

高性能计算环境下陆面建模系统的关键技术*

罗立辉^{1,2,3}, 张耀南^{1,2,3}

¹(中国科学院 寒区旱区环境与工程研究所, 甘肃 兰州 730000)

²(甘肃省资源环境科学数据工程技术研究中心, 甘肃 兰州 730000)

³(中国科学院 超级计算兰州分中心, 甘肃 兰州 730000)

通讯作者: 张耀南, E-mail: yaonan@lzb.ac.cn, http://www.career.cas.cn

摘要: 为了从多角度精确评估陆面特征,改善陆面过程模型的模拟性能,并为研究者提供一套完整的从数据处理到模拟分析的陆面建模系统.采用多种脚本语言和模型数据融合方法来构建陆面建模系统.此建模系统集成观测数据、陆面过程模型、高性能计算、数据处理和分析方法,以及可视化等技术手段.在此系统内针对两种不同的陆面过程模型进行了应用示范,证明了不同脚本语言的建模系统在目前高性能计算环境中的应用潜力,以及不同可视化方案在陆面建模系统的作用.

关键词: 高性能计算;陆面过程模型;陆面建模系统;脚本语言

中文引用格式: 罗立辉,张耀南.高性能计算环境下陆面建模系统的关键技术.软件学报,2013,24(Suppl.(2)):80-88. <http://www.jos.org.cn/1000-9825/13026.htm>

英文引用格式: Luo LH, Zhang YN. Key technologies of land surface modeling system in high performance computing environment. Ruan Jian Xue Bao/Journal of Software, 2013,24(Suppl.(2)):80-88 (in Chinese). <http://www.jos.org.cn/1000-9825/13026.htm>

Key Technologies of Land Surface Modeling System in High Performance Computing Environment

LUO Li-Hui^{1,2,3}, ZHANG Yao-Nan^{1,2,3}

¹(Cold and Arid Regions Environmental and Engineering Research Institute, The Chinese Academy of Sciences, Lanzhou 730000, China)

²(Gansu Data Engineering and Technology Research Center for Resources and Environment, Lanzhou 730000, China)

³(Lanzhou Super Computing Center, The Chinese Academy of Sciences, Lanzhou 730000, China)

Corresponding author: ZHANG Yao-Nan, E-mail: yaonan@lzb.ac.cn, http://www.career.cas.cn

Abstract: In order to assess accurately the characteristics of the land surface from different angles, and provides guidance to improve the simulation performance of land surface models, the framework of LSMS (Land Surface Modeling System) is constructed on the basis of scripting languages and Model-Data Fusion methods to provide researchers a land surface modeling environment from data processing to simulation analysis. The modeling system is an integration of observing data, land surface simulation, high-performance computing, data processing and analysis, and visualization techniques. Different land surface models were used in the framework to prove the utilization potentials of modeling system based on scripting language in high-performance computing environment and the role of various visualization schemes in the modeling system.

Key words: high performance computing; land surface model; land surface modeling system; scripting language

近年来,陆面建模系统(land surface modeling system,简称 LSMS)已日趋成为陆面过程模型研究中的关键部分.中国南方洪涝和西南干旱等一系列的自然灾害相继发生,使得研究者必须能够为上述问题的解决提供实

* 基金项目: 国家自然科学基金(J1210003, 41301508, 91125005); 中国科学院寒区旱区环境与工程研究所青年基金(51Y351061)

收稿时间: 2012-08-05; 定稿时间: 2013-07-22

时、准确的数据、信息和知识的支撑,这就必然要求能够提供一套从数据采集到结果可视化的即时的(或者延时的)软件平台.如何将数据、知识、模型算法、软件工具等联合在一起构建陆面特征信息化的科学研究环境,实现高性能计算与协同研究,是对当前陆面建模系统的必然要求与发展目标^[1-3].

陆面建模系统是地学建模系统的一个分支.地学建模系统起步于 21 世纪初,在高分辨率时空数据、高性能计算及分布式可视化等需求的基础上发展而来,早期的工作主要包括地球系统科学 workflow 环境^[4,5]和陆地信息系统^[6,7],其设计与实现都是通过将各种已有的专用软件框架和工具包加以统一、标准化和扩展,从而达到充分共享资源、快速构建地学建模应用的目的.陆面建模系统是在陆面过程模型的基础上,集成地球观测系统、陆面过程模型、高性能计算、数据处理和分析方法以及可视化等技术手段,从多角度来精确评估陆面动量、能量和水分状况等陆面特征,为改善陆面过程模型的模拟性能提供指导,更为重要的是为研究者提供一套完整的从数据处理到模拟分析的陆面建模环境^[7-9].

1 陆面过程模型

陆面过程是指发生在地表,控制地气之间水分、热量和动量交换的作用过程,包括地面上的热力过程、水文过程和生物过程,地气间的能量和物质交换以及地面以下土壤中的热传导和水热输送过程等.它涉及了地球系统五大圈(大气圈、岩石圈、生物圈、水圈和冰冻圈)中几乎所有的圈层^[10].陆面作为全球气候系统的重要组成部分,在气候变化中的作用显著.目前的陆面过程模型基本上都包含了如下的通用的组件:(1) 系统边界;(2) 强迫输入;(3) 初始状态;(4) 参数;(5) 模型结构;(6) 模型状态;(7) 模型输出.陆面过程模型包含了物理、化学和生物过程,其与能量、水和碳密切相关,它们对不同时间尺度的环境变化相当敏感.陆面过程模型包括不同时间尺度的一系列过程,大气观测及其分析数据驱动陆面过程模型的模拟应用,或者气候模型直接耦合陆面过程模型.

陆面过程模型一般面向站点和区域的模拟.区域网格的大小及时间分辨率的尺度决定了陆面过程模型的模拟需要在高性能计算平台下计算.高时空分辨率数据的模拟计算是目前面临的一个不小的挑战.一般来说,陆地表面分为二维的区间或者网格单元,例如,1km×1km 尺度的全球范围大约是 5×10^8 个网格单元.假设时间步长为 15 分钟的单点,采用单 CPU 用陆面过程模型模拟一天的时间为 15ms,那么,模拟 1km 的空间范围,用单 CPU 模拟 1 天的时间约为 3 个月.陆面建模系统将采用全球 1km 的数据集,包括 LAI、土壤、地形和其他陆地表面参数.强迫数据来自于不同的数据源,如全球通量观测网络 FLUXNET^[11]和气候模型 WRF^[12]的输出.随着网格点和高分辨率数据的增长,内存和硬盘存储的需求也显著增加,全球 1km 的模拟大约所需的内存和磁盘空间为 T 级别.陆面过程模型采用高性能计算技术来迎接诸如此类的挑战,陆面建模系统在模拟时提供有效的管理、存储、高吞吐量的访问,且提供一系列的工具来管理模拟计算.

陆面过程模型的模拟计算可以分为 3 个阶段:前处理、模拟计算和后处理.前处理主要包括数据的制备、处理和分析、参数的制备、模型的全局、本地和运行选项设置等;模拟计算包括模拟的启动、模拟的初始化和模拟运行中的数据读取、数据交换和运行次序的控制与管理;后处理主要包括统计分析、模拟输出数据的处理及可视化.对陆面过程模型通用组件的认识,以科学 workflow 方式构建陆面建模系统提供了理论和技术支持.将陆面过程模型模拟计算的 3 个阶段进行了细分,分为 6 个小阶段作为陆面建模系统的工作流(如图 1 所示),分别为:(1) 模拟配置;(2) 模拟启动;(3) 模拟初始化;(4) 模拟运行;(5) 后处理;(6) 模拟结束.

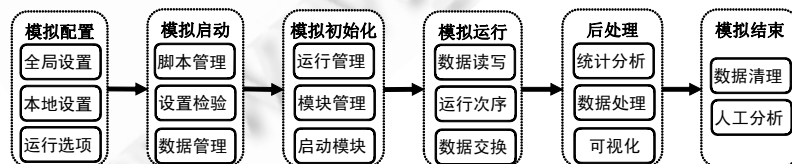


Fig.1 The workflow of land surface modeling system

图 1 陆面建模系统的工作流

2 陆面建模系统

陆面过程模型的模拟计算需要消耗大量的计算资源及存储空间,所以必须在大型计算机上进行模拟计算.众多的陆面过程模型在大型计算机上进行模拟计算有如下的4个特点:(1)以Linux或类Linux的SHELL为核心的命令行下的数据存取和模拟计算;(2)模型的复杂度不仅决定了其计算需求,而且决定了其从模型所需数据的处理到模拟运行结束后的后处理的复杂度;(3)模型由许多子模块组成,模拟计算时可以对运行的子模块进行调用和控制;(4)大型机计算机提供了各种开源的工具及脚本语言,而且SHELL能够方便地调用这些资源,集成不同的脚本语言和程序语言.在已有陆面过程模型的基础上,模型的模拟计算如何实现全自动(automatic)?模型的各个功能模块(从前处理到后处理)以及子模块之间如何衔接?如何将复杂的模拟计算过程简化?如何将模拟结果可视化?如何改善模型的模拟性能?如何制定其各个步骤之间的策略?这些问题是地学模型使用者目前所面临的问题.

2.1 脚本语言

陆面过程模型模拟应用的特点决定了陆面建模系统框架的建立,考虑现有的高性能计算的访问和计算模式,拥有大型计算机的机构和单位一般要求使用SSH等进行登录,然后采用作业调度系统提交计算作业,用户一般不能使用GUI接口进行操作,且对大型机操作系统没有控制权限,而且陆面过程模型的模拟计算过程中数据的访问与处理、作业提交、模拟后处理等工作往往需要多个脚本语言分工合作完成.

陆面建模系统其前处理主要由NCO,CDO,R,GrADS和NCL脚本语言处理分析数据,NCO,CDO和GrADS适于处理如NetCDF,GRIB,HDF等格式的输入、输出数据,对站点和区域的模拟而言,R语言对站点的统计分析和数据处理方面具有强大的功能,而NCL语言则更擅长于区域的统计分析与可视化,LSF作为高性能计算环境中的模拟计算的作业调度软件,可以封装在SHELL脚本中自动调用.为了集成和统一调度陆面过程模型的模拟,采用SHELL(适合KSH和BASH)的方式封装和衔接CDO,NCO,R,GrADS,NCL和LSF脚本程序,在陆面建模系统中配置模拟计算的各个功能模块之间的衔接.SHELL封装的脚本具有模拟运行的前处理、模拟运行、后处理等功能,使其衔接成一个整体,与陆面建模系统的科学 workflow 方式结合,从数据的制备到模拟计算再到可视化等模拟应用所需要的功能模块都集成在此封装脚本中,从而构建了一套自动的基于脚本语言的陆面建模系统(如图2所示).陆面建模系统的核心功能主要由R和NCL这两种脚本语言编写相应的功能函数来实现.

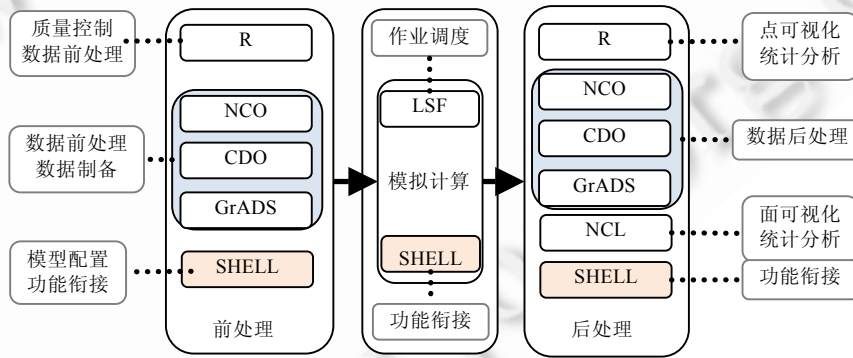


Fig.2 The conceptual graph of modeling environment based on scripting languages

图2 基于脚本语言的建模环境的概念图

2.1.1 构建R功能函数

在模型模拟的前处理部分,R用于制备数据、质量控制和空值填补,数据制备加载了RNetCDF包,采用相邻7个值的方法对数据进行了质量控制,而且对每个变量都设置了旗标(flag)来限制其值域的范围.模型模拟的后处理部分采用R语言来实现其功能(见表1).后处理中加载了R语言的ncdf,gdata,gplots,plotrix,Hmisc,lattice和泰勒R软件包.在此基础上构建了陆面过程模型模拟后处理的R语言功能函数,其代码特点具有普适性,能够根

据模拟结果参数等的变化调整其代码,R 脚本也嵌套进入 SHELL 脚本里,SHELL 可以根据实际需要调用 R 脚本,然后将可视化以 pdf,eps,emf,ps 等格式文件输出,而且可以用 x11()功能,将所有图像输出到 RGui 的图形界面上,其实现的主要功能及函数见表 1.

Table 1 Main functions of the R language for land surface modeling system

表 1 R 语言构建陆面建模系统的主要功能函数

功能函数	说明
图形函数 LSM_FLUXNET_GRAPHIC LSM_FLUXNET_GRAPHIC_FAPAR LSM_FLUXNET_GRAPHIC_FORCING LSM_FLUXNET_GRAPHIC_ERR LSM_FLUXNET_STATS_GRAPHIC LSM_FLUXNET_TAYLOR GWS_ANALYSIS	基本的图像函数,包括绘图函数 plot,方差分析,箱图等功能 专门为 fAPAR 数据定制的图像函数 与强迫数据集进行对比分析的图形函数 误差分析的图形函数 统计分析的图形函数 泰勒图 小波分析
数据函数 FLUXNET_2_NETCDF LSM_DATA FLUXNET_DATA	数据制备,将 FLUXNET 数据转换为 NetCDF 格式 读写模拟结果 读写 FLUXNET 数据
时间操作 LSM_FLUXNET_COMPUTE_GROUPING	不同时间分辨率的转换
模型评估 LSM_FLUXNET_EVALUATION	15 种统计分析评估
常规函数 CONVERT_SITE_TO_PARAMETER DIFF_VERSION_COMPARE CALC_RELATIVE_HUMIDITY_IN_SOIL LONLAT	站点与参数之间的数据类型转换 比较不同的模型版本的模拟结果 计算不同土壤模型的相对湿度(水箱和五层土壤模型) 经纬度
主程序 MAIN_PROGRAM DIFF_PARAMETER DIFF_SITE	控制评估流程及输出模式 不同参数间的对比分析 不同站点间的对比分析

例如,LSM_FLUXNET_EVALUATION 函数集成了 15 种不同数学统计分析方法来评估陆面过程模型.因为每一种数学统计方法应用于模型模拟性能的评价都有其局限性,所以在陆面建模系统的模型评估中采用的评价方法有斜率(slope)、截距(intercept)、相关系数(correlation coefficient, r)、决定系数(coefficient of determination, R^2)、均方根误差(root mean square error,简称 RMSE)、离散率(ratio of scatter)、标准化均方根误差(normalized RMSE,简称 NRMSE)、标准化平均误差(normalized average error,简称 NAE)、方差率(variance Ratio,简称 VR)、标准偏差(standard deviation,简称 SD)、偏差百分比(percent bias,简称 PBIAS)、RMSE 观测标准偏差率(RMSE-observations standard deviation ratio,简称 RSR)和 Nash 效率系数(Nash-Sutcliffe efficiency,简称 NSE)等 15 种方法^[13].

其中,偏差百分比(PBIAS)(如式(1)所示)衡量模拟与观测值的平均趋势,PBIAS 的最佳值为 0,PBIAS 的值越少表明所具有的模拟性能越好.正值表明模型低估了偏差,负值表明模型高估了偏差. S_i 表示 i 时刻的模拟值, O_i 表示 i 时刻的观测值, \bar{S} 是模拟值的平均值, \bar{O} 是观测值的平均值; n 为样本数.

$$PBIAS = \left[\frac{\sum_{i=1}^n (O_i - S_i) \times 100}{\sum_{i=1}^n (O_i)} \right] \quad (1)$$

RMSE 经常用于模型评估,尽管较低的 RMSE 值具有较好的模拟性能,但是不能直观地反映模拟性能,基于 RMSE,与观测的标准差相除得到 RSR(如式(2)所示).RSR 吸收了误差索引统计的优点,其中包含比例/标准化因素.RSR 的最佳值为 0,它显示 RMSE 或者残差的变化为 0.通常当 RSR 和 RMSE 值较低时,说明模拟性能较好.

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

Nash 效率系数(NSE)是残差(噪音)与观察数据的方差(信息)之比的标准化统计方法(如式(3)所示).NSE 的值为 $-\infty \sim 1$,1 是 NSE 的最佳值,NSE 在 0 和 1 之间经常被认为是模拟性能可接受的范围.当 NSE 值小于 0 时,表明模拟性能较差.

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - S_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (3)$$

2.1.2 构建 NCL 功能函数

新增 NCL 脚本语言作为陆面过程模型区域模拟的可视化方案,主要在数学统计和可视化中构建了以下的 NCL 函数及功能:(1) 经纬度与网格点的转换;(2) 时间序列数据的累计计算,包括求不同时间尺度的平均值、和、偏差等;(3) 针对净辐射、感热通量、潜热通量和地热通量,制定了其区域的 2D 的可视化功能;(4) 值域函数,主要是针对几个主要的变量定制其值域范围,以方便 ColorBar 的输出;(5) 数据的读取和写入功能,因为潜热通量、净辐射和地热通量在 CLM 的模拟中并没有直接地以一个变量的形式直接地输出,需要进行计算,所以在计算之后用 NCL 脚本语言写入计算后的变量,并保存为 NetCDF 文件,以便下一次直接调用;(6) 通用的图形函数,针对 CLM 的模拟定制了 NCL 图形绘画中的基本图形的设置.

2.2 模型数据融合

模型数据融合(model-data fusion,简称 MDF)通过参数估计、敏感性分析、误差权重、价值函数和数据同化等方法(序贯和非序贯)联合模型和观测数据,以提高模型模拟性能的能力.模型数据融合是一种改善数据、模型、模拟性能的方法集合.它以模型评估为基准,在观测数据和模型之间集成不同的方法,比如,现今比较流行的数据同化算法也是模型数据融合的方法之一.模型数据融合不仅仅是集成模型和各种观测因子,并不断依靠观测数据来自动调整模型的轨迹以减少误差的方法^[14],而且还考虑了数据的不确定性、模型参数中的敏感性、模型模拟结果的评估.模型数据融合是一个多阶段的过程,在每一个阶段,数据、模型参数、模型结构和模型开发者之间相互影响.这样就避免了数据、参数和模型不确定性带来的误差,并将不同的模拟评估技术应用于其中,从观测、分析和预测为模型开发者和使用者提供全方位的模型优化手段.

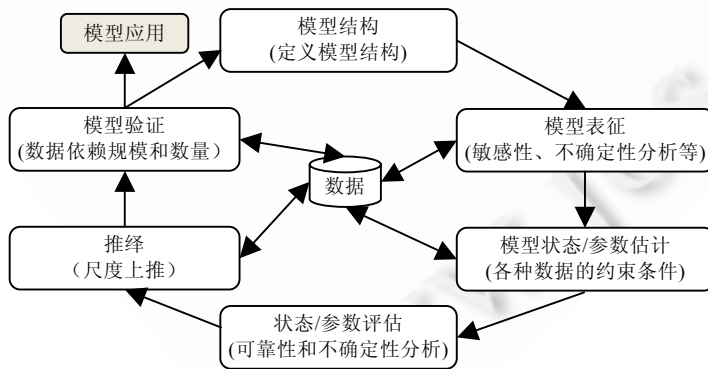


Fig.3 The multi-stage process for model-data fusion

图3 多阶段模型数据融合过程

统计分析 R 函数等构成了此陆面建模系统的模型数据融合方法集.

图3的MDF的概念图显示了主要的步骤,以及这些步骤的迭代属性.MDF包含一系列的程序以联合观测数据和模型,定量结合模型和数据的不确定性.MDF评估模型状态和参数,以及各自的不确定性.MDF需要考虑以下几点:(1) 随机和系统误差;(2) 模型和数据的初始状态;(3) 参数的验证和不确定性的繁殖;(4) 时空参数的变化性;(5) 模型缺陷和基准测试;(6) 参数优化或者模型调整导致的等效原则.R的MCMC包(马尔可夫链蒙特卡罗)、dlm包(卡尔曼滤波)、自编写的

2.3 陆面建模系统核心功能

陆面建模系统核心功能由脚本语言来实现,除了构建 R 和 NCL 功能函数以及 NCO 和 CDO 的数据处理之外,其部分核心功能由 Fortran 等程序语言组成的插件所实现,主要是目前的大部分地学模型(包含陆面过程模型)由 Fortran 代码编写.例如,利用 Fortran 语言在调用 NetCDF,HDF5 的库文件后,进行数据输入、输出的读写,子模块实现和模块之间功能的衔接等等.陆面建模系统的核心功能主要包括(如图 4 所示):(1) 运行控制,包括模型状态及初始状态、次网格变异性、运行时间统计以及子模块和功能模块的支持等;(2) 高性能计算,包括计算的操作模式,研究区域的分解,并行输入、输出,计算的负载均衡,以及故障容错等功能;(3) 时间管理,提供了时间管理的一些工具集、时间步长的设置、日历函数(在 CLM 陆面过程模型里,2 月统一为 28 天)、错误处理;(4) 输入、输出,使用通用的输入程序来读取输入文件,采用分布式存储输出文件,支持 NetCDF,GRIB,二进制等不同的格式输出;(5) 工具集,提供时空数据的插值、重投影、错误日志的记录、动态内存的管理、数据分析与管理、模型数据融合方法集、可视化、数理统计分析等工具及软件包^[15];(6) 用户定义,用户需要自定义研究区域、陆面过程模型、强迫数据及其相关参数.

陆面建模系统的核心功能构成了建模系统的工具库和方法库,还需要了解研究区范围,选择陆面过程模型、参数及其气象输入数据集(大气强迫),陆面建模系统需要一些通用的功能函数,也就是系统的中间层,主要由 4 个部分组成:(1) 范围插件;(2) 模型插件;(3) 参数插件;(4) 强迫插件.这些插件集成了参数、数据和模型的接口.在中间层的下一层是用户自定义的层,定义了研究区域的经纬度,模型的选择,土壤、叶面积指数等参数的定义,以及采用何种大气强迫数据集.

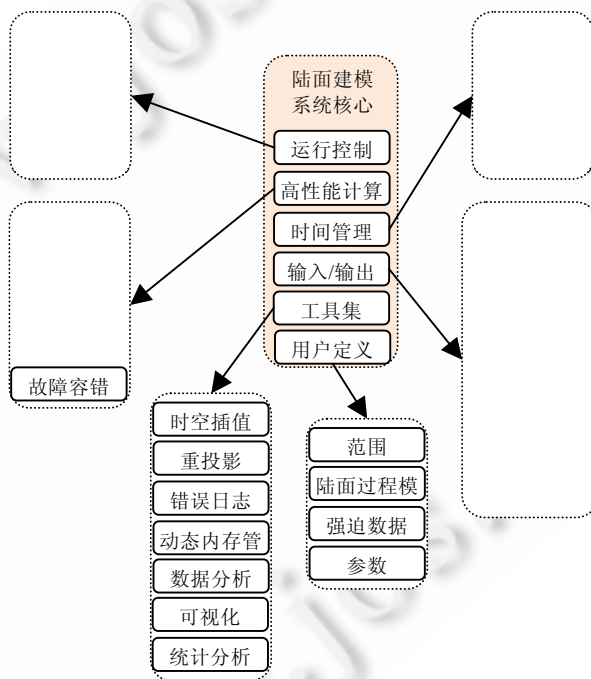


Fig.4 Main functions of land surface modeling system

图 4 陆面建模系统核心功能

3 陆面建模系统的应用研究

陆面建模系统将随时随地获取所需要的地学数据资源,利用陆面过程模型、软件工具和脚本语言,以及高性能计算环境解决陆面科学问题.以陆面过程模型模拟应用为例构建了陆面建模系统,并在此系统里开展了两

种陆面过程模型的模拟实验.

在陆面建模系统下,针对站点的模拟,采用 FLUXNET 数据集,模拟了全球 52 个通量观测站点的陆面过程,并采用 15 种数学统计方法,从不同的强迫场分辨率、不同的模拟输出分辨率、不同站点不同参数之间进行了全方位的模拟性能评估,并改善了模型的模拟性能,为模型的开发提供指导.在陆面建模系统下,针对区域面上的模拟,分别采用 NCEP 和 WRF 气候数值模型的动力降尺度输出作为陆面过程模型的强迫场对全中国和青藏高原区域进行了模拟分析.模拟结果的可视化因站点和区域的模拟而有所不同,它们都以 PostScripts 格式自动输出成一个文件以便于直接用于科学出版,科学可视化主要包括箱图、对比分析图、点线图、泰勒图、2D 平面图等(如图 5 和图 6 所示),而且模拟结果的统计值不仅在绘图上展示,而且以文本文件的形式输出,以便于后续的分析.

陆面建模系统采用各种统计分析方法和全自动的代码,将极大地改进现有模型的模拟计算方式,为模型的模拟应用提供可视化的技术手段,而且这种计算与模拟方式将为模型开发者和使用者带来便利.只有对模型的模拟进行全方位的评估与可视化才能在了解模型模拟性能的基础上改进模型(如图 7 所示).

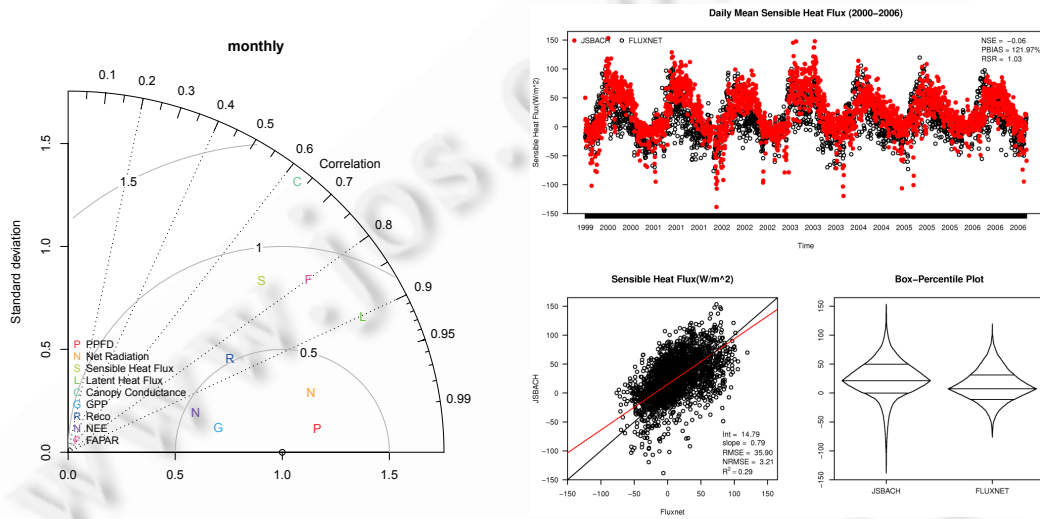


Fig.5 The visualization of site in land surface modeling system

图 5 陆面建模系统下的站点可视化输出

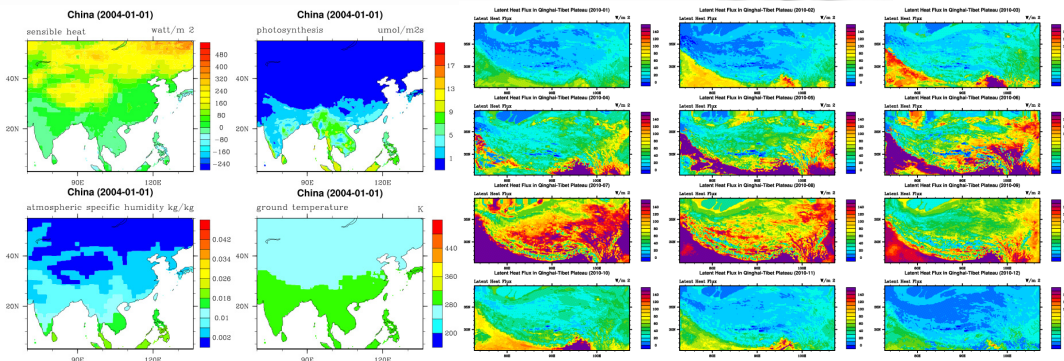


Fig.6 The visualization of region in land surface modeling system

图 6 陆面建模系统下的区域可视化输出

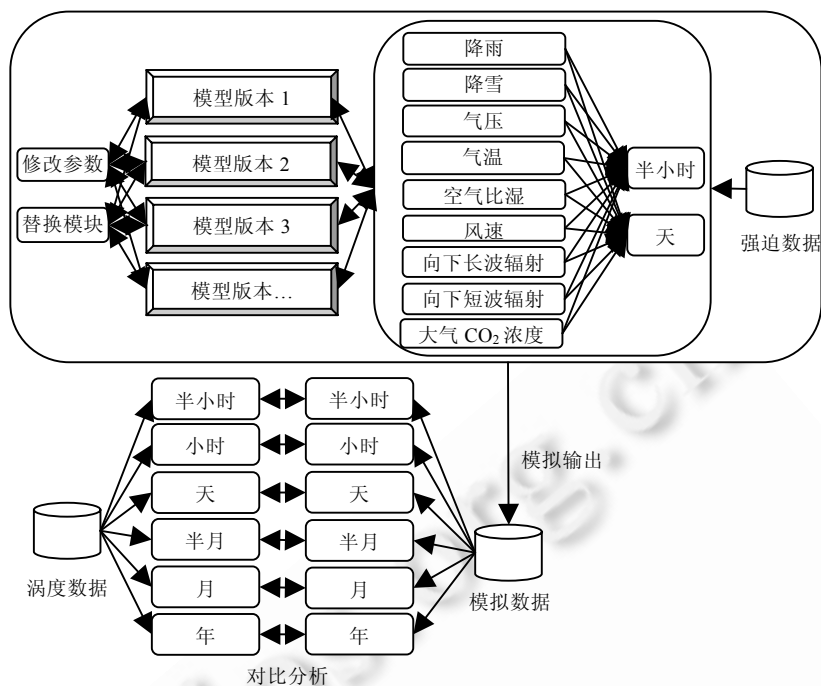


Fig.7 Comparison in different simulations

图 7 不同模拟下的对比分析

4 结束语

本文介绍了陆面建模系统的几个主要的研究内容,从陆面建模系统的结构入手,从高性能计算的访问模式及特点上开发了基于脚本语言的陆面建模系统.陆面建模系统为用户提供一套完整的从数据处理到模拟分析的陆面建模环境.这种环境需要联合不同来源的数据集、不同的陆面过程模型、通用的高性能计算的访问和计算模式、多种可视化方案和陆面过程模型模拟的工作流程等数据和技术手段来实现.并在此陆面建模系统内,针对站点和区域模拟这两种不同的陆面过程模型进行了应用示范,证明了基于脚本语言的建模系统在目前高性能计算环境中的应用潜力,以及自动化评估方法、不同可视化方案在陆面建模系统的作用.

本文虽然在陆面建模系统的开发及应用方面取得了一些进展,但是在研究中也发现了一些不足.总的说来,基于高性能计算环境的陆面建模系统才刚刚起步,在陆面过程模型、脚本语言、模型数据融合等方面的研究工作仍然在继续,这些研究工作都在不同程度上促进了陆面建模系统的发展.

未来有以下几个方面需要在今后的研究工作中进一步深入.陆面建模系统的系统设计还面临着诸多的难题,如模型数据融合方法的集成、不同来源数据和不同格式数据的转换与处理、区域模拟与站点的对比分析的误差、可视化需求的差异等,因此需要依靠多学科的知识融合和集成才能进一步地改进和完善陆面建模系统.目前只是采用了两种不同的陆面过程模型为示范构建了陆面建模系统,虽然对陆面过程模型的输入/输出、模拟分析等提高了较好的解决办法,但是在实际应用中会遇到更多的困难.因此,发展一个通用、快速、稳健的陆面建模系统对于促进陆面过程的研究是非常必要的.

致谢 本文撰写过程中得到了中国科学院寒区旱区环境与工程研究所赵国辉工程师、叶柏生研究员、周剑副研究员、潘小多副研究员、孙维君博士和庄艳丽博士的帮助,在此表示感谢.

References:

- [1] Hill C, DeLuca C, Balaji, Suarez M. The architecture of the earth system modeling framework. *Computing in Science & Engine*, 2004,6(1):18–28. [doi: 10.1109/MCISE.2004.1255817]
- [2] Peters-Lidard CD, Houser PR, Tian Y, Kumar SV, Geiger J, Olden S, Lighty L, Doty B, Dirmeyer P, Adams J, Mitchell K, Wood EF, Sheffield J. High-Performance earth system modeling with NASA/GSFC's land information system. *Innovations in Systems and Software Engineering*, 2007,3(3):157–165. [doi: 10.1007/s11334-007-0028-x]
- [3] Schmitza O, Karssenberga D, Deursenb WPA, Wesselingb CG. Linking external components to a spatio-temporal modelling framework: Coupling MODFLOW and PCRaster. *Environmental Modelling & Software*, 2009,24(9):1088–1099. [doi: 10.1016/j.envsoft.2009.02.018]
- [4] Grimme C, Papaspyrou A. Cooperative negotiation and scheduling of scientific workflows in the collaborative climate community data and processing grid. *Future Generation Computer Systems*, 2009,25(3):301–307. [doi: 10.1016/j.future.2008.05.002]
- [5] Turuncoglu UU, Murphy S, DeLuca C, Dalfes N. A scientific workflow environment for Earth system related studies. *Computers & Geosciences*, 2011,37(7):943–952. [doi: 10.1016/j.cageo.2010.11.013]
- [6] Kumar S, Peters-Lidard C, Tian Y, Houser PR, Geiger J, Olden S, Lighty L, Eastman JL, Doty B, Dirmeyer P. Land information system: An interoperable framework for high resolution land surface modeling. *Environmental modelling & Software*, 2006,21(10):1402–1415. [doi: 10.1016/j.envsoft.2005.07.004]
- [7] Holt TR. A high-resolution urban canopy/land-surface modeling system. *Ocean and Atmospheric Science Technology*, 2006, 1–3.
- [8] Kumar SV, Reichle RH, Peters-Lidard CD, Koster RD, Zhan X, Crow WT, Eylander JB, Houser PR. A land surface data assimilation framework using the land information system: Description and applications. *Advances in Water Resources*, 2008, 31(11):1419–1432. [doi: 10.1016/j.advwatres.2008.01.013]
- [9] Wood EF, Roundy JK, Troy TJ, van Beek LPH, Bierkens MFP, Blyth E, de Roo AA, Doll P, Ek M, Famiglietti J. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring earth's terrestrial water. *Water Resources Research*, 2011, 47(5). [doi: 10.1029/2010WR010090].
- [10] Sun S. Advance in land surface process study. *Bimonthly of Xinjiang Meteorology*, 2002,25(6):1–6 (in Chinese with English abstract).
- [11] Baldocchi D, Falge E, Gu L, Olson R, Holinger D, Running S, Anthoni P, Bernhofer C, Davis K, Evans R. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, 2001,82(11):2415–2434. [doi: 10.1175/1520-0477%282001%29082<2415%3AFANTTS>2.3.CO%3B2]
- [12] Michalakes J, Chen S, Dudhia J, Hart L, Klemp J, Middlecoff J, Skamarock W. Development of a next generation regional weather research and forecast model in developments in teracomputing. In: Zwiefelhofer W, Kreitz N, eds. *Proc. of the 9th ECMWF Workshop on the Use of High Performance Computing in Meteorology*. Singapore: World Scientific, 2001. 269–276.
- [13] Legates DR, Jr McCabe GJ. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 1999,35(1):233–241. [doi: 10.1029/1998WR900018]
- [14] Reichle RH, McLaughlin DB, Entekhabi D. Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, 2002,130(1):103–114. [doi: 10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2]
- [15] Balaji V. The flexible modeling system. *Earth System Modelling*, 2012,3:33–41.

附中中文参考文献:

- [10] 孙淑芬. 陆面过程研究的进展. *新疆气象*, 2002,25(6):1–6.



罗立辉(1980—),男,湖南常德人,博士,助理研究员,主要研究领域为陆面过程模型及地学 e-Science 应用研究.
E-mail: luolh@lzb.ac.cn



张耀南(1966—),男,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为基于数据、模型、计算、可视化的地学 e-Science 研究.
E-mail: yaonan@lzb.ac.cn