A spatially hybrid hydrological modeling approach using subbasin-specific model structures

Yu-Jing Wang[a,b,c,d], Liang-Jun Zhu[a,b,f,*], Cheng-Zhi Qin[a,b,e], A-Xing Zhu[f]

[a] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing, China

[b] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China

[c] Research Center on Flood and Drought Disaster Reduction of the Ministry of Water Resources, Beijing, China

[d] China Institute of Water Resources and Hydropower Research, Beijing, China

[e] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, China

[f] Department of Geography, University of Wisconsin–Madison, Madison, WI, USA

Corresponding author: Liang-Jun Zhu

E-mail addresses: zlj@lreis.ac.cn

Highlights

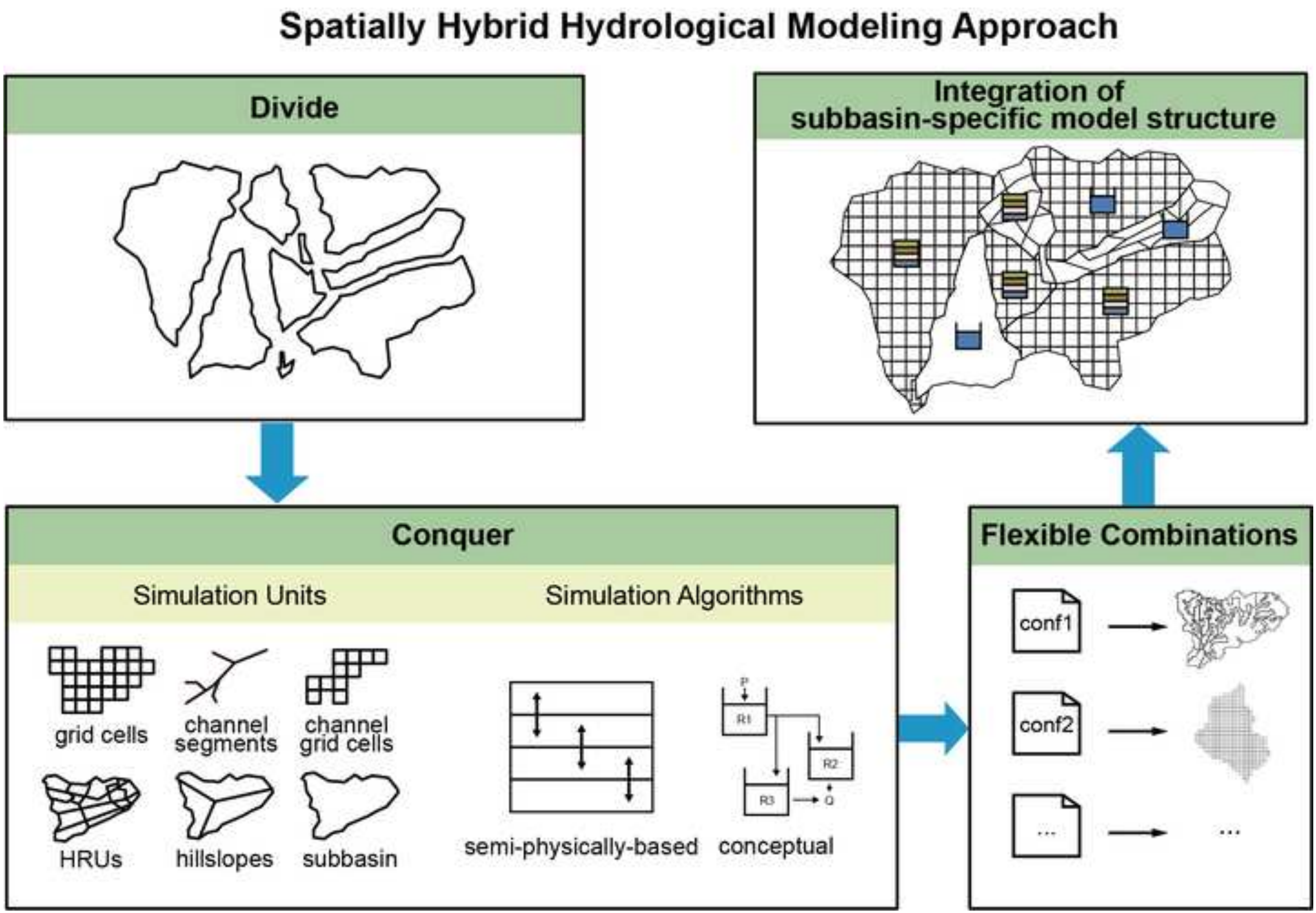
- A new approach ~~to~~ supporting spatially hybrid hydrological model structures is proposed.
- Diverse subbasin-specific model structures are integrated for watershed-scale simulation.
- The approach~~Model structure~~ flexibly combines compatible spatial units and simulation algorithms.
- Spatial units ~~of~~ (subbasin, hillslope, HRU, and cell)~~s~~ support conceptual or semi-physical algorithms.
- The s~~S~~patially hybrid model structure ~~from the proposed approach~~ achieves a balanced ~~improves general~~ performance improvement.

Highlights

- A new approach supporting spatially hybrid hydrological model structures is proposed.
- Diverse subbasin-specific model structures are integrated for watershed-scale simulation.
- The approach flexibly combines compatible spatial units and simulation algorithms.
- Spatial units (subbasin, hillslope, HRU, and cell) support conceptual or semi-physical algorithms.
- The spatially hybrid model structure achieves a balanced performance improvement.

Graphical Abstract

Click here to access/download;Graphical Abstract;abstract-v1.png



Abstract:

Hydrological models ~~that typically~~ adopting spatially consistent model structures are often ill-suited ~~unfit~~ for complex application contexts with significant spatial heterogeneity. Although existing hydrological modeling frameworks support spatially varying lumped or semi-distributed conceptual ~~model~~ structures, they face ~~a significant~~ challenges in integrating distributed physically-based ~~model~~ structures. This paper proposes a new spatially hybrid hydrological modeling approach ~~that, which flexibly~~ combining~~es~~ compatible spatial units and simulation algorithms to construct distinct~~different~~ model structures for individual subbasins within ~~interest~~ a watershed, ~~and then coherently~~ integrated~~s them~~ through channel routing f~~t~~or ~~perform~~ watershed-scale simulation. Implemented within the open-source Spatially Explicit Integrated Modeling System (SEIMS), this approach was evaluated through a proof-of-concept case study ~~in a medium-sized watershed~~ using ~~with~~ two distinct model structures: a lumped conceptual model ~~structure~~ for gently sloping subbasins and a fully-distributed semi-physically-based model ~~structure~~ for mountainous subbasins. Comparative experiments demonstrate~~d~~ that the hybrid model inherits ~~combines~~ both the ~~the~~ strengths and limitations of its constituent structures~~of both~~, achieving a balanced performance improvement~~ing general performance~~. This~~e~~ ~~proposed~~ approach enhances modeling flexibility, advancing ~~towards~~ knowledge-driven intelligent modeling in spatially heterogeneous environments.

Keywords: Watershed simulation; hydrological modeling framework; spatially hybrid modeling; model structure; spatial heterogeneity; SEIMS

Software availability statement

Software: SEIMS (Spatially Explicit Integrated Modeling System)

Developers: Yu-Jing Wang, Liang-Jun Zhu

Contact: zlj@lreis.ac.cn

Operating systems supported: Windows, Linux, and macOS

Language: C++ and Python

Source code: The archived version supporting this paper can be accessed ~~vis~~ via the OSF site at https://doi.org/10.17605/OSF.IO/D5WUR. For the continuous development version, please visit the GitHub repository at https://github.com/lreis2415/SEIMS.

Documentation: Documentation for running the proposed approach is available at https://osf.io/vwxkr, with the continuous user manual for the SEIMS framework accessible at https://lreis2415.github.io/SEIMS/.

1 Introduction

Hydrological models serve as effective tools for simulating and enhancing our understanding of complex hydrological processes in real-world watersheds. A variety of hydrological models have been developed, each suited to different application contexts depending on ~~multiple~~ factors such as watershed characteristics and data availability~~e~~ ~~data~~ (Beven, 2000; Gharari et al., 2021; Wagener et al., 2001). Building an application-specific model that is both hydrologically reasonable and practically executable ~~stands~~ is a~~s the~~ primary and critical step for hydrological modeling and applications such as scenario analysis of watershed management (Butts et al., 2004; David et al., 2022; Pilz et al., 2020; Qin et al., 2025; van Esse et al., 2013).

The model structure of a hydrological model can be defined by its spatial discretization scheme and the representation of hydrological processes (Chow, 1988; Milad et al., 2012; Hrachowitz and Clark, 2017). The spatial discretization scheme refers to the arrangement of one or more types of spatial units (or ~~also known as~~ simulation units) within the watershed ~~for hydrological simulation~~ (Dehotin and Braud, 2008). Lumped spatial units typically ignore or ~~broadly~~ generalize the spatial heterogeneity of watershed characteristics, including either coarse-scale units (such as a subbasin or ~~even~~ an entire watershed), or units defined by dominant combinations of selected geographic attributes (Arnold et al., 2010). In contrast, semi- or fully-distributed spatial units can fully cover the entire ~~geographic~~ spatial domain of the watershed at varying levels of detail~~s~~, such as spatially explicit HRUs (hydrological response units; Teshager et al., 2016) and grid cells (Wigmosta et al., 1994). The representation of hydrological processes ~~denotes~~ refers to the simulation algorithms used ~~for~~ to describe hydrological phenomena (e.g., infiltration and surface flow routing). Conceptual representations simplify hydrological processes with ~~a few~~ lumped components, typically using conceptual reservoirs (or buckets) (Fenicia et al., 2011; Knoben et al., 2019). In contrast, fully-physically-based representations apply fundamental physical laws (e.g., conservation of mass, momentum, and energy ~~conservation~~) by solving coupled partial differential equations (Abbott et al., 1986; Shu et al., 2020). To balance physical realism with computational feasibility, semi-physically-based representations employ simplified physical equations for individual hydrological processes that are generally simulated independently (Arnold et al., 1998; Tague and Band, 2004; Tang et al., 2019).

Existing methods ~~o~~for determining model structures can be categorized into~~as~~ spatially consistent methods and spatially varying methods. The spatially consistent model structure means the types of simulation units and simulation algorithms for ~~interested~~ the hydrological processes of interest are uniform across the entire~~the same for the whole~~ watershed. Most research involving the~~on~~ direct selection of~~ly choosing~~ a single existing hydrological model with a fixed model structure falls into this category. A more flexible ~~choice~~ alternative is utilizing modular hydrological modeling frameworks, which emphasize various aspects of flexibility in customizing application-specific model structures, such as extensible types of simulation units and associated simulation algorithms (Kneis, 2015), and alternative simulation algorithms for individual hydrological processes (Clark et al., 2015; Craig, 2020; Knoben et al., 2019; Zhu et al., 2019). However, their default design or most common application of these frameworks remain focused~~s~~ on customizing and applying a spatially consistent model structure. This method inherently assumes that the spatial variability of watershed characteristics can be adequately captured by the required input data and parameters alone. However, using a fixed model structure may fail to effectively accommodate diverse and complex application contexts that exhibit~~ing~~ significant spatial heterogeneity (Ley et al., 2016; Savenije, 2009; Sui and Turner, 2021).

Recognizing the limitations of spatially consistent model structures, researchers have explored two principal approaches to constructing spatially varying model structures within watersheds~~ in two principal approaches~~. The first involves externally integrating multiple individual models, either manually (Liu et al., 2020; Li et al., 2021; Wang et al., 2021) or through unified model interoperability~~e~~ interfaces such as ESMF (Earth System Modeling Framework; Hill et al., 2004) and OpenMI (Open Modeling Interface; Harpham et al., 2019). However, the flexibility of such model-level integration may be limited by the inherently fixed model structures of each integrated model. Furthermore, ~~different~~ interoperable interfaces often require~~need to be~~ separate~~ly~~ implementation~~ed~~ for each model, posing considerable challenges for hydrologists lacking programming expertise.

To further enhance ~~the~~ flexibility~~ of spatially varying model structures~~, a second approach~~another idea~~ involves internally constructing and integrating multiple model structures within a single modular hydrological modeling framework such as FLEX-Topo (Topography-driven Flux Exchange hydrological model; Gao et al., 2014; Savenije, 2010), SUPERFLEX (SUPER Flux Exchange hydrological model; Fenicia et al., 2011, 2016), and airGR (Suite of GR Hydrological Models; Coron et al., 2017; Thébault et al., 2024). These frameworks allow ~~the assignment of~~ different model structures to be assigned to different modeling areas within a ~~the~~ watershed with ~~only~~ minimal changes to configuration or source code (Thébault et al., 2024). Nevertheless, the~~yse frameworks~~ are fundamentally limited to conceptual representations of hydrological processes, with simulation units typically restricted to~~being~~ subbasins and HRUs. Therefore, these state-of-the-art flexible frameworks primarily support lumped or semi-distributed conceptual model structures, ~~and thus~~ facing~~e~~ ~~a~~ significant challenges in integrating physically-based simulation algorithms and spatially explicit distributed simulation units.

To address this limitation, we propose a new ~~and practically significant~~ spatially hybrid modeling approach. This approach enables the flexible combination of compatible spatial units and simulation algorithms to construct distinct model structures for individual subbasins within the watershed, such as lumped conceptual and fully-distributed semi-physically-based model structures. Subbasin-level model structures are then coherently integrated through channel routing to perform watershed-scale simulation. ~~The remainder of this paper is organized as follows:~~

Section 2 outlines the design of the proposed approach and its implementation within the SEIMS (Spatially Explicit Integrated Modeling System; Zhu et al., 2019) framework.~~,~~ Section 3 presents an experimental case study, followed by the discussion in Section 4, and the conclusion in Section 5.

2 Method design and implementation

2.1 Basic idea

The basic idea of the proposed spatially hybrid hydrological modeling approach is to employ~~ing~~ a divide-and-conquer strategy to construct and execute distinct~~different~~ model structures for relatively independent areas within a~~the~~ watershed. A watershed can be naturally divided into distinct ~~closed~~ catchment areas connected by a ~~the~~ channel network, i.e., subbasins. Thus, subbasins are naturally suited to serve as independent~~the divide-and-conquer~~ modeling units, each configurable with a specific~~n individual~~ model structure. Therefore, the core design of this approach involves two key aspects: 1) flexibility in constructing model structures, and 2) the capability to integrate these subbasin-specific model structures for~~to perform~~ watershed-scale simulation.

The idea in designing a flexible model structure is to ensure compatibility between the spatial discretization scheme (i.e., simulation units) and hydrological process representations (i.e., simulation algorithms). Both components can be classified ~~according~~ based onto~~to~~ their dependencies, which determine their compatibility (Figure 1a). For simulation units, spatial dependency (or connectivity) refers to whether spatially explicit relationships exist between units. For simulation algorithms, computational dependency refers to whether a simulation on one unit relies on outputs (e.g., water flows or state variables) from adjacent or upstream units (Figure 1b). Therefore, a valid combination requires that the computational dependencies of selected simulation algorithms be ~~are~~ satisfied by the spatial connectivity of the simulation units (Figure 1b). For instance, a conceptual reservoir-based surface runoff generation algorithm combined with a physically-based kinematic wave~~-based~~ surface routing algorithm can be applied to grid cells, where the computational dependency is satisfied by the flow directions between ~~grid~~ cells. Conversely, ~~Another example is the~~ combining~~ation of~~ a physically-based Green-Ampt surface runoff generation algorithm with a conceptual unit hydrograph-based surface routing algorithm for application on HRUs (Craig, 2020)~~, which~~ requires no computational dependency. It is crucial to note that while computational dependency acts a~~i~~s a fundamental constraint, the ultimate scientific validity~~rationality~~ of a customized model structure relies on the modeler's hydrological knowledge and application-specific requirements.

Each subbasin model simulates hillslope processes and ~~the~~ channel routing~~ process~~. In this study, it is assumed that mass and energy ~~are~~ exchanges~~d~~ between subbasins ~~only~~ occur exclusively through the channel network, aligning with standard practices in most semi-physically-based hydrological models, such as SWAT. Thus, ~~the~~ channel routing components within each subbasin model structure (e.g., the

Muskingum method) serve ~~are used~~ to couple with ~~their~~ upstream and downstream subbasins, integrating all subbasin model structures into a comprehensive watershed model (Figure 1c). The channel routing of a downstream subbasin relies on outputs from its upstream channels. This dependency implies that the hillslope processes of each subbasin~~s~~ can be parallelized to improve computational~~ng~~ efficiency, while channel routing processes ~~of each subbasins~~ must be executed sequentially following~~according to~~ the upstream–downstream network structure (Liu et al., 2016).

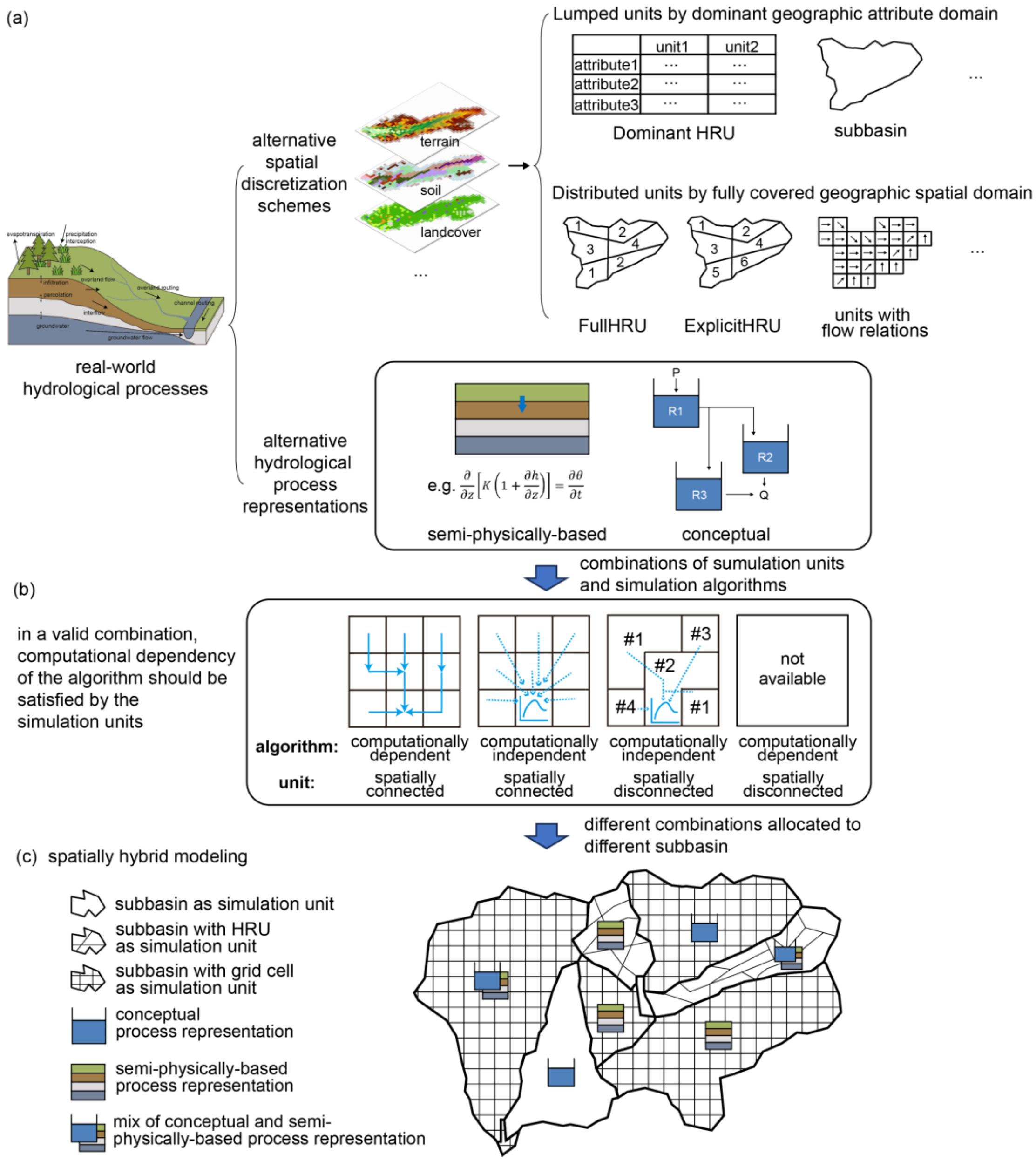


Figure 1. Design of the proposed spatially hybrid hydrological modeling approach:~~.~~ a) definition of~~g~~ model structure via~~by~~ spatial discretization schemes and hydrological process representations; b) support for~~ing~~ compatible combinations of simulation units and simulation algorithms; and c) ~~enabling the~~ spatial varying allocation and integration of different model structures at the subbasin level.

2.2 Design of the spatially hybrid hydrological modeling approach using the SEIMS framework

Following the above basic idea, the proposed spatially hybrid hydrological modeling approach was designed using the SEIMS framework, leveraging ~~for~~ its flexible modular design and two-level parallelization strategy (Liu et al., 2016; Zhu et al., 2019). To support the flexible model construction~~ of model structures~~, the framework underwent three enhancements: 1) the spatial discretization schemes of SEIMS were expanded (Figure 2a; Section 2.2.2); 2)~~,~~ the simulation modules ~~o~~for hillslope processes were adapted~~upgraded~~ to accommodate compatible simulation units (Figure 2b; Section 2.2.2);~~,~~ and 3) the model execution mechanism ~~of SEIMS-based model~~ was ~~upgraded~~ modified to allocate different model structures to subbasins and execute them~~d~~ in an integrated ~~way~~ manner (Figure 2c; Section 2.2.3).

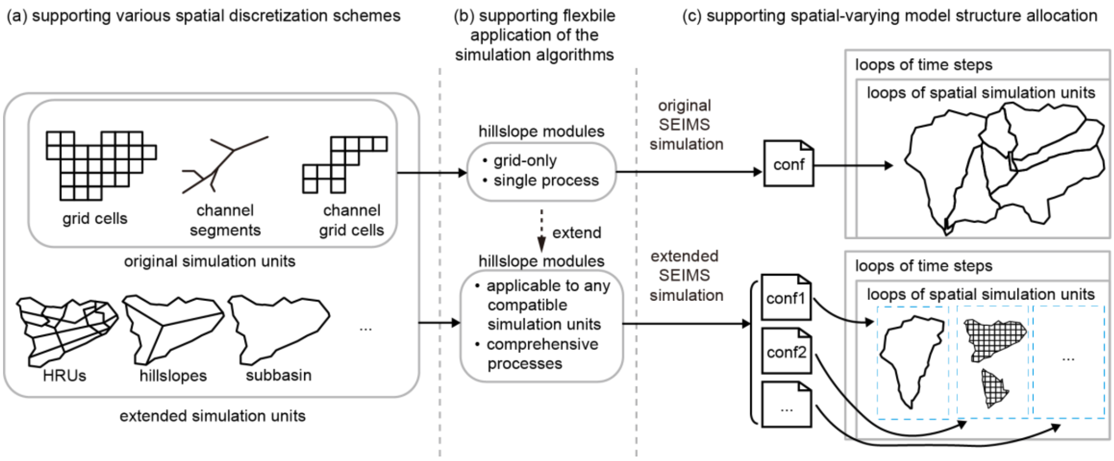


Figure 2. Overall design ~~o~~for extending ~~the~~ SEIMS (Spatially Explicit Integrated Modeling System) to implement the proposed spatially hybrid hydrological modeling approach: a) expansion of~~supporting various~~ spatial discretization schemes; b) adaptation of ~~supporting applying simulation modules of~~ hillslope process ~~e~~simulation modules for ~~to~~ compatible simulation units; and c) ~~supporting~~ allocation and integration of spatially varying model structures at the ~~allocation on~~ subbasin levels~~s~~ to construct a watershed model.

2.2.1 Brief introduction to SEIMS

As a hydrological modeling framework, SEIMS provides users with a ~~the~~ comprehensive~~lete~~ toolchain for~~to~~ data preprocessing~~ modeling data~~, model construction and execution~~e the SEIMS-based model~~, ~~analyze the~~ parameter sensitivity analysis, and more~~so on~~ (Liu et al., 2021; Zhu et al., 2019). More importantly, SEIMS supports developers in adding simulation modules ~~o~~for hydrologic processes, adher~~following~~ing to three core architectural~~developing~~ principles: the spatial discretization scheme, the modular modeling design, and the parallelization strategy.

(1) Spatial discretization scheme

SEIMS primarily adopts a two-level spatial discretization scheme, i.e., the "subbasin–basic simulation unit." During ~~In~~ the ~~stage of~~ data preprocessing stage~~–~~ ~~modeling data~~, SEIMS delineates the watershed into subbasins, each containing~~with one~~ a single channel. Within each subbasin, ~~the~~ basic simulation units (i.e., grid cells in the original version) are delineated and organized as a one-dimensional array. Most ~~simulation modules of~~ hillslope process simulation modules use grid cells as simulation units, while a few modules ~~use~~ operate at the subbasin ~~units~~ level such as the linear reservoir algorithm for ~~the~~ groundwater~~ process~~. C~~Modules of c~~hannel routing modules~~processes~~ use channel lines or grid cells as simulation units. All parameters required by the modules on these simulation units are prepared using the data preprocessing tools. G~~The g~~rid cells are organized into topological layers~~ed~~ based on ~~the~~ flow direction ~~of each cell~~ (i.e., the single flow direction algorithm in the original version; Liu et al., 2014). These layers determine the simulation sequences of grid cells ~~are used~~ in ~~modules of~~ surface or subsurface flow routing modules~~processes to determine the simulation sequences of grid cells~~. ~~The same idea is used for~~ A similar layering strategy is applied to subbasins and channels.

(2) Modular modeling design

A SEIMS-based hydrological model is constructed by the SEIMS main program and several user-configured ~~modules of~~ hydrologic process modules, defined via configuration files~~ in a text-based format~~ rather than hard-coding~~ed~~. The main program controls the simulation time loop ~~of the simulation~~ and repeatedly invokes the selected modules in sequence. Each module is responsible for ~~the~~ simulati~~o~~n~~ of~~g one or ~~several~~ more hydrologic processes on the corresponding simulation units in ~~different~~ a specific ~~simulation~~ order~~s~~. For instance, the potential evapotranspiration module executes on every grid cell in an arbitrary order, the subsurface flow (~~also referred to as~~or interflow) routing module executes on every grid cell following topological~~by~~ layering orders (Liu et al., 2014), and the channel routing module executes on each channel following the~~by~~ upstream–downstream ~~orders~~sequence.

In SEIMS, each module must ~~should~~ handle its required simulation conditions independently. Users are responsible for ensuring~~should decide~~ the compatibility ~~integration feasibility~~ of ~~the~~ module combinations to meet the requirements of each module, such as the boundary conditions ~~when calculating~~for vertical and lateral water movements. For instance, to implement a semi-physically-based interflow module based on the Darcy's Law, the boundary condition could be set either with the hydraulic gradient equal to the slope at each cell, or dynamically determined using the outputs of adjacent cells provided by other modules.

Each module inherits from a standard module interface, which includ~~ing the~~es definitions of input data, parameters, and output data, and the core execution code. The input data and parameters of one module can be either read from the database created during the preprocessing stage or ~~referred~~ retrieved from other modules during the runtime. Each module manages the ~~itself~~ initializ~~es~~ation of its own output data. In this way, ~~during the execution of the SEIMS-based model,~~ the same variable is shared among modules during the model execution ~~of a SEIMS-based model~~.

(3) Two-level parallelization strategy

The first level of the two-level parallelization strategy ~~dispatches~~ distributes the simulation of subbasins to different computing processes (or nodes) through the Message Passing Interface (MPI) (Liu et al., 2016). The second level is achieved within each subbasin by dispatching the simulations of independent grid cells ~~without mutual dependencies~~ to computing threads via Open Multi-Processing (OpenMP) (Liu et al., 2014). Based on the design of the modular structure, variables required to be communicated among subbasins only need to be defined in the metadata of each module and will be handled by the SEIMS main program (Zhu et al., 2019). This~~e~~ two-level parallelization strategy provides the foundational capability ~~potential~~ to implement the proposed approach, enabling the ~~in this study that~~ application of des~~-~~ ~~different~~ istinct model structures to ~~distinct~~ different subbasins.

2.2.2 Constructing model structures with diverse simulation units and algorithms

To enable the flexible construction of model structures ~~based on~~within SEIMS, the diversity of both simulation units and simulation algorithms should be ~~guaranteed~~ensured. For simulation units, hydrological response units (HRUs) were added as simulation units ~~o~~for hillslope processes ~~in SEIMS~~, including DominantHRU (defined by dominant combination of soil types, land use types, and slope classes), FullHRU (where one HRU may consist of several spatially ~~separated~~ disjoint areas), and ExplicitHRU (where one HRU represents ~~is~~ a spatially independent and contiguous area) (Figure 1a; Figure 2a). P~~The p~~hysically-based parameters for these units are ~~can be~~ derived from actual properties of the soil, landuse/landcover, and other spatial data, while conceptual parameters are directly specified by lookup tables of soil and land-use types. For instance, the storage capacity of an HRU can be ~~represented~~ defined by the soil depth associated with that specific HRU area, whereas empirical ~~while some experimental~~ coefficients ~~could be specified~~are typically provided in ~~the~~ lookup tables. Under such an HRU discretization scheme, ~~the~~ hillslope ~~unit~~ and subbasin units can be regarded as specific instances of DominantHRUs. Therefore, the extended SEIMS ~~will~~ offers the flexibility to~~of~~ utilize~~ing~~ subbasins, hillslopes, HRUs, and grid cells as simulation units ~~o~~for hillslope processes.

For hydrological process representations, SEIMS was ~~initially~~ originally designed to primarily implement ~~the~~ individual simulation algorithm of one hydrological process as a single module, ~~while it~~ lacking a mechanism to handle lumped~~ed~~ ~~a~~ conceptual models~~ capable of~~ that inherently integrate~~considering~~ multiple hydrological processes ~~comprehensively~~ (Figure 2b). In this study, lumped conceptual models were ~~designed~~ adapted by decomposing them into~~to be integrated as~~ two separate hillslope modules: (1) a surface runoff module that generates runoff at each basic simulation unit; and (2) a surface routing module that conveys runoff to the subbasin outlet. For instance, the representative widely used ~~conceptual model,~~ GR4J (Perrin et al., 2003) can be adapted~~, is suitable to be incorporated~~ as a surface runoff generation module, coupled with a separate routing module such as a simple accumulation method. Such a conceptual module could be applied to any of the aforementioned simulation unit types~~ stated above~~.

To ensure ~~the combination~~ compatibility between~~of~~ simulation algorithms and ~~with~~ simulation units~~ applied to subbasins~~, the SEIMS module interface was extended to explicitly flag~~mark its~~ computational dependency requirements. C~~A c~~omputationally dependent simulation modules are~~ is~~ exclusively applicable to grid cells, whereas~~and a~~ computationally independent modules are compatible with~~ is applicable to any types of~~ all simulation unit types. Therefore, any~~A~~ model structures incorporating~~ containing any~~ a computationally dependent module is restricted~~can only be applied~~ to ~~the~~ subbasins that utilize~~ using~~ grid cells as their basic simulation units.

2.2.3 Allocating individual model structures to subbasins and integrating as a spatially hybrid watershed model structure

An essential part of the spatially hybrid hydrological modeling approach is ~~to~~ enabling~~e~~ the separate allocation and execution of ~~individual~~ distinct model structures for~~to~~ different subbasins (referred to as subbasin models) within a ~~the~~ watershed. All subbasin models are subsequently integrated to form ~~as~~ a complete watershed model. ~~The~~ While the requirements of flexible model configuration and subbasin-separate simulation are compatible with the text-based module configuration method and two-level parallelization strategy of SEIMS, the existing architecture required specific extensions~~but still need improvement~~. The SEIMS module configuration file ~~of SEIMS~~ was extended to allow the assignment of~~designate~~ a specific model structure to ~~the specific~~ each subbasin~~, and the adopted type of simulation units~~ (Figure 3). Accordingly~~In this way~~, the SEIMS main program was ~~extended~~ modified to read the configuration file for each subbasin dynamically, load the appropriate simulation modules, and retrieve the modeling parameters according to the specified basic simulation units.

This subbasin-specific configuration mechanism naturally facilitates ~~S~~subbasin-specific parameter calibration. To manage parameter consistency across different subbasins, the framework allows users to define “parameter groups.” Within each group, a designated set of subbasins can be linked such~~o~~ that a shared parameter is constrained to a single, synchronized value across all member subbasins during the optimization process. This feature enables users to enforce cross-subbasin parameter consistency based on hydrological understanding or data quality, while still permitting differential calibration in other areas to account for spatial heterogeneity or compensation for source data uncertainties.

The two-level parallelization strategy was improved in two aspects. The first aspect involves the ~~is~~ task scheduling of ~~the~~ subbasin-level parallelization for load balancing, handled by the SEIMS main program. D~~The d~~omain decomposition of subbasins is now~~was~~ determined at runtime based on~~by~~ upstream–downstream relationships between subbasins and the number of basic simulation units within~~of~~

each subbasin. This replaces the original method which relied on the static count ~~in the runtime, rather than using the numbers~~ of grid cells of each subbasin during~~in~~ the preprocessing stage ~~in the original version of SEIMS~~ (Liu et al., 2016). The second aspect concerns ~~the~~ parallelization at the basic simulation unit level, which is handled in computationally independent modules. This is inherently supported and ~~easy~~ straightforward to implement, as ~~the~~ OpenMP ~~for~~ loop constructs can be ~~also~~ applied~~y~~ to newly added basic simulation units such as HRUs.

2.3 Implementation

Based on the above method design, the implementation of the proposed approach within SEIMS involved~~s~~ modifications to the data preprocessing tools, SEIMS main programs, and SEIMS module library. The data preprocessing tools mainly include a ~~collection~~ suite of Python scripts and C++ programs designed to delineate spatial units at ~~different~~ various scales (e.g., subbasins and hillslopes), extract spatial parameters of spatial units, and create watershed modeling database (Zhu et al., 2019).

In this study, a configurable tool was implemented to support the delineation and parameterization of HRU-based spatial units. This tool allows users to specify environmental variables (e.g., the default land-use and soil types) for defining~~to delineate~~ HRU-based units. This tool also allows specifying base spatial units and delineation strategy to generate different types of HRU such as taking subbasins or hillslopes as base units to generate DominantHRU or FullHRU. Therefore, the data preprocessing tools could prepare various types of spatial unit and associated spatial parameters for each subbasin according to the user-defined configurations~~ file~~ to meet various hydrological modeling scenarios. The ~~procedures ~~workflow ~~o~~for spatially hybrid modeling with SEIMS ~~are ~~is depicted in Figure 3.

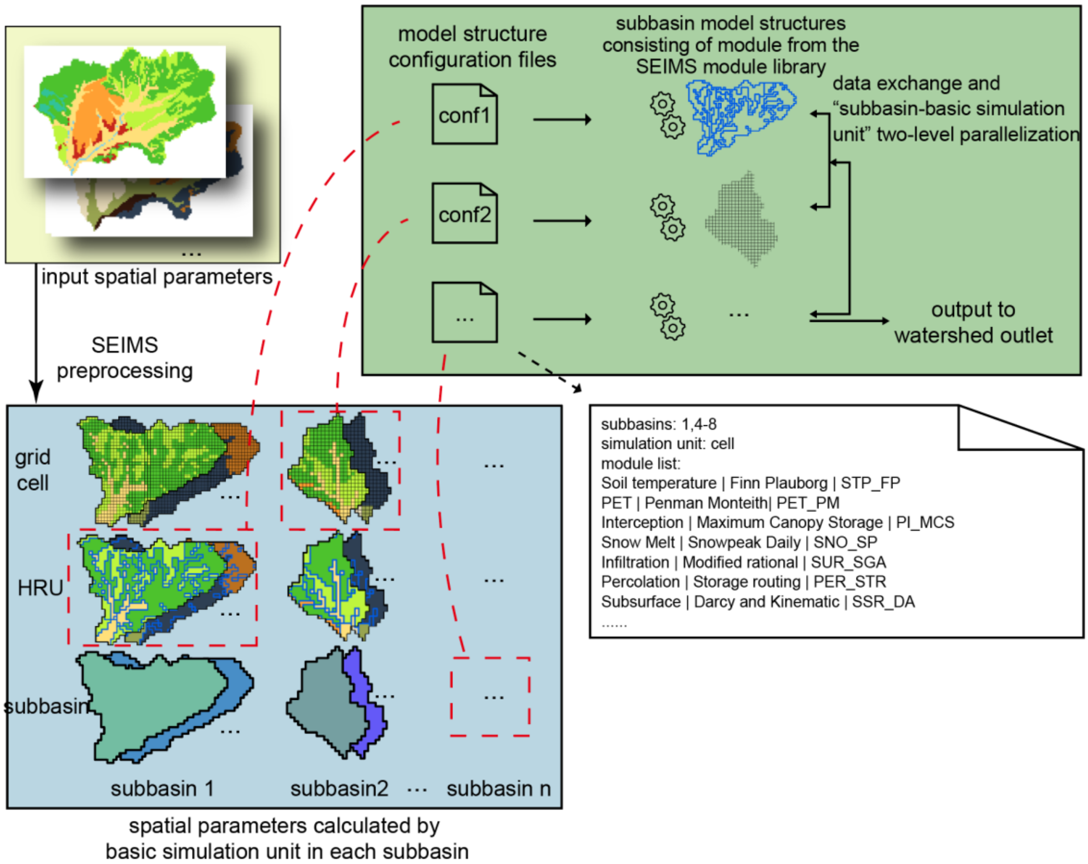


Figure 3. Workflow~~Procedures~~ ~~o~~for spatially hybrid modeling with~~ the~~ SEIMS (Spatially Explicit Integrated Modeling System)

The SEIMS main program, written in C++, was extended to dynamically read the configuration file for each subbasin, load simulation modules and retrieve the modeling parameters based on~~according to~~ the specified simulation units. Existing s~~S~~imulation modules in~~of~~ the SEIMS module library were updated to declare their

~~type of~~ computational dependency types. New conceptual simulation modules were added such as GR4J for simulating hillslope runoff generation.

With these~~ above briefly introduced~~ implementations~~ of the proposed approach~~, SEIMS ~~can~~ now supports the ~~users in~~ construction of~~g~~ spatially hybrid model structures that explicitly account for watershed ~~for considering~~ spatial heterogeneity.~~ of the watershed,~~ E~~where e~~ach subbasin ~~may~~ can be simulated using distinct~~with different~~ combinations of conceptual or semi-physically-based simulation modules and spatial units. ~~Consequently~~Currently, ~~more than~~ over 40 modules are available ~~after the implementation~~following~~ of~~ this study (detailed in Table S1 of the supplementary material), supporting a comprehensive range ~~simulation~~ of processes:~~ including~~ snowfall, atmospheric deposition, snow balance, interception, soil temperature, glaciers, surface runoff, evapotranspiration, infiltration, percolation, interflow, groundwater, channel routing, plant growth, and soil erosion.

~~The~~ SEIMS is open-source and hosted on GitHub (https://github.com/lreis2415/SEIMS), where it remains~~ and is~~ under continuous development.

3 Case study

3.1 Study area and data

The case study area is the Babao River watershed in~~at~~ Qilian, Qinghai Province, China (Figure 4), a high-altitude, cold and mountainous region with an area of approximately 2,511 km^2. The average elevation is 3,565 m, and the region features glaciers, snow cover, and frozen soil. The datasets used in the case study are ~~listed~~ described below.

The MERIT DEM with ~~the~~ a resolution of 90 m (Yamazaki et al., 2017) was selected for watershed delineation and the calculation of terrain attributes. Meteorological data was obtained from the China Meteorological Assimilation Driving Datasets for the SWAT model (CMADS) version 1.2, provided as gridded data with a~~in the form of~~ resolution of approximately 0.125° ~~resolution gridded station data~~ (Meng et al., 2017). Land cover data was derived from GLOBELAND30 with a resolution of 30 m (http://www.globallandcover.com). Soil attribute data was sourced from the Harmonized World Soil Database (HWSD) Version 2.0 (Nachtergaele, 2023). The observed daily discharge at~~of~~ the outlet gauge from 2013 to 2018 was obtained from the National Hydrological Yearbook.

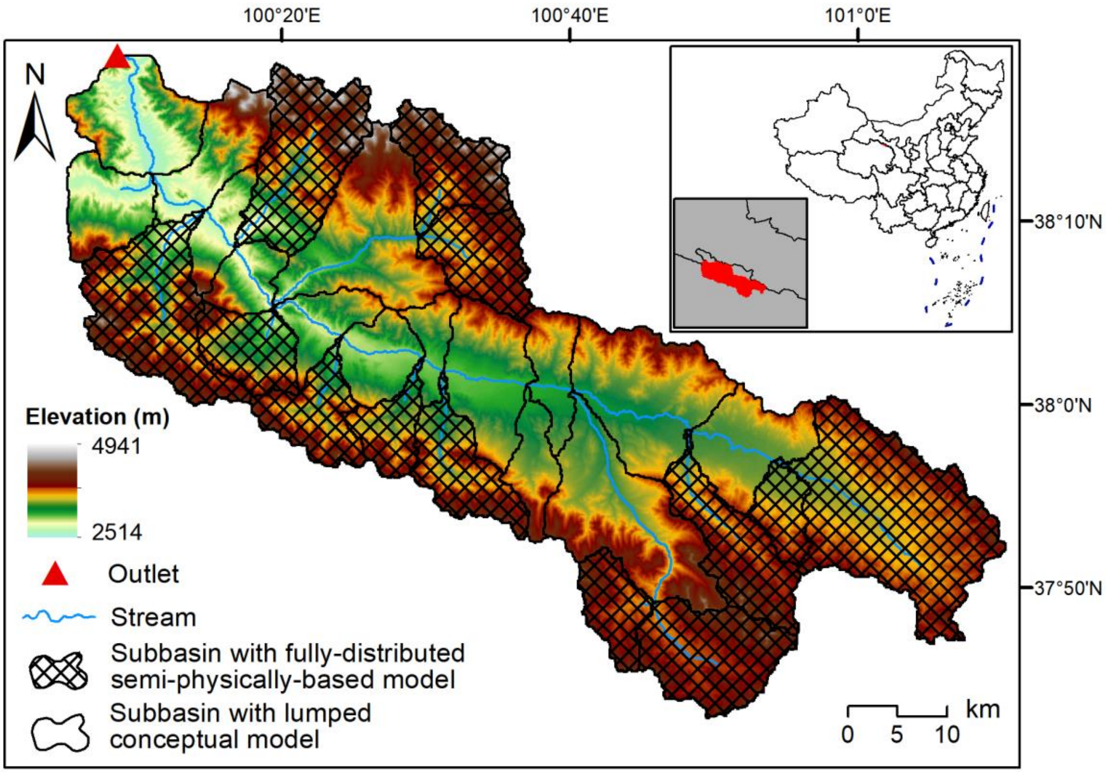


Figure 4. The Babao River watershed and the spatial constitution of the spatially hybrid watershed model.

3.2 Experimental design

To ~~verify~~ evaluate the integration feasibility and simulation effectiveness of the proposed spatially hybrid hydrological modeling approach and its implementation ~~using~~ within SEIMS, we constructed one fully-distributed semi-physically-based model structure (S~~s~~ection 3.2.1) and one lumped conceptual model structure (S~~s~~ection 3.2.2). These were used to build one spatially hybrid model and two spatially consistent models (S~~s~~ection 3.2.3) for comparison after parameter calibration (S~~s~~ection 3.2.4). It is ~~worth~~ important to note~~ing~~ that the watershed models constructed in this comparative experiment are ~~not~~ intended primarily as a proof-of-concept. They do not aim to achieve a perfect reproduction of~~model the~~ hydrological processes in the study area~~precisely~~, nor to enumerate ~~the~~ all possible model structures applicable as illustrated in Figure 1.

3.2.1 The fully-distributed semi-physically-based model structure

The fully-distributed semi-physically-based model structure uses grid cells as the basic simulation units and encompasses simulation algorithms based on physical laws. For instance, ~~the algorithm for~~ percolation is calculated as the product of hydraulic conductivity and the gradient of the hydraulic potential, and ~~the~~ interflow is calculated from the kinematic approximation of Darcy's Law, ~~and,~~ with the hydraulic gradient assumed equal to the slope at each grid cell. These two simulation algorithms are the same as those used in WetSpa (Water and Energy Transfer between Soil, Plants and Atmosphere; Liu and Smedt, 2004). Table 1 lists the ~~considered~~ hydrological processes considered,~~and~~ their simulation algorithms, and ~~associated with~~ key parameters for calibration. The physically based simulation algorithms primarily utilize spatial parameters with clear physical meaning, such as soil hydraulic conductivity. Besides, this model structure also includes conceptual simulation algorithms such as ~~the~~ a unit hydrograph-based surface routing module ~~based on unit hydrograph~~ (Table 1).

Table 1. Simulation algorithms adopted for the fully-distributed semi-physically-based model structure and the parameters involved in model calibration

Hydrological process	Simulation algorithm	Parameters involved in calibration
Potential evapotranspiration	Penman-Monteith	K_pet (correction factor)
Interception	Maximum canopy storage	Interc_max (maximum interception storage), Interc_min (minimum interception storage), Pi_b (interception storage capacity exponent)
Glacier	HBV method	-
Snow melt	HBV method	T_rain_snow_delta (rain/snow mixture temperature range)
Infiltration and surface runoff	Modified coefficient method	K_run (runoff exponent), P_max (maximum precipitation corresponding to potential runoff coefficient), Runoff_co (potential runoff coefficient)
Depression	Linsley	Depression (depression storage capacity)
Percolation	Storage routing	-
Interflow	One-dimension kinematic wave	Ki (interflow scale factor)
Actual evaporation	Thornthwaite and Mather method	-
Plant growth	Simplified EPIC	-
Groundwater	Linear reservoir	Base_ex (baseflow recession exponent), df_coef (deep percolation coefficient), gwmax (maximum ground water storage), Kg (baseflow recession coefficient)
Surface routing	Geomorphology-based unit hydrograph	-
Channel routing	Muskingum	MSK_X, MSK_K

3.2.2 The lumped conceptual model structure

The lumped conceptual model structure adopts DominantHRUs as the simulation units and the GR4J as the ~~main~~ core simulation algorithm (Table 2). The Hargreaves method is used to estimate the potential evapotranspiration for the GR4J. The GR4J receives the potential evapotranspiration and precipitation as input to generate the hillslope runoff at the outlet of the simulation unit (i.e., DominantHRUs), ~~which~~ implicitly lumping ~~generalizes~~ the internal surface routing process. The runoff from all DominantHRUs is then aggregated to the subbasin outlet using a simple accumulation method. The inter-subbasin channel routing process is modeled using the Muskingum algorithm. In this case study, each subbasin corresponds to a single DominantHRU, defined by the land cover and soil type combination ~~that~~ occupying~~ies~~ the largest area within the subbasin. As a result, GR4J functions effectively~~can also be applied~~ without requiring an additional surface routing module.

Table 2. Simulation algorithms adopted for the lumped conceptual model structure and the parameters involved in model calibration

Hydrological process	Simulation algorithm	Parameters involved in calibration
Potential evapotranspiration	Hargreaves	-
Surface runoff	GR4J	SOILTHICK, X2, X3, X4
Surface routing	Simple accumulation method	-
Channel routing	Muskingum	MSK_X, MSK_K

3.2.3 One spatially hybrid and two spatially consistent model structures for comparison

~~Ideally, physically based models can better simulate hydrological processes with explicit physical representations. However, due to practical limitations, such as incomplete understanding of hydrological processes and the difficulty of implementing complex simulation algorithms and deriving essential modeling data, a conceptual model can serve as a valuable complement.~~ In this case study, we constructed a spatially hybrid model structure (hereafter referred to as ~~the~~ HybM;

Figure 4) by combining both the fully-distributed semi-physically-based (hereafter referred to as PhyM, see Section 3.2.1) and the lumped conceptual models (hereafter referred to as ConM, see Section 3.2.2). S~~The s~~Subbasins were first ~~classified~~categorized ~~into two categories~~ based on the topographic statistics. Specifically, ~~-a preliminary assessment of topography:~~ headwater subbasins with a mean elevation above 3400 m and a mean~~n average~~ slope greater than 10° were classified as mountainous subbasins (13 subbasins, shown in Figure 4), while the remaining downstream subbasins were classified as gently sloping subbasins (16 subbasins). Accordingly, ~~We considered the~~ PhyM ~~fully-distributed semi-physically-based model structure~~ was assigned to the~~more suitable for~~ mountainous subbasins~~ (i.e., 13 subbasins shown in Figure 4)~~, while ~~applying the lumped conceptual model structure~~ ConM was applied to the gently sloping subbasins~~ (i.e., 16 subbasins shown in Figure 4)~~. We acknowledge that these specific classification thresholds and corresponding model assignments serve primarily as a proof-of-concept for the proposed approach and may require~~s~~ further verification in practical modeling contexts.

For comparison with ~~the~~ HybM, ~~we also applied~~ PhyM and ConM were also applied individually to the entire watershed as ~~constructed two~~ spatially consistent model structures~~: PhyM, which uses only the fully-distributed semi-physically-based model structure, and ConM, which uses only the lumped conceptual model structure~~.

3.2.4 Comparative experiments

The comparative experiments ~~primarily~~ aim to assess the post-calibration performance of the three watershed model structures. Parameters for model calibration were chosen through prior sensitivity analyses (listed in Table 1 and Table 2). In principle, the proposed approach adheres to the atomistic forward modeling concept: each subbasin-level model structure should be determined based on the appropriate representation of hydrological processes and available data, ensuring a physically plausible~~reasonable~~ and practically executable watershed model. Ideally, model calibration should utilize multi-scale observations (e.g., at both subbasin and watershed outlets) to avoid reliance~~ying~~ solely on the inverse calibration assumption~~–that successfully watershed-scale reproduction validates subbasin-scale model structures~~. Successfully reproducing watershed outlet dynamics does not necessarily imply correctness of ~~that the~~ subbasin~~-scale~~ models~~-structures are correct; multi-site data are needed to constrain the model's internal consistency~~.

In the absence of multi-scale observations, it is a common practice in current hydrological modeling to calibrate based solely on the watershed outlet data. In this case study, t~~only the watershed outlet discharge data was available. T~~o partially mitigate this limitation, we designed two calibration strategies (i.e., universal and regional). The universal calibration strategy involves a uniform adjustment of parameters ~~for~~ across all subbasin models, ~~while~~ whereas the regional calibration strategy applies separate parameter sets to ~~models of~~ mountainous ~~subbasins~~ and gently sloping subbasins. Consequently, ~~the~~ HybM was calibrated using the regional strategy, while both PhyM and ConM were calibrated using both universal and regional strategies (Table 3).

Table 3. Experiments of parameter calibration with different calibration strategies for different watershed model structures

Calibration strategy	Name of experimental cases		
	ConM	PhyM	HybM
Universal	ConM1	PhyM1	-
Regional	ConM2	PhyM2	HybM

All three models were executed over a six-year simulation (January 1, 2013–December 30, 2018). The first year (2013) served as a warm-up period, followed by a three-year calibration period (2014–2016) and a two-year validation period (2017–2018). Model performance was evaluated using three standard indices: the Nash-Sutcliffe coefficient (NSE; Eq. 1), root mean square error-standard deviation ratio (RSR; Eq. 2), and percent bias (PBIAS; Eq. 3).

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \quad (1)$$

$$RSR = \frac{\sqrt{\sum_{i=1}^{n}(O_i - P_i)^2}}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2}} \quad (2)$$

$$PBIAS = \frac{\sum_{i=1}^{n}(O_i - P_i) \times 100}{\sum_{i=1}^{n} O_i} \quad (3)$$

where O_i and P_i are *i*-th observed value and predicted value, respectively. $\bar{O}$ is the averaged observed value, and *n* is the size of simulated time series.

~~The~~ P~~p~~arameter calibration experiments were conducted using the NSGA-II (non-dominated sorting genetic algorithm-II; Deb et al., 2002) integrated within the SEIMS framework (Zhu et al., 2019). The optimization parameters setting for each calibration experiment included a generation of 25 and a population of 360. The multi-objective optimization aimed to maximize NSE and minimize RSR and absolute value of PBIAS. Experiments were carried out on a server equipped with dual ~~a 2.70 GHz~~ Intel Xeon Gold 6150 ~~dual~~ CPUs (2.70 GHz, 36 cores).

3.3 Experimental results and discussion

The integration feasibility of the ~~constructed~~ spatially hybrid watershed model structure is demonstrated by the successful results of the comparative experiments. To evaluate the simulation effectiveness of the proposed approach, a representative ~~one~~ set of calibrated parameters from~~of~~ each experiment (detailed in Table S2 of the supplementary material) was selected for analysis~~comparison~~, focusing on ~~analyzing~~ both model performance metrics and the hydrograph characteristics.

3.3.1 Model performance metrics of calibrated model structures

Table 4 lists the model performance metrics for both the calibration and validation periods of calibrated model structures under the universal and regional calibration strategies.

Table 4. Model performance metrics of calibrated model structures under different calibration strategies (PhyM1 and PhyM2 represent fully-distributed semi-physically-

based model structure using universal and regional calibration, respectively; ConM1 and ConM2 represent lumped conceptual model structure using universal and regional calibration, respectively; HybM represents the spatially hybrid model structure using regional calibration; NSE: Nash-Sutcliffe coefficient; RSR: root mean square error-standard deviation ratio; PBIAS: percent bias; values are reported as calibration/validation)

Experimental case	**Performance metrics of calibration/validation periods**		
	NSE	RSR	PBIAS (%)
PhyM1	0.58/0.60	0.64/0.63	10.63/2.58
PhyM2	0.59/0.60	0.64/0.63	17.34/13.90
ConM1	0.41/0.32	0.77/0.82	2.89/-42.87
ConM2	0.45/0.50	0.74/0.71	10.96/-21.03
HybM	0.72/0.60	0.53/0.63	10.02/-18.28

The spatially consistent fully-distributed semi-physically-based models (PhyM) exhibited similar NSE (~0.59) and RSR (~0.63) in both calibration and validation periods, regardless of the calibration strategy. However, PhyM1 (universal calibration) yielded a lower absolute PBIAS than PhyM2 (regional calibration). While regional calibration theoretically ~~generally~~ allows for more flexible parameter adjustments across heterogeneous subbasins, the higher overall bias observed in PhyM2 may be attributed to the lack of internal subbasin observation data. Without such internal constraints, regional calibration may lead to spatially inconsistent simulation errors that accumulate toward the watershed outlet, thereby amplify~~increas~~ing the overall bias. In contrast, universal calibration often benefits from compensatory effects among subbasins, resulting in a lower overall bias at the watershed scale.

In contrast to the stability of ~~Unlike the effect of calibration strategies on~~ PhyM, the spatially consistent lumped conceptual models (ConM) showed moderate improvements under the regional calibration strategy (ConM2) compared to the universal calibration strategy (ConM1). Specifically, for the calibration/validation periods, the NSE increased from 0.41/0.32 to 0.45/0.50 and the RSR decreased from 0.77/0.82 to 0.74/0.71. For PBIAS, ConM2 yielded 10.96%/-21.03%, showing improved validation bias compared to ConM1 (2.89%/-42.87%), despite a higher bias during the calibration ~~-~~period ~~bias~~. Conceptual models like ConM, characterized by~~with~~ simple model structure and few parameters, are often considered easier to calibrate and potentially benefit more from~~perform better under~~ regional calibration. However, in this study, both ConM1 and ConM2 exhibited relatively low NSE and high RSR values in both periods, indicating the limited capability of the GR4J-based conceptual model structure to represent the hydrological processes in this mountainous watershed, particularly because it does not account for glacier and snowmelt processes. The noticeable decline in validation performance for both ConM1 and ConM2 (e.g., the substantial shift in PBIAS from overestimation to underestimation) further reflects the limited generalization capability of ConM. The contrasting responses of ~~results between~~ ConM and PhyM to calibration strategies may be attributed to their structural differences in handling spatial variability. ConM relies ~~more~~ heavily on regional calibration to ~~consider~~ account for spatial heterogeneity ~~heterogenous~~ across subbasins by adjusting its less physically constrained parameters. In contrast, PhyM captures spatial heterogeneity directly through spatially distributed input data and parameters, implying a low dependency on calibration strategies to represent spatial variations in hydrological processes.

The spatially hybrid model structure (HybM) demonstrated ~~a~~ balanced ~~model~~ performance by leveraging the strengths of both PhyM and ConM, achieving the highest NSE (0.72/0.60) and lowest RSR (0.53/0.63) among all experimental cases. However, its PBIAS performance (10.02%/-18.28%) fell between ConM and PhyM, i.e., it was superior to ~~was lower than (i.e., better overall bias) that of~~ both ConM cases but inferior to ~~higher than (i.e., worse overall bias)~~ both PhyM cases. This reflects that ~~the spatially hybrid model structure~~HybM provides a promising compromise between modeling flexibility and simulation accuracy. That is, while HybM combined the calibration responsiveness of ConM and the process realism of PhyM, it also inherited some of their respective limitations, such as biases associated with ~~the~~ conceptual representations. Crucially, the NSE improvement of HybM is driven by the complementary mechanisms of the hybrid structure rather than simply increased parameter degrees of freedom (i.e., more calibrated parameters). This is~~,~~ evidenced by the fact that HybM achieved a significantly higher NSE (0.72) using fewer calibrated parameters (26) than the regional PhyM2 (40 parameters). Overall, these findings highlight the importance of careful model structure construction and the necessity of multi-site and long-term observation data to effectively constrain calibration and fully realize the benefits of the proposed approach.

3.3.2 Hydrographs of calibrated model structures

Figure 5 compares the outlet hydrographs simulated by PhyM1, ConM2, and HybM against the observed discharge. During the dry seasons (typically November to April), when baseflow dominates streamflow, PhyM1 better simulates baseflow magnitudes, benefiting from its process-based representation of hydrological processes. In contrast, ConM2 tends to overestimate baseflow, likely due to its simplified recession formulation and lack of explicit representation of groundwater storage and release dynamics. During the wet season, characterized by high-intensity precipitation events, PhyM1 produces sharper rising and falling limbs, generally capturing the dynamics of observed peaks and recessions. This behavior reflects the physical realism of PhyM in representing rapid surface and subsurface responses in the study area. Conversely, ConM2 often either overestimates or underestimates peak flows, likely due to limitations in its conceptual routing and storage structure.

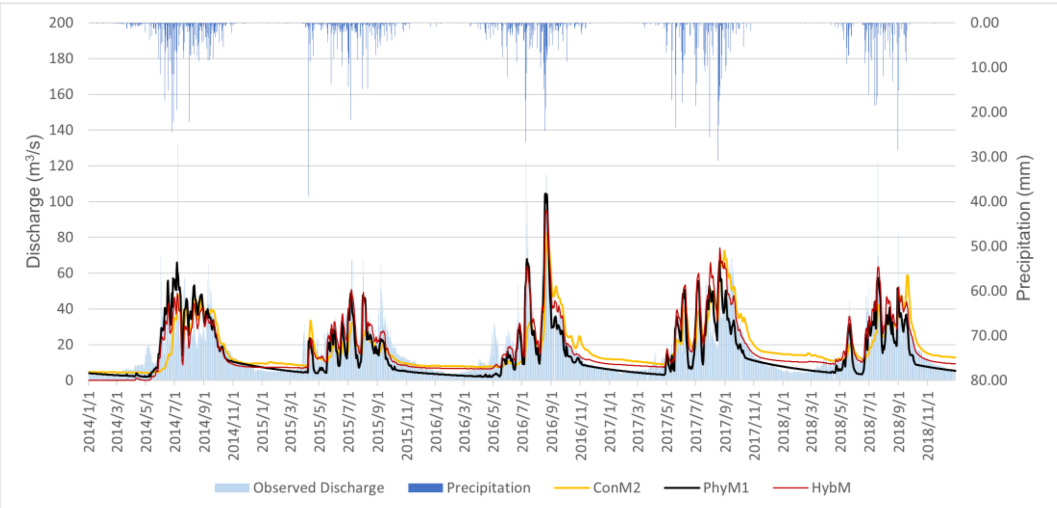


Figure 5. Simulated outlet hydrographs of best-performing cases for each model structure: PhyM1 (fully-distributed semi-physically-based model structure using universal calibration strategy), ConM2 (lumped conceptual model structure using

regional calibration strategy), and HybM (spatially hybrid model structure using regional calibration strategy)

The spatially hybrid model (HybM) effectively integrates the strengths of both model structures. HybM more accurately captures peak magnitudes and recession processes throughout the simulation period, exhibiting rising limbs similar to PhyM1 and a moderated baseflow response that is less overestimated than in ConM2. Although HybM still overestimates baseflow compared to observations, its improved flood response and recession simulation suggest a more balanced simulation across different hydrological conditions.

4 Discussion

The case study results ~~have~~ demonstrate~~d~~ the integration feasibility and simulation effectiveness of the proposed spatially hybrid hydrological modeling approach. These results ~~reflect~~ highlight an inherent trade-off in this approach, that is, enhanced flexibility and local adaptability often come at the cost of increased structural complexity and potential uncertainty. Nevertheless, the proposed approach offers a practical and adaptable framework to address spatial heterogeneity in diverse hydrological modeling application contexts, particularly where data availability and watershed characteristics vary spatially~~across space~~. This section further discusses the~~its~~ broader implications of the study, focusing on the ~~significance of the~~ advantages of SEIMS in implementing the proposed approach, as well as its ~~and~~ limitations and future directions.

4.1 Advantages of SEIMS in implementing the proposed approach

SEIMS effectively facilitates ~~supports~~ the implementation of the spatially hybrid modeling approach through its modular and subbasin-independent architecture, offering both modeling flexibility and computational efficiency.

The architecture of SEIMS~~'s design~~ enables model developers to reuse or adapt existing model codes as SEIMS modules to simulate individual or multiple hydrological processes. For instance, the GR4J module was adapted from the Raven framework (Craig, 2020). Newly developed modules can be flexibly combined with existing ones to construct ~~various~~ diverse model structures. This capability is particularly valuable for large and heterogeneous watersheds where modeling application contexts vary across regions. For instance, in subbasins with limited spatial data, lumped conceptual components can re~~be employed in~~ place ~~of~~ data-intensive distributed model structures to ensure model feasibility. It is worth noting that while SEIMS provides this flexibility, module developers are responsible for specifying input requirements and boundary conditions~~ for their modules~~, and users should assess the hydrological ~~rationality~~ plausibility and integration feasibility of their combined model structure.

Regarding computational efficiency, in the case study, the ~~fully-distributed semi-physically-based~~PhyM model ~~structure~~ required approximately 75 hours to complete 9000 model runs for the calibration experiment, whereas the ConM~~lumped~~

~~conceptual~~ model ~~structure~~ took only about 1 hour. The HybM~~spatially hybrid~~ model, incorporating both structures, achieved a balanced efficiency of around 30 hours. This computational efficiency, enabled by the parallelization strategies of SEIMS, supports ~~more~~ scalable applications in large watersheds, particularly those involving extensive model runs such as uncertainty analysis, parameter calibration, and scenario analysis.

4.2 Limitations and future directions

This section highlights key limitations and outlines future directions to enhance the ~~its~~ practical applicability of the proposed approach.

First, the successful application of the proposed approach implemented within SEIMS largely depends on the breadth and quality of its module library. Though extensible, the module library lacks modules for several critical hydrological processes. For example, while simplified representations of cold-region processes (e.g., disabling subsurface flow when soil temperature is below freezing) were used in this study, a dedicated, physically-based frozen soil dynamics module is not yet available. The library also lacks other specific representations, such as fully-physically-based simulation algorithms that tightly couple hillslope hydrological processes.~~(e.g., frozen soil dynamics) and specific representations (e.g., fully-physically-based simulation algorithms that tightly couple hillslope hydrological processes).~~ Recent research efforts are ~~continue to~~ expanding the SEIMS module library to support more complex ~~hydrological~~ processes, for instance, by modeling groundwater–surface water interactions between alpine runoff and alluvial aquifers in high mountain watersheds (Liu et al., 2025). We plan to gradually implement a variety of fully-physically-based simulation modules, such as a 2D hillslope overland flow module inspired by FullSWOF (Full Shallow-Water equations for Overland Flow; Delestre et al., 2017), and a fully coupled 3D module of overland flow, subsurface flow, and groundwater flow following approaches from SHUD (Shu et al., 2020) and ParFlow (Kollet and Maxwell, 2006). We encourage the community ~~model developers~~ to contribute to the enrichment of the SEIMS module library, either by ~~transplanting~~ porting existing code or developing new modules based on established hydrological equations and the SEIMS module template.

Second, while the subbasin-independent architecture of SEIMS is a key feature for~~to~~ implementing the proposed approach, it ~~also~~ raises concerns ~~about~~ regarding inter-subbasin connectivity in ~~for more~~ complex modeling ~~circumstances~~scenarios. For instance, when multiple flow direction algorithms are used to construct the hillslope flow routing network, flow paths may cross subbasin boundaries, complicating~~making and the~~ inter-subbasin connectivity~~ between subbasins becomes more complexfrequent than only connected by the channel network~~. Future work should explore flexible inter-subbasin connectivity mechanisms to ~~consider~~ account for ~~more~~ realistic ~~hydrological~~lateral interactions~~ across subbasins~~, such as integrating watershed-scale grid-based groundwater module similar to *gwflow* in SWAT+ (Bailey et al., 2022). Furthermore~~Besides~~, while the inter-subbasin coupling operates at a consistent subbasin scale~~ through the same or different channel routing algorithms~~, this study did not explicitly address the coupling of processes across nested spatial and temporal scales within each subbasin and the transfer of uncertainties between them, which remain~~s~~ central challenges in the hydrology community. The proposed approach and the implemented modeling framework are extensible to support such future research~~ in the future~~.

Third, while the proposed approach enables flexible configuration of model structure across subbasins, it currently lacks intelligent methods and tools to reduce the substantial modeling burdens on users (particularly non-expert users), when constructing appropriate model structures for specific application contexts. –For instance, i~~I~~n this study, ~~we manually classified~~ subbasins were manually classified ~~were classified~~ using specific topographic thresholds to assign different model structures. While this– ~~(e.g., headwater subbasins with mean elevation > 3400 m and mean slope > 10° as “mountainous”); this~~ servess as one feasible example, it~~ but~~ may not represent the optimal or universally applicable criteria. Future efforts should focus on developing domain knowledge-driven methods that incorporate various knowledge types, such as: (1) integration knowledge of SEIMS modules derived from module metadata; (2) rule-based relationships linking application contexts (e.g., watershed topographic characteristics and available data) to suitable simulation algorithms; and (3) case-based reasoning to determine appropriate simulation units, algorithms, and parameter settings (Jiang et al., 2019; Qin et al., 2025).

Fourth, the current case study was designed as a simplified proof-of-concept

~~application~~ to demonstrate ~~the~~ methodological feasibility ~~of the proposed approach~~. Nevertheless, more comprehensive case studies in watersheds with multi-scale observed data and diverse climatic, topographic, and landuse conditions are needed to further evaluate the robustness, generalizability, and uncertainty implications of the proposed approach. Specifically, –incorporating multi-site observational data would help constrain individual subbasin simulations, thereby reducing the reliance on potential error compensation between upstream and downstream subbasins. Furthermore~~or instance~~, future applications should incorporate uncertainty analysis frameworks to better understand and quantify the trade-offs between modeling flexibility and simulation accuracy.

5 Conclusion

This study ~~proposed~~ presented a new spatially hybrid hydrological modeling approach to address the challenge of representing spatial heterogeneity in complex hydrological modeling application contexts. Adopting a divide-and-conquer strategy, the approach ~~allows~~ enables the~~for~~ flexible combination of compatible spatial discretization schemes (i.e., simulation units) and hydrological process representations (i.e., simulation algorithms) across subbasins. This approach provides a pragmatic and flexible modeling solution, facilitating ~~to enable~~ the integration of existing hydrological modeling methods at the subbasin scale.

The approach was implemented within the SEIMS framework, leveraging its modular, subbasin-independent architecture and parallel computing capabilities. Comparative experiments in a medium-sized watershed demonstrated the~~its~~ integration feasibility and simulation effectiveness of the proposed approach. The spatially hybrid model structure successfully combined~~s~~ the strengths of both spatially consistent model structures, achieving balanced ~~improving general~~ performance improvements, while ~~also~~ inheriting ~~some~~ limitations from its constituent components. The experimental results highlight ~~its~~ the capability of the approach to balance modeling flexibility and simulation accuracy under heterogeneous watershed conditions.

This study does not ~~intend~~ claim to resolve the fundamental hydrological issues such as the scaling issues (Dooge, 1986; Beven, 2000), nor to ~~ignore~~ overlook their ~~existence~~significance, but rather to provide a practical and extensible framework for hydrological modeling across diverse application contexts. The proposed approach and its implementation are extensible through the reuse and integration of existing modeling efforts, thus holding the potential to support investigations into these theoretical issues as new theories and data become available. Moreover, this study lays the groundwork for future research on knowledge-driven intelligent modeling methods and tools for complex watershed modeling and management in spatially heterogeneous environments.

Acknowledgments

This work was supported by National Natural Science Foundation of China [42101480, 42471499], Strategic Priority Research Program of the Chinese Academy of Sciences [XDB0740200], and LREIS [YPI005, KPI003]. We would like to acknowledge Dr. Kaiwen Wang from the IGSNRR, CAS, for his valuable discussions on physically-based hydrological modeling.

CRediT authorship contribution statement

Yu-Jing Wang: Writing – original draft, Software, Methodology, Investigation. **Liang-Jun Zhu**: Writing – original draft, Writing – review & editing, Software, Methodology, Investigation, Funding acquisition, Conceptualization. **Cheng-Zhi Qin**: Writing – review & editing, Methodology, Investigation, Funding acquisition, Conceptualization. **A-Xing Zhu**: Supervision, Conceptualization.

Declarations of interest: none.

Reference

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O’Connell, P.E., Rasmussen, J., 1986. An introduction to the European Hydrological System — Systeme Hydrologique Europeen, “SHE”, 2: Structure of a physically-based, distributed modelling system. Journal of Hydrology 87, 61–77. https://doi.org/10.1016/0022-

1694(86)90115-0

Arnold, J. G., Allen, P. M., Volk, M., Williams, J. R., & Bosch, D. D., 2010. Assessment of different representations of spatial variability on SWAT model performance. Transactions of the ASABE 53, 1433–1443. https://doi.org/10.13031/2013.34913

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large Area Hydrologic Modeling and Assessment Part I: Model Development. JAWRA Journal of the American Water Resources Association 34, 73–89. https://doi.org/10.1111/j.1752-1688.1998.tb05961.x

Bailey, R.T., Abbas, S., Arnold, J., White, M., Gao, J., Cerkasova, N., 2022. Augmenting the National agroecosystem model with physically based spatially distributed groundwater modeling. Environmental Modelling & Software 105589. https://doi.org/10.1016/j.envsoft.2022.105589

Beven, K.J., 2000. Uniqueness of place and process representations in hydrological modelling. Hydrology and Earth System Sciences 4, 203–213. https://doi.org/10.5194/hess-4-203-2000

Butts, M.B., Payne, J.T., Kristensen, M., Madsen, H., 2004. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. Journal of Hydrology 298, 242–266. https://doi.org/10.1016/j.jhydrol.2004.03.042

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied hydrology, 2nd ed.

818 McGraw-Hill, New York.

819 Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A.,
820 Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D., Arnold, J.R., Gochis,
821 D.J., Rasmussen, R.M., 2015. A unified approach for process-based hydrologic
822 modeling: 1. Modeling concept. Water Resources Research 51, 2498–2514.
823 https://doi.org/10.1002/2015WR017198

824 Coron, L., Thirel, G., Delaigue, O., Perrin, C., Andréassian, V., 2017. The suite of
825 lumped GR hydrological models in an R package. Environmental Modelling &
826 Software 94, 166–171. https://doi.org/10.1016/j.envsoft.2017.05.002

827 Craig, J.R., 2020. Flexible watershed simulation with the Raven hydrological
828 modelling framework. Environmental Modelling & Software 129, 104728.

829 David, P.C., Chaffe, P.L.B., Chagas, V.B.P., Dal Molin, M., Oliveira, D.Y., Klein,
830 A.H.F., Fenicia, F., 2022. Correspondence Between Model Structures and
831 Hydrological Signatures: A Large-Sample Case Study Using 508 Brazilian
832 Catchments. Water Resources Research 58, e2021WR030619.
833 https://doi.org/10.1029/2021WR030619

834 Deb K, Pratap A, Agarwal S, et al., 2002. A fast and elitist multiobjective genetic
835 algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6: 182–
836 197.

837 Dehotin, J., Braud, I., 2008. Which spatial discretization for distributed hydrological
838 models? Proposition of a methodology and illustration for medium to large-scale

catchments. Hydrology and Earth System Sciences 12, 769–796. https://doi.org/10.5194/hess-12-769-2008

Delestre, O., Darboux, F., James, F., Lucas, C., Laguerre, C., Cordier, S., 2017. FullSWOF: Full Shallow-Water equations for Overland Flow. Journal of Open Source Software 2, 448. https://doi.org/10.21105/joss.00448

Dooge, J.C.I., 1986. Looking for hydrologic laws. Water Resources Research 22, 46S-58S. https://doi.org/10.1029/WR022i09Sp0046S

Fenicia, F., Kavetski, D., Savenije, H.H.G., 2011. Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. Water Resources Research 47. https://doi.org/10.1029/2010WR010174

Fenicia, F., Kavetski, D., Savenije, H.H.G., Pfister, L., 2016. From spatially variable streamflow to distributed hydrological models: Analysis of key modeling decisions. Water Resources Research 52, 954–989. https://doi.org/10.1002/2015WR017398

Gao, H., Hrachowitz, M., Fenicia, F., Gharari, S., Savenije, H.H.G., 2014. Testing the realism of a topography-driven model (FLEX-Topo) in the nested catchments of the Upper Heihe, China. Hydrology and Earth System Sciences 18, 1895–1915. https://doi.org/10.5194/hess-18-1895-2014

Gharari, S., Gupta, H.V., Clark, M.P., Hrachowitz, M., Fenicia, F., Matgen, P., Savenije, H.H.G., 2021. Understanding the Information Content in the Hierarchy of Model Development Decisions: Learning From Data. Water Resources

Research 57, e2020WR027948. https://doi.org/10.1029/2020WR027948

Harpham, Q.K., Hughes, A., Moore, R.V., 2019. Introductory overview: The OpenMI 2.0 standard for integrating numerical models. Environmental Modelling & Software 122, 104549. https://doi.org/10.1016/j.envsoft.2019.104549

Hill, C., DeLuca, C., Balaji, Suarez, M., Da Silva, A., 2004. The architecture of the earth system modeling framework. Computing in Science & Engineering 6, 18–28. https://doi.org/10.1109/MCISE.2004.1255817

Hrachowitz, M., Clark, M.P., 2017. HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. Hydrology and Earth System Sciences 21, 3953–3973. https://doi.org/10.5194/hess-21-3953-2017

Jiang, J., Zhu, A.-X., Qin, C.-Z., Liu, J., 2019. A knowledge-based method for the automatic determination of hydrological model structures. Journal of Hydroinformatics 21, 1163–1178. https://doi.org/10.2166/hydro.2019.029

Kiraz, M., Coxon, G., Wagener, T., 2023. A priori selection of hydrological model structures in modular modelling frameworks: application to Great Britain. Hydrological Sciences Journal 68, 2042–2056. https://doi.org/10.1080/02626667.2023.2251968

Kneis, D., 2015. A lightweight framework for rapid development of object-based hydrological model engines. Environmental Modelling & Software 68, 110–121. https://doi.org/10.1016/j.envsoft.2015.02.009

Knoben, W.J.M., Freer, J.E., Fowler, K.J.A., Peel, M.C., Woods, R.A., 2019. Modular

Assessment of Rainfall–Runoff Models Toolbox (MARRMoT) v1.2: an open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations. Geoscientific Model Development 12, 2463–2480. https://doi.org/10.5194/gmd-12-2463-2019

Kollet, S.J., Maxwell, R.M., 2006. Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. Advances in Water Resources 29, 945–958. https://doi.org/10.1016/j.advwatres.2005.08.006

Ley, R., Hellebrand, H., Casper, M.C., Fenicia, F., 2016. Is catchment classification possible by means of multiple model structures? A Case Study Based on 99 Catchments in Germany. Hydrology 3, 22. https://doi.org/10.3390/hydrology3020022

Li, X., Zhang, L., Zheng, Y., Yang, D., Wu, F., Tian, Y., Han, F., Gao, B., Li, H., Zhang, Y., Ge, Y., Cheng, G., Fu, B., Xia, J., Song, C., Zheng, C., 2021. Novel hybrid coupling of ecohydrology and socioeconomy at river basin scale: A watershed system model for the Heihe River basin. Environmental Modelling & Software 141, 105058. https://doi.org/10.1016/j.envsoft.2021.105058

Liu, Jiaojiao, Liu, Junzhi, Liu, M., Tian, S., Liu, Yongbo, Yang, W., Liu, Yonqin, 2025. Development of an alpine hydrological model considering the recharge of stream water to alluvial plain aquifers. Environmental Modelling & Software 192, 106567. https://doi.org/10.1016/j.envsoft.2025.106567

Liu, J., Liu, Z., Yin, Y., Croke, B. F. W., Chen, M., Qin, C.-Z., Tang, G., Zhu, A.-X, 2021. A hybrid vector-raster approach to drainage network construction in agricultural watersheds with rice terraces and ponds. Journal of Hydrology 601, 126585. https://doi.org/10.1016/j.jhydrol.2021.126585

Liu, Y., De Smedt, F., 2004. WetSpa extension, a GIS-based hydrologic model for flood prediction and watershed management. Vrije Universiteit Brussel, Brussel, Belgium.

Liu, J., Zhu, A.-X., Qin, C.-Z., Wu, H., Jiang, J., 2016. A two-level parallelization method for distributed hydrological models. Environmental Modelling & Software 80, 175–184. https://doi.org/10.1016/j.envsoft.2016.02.032

Liu, J., Zhu, A.-X., Liu, Y., Zhu, T., Qin, C.-Z., 2014. A layered approach to parallel computing for spatially distributed hydrological modeling. Environmental Modelling & Software 51, 221–227. https://doi.org/10.1016/j.envsoft.2013.10.005

Liu, Y., Zhang, K., Li, Z., Liu, Z., Wang, J., Huang, P., 2020. A hybrid runoff generation modelling framework based on spatial combination of three runoff generation schemes for semi-humid and semi-arid watersheds. Journal of Hydrology 590, 125440. https://doi.org/10.1016/j.jhydrol.2020.125440

Meng, X., Wang, H., 2017. Significance of the China meteorological assimilation driving datasets for the SWAT model (CMADS) of East Asia. Water, 9(10), 765. https://doi.org/10.3390/w9100765

Milad, J., Sobri, H., Mohsen, S., 2012. A Review on Theoretical Consideration and Types of Models in Hydrology. Journal of Environmental Science and Technology 5: 249–261.

Nachtergaele, F., van Velthuizen, H., Verelst, L., Wiberg, D., Henry, M., Chiozza, F., Yigini, Y., Aksoy, E., Batjes, N., Boateng, E., others, 2023. Harmonized world soil database version 2.0. Food and Agriculture Organization of the United Nations.

Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. Journal of Hydrology 279, 275–289. https://doi.org/10.1016/S0022-1694(03)00225-7

Pilz, T., Francke, T., Baroni, G., Bronstert, A., 2020. How to Tailor My Process-Based Hydrological Model? Dynamic Identifiability Analysis of Flexible Model Structures. Water Resources Research 56, e2020WR028042. https://doi.org/10.1029/2020WR028042

Qin, C.-Z., Zhu, L.-J., Zhu, A.-X., 2025. Position paper: Domain knowledge-driven intelligentization of watershed modeling and scenario analysis for precise watershed management. Information Geography 1, 100016. https://doi.org/10.1016/j.infgeo.2025.100016

Savenije, H.H.G., 2010. HESS Opinions “Topography driven conceptual modelling (FLEX-Topo).” Hydrology and Earth System Sciences 14, 2681–2692. https://doi.org/10.5194/hess-14-2681-2010

Savenije, H.H.G., 2009. HESS Opinions “The art of hydrology.” Hydrology and Earth System Sciences 13, 157–161. https://doi.org/10.5194/hess-13-157-2009

Shu, L., Ullrich, P.A., Duffy, C.J., 2020. Simulator for Hydrologic Unstructured Domains (SHUD v1.0): numerical modeling of watershed hydrology with the finite volume method. Geoscientific Model Development 13, 2743–2762. https://doi.org/10.5194/gmd-13-2743-2020

Sui, D., Turner, M, 2022. General theories and principles in geography and GIScience: Moving beyond the idiographic and nomothetic dichotomy. Annals of GIS, 28, 1–4. https://doi.org/10.1080/19475683.2022.2030939

Tague, C.L., Band, L.E., 2004. RHESSys: Regional Hydro-Ecologic Simulation System—An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling. Earth Interactions 8, 1–42. https://doi.org/10.1175/1087-3562(2004)8<1:RRHSSO>2.0.CO;2

Tang, G., Li, S., Yang, M., Xu, Z., Liu, Y., Gu, H., 2019. Streamflow response to snow regime shift associated with climate variability in four mountain watersheds in the US Great Basin. Journal of Hydrology 573, 255–266. https://doi.org/10.1016/j.jhydrol.2019.03.021

Teshager, A. D., Gassman, P. W., Secchi, S., Schoof, J. T., & Misgna, G., 2016. Modeling agricultural watersheds with the soil and water assessment tool (SWAT): Calibration and validation with a novel procedure for spatially explicit HRUs. Environmental Management 57, 894–911.

https://doi.org/10.1007/s00267-015-0636-4

Thébault, C., Perrin, C., Andréassian, V., Thirel, G., Legrand, S., Delaigue, O., 2024. Multi-model approach in a variable spatial framework for streamflow simulation. Hydrology and Earth System Sciences 28, 1539–1566.

van Esse, W.R., Perrin, C., Booij, M.J., Augustijn, D.C.M., Fenicia, F., Kavetski, D., Lobligeois, F., 2013. The influence of conceptual model structure on model performance: a comparative study for 237 French catchments. Hydrology and Earth System Sciences 17, 4227–4239. https://doi.org/10.5194/hess-17-4227-2013

Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V., Sorooshian, S., 2001. A framework for development and application of hydrological models. Hydrology and Earth System Sciences 5, 13–26. https://doi.org/10.5194/hess-5-13-2001

Wang, C., Hua, W., Chen, G., Fang, X., Li, X., 2021. Distributed-Framework Basin Modeling System: I. Overview and Model Coupling. Water 13, 678. https://doi.org/10.3390/w13050678

Wigmosta, M.S., Vail, L.W., Lettenmaier, D.P., 1994. A distributed hydrology-vegetation model for complex terrain. Water Resources Research 30, 1665–1679.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., et al., 2017. A high‐accuracy map of global terrain elevations. Geophysical Research Letters 44, 5844‐5853. https://doi.org/10.1002/2017GL072874

986 Zhu, L.-J., Liu, J., Qin, C.-Z., Zhu, A.-X., 2019. A modular and parallelized watershed modeling framework. Environmental Modelling & Software 122, 104526. https://doi.org/10.1016/j.envsoft.2019.104526