



集成建模环境研究及其在黑河流域的初步应用

南卓铜*, 舒乐乐, 赵彦博, 李新, 丁永建

中国科学院寒区旱区环境与工程研究所, 兰州 730000

* E-mail: nztong@lzb.ac.cn

收稿日期: 2010-05-14; 接受日期: 2011-03-24

中国科学院重要方向项目(批准号: KZCX2-YW-Q10-1)、国家高技术研究发展计划(“863”计划)(批准号: 2008AA12Z205)和中国科学院西部行动计划(批准号: KZCX2-XB2-09)资助项目

摘要 环境和水问题是跨学科问题, 解决这些问题需要集成多学科模型. 建模环境提供了有效的模型集成技术手段. 现有建模环境由于技术设计和学科侧重点的原因, 无法满足当前的需求, 比如无法集成知识系统形成智能建模环境. 本文建立了用于水文和地表过程领域的建模环境, 采用高效灵活的模块控制和数据传递机制, 实现了灵活的模块重用性和扩展性. 本文介绍了本建模环境的体系结构、模块组件化表示、模块控制和数据传递、模块连接和重用等关键技术. 最后, 作为一个应用实例, 在建模环境的支持下使用 Noah 陆面过程模型的蒸散发模块替换了 TOPMODEL 的蒸发模块, 演示了基于建模环境的模型集成和模型比较是高效、可靠、有意义的.

关键词

建模环境
黑河流域
TOPMODEL
Noah 陆面过程模型

环境和水问题本质上是多学科交叉的复杂问题^[1, 2]. 尽管在过去的几十年, 国内外已经发展了许多水文、生态和陆面过程模型, 用于回答环境和水问题. 但是, 这些模型建立之初往往有其特定的学科背景, 针对研究区当时的数据条件和流域特点, 尝试回答当时具体的科学问题. 比如, 水文模型通常侧重模拟水循环系统中的一个或几个物理过程, 而其他过程被适当的简化. 当面临一个不同类型的流域或者更综合的流域问题时, 现有模型往往不能全面和准确地进行模拟^[3]. 针对这种情况, 模型的发展有 2 种趋势, 一种是走“大模型”路线, 即在现有模型基础上, 将更多的物理过程模块添加进来, 以增加模型的模拟能力; 另一种是走“集成”路线, 即认为不同的模型有不同的模拟优势, 通过形成集成模型, 发展跨学科的综合模拟能力^[4]. 近年来模型集成的研究工作得到广泛关注, 比如 Oxley 等人^[5]集成自然、经济和社会

有关的 10 个模型用于地中海区域土地退化决策支持; Twarakavi 等人^[6]连接土壤水模型 Hydrus 和地下水模型 MODFLOW 以改善地下水模拟能力; Zhang 和 Xia^[7]耦合水文和生态过程更好地实现汉江流域的可持续水资源管理. 为了解决黑河流域日益严峻的生态环境和水资源危机, 李新等人^[8]建议在黑河流域形成水-土-气-生-人集成模型.

由于现有模型是被独立开发, 在数据和参数要求、时空尺度、文件格式、运行环境等方面都存在差异, 模型间的集成不但需要对模型所涉及专业知识的深入认识, 还需要软件系统设计和开发方面的大量技术工作, 尤其是多学科模型间的集成, 相当具有挑战性. 集成建模环境提供了一种模型集成的技术手段, 借助最新的计算机技术, 提供了快速生成新模型和模型比较的能力^[9]. 尽管建模环境的研究在国际上已获关注并进行了近 30 年的探索, 但其发展尚且

停留在实验室原型阶段. 最近的几年内一些大型实用建模环境开始获得资助, 并应用于气候、海洋、水文水资源等领域, 比如美国宇航局(NASA)支持的地球系统模拟平台(ESMF)^[10]、英国 e-Science 支持的“网格计算支持的地球系统模拟平台”(GENIE)^[11, 12]等. 国内对这一问题的关注始于最近 5 年, 相关研究才刚刚起步. 冯克庭等人^[13]形成了黑河流域建模环境原型, 在图形界面支持下模块化了新安江模型和 TOPMODEL 模型; 刘昌明等人^[14]开发了模块化的水文水资源模拟系统(HIMS), 包含了流域水循环主要过程, 并使用黄河流域和澳大利亚数据进行了验证; 香港中文大学和南京师范大学开展的虚拟地理环境是个更广义上的建模环境^[15]; 李新等人^[16]归纳了黑河流域模型集成的进展, 将集成方案分解为知识途径和技术途径, 其中技术途径主要是建模环境.

国际上生态和水文领域的建模环境依设计思想和应用方法可以分为 3 个类型. 第一类, 模型的分解、重构、运行和分析均在一个软件框架下进行, 往往提供良好的可视化界面; 第二类, 以类库的方式提供集成建模的主要功能和函数接口, 用户在这些类库基础上进行开发, 形成新模型, 此类往往追求卓越的运行性能; 第三类, 模型工具箱, 通常将多个水文模型集成到一个软件, 提供统一操作界面和辅助分析功能, 模型间集成能力有限. 三类建模环境优缺点归纳见表 1.

现有建模环境具备各自的学科背景, 比如考虑到生态和经济模型多应用通用代数建模系统(GAMS)进行编程, SME 提供了对 GAMS 代码的解构功能^[21]; MMS 主要是水文建模的支持, 而对分布式水文模型的支持不够^[18]; OMS 运行在 Java 平台, 计算效率是个瓶颈^[19]. 这决定了很难在现有建模环境基础上, 扩

展更多的功能, 或者满足将来的需求, 比如将知识库连接进建模环境, 以提供建模过程的智能化支持. 在此背景下, 中国科学院重点方向项目“地表过程建模环境和模型集成研究”致力于发展一个新的高效、跨平台的建模环境, 具备第一类的易用性和第二类的扩展性, 且能满足将来建模过程智能化的要求.

本文介绍此新的建模环境的系统结构、关键技术, 最后给出一个简单的应用实例, 展示建模环境在模型集成上的优势.

1 集成建模环境总体设计

建模环境面向具备专业知识的建模人员, 通过建模环境的软件平台支持, 快速实现功能模块的连接组合, 创建能够应对当前复杂科学问题的集成模型. 模块化的建模环境也提供了方便的模型比较分析功能. Argent^[9]认为集成建模环境应当具备以下基本功能: 1) 具备一系列可用的表示不同物理过程的模块; 2) 组织和管理模块和模型; 3) 通过连接不同模块, 形成模型; 4) 运行模型; 5) 分析模型运行结果. 本文发展的集成建模环境(HIME)整体框架如图 1 所示.

HIME 包括本地和网络模块/组件库. 一个模块通常对应于一个物理子过程, 具有独立的计算过程, 在软件层次上被实现为一个组件. 模块通常由成熟的水文、陆面过程或者生态模型分解得到, 模块有不同的粒度, 一个完整的模型也可作为一个模块. 模块/组件库组织和管理模块, 模块按类型进行归类, 每个模块有详细的元信息说明, 以表明模块来源、运行的时间和空间尺度、输入输出等. 模块可以存放在任意可访问的网络结点上, 通过网络服务(Web Services)

表 1 三类建模环境优缺点比较

	Advantage	Weakness	Typical work
First type, framework type	<ol style="list-style-type: none"> 1. good user interface, easy to use and learn 2. various models and functional modules provided to facilitate model creation 	<ol style="list-style-type: none"> 1. not easy for second development, limited open APIs 2. hard to be reused, developed for specific OS 3. low computational efficiency 4. limited analysis functions 	<ol style="list-style-type: none"> 1. MMS [17, 18] 2. OMS [19] 3. Tarsier [20] 4. SME [21]
Second type, shared library type	<ol style="list-style-type: none"> 1. flexible for development 2. high computation efficiency, good for Earth system modeling 3. shared on the library level, easy migration to other systems 	<ol style="list-style-type: none"> 1. long learning curve, high requirement to compute skills and model knowledge 2. no or simple user interface, hard to use 	<ol style="list-style-type: none"> 1. OpenMI [22] 2. ESMF [10]
Third type, model kit type	<ol style="list-style-type: none"> 1. easy to use, uniform user interface 2. rich analysis functions, GIS incorporated 	<ol style="list-style-type: none"> 1. few or limited model linkage support 2. few or limited APIs opened for development 3. not real modeling environment 	<ol style="list-style-type: none"> 1. WMS [23]

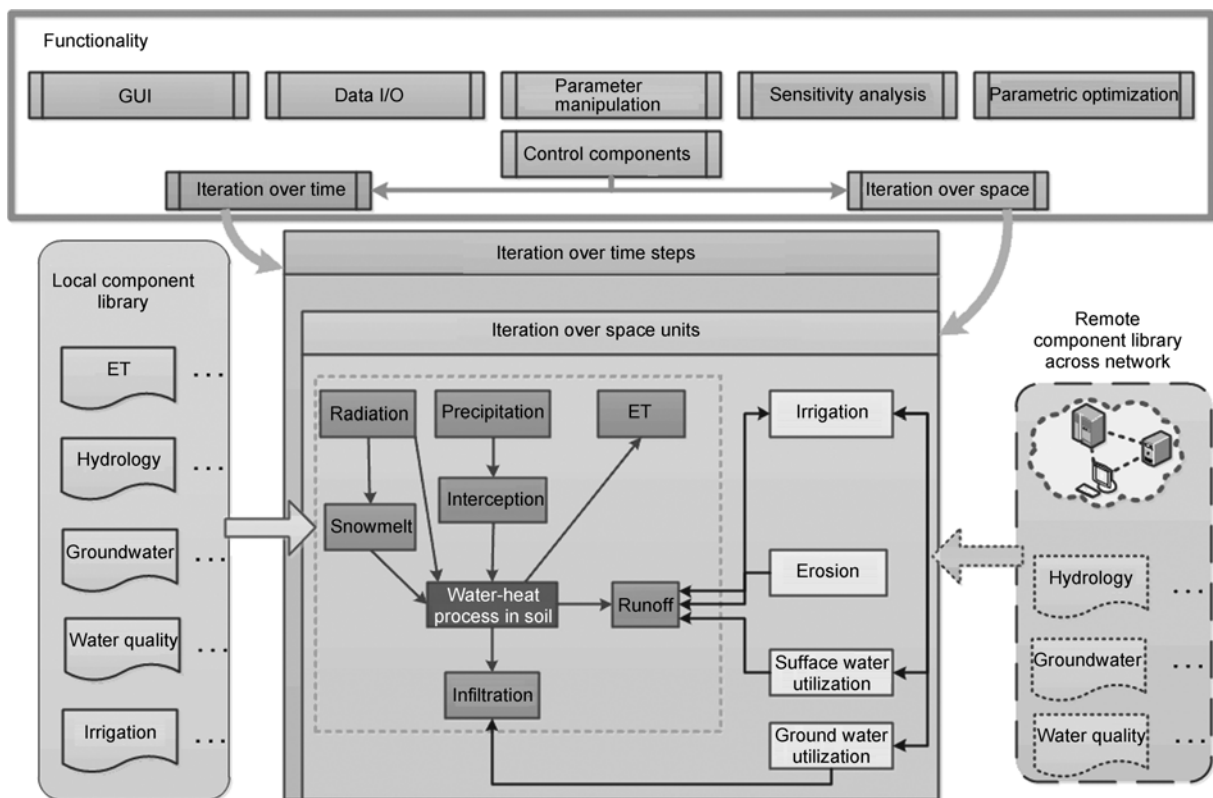


图1 集成建模环境整体框架

被建模环境访问。

本地或者网络上的模块在建模环境中进行连接和组装。模块的执行次序通过控制组件进行协调，控制组件包括时间步长上的迭代和空间单元上的计算过程控制。连接后的多个模块形成一个新模型，用户需要指定模块间的数据交换，并提供必要的驱动数据、参数和模型控制参数，在建模环境中运行并分析结果。借助于丰富的模块库，模型的创建和模型中物理子过程的替代就变得十分简单。

HIME 被设计为支持整个建模过程，包括参数的敏感性分析、参数率定、模拟运行、模型验证。需要明确的是，建模环境提供了建模的技术支持，自动完成很多复杂但通用的模块连接、组合、通讯等技术问题，使建模人员专注于模型背后的科学问题。然而建模过程仍需要大量的专家知识，比如如何选择满足研究目标、时空尺度合适的模块，如何合理连接模块，以及参数操作、模型运行和分析，都需要专家知识的输入。集成建模环境与专家知识系统从而形成智能化的建模环境，是未来的重要发展方向^[24]。

2 关键技术

2.1 模块组件化表示

建模环境里的模块通常对应于一个物理子过程。基本物理子过程，比如蒸发，是一个模块；几个模块也可以形成一个更大粒度上的模块存储在模块库中；甚至一个现有模型也可以做为一个模块。建模环境中，模型是建模过程的最终产物，模块则是建模过程中的计算单元。模块一般分解自现有的水文、生态和陆面过程模型，或者由建模人员自行编写。

在实现层次上，模块被封装，进而表示为组件。如图2所示，物理过程 X, Y 是 2 个组件，组件在建模环境中通过过程控制组件进行组装，形成模型。组件具备 3 个必要的接口，即初始化接口(init)，用于初始化模块内的变量和状态；运行接口(run)，具体执行模块的计算过程；清理接口(finalise)，在结束该模块前，释放内存和文件句柄等计算机资源。从现有模型中提取模块，需要少量的技术工作，包括按 3 个功能接口对已有代码进行封装，以及将模块的输入输出按

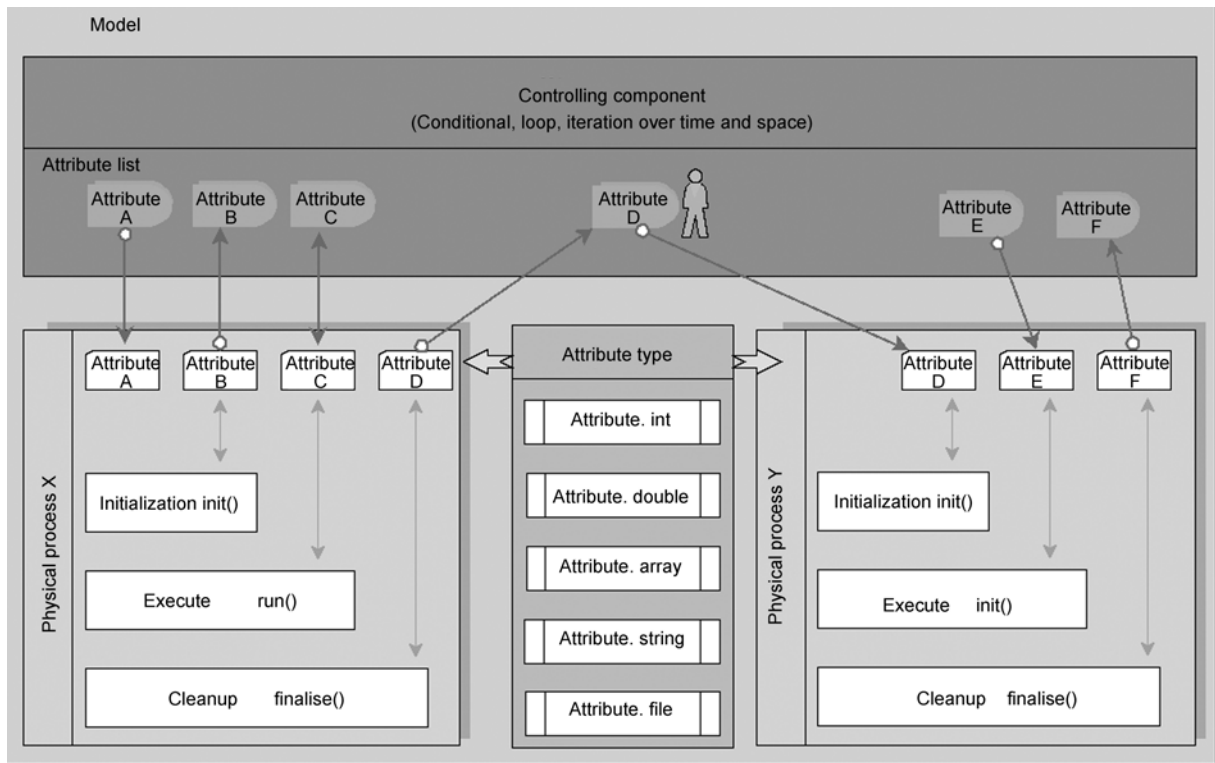


图 2 模块组装形成模型的示意图

属性封装. 当有模型的源代码时, 创建一个类(Class), 包括组件描述、作者、来源、参考文献、状态、版本、关键词、标签等模块描述信息, 每个属性的读写类型、单位、阈值等属性描述信息, 以及指定 init, run, finalise 接口对应的具体函数, 目前通过 Qt/C++^[25]的 Q_CLASSINFO 宏实现反射机制. 当我们只有模型的二进制代码时, HIME 系统支持通过一个 XML 对二进制代码进行解释, XML 文件里包括该模块对应的链接库名及库内的函数名. 图 3 示例了通过 XML 文件封装了一个来自 Noah 陆面过程模型链接库(simple-

driver.dll, 源码是基于 Fortran, 通过 GFortran 编译得到)的独立于时间步长设定的计算陆面水和能量平衡的模块. 鉴于不同平台上对链接库的物理表示不一致, 基于 XML 解释的方法目前只支持 GNU(如 GNU C++, GFortran)编译的链接库.

2.2 数据传递方案

多数建模环境如 Tarsier 采用统一数据模型进行模块间的数据传递^[26]. 统一数据模型在建模环境里构建一个通用的数据模型, 每个模块通过将需求的

```
<?xml version="1.0" encoding="UTF-8"?>
<library name="simpledriver.dll" description="" author="" bibliography="Natural" status=""
versionInfo="3.1" sourceInfo="simpledriver v3.1" keywords="Noah" label="Noah">
  <module name="module_sf_noahism">
    <function name="NoahSFLX">
      <argument name="ffrozp" intent="IN" type="double" unit="" range=""
arrayBound="" description="Fraction of precip which is frozen (0.0 - 1.0)."/>
      <argument name="ice" intent="IN" type="int" unit="" range=""
arrayBound="" description="Flag for sea-ice (1) or land (0)."/>
      <!-- More arguments here -->
      <result name="" type="" unit="" range="" arrayBound="" description=""/>
    </function>
  </module>
</library>
```

图 3 一个简化的 XML 文件, 用于解释二进制代码形成被 HIME 识别的模块

数据格式转换为此数据模型从而实现模块间的数据通讯. 考虑到模块对数据格式需求的多样性, 统一数据模型很难真正做到通用, 或者变得很复杂. HIME里采用与 OMS 类似的灵活的数据传递方案, 即通过组件的属性来实现. 此方案的优势在于避免了构建复杂的统一数据模型, 而在变量层次上需要交换的数据类型就相应简单(图 4).

属性类型包括整型、双精度浮点型、数组型、字符串型和规则文件型. 如果将组件想象为函数, 属性可以理解为函数的输入输出变量(图 4). 组件内部临时变量则不作为属性向外部公开. 属性有不同的读写类型, 图 2 中属性 A 具有读类型, 属性 B 具有写类型, 属性 C 具有读写类型.

模块间的数据传递通过模块间的属性连接实现. 在图 2 的简单示例中, 模型包括物理过程 X 和 Y(分别封装为模块 X 和 Y). 当 X, Y 添加时, X, Y 将各自的属性向上一级控制组件(图 2 例子中的上一级组件是根节点)申报, 根节点的属性列表具备了来自 X 的属性 A, B, C 和 D, 以及来自 Y 的属性 D, E 和 F. 图 2 例子中 Y 需要将 X 的 D 输出作为输入. HIME 系统将同名同属性类型的属性(如 D)进行自动连接(即自动连接 X, Y 模块中的属性 D), 也允许人工连接(比如需要将 X 的 D 输出作为 Y 的 E 输入). HIME 系统限定连

接的属性须是同一属性类型.

模块间的数据交换发生在上一级控制组件(或者模型根节点), 而不是直接在前后模块间发生交换, 确保了模块的独立性, 每模块无须关心与之相连的模块的输出输入需求. 在一个复杂的实际应用中, 需要交换数据的模块可能位于不同的位置, 比如一个位于时间循环内的模块 A 需要读取时间循环体外的模块 B 的输出, 这时, 模块 A 的上级时间循环控制组件除了包括来自循环体内的全部模块的属性列表, 也复制了更上一级控制组件的属性列表. 通过这种级联机制保证了一个新组件可以与此前任一位置的组件进行数据通讯.

网络模块利用网络服务(Web service)与本地或者其他网络模块进行关联, 采用本地模块一致的属性交换方案, 区别是数据通过网络在本地和网络模块所在的远程机器进行数据交换.

2.3 过程控制

分析现有水文和陆面过程模型, 子过程的执行模式, 无非是时间空间上的不同嵌套方式. 比如多数模型在每一空间计算单元(可能是网格、水文响应单元或者子流域等)进行逐个计算, 每个计算过程中再行考虑时间上的迭代, 每个时间单元上包括具体的物理计算过程. 对于用于实时预报模型, 往往又采用相反的时间空间嵌套方式, 即在一个时间步长内, 计算不同网格内的过程. 建模环境需要提供灵活的时间空间嵌套方式以满足不同的需求. 传统建模环境将时间和空间迭代放在模块体内, 使得模块不能被应用于不同时间步长或者不同的空间离散方法的模型里. 这种考虑基本上是正确的, 很多模块有其自身的时空尺度的限制. 然后这种底层框架上的限制, 也使得一些基本的物理过程, 原本是适用于更灵活的时间和空间尺度上, 也不能被直接使用, 同时也使得建模环境不能满足建立变时空步长模型的需求. HIME 建立了专门的过程控制组件, 用以控制物理过程模块的时间迭代和空间单元的计算, 从而将计算逻辑与时间空间迭代分离表示. 模块的时空尺度限制则通过模块元信息进行说明. 过程控制支持以下 3 种执行顺序.

- 1) 顺序执行. 依次执行各个组件.
- 2) 条件执行. 满足条件时依次执行条件内部各组件, 条件不满足时, 跳过执行.
- 3) 循环执行. 包括时间迭代过程中的时间条件

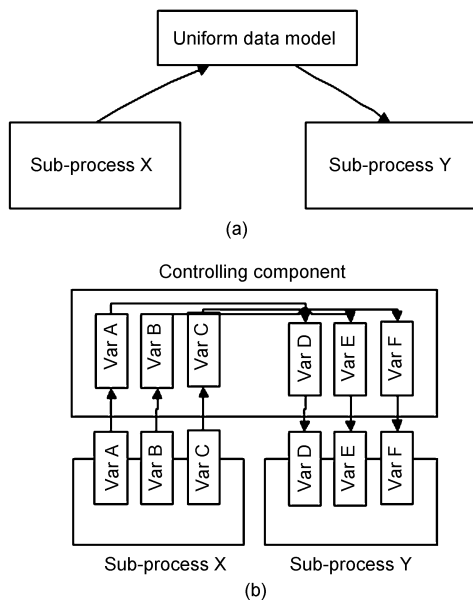


图 4 用于模型间数据交换的统一数据模型(a)和 HIME 采用的基于属性的数据传递方案(b)

控制和空间计算过程中的空间单元条件控制. HIME 支持基于格网和基于水文响应单元两种空间划分.

2.4 模块连接、重用和跨平台

模块在建模环境图形界面上表示为图标, 模块间关系表示为连线, 通过拖拉图标的方式形成新模型. 要创建一个新模型, 首先从本地和网络模块库中选择合适的表示不同物理过程的模块, 在图形化建模环境中组织模块, 指定模块间的属性连接. 组成模型的模块、模块间的连接关系、模型的初始输入存储在一个 XML(扩展标记语言)文件, 称为模型配置文件.

图 5 示例了一个简单的模型配置文件, 通过状态参数初始化、时间循环控制组件、状态参数更新、陆面过程计算、输出等几个模块, 连接为一个完整的 Noah 陆面过程模型. HIME 系统解析模型配置、组件配置、动态生成组件类, 合并生成模型工程文件, 调用编译器进行编译生成新模型. 解析和生成模型的过程是自动完成, 无须用户的参与. 最后, 在用户指定模型运行所需的驱动数据和参数后, 运行、查看和分析模型结果.

HIME 支持从源代码和二进制代码级别上的模块重用. 国内外已经有很多优秀的基于物理过程的水文和地表过程模型, 这些模型经过诸多的测试和应用, 在实现上一般都十分复杂, 重用现有模型和模

型里的物理子过程便显得十分重要. 多数建模环境支持具备模型源代码情况下的封装重用, 如 2.1 节的描述, HIME 经过少量的技术工作实现代码重用. 此外, 区别于多数建模环境, 由于 HIME 通过 GNU C++ 编译, 在二进制级别上兼容于 GNU 编译标准的任何代码, 从而实现了不同编程语言的混合调用. 对于一些闭源的科学模型, 二进制级别的重用是十分必要的. HIME 对模块重用的支持也体现在, HIME 与当前流行的建模环境 OMS, ESMF 采用类似的 XML 方式, 通过建立 XML 元素间的对应关系, 使得不同建模环境的模块重用成为可能.

模块解构和重组实践表明, 从现有模型中重用模块的主要耗时在于对模型的理解和物理过程逻辑关系的分解上, 技术上的封装往往只占整个解构过程的 10%~20% 时间. 因此, 在进行多模型集成时, HIME 推荐建模技术人员应当与模型专家进行紧密的合作, 可以极大加速建模进程.

考虑到科学模型开发和应用通常在不同操作系统平台上开展, 建模环境和模块的实现采用开源跨平台的 Qt/C++. Qt 是一个标准 C++ 的应用程序开发库, 可以在目前主要的操作系统包括 Windows, Mac OS X, Linux 和 Unix 等进行编译运行^[25], 而无需修改源代码. 建模环境辅助功能, 如 GIS 由基于 Qt 的 QGIS¹⁾实现, 绘图功能由同样基于 Qt 的 QtPlot²⁾

```
<?xml version="1.0" encoding="UTF-8"?>
<LSMS>
  <control type="base">
    <component name="NoahParamsInit" source="../noah/noahparamsinit/bin/libnoahparamsinit.a" underlaidobj="
simplifiedriver"/>
    <control type="loop" operator="&lt;" leftComp="NoahParamsUpdate" leftAttr="nowdate" rightComp=
"NoahParamsInit" rightAttr="enddate">
      <component name="NoahParamsUpdate" source="../noah/noahparamsupdate/bin/libnoahparamsupdate.a"
underlaidobj="simplifiedriver" />
      <component name="NoahSFLX" source="noahsflx.xml" underlaidobj="simplifiedriver" />
      <component name="NoahOutput" source="../noah/noahoutput/bin/libnoahoutput.a" underlaidobj="simplifiedriver"
/>
    </control>
  </control>
  <connection componentA="NoahParamsInit" attributeA="" componentB="NoahParamsUpdate" attributeB="" />
  <connection componentA="NoahParamsUpdate" attributeA="" componentB="NoahSFLX" attributeB="" />
  <connection componentA="NoahParamsInit" attributeA="" componentB="NoahSFLX" attributeB="" />
  <connection componentA="NoahSFLX" attributeA="" componentB="NoahParamsUpdate" attributeB="" />
  <connection componentA="NoahParamsInit" attributeA="" componentB="NoahOutput" attributeB="" />
  <connection componentA="NoahParamsUpdate" attributeA="" componentB="NoahOutput" attributeB="" />
  <connection componentA="NoahSFLX" attributeA="" componentB="NoahOutput" attributeB="" />
  <initialization component="NoahParamsInit" attribute="forcing_filename" value="bondville.dat" />
  <initialization component="NoahParamsInit" attribute="iunit" value="10" />
</LSMS>
```

图 5 简单解构后形成的 Noah 陆面过程模型的模型配置文件

1) <http://www.qgis.org/>, 访问日期 2011 年 3 月 3 日.

2) <http://soft.proindependent.com/qtiplot.html>, 访问日期 2011 年 3 月 3 日.

实现, 或者采用第三方开源标准 C++代码. 对于如 Fortran 语言开发的水文、陆面过程模型, 在二进制级别上实现跨语言调用. 所有这些努力, 使得 HIME 以一种开源免费方式运行于不同的操作系统.

3 验证与应用实例

为验证 HIME 能按预期设想实现建模目标, 本研究设计了 2 组实验. 实验一将 TOPMODEL^[27]模型解构, 形成物理子过程模块, 在建模环境中连接, 形成模块化后的 TOPMODEL, 比较模拟结果, 验证建模环境和 TOPMODEL 的模块化过程. 实验二使用 Noah 陆面过程模型^[28]中的蒸散发过程替换 TOPMODEL 蒸发过程, 验证在 HIME 支持下模块替换和创建新模型.

3.1 研究区和数据说明

实验一使用 Beven 在其 TMOD9701 版 TOPMODEL 所提供的英国 Slapton Wood 流域的数据. Slapton Wood 流域面积约 1 km², 数字高程(DEM)分辨率为 20 m. 60%流域面积在位于海拔 90 m 以上, 流域内地形起伏较小(一般小于 5%), 土地耕作密集, 主要是草原、谷物和块根作物. 低于 90 m 地区地形起伏较大(最大 25%), 主要是草原和森林(占流域 13.5%). 主要土壤类型为粘壤土^[29].

实验二使用黑河干流上游山区数据. 黑河干流上流面积 10018 km², DEM 来自 90 m 的 SRTM 高程数据³⁾. 地形指数最大 27.73, 最小 3.34, 分为 33 个级

别, 平均地形指数为 7.47. 用于汇流计算的河流距离分为 11 个级别, 最长距离为 304.4 km. 两者均大于 TOPMODEL 默认允许范围, 本研究在代码上做了相应更改. 地形指数和河流距离不同于文献[30]的描述, 文献[30]使用了 1500 m 分辨率的 DEM 可能无法充分地描述研究区地形情况. 降水数据来自祁连、札马什克、莺落峡、冰沟、新地、肃南、大野口和康乐等 8 个站点的平均值. 蒸发数据则来自祁连、札马什克、莺落峡、冰沟、新地和肃南等 6 个站点. 径流资料为莺落峡站实测流量. 单位均换算成 m h⁻¹. 使用了 1990 年 1 月 1 日至 2000 年 12 月 31 日资料, 其中 1990 年用于模型预热, 不参与参数标定, 1991 年至 1997 年用于参数标定, 1998 年至 2000 年数据用于模型验证.

3.2 模块代换方案

TOPMODEL 是由 Beven 和 Kirkby 发展的以地形为基础的半分布式水文模型, 用地形指数描述水文特性的空间不均匀性, 基于重力作用径流沿坡向运动原理, 模拟了径流产流过程, 在集总式水文模型和分布式水文模型之间起到了一个承上启下的作用^[27]. Noah 模型是一个优秀的陆面过程模型, 被用于模拟土壤-植被-大气连续体中的水量与能量通量过程, 也可以与大气模型耦合, 进行全球或局地尺度上的环境与气候情景模拟^[28]. 相比 TOPMODEL, Noah 能更精细的刻画水文过程的物理机制, 然而也需要更多的参数和驱动数据输入.

依据 TOPMODEL 的基本物理过程, 具有地形指数计算、植被截流、蒸发、产流、汇流等基本模块(图 6).

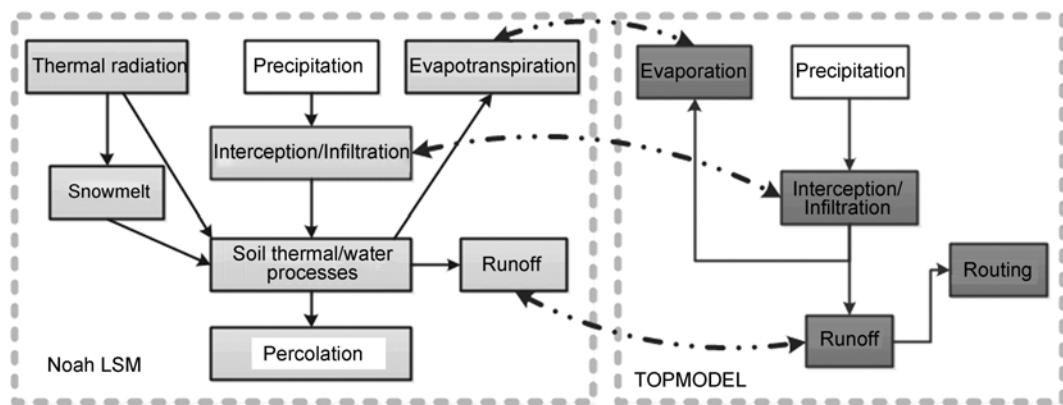


图 6 Noah 陆面过程模型和 TOPMODEL 模型的模块可替换性

3) <http://www2.jpl.nasa.gov/srtm/>, 访问日期 2011 年 3 月 3 日.

Noah 模型主要包括 7 个模块. 如图 6 所示, 在至少 3 个有近似物理目的而基于不同实现的模块(蒸散发、植被截流、产流)上可以进行替代. 结合陈仁升等人^[30]的研究结果, 实验二采用 Noah 模型的蒸散发模块改进 TOPMODEL 的蒸发模块, 一方面演示建模环境作为模型集成工具形成新模型的能力, 另一方面调查在 TOPMODEL 里显式增加蒸散发的表述, 能否提高 TOPMODEL 的总体模拟性能.

在 TOPMODEL 的蒸发计算中, 重力排水区或饱和和区域蒸发强度与潜在蒸发强度相等; 当重力排水区的水被消耗完之后, 流域内任意 i 处的植被根系层实际蒸发量 $E_{a,i}$ 计算公式如下

$$E_{a,i} = E_p (1 - S_{r,z,i} / S_{r,max,i}), \quad (1)$$

其中, $S_{r,z,i}$ 表示点 i 处植被根系层的缺水(m); $S_{r,max,i}$ 为根系区最大容水量(m), 可由田间持水量和凋萎含水量求出; E_p 为潜在蒸发能力(m/h).

Noah 模型中蒸散发量计算基本处理使用以下公式

$$E = E_{dir} + E_t + E_c + E_s, \quad (2)$$

其中, E_{dir} 由地表直接蒸发水量(W/m^2); E_t 由植被根系吸水后的蒸腾量(W/m^2); E_c 植被冠层截流的蒸发量(W/m^2); E_s 积雪升华(W/m^2).

$$E_{dir} = E_p (1 - \sigma_f) \frac{\theta_i - \theta_w}{\theta_{ref} - \theta_w}, \quad (3)$$

$$E_t = \sigma_f E_p B_c \left[1 - \left(\frac{W_c}{S} \right)^n \right], \quad (4)$$

$$E_c = \sigma_f E_p \left(\frac{W_c}{S} \right)^n, \quad (5)$$

$$E_s = \begin{cases} E_p & D_{snow} \geq E_p \delta_t, \\ \frac{D_{snow}}{\delta_t} & D_{snow} < E_p \delta_t, \end{cases} \quad (6)$$

其中, E_p 为潜在蒸发能力; σ_f 为植被覆盖度; θ_{ref} , θ_w , θ_i 分别为田间持水量、凋萎点和当前土壤体积含水量($m^3 m^{-3}$); B_c 为冠层阻力函数; W_c 为植被冠层截流水量(m); S 是冠层最大截流能力, n 取为 0.5^[31]; D_{snow} 为积雪深度(m).

因此在 TOPMODEL 中引入 Noah 蒸散发模块, 将引进 3 个新参数, 即植被覆盖度 σ_f 、冠层阻力函数 B_c 、冠层截流能力 W_c/S . 积雪升华 E_s 因为缺少合适的

雪深数据, 计算中暂不作考虑. 在 HIME 支持下, 原 TOPMODEL 蒸发模块替换后创建的新模型 MyTopmodel 可视化表示为图 7. 其中 Inputs 是读入降水、潜在蒸发、观测流量、地形指数等数据的模块; 3 个循环模块分别对子流域、时间和地形指数进行迭代计算; Initialize 初始化了时间循环体内的主要变量如流量、实际蒸发、总降水等; Infiltration excess 计算植被截流和土壤入渗; 在每个地形指数区间, Local storage deficit 计算水量亏损; Root zone, Unsaturated zone, Saturated area 分别是根系区、非饱和区、饱和区的流量; NOAH ET 即来自 Noah 模型的蒸散发模块; Cleanup 对地形指数区间进行必要的积累计算; Routing 对每时间步长上进行汇流; Balance calc 计算水量平衡各分量; Outputs 是输出模块.

选择模块后, 需要对模块属性进行连接, 比如 NOAH ET 模块需要输入 Local storage deficit 模块计算的水量亏损. 指定模块间的连接关系后, HIME 对模型进行编译, 形成可执行模型程序. 与 TOPMODEL 一样, 新模型可以执行参数率定、运行和结果分析.

3.3 参数标定

TOPMODEL 参数多是流域平均值, 无法直接得到, 因此在经验基础上给定参数取值范围, 应用模拟

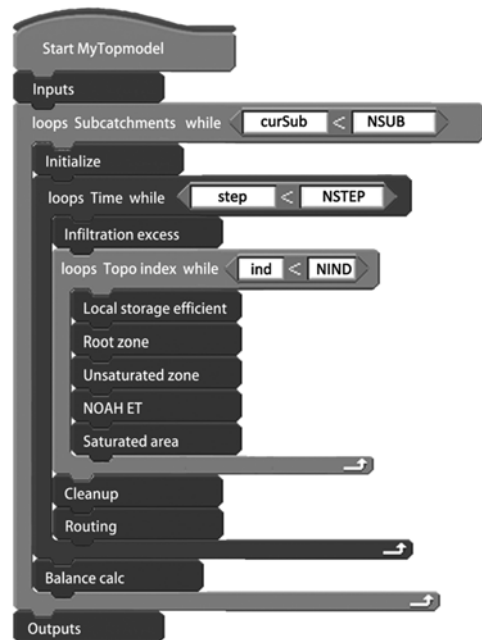


图 7 HIME 支持下的替换后 TOPMODEL 的模块可视化

退火算法^[32, 33]进行参数标定. 模拟退火算法设定如下: 初始温度 1000°, 终止温度 1°; 发生函数选定为自适应高斯正态函数; 分别计算了文献[33]描述的 9 种降温曲线; 最大迭代次数设定为 2000; 目标函数为 Nash-Sutcliffe 函数(NSE)^[34]. 首先对原 TOPMODEL 进行参数优化, 选择最优参数; 对于修改后的 TOPMODEL 模型, 保持 TOPMODEL 原参数, 对新增加的 3 个 Noah 蒸散发模块参数再次进行优化. 标定后参数取值列于表 2. 其中 Q_0 不参于标定, 取模拟起始(1990 年 1 月 1 日)莺落峡出口流量. 需要指出的是, 这里给出的参数取值与文献[30]不一致, 这是因为陈仁升等人的工作使用 2000 年 1 年的日资料, 参数是由经验给出.

3.4 模拟结果分析

实验一应用 Slapton Wood 流域的数据, HIME 支持下的模块化后 TOPMODEL 与未拆分前的 TOPMODEL 有一致的模拟结果, 证明了建模环境与模块化过程是准确的, 能按预期目标运行.

实验二使用 Noah 的蒸散发过程替换 TOPMODEL 蒸发过程形成新模型, 与原 TOPMODEL 的模拟结果进行对比分析. 莺落峡出山口径流观测和模拟结果如图 8 所示. 上行是原 TOPMODEL 在标定期和验证期的模拟结果; 下行是替换 Noah 蒸散发过程后的 TOPMODEL 模拟结果. 绿色表示的降水是输入模型的流域平均降水. 图 8(a)和(c)的标定期 NSE 效率系数分别是 0.521 和 0.558; 图 8(b)和(d)验证期的 NSE 分别是 0.697 和 0.698. 总体讲, 原 TOPMODEL 模型和显式考虑植被蒸腾的替换后模型均能较好反映水

文过程. 在标定期, 由于修改后 TOPMODEL 新增加了 3 个参数, 通过针对此三参数的优化, NSE 得到一定提高, 然而, 验证期的效率系数表明, 在 TOPMODEL 中显式考虑植被蒸腾并没有提升总体模拟性能.

对比 2 个模型的水文过程线, 显式考虑了植被蒸腾作用的模拟结果(图 8(c)和(d))表现出植被对河川径流的调节作用, 主要表现在洪峰流量的消减和枯水期流量略有增加 2 个方面, 在洪峰的拟合上优于原 TOPMODEL, 文献[35]也有相似的结论. 这可以解释为, 添加了植被蒸腾作用的 TOPMODEL 蒸发量有变化, 影响到流域的蓄水量, 从而导致产流量以及产流中地表径流和基流比例变化, 最终影响流量过程. 然而, 原模型和替换后模型对于春季径流模拟均不好, 结合祁连山区实际情况, 主要是两模型缺少对融雪径流的模拟.

实验二完整演示了基于建模环境的模型创建过程, 从模块解构、组合到参数率定和结果分析, 表明基于建模环境的模型集成是高效、可靠、有意义的.

4 结论

建模环境是一种有效的模型集成技术手段, 通过提供技术平台辅助建模人员快速集成模型, 应对复杂的涉及多学科模型的水文和地表过程科学问题以及流域/区域的管理需求. 尽管国际上已经发展了诸多用于不同学科领域的建模环境, 但都由于技术设计和学科侧重点的原因, 现有建模环境无法满足扩展性的要求, 比如无法将知识系统连接进建模

表 2 模型参数取值

Parameter	Definition	Range	Optimized
M	exponential storage parameter (m)	0.0001–0.03	0.0005983
$\ln(T_0)$	logarithmic saturated effective conductivity \ln (m^2/h)	0.5–1.5	0.853
T_d	unsaturated zone time delay per unit storage deficit (h)	0.001–50	0.8593
CHV	main channel routing velocity (m/h)	500–9000	5500
RV	internal subcatchment routing velocity (m/h)	500–5000	4500
S_{max}	the root zone available water capacity (m)	0.0–0.5	0.00012
Q_0	initial stream discharge (m/h)		0.000005
SR_0	initial value of root zone deficit (m)	0.0–0.5	0.001
σ	average vegetation coverage fraction	0.2–0.6	0.21918
B_c	canopy resistance function	0.2–0.8	0.22316
W_c/S	canopy interception capacity	0–0.95	0.12771

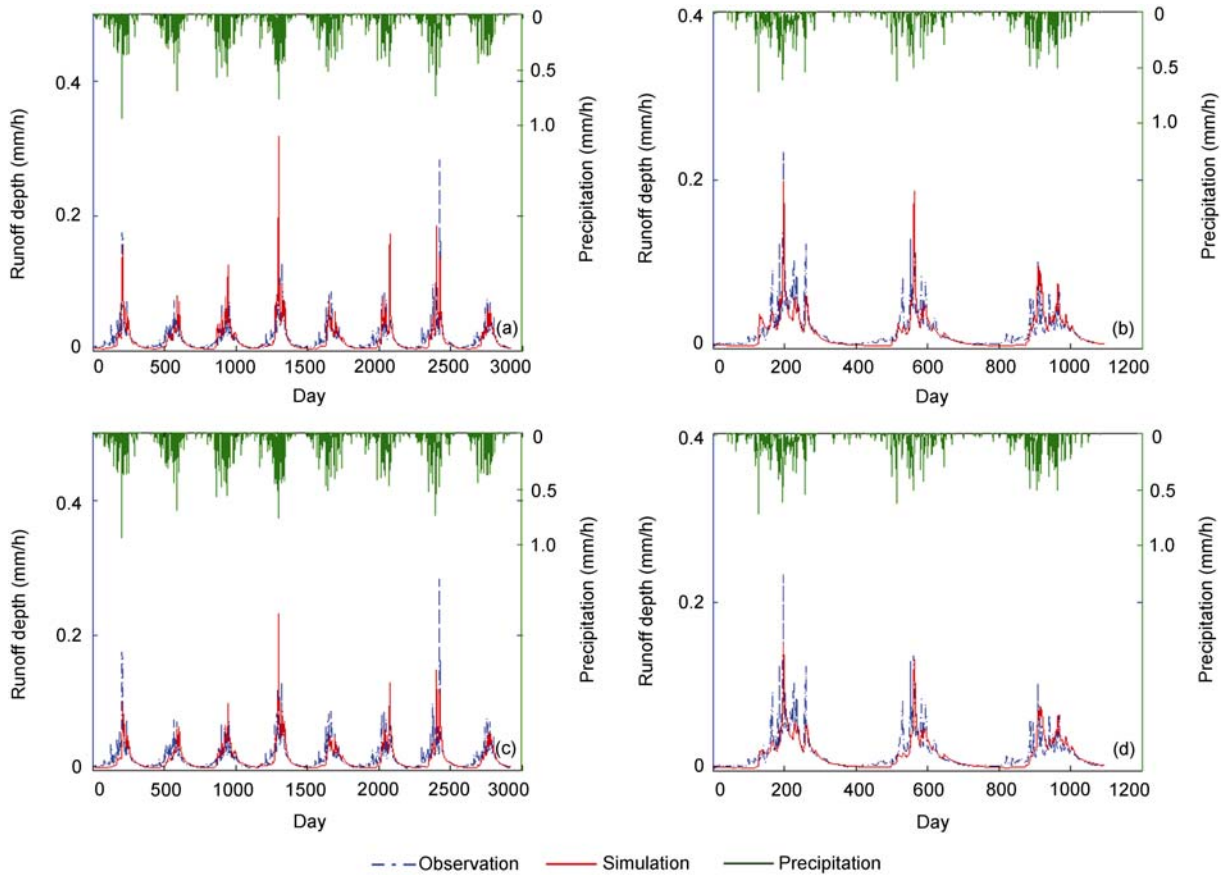


图 8 原 TOPMODEL 和修改后 TOPMODEL 模拟结果

(a) 原 TOPMODEL 预热期(1990.1.1~1990.12.31)和标定期(1991.1.1~19997.12.31); (b) 原 TOPMODEL 验证期(1998.1.1~2000.12.31); (c) 修改后 TOPMODEL 标定期; (d) 修改后 TOPMODEL 验证期.

环境, 以实现智能化的建模平台. 我们设计和实现的用于水文和地表过程领域的建模环境, 在学习现有建模环境结构和实现策略基础上, 具有如下特点.

- 1) 组件化的建模方法, 通过图标拖放等可视化手段连接物理过程模块, 形成集成模型;
- 2) 依托灵活的模块控制和数据传递机制, 模块间通讯和连接灵活且高效;
- 3) 通过解释元信息(实现为 XML 文件), 实现源码级别和二进制级别的模块重用, 第三方模块可以方便地被重用, 同时建模环境支持多种操作系统平台;
- 4) 开放的建模环境接口, 侧重于水文和地表过程建模, 易于接入知识系统以实现智能建模.

应用初步建立起来的水文建模环境, 我们进行了 2 组实验, 测试了建模环境的模块解构和组合功能. 并使用 Noah 陆面过程模型的蒸散发模块替换了 TOPMODEL 的蒸发模块, 结果表明, 在 TOPMODEL

中显式考虑植被蒸腾并没有提升总体模拟性能, 然而, 增加显式植被蒸腾影响, 使得基流在时间上进行了重分配, 在洪峰的拟合上优于原 TOPMODEL, 同时使得枯水期基流稍为增大. 此实验表明了基于建模环境的模型集成和模型比较是高效、可靠、有意义的.

目前进行的工作是将更多的气文和陆面过程模型分解到建模环境, 在黑河流域上游主要考虑现有模型对积雪、冻土模拟的改进, 中游主要是地表地下水交互, 农田灌溉的影响, 并且集成土地利用等人文过程. 考虑到科学模型运行在高时空分辨率下计算量往往十分大, 建模环境将逐步集成高性能计算的支持, 一方面是框架级别上集成 OpenMPI 等并行支持库, 另一方面在模块级别上添加必要的并行代码以支持模块的并行运算. 未来云计算的支持也是工作方向之一.

致谢 感谢中科院寒旱所陈仁升研究员提供了黑河干流上游山区部分研究数据;感谢美国地质调查局 George H Leavesley 教授在模块化建模环境 MMS 和对象建模环境 OMS 学习和应用上的极大帮助;感谢甘肃水文勘测设计院的胡兴林高工提供了黑河流域相关水文数据;感谢中国科学院北京地理所周成虎研究员在集成建模环境设计和实现上的宝贵意见;感谢美国匹兹堡大学王书功博士提供了 Noah 模型应用的相关帮助和对模拟结果的讨论.

参考文献

- 1 Westervelt J D. Simulation Modeling for Watershed Management. New York: Springer-Verlag, 2001. 1–189
- 2 Khan S, Malano H. Integrated modelling approaches to support water resource decision making: Crossing the chasm. In: Anderssen R S, Braddock R D, Newham L T H, eds. 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Cairns, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, 2009. 3838–3843
- 3 Pomeroy J W, Gray D M, Brown T, et al. The cold regions hydrological process representation and model: A platform for basing model structure on physical evidence. *Hydrol Processes*, 2007, 21: 2650–2667
- 4 Brandmeyer J E, Karimi H A. Coupling methodologies for environmental models. *Environ Modelling Software*, 2000, 15: 479–488
- 5 Oxley T, Mcintosh B S, Winder N, et al. Integrated modelling and decision-support tools: A Mediterranean example. *Environ Modelling Software*, 2004, 19: 999–1010
- 6 Twarakavi N K C, Simunek J, Seo S. Evaluating interactions between groundwater and vadose zone using the HYDRUS-based flow package for MODFLOW. *Vadose Zone J*, 2008, 7: 757–768
- 7 Zhang X, Xia J. Coupling the hydrological and ecological process to implement the sustainable water resources management in Hanjiang River Basin. *Sci China Ser E-Tech Sci*, 2009, 52: 3240–3248
- 8 李新, 程国栋. 流域科学研究中的观测系统和模型系统建设. *地球科学进展*, 2008, 23: 756–764
- 9 Argent R M. An overview of model integration for environmental application-components, frameworks and semantics. *Environ Modelling Software*, 2004, 19: 219–234
- 10 Hill C, Deluca C, Balaji, et al. The architecture of the earth system modeling framework. *Comput Sci Eng*, 2004, 6: 18–28
- 11 Lenton T M, Myerscough R J, Marsh R, et al. Using GENIE to study a tipping point in the climate system. *Philos Transact R Soc A-Math Phys Eng Sci*, 2009, 367: 871–884
- 12 Price A, Lenton T, Cox C, et al. GENIE: Grid enabled integrated earth system model. *ERCIM News*, 2005, 61: 15–16
- 13 冯克庭, 南卓铜, 赵彦博, 等. 基于插件的集成建模环境原型开发研究. *遥感技术与应用*, 2008, 23: 587–591
- 14 刘昌明, 王中根, 郑红星, 等. HIMS 系统及其定制模型的开发与应用. *中国科学 E 辑: 技术科学*, 2008, 38: 350–360
- 15 林琿, 黄凤茹, 闫国年. 虚拟地理环境研究的兴起与实验地理学新方向. *地理学报*, 2009, 64: 7–20
- 16 李新, 程国栋, 康尔泗, 等. 数字黑河的思考与实践 3: 模型集成. *地球科学进展*, 2010, 25: 851–865
- 17 Leavesley G H, Markstrom S L, Brewer M S, et al. The modular modeling system (MMS)—The physical process modeling component of a database-centered decision support system for water and power management. *Water Air Soil Pollu*, 1996, 90: 303–311
- 18 Leavesley G H, Markstrom S L, Restrepo P J, et al. A modular approach to addressing model design, scale, and parameter estimation issues in distributed hydrological modelling. *Hydrol Proc*, 2002, 16: 173–187
- 19 David O, Markstrom S L, Rojas K W, et al. In: Ahuja L, Ma L, Howell T A, eds. *Agricultural System Models in Field Research and Technology Transfer*. Boca Raton: CRC Press, 2002. 317–331
- 20 Watson F, Rahman J M. Tarsier: A practical software framework for model development, testing and deployment. *Environ Model Software*, 2004, 19: 245–260
- 21 Voinov A, Costanza R, Wainger L, et al. Patuxent landscape model: Integrated ecological economic modeling of a watershed. *Environ Model Software*, 1999, 14: 473–491
- 22 Blind M, Gregersen J B. Towards an open modelling interface (OpenMI): The HarmonIT project. In: *Complexity and Integrated Resources Management, Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society*, vol 1. Manno, Switzerland: iEMSs, 2004. 346–351
- 23 Erturk A, Gurel M, Baloch M A, et al. Application of watershed modeling system (WMS) for integrated management of a watershed in turkey. *J Environ Sci Health, Part A: Tox/Haz Subst Environ Eng*, 2006, 41: 2045–2056

- 24 Villa F, Athanasiadis I N, Rizzoli A E. Modelling with knowledge: A review of emerging semantic approaches to environmental modelling. *Environ Model Software*, 2009, 24: 577–587
- 25 Molkenin D. *The Book of Qt 4: The Art of Building Qt Applications*. San Francisco: No Starch Press, 2007. 1–440
- 26 Rahman J M, Cuddy S M, Watson F. Tarsier and ICMS: Two approaches to framework development. *Math Comput Simul*, 2004, 64: 339–350
- 27 Beven K J, Kirkby M J. A physically based variable contributing area model of basin hydrology. *Hydrol Sci Bull*, 1979, 24: 43–69
- 28 Chen F, Mitchell K, Schaake J, et al. Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *J Geophys Res-Atmos*, 1996, 101: 7251–7268
- 29 Birkinshaw S J, Ewen J. Modelling nitrate transport in the Slapton Wood catchment using SHETRAN. *J Hydrol*, 2000, 230: 18–33
- 30 陈仁升, 康尔泗, 杨建平, 等. Topmodel 模型在黑河干流出山径流模拟中的应用. *中国沙漠*, 2003, 23: 428–434
- 31 Chen F, Dudhia J. Coupling an advanced land surface-hydrology model with the Penn State -NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Rev*, 2001, 129: 569–585
- 32 Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by simulated annealing. *Science*, 1983, 220: 671–680
- 33 王书功. 水文模型参数估计方法及参数估计不确定性研究. 博士学位论文. 兰州: 中国科学院寒区旱区环境与工程研究所, 2006. 1–230
- 34 Nash J E, Sutcliffe J V. River flow forecasting through conceptual models. 1. A disussion of principles. *J Hydrol*, 1970, 10: 282–290
- 35 王礼先, 张志强. 干旱地区森林对流域径流的影响. *自然资源学报*, 2001, 16: 454–459