

1 **A spatially hybrid hydrological modeling approach based on**
2 **spatial heterogeneity of watershed characteristics**

3

4 **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}**

5 ^a State Key Laboratory of Resources and Environmental Information System, Institute of
6 Geographic Sciences and Natural Resources Research, CAS, Beijing, China

7 ^b College of Resources and Environment, University of Chinese Academy of Sciences,
8 Beijing, China

9 ^c School of Geography and Tourism, Shaanxi Normal University, Xi'an, China

10 ^d Jiangsu Center for Collaborative Innovation in Geographical Information Resource
11 Development and Application, Nanjing, China

12 ^e Department of Geography, University of Wisconsin-Madison, Madison, WI, USA

13 ^f Key Laboratory of Virtual Geographic Environment (Ministry of Education of PRC),
14 Nanjing Normal University, Nanjing, China

15

16 Corresponding author: Liang-Jun Zhu

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

18

1 **Abstract:**

2 The model structure of a hydrological model is primarily determined by the spatial
3 discretization scheme and the representation of hydrological processes. Using a
4 spatially consistent model structure is often unreasonable for complex watersheds
5 with significant spatial heterogeneity. Hydrological modeling framework is a
6 powerful tool to integrate multiple model structures, each tailored to a specific area of
7 the watershed. The state-of-the-art frameworks only support lumped conceptual
8 model structures and do not support model structures composed of spatially explicit
9 distributed units and physically-based simulation algorithms. To fill this gap, this
10 paper proposed an innovative spatially hybrid hydrological modeling approach where
11 compatible spatial units and simulation algorithms can be combined to construct
12 different model structures for each subbasin within the watershed such as lumped
13 conceptual and distributed physically-based model structures. Model structures in
14 each subbasin can be integrated to perform the entire watershed simulation. The
15 proposed approach is implemented by extending the Spatially Explicit Integrated
16 Modeling System (SEIMS) and evaluated by a case study in a medium-sized
17 headwater of the Heihe River Basin, China. Two distinct model structures were
18 designed to construct the spatially hybrid hydrological model of the study area: a
19 lumped conceptual model structure for gently sloping subbasins and a distributed
20 physically-based model structure for mountainous subbasins. The two model
21 structures were also separately used to build two spatially consistent hydrological
22 models for comparison. Experiment results showed that the spatially hybrid model
23 structure could harness the strengths of both model structures on the simulation
24 performance. This approach holds promise for improving the rationality of
25 hydrological modeling in watersheds with significant spatial heterogeneity.

1 **Highlight**

- 2 • We proposed an approach for spatially hybrid hydrological model structure.
- 3 • Model structures can be flexibly constructed in subbasins and then integrated.
- 4 • We implemented the approach by extending SEIMS as an example.
- 5 • The spatially hybrid model structure outperformed two spatially consistent ones.

19 **1 Introduction**

20 Hydrological models provide an effective way to simulate and understand
21 complex hydrological processes in real-world watersheds. Various hydrological
22 models have been developed, with each suitable for different application contexts
23 according to watershed characteristics, available data, and other factors (Beven, 2000;
24 Gharari et al., 2021; Wagener et al., 2001). Hydrological models can be classified
25 from two perspectives that define the model structure: the spatial discretization
26 scheme (i.e., lumped and semi- or fully-distributed models) and the representation of
27 hydrological processes (i.e., conceptual and physically-based models) (Chow, 1988;
28 Milad et al., 2012; Dingman, 2015; Hrachowitz and Clark, 2017). The determination
29 of an application-specific model structure, which involves deciding the suitable
30 spatial discretization scheme and hydrological process representations, stands as the
31 primary and critical step in hydrological modeling (Butts et al., 2004; David et al.,
32 2022; Pilz et al., 2020; van Esse et al., 2013).

33 The spatial discretization scheme refers to the organization of one or several
34 types of spatial units within the watershed for hydrological modeling deriving from
35 spatial data including terrain, land use, soil, and others (Dehotin and Braud, 2008).
36 Spatial units for lumped hydrological models typically ignore or broadly describe the
37 spatial heterogeneity of watershed characteristics using dominant geographic attribute
38 space, such as the dominant hydrological response unit (HRU) (DominantHRU for
39 short, i.e., HRUs defined by dominant soil types, land use types, and slope classes
40 within each subbasin; Arnold et al., 2010). In this regard, the entire watershed or
41 subbasins could also be defined as DominantHRU for lumped hydrological models. In
42 contrast, spatial units of semi- or fully-distributed hydrological models can fully cover
43 the entire geographic spatial space of the watershed with different detailed levels,
44 such as HRU defined by unique combinations of selected geographic attributes
45 (FullHRU for short and one HRU may consist of several spatially separated areas;
46 Dile et al., 2016), spatially explicit HRUs (ExplicitHRU for short and each HRU is a

47 spatially independent and continuously area; Teshager et al., 2016), fields or patches
48 (Tague and Band, 2004), and grid cells (Wigmosta et al., 1994). It is noted that the
49 HRU is actually a broad term that can be specifically defined as lumped units
50 (DominantHRUs), semi-distributed units (FullHRUs), and fully-distributed units
51 (ExplicitHRUs). From this perspective, the widely used SWAT (Soil and Water
52 Assessment Tool; Arnold et al., 1998) model can be used as a lumped or semi-
53 distributed hydrological model according to different HRU definitions. However, it
54 cannot be a fully-distributed model since it does not consider lateral flow processes
55 between HRUs.

56 The representation of hydrological processes denotes the simulation algorithms
57 abstracting the real-world hydrological processes (e.g., evaporation, infiltration, and
58 surface and subsurface flow routing) using different mathematical equations,
59 numerical solutions, temporal scales, and spatial units. A conceptual representation
60 would simplify the hydrological processes within a watershed as three basic
61 processes, which are water storage, water loss, and flow routing processes, based on
62 conceptual reservoirs (also called buckets), such as the linear method or two-reservoir
63 method for runoff prediction. The parameters of such conceptual representations often
64 lack explicit physical meanings (Fenicia et al., 2011; Knoben et al., 2019). In contrast,
65 a physically-based representation uses known scientific principles to model water
66 movements in vertical or lateral directions by the partial differential equation
67 representing the mass, momentum, and energy balance that solved by finite difference
68 approximations or empirical equations (Abbott et al., 1986), such as the Richards'
69 equation or kinematic wave method for infiltration (Graham and Butts, 2005). The
70 rigorous physically-based models with partial differential equations for water
71 movement such as the MIKE SHE (Systeme Hydrologique Europeen; Graham and
72 Butts, 2005) often faces challenges of data requirements and computational
73 complexity. Most physically-based models today, such as the SWAT (Arnold et al.,
74 1998), TOPMODEL (TOPography based hydrological MODEL; Beven and Kirkby,

75 1979), and WetSpa (Water and Energy Transfer between Soil, Plants and Atmosphere;
76 Wang et al., 1996), also use empirical formulations based on simplified
77 representations of physical principles and adhere to mass and energy balance.
78 Parameters (especially terrain and soil attributes) used in such physically-based
79 models primarily have clear physical meanings and may be derived from field
80 measurements. Therefore, we regarded “physically-based” as a broad term in this
81 study.

82 Existing methods of determining model structures can be categorized as
83 spatially consistent method and spatially varying method. The spatially consistent
84 model structure means the types of spatial units (also referred to as simulation units)
85 and simulation algorithms of hydrological processes are the same for the whole
86 watershed. Most research on directly choosing a single existing hydrological model
87 with a fixed or near-fixed model structure fall into this category. Such model
88 structures are often developed with specific assumptions and suitable for constrained
89 application contexts. They vary from lumped to distributed spatial discretization
90 schemes, and from conceptual to physically-based representation of hydrological
91 process, such as HBV (Hydrologiska Byråns Vattenbalansavdelning; Lindström et al.,
92 1997), GR4J (modèle du Génie Rural à 4 paramètres Journalier; Perrin et al., 2003),
93 TOPMODEL, SWAT, RHESSys (Regional Hydro-Ecologic Simulation System;
94 Tague and Band, 2004), and DHSVM (Distributed Hydrology Soil Vegetation Model;
95 Wigmosta et al., 1994). Adopting a spatially consistent model structure assumes that
96 the spatial variability of watershed characteristics can be sufficiently represented by
97 input data and model parameters required by the determined model structure.
98 Nevertheless, the fixed or near-fixed model structure lacks flexibility in changing
99 simulation units or customizing alternative simulation algorithms of considered
100 hydrological processes. Thus, it cannot accommodate diverse and complex
101 application contexts (Ley et al., 2016; Savenije, 2009).

102 To accommodate complex application contexts with flexibility in determining

103 the model structure, researchers have increasingly turned to modular hydrological
104 modeling frameworks, such as SUMMA (Structure for Unifying Multiple Modeling
105 Alternatives; Clark et al., 2015), ECHSE (ECo-Hydrological Simulation
106 Environment; Kneis, 2015), MARRMoT (Modular Assessment of Rainfall–Runoff
107 Models Toolbox; Knoben et al., 2019), SEIMS (Spatially Explicit Integrated
108 Modeling System; Liu et al., 2016; Zhu et al., 2019), and Raven (Craig, 2020).
109 These frameworks focus on different perspectives of flexibility in customizing
110 application-specific model structures, such as extensible types of simulation units and
111 associated simulation algorithms (Kneis, 2015), alternatives of simulation algorithms
112 of each specific hydrological process (Zhu et al., 2019), and alternatives of simulation
113 algorithms of each conceptual part of hydrological processes (Craig, 2020). These
114 frameworks provide supports for the concept of “multiple working hypotheses” (Clark
115 et al., 2011), facilitating studies including the suitability and uncertainty of different
116 model structures (David et al., 2022; Kiraz et al., 2013; Knoben et al., 2020). The
117 focus of these frameworks is primarily customizing and applying a consistent model
118 structure across the whole watershed. However, the spatially consistent model
119 structure may be unreasonable and inaccurate to capture the characteristics of the
120 dominant hydrological processes when the application context exhibits significant
121 spatial heterogeneity (Gao et al., 2018; Sui and Turner, 2021).

122 Recognizing the limitation of utilizing spatially consistent model structures,
123 researchers have explored constructing spatially varying model structures within the
124 watershed in two ways. The first is externally integrating multiple individual models
125 manually (Liu et al., 2020; Li et al., 2021; Wang et al., 2021) or through a unified
126 model interoperable interface such as ESMF (Earth System Modeling Framework;
127 Hill et al., 2004) and OpenMI (Open Modeling Interface; Harpham et al., 2019).
128 However, the flexibility of such model-level integration may be limited because each
129 integrated hydrological model has fixed or near-fixed model structures. Besides,
130 different interoperable interfaces often need to be implemented for each hydrological

131 model to be integrated, making it challenging for hydrologists without programming
132 skills.

133 To augment the flexibility of spatially varying model structures, the other idea
134 was to internally construct and integrate multiple model structures within one modular
135 hydrological modeling framework such as FLEX-Topo (Topography-driven Flux
136 Exchange hydrological model; Gao et al., 2014; Savenije, 2010), SUPERFLEX
137 (SUPER Flux Exchange hydrological model; Fenicia et al., 2011, 2016), and airGR
138 (suite of GR Hydrological Models; Coron et al., 2017; Thébault et al., 2023). In this
139 way, model structures can be distributed to different modeling areas in the watershed
140 with only minor changes in configuration or source code. Nevertheless, modeling
141 frameworks employed in these studies are inherently limited to conceptual
142 representations of hydrological processes, with the simulation units being only
143 subbasins, DominantHRUs, or FullHRUs. Therefore, these state-of-the-art flexible
144 frameworks only support lumped or semi-distributed conceptual model structures and
145 thus face a substantial challenge when attempting to integrate physically-based
146 simulation algorithms and spatially explicit distributed simulation units, such as grid
147 cells and patches, where the spatial relationship between units are kept to enable the
148 execution of physically-based hydrological process representation.

149 To address the inability of existing hydrological modeling frameworks to support
150 distributed physically-based model structures in a spatially varying manner, this paper
151 proposed a novel spatially hybrid modeling approach. In this new approach,
152 compatible types of simulation units and simulation algorithms can be combined to
153 construct different model structures for each subbasin within the watershed, such as
154 lumped conceptual and distributed physically-based model structures. Model
155 structures in each subbasin can then be integrated to perform the entire-watershed
156 simulation. The remainder of this paper was organized as follows: Section 2 outlined
157 the design of the proposed approach and its implementation based on the SEIMS
158 framework. Section 3 presented an experimental case study, followed by results and

159 discussion in Section 4, and conclusion in Section 5.

160 **2 Method design and implementation**

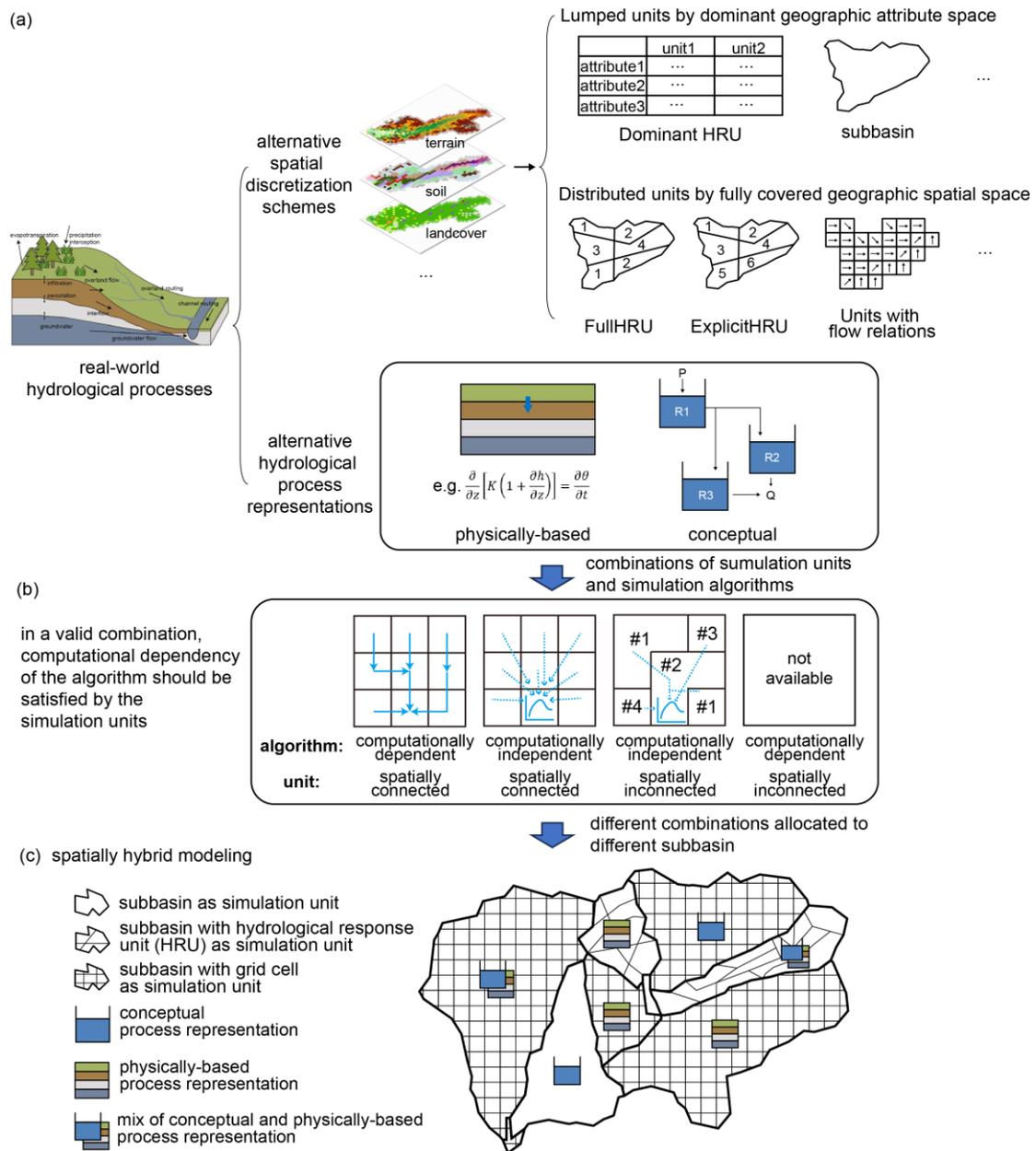
161 **2.1 Basic idea**

162 The basic idea of designing a spatially hybrid hydrological modeling approach is
163 constructing and executing different model structures on relatively independent areas
164 within the watershed in a divide-and-conquer way. Subbasins are relatively closed
165 catchment areas connected by the watershed drainage network and thus can be
166 regarded as the modeling unit to be configured with an individual model structure.
167 Therefore, the core design of the proposed approach involves two aspects: 1) the
168 flexibility for model structure construction, which enables the combination of
169 different spatial discretization schemes and hydrological process representations, and
170 2) the ability to distribute individual model structures to subbasins of the watershed
171 and integrate them to be executed under the same time loop.

172 The basic idea of designing the flexibly constructed model structure is to
173 decouple the hydrological process representations (specifically, the simulation
174 algorithms) from being bound to specific spatial discretization scheme (specifically,
175 the simulation units). For instance, the physically-based algorithms are typically
176 applied to fine-scale simulation units such as grid cells for both are assumed to
177 describe the hydrological process in detail. Similarly, conceptual algorithms are
178 typically applied to lumped or coarse-scale units such as DominantHRUs, hillslopes,
179 or subbasins. Such a binding is not necessary, and conversely, the combination of
180 model structure components have been acknowledged as effective, as demonstrated
181 by various studies (Gunduz and Aral, 2005; Liu et al., 2020; Sidle, 2021) and popular
182 models such as the SWAT. The simulation algorithms can be categorized as
183 computationally dependent and independent according to their computational
184 characteristics between simulation units. The simulation of a computationally
185 dependent algorithm on one simulation unit often relies on the water flows or state

186 variables of its adjacent or upstream units, while the simulation of computationally
187 independent algorithm on each simulation unit is independent with others (Figure 1b).
188 Similarly, existing simulation units can be categorized as spatially connected and
189 disconnected depending on whether spatially explicit relationships exist between units
190 (Figure 1a).

191 Therefore, three combinations of compatible simulation algorithms and
192 simulation units can be summarized: computationally dependent algorithm with
193 spatially connected units, computationally independent algorithm with spatially
194 connected units, and computationally independent algorithm with spatially
195 disconnected units (Figure 1b). For example, the conceptual reservoir-based surface
196 runoff generation algorithm with the physically-based kinematic wave-based surface
197 routing can be integrated to be applied to grid cells, where computational dependency
198 is satisfied by the flow directions between grid cells; the physically-based Green-
199 Ampt surface runoff generation algorithm and the conceptual unit hydrograph-based
200 surface routing algorithm can be integrated to applied on HRUs (Craig, 2020), where
201 no computational dependency is required. Note that the computational dependency is
202 only a basic constraint, the rationality of the model structure is to be determined by
203 the modeler's knowledge and requirements.



204

205 Figure 1. Design of the proposed spatially hybrid hydrological modeling approach. a)

206 decoupling the simulation of hydrological processes into spatial discretization

207 schemes and hydrological process representations; b) supporting combinations of

208 compatible simulation algorithms and simulation units; and c) enabling the spatial

209 varying allocation of different model structures at the subbasin level.

210

211 Based on the above idea of flexible model structure construction approach, in

212 this study, considering the spatial heterogeneity of watershed characteristics, each

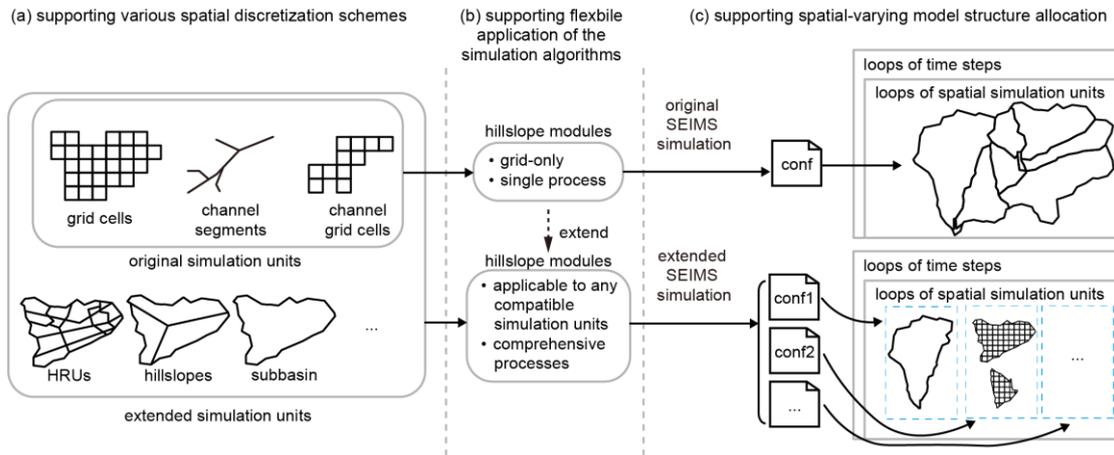
213 subbasin unit is regarded as the modeling unit to be configured with an individual

214 model structure (namely a subbasin model structure) to simulate hillslope processes,
215 and then integrated as a watershed model through channel routing processes (Figure
216 1c). A subbasin model structure can take all hydrologic process as a whole, or more
217 commonly, separately consider hillslope processes and channel routing processes.
218 Therefore, different subbasin model structures can be constructed, such as conceptual
219 algorithms with subbasin as simulation unit, physically-based algorithms with grid
220 cells as simulation units, and a mix of conceptual and physically-based algorithms
221 with HRU as simulation units for hillslope process and conceptual algorithms for
222 channel routing processes (Figure 1c). All subbasin model structures will collectively
223 constitute the watershed model structure by interconnecting through the drainage
224 network with upstream-downstream relationships. The execution of a downstream
225 subbasin relies on the output of its upstream subbasins. Thus, executions of subbasins
226 without upstream-downstream relationships could be parallelized; otherwise, they
227 must be executed sequentially.

228 **2.2 Design of the spatially hybrid hydrological modeling** 229 **approach based on the hydrological modeling framework**

230 **SEIMS**

231 Following the above basic idea, the proposed spatially hybrid hydrological
232 modeling approach was designed based on the hydrological modeling framework
233 SEIMS for its flexible modular design and two-level parallelization strategy (Liu et
234 al., 2016; Zhu et al., 2019). To support the flexible model structure construction of the
235 proposed approach, the spatial discretization schemes of SEIMS should be expanded
236 (Figure 2a; Section 2.2.2), the simulation modules of hillslope processes should be
237 upgraded to accommodate compatible simulation units (Figure 2b; Section 2.2.2), and
238 the execution of SEIMS-based model should be upgraded to allocate different model
239 structures to subbasins and executed in an integrated way (Figure 2c; Section 2.2.3).



240

241

242

243

244

245

246

247

Figure 2. Overall design of extending the SEIMS (Spatially Explicit Integrated Modeling System) to implement the proposed spatially hybrid hydrological modeling approach: a) supporting various spatial discretization schemes; b) supporting applying simulation algorithms of hillslope processes to compatible simulation units; and c) supporting spatial varying model structure allocation on the subbasin level to construct a watershed model.

248

2.2.1 Brief introduction to SEIMS

249

250

251

252

253

254

255

As a hydrological modeling framework, SEIMS provides users with the complete toolchain to preprocess modeling data, construct and execute the SEIMS-based model, analyze the parameter sensitivity, and so on (Liu et al., 2021; Zhu et al., 2019). More importantly, SEIMS supports developers in adding simulation modules of hydrologic processes following the developing principles. The principles are briefly introduced in three aspects: the spatial discretization scheme, the modular modeling design, and the parallelization strategy.

256

(1) Spatial discretization scheme

257

258

259

260

SEIMS adopts a two-level spatial discretization scheme, i.e., the “subbasin-basic simulation unit.” In the stage of preprocessing modeling data, SEIMS delineates the watershed into subbasins, each with one channel. Within each subbasin, the basic simulation units (i.e., grid cells in the original implementation) are delineated and

261 organized as a one-dimension array with their actual positions recorded. Most
262 simulation modules of hillslope processes use grid cells as simulation units, while a
263 few modules use subbasin units such the linear reservoir algorithm for the
264 groundwater process. Modules of channel routing processes use channel lines or grid
265 cells as simulation units. All parameters required by the modules on these simulation
266 units are prepared using the preprocessing scripts of SEIMS. The grid cells are
267 layered within each subbasin based on the flow direction of each cell (i.e., the single
268 flow direction in the original implementation; Liu et al., 2014). The layers are used in
269 modules of hillslope surface or subsurface flow routing processes to determine the
270 simulation sequences of grid cells. The same idea is used for layering subbasins (i.e.,
271 channels).

272 **(2) Modular modeling design**

273 A SEIMS-based hydrological model is constructed by the SEIMS main program
274 and several user-configured modules of hydrologic processes in a text-based format
275 rather than hard-coded. The main program controls the time loop of the simulation
276 and repeatedly invokes the modules in sequence. Each module is responsible for the
277 simulation of one or several hydrologic processes on corresponding simulation units
278 in different orders. For example, the potential evapotranspiration module executes on
279 every grid cell in an arbitrary order, the subsurface flow (also referred to as interflow)
280 routing module executes on every grid cell by layering orders (Liu et al., 2014), and
281 the channel routing module executes on each channel by upstream-downstream
282 orders.

283 In SEIMS, each module should handle its required simulation conditions. Users
284 should decide the feasibility of the module combinations to meet the requirements of
285 each module, including the boundary conditions when calculating vertical and lateral
286 water movements. For example, to implement a physically-based interflow module
287 based on the Darcy's Law, the boundary condition could be set either with the
288 hydraulic gradient equal to the slope at each cell, or with the output of other

289 cooperating modules.

290 Each module inherits from a standard module interface, including the definitions
291 of input data, parameters, and output data. The input data and parameters of one
292 module can be either read from the database created during the preprocessing stage or
293 referred from other modules during the runtime. Each module itself initializes its
294 output data. In this way, during the execution of the SEIMS-based model, the same
295 variable is shared among modules.

296 **(3) Two-level parallelization strategy**

297 Since SEIMS treated subbasins as relatively independent modeling units for
298 hydrological modeling, the first level of the two-level parallelization strategy
299 dispatches the simulation of subbasins to different computing processes (or nodes)
300 through the Message Passing Interface (MPI) (Liu et al., 2016). The second level is
301 achieved within each subbasin by dispatching the simulations of grid cells without
302 mutual dependencies to computing threads via Open Multi-Processing (OpenMP)
303 (Liu et al., 2014). Based on the design of the modular structure, variables required to
304 be communicated among subbasins only need to be defined in the metadata of each
305 module and will be handled by the SEIMS main program (Zhu et al., 2019). That
306 means the complicated MPI programming details are hidden from module developers.
307 Besides, the implementation of OpenMP-based parallelization only needs one line of
308 preprocessor directive code before the loop of computationally independent code.
309 Therefore, module developers can easily develop modules and build high-
310 performance hydrological models in a nearly serial programming way (Zhu et al.,
311 2019). The two-level parallelization strategy provides the potential to implement the
312 proposed approach in this study that applies different model structures to distinct
313 subbasins.

314 **2.2.2 Constructing model structures with diverse simulation units and algorithms**

315 To enable the flexible construction of model structures based on SEIMS, the

316 diversity of both spatial discretization schemes and hydrological process
317 representations should be guaranteed. For spatial discretization, the idea of
318 constructing hydrological response units (HRUs) could be added as simulation units
319 of hillslope processes in SEIMS, including DominantHRU, FullHRU, and
320 ExplicitHRU (see the Introduction section for details) (Figure 2a). In terms of
321 parameters, the physically-based parameters can be derived from actual properties of
322 the soil, landuse/landcover, and other spatial data, while conceptual parameters are
323 directly specified by lookup tables of soil and land-use. For instance, the reservoir
324 capacity of an HRU can be represented by the soil depth associated with that specific
325 HRU area, while some experimental coefficients could be specified in the lookup
326 tables.

327 Under such an HRU discretization scheme, the hillslope unit and subbasin unit
328 can be regarded as specific DominantHRU to be used as simulation units. Therefore,
329 the extended SEIMS will offer the flexibility of utilizing subbasins, hillslopes, HRUs,
330 and grid cells as simulation units of hillslope processes for various modeling needs.

331 For hydrological process representations, SEIMS was initially designed to
332 primarily integrate simulation algorithms of one hydrological process into each single
333 module, while it lacked a conceptual model capable of considering multiple
334 hydrological processes comprehensively (Figure 2b). To bridge this gap, the lumped
335 conceptual models could be integrated as two separate process-based hillslope
336 modules, i.e., the surface runoff module that generate runoff at each simulation unit
337 and the surface routing module that directly convey water to the subbasin outlet. For
338 example, the representative widely used conceptual models, GR4J (Perrin et al.,
339 2003), is suitable to be incorporated as SEIMS modules. Such modules could be
340 applied to any simulation unit types stated above.

341 To ensure the combination compatibility of simulation algorithms with
342 simulation units applied to subbasins, the SEIMS module interface should be
343 extended to mark its computational dependency requirement as computationally

344 dependent or independent. A computationally dependent simulation module is
345 exclusively applicable to grid cells, and a computationally independent module is
346 applicable to any types of simulation unit. A model structure containing any
347 computationally dependent module can only be applied to the subbasin using grid
348 cells as the basic simulation units.

349 **2.2.3 Allocating individual model structures to subbasins and integrating as a** 350 **hybrid watershed model structure**

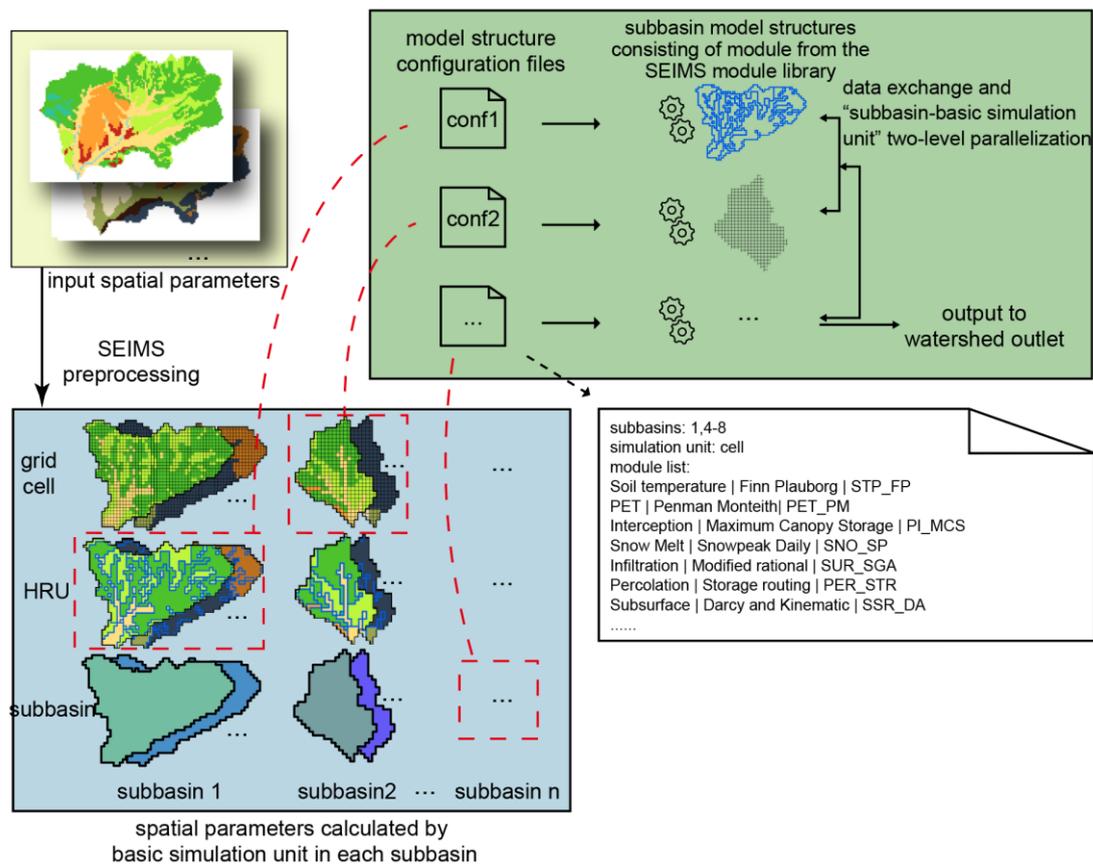
351 An essential part of the spatially hybrid hydrological modeling approach is to
352 enable the separate allocation and execution of individual model structures to different
353 subbasins (referred to as the subbasin models) within the watershed. All subbasin
354 models are subsequently integrated as a watershed model. The requirements of
355 flexible model configuration and subbasin-separate simulation are compatible with
356 the text-based module configuration method and two-level parallelization strategy of
357 SEIMS (see section 2.2.1) but still need improvement. The module configuration file
358 of SEIMS should be extended to designate a model structure to the specific subbasin,
359 and the adopted type of basic simulation units (Figure 3). In this way, the SEIMS
360 main program can read the configuration file for each subbasin dynamically, load
361 simulation modules, and retrieve the modeling parameters according to the specified
362 basic simulation units.

363 The two-level parallelization strategy should be improved in two aspects. The
364 first is task scheduling of the subbasin-level parallelization for load balancing handled
365 by the SEIMS main program. The domain decomposition of subbasins should be
366 determined by upstream-downstream relationships between subbasins and the number
367 of basic simulation units of each subbasin in the runtime, rather than using the
368 numbers of grid cells of each subbasin in the preprocessing stage in the original
369 implementation of SEIMS (Liu et al., 2016). The second aspect concerns the basic
370 simulation unit-level parallelization handled in computationally independent modules,

371 which could be easily extended to support newly added simulation units such as
372 HRUs since the OpenMP for-loop code structure does not need to be changed.

373 **2.3 Implementation**

374 Based on the above method design, the implementation of the proposed approach
375 with SEIMS involves modifications to the data preprocessing tools, SEIMS main
376 programs, and SEIMS module library. The data preprocessing tools mainly include a
377 collection of Python scripts and C++ programs to delineate spatial units at different
378 scales (e.g., subbasins and hillslopes), extract spatial parameters of spatial units, and
379 create watershed modeling database (Zhu et al., 2019). In this study, a configurable
380 tool was implemented to support the delineation and parameterization of HRU-based
381 spatial units. This tool allows users to specify environmental variables (e.g., the
382 default land-use and soil types) to delineate HRU-based units. This tool also allows
383 specifying base spatial units and delineation strategy to generate different types of
384 HRU such as taking subbasins or hillslopes as base units to generate DominantHRU
385 or FullHRU. Therefore, the data preprocessing tools could prepare various types of
386 spatial unit and associated spatial parameters for each subbasin according to the user-
387 defined configuration file to meet various hydrological modeling scenarios. The
388 procedures of spatially hybrid modeling with SEIMS is depicted in Figure 3.



389

390 Figure 3. Procedures of spatially hybrid modeling with the SEIMS (Spatially Explicit
 391 Integrated Modeling System)

392

393 The SEIMS main program, written in C++, was extended to read the
 394 configuration file for each subbasin, load simulation modules and retrieve the
 395 modeling parameters according to the specified simulation units. Simulation modules
 396 of the SEIMS module library were updated to declare the type of computational
 397 dependency. New conceptual simulation modules were added such as GR4J for
 398 simulating hillslope runoff generation. With the above briefly introduced
 399 implementation of the proposed approach, SEIMS can now support users in
 400 constructing spatially hybrid model structures for considering spatial heterogeneity of
 401 the watershed, where each subbasin may be simulated with different combinations of
 402 conceptual or physically-based simulation modules and spatial units. Consequently,
 403 42 modules are available after the implementation of this study (detailed in Table S1

404 of the supplementary material), supporting simulation of processes including
405 snowfall, atmospheric deposition, snow balance, interception, soil temperature,
406 glacier, surface runoff, evapotranspiration, infiltration, percolation, interflow,
407 groundwater, channel routing, plant growth, and soil erosion.

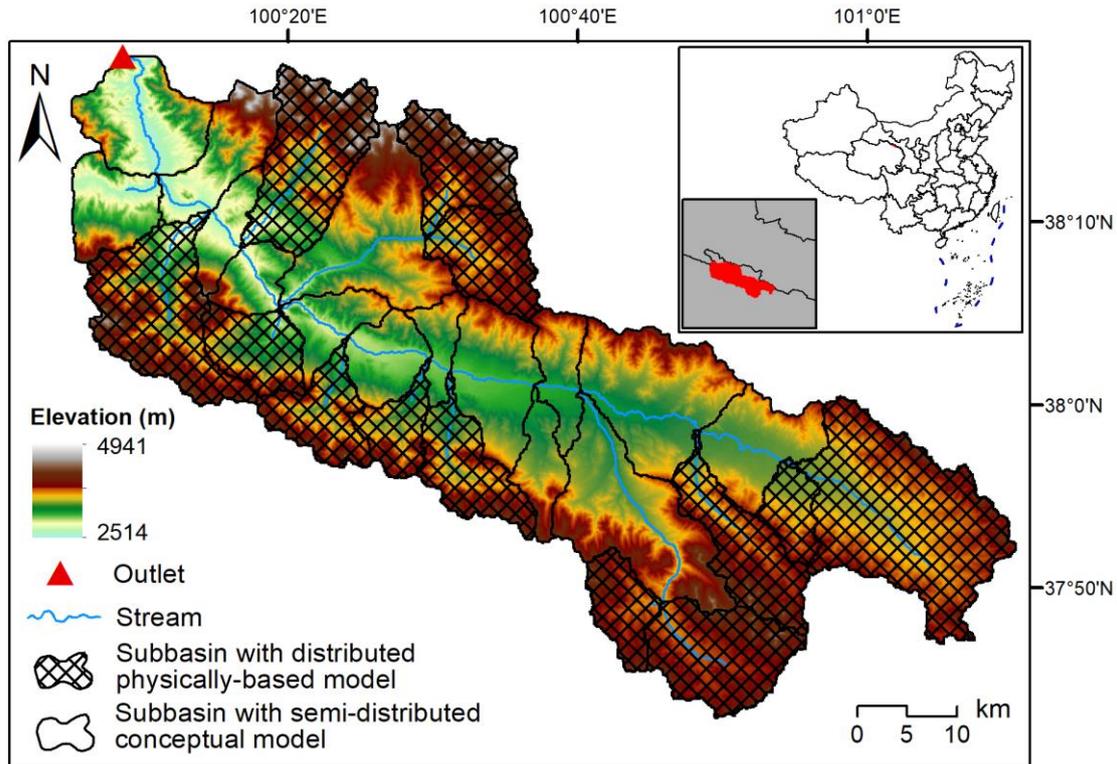
408 The SEIMS is open-source on GitHub (<https://github.com/lreis2415/SEIMS>) and
409 is under continuous development, such as adding conceptual and physical-based
410 simulation modules for various geographic environments.

411 **3 Case study**

412 **3.1 Study area and data**

413 In this study, we selected the Babao River watershed at Qilian, Qinghai Province,
414 China (Figure 4) as the case study area to verify the proposed approach through daily-
415 scale runoff simulation. The Babao River watershed is one of the headwaters of the
416 Heihe River Basin. It is located in a high-altitude, cold and mountainous region with
417 an area of approximately 2,511 km². The average elevation is 3,565 m, and the region
418 features glacier, snow cover and frozen soil.

419 The data for hydrological modeling in this study are all publicly available. The
420 MERIT DEM with the resolution of 90 m (Yamazaki et al., 2017) was selected for
421 watershed delineation and calculation of terrain attributes. Meteorological data was
422 obtained from the China Meteorological Assimilation Driving Datasets for the SWAT
423 model (CMADS) version 1.2, provided in the form of approximately 0.125°
424 resolution gridded station data (Meng et al., 2019). Land cover data was derived from
425 GLOBELAND30 with a resolution of 30 m (<http://www.globallandcover.com>). Soil
426 attribute data was sourced from the Harmonized World Soil Database (HWSD)
427 Version 2.0 (Nachtergaele, 2023).



428

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

431 3.2 Experimental design

432 To verify the feasibility and effectiveness of the proposed spatially hybrid
 433 hydrological modeling approach and its implementation based on SEIMS, we
 434 constructed one distributed physically-based model structure (section 3.2.1) and one
 435 lumped conceptual model structure (section 3.2.2) to build spatially consistent and
 436 spatially hybrid watershed models (section 3.2.3) for comparison after parameter
 437 calibration (section 3.2.4). It is worth noting that the watershed models constructed in
 438 this comparative experiment are not intended to model the hydrological processes in
 439 the study precisely nor to enumerate the possible model structures applicable as
 440 illustrated in Figure 1.

441 3.2.1 The distributed physically-based model structure

442 The distributed physically-based model structure uses grid cells as the basic

443 simulation units and encompasses simulation algorithms based on physical laws. For
444 example, the algorithms of percolation and interflow are based on Darcy's law. The
445 percolation is calculated as the product of hydraulic conductivity and the gradient of
446 the hydraulic potential, and the interflow is calculated from the kinematic
447 approximation of Darcy's Law and, with the hydraulic gradient equal to the slope at
448 each grid cell. These two methods are the same as those used in WetSpa [please refer
449 to Liu and Smedt (2004) and Safari et al. (2012) for more details]. Table 1 listed the
450 considered hydrological processes and their simulation algorithm associated with key
451 parameters for model calibration. The physically-based simulation algorithms
452 primarily utilize spatial parameters with clear physical meaning, such as soil hydraulic
453 conductivity and leaf area index. The geomorphology-based unit hydrograph is
454 derived from terrain data. Besides, this model structure also includes conceptual
455 simulation algorithms such as the surface routing module based on unit hydrograph
456 (Table 1). A model structure combined physically-based and conceptual hydrological
457 process representations is often regarded as a physically-based model, as stated by
458 existing hydrological model with a fixed or near-fixed model structure such as SWAT,
459 WetSpa, and RHESSys.
460

461 Table 1. Simulation algorithms adopted for the distributed physically-based model
 462 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

463 **3.2.2 The lumped conceptual model structure**

464 The lumped conceptual model structure has DominantHRUs as the simulation
 465 units and has GR4J as the main simulation algorithm, which offers simplicity and can
 466 generalize multiple hydrological processes (Table 2). The DominantHRUs are derived
 467 from the overlay of the categorized land cover and soil maps, with the HRU with
 468 occupying the largest area being the DominantHRU. The Hargreaves method is used
 469 to provide the potential evapotranspiration for the GR4J. And the GR4J receives the

470 potential evapotranspiration and precipitation as input to generate the hillslope runoff
 471 at subbasin outlets, which are then aggregated by the Muskingum method.

472

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

Hydrological process	Simulation algorithm	Parameters involved in calibration
Potential evapotranspiration	Hargreaves	-
Conceptual and comprehensive hillslope runoff	GR4J	SOILTHICK (X1), X2, X3, X4
Channel routing	Muskingum	MSK_X, MSK_K

475

476 **3.2.3 One spatially hybrid watershed model structure and two spatially**
 477 **consistent model structures for comparison**

478 Ideally, a physically-based model could better simulate hydrological processes
 479 with explicit physical meaning. While limited to the insufficiency of the cognization
 480 of hydrological processes and difficulty of implementing complicated simulation
 481 algorithms, a conceptual model could be a valuable complementary, which is often
 482 come up with mathematical fitting under some degree of generalization. Therefore, a
 483 spatially hybrid model structure (referred to as the HybM) combining both physically-
 484 based and conceptual models were constructed in this case study. Most parameters of
 485 the distributed physically-based model structure were derived from actual terrain,
 486 landuse, and soil data, which would vary more significantly in spatial. Thus, the
 487 distributed physically-based model structure is considered more suitable for
 488 mountainous areas than the lumped conceptual model structure. Under this
 489 assumption, the HybM in the Babao River watershed was constructed, as depicted in
 490 Figure 4. In the HybM, 16 subbasins were assigned with the lumped conceptual
 491 model structure, and the other 13 subbasins with the distributed physically-based

492 model structure.

493 The two model structures were then used to build spatially consistent model
494 structures to be compared with the HybM. The model utilizing only the distributed
495 physically-based model structure was referred to as PhyM, and the model utilizing
496 only the lumped conceptual model structure was referred to as ConM.

497 **3.2.4 Comparative experiments of parameter calibration**

498 The parameters involved in model calibration that listed in Table 1 and Table 2
499 were selected by parameters sensitivity analysis. In the calibration of the HybM,
500 parameters in subbasins with the same model structure (the distributed physically-
501 based or the lumped conceptual) were changed together (i.e., add or multiply). That is,
502 parameters in the mountainous subbasins share the same change, and those in gently
503 sloping subbasins share the same change (Figure 4). For comparison to the HybM,
504 two calibration strategies are used for both the PhyM and ConM: 1) universal
505 calibration, where parameters in all subbasins were changed together. 2) regional
506 calibration, where parameters in the mountainous subbasins were changed together,
507 and those in gently sloping subbasins were changed together, i.e., the same calibration
508 strategy as the HybM.

509 The three models were executed in a six-year simulation, spanning from January
510 1, 2013 to December 30, 2018. The first year (2013) served as a warm-up period,
511 followed by three years (2014–2016) designated for calibration, and two years (2017–
512 2018) for validation. Four model performance indices, including the Nash-Sutcliffe
513 coefficient (NSE; Eq. 1), root mean square error-standard deviation ratio (RSR; Eq.
514 2), percent bias (PBIAS; Eq. 3), and R^2 (Eq. 4), were selected to evaluate the model
515 performance.

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

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

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

$$R^2 = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (4)$$

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

518 The parameter calibration was conducted using the NSGA-II (non-dominated
519 sorting genetic algorithm-II; Deb et al., 2002) integrated in the SEIMS framework
520 (Zhu et al., 2019), with each experiment having an optimization of 360 population and
521 25 generation. The calibration experiments were conducted on a personal computer
522 with a 2.70 GHz Intel Xeon Gold 6150 dual CPU (36 cores). Under the NSGA-II
523 algorithm, the Pareto front of each generation would be kept and compared with the
524 next generation. The objective function for parameter calibration is maximum NSE,
525 minimum absolute values of RSR and PBIAS, with each component having the same
526 weight. Comparative analysis was carried out to assess the simulation performance of
527 the three models based on selected optimal parameters from the Pareto front.

528

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

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

531

532 **4 Results and discussion**

533 The successful execution of the constructed spatially hybrid watershed model
534 structure proved the feasibility of integrating lumped conceptual model structure and
535 distributed physically-based model structure in a spatially varying manner. According
536 to the comparative experiments in this case study, the effectiveness of the proposed
537 approach and its implementation based on SEIMS can be discussed from three
538 aspects: 1) the performance of different model structures after automatic parameter
539 calibration under different calibration strategies; 2) the rationality of the spatially
540 hybrid model structure; and 3) the advantage of SEIMS in implementing the proposed
541 approach.

542 **4.1 The model performance of different model structures under** 543 **different calibration strategies**

544 The model performances of the selected optimal individuals of all calibration
545 strategies of the three model structures are listed in Table 4. According to the
546 performance metrics of both calibration and validation periods, the spatially
547 consistent distributed physically-based model (i.e., PhyM) exhibits similar
548 performance under both the universal and the regional calibration strategies (i.e.,
549 PhyM1 versus PhyM2). In contrast, for the spatially consistent lumped conceptual
550 model, the ConM2 that utilized the regional calibration strategy outperformed the
551 ConM1. The difference may attribute to the characteristics of modeling parameters.
552 The parameters of the PhyM are derived from spatially distributed input data,
553 implying a low dependency on regional calibration. In contrast, the initial parameters
554 of the ConM were the same across the watershed since their lack of clear physical
555 meaning, resulting in the dependency on distributed calibration.

556

557 Table 4. Performance metrics of selected optimal calibrated individuals with different
 558 model structures under different calibration strategies (PhyM1 and PhyM2:
 559 physically-based model structure with universal and regional calibration; ConM1 and
 560 ConM2: conceptual model structure with universal and regional calibration,
 561 respectively; HybM: the spatially hybrid model structure with regional calibration)

Experimental case	Best Representative individual performance selected from the Pareto front (calibration / validation periods)			
	NSE	RSR	PBIAS (%)	R ²
PhyM1	0.58/0.60	0.64/0.63	10.63/2.58	0.66/0.67
PhyM2	0.59/0.60	0.64/0.63	17.34/13.90	0.67/0.66
ConM1	0.41/0.32	0.77/0.82	2.89/-42.87	0.42/0.61
ConM2	0.45/0.50	0.74/0.71	10.96/-21.03	0.48/0.60
HybM	0.72/0.60	0.53/0.63	10.02/-18.28	0.73/0.71

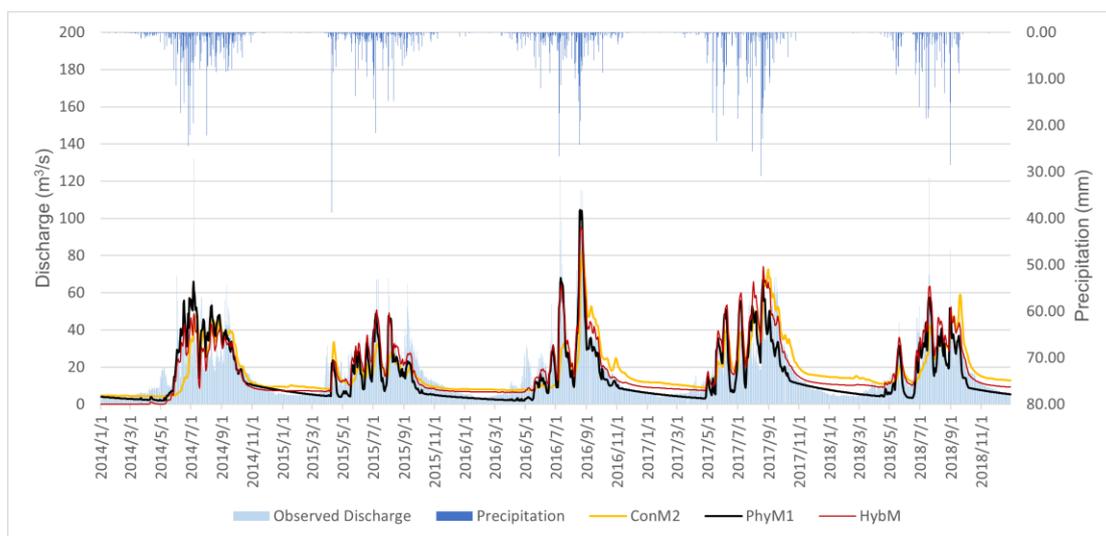
562

563 The HybM demonstrated an overall better performance over the PhyM and
 564 ConM. In terms of NSE, the HybM showed the highest performance in the calibration
 565 period (0.72), and one of the highest performances in the validation period (0.6). The
 566 HybM also had lower RSR (0.53/0.63) and higher R² (0.73/0.71) than the PhyMs and
 567 ConMs, while its PBIAS (10.02%/-18.28%) did not show advantage. This
 568 demonstrated that the proposed approach could achieve higher performance compared
 569 to the spatially consistent model structures.

570 4.2 The rationality of the spatially hybrid model structure

571 Figure 4 showed the hydrographs of simulation results of ConM2, PhyM1, and
 572 HybM, compared with the observed discharge (Figure 4). In dry seasons, typically
 573 from November to the following April when baseflow is the primary contributor to

574 discharge, the PhyM1 exhibits an accurate estimation of baseflow. In contrast, the
 575 ConM2 provides an overall overestimation of the baseflow simulation. In the wet
 576 seasons, the PhyM1 produces steeper rising and falling limbs in the hydrographs,
 577 corresponding to the intensive precipitation, as depicted in Figure 4. This behavior
 578 aligns with the physical basis of the study area. The ConM2, on the other hand, tend
 579 to either overestimate or underestimate the peak flows.
 580



581
 582 Figure 4. The simulation results of the individuals with best performance after
 583 calibration of the ConM2, PhyM1, and HybM. (PhyM1: physically-based model
 584 structure with universal calibration; ConM2: conceptual model structure with regional
 585 calibration; HybM: the spatially hybrid model structure with regional calibration)
 586

587 The spatially hybrid model HybM, as a result, combines the strengths of both the
 588 physically-based and the conceptual model structures. Especially, the HybM exhibited
 589 a more accurate simulation of peak flows and recession all through the simulation
 590 periods, better than both the PhyM1 and the ConM2. And for the rising limbs, the
 591 HybM reacted close to the PhyM1, which fitted better than the ConM2. While for the
 592 baseflow, the HybM showed a compromising between the PhyM1 and the ConM2,
 593 which turns out to be a n overestimation but better than the PhyM1.

594 **4.3 The advantage of SEIMS in implementing the proposed** 595 **approach**

596 The selection of SEIMS in implementing the proposed approach exhibited
597 effectiveness in computing efficiency and modeling flexibility. Due to the detailed
598 representation of various hydrological processes in the physically-based model, its
599 calibration took approximately 75 hours for 9000 model runs (360 population and 25
600 generations). Such a high calibration demand limits its availability to larger
601 watersheds. In contrast, the calibration time for the conceptual model was
602 significantly shorter, approximately 1 hour. Therefore, the spatially hybrid model,
603 which contains both models, exhibited an intermediate calibration time of around 30
604 hours, average of the two models. This efficiency enhancement suggests a broader
605 applicability to larger watersheds, where specific subbasins of interest can be treated
606 individually. This approach not only saves time but also enables the production of
607 more model runs for parameter calibration or uncertainty analysis.

608 The subbasin-independent design also gives rise to the modeling flexibility.
609 Model developers could implement their models in module-level to simulate a single
610 process or multiple processes, which can then be combined easily with other modules
611 into a customized spatially hybrid model structure. While it is worth noting that the
612 module developers should handle the input requirements and boundary conditions for
613 the single module, and the module users should examine the feasibility of the
614 combined model structure.

615 **5 Conclusion**

616 This paper introduced a novel spatially hybrid hydrological modeling approach,
617 offering a versatile solution to address the modeling challenges posed by the spatial
618 heterogeneity of complex watersheds. Unlike existing distributed hydrological models
619 that allow only spatial variability of parameters, the proposed approach takes a divide-

620 and-conquer idea to accommodate spatially varied and hybrid model structures. This
621 approach advocates the hybrid combination of different spatial discretization schemes
622 and hydrological process representations, enhancing structural flexibility on each
623 subbasin. The constraints of integrating subbasin model structures were also discussed
624 to ensure the successful construction of the final watershed model.

625 Based on the implementation on the hydrological modeling framework SEIMS,
626 the comparative experiments were conducted to validate the feasibility and
627 effectiveness of the proposed approach. The spatially hybrid model structure,
628 compared to the spatially consistent distributed physically-based and lumped
629 conceptual model structures, demonstrated not only the capability of modeling with
630 spatially varied and flexible structure, but also the ability to synergize the strengths of
631 the spatially consistent model structures, potentially leading to reduced uncertainty.

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

638 However, there are also limitations in the current implementation. Firstly, the
639 SEIMS module library should undergo further development to encompass a broader
640 range of simulation algorithms and hydrological processes, for example,
641 implementing tightly coupled simulation methods of several hillslope hydrological
642 processes using rigorously physically-based algorithms as one simulation module,
643 implementing simulation modules for glacier and frozen soil processes, etc. Secondly,
644 as the spatial constitution of the spatially hybrid model structure can be freely
645 configured and revised, determining the optimal spatial constitution for specific
646 application contexts remains a subject for future study.

647 **Acknowledgments**

648 This work was supported by National Key Research and Development Program
649 of China [2021YFB3900904], National Natural Science Foundation of China
650 [42101480], LREIS [YPI005, KPI003], and Shaanxi Normal University. The support
651 for A-Xing Zhu through the Vilas Associate Award, the Hammel Faculty Fellow
652 Award, and the Manasse Chair Professorship from the University of Wisconsin-
653 Madison is greatly appreciated. We want to acknowledge the discussion regarding the
654 physically-based hydrological modeling algorithms with Dr. Kaiwen Wang from the
655 IGSNRR, CAS.

656 **Declarations of interest:** none.

657

658

659 **Reference**

- 660 Abbott, M.B., Bathurst, J.C., Cunge, J.A., O’Connell, P.E., Rasmussen, J., 1986. An
661 introduction to the European Hydrological System — Systeme Hydrologique
662 Europeen, “SHE”, 2: Structure of a physically-based, distributed modelling
663 system. Journal of Hydrology 87, 61–77. [https://doi.org/10.1016/0022-](https://doi.org/10.1016/0022-1694(86)90115-0)
664 [1694\(86\)90115-0](https://doi.org/10.1016/0022-1694(86)90115-0)
- 665 Arnold, J. G., Allen, P. M., Volk, M., Williams, J. R., & Bosch, D. D., 2010.
666 Assessment of different representations of spatial variability on SWAT model
667 performance. Transactions of the ASABE 53, 1433–1443.
668 <https://doi.org/10.13031/2013.34913>
- 669 Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large Area
670 Hydrologic Modeling and Assessment Part I: Model Development. JAWRA
671 Journal of the American Water Resources Association 34, 73–89.
672 <https://doi.org/10.1111/j.1752-1688.1998.tb05961.x>
- 673 Beven, K.J., 2000. Uniqueness of place and process representations in hydrological
674 modelling. Hydrology and Earth System Sciences 4, 203–213.
675 <https://doi.org/10.5194/hess-4-203-2000>
- 676 Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model
677 of basin hydrology. Hydrological Sciences Bulletin 24, 43–69.
678 <https://doi.org/10.1080/02626667909491834>
- 679 Butts, M.B., Payne, J.T., Kristensen, M., Madsen, H., 2004. An evaluation of the

680 impact of model structure on hydrological modelling uncertainty for streamflow
681 simulation. *Journal of Hydrology* 298, 242–266.
682 <https://doi.org/10.1016/j.jhydrol.2004.03.042>

683 Chow, V.T., Maidment, D.R., Mays, L.W., 1988. *Applied hydrology*, 2nd ed.
684 McGraw-Hill, New York.

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

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

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

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

698 David, P.C., Chaffe, P.L.B., Chagas, V.B.P., Dal Molin, M., Oliveira, D.Y., Klein,
699 A.H.F., Fenicia, F., 2022. Correspondence Between Model Structures and
700 Hydrological Signatures: A Large-Sample Case Study Using 508 Brazilian

701 Catchments. Water Resources Research 58, e2021WR030619.
702 <https://doi.org/10.1029/2021WR030619>

703 Deb K, Pratap A, Agarwal S, et al., 2002. A fast and elitist multiobjective genetic
704 algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6: 182–
705 197.

706 Dehotin, J., Braud, I., 2008. Which spatial discretization for distributed hydrological
707 models? Proposition of a methodology and illustration for medium to large-scale
708 catchments. Hydrology and Earth System Sciences 12, 769–796.
709 <https://doi.org/10.5194/hess-12-769-2008>

710 Dile, Y.T., Daggupati, P., George, C., Srinivasan, R., Arnold, J., 2016. Introducing a
711 new open source GIS user interface for the SWAT model. Environmental
712 Modelling & Software 85, 129–138.
713 <https://doi.org/10.1016/j.envsoft.2016.08.004>

714 Dingman, S.L., 2015. Physical hydrology, 3rd ed. Waveland Press, Inc, Long Grove,
715 Illinois.Fenicia, F., Kavetski, D., Savenije, H.H.G., 2011. Elements of a flexible
716 approach for conceptual hydrological modeling: 1. Motivation and theoretical
717 development. Water Resources Research 47.
718 <https://doi.org/10.1029/2010WR010174>

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

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

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

730 Gharari, S., Gupta, H.V., Clark, M.P., Hrachowitz, M., Fenicia, F., Matgen, P.,
731 Savenije, H.H.G., 2021. Understanding the Information Content in the Hierarchy
732 of Model Development Decisions: Learning From Data. *Water Resources*
733 *Research* 57, e2020WR027948. <https://doi.org/10.1029/2020WR027948>

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

737 Harpham, Q.K., Hughes, A., Moore, R.V., 2019. Introductory overview: The OpenMI
738 2.0 standard for integrating numerical models. *Environmental Modelling &*
739 *Software* 122, 104549. <https://doi.org/10.1016/j.envsoft.2019.104549>

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

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

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

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

753 Knoben, W.J.M., Freer, J.E., Fowler, K.J.A., Peel, M.C., Woods, R.A., 2019. Modular
754 Assessment of Rainfall–Runoff Models Toolbox (MARRMoT) v1.2: an open-
755 source, extendable framework providing implementations of 46 conceptual
756 hydrologic models as continuous state-space formulations. *Geoscientific Model*
757 *Development* 12, 2463–2480. <https://doi.org/10.5194/gmd-12-2463-2019>

758 Knoben, W.J.M., Freer, J.E., Peel, M.C., Fowler, K.J.A., Woods, R.A., 2020. A Brief
759 Analysis of Conceptual Model Structure Uncertainty Using 36 Models and 559
760 Catchments. *Water Resources Research* 56, e2019WR025975.
761 <https://doi.org/10.1029/2019WR025975>

762 Ley, R., Hellebrand, H., Casper, M.C., Fenicia, F., 2016. Is Catchment Classification
763 Possible by Means of Multiple Model Structures? A Case Study Based on 99

764 Catchments in Germany. *Hydrology* 3, 22.
765 <https://doi.org/10.3390/hydrology3020022>

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

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

774 Liu, J., Liu, Z., Yin, Y., Croke, B. F. W., Chen, M., Qin, C.-Z., Tang, G., Zhu, A.-X.,
775 2021. A hybrid vector-raster approach to drainage network construction in
776 agricultural watersheds with rice terraces and ponds. *Journal of Hydrology* 601,
777 126585. <https://doi.org/10.1016/j.jhydrol.2021.126585>

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

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

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

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

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

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

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

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

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

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

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

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

818 Teshager, A. D., Gassman, P. W., Secchi, S., Schoof, J. T., & Misgna, G., 2016.
819 Modeling agricultural watersheds with the soil and water assessment tool
820 (SWAT): Calibration and validation with a novel procedure for spatially explicit
821 HRUs. Environmental Management 57, 894–911.
822 <https://doi.org/10.1007/s00267-015-0636-4>

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

826 van Esse, W.R., Perrin, C., Booij, M.J., Augustijn, D.C.M., Fenicia, F., Kavetski, D.,

827 Lobligeois, F., 2013. The influence of conceptual model structure on model
828 performance: a comparative study for 237 French catchments. Hydrology and
829 Earth System Sciences 17, 4227–4239. [https://doi.org/10.5194/hess-17-4227-](https://doi.org/10.5194/hess-17-4227-2013)
830 2013

831 Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V., Sorooshian, S.,
832 2001. A framework for development and application of hydrological models.
833 Hydrology and Earth System Sciences 5, 13–26. [https://doi.org/10.5194/hess-5-](https://doi.org/10.5194/hess-5-13-2001)
834 [13-2001](https://doi.org/10.5194/hess-5-13-2001)

835 Wang, Z.M., Batelaan, O., De Smedt, F., 1996. A distributed model for water and
836 energy transfer between soil, plants and atmosphere (WetSpa). Physics and
837 Chemistry of the Earth 21, 189–193. [https://doi.org/10.1016/S0079-](https://doi.org/10.1016/S0079-1946(97)85583-8)
838 [1946\(97\)85583-8](https://doi.org/10.1016/S0079-1946(97)85583-8)

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

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

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