

Spatially hybrid hydrological modeling based on spatial heterogeneity of watershed characteristics

Yu-Jing Wang[a,b], Liang-Jun Zhu[a,b,*], Cheng-Zhi Qin[a,b,c,d,*], A-Xing Zhu[a,b,d,e,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] School of Geography and Tourism, Shaanxi Normal University, Xi'an, China

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

[e] Department of Geography, University of Wisconsin-Madison, Madison, WI, USA

[f] Key Laboratory of Virtual Geographic Environment (Ministry of Education of PRC), Nanjing Normal University, Nanjing, China

Corresponding author: Liang-Jun Zhu and Cheng-Zhi Qin

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

Abstract:

Hydrological modeling with a uniform model structure is often unreasonable for complex watersheds with significant spatial heterogeneity. For such a situation, a reasonable way is to integrate multiple hydrological model structures, each tailored to a specific area in the watershed of interest. While some pioneering hydrological modeling frameworks have enabled such integration of different hydrological model structures, they have primarily focused on lumped and conceptual models. The absence of distributed and physically-based models in such integration limits the capabilities of those frameworks to represent the spatially heterogeneous hydrological processes in detail. This paper proposes an innovative spatially hybrid hydrological modeling approach based on a divide-and-conquer idea. In this approach, the watershed is initially divided into model structure allocation units (MSAUs) at the subbasin level. Individual model structure can be flexibly composed of conventional conceptual and physically-based models, and then allocated to MSAU deemed suitable. Finally, these model structures at MSAU level are integrated into a spatially hybrid watershed model. The proposed approach is implemented through extending the Spatially Explicit Integrated Modeling System (SEIMS). A case study was conducted in a medium-sized natural watershed in Heihe River Basin, China, where two distinct model structures, containing the components from both lumped conceptual model and distributed physically-based model, were integrated into a spatially hybrid model, so to evaluate the effectiveness of the proposed approach. The results demonstrate that the spatially hybrid hydrological model integrated using the proposed approach could harness the strengths of both model structures on the simulation performance. This approach holds promise for improving the rationality of hydrological modeling in watersheds with significant spatial heterogeneity thus benefiting watershed research and decision supporting.

Highlights (3 to 5 bullet points (maximum 85 characters including spaces per bullet point)

- We proposed a spatially hybrid-structure approach for hydrological modeling.
- Model structures are spatially varied and flexibly constructed in the approach.
- The approach is implemented based on extensions to SEIMS.
- The spatially hybrid model outperformed two conventional uniform models.

Abstract:

Hydrological modeling with a uniform model structure is often unreasonable for complex watersheds with significant spatial heterogeneity. For such a situation, a reasonable way is to integrate multiple hydrological model structures, each tailored to a specific area in the watershed of interest. While some pioneering hydrological modeling frameworks have enabled such integration of different hydrological model structures, they have primarily focused on lumped and conceptual models. The absence of distributed and physically-based models in such integration limits the capabilities of those frameworks to represent the spatially heterogeneous hydrological processes in detail. This paper proposes an innovative spatially hybrid hydrological modeling approach based on a divide-and-conquer idea. In this approach, the watershed is initially divided into model structure allocation units (MSAUs) at the subbasin level. Individual model structure can be flexibly composed of conventional conceptual and physically-based models, and then allocated to MSAU deemed suitable. Finally, these model structures at MSAU level are integrated into a spatially hybrid watershed model. The proposed approach is implemented through extending the Spatially Explicit Integrated Modeling System (SEIMS). A case study was conducted in a medium-sized natural watershed in Heihe River Basin, China, where two distinct model structures, containing the components from both lumped conceptual model and distributed physically-based model, were integrated into a spatially hybrid model, so to evaluate the effectiveness of the proposed approach. The results demonstrate that the spatially hybrid hydrological model integrated using the proposed approach could harness the strengths of both model structures on the simulation performance. This approach holds promise for improving the rationality of hydrological modeling in watersheds with significant spatial heterogeneity thus benefiting watershed research and decision supporting.

Keywords: watershed simulation; hydrological modeling framework; hydrological model structure; spatial heterogeneity; SEIMS

1 Introduction

Hydrological models provide an effective way to simulate and understand complex hydrological processes in real-world watersheds. A variety of hydrological models exist, with each suitable for different modeling requirements dictated by the unique characteristics of a watershed and available data (Beven, 2000; Gharari et al., 2021; Wagener et al., 2001). This suitability is fundamentally shaped by the model structures they adopt. The structure of most hydrological models encompasses two aspects: the spatial discretization, and the hydrological processes representation. The spatial discretization involves simulating hydrological processes (e.g., evaporation, infiltration, and channel routing) on specific spatial units (e.g., the entire watershed, subbasin, channel lines, and grid cells). Meanwhile, the hydrological process representations, which encapsulates the translation of real-world hydrological processes into simulation algorithms, involves the selection of equations, numerical solutions, and the temporal scale (e.g., annual, monthly, daily, and hourly) of simulation. Therefore, the selection of an application-specific model structure stands as the primary and critical step in hydrological modeling, as implied in many model structure evaluation researches (Butts et al., 2004; David et al., 2022; Pilz et al., 2020; van Esse et al., 2013).

The model structure determined for a watershed can be categorized into two types, i.e., the uniform model structure, and the spatially varying model structure, according to whether they are spatially varying within the watershed of interest. Most research directly choose an existing model with a fixed or near-fixed model structure which is uniform for the whole study area. Such a model as well as its model structure are often developed with specific assumptions and intended for application in constrained catchment characteristics. Examples of such single models include HBV (Lindström et al., 1997), GR4J (Perrin et al., 2003), SWAT (Arnold et al., 1998), TOPMODEL (Beven and Kirkby, 1979), RHESSys (Tague and Band, 2004), and DHSVM (Wigmosta et al., 1994). Nevertheless, it is recognized that the uniform model structure lacks flexibility in changing simulation units or customizing alternative simulation algorithms of considered hydrological processes and thus cannot accommodate watershed applications under diverse circumstances (Ley et al., 2016; Savenije, 2009).

To accommodate complex application context with flexibility in controlling the model structure compositions, researchers have increasingly turned to modular hydrological modeling frameworks, such as ECHSE (Kneis, 2015), SUMMA (Clark et al., 2015), MARRMoT (Knoben et al., 2019), SEIMS (Spatially Explicit Integrated Modeling System; Liu et al., 2016; Zhu et al., 2019), and RAVEN (Craig, 2020). These frameworks have been developed to integrate the compositions of multiple existing models, allowing for easily customizing application-specific model structures. It is typical that each of the interdependent modules within a hydrological modeling framework simulates a specific part of the hydrological processes, and there are often multiple alternative algorithms for simulating one hydrological process. These modeling frameworks well support the concept of “multiple working hypotheses”, as advocated by Clark et al. (2011). However, it is important to note that these modeling frameworks primarily emphasize the application of a uniform model structure across the entire application area, with spatial variability confined to input data and model parameters. This way assumes that spatially heterogeneous data and parameters are sufficient to capture the behavior of hydrologic processes, while the model structure remains spatially uniform. While such a modeling way with spatially uniform model structure yields satisfactory results in many applications, the uniform model structure may be unreasonable and inaccurate to capture the characteristics of the dominant hydrological processes when the application context (i.e., the watershed characteristic or data availability) exhibit significant spatial heterogeneity.

Recognizing the limitation of the uniform-structure modeling approach, some approaches have explored to construct spatially varying model structure within the watershed under simulation. These approaches are mainly based on one of the two different ideas: externally integrating individual models through unified model interoperable interface, and internally integration within a hydrological modeling framework. Examples of the external integration are the works of Liu et al. (2020), Li et al. (2021), and Wang et al. (2021), as well as those works done through model integration frameworks such as ESMF (Hill et al., 2004) and OpenMI (Harpham et al., 2019). What they integrated are commonly individual models that simulates entire hydrological processes, from rainfall to runoff. However, as a great deal of interoperable interfaces often need to be implemented to integrate a model, the flexibility of the external integration may be limited. The structures constructed in this way are often fixed, and once the integrated model is established, it will require further coding efforts to change the spatial constitution of the model structure, which is difficult and inconvenient for most hydrologists.

To augment the modeling flexibility of spatially varying model structure, several studies suggested the internal integration approach based on the modular hydrological modeling frameworks. In this approach, model structures can be distributed to different modeling areas in the watershed, with only minor change in configuration or the main functions. This approach has been exemplified by works based on the FLEX-Topo framework (Gao et al., 2014; Savenije, 2010), SUPERFLEX framework (Fenicia et al., 2011, 2016), and airGR framework (Coron et al., 2017; Thébault et al., 2023). Nevertheless, the hydrological modeling frameworks employed in these studies are inherently limited to conceptual hydrological process representations, and their spatial discretization is also confined to subbasins or hydrologic response units (HRUs). Such a constraint on model structure poses a substantial challenge when attempting to integrate and utilize distributed and physically-based models or algorithms based on grid cells or patches.

As the state-of-the-art flexible approach to achieving spatially varying model structure, existing internal integration studies based on modular hydrological modeling framework only support lumped and conceptual model structures. This inability of framework in supporting distributed and physically-based model structures limits the flexibility and applicability during hydrological experimentation, comparative analyses, and applications pertaining to diverse model structures. To relieve such an inability, this paper proposes a novel spatially hybrid hydrological modeling approach by enabling the spatial allocation of various simulation units and simulation algorithms, comprising both conceptual and physically-based models, to different areas within a watershed. The remainder of the paper is organized as follows: Section 2 outlines the design of the proposed spatially hybrid modeling approach and its implementation based on the SEIMS framework. Section 3 discusses the case study, followed by presentation of results and discussion in Section 4, and conclusion in Section 5.

2 Method design and implementation

2.1 Basic idea

The core design of the spatially hybrid modeling involves two aspects: the first is the flexibility for model structure construction, which enables combination of different spatial discretization methods and hydrological process representations across the watershed of interest; the second is the ability to distribute individual model structures to different areas within the watershed, and integrate them to be executed under the same time loop.

The basic idea of designing the flexibly constructed model structure is to decouple the hydrological process representations (or specifically, the simulation algorithms) from being bound to particular spatial discretization methods (specifically, the simulation units). For instance, the physically- and process-based algorithms are typically applied to fine-scale simulation units such as grid cells for both are assumed to describe the hydrological process in detail. Similarly, conceptual algorithms are typically applied to lumped or coarse-scale units such as HRUs, hillslopes, or subbasins.

Such a binding is not necessary, and conversely, the combination of these model structure components have been acknowledged as effective. Hydrological modeling can be approached through two main paradigms: the bottom-up way, based on physical theories, and the top-down way, founded on empirical and conceptual deductions. Each of these approaches has its own set of advantages and disadvantages, depending on the circumstances (Clark et al., 2011; Hrachowitz and Clark, 2017). The fusion and compromise between the two ways could often yield advantages such as the ease of implementing, computational efficiency, and the data requirements, as demonstrated by various studies (Gunduz and Aral, 2005; Liu et al., 2020; Sidle, 2021) and popular models such as the SWAT and TOPMODEL.

The combination of distinct simulation units and algorithms is possible when the simulation algorithms are categorized as computational independent or dependent, according to whether the interaction among adjacent or upstream-downstream simulation units is required or not. An essential constraint on the applicability of simulation algorithms is that their computational dependencies must be satisfied by the spatial relationship of simulation units provided by chosen spatial discretization method. For example, the conceptual reservoir-based surface runoff generation algorithm with the physically-based kinematic wave-based surface routing can be integrated to be applied to grid cells, where computational dependency is satisfied by the flow directions between grid cells; the physically-based Richards equation-based surface runoff generation algorithm and the conceptual unit hydrograph-based surface routing algorithm can be integrated to applied on HRUs, where no computational dependency is required. Note that the computational dependency is only a basic constraint, the rationality of the model structure is to be determined by the modeler’s knowledge and requirements. Thus, diverse model structures could be formed and customized through combinations of compatible simulation units and simulation algorithms under the same hydrological modeling framework (Fig. 1a and Fig. 1b).

Figure 1 Design of the proposed spatially hybrid modeling approach, which involves a) incorporating diverse spatial discretization methods and hydrological process representations and b) supporting their flexible combinations into model structure on the MSAU level; c) enabling the spatial-varied allocation of these MSAU models.

Based on the flexible model structure construction, different areas within the watershed should then be assigned distinct application-specific model structures and executed after integration (Fig. 1c). In this study, the minimal unit for model structure allocation is named as a model structure allocation unit (MSAU), which is currently defined to be coincident with the subbasin unit. In other words, each subbasin adopts a model structure and all subbasin-level model structures collectively constitute the watershed model structure. These subbasin models are interconnected by the stream network with upstream-downstream relationships. The execution sequence of subbasin models can be determined accordingly. When the input requirements of all downstream MSAUs are satisfied with the output of upstream MSAUs, the models on each MSAUs can be integrated. Models on MSAUs without the upstream-downstream relationship could be parallelly executed, otherwise sequentially.

2.2 Design of the spatially hybrid hydrological modeling approach based on the hydrological modeling framework SEIMS

Following the basic idea, the proposed spatially hybrid modeling approach is designed based on the modular hydrological modeling framework SEIMS (Liu et al., 2016; Zhu et al., 2019), for its flexible modular design and “subbasin-simulation unit” two-level parallelization strategy. The detailed design of the proposed approach based on SEIMS includes two aspects (Fig. 2): (1) Enabling the construction of model structures with diverse simulation units and algorithms, which serves as the basis of the flexibility and diversity of the watershed model constructed. This requires SEIMS to be extended on the spatial discretization methods and upgrading the simulation modules to accommodate to different simulation units. (2) Enabling the spatial-varied allocation of individual model structures and their integrated execution.

Figure 2. Overall design of the spatially hybrid hydrological modeling approach based on extension to the Spatially Explicit Integrated Modeling System (SEIMS), which includes a) supporting various spatial discretization methods; b) supporting flexible application of the simulation unit to compatible simulation units; and c) supporting spatial-varied model structure allocation.

2.2.1 Brief introduction to the modular design and parallelization strategy of Spatially Explicit Integrated Modeling System (SEIMS)

SEIMS provides standard and concise interfaces to implement the simulation algorithms into modules, which involves hillslope process simulation modules on grid cells and channel routing process simulation modules on channel segments or channel grid cells. Each module exposes input and output information via metadata for the integration with other modules. Therefore, a list of user-configured modules comprises a SEIMS-based model in a loosely coupled manner. To support the flexible model structure configuration required in the proposed approach, the spatial discretization methods as well as the module library of SEIMS should be expanded (Fig.2a and 2b; detailed in Section 2.2.2).

The two-level parallelization strategy of SEIMS treated subbasin as relative independent spatial units for hydrological modeling dispatched to dedicated computing processes through the Message Passing Interface (MPI). The simulation tasks on grid cells within a subbasin are dispatched among computing threads via the Open Multi-Processing (OpenMP). Although the parallelization strategy underpins the relative independence of computations within each subbasin, thereby providing the potential to apply different model structures to distinct subbasins, the simulation modules and simulation units are predefined to be uniform for the whole watershed. These modules are loaded based on a configuration text file that sequentially records the modules involved. Therefore, the execution of SEIMS-based model should be upgraded to allocate different model structures to subbasins and executed independently (Fig. 2c; detailed in Section 2.2.3).

2.2.2 Constructing model structures with diverse simulation units and algorithms based on extension to SEIMS

To enable the flexible customization for a hybrid structure, the diversity of both spatial discretization methods and hydrological process representations should be guaranteed. For spatial discretization, the hydrological response unit (HRU) could be added as a simulation unit type in SEIMS, thus enabling the simulation based on not only grid cells but also HRUs. The HRUs, generated using soil and land-use maps, could follow the parameterization methodology of semi-distributed models and support both conceptual and physically-based parameters. Physically-based parameters can be derived from the actual properties of the soil or land-use data, while conceptual parameters are directly specified within the soil and land-use lookup tables. For instance, the reservoir capacity within an HRU can be represented by the soil depth associated with that specific HRU area.

Under such a HRU discretization method, the hillslope and subbasin simulation unit can be supported with SEIMS, as specialized cases of HRUs. For example, subbasin simulation is achieved when all HRUs within a subbasin share identical parameter values. With the above extension for spatial discretization method, SEIMS will offer the flexibility of utilizing subbasin, hillslopes, HRUs, and grid cells as spatial discretization options, allowing modelers to choose the most appropriate option for specific modeling needs.

The expansion of diverse discretization methods within SEIMS necessitates a corresponding extension of the hydrological process library to support the flexible combination of conceptual and physically-based processes across different discretization methods. SEIMS was designed to integrate algorithms simulating one specific processes, while it previously lacked a conceptual model capable of considering multiple hydrological processes comprehensively. To bridge this gap, the lumped conceptual models could be integrated as separate process-based modules: surface runoff modules that generate runoff at each simulation unit, and surface routing that directly convey water to the subbasin outlet. Such modules could be applied to any simulation unit in SEIMS, including grid cells, HRUs, hillslopes, and subbasins. Following this schema, a representative of widely used conceptual models, GR4J (Perrin et al., 2003), was incorporated into the SEIMS module library.

To ensure the compatibility of module combinations with the spatial discretization method applied to subbasins, the module interface of SEIMS is extended to mark its computational dependency requirement as one of the following two types.

- computational dependent: exclusively applicable to grid cells, where water is routed based on flow directions derived from digital terrain analysis. A model structure containing any computational dependent module can only be applied to the MSAU with the simulation unit of grid cell.
- computational independent: applicable to any simulation unit, including grid cells, HRUs, subbasins, and the entire watershed within SEIMS.

2.2.3 Allocating individual model structures to MSAUs based on SEIMS

An essential aspect of implementing the spatially hybrid hydrological modeling method is to enable the allocation of individual model structures to different MSAUs within the watershed. This entails allowing the separate construction and execution of individual models on each MSAU, referred to as MSAU models (or subbasin models in this study), which are subsequently integrated into a watershed model. This requires the MSAUs to be treated individually in an objective-oriented manner. To achieve flexible and swift model construction and modification, the structure configuration should be text-based rather than hard-coded.

Among the above-mentioned requirements, the subbasin-separate simulation and flexible model configuration are compatible to the two-level parallelization strategy and text-based module configuration of SEIMS, as illustrated in section 2.2.1. While extension should still be implemented on its module configuration method to support spatial-varying model structure allocation. To facilitate such an ability, the configuration files should be extended to include the subbasin number it is assigned to, and the discretization method the subbasin adopts, given that the MSAU is defined to be subbasins in this study. This allows an individual configuration when the computing processes of each subbasin dynamically read the configuration files, load each module, and retrieve the spatial parameters of each subbasin with the specified discretization method.

Extension should also be made on the design of the two-level parallel computing strategy to enable the execution of the individual subbasin models in SEIMS. In the original strategy, the simulation unit within a subbasin particularly refers to the grid cells, while simulation at the HRU level was not fully considered and implemented. To extend this strategy, the HRUs could be designed in a manner similar to grid cells but without topological orders, and are directly distributed to different computing threads using OpenMP. the subbasin-level parallelization could remain unchanged, with subbasins hierarchically sorted based on their flow relationships and subsequently assigned to different computing processes using MPI.

This enhanced strategy allows subbasin models to operate within the same time loop, enabling the parallel execution of hillslope processes among subbasins, with the output at each subbasin’s outlet exchanged with its downstream subbasin using the configured channel routing module. The routing modules of SEIMS are retained and not extended to allow for spatial variation, as the hillslope processes with spatial variability are considered sufficient to represent the heterogeneity of hydrological processes.

2.3 Current implementation

The implementation of the proposed approach with SEIMS involves modifications to both the preprocessing and the main modeling programs of SEIMS. In the original Python-based preprocessing program, spatial parameters of a watershed, including terrain, land-use, soil, and meteorological data, are distributed to grid cells organized by subbasins. The new implementation introduces support for HRU-based delineations, such as subbasins and hillslopes. This entails calculations of spatial parameters for multiple times for every subbasin and every type of simulation unit specified in the configuration file. The main modeling program and the module library of the SEIMS are based on the C++ language. They have been extended to distinguish and retrieve the spatial parameters, which are stored in MongoDB, corresponding to the simulation unit type of subbasin. The declaration of computational dependency is also added to every module in the module library.

With the SEIMS extensions introduced in this study, users can now construct diverse model structures by selecting alternative simulation modules, and assign these structures to any subbasin within a watershed all through text-based configuration editing. The SEIMS is open-source on GitHub (https://github.com/lreis2415/SEIMS) and under continuous development.

3 Case study

3.1 Study area and data

In this study, we selected the Babao River watershed at Qilian, Qinghai Province, China, in the upper reaches of the Heihe River Basin as the case study area to verify the proposed approach through daily-scale runoff simulation. The Babao River watershed Figure 3 is one of the headwaters of the Heihe River Basin. It is located in a high-altitude, cold and mountainous region with an area of approximately 2511 km^2. The average elevation is 3565 m, and the region features glacier, snow cover and frozen soil.

This study extensively relies on publicly available data. We used the 3" resolution (approximately 90 m) MERIT DEM sourced from Yamazaki et al. (2017) for watershed delineation and calculation of terrain attributes. Meteorological input data for the model were obtained from the China Meteorological Assimilation Driving Datasets for the SWAT model (CMADS) version 1.2, provided in the form of approximately 0.125 ° resolution gridded station data (Meng et al., 2019). Land cover data for the model were derived from GLOBELAND30 with a resolution of 30 m (http://www.globallandcover.com). Soil attribute input data for the model were sourced from the Harmonized World Soil Database (HWSD; Wieder, 2014).

Figure 3. The Babao River watershed in the upper reaches of the Heihe River Basin and the spatial constitution of the spatially hybrid model.

3.2 Experimental setting

Two individual model structures were adopted to construct the spatially hybrid model. One is a physically-based model structure using grid cells as the simulation unit, providing a more detailed representation of spatial and hydrological processes. This structure holds promise for simulating regions with significant elevation variations, such as mountainous areas within the watershed. The other was a conceptual model structure with HRUs as the simulation units, which offers simplicity and can generalize multiple hydrological processes. The HRU are derived from the land-use and soil maps. This structure was assumed to be well-suited for large, gently sloping terrain within the watershed, compared to the physically-based structure. The algorithms for hydrological process description of the two model structures are listed in Table 1. Note that the physically-based structure includes both physically-based modules (such as the interflow module that calculates the flow cell by cell) and conceptual modules (such as the surface routing module based on unit hydrograph), which reflects the flexibility of the supported model structures.

Table 1. The modules adopted for the physically-based model and the conceptual model.

However, since the lack of modules for permafrost and glaciers in SEIMS, a purely physically-based model, if applied, may be confronted with inaccuracies in representing certain processes. Therefore, a spatially hybrid model structure combining both physically-based and conceptual models can be considered in this watershed to achieve more accurate runoff simulations. Under this assumption, a combination of conceptual and physically-based model structures was allocated to different subbasins in the Babao River watershed, resulting in a spatially hybrid model, as depicted in Figure 3. Another two model structures were established for comparison. The first model exclusively utilized the physically-based model for all subbasins, referred to as the physically-based model. The second model employed the conceptual model structure for all subbasins, referred to as the conceptual model.

The three model structures described above were utilized in a four-year simulation, spanning from January 1, 2013, to December 30, 2016. The first year (2013) served as a warm-up period, followed by two years (2014 and 2015) designated for calibration, and the final year (2016) for validation. Various performance indices, including the Nash-Sutcliffe coefficient (NSE), root mean square error-standard deviation ratio (RSR), percent bias (PBIAS), and R-square, were selected to evaluate the modeling performance. Parameter calibration was conducted using the NSGA-II genetic algorithm integrated in the SEIMS framework, with a population size of 200 and optimization for 15 generations. The calibration objectives are maximum NSE, minimum absolute values of RSR and PBIAS. The P-factor and R-factor were used as additional indicators on the prediction uncertainty, where P-factor is the percentage of observations fall within the 95% prediction uncertainty (PPU) interval, and R-factor is the average width of the 95% PPU interval, normalized by the standard deviation of the observed data. The two factors are utilized to evaluate the final generations of the experimental model structures. Comparative analysis was carried out to assess the simulation performance of the three model sets based on the optimal parameters obtained during calibration.

4 Results and discussion

4.1 Quantitative performance of the spatially hybrid model structure constructed by the proposed approach

Figure 4 illustrates the calibration results of the three model structures, revealing that the spatially hybrid model (Figure 4c), established using the proposed approach, outperforms the others in terms of the best-fit simulations. The NSE of the best-fit spatially hybrid model is the highest (0.63 / 0.81 for calibration / validation period, the same below), compared to the physically-based model (0.34 / 0.73) and conceptual model (0.45 / 0.61). The lowest RSR (0.61 / 0.43), lowest PBIAS (3.66% / -4.62%) and highest R-square (0.64 / 0.83) are also seen in the spatially hybrid model, which indicates a best fit among the experimental model structures.

Figure 4. The simulation results of the final generation in calibration of the (a) physically-based model, (b) conceptual model, and (c) the spatially hybrid model from the proposed approach.

In terms of P-factor and R-factor, the spatially hybrid model shows a medium P-factor (0.42 / 0.30) and the lowest R-factor (0.60 / 0.49). In comparison, the physically-based models show the lowest P-factor (0.26 / 0.37) with a medium R-factor (0.65 / 0.57), which indicates a lower prediction uncertainty among the population compared to the conceptual models, while the accuracy is also low. The conceptual models, while having the highest P-factor (0.57 / 0.51), show a wide prediction interval with the highest R-factor (1.36 / 1.03). This indicates that although the observed discharge values are more within the prediction interval of conceptual models, the uncertainty is also high. The result shows the spatially hybrid model structure has a more concentrated prediction with the lowest uncertainty in the last generation during optimization. While uncertainty may originate from multiple sources including parameters and model structure, this result also implies a potential advantage in achieving more possibly accurate results through parameter optimization.

4.2 Discussion on the simulation results in terms of the hydrographs from the proposed approach

The strength of the spatially hybrid model becomes more evident when examining the visual interpretation of the simulation results (Figure 5). To elucidate this, we begin by discussing the disparities between the physically-based model and the conceptual model (Figure 4a and 4b). In dry seasons, typically from November to the following April when baseflow is the primary contributor to discharge, the physically-based model exhibits a noticeable underestimation of baseflow during the dry season between 2015 and 2016. This discrepancy suggests potential inaccuracies in representing groundwater processes. Conversely, the conceptual model provides an overall reasonable simulation but at the cost of overestimating baseflow. This overestimation can serve as compensation for low flows in wet seasons, which typically occur from May to October. In the wet seasons, the physically-based model produces steeper rising and falling limbs in the hydrographs, corresponding to the intensive precipitation during wet seasons, as depicted in Figure 5. This behavior aligns with its physical basis. The conceptual model, on the other hand, struggles to simulate the peak flows accurately, especially during the validation period.

Figure 5. Hydrograph of the best-fit spatially hybrid model.

The spatially hybrid model, as a result, combines the strengths of both models. While it tends to slightly overestimate baseflow, it outperforms both the physically-based model (which underestimates) and the conceptual model (which overestimates) and provides more precise and steeper peak flows during the wet season. This represents an enhancement to the advantages of the physically-based model. In areas with gentle terrain, groundwater tends to accumulate. However, during the period from October 2015 to May 2016, when overall precipitation was low, the physically-based model's process-based simulations may have resulted in an underestimation of groundwater levels. In contrast, the conceptual model, through adjustments to parameters governing groundwater outflow, can reconcile these levels with observed values. In the mountainous regions within the watershed, characterized by rapid runoff generation and routing, the conceptual model may not accurately capture these dynamics, resulting in less precise simulations of rainfall-induced runoff. However, this limitation is offset within the hybrid model due to the inclusion of the physically-based model.

An additional merit of the spatially hybrid model lies in the computational efficiency. Due to the detailed representation of various hydrological processes in the physically-based model, its calibration took approximately 100 hours for 3000 model runs (200 population and 15 generations). Such a high calibration demand limits its availability to larger watersheds. In contrast, the calibration time for the conceptual model was significantly shorter, approximately 15 minutes. Therefore, the spatially hybrid model, which contains both models, exhibited an intermediate calibration time of around 50 hours, proportionate to the participation of the two models. This efficiency enhancement suggests a broader applicability to larger watersheds, where specific subbasins of interest can be treated individually. This approach not only saves time but also enables the production of more model runs for parameter optimizations or uncertainty analysis.

5 Conclusion

This paper introduces a novel spatially hybrid hydrological modeling approach, offering a versatile solution to address the challenges posed by the spatial heterogeneity of complex watersheds. Unlike traditional distributed hydrological models that allow only spatial variability of parameters, the proposed approach takes a divide-and-conquer idea to accommodate spatially varied and hybrid model structures. It introduces the concept of Model Structure Allocation Units (MSAUs) for allocating individual model structures. This approach advocates the hybrid combination of different spatial discretization methods and hydrological process representations, enhancing structural flexibility on each MSAU. Constraints on the MSAU model structures are discussed to ensure their integration into the final watershed model.

Based on the implementation on the hydrological modeling framework SEIMS, an experiment was conducted to validate the effectiveness of the proposed approach. The spatially hybrid model structure, compared to uniform physically-based and conceptual model structures, demonstrated not only the capability of modeling with spatially varied and flexible structure, but also the ability to synergize the strengths of the constituent model structures, potentially leading to reduced uncertainty.

This approach empowers researchers to scrutinize and fine-tune model structures with precision, aligning them with the unique characteristics of specific areas within a watershed. It offers enhanced flexibility and a wider range of available model structures, extending the horizons of hydrological modeling for spatially varying model structures. This extension facilitates the exploration of multiple working hypotheses, ultimately enhancing our understanding of complex watershed systems.

However, there are also limitations in the current implementation. First, while the concept of MSAUs (model structure allocation units) is theoretically applicable to other simulation units, such as HRUs or coarse grid cells, further exploration is needed to assess its suitability on these units. Second, the SEIMS module library must undergo further development to encompass a broader spectrum of models and algorithms, particularly those related to glacier and frozen soil processes, in order to construct model structures for a wider range of applications in watersheds with significant spatial heterogeneity. Lastly, as the spatial constitution of the spatially hybrid model can be freely configured and revised, determining the optimal spatial constitution remains a subject for future study.

Acknowledgments

The support for A-Xing Zhu through the Vilas Associate Award, the Hammel Faculty Fellow Award, and the Manasse Chair Professorship from the University of Wisconsin-Madison is greatly appreciated.

Funding: This work was supported by National Key Research and Development Program of China [2021YFB3900904], National Natural Science Foundation of China [42101480], LREIS [YPI005, KPI003], and Shaanxi Normal University.

Declarations of interest: none.

Reference

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

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

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin 24, 43–69. https://doi.org/10.1080/02626667909491834

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

Clark, M.P., Kavetski, D., Fenicia, F., 2011. Pursuing the method of multiple working hypotheses for hydrological modeling. Water Resources Research 47. https://doi.org/10.1029/2010WR009827

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

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

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

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

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

Gunduz, O., Aral, M., 2005. Large-Scale Hybrid Watershed Modeling, in: Frevert, D., Singh, V. (Eds.), Watershed Models. CRC Press, pp. 74–95. https://doi.org/10.1201/9781420037432.ch4

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

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

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

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. Journal of Hydrology 201, 272–288. https://doi.org/10.1016/S0022-1694(97)00041-3

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, 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

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

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

Sidle, R.C., 2021. Strategies for smarter catchment hydrology models: incorporating scaling and better process representation. Geoscience Letters 8, 24. https://doi.org/10.1186/s40562-021-00193-9

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

Thébault, C., Perrin, C., Andréassian, V., Thirel, G., Legrand, S., Delaigue, O., 2023. Multi-model approach in a variable spatial framework for streamflow simulation. EGUsphere. https://doi.org/10.5194/egusphere-2023-569

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.

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

Declaration of interests

☑ 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.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Figure

Figure 1. Design of the proposed spatially hybrid modeling approach, which involves a) incorporating diverse spatial discretization methods and hydrological process representations and b) supporting their flexible combinations into model structure on the MSAU level; c) enabling the spatial-varied allocation of these MSAU models.

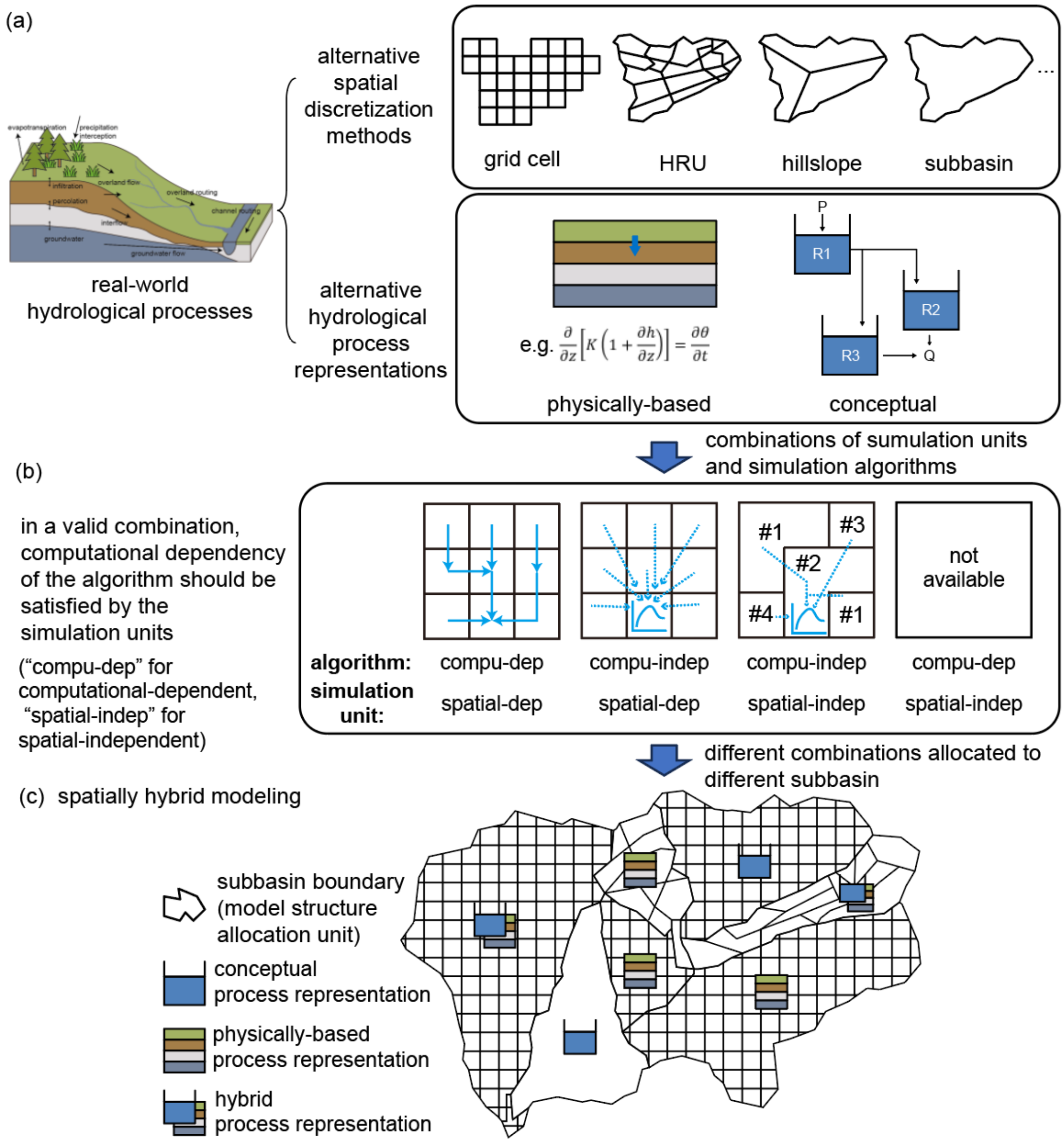

Figure 2. Overall design of the spatially hybrid hydrological modeling approach based on extension to the Spatially Explicit Integrated Modeling System (SEIMS), which includes a) supporting various spatial discretization methods; b) supporting flexible application of the simulation unit to compatible simulation units; and c) supporting spatial-varied model structure allocation.

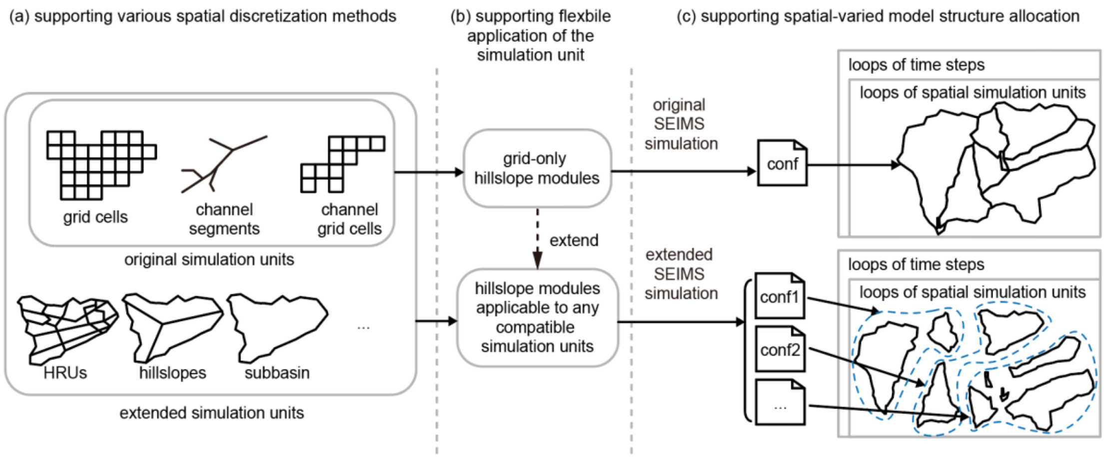

Figure 3. The Babao River watershed in the upper reaches of the Heihe River Basin and the spatial constitution of the spatially hybrid model.

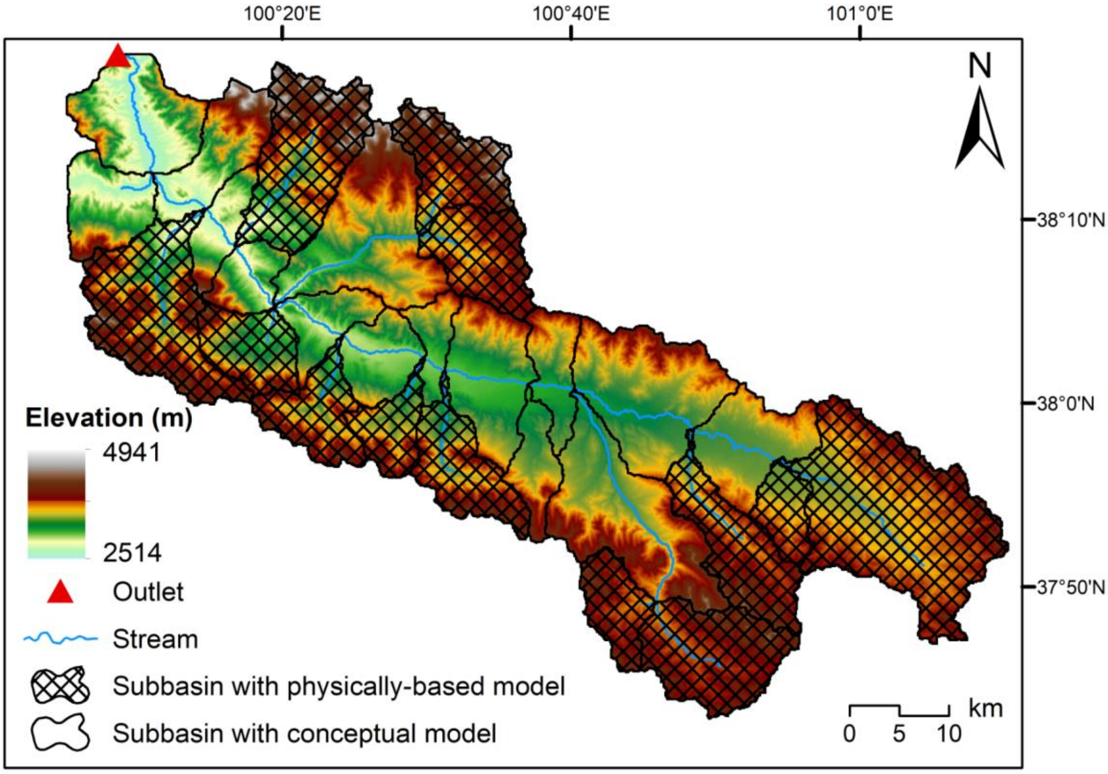

Figure 4. The simulation results of the final generation in calibration of the (a) physically-based model, (b) conceptual model, and (c) the spatially hybrid model from the proposed approach.

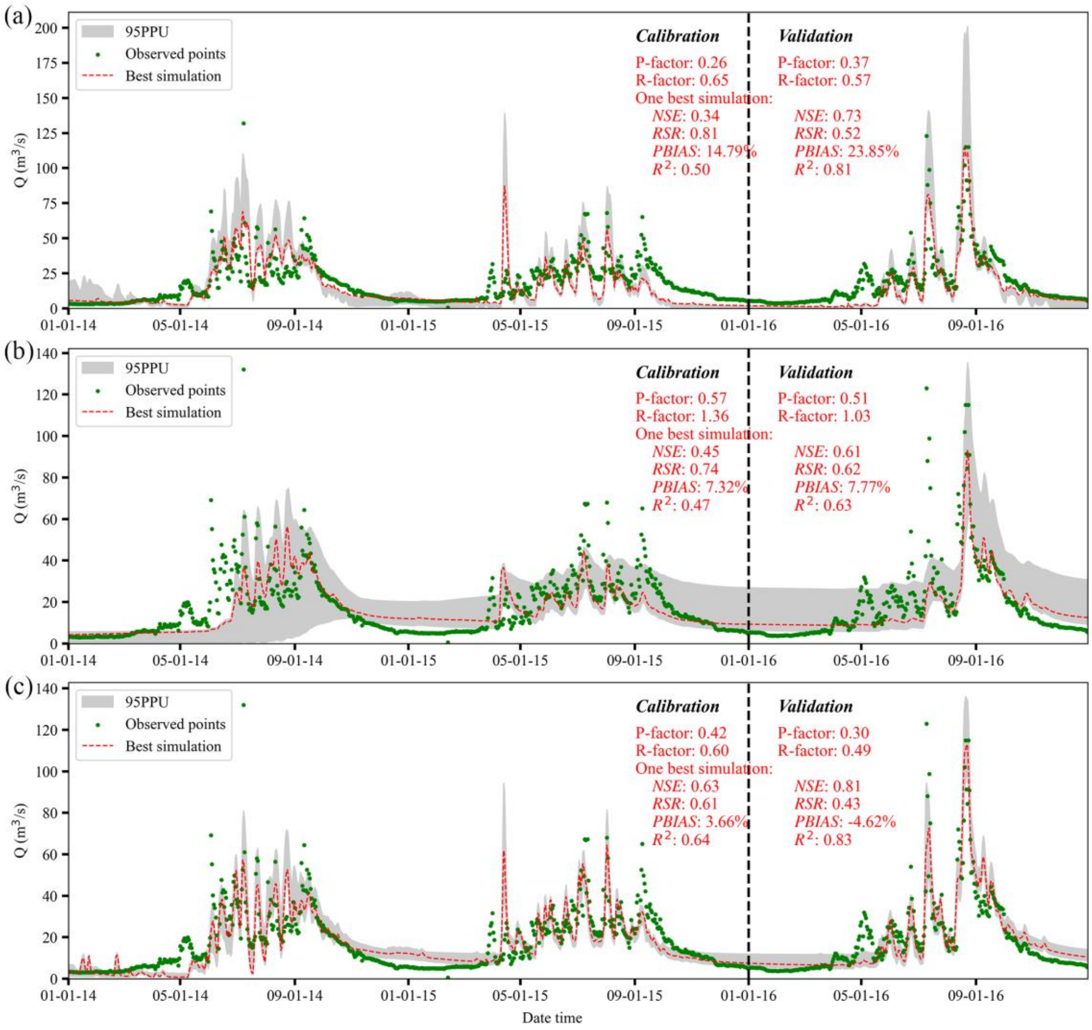

Figure 5. Hydrograph of the best-fit spatially hybrid model.

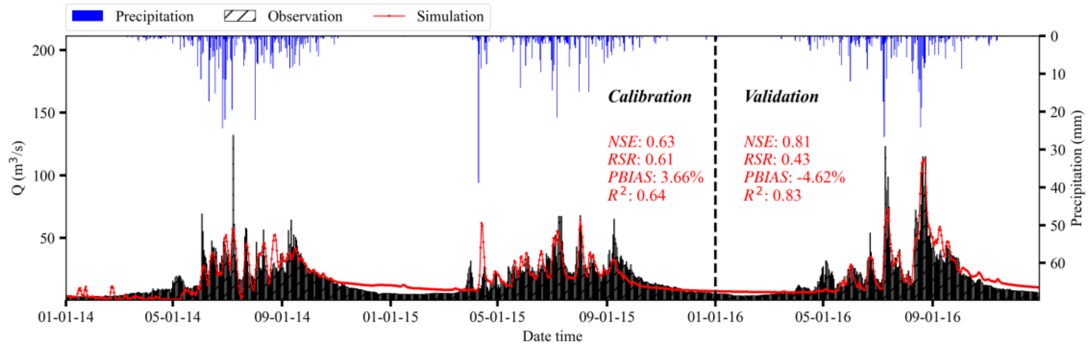

Figure 1. Design of the proposed spatially hybrid modeling approach, which involves a) incorporating diverse spatial discretization methods and hydrological process representations and b) supporting their flexible combinations into model structure on the MSAU level; c) enabling the spatial-varied allocation of these MSAU models.

Figure 2. Overall design of the spatially hybrid hydrological modeling approach based on extension to the Spatially Explicit Integrated Modeling System (SEIMS), which includes a) supporting various spatial discretization methods; b) supporting flexible application of the simulation unit to compatible simulation units; and c) supporting spatial-varied model structure allocation.

Figure 3. The Babao River watershed in the upper reaches of the Heihe River Basin and the spatial constitution of the spatially hybrid model.

Figure 4. The simulation results of the final generation in calibration of the (a) physically-based model, (b) conceptual model, and (c) the spatially hybrid model from the proposed approach.

Figure 5. Hydrograph of the best-fit spatially hybrid model.

Table 1. The modules adopted for the physically-based model and the conceptual model.

<table>
<tr><th>Hydrological processes</th><th>Modules of the physically-based model</th><th>Modules of the conceptual model</th></tr>
<tr><td>Conceptual and comprehensive surface runoff</td><td>(NA)</td><td>GR4J</td></tr>
<tr><td>Potential evapotranspiration</td><td>Penman Monteith</td><td>Hargreaves</td></tr>
<tr><td>Interception</td><td>Maximum canopy storage</td><td rowspan="10">(NA)</td></tr>
<tr><td>Snow melt</td><td>Snowpack Daily</td></tr>
<tr><td>Infiltration and surface runoff</td><td>Modified rational</td></tr>
<tr><td>Depression</td><td>Linsley</td></tr>
<tr><td>Percolation</td><td>Storage routing</td></tr>
<tr><td>Interflow</td><td>One-dimension kinematic wave</td></tr>
<tr><td>Actual evaporation</td><td>Hargreaves and Priestley Taylor</td></tr>
<tr><td>Plant growth</td><td>Simplified EPIC</td></tr>
<tr><td>Groundwater</td><td>Linear reservoir</td></tr>
<tr><td>Surface routing</td><td>Geomorphology-based unit hydrograph</td><td>GR4J</td></tr>
<tr><td>Channel routing</td><td>Muskingum</td><td>Muskingum</td></tr>
</table>