, respectively. LAI represents the vegetation growth module, which regulates DOC production via its control over plant growth, carbon fixation, and organic matter input. RH reflects relative humidity, which influences DOC levels indirectly by affecting vegetation activity, soil water content, and the decomposition of organic material.

3.3.3. Comprehensive Diagnosis of Module Performance

By integrating the results from global SHAP analysis and residual analysis, we identified the key factors influencing DOC simulation and conducted a comprehensive diagnosis of module performance.

The carbon cycle module, represented by DOC_Simulate, was identified as a dominant factor in both SHAP and residual analyses. This highlights its central role in DOC simulation and suggests that

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

errors within this module are a major source of DOC prediction uncertainty. Therefore, targeted improvement of the carbon module is essential to enhance the performance of DOC simulations in SWAT-C model. The pollutant transport module, reflected by TOT_P, exhibited high global importance in SHAP analysis but contributed little to residuals, indicating relatively stable performance in the DOC simulation process. Similarly, PRE was recognized as a key driver in SHAP analysis but had limited residual influence, suggesting that rainfall data were sufficiently accurate and contributed reliably to the DOC simulation.

The vegetation growth module, indicated by LAI, showed low importance in the global SHAP analysis yet emerged as a major factor in the residual analysis. To investigate this discrepancy, a single-feature SHAP analysis was conducted for LAI (Text S4 in the Supplementary Materials), revealing a negative effect on model output. Given LAI's relatively low temporal variability and stable contribution over time, its global SHAP importance remained limited. However, the high residuals associated with LAI suggest structural issues in the vegetation growth module, indicating a need for further refinement to improve the accuracy of DOC simulations.

Likewise, RH, while showing limited global SHAP importance, appeared as a key factor in the residual analysis. A single-feature SHAP analysis of RH (Text S4 in the Supplementary Materials) indicated a negative impact on DOC prediction. Although RH exhibited considerable temporal variability, its limited global contribution and substantial residual influence imply that RH likely affects DOC outcomes indirectly by influencing other modules. As such, improving the quality and integration of RH data should be prioritized to reduce prediction error in DOC simulations.

4. Discussion

4.1. Advantages of Coupling SWAT-C with Data-driven Learning

Machine learning (ML) models, as a core approach to data-driven learning, exhibit strong capabilities in capturing complex nonlinear patterns from large datasets and are particularly effective in modeling long-term temporal variability. However, they often lack interpretability and struggle to clearly explain how input variables interact to produce outputs (Azzam et al., 2022). In contrast, process-based models such as SWAT-C model offer explicit representations of physical mechanisms

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

and temporal dynamics but tend to underperform in data-scarce or environmentally complex basins due to limitations in structural completeness and parameter uncertainty (Zhao et al., 2024).

Integrating data-driven learning with the process-based SWAT-C model allows these two paradigms to complement each other. In the SWAT-MDF framework, SWAT-C model provides a rich set of physically meaningful feature parameters, which serve as external representations of internal watershed dynamics. These parameters offer valuable physical context for data-driven learning. The ML component, as the core of this data-driven module, learns the nonlinear relationships between these parameters and DOC outputs, enabling reliable and effective simulation of DOC under complex hydrological and environmental conditions. Essentially, this hybrid framework reflects a data-driven approach grounded in explicit physical processes, establishing a robust foundation for understanding DOC generation and transport mechanisms.

Comparative analysis of eight data-driven algorithms demonstrated that LSTM-based models, particularly Bi-LSTM, achieved the best performance. This can be attributed to Bi-LSTM ability to incorporate both past and future information in a time series, allowing it to better capture long-term dependencies and improve predictive accuracy. Additionally, Bi-LSTM demonstrated superior performance in modeling nonlinear and highly variable DOC time series, effectively extracting relevant patterns from complex feature inputs (Siami-Namini et al., 2019). In contrast, models such as XGBOOST, MLP, and SVR performed reasonably well on the training set but exhibited substantial performance degradation on the test set. This decline likely reflects the difficulty of traditional ML models in capturing long-term temporal dependencies and complex nonlinearities, resulting in reduced generalizability during testing (Zoremsanga and Hussain, 2024).

4.2. Advantages of Using SHAP and Residual Analysis for Module Diagnosis

To enhance the interpretability and diagnostic capacity of the coupled modeling framework, this study incorporates SHAP and residual analysis as two complementary approaches for module diagnosis. SHAP quantifies the contribution of each input variable to the predicted DOC outputs, thereby revealing the internal decision logic of the ML model (Mosca et al., 2022). Its ability to provide both global and local interpretability makes it particularly effective for identifying key drivers in

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

watershed-scale DOC simulations. In contrast, residual analysis focuses on the discrepancies between predicted and observed values, facilitating the detection of systematic biases in model outputs (Hantush and Kalin, 2008). More importantly, when residuals are decomposed by module or variable, they can reveal the specific contributions of individual sub-modules to the overall simulation error in the process-based model.

These two approaches offer complementary interpretive perspectives. SHAP highlights how input features influence model outputs, allowing the identification of dominant predictors and enhancing model transparency. Residual analysis, by quantifying and characterizing prediction errors, links these errors to specific model components and reveals structural deficiencies. The complementarity between these methods makes their integration a powerful strategy for achieving both interpretability and diagnostic insight. Together, they provide a comprehensive framework for module diagnosis that attributes output behavior to input drivers while diagnosing errors at the module level. In this study, the integrated framework successfully identified the dominant contribution and error source of the carbon module in DOC simulation, offering clear guidance and theoretical justification for optimizing the SWAT-C model. Furthermore, the proposed methodology exhibits strong generalizability and can be extended to module diagnosis and improvement in other coupled modeling systems.

4.3. Sources of DOC Prediction Biases

Previous studies have generally indicated that the NSE values for daily DOC simulation using SWAT-C typically remain around 0.5 (Du et al., 2023; Mukundan et al., 2023; Qi et al., 2020a), even after extensive parameter calibration. This indicates that the limited performance observed in this study is more likely due to structural deficiencies in the model, even after good calibration. Therefore, the modular diagnostic framework proposed here aims not merely to improve model outputs, but also to identify and quantify structural bottlenecks that hinder the DOC simulation accuracy and inform future model refinement.

A combined analysis of SHAP global importance and residual contributions shows that DOC simulation errors mainly originate from the carbon module and its nonlinear interactions with meteorological and hydrological components. The vegetation module, particularly LAI, also introduces

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

significant errors due to its influence on carbon input and surface processes (Li et al., 2025a).

Among interaction terms, DOC-TEMIN underscores the importance of minimum temperature in regulating carbon mineralization, while DOC-SOLAR and ET-SOLAR reflect the indirect effects of radiation via soil heat and evapotranspiration (Slavik et al., 2023). DOC-WIND indicates that WIND influences the transport of DOC mainly by inducing wind erosion, which alters surface conditions and facilitates DOC mobilization (Guo et al., 2023). Although LAI shows modest global SHAP importance, its high residual contribution reveals temporal or structural mismatches in the vegetation module. A negative correlation between LAI and DOC output suggests that the model fails to capture seasonal vegetation dynamics or carbon input variation (Gier et al., 2024; Li et al., 2025a), making LAI a latent but important error source that requires improved temporal representation and coupling with carbon processes. Moreover, climate warming has significantly altered vegetation phenology in recent decades, leading to earlier start-of-season (SOS) and delayed end-of-season (EOS) trends (Fu et al., 2019; Geng et al., 2020). These phenological shifts amplify the temporal complexity of LAI evolution, further challenging the model’s ability to simulate realistic canopy dynamics and accurately represent its regulatory role in DOC production and transport. Other modules, such as pollutant transport and precipitation input, demonstrated high global relevance but minimal residual impact, indicating reliable structure and data quality. In contrast, relative humidity had low SHAP importance but substantial residual contribution, likely due to indirect influences via evapotranspiration and soil moisture.

DOC simulation errors stem from complex, nonlinear interdependencies across modules. The carbon and vegetation modules-especially LAI-are the dominant error sources, while meteorological and hydrological variables contribute through both direct and indirect pathways. Refining the carbon module and improving inter-module coupling are key to enhancing DOC simulation performance.

4.4. Significance and Future Perspectives of the SWAT-MDF Framework

Accurately simulating and understanding the dynamics of DOC in data-scarce watersheds under continuous environmental change remains a critical yet challenging task (Li et al., 2025b). This study developed the SWAT-MDF diagnostic framework, which integrates the process-based SWAT-C model, a data-driven learning module implemented through ML algorithms, and SHAP- and residual-based interpretation to enhance the transparency and reliability of DOC simulations. By combining the physically based process mechanisms of SWAT-C model with the nonlinear learning capacity of ML, the framework effectively represents DOC variability under complex hydrological and environmental conditions. In addition, the integrated SHAP and residual analysis system provides interpretable insights into variable importance and sources of structural uncertainty.

More importantly, the SWAT-MDF diagnostic framework integrating SHAP and residual analysis offers a systematic and quantitative diagnostic tool for model error identification. This framework not only detects performance bottlenecks in SWAT-C model for DOC simulation but also exhibits strong generalizability, making it applicable to module diagnosis across other hydrological and biogeochemical processes. As such, it provides both a theoretical foundation and a practical pathway for structural improvement of process-based models that is applicable across regions and environmental conditions. In principle, the proposed diagnosis strategy is not restricted to SWAT-C and has the potential to be transferred to other models, such as SWAT+. Although SWAT+ represents the newest version of the SWAT family, carbon-related functionalities comparable to SWAT-C are still under development. In this context, the SWAT-MDF framework could serve as a transferable diagnostic reference to help identify model weaknesses and guide process improvement during the development and validation of future SWAT+ carbon modules. This also suggests that the framework can be extended to other models when needed. However, since the primary objective of this study is to diagnose and improve DOC simulations, we adopted the established SWAT-C model as the process-based backbone.

Despite these contributions, limitations remain. While SHAP and residual analysis are effective for identifying critical modules and influential variables, they serve primarily as diagnostic tools. They do not directly modify model structures or parameterization schemes. Substantive model improvement still requires the integration of experimental observations, long-term monitoring data, and theoretical insight to refine key processes and complete the cycle from error identification to resolution.

Future research should focus on: (1) incorporating anthropogenic activity datasets into the modeling framework to improve its responsiveness to real-world environmental changes; (2) using the current DOC module diagnosis system to guide field-based experiments and mechanistic investigations,

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

thereby supporting structural improvements of the SWAT-C model DOC simulation modules; (3) extending the proposed diagnosis strategy to other key hydrological and ecological variables, and developing standardized criteria for module diagnosis and cross-watershed model comparison.

5. Conclusion

In this study, we developed a SWAT-MDF coupled modeling framework that integrates the strengths of a process-based hydrological model and a data-driven deep learning model to simulate and diagnose daily-scale DOC dynamics in the Yalong River Basin. SHAP interpretation and residual analysis were incorporated to conduct a comprehensive diagnosis of model components. The aim of this study was to identify and quantify the structural sources of uncertainty in DOC simulation under complex and data-scarce conditions, thereby supporting targeted refinement of the SWAT-C model. Our main findings can be summarized as follows:

(1) Among the tested data-driven models, the Bi-LSTM-based SWAT-MDF configuration showed the most reliable performance (training *NSE* = 0.76, *KGE* = 0.80; testing *NSE* = 0.48, *KGE* = 0.67), confirming the advantage of combining process-based features with data-driven learning. However, the testing performance remained limited, indicating that structural deficiencies in SWAT-C continue to constrain DOC simulation and require further structural improvement.

(2) SHAP-based global interpretation identified DOC_Simulate, TOT_P, and PRE as dominant predictors of DOC, highlighting the strong interactions among carbon cycling, pollutant transport, and hydrological processes in the watershed.

(3) Residual analysis revealed that LAI, RH, and DOC_Simulate were the most significant contributors to prediction errors, corresponding to the vegetation growth module, relative humidity input, and carbon module, respectively. These results suggest that these components currently represent performance bottlenecks and should be prioritized in future model optimization.

(4) By integrating SHAP and residual analysis, this study established a systematic module diagnosis framework. The results indicate that rainfall inputs and the pollutant module perform relatively well in DOC simulation, whereas the carbon module, vegetation growth module, and relative humidity inputs exhibit larger errors. These components represent key targets for improving the

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

accuracy and stability of future DOC simulations.

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

Data availability. The DOC dataset used in this study was obtained from the publication by Xu et al. (2024), titled Escalating carbon export from high-elevation rivers in a warming climate, published in Environmental Science & Technology. The data is available at: [https://doi.org/10.1021/acs.est.3c06777]. All analysis was conducted in Python. The Python code used in this study is available on Zenodo at [https://doi.org/10.5281/zenodo.16599142].

Supplementary. Supplementary materials are provided in Appendix A.

Author contributions. Zehong Huang: Methodology, Software, Formal analysis, Visualization, Writing-original draft. Shouzhi Chen: Data curation, Validation, Investigation, Writing-review & editing. Yufeng Gong: Resources, Data curation, Investigation. Zheng Wang: Validation, Writing-review & editing. Zheng Duan: Methodology, Validation, Writing-review & editing. Yongshuo Fu*: Conceptualization, Supervision, Project administration, Funding acquisition, Writing-review & editing.

Competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Disclaimer. Publisher's note: 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.

Acknowledgments. The authors are grateful to the editor and reviewers for their valuable comments and suggestions, which have greatly helped us improve this paper.

Financial support. This study was supported by the Key Program of the National Natural Science Foundation of China (Grant 42430504), and the National Funds for Distinguished Young Youths (Grant 42025101), the Key Program of the National Natural Science Foundation of China (Grant 42330515), the Fundamental Research Funds for the Central Universities (2243300004) and the 111 Project (Grant B18006). We thank the reviewers and

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

associate editor of this manuscript for their valuable comments and suggestions that helped us to substantially improve the paper.

References

Abbaspour, K., Vaghefi, S., and Srinivasan, R.: A guideline for successful calibration and uncertainty analysis for soil and water assessment: a review of papers from the 2016 international SWAT conference, Water, 10, 6, https://doi.org/10.3390/w10010006, 2017.

Azzam, A., Zhang, W., Akhtar, F., Shaheen, Z., and Elbeltagi, A.: Estimation of green and blue water evapotranspiration using machine learning algorithms with limited meteorological data: A case study in Amu Darya River Basin, Central Asia, Comput. Electron. Agric., 202, 107403, https://doi.org/10.1016/j.compag.2022.107403, 2022.

Carvalho, D. V., Pereira, E. M., and Cardoso, J. S.: Machine learning interpretability: a survey on methods and metrics, Electronics, 8, 832, https://doi.org/10.3390/electronics8080832, 2019.

Cole, J. J., Prairie, Y. T., Caraco, N. F., McDowell, W. H., Tranvik, L. J., Striegl, R. G., Duarte, C. M., Kortelainen, P., Downing, J. A., Middelburg, J. J., and Melack, J.: Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget, Ecosystems, 10, 172–185, https://doi.org/10.1007/s10021-006-9013-8, 2007.

Ding, B., Yu, X., and Jia, G.: Exploring the controlling factors of watershed streamflow variability using hydrological and machine learning models, Water Resour. Res., 61, https://doi.org/10.1029/2024wr039734, 2025.

Du, X., Zhang, X., Mukundan, R., Hoang, L., and Owens, E. M.: Integrating terrestrial and aquatic processes toward watershed scale modeling of dissolved organic carbon fluxes, Environ. Pollut., 249, 125–135, https://doi.org/10.1016/j.envpol.2019.03.014, 2019.

Du, X., Faramarzi, M., Qi, J., Lei, Q., and Liu, H.: Investigating hydrological transport pathways of dissolved organic carbon in cold region watershed based on a watershed biogeochemical model, Environ. Pollut., 324, 121390, https://doi.org/10.1016/j.envpol.2023.121390, 2023.

Evans, C. D., Monteith, D. T., and Cooper, D. M.: Long-term increases in surface water dissolved organic carbon: observations, possible causes and environmental impacts, Environ. Pollut., 137, 55–71, https://doi.org/10.1016/j.envpol.2004.12.031, 2005.

Fan, H., Jiang, M., Xu, L., Zhu, H., Cheng, J., and Jiang, J.: Comparison of Long Short Term Memory Networks and the Hydrological Model in Runoff Simulation, Water, 12, 175, https://doi.org/10.3390/w12010175, 2020.

Fu, Y. H., Piao, S., Delpierre, N., Hao, F., Hänninen, H., Geng, X., Peñuelas, J., Zhang, X., Janssens, I. A., and Campioli, M.: Nutrient availability alters the correlation between spring leaf-out and autumn leaf senescence dates, Tree Physiol., 39, 1277–1284, https://doi.org/10.1093/treephys/tpz041, 2019.

Geng, X., Fu, Y. H., Hao, F., Zhou, X., Zhang, X., Yin, G., Vitasse, Y., Piao, S., Niu, K., De Boeck, H. J., Menzel, A., and Peñuelas, J.: Climate warming increases spring phenological differences among temperate trees, Global Change Biol., 26, 5979–5987, https://doi.org/10.1111/gcb.15301, 2020.

Gier, B. K., Schlund, M., Friedlingstein, P., Jones, C. D., Jones, C., Zaehle, S., and Eyring, V.: Representation of the terrestrial carbon cycle in CMIP6, Biogeosciences, 21, 5321–5360, https://doi.org/10.5194/bg-21-5321-2024, 2024.

Grusson, Y., Sun, X., Gascoin, S., Sauvage, S., Raghavan, S., Anctil, F., and Sáchez-Pérez, J.-M.: Assessing the capability of the SWAT model to simulate snow, snow melt and streamflow dynamics over an alpine watershed, J. Hydrol., 531, 574–588, https://doi.org/10.1016/j.jhydrol.2015.10.070, 2015.

Guo, Y., Ni, Z., Dong, Y., Wang, S., Wu, Y., Liu, S., and Huang, Y.: Overlooked great role of wind erosion in terrestrial dissolved organic matter input to lake ecosystem in cold and arid regions, Sci. Total Environ., 890, 164272, https://doi.org/10.1016/j.scitotenv.2023.164272, 2023.

Guo, Z., Feng, C., Yang, L., and Liu, Q.: Bridging the gap: An interpretable coupled model (SWAT-ELM-SHAP) for blue-green water simulation in data-scarce basins, Agric. Water Manag., 306, 109157, https://doi.org/10.1016/j.agwat.2024.109157, 2024.

Hantush, M. M. and Kalin, L.: Stochastic residual-error analysis for estimating hydrologic model predictive uncertainty, J. Hydrol. Eng., 13, 585–596, https://doi.org/10.1061/(ASCE)1084-0699(2008)13:7(585), 2008.

Harkort, L. and Duan, Z.: Estimation of dissolved organic carbon from inland waters at a large

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

scale using satellite data and machine learning methods, Water Res., 229, 119478, https://doi.org/10.1016/j.watres.2022.119478, 2023.

Hongyi Li, G. S. and Maddala: Bootstrapping time series models, Econom. Rev., 15, 115–158, https://doi.org/10.1080/07474939608800344, 1996.

Lawlor, A. J. and Tipping, E.: Metals in bulk deposition and surface waters at two upland locations in northern England, Environ. Pollut., 121, 153–167, https://doi.org/10.1016/S0269-7491(02)00228-2, 2003.

Lee, D., Karki, R., Kalin, L., Isik, S., Srivastava, P., and Zhang, X.: Management Strategies for Dissolved Organic Carbon Reduction from Forested Watersheds using the SWAT-C model, Environ. Manag., 75, 1181–1200, https://doi.org/10.1007/s00267-025-02128-y, 2025.

Lee, J., Abbas, A., McCarty, G. W., Zhang, X., Lee, S., and Hwa Cho, K.: Estimation of base and surface flow using deep neural networks and a hydrologic model in two watersheds of the Chesapeake Bay, Journal of Hydrology, 617, 128916, https://doi.org/10.1016/j.jhydrol.2022.128916, 2023a.

Lee, J., Abbas, A., McCarty, G. W., Zhang, X., Lee, S., and Hwa Cho, K.: Estimation of base and surface flow using deep neural networks and a hydrologic model in two watersheds of the Chesapeake Bay, J. Hydrol., 617, 128916, https://doi.org/10.1016/j.jhydrol.2022.128916, 2023b.

Li, M., Chen, S., Hao, F., Wang, N., Wu, Z., Xu, Y., Zhang, J., Zhang, Y., and Fu, Y. H.: Integration of the vegetation phenology module improves ecohydrological simulation by the SWAT-carbon model, Hydrol. Earth Syst. Sci., 29, 2081–2095, https://doi.org/10.5194/hess-29-2081-2025, 2025a.

Li, S.-L., Chetelat, B., Yue, F., Zhao, Z., and Liu, C.-Q.: Chemical weathering processes in the Yalong River draining the eastern Tibetan Plateau, China, J. Asian Earth Sci., 88, 74–84, https://doi.org/10.1016/j.jseaes.2014.03.011, 2014.

Li, Y., Wang, G., Wang, W., Sun, X., Li, Y., Xiao, J., Xie, W., Ding, J., and Song, C.: Ultraviolet radiation stimulates the degradability of groundwater-fed DOC during the baseflow period of streams on the qinghai-tibet plateau permafrost region, J. Hydrol., 653, 132751, https://doi.org/10.1016/j.jhydrol.2025.132751, 2025b.

Liu, J., Liu, J., Du, X., Guo, R., Duan, Z., Yuan, B., and Liu, Y.: The importance of impoundment

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

interception in simulating riverine dissolved organic carbon, Water Resour. Res., 60, https://doi.org/10.1029/2024wr038133, 2024.

Liu, X., Yang, M., Meng, X., Wen, F., and Sun, G.: Assessing the impact of reservoir parameters on runoff in the yalong river basin using the SWAT model, Water, 11, 643, https://doi.org/10.3390/w11040643, 2019.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell., 2, 56–67, https://doi.org/10.1038/s42256-019-0138-9, 2020.

Mosca, E., Szigeti, F., Tragianni, S., Gallagher, D., and Groh, G., 2022. SHAP-based explanation methods: a review for NLP interpretability. Proceedings of the 29th International Conference on Computational Linguistics, pp. 4593–4603.

Mukundan, R., Gelda, R. K., Moknatian, M., Zhang, X., and Steenhuis, T. S.: Watershed scale modeling of dissolved organic carbon export from variable source areas, J. Hydrol., 625, 130052, https://doi.org/10.1016/j.jhydrol.2023.130052, 2023.

Pelletier, G. J., Chapra, S. C., and Tao, H.: QUAL2Kw – A framework for modeling water quality in streams and rivers using a genetic algorithm for calibration, Environ. Model. Softw., 21, 419–425, https://doi.org/10.1016/j.envsoft.2005.07.002, 2006.

Qi, J., Du, X., Zhang, X., Lee, S., Wu, Y., Deng, J., Moglen, G. E., Sadeghi, A. M., and McCarty, G. W.: Modeling riverine dissolved and particulate organic carbon fluxes from two small watersheds in the northeastern United States, Environ. Model. Softw., 124, 104601, https://doi.org/10.1016/j.envsoft.2019.104601, 2020a.

Qi, J., Zhang, X., Yang, Q., Srinivasan, R., Arnold, J. G., Li, J., Waldholf, S. T., and Cole, J.: SWAT ungauged: Water quality modeling in the Upper Mississippi River Basin, J. Hydrol., 584, 124601, https://doi.org/10.1016/j.jhydrol.2020.124601, 2020b.

Rathjens, H., Oppelt, N., Bosch, D. D., Arnold, J. G., and Volk, M.: Development of a grid-based version of the SWAT landscape model, Hydrol. Process., 29, 900–914, https://doi.org/10.1002/hyp.10197, 2015.

Razavi, T. and Coulibaly, P.: Streamflow Prediction in Ungauged Basins: Review of

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

Regionalization Methods, J. Hydrol. Eng., 18, 958–975, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000690, 2013.

Santos Nobre, J. and Da Motta Singer, J.: Residual Analysis for Linear Mixed Models, Biom. J., 49, 863–875, https://doi.org/10.1002/bimj.200610341, 2007.

Siami-Namini, S., Tavakoli, N., and Namin, A. S.: The performance of LSTM and BiLSTM in forecasting time series, in: 2019 IEEE International Conference on Big Data (Big Data), 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 3285–3292, https://doi.org/10.1109/BigData47090.2019.9005997, 2019.

Slavik, I., Kostrowski, D., and Uhl, W.: Effect of solar radiation on natural organic matter composition in surface waters and resulting impacts on drinking water treatment, Environ. Technol., 2023.

Srinivasan, R., Ramanarayanan, T. S., Arnold, J. G., and Bednarz, S. T.: LARGE AREA HYDROLOGIC MODELING AND ASSESSMENT PART II: MODEL APPLICATION [1], JAWRA J. Am. Water Resour. Assoc., 34, 91–101, https://doi.org/10.1111/j.1752-1688.1998.tb05962.x, 1998.

Tan, M. L., Gassman, P. W., Srinivasan, R., Arnold, J. G., and Yang, X.: A Review of SWAT Studies in Southeast Asia: Applications, Challenges and Future Directions, Water, 11, 914, https://doi.org/10.3390/w11050914, 2019.

Todini, E.: Hydrological catchment modelling: past, present and future, Hydrology and Earth System Sciences, 11, 468–482, https://doi.org/10.5194/hess-11-468-2007, 2007.

Tyagi, K., Rane, C., Harshvardhan, and Manry, M.: Chapter 4 - regression analysis, in: Artificial Intelligence and Machine Learning for EDGE Computing, edited by: Pandey, R., Khatri, S. K., Singh, N. kumar, and Verma, P., Academic Press, 53–63, https://doi.org/10.1016/B978-0-12-824054-0.00007-1, 2022.

Wang, K., Mukundan, R., Gelda, R. K., and Frei, A.: Modeling dissolved organic carbon export from water supply catchments in the northeastern United States, Sci. Total Environ., 963, 178532, https://doi.org/10.1016/j.scitotenv.2025.178532, 2025.

Xu, S., Li, S.-L., Bufe, A., Klaus, M., Zhong, J., Wen, H., Chen, S., and Li, L.: Escalating carbon export from high-elevation rivers in a warming climate, Environ. Sci. Technol., 58, 7032–7044,

https://doi.org/10.5194/egusphere-2025-5503
Preprint. Discussion started: 27 December 2025

https://doi.org/10.1021/acs.est.3c06777, 2024.

Yang, Q. and Zhang, X.: Improving SWAT for simulating water and carbon fluxes of forest ecosystems, Sci. Total Environ., 569–570, 1478–1488, https://doi.org/10.1016/j.scitotenv.2016.06.238, 2016.

Yao, Z., Wang, Z., Wang, D., Wu, J., and Chen, L.: An ensemble CNN-LSTM and GRU adaptive weighting model based improved sparrow search algorithm for predicting runoff using historical meteorological and runoff data as input, J. Hydrol., 625, 129977, https://doi.org/10.1016/j.jhydrol.2023.129977, 2023.

Zhang, X., Izaurralde, R. C., Arnold, J. G., Williams, J. R., and Srinivasan, R.: Modifying the Soil and Water Assessment Tool to simulate cropland carbon flux: Model development and initial evaluation, Sci. Total Environ., 463–464, 810–822, https://doi.org/10.1016/j.scitotenv.2013.06.056, 2013.

Zhao, J., Zhang, N., Liu, Z., Zhang, Q., and Shang, C.: SWAT model applications: from hydrological processes to ecosystem services, Sci. Total Environ., 931, 172605, https://doi.org/10.1016/j.scitotenv.2024.172605, 2024.

Zhao, Y., Dong, N., Li, Z., Zhang, W., Yang, M., and Wang, H.: Future precipitation, hydrology and hydropower generation in the Yalong River Basin: Projections and analysis, J. Hydrol., 602, 126738, https://doi.org/10.1016/j.jhydrol.2021.126738, 2021.

Zoremsanga, C. and Hussain, J.: Particle Swarm Optimized Deep Learning Models for Rainfall Prediction: A Case Study in Aizawl, Mizoram, IEEE Access, 12, 57172–57184, https://doi.org/10.1109/ACCESS.2024.3390781, 2024.