

A tightly coupled GIS and distributed hydrologic modeling framework



Gopal Bhatt*, Mukesh Kumar, Christopher J. Duffy

Department of Civil and Environmental Engineering, The Pennsylvania State University, 212 Sackett Building, University Park, PA 16802, USA

ARTICLE INFO

Article history:

Received 1 September 2013

Received in revised form

4 August 2014

Accepted 5 August 2014

Available online xxx

Keywords:

Distributed hydrologic model

GIS

PIHM

PIHMgis

Shared geodata model

ABSTRACT

Distributed, physics-based hydrologic models require spatially explicit specification of parameters related to climate, geology, land-cover, soil, and topography. Extracting these parameters from national geodatabases requires intensive data processing. Furthermore, mapping these parameters to model mesh elements necessitates development of data access tools that can handle both spatial and temporal datasets. This paper presents an open-source, platform independent, tightly coupled GIS and distributed hydrologic modeling framework, PIHMgis (www.pihm.psu.edu), to improve model-data integration. Tight coupling is achieved through the development of an integrated user interface with an underlying shared geodata model, which improves data flow between the PIHMgis data processing components. The capability and effectiveness of the PIHMgis framework in providing functionalities for watershed delineation, domain decomposition, parameter assignment, simulation, visualization and analyses, is demonstrated through prototyping of a model simulation. The framework and the approach are applicable for watersheds of varied sizes, and offer a template for future GIS-Model integration efforts.

© 2014 Elsevier Ltd. All rights reserved.

Software availability

PIHMgis

Developers: Gopal Bhatt, Mukesh Kumar

Contact Address: Department of Civil & Environmental Engineering, 212 Sackett Building, University Park, PA 16802, USA

Email: gopal.bhatt@psu.edu, mukesh.kumar@duke.edu

Year First Available: 2006

Hardware Required: Desktop/Laptop with 2 GHz CPU, 2 GB RAM or more

Operating System Required: Macintosh OSX 10.4 or newer; Windows XP or newer; Linux

Libraries Required: SUNDIALS, Qt4, GDAL, SQLite, GEOS, GSL, Expat, and PostgreSQL

Availability: <http://www.pihm.psu.edu>

Cost: Free

Source Code: <http://sourceforge.net/projects/pihmgis/>

Program Language: C/C++

1. Introduction

Physics based, fully coupled distributed hydrologic models such as FIHM/PIHM3D (Kumar et al., 2009a), InHM (VanderKwaak,

1999), MIKE-SHE (Abbott et al., 1986a,b; DHI, 2005), MODFLOW-SURFACT (McDonald and Harbaugh, 1988; Panday and Huyakorn, 2008), ParFlow (Kollet and Maxwell, 2006), Penn State Integrated Hydrologic Model (PIHM) (Qu and Duffy, 2007; Kumar 2009), PREVAH (Viviroli et al., 2009), WASH123D (Yeh and Huang, 2003), and WATFLOOD (Kouwen, 1988), simulate spatially explicit hydrologic response. These models also capture the heterogeneities in hydrogeologic and meteorological parameters within the watershed at finer spatial resolutions (Freeze and Harlan, 1969; Clark, 1998; Singh and Fiorentino, 1996; Entekhabi and Eagleson, 1989; Haverkamp et al., 2005; Pitman et al., 1990; Kollet and Maxwell, 2006; Kumar et al., 2009b), and have been demonstrated to enhance understanding and prediction of hydrologic processes (Refsgaard, 1997; Boyle et al., 2003; Meixner et al., 2003; Kirchner, 2006; Shrestha and Rode, 2008; Lu et al., 2009; Kumar et al., 2013). One of the key challenges in application of distributed, physics based models is the lack of a framework for efficient prototyping of model simulations, evaluation of a-priori parameters, and for simulation, analysis and visualization (Goodchild, 1992; DeVantier and Feldman, 1993; Nyerges, 1993; McDonnell, 1996; Moore, 1996; Sui and Maggio, 1999; Duffy et al., 2011).

Distributed hydrologic models require assignment of watershed parameters related to geology, soil, topography, land use, land cover, initial conditions, precipitation, and meteorological conditions, to all mesh elements within the model domain. Because of the inherent heterogeneity in geospatial data sets, the process of assigning parameters to large number of mesh elements is an error

* Corresponding author.

E-mail address: gopal.bhatt@psu.edu (G. Bhatt).

prone and a time intensive step. Traditionally, Geographic Information Systems (GISs) have been used for offline data processing, and post-simulation data analysis and visualization to support hydrologic models (DeVantier and Feldman, 1993; Maidment, 1996). However, faster prototyping and implementation of large-scale distributed hydrologic models necessitates more than pre-processing of geospatial data to generate model inputs. Distributed hydrologic models also require: generation of numerically efficient mesh network, definition of mesh network topology, derivation of watershed parameters from geospatial data, development of mapping functions for assigning both spatial and temporal data to mesh elements, and simulation, analyses and visualization of multiple hydrologic states. Satisfying these functions from within a GIS framework is challenging. The required effort is daunting, especially for GISs that are platform dependent with closed-source architectures and proprietary data structures, as they limit capabilities to add libraries and tools including those for transient watershed modeling.

Hence there is a strong motivation for coupling GISs and modeling systems to support and improve watershed research and decision-making (Maidment, 1993; Abel and Kilby, 1994; Kopp, 1996; Deckmyn et al., 1997). One of the several efforts in this context was from Smith and Maidment (1995) who developed Hydrologic Data Development System (HDDS), where ARC/INFO was used to generate parameters such as average sub-basin slopes, time of concentration, and SCS curve numbers, as inputs to the Texas Department of Transportation Hydraulic Program (THYSYS). Other examples include the Water Erosion Prediction Project (WEPP) interface on Geographic Resources Analysis Support System (GRASS) (Engel et al., 1993); the interface between ARC/INFO and the Hydrologic Engineering Center (HEC) modeling system (Hellweger and Maidment, 1999); the Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) by Environmental Protection Agency (EPA) (Lahlou et al., 1998) for Soil and Water Assessment Tool (SWAT) (Luzio Di et al., 2002); and ArcGIS-SWAT as an extension to ArcMap (Olivera et al., 2006). Several interfaces have also been developed using Map Window, a non-proprietary desktop GIS for Microsoft Windows operating system. These include a model setup interface to SWAT, Map Window SWAT (MWSWAT) (Leon, 2009), and version 4.0 of BASINS (Kinerson et al., 1999). In contrast to other attempts, development of Watershed Modeling System (WMS) (Nelson, 1997) was notable in the sense that it was a GIS developed with a dedicated purpose of conducting hydrological analyses.

In this paper, we present the development and structure of PIHMgis, an open-source, platform independent, extensible, tightly coupled GIS and distributed hydrologic modeling framework. PIHMgis establishes a formal strategy for prototyping a distributed hydrologic model simulation supported by national geospatial data products, simulation, analysis, and visualization of model results. Tight coupling between the GIS with the distributed hydrologic model was achieved by the development of an integrated user interface that uses a shared data model paradigm explained in Kumar et al. (2010). The underlying architecture that integrates the two systems and the procedural data flow between PIHMgis tools are also discussed. Advantages of the framework in terms of ease in prototyping of a physics based, fully coupled, distributed hydrologic model simulation using national geospatial data products is demonstrated using a case study application at Mahantango Creek Watershed (430 sq. km.) in Pennsylvania, USA. It is noted that the scope of this paper is limited to demonstrating the capability and effectiveness with which PIHMgis framework enables setup, operation, and analysis of PIHM simulations. The model results shown in this paper were neither calibrated nor validated against field observations. Readers are referred to Yu et al. (2013) for details

Table 1
Characteristics of different possible strategies for GIS and hydrologic model integration.

Characteristics	Coupling type		
	Loose	Tight	Embedded
Shared user interface	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Shared data and method base	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Intra-simulation model modification	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Intra-simulation query and control	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Advantages^a & disadvantages</i>	Distinct GIS and Hydrologic Modeling packages with independent interfaces Information sharing through exchange of files, which can be inefficient and error prone <i>Different tools and libraries facilitate independent development</i>	<i>Data exchange is automatic between GIS and Hydrologic Modeling System</i> <i>Merges different tools in a single powerful system</i> <i>Avoids inconsistency and data loss originating from redundancy and heterogeneity of method base</i>	<i>Steerable simulation in terms of possibility of changes in parameters or processes</i> Significantly complex programming and data management Changes to the code base are not easy because of its embedded large source code structure
Examples	Coroza, 1997; Storck et al., 1998; Hellweger and Maidment, 1999	Lahlou et al., 1998; Liu, 2005; Olivera et al., 2006;	Huang and Jiang, 2002; Kinerson et al., 1999

^a Italicized texts in the table indicate advantages.

about an automated calibration strategy for PIHM simulations. The approach and structure of PIHMgis, especially its procedural and architectural framework, and the integration of relevant software libraries needed for data access, domain decomposition, and model simulation and visualization of the results, can be used in similar model and GIS integration efforts. Other models can also be integrated within the framework, after making relevant model-specific modifications.

2. Model-GIS coupling strategy

Model-GIS integration can be achieved using one of the three potential coupling strategies outlined in Table 1 (Goodchild et al., 1992; Sui and Maggio, 1999; Brimicombe, 2003). Here, we use a tight coupling approach to integrate PIHM in Quantum GIS (QGIS, <http://www.qgis.org>) framework. Tight coupling requires development of a shared data model and data access methods so that both GIS and model functions are able to access the shared

geodatabase. Shared data model concept has been previously used in ArcHydro (Maidment, 2002) to process and develop geospatial data that can be henceforth used in hydraulic and hydrologic analyses. The tight coupling strategy has the advantage of independent development of both GIS and the hydrologic model, as is the characteristic of loosely coupled systems, while also allowing shared geospatial data access between the GIS and hydrologic model. The two systems are linked through object oriented tools that mimic many of the advantages of an embedded system, such as providing access to model parameters.

2.1. GIS framework

QGIS was used as the GIS platform to perform GIS-Model integration. The selection of QGIS was based on multiple factors including: its compatibility with multiple operating systems (Microsoft Windows, Macintosh, Linux etc.), an uncluttered user friendly interface, easy plugin-development functionality, support of wide range of data formats, an active developer community to support state-of-the-art software and hardware developments, and finally its open-source structure that allow function calls to models, data and numerical libraries. QGIS has been developed in C++ and extensively uses Qt (<http://trolltech.com/products/qt>) and Python libraries. It is comprised of four major subsystems: (1) input/capture; (2) management; (3) manipulation/analysis; and (4) output/display. The data input/capture subsystem, provides operational functions for reading, collecting, capturing, and querying geospatial data. The data management subsystem organizes and stores spatial data and their attributes to enable efficient query and quick retrieval for display, processing, and analysis. It also manages the modification and update of existing databases through editing tools. The manipulation and analysis subsystem executes the transformation of data from one form to another depending on model applications. The data display subsystem provides visual aids for quick interpretation of geospatial data in the form of diagrams, maps, or tables as well as data output providing access to the analyzed data in one of the several supported file formats.

The data management subsystem of QGIS offers an easy link point for integration with the hydrologic model through development of shared data structures that define both GIS feature objects and hydrologic features in the model. The data model for a GIS includes constructs for spatial data, topological data, and attribute data (Nyerges, 1987). In contrast, data structure of hydrologic features is determined by the topological relations between physiographic prototypes and coupling between processes. Open-source access to a QGIS facilitates the development and use of native QGIS classes and functions, which offer the interface and linkages necessary for supporting data structures of both GIS and hydrologic feature objects.

2.2. Hydrologic modeling framework

Penn State Integrated Hydrologic Model (PIHM) (Qu and Duffy, 2007; Kumar, 2009) is a physics-based, fully coupled, distributed hydrologic model that uses finite volume method (FVM) based discretization of the governing physical conservation equations and constitutive relationships on each control volume element. The FVM based formulation in PIHM guarantees mass balance locally and globally (Leveque, 2002). The local system of ordinary differential equations (ODEs) defined on each control volume are combined to form a global system of ODEs and solved altogether using CVODE (Cohen and Hindmarsh, 1994), a state of the art numerical solver in SUNDIALS (Hindmarsh et al., 2005). Fig. 1 shows discretization of a watershed (with triangular and rectangular control volume elements) in PIHM. Major hydrologic processes operating

on prismatic (triangular or rectangular) control volumes are illustrated. PIHM simulates transient dynamics of numerous hydrologic processes including channel routing, overland flow, infiltration, groundwater recharge, groundwater flow, interception, evapotranspiration, and snowmelt as a fully coupled system. Table 2 shows the requirements of geospatial data sets (such as state variables, hydrologic fluxes and hydrogeologic parameters) in simulating respective hydrologic processes in a PIHM simulation. The model has been previously applied at multiple scales by Qu (2004), Kumar (2009), Li and Duffy (2011), Bhatt (2012), Shi et al. (2013), Wang et al. (2013), Chen et al. (2014), Yu et al. (2014) etc.

2.3. A shared data model for distributed watershed modeling

To facilitate seamless translation of hydrogeologic and climatological information from GIS to hydrologic model, a shared data model that uses common ontology to describe both GIS and hydrologic model objects (Kumar et al., 2010) is used. The shared data model allows consideration of numerous geospatial feature objects, associated data attributes and time series datasets, e.g., precipitation, meteorology, and boundary conditions. It also provides a template for storing spatial-hydrographic and time series data in the form of a database that is accessible by the hydrologic model during simulation and GIS using access methods. Furthermore, because of its object-oriented design, the shared data model is extensible and can incorporate new data descriptors for physical, chemical or biological processes. For further details about the shared data model, readers are referred to Kumar et al. (2010). Using a collection of conceptual tools (see Section 3) for describing data, data relationships, data semantics, and consistency constraints, PIHMgis uses the shared data model to parameterize all mesh elements and to organize the simulation results for visualization and analyses. Consistency constraints of the data model ensure that certain minimum standards for compatibility are met. It is to be noted that the shared data model strategy can be used to couple other hydrologic models with the GIS. However such efforts will require model specific modifications in the shared data model, such as in definition of mesh topology, attribute information, and model input/output data structure.

3. PIHMgis user interface

PIHMgis user interface is integrated with the QGIS, as a plugin based toolbar (Fig. 2). PIHMgis tools support organization, development, and assimilation of spatial and temporal data into the hydrologic model, while the GIS provides tools for display, navigation and editing of geospatial data layers. Therefore, the hydrologic model becomes one of the analytical functions of the GIS. User interface of PIHMgis tools are accessible from the dropdown menus of the toolbar.

3.1. PIHMgis design: architectural and procedural details

PIHMgis tools incorporated in the shared model-GIS user interface accomplish three functions: Data Development, Hydrologic Model and Data Analysis (Fig. 3). Data Development includes functionalities necessary for a hydrologic model setup. The Hydrologic Model and Data Analysis components provide the ability to run hydrologic model simulations, and to perform spatio-temporal analyses and visualization of the simulated outputs, respectively. Bidirectional arrows in Fig. 3 illustrate the possible data flow (input and output) between the processing tools and the shared geodatabase. For example, Data Development functionalities when executed harvests the input data from the geodatabase

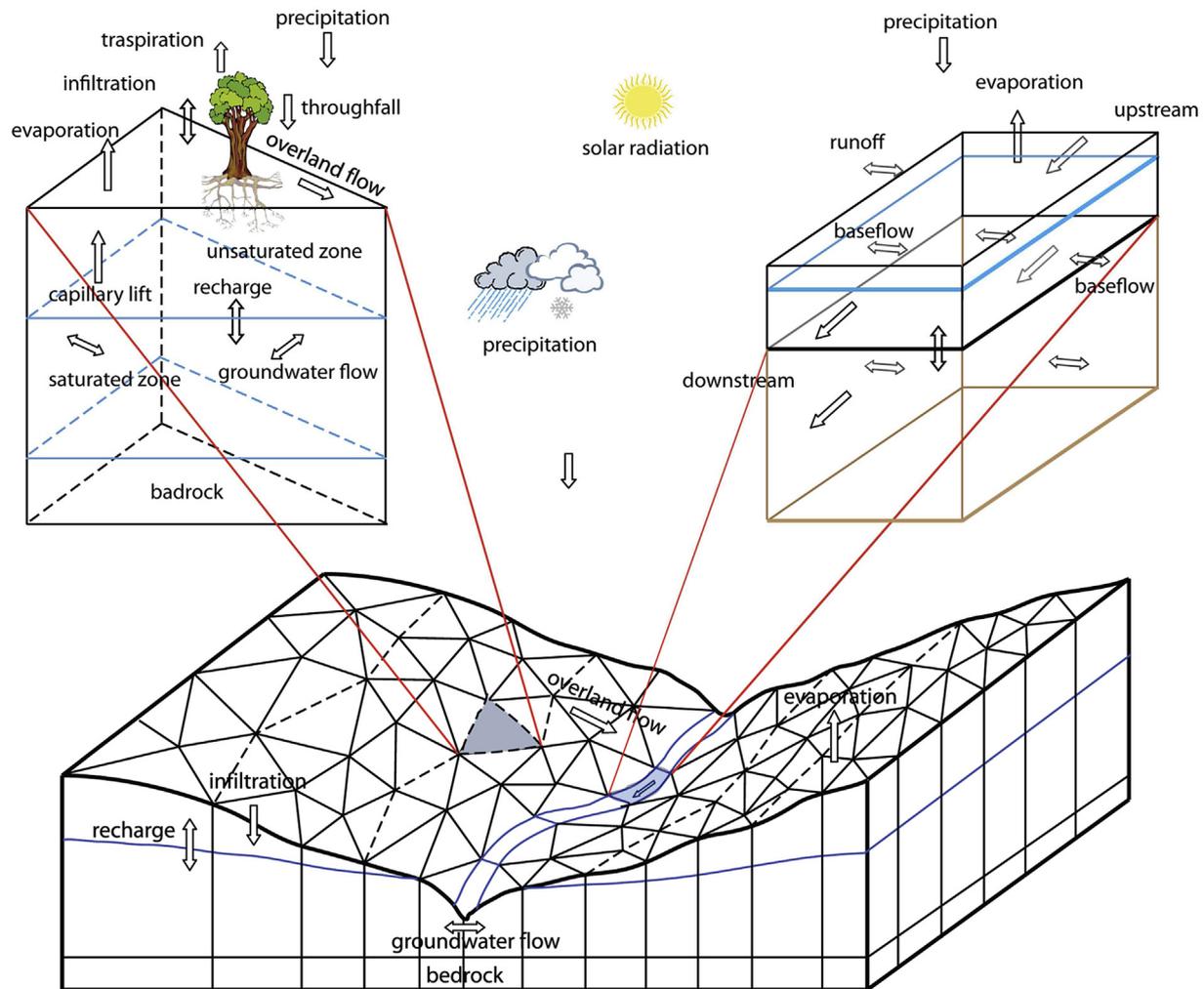


Fig. 1. Discretization of a watershed into triangular and linear mesh elements in PIHM. Interacting hydrologic processes are coupled within each control volume and between neighboring mesh elements.

using Data Access libraries and enriches the shared geodatabase back with processed data products. This is achieved by synthesizing a predefined structure of data flow between PIHMgis tools and the geodatabase. Overall the shared geodatabase makes the user interface less error prone, removes data redundancy, and enhances accuracy during data processing. This reduces complexity involved in model setup, operation, and analysis of simulation results. It is noted that subsystems of the integrated framework, data flow and the shared geodatabase strategy can be used in other similar model-GIS integration efforts, after incorporating model-specific modifications in definition of mesh topology, attribute information, and model input/output data structure.

PIHMgis functionalities detailed in Fig. 3 are organized as a sequence of procedural tools (Fig. 4) to (a) create and manage a model application; (b) perform raster processing steps on digital elevation data to delineate watershed; (c) prepare watershed polygons and other vector constraints for generation of quality irregular triangular mesh for the watershed; (d) perform domain discretization and assign topology information to each triangular and linear mesh elements; (e) estimate and map hydrogeologic and meteorological parameters to each mesh elements; (f) perform model simulation; and (g) analyze simulation outputs. Functional details of each PIHMgis toolsets are summarized in Table 3. More details of the procedural tools are provided in Section 4.

3.2. PIHMgis development

PIHMgis source code has been integrated within the QGIS source code (<http://download.qgis.org/qgis/src/>) using an internal plugin-based architecture. This involves calling a PIHMgis object and its dependent libraries from within the QGIS open-source code, which provides a template function for integration of new plugins. Development of PIHMgis tools includes implementation of many new geospatial data processing functions that can read and create geospatial data layers. These geospatial data processing functions use dependent libraries such as Geospatial Data Abstraction Library (GDAL) and OGR (<http://www.gdal.org/ogr/>) to access geospatial data in raster and vector formats respectively, and PostgreSQL (<http://www.postgresql.org>) to query and support relational databases.

The user interface of PIHMgis tools has been implemented in Qt (<http://trolltech.com/products/qt>), which is a standard toolkit for high performance, cross-platform graphical widget development. Qt uses an extended version of the C++ programming language, but has bindings for Python, Ruby, C, Perl, and Pascal. It supports development on all major platforms, and has extensive internationalization support. The graphic interface of PIHMgis tool is a Qt dialog object, again called from within the plugin template function. Multiple input/output objects are incorporated within a

Table 2
Summary of data requirements in each triangular mesh element in PIHM.

Data requirements	
Interception	State variables: initial interception storage Fluxes: precipitation, evaporation, throughfall Physical parameters: leaf area index (LAI), vegetation fraction, interception storage capacity
Snow melt	State variables: initial snow depth, snow density, snow cover temperature, snow water equivalent, soil temperature Fluxes: precipitation, melt rate Physical parameters: short wave radiation, long wave radiation, net solar radiation, air temperature, wind velocity, vapor pressure, melt factor
Infiltration/Exfiltration	State variables: surface water, soil moisture, groundwater storage Fluxes: precipitation, net precipitation Physical parameters: soil hydraulic conductivity, macropore hydraulic conductivity, macropore area fraction, maximum infiltration capacity, soil porosity, infiltration depth
Evapotranspiration	State variables: interception storage, snow storage on canopy and ground, soil moisture, soil temperature, groundwater storage Fluxes: precipitation, throughfall, net precipitation, infiltration Physical parameters: net solar radiation, air temperature, air density, wind velocity, relative humidity, vapor pressure, LAI, vegetation fraction, displacement height, resistance length, albedo, rooting depth
Surface flow	State variables: initial storage, surface water head in land and river mesh elements, boundary conditions Fluxes: net precipitation, evaporation, infiltration/exfiltration Physical parameters: nodal elevations, roughness coefficient, discharge coefficient, vegetation fraction
Unsaturated	State variables: initial soil moisture, groundwater storage, boundary conditions Fluxes: infiltration, evaporation, transpiration, recharge Physical parameters: soil hydraulic conductivity, macropore hydraulic conductivity, macropore area fraction, van Genuchten soil hydraulic parameters, soil porosity
Groundwater	State variables: initial groundwater storage, groundwater head in the neighboring triangular mesh elements, boundary conditions Fluxes: recharge, evaporation, transpiration Physical parameters: hydrologic conductivity, porosity, nodal elevations, regolith thickness, van Genuchten hydraulic parameters, LAI, vegetation fraction
Stream flow	State variables: initial stream water storage, stream water storage upstream and downstream, surface and groundwater storage in neighboring triangular mesh elements, boundary conditions Fluxes: inflows from upstream segments, outflow to downstream segment, surface runoff from banks, baseflow from adjacent aquifers, seepage through stream bed. Physical parameters: elevation of the river nodes, depth and width of channel, roughness, hydraulic conductivity

PIHMgis dialog along with push button objects. Depending on the objectives and functionalities extended by a dialog interface, necessary *slots* are created and attached to objects. To complete the integration of a new slot, a unique combination of *signal* and *action* is linked. Operationally a slot executes the associated action (e.g., data processing function) when an acceptable signal (e.g., click,

toggle) is received from the interface. Functions for handling and processing of geospatial data and for access and manipulation of data files have been implemented in C and C++. The hydrologic modeling capability of PIHM, which was developed in C, is also available as an action from the PIHMgis user interface. It is noted that the modularity in development of PIHMgis through the use of plugin architecture allows flexibility to extend core QGIS applications, while maintaining independent development paths. New tools such as a different hydrologic model or support for new data formats, hotlink handlers, and data editors can be added to the PIHMgis framework without the need to modify the core QGIS source code.

4. PIHMgis application: a demonstration of PIHMgis tools

To demonstrate the capability and effectiveness of the PIHMgis framework, we prototype a new model simulation for a meso-scale Mahantango Creek watershed in the Susquehanna River Basin, USA. The watershed has a surface drainage area of approximately 430 square-kilometers, and partially spans the Dauphin, Northumberland, and Schuylkill counties in Pennsylvania, USA. The topographic elevation in the watershed ranges from 113 to 552 m. Geology of the watershed is predominantly Munch Chunk, Catskill, Trimmers Rock, Pottsville, and Pocono formations. There are 13 different standard land cover types present across the watershed. Dominant land cover types in the watershed include deciduous forest (53%), pasture/hay (16%), cultivated crops (14%), and developed/open-space (6.5%). The national data products that were used in prototyping of the distributed model simulation include: (a) USGS seamless 30 m resolution digital elevation data; (b) Soil classification map and related tabular texture data from Soil Survey Geographic Database 2 (SSURGO); (c) National Land Cover Database (NLCD) land cover map; and (d) precipitation and meteorological parameters from Phase-2 North America Data Assimilation System (NLDAS-2). Readers are referred to the HydroTerre (Leonard and Duffy, 2013) website (www.hydroterre.psu.edu) for access to basic geospatial data for USA. Next five subsections provide detailed descriptions of the PIHMgis tools. From here onwards use of italics indicate a reference to PIHMgis tool or toolset.

4.1. Raster processing

Raster Processing toolset facilitates subwatershed delineation and stream definition. These vector layers are critical for generation of conformed triangular irregular mesh. Generation of sub-watershed boundaries and streams is achieved through seven procedural steps (Fig. 4). The *Fill-Pits* tool performs hydrologic conditioning of DEM by removing sinks. The *Flow Direction Grid* tool calculates direction of steepest descent by evaluating slope at each grid cell with respect to its eight (maximum possible) neighboring grid cells using eight-direction (D8) algorithm (O'Callaghan and Mark, 1984; Jenson and Domingue, 1988). The *Flow Accumulation Grid* tool computes the cumulative number of grid cells draining to a grid cell based on the flow direction grid. Stream definition threshold is applied to the flow accumulation grid to compute stream grids using the *Stream Grid* tool. A threshold, indicating the least number of grid cells draining to a grid cell for it to be classified as a stream grid cell, provides ability to control the resolution of delineated stream network. The *Link Grid* tool divides the stream network into reaches by identifying confluence points. The *Catchment Grid* tool partitions the watershed based on the stream reaches to which its grid cells are draining. The stream grids and the catchment grids are converted to vector data types using the *Stream Polyline* and the *Catchment Polygon* tools to enable further processing. Fig. 5 shows the DEM of the Mahantango Creek watershed,

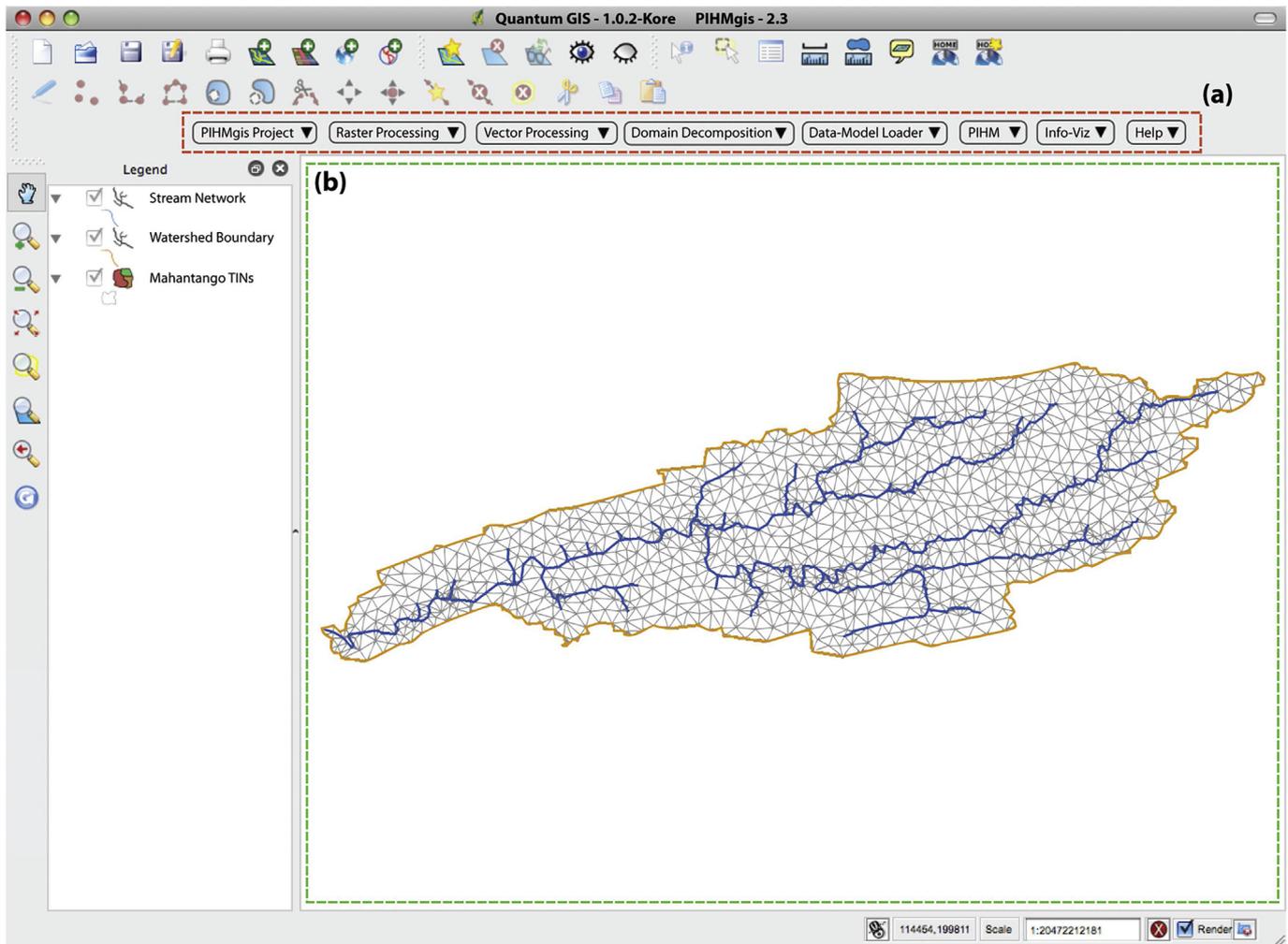


Fig. 2. The PIHMgis user interface within QGIS. (a) PIHMgis toolbar includes data processing and modeling system components. (b) The data frame displays raster/vector geospatial datasets. Shown in the data frame are stream network, watershed boundary, and triangular irregular mesh for Mahantango Creek watershed.

which is the only input required in *Raster Processing* apart from the streams definition threshold. Fig. 5 shows the catchment boundary and stream network obtained from digital elevation dataset using *Raster Processing* toolset. Here, the stream grids were generated using 4000 grid cells (equivalent to a drainage area of 3.6 sq. km.) as the threshold. *RunAll* combines all of the procedural data processing steps of *Raster Processing* into a single tool.

4.2. Vector processing

Vector Processing provides tools to condition and consolidate multiple hydrologic features into one vector layer, so that they can be used for generating quality triangular mesh. Consolidation includes merging of watershed boundary, stream network, and other relevant vector objects such as nodes (stream gauge, groundwater observation well locations), polygons (soil classification, land use, political boundaries) and polylines (artificial channels or drainage networks). Data processing steps involved in the *Vector Processing* are illustrated in Fig. 4. The *Dissolve Polygon* is an optional data processing tool that may be used to remove the internal (e.g., subwatershed) boundaries, if needed. The *Polygon to Polyline* tool converts the input polygon watershed features, which are polygon vector objects, to their polyline equivalents. This step is needed because a simplification operation only works on watershed

features that are polylines. The *Polygon to Polyline* tool preserves the external boundaries of polygons as polyline equivalents. The *Simplify Line* tool removes unwanted irregularities from a polyline by removing unnecessary nodes (Douglas and Peucker, 1973) based on specified simplification tolerance length as shown in Fig. 6a. The reduction in number of nodes for representation of polylines decreases the number of generated mesh elements around it. It is to be noted that simplification tolerance is selected such that semblance of the constraining polyline is maintained. For illustration purposes a qualitative comparison between unsimplified and simplified watershed features has been shown in Fig. 6b. The illustration shows that the aliasing effects that are introduced in watershed boundary and stream features during the *Raster Processing* are overcome by simplification. This results in significant decrease in the number of triangular elements in generated mesh as compared to one generated using unsimplified watershed features (Fig. 6b). Next, polylines are split into lines by separating them at the vertices through the *Split Line* tool. The *Vector Merge* tool combines all feature objects into a single vector layer. Fig. 7 shows simplified and consolidated output from *Vector Processing* for the Mahantango Creek watershed. Here, only the external catchment boundary and stream network were considered as constraints. Streams and watershed boundary were preferentially simplified using a tolerance length of 50 m and 100 m respectively.

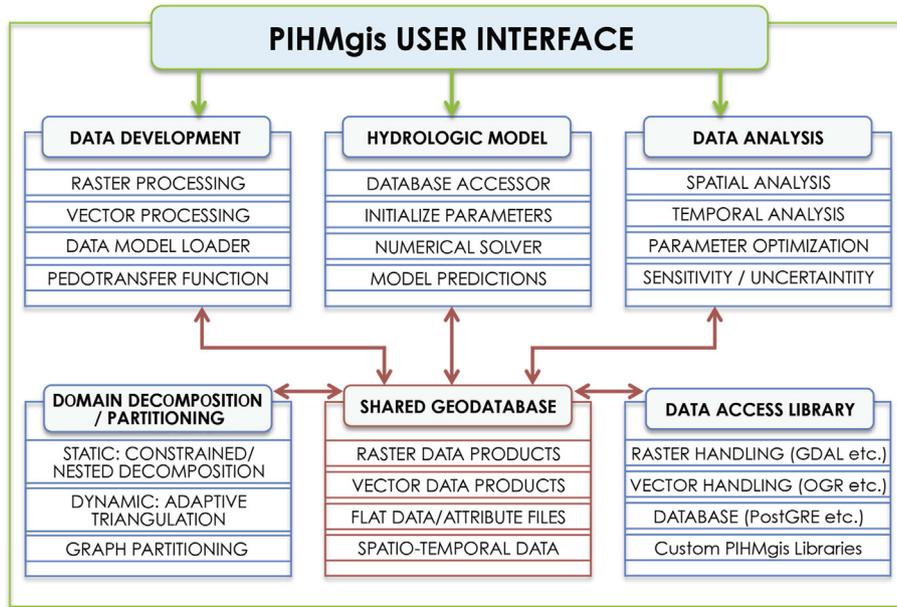


Fig. 3. The architectural framework of PIHMgis. The graphic shows key subsystems and processing components in PIHMgis. Arrows depict data flow directions between different subsystems.

Simplification reduced the number of nodes used in the representation of stream and watershed boundaries from 2945 to 492 nodes (~83% reduction) and 2937 to 177 nodes (~93% reduction) respectively.

4.3. Domain decomposition

Model simulation time is generally proportional to the number of triangular mesh elements used in representing a watershed.

Therefore, quality spatial decomposition of the model domain, which sufficiently captures important hydrologic features and hydrogeologic data while also ensuring a minimum number of quality mesh elements (Kumar et al., 2009b), is important for attaining an efficient simulation. In the case of irregular triangular mesh, quality is determined by generation of well-rounded (or not skinny) triangular elements that also conform to hydrologic features. These features include internal boundaries such as lakes or wetlands, external boundaries such as watershed boundary,

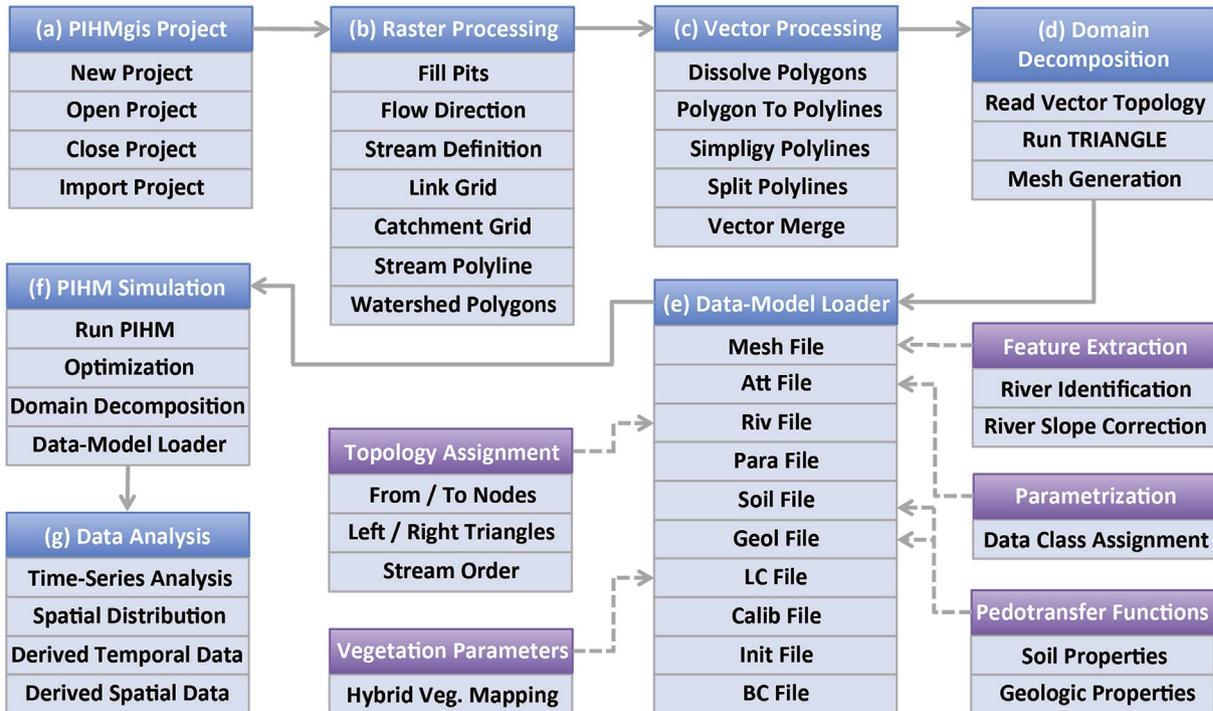


Fig. 4. Procedural framework of PIHMgis. Blue boxes identify key functionalities, while purple boxes identify support functions. Arrows show the default sequence of execution of PIHMgis tools. Dashed arrows show the associations between PIHMgis tools with support functions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
PIHMgis toolsets and summary of their data processing functions.

PIHMgis toolsets	Data processing function
(a) PIHMgis project	Native PIHMgis project management handles input/output options in response to selections made in the preceding step.
(b) Raster (DEM) processing	A set of seven tools that is used for watershed delineation and stream definition from gridded topographic data.
(c) Vector Processing	A set of five tools that provides flexibility to consolidate multiple geospatial layers (e.g., streams, watershed boundary) in form of line or point feature type. These tools aid in generation of constraints for domain decomposition.
(d) Domain decomposition	A set of three tools to produce quality triangular mesh network.
(e) Data model loader	A set of ten tools for automated assignment of topology information and for mapping spatio-temporal watershed and climatological properties to mesh elements and river reaches.
(f) Run PIHM	This tool provides ability to perform concurrent execution of multiple instances of PIHM simulations.
(g) Data analysis	Data analysis toolset provides the ability to visualize simulated outputs for a user-selected group of triangular or linear mesh element over the simulation period.
(h) Help	Provides information regarding the objective of each PIHMgis tool, input and output datasets, and data processing algorithms.

constraining points such as weirs or monitoring stations and constraining lines such as rivers, man made drainage systems, and faults.

PIHMgis provides a set of three tools under *Domain Decomposition* to divide the watershed into triangular mesh elements. The *Read Shape Topology* tool transforms the consolidated vector file created by *Vector Processing* toolset into Planar Straight Line Graphs (PSLGs). The PSLG file contains the coordinate of nodes and topological information of the lines connecting two nodes. The *Run TRIANGLE* tool uses the PSLGs and selected quality options to

generate triangular mesh using TRIANGLE (Shewchuk, 1996). TRIANGLE provides options to decompose the domain into quality constrained or conformed Delaunay triangulations (Ruppert, 1995). The *Run TRIANGLE* tool provides options for quality controls, such as minimum angle and maximum area of the triangular elements, to strike a balance between accuracy in representation of geospatial data and computation demand. Statistics regarding quality of the triangular elements are also generated to facilitate fine-tuning of the spatial decomposition. Geometric and topology information of the mesh is stored in four separate output files (poly, element, node, and neighbor files). *TIN Generation* tool uses these four files to create a vector (shapefile) of the mesh. The shapefile can be loaded in the data frame and visually inspected for quality. Fig. 8 shows a quality mesh of Mahantango Creek watershed generated by *Domain Decomposition* toolset. The mesh contains 2606 triangular elements and 509 linear river elements. A 20-degree minimum angle and 400,000 sq. meters as maximum area quality constraints were used. PIHMgis is also capable of generating fairly complex spatial decompositions of a watershed such as nested meshes (Kumar et al., 2009a,b), where the spatial resolution of mesh elements is purposely varied in different regions of the watershed.

4.4. Data model loader

Data Model Loader toolset performs automated assignment of spatial and temporal watershed properties such as climate forcing, soil hydrologic parameters, land cover, geology, boundary conditions, surface and subsurface topography to mesh elements and its nodes based on their geographic location. Fig. 9a shows conceptualization of how raster data layers and associated properties are mapped to individual triangular mesh elements. *Generate Att* tool assigns spatial classification information associated with various hydrogeologic layers to individual triangular elements (Fig. 9b). Hydrogeologic parameters describing the properties of classes are estimated using pedo-function subroutines. *Generate Soil* and *Generate Geol* tools use pedo-transfer functions (Wosten et al., 1999) to estimate hydrologic parameters such as saturated

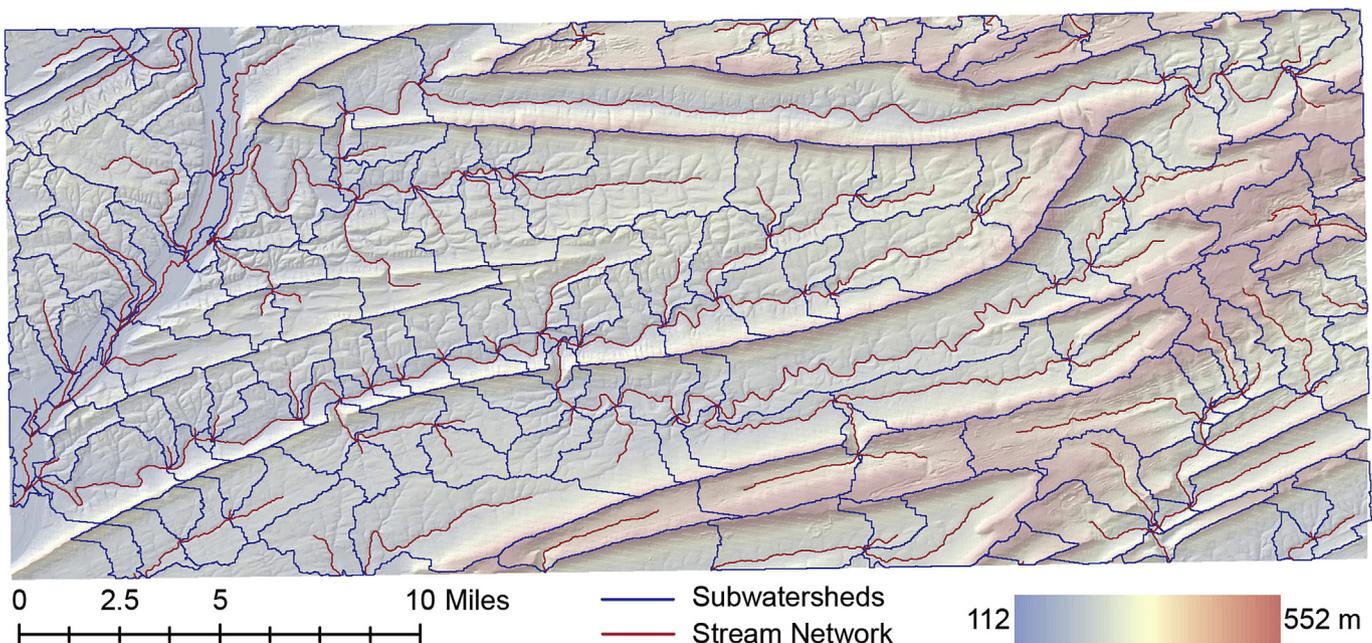


Fig. 5. The 30-m USGS DEM encompassing Mahantango Creek watershed. Subwatershed boundaries and stream network were delineated using the *Raster Processing* toolset. A stream definition threshold of 4000 grid cells (equivalent to a drainage area of 3.6 sq. km.) was used.

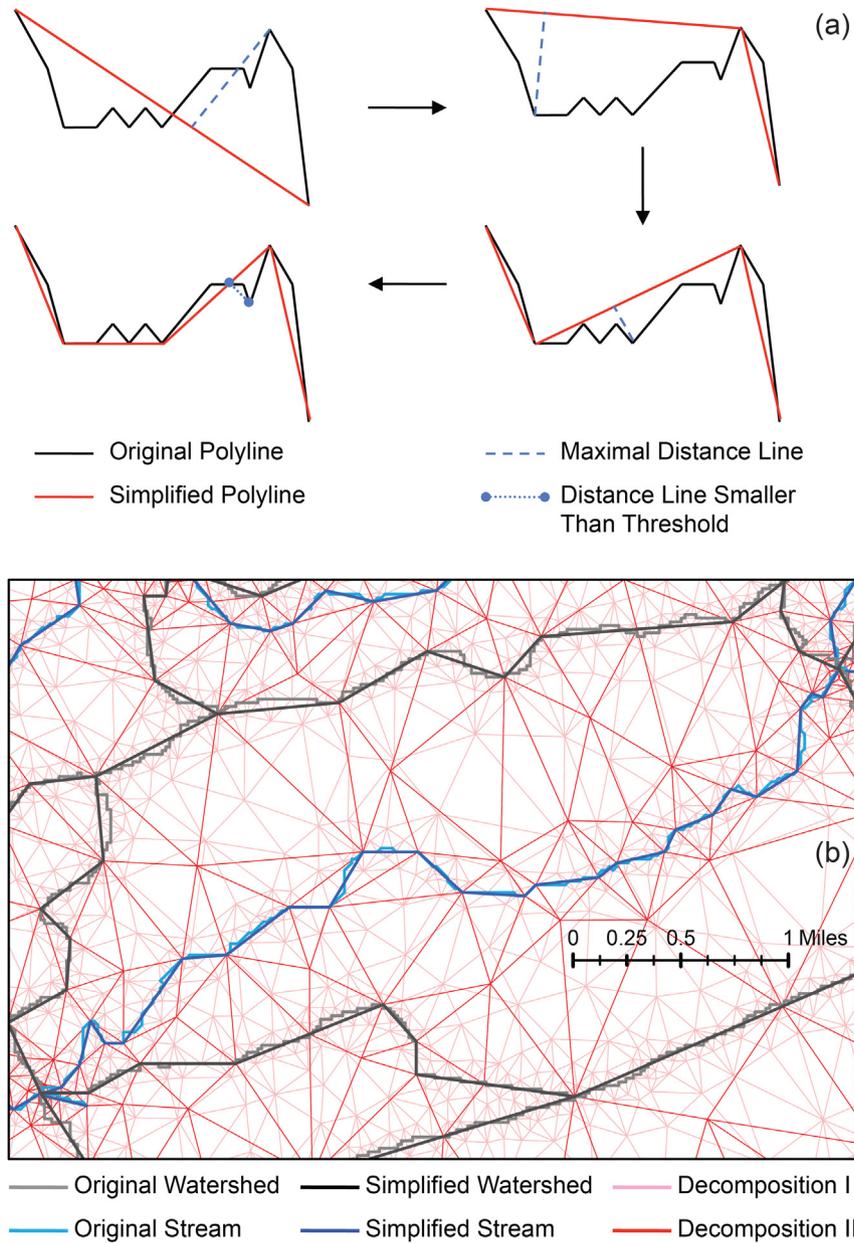


Fig. 6. (a) The schematic depiction of intermediate steps involved in the simplification of a polyline in PIHMgis. (b) Comparison between domain decompositions using (i) original and (ii) simplified subwatershed boundaries and stream reaches. The simplification algorithm aids in reducing the number of triangular elements in the mesh by removing unnecessary nodes, while preserving the semblance of constraining layers.

hydraulic conductivity, porosity, residual porosity and van Genuchten parameters based on available texture properties (sand, silt, clay, organic matter, bulk density, and organic content) obtained from SSURGO database (websoilsurvey.nrcs.usda.gov). The *Generate LC* tool estimates monthly vegetation parameters such as leaf area index, roughness length, stomata resistance, albedo, and vegetation fraction for the NLDAS classes using functions that map NLDAS major land cover parameters (NLDAS Vegetation Parameters, 2013) based on percent land cover description. The *Generate Mesh* tool maps the digital terrain and regolith elevations to the surface and subsurface nodes of the prismatic control volumes (vertical projection of the triangular mesh elements). Topology information of individual triangular mesh elements is generated as well by identifying shared nodes, edges, and neighboring elements. The *Generate Riv* tool establishes the topologic

information of individual stream segments and their association with the triangular mesh elements as shown in Fig. 9c. Topology definition includes identification of To-Node, From-Node, Downstream Segment, Left-Triangle, and Right-Triangle for each river element. The *Generate Para* tool assigns model operation and solver parameters.

4.5. PIHM simulation

PIHMgis provides the ability to run multiple instances of PIHM simulations using the user interface. The *RunPIHM* tool interacts directly with the shared geodatabase, to retrieve data required by PIHM simulations. PIHMgis also provides the ability to calibrate the model against observed data. This is achieved by uniformly adjusting the key hydrologic parameters globally over the entire

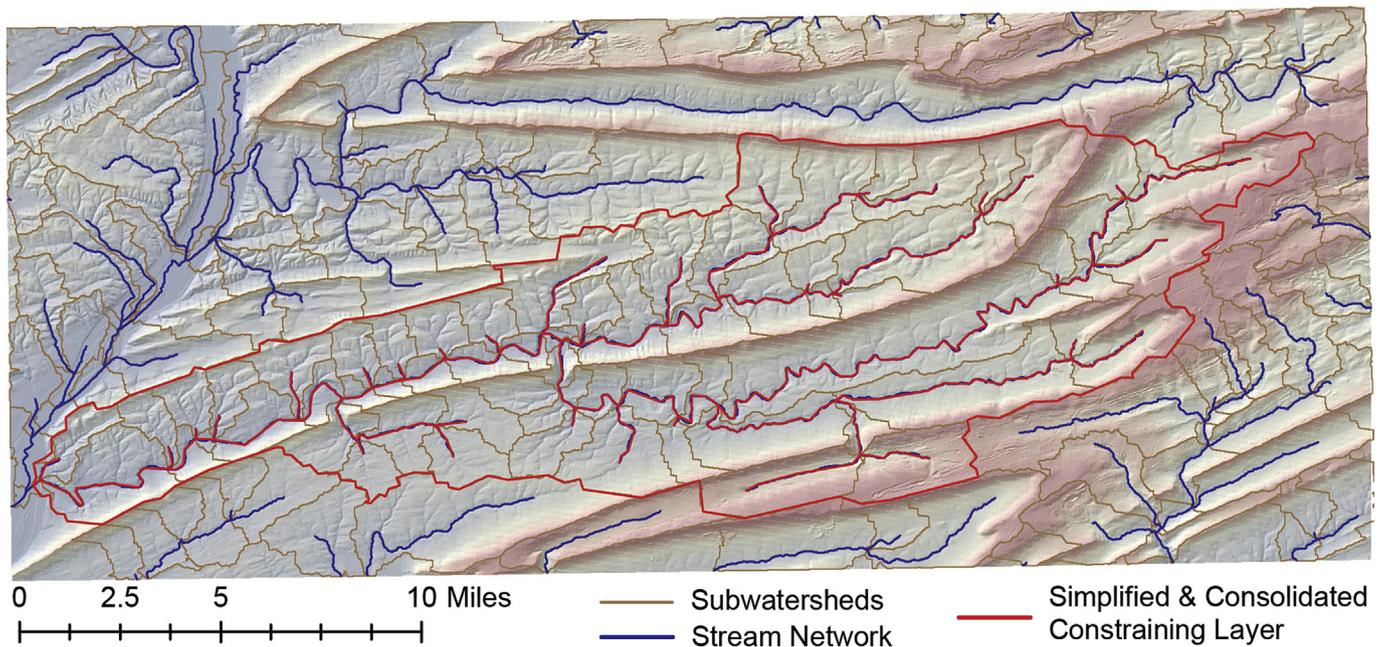


Fig. 7. A consolidated vector layer (a line feature type) is generated using the *Vector Processing* toolset. Streams and watershed boundary were preferentially simplified using a tolerance length of 50 m and 100 m respectively. The number of nodes was reduced from 2945 to 492 (83% reduction). In watershed boundaries, number of nodes reduced from 2937 to 177 (93% reduction).

watershed through the *Generate Calib* tool. The calibration strategy for PIHM is presented in detail in [Bhatt \(2012\)](#) and [Yu et al. \(2013\)](#). The methodology preserves a-priori spatial variability of parameters. Simulated hydrologic states and fluxes from model simulations are archived in the shared geodatabase. PIHM also supports Network Common Data Form (NetCDF) data format for storing simulation outputs, which provides machine-independent easy access and sharing, and also saves storage space as compared to traditional text files. Model operation parameters control the output variables and their temporal resolution at which they are archived.

4.6. Analysis and visualization

PIHMgis provides visualization and analysis of model outputs with the help of several basic GIS display functionalities provided by QGIS. The *Data Analysis* toolset provides ability to analyze model outputs, by building query to aggregate or average hydrologic states or fluxes, at a particular triangular or linear mesh element or group of elements. The *Time Series* tool allows visualization of temporal variability of a predicted states/fluxes in any mesh element. Aggregation and averaging can also be performed on the time series data for a specific region of interest. [Fig. 10](#) shows a representative time series plots of some of the simulated fluxes over a one year period. The graphic demonstrates the ability of the tool in providing estimates (as spatial aggregates) over a specific region for either streams or watershed mesh elements. The *Spatial Plot* tool on other hand provides ability to visualize spatial variability of fluxes/states within the watershed. The tool generates spatial plots of states and fluxes as shapefiles that can be visualized in any GIS including QGIS. Data representing simulated hydrologic states/fluxes (e.g., groundwater storage across the watershed or baseflow in the stream network) are stored as attributes in the shapefiles. The *Spatial Plot* tool also provides ability to aggregate data over a specified time interval. [Fig. 11](#) shows representative spatial maps of groundwater recharge during the selected year. The groundwater recharge rates simulated by the model were averaged at each

triangular mesh element over four separate seasonal intervals as shown in the figure to examine seasonal variations in response of the watershed.

4.7. Timing a representative simulation using PIHMgis tools

[Table 4](#) shows the processing time required by PIHMgis tools in prototyping a PIHM application for the Mahantango Creek watershed. Before the development of PIHMgis tools, a manual setup, simulation, and analysis of PIHM required operation of multiple software tools to accomplish data processing tasks. Because of the broken data flow between the processing steps, this used to make the process cumbersome, error prone and time consuming. As a result, manual setup typically required number of hours to days in prototyping a PIHM simulation. Platform dependence of many of the softwares used in this process posed additional challenges. In contrast, significantly small amount of time required by the PIHMgis framework in prototyping a distributed model simulation in a new watershed demonstrates the efficiency of PIHMgis framework as compared to the offline manual setup. It is noted that the efficiency is primarily gained through existence of simulation and data management, analyses, visualization and processing tools, within a same framework. Native project management also plays a critical role by reducing the delay in transition from one tool to the next, by auto-filling of input/output file names and directory paths in response to selections made in preceding steps. It took only 3.7 min to setup a new model simulation in Mahantango Creek watershed ([Table 4](#)), while execution and analysis took 147 min and 15.4 min respectively. PIHMgis reduces the model setup time to minutes, thus allowing users to focus on core hydrologic modeling goals of answering science and management challenges.

5. Discussions and conclusions

In this paper, we present a tightly coupled GIS and distributed hydrologic model framework called PIHMgis. The framework is an open-source, platform independent, extensible (due to pluggable

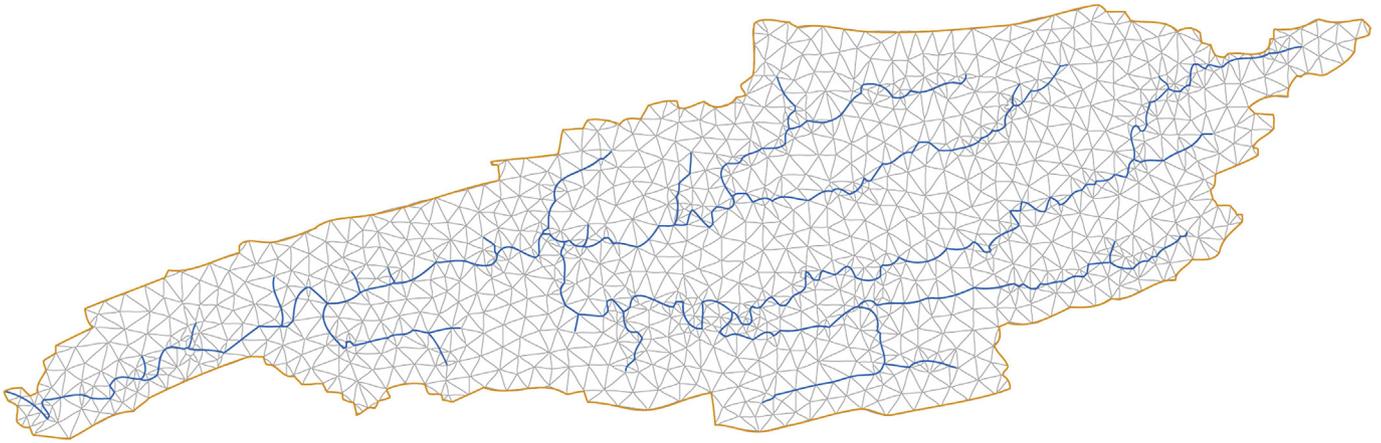
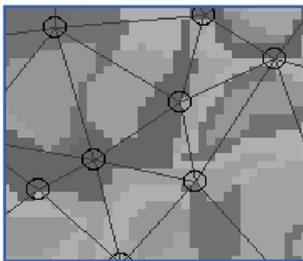
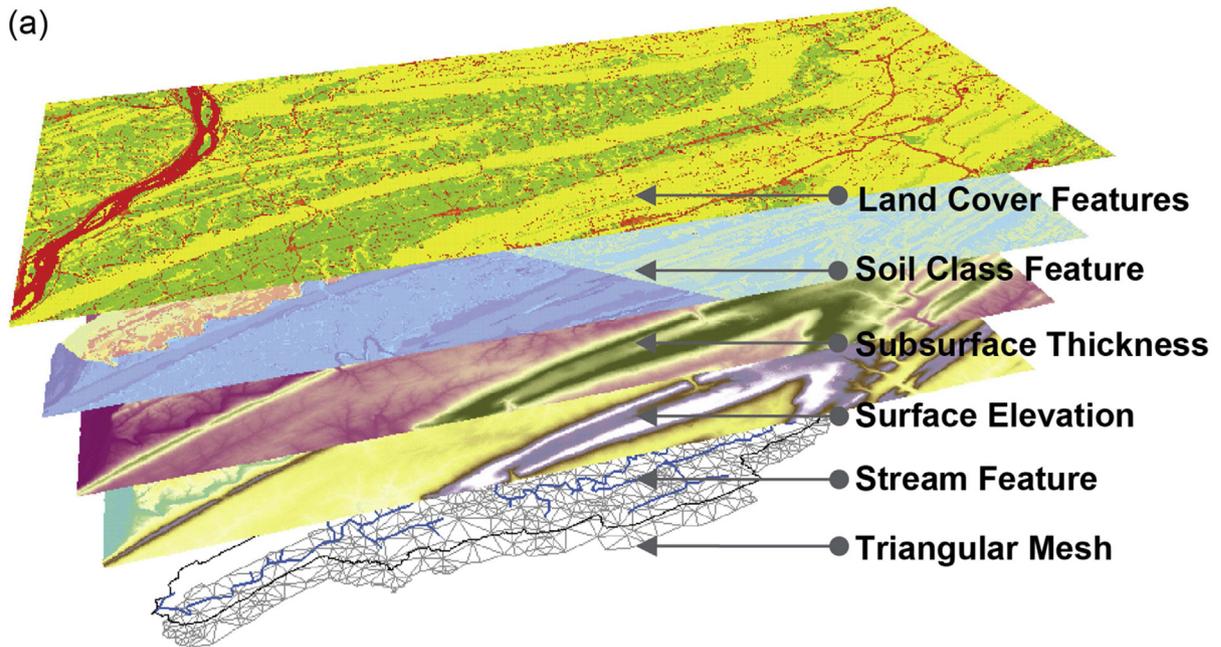
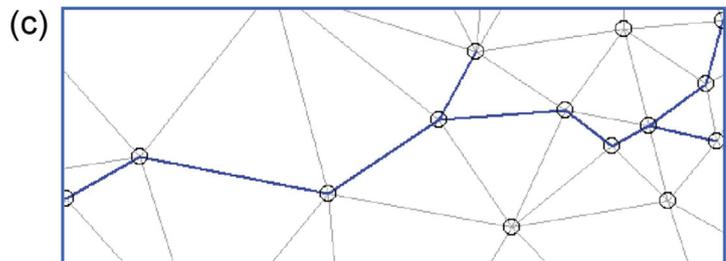


Fig. 8. Domain decomposition (conforming Delaunay triangulation) of Mohantango Creek watershed into 2606 triangular mesh elements and 509 linear stream elements using *Domain Decomposition* toolset. Quality constraint of 20° for minimum angle of a triangle, and 400,000 square meters for maximum area were used.



(b)

Data Parameterization: assigns classification of each geospatial data layer to triangular mesh elements.



(c)

Topology for stream segments: established by From- and To- Node, Up- and Down- stream segment, Left- and Right- mesh element.

Fig. 9. (a) The schematic illustrates the process of assigning hydrogeologic parameters to individual triangular mesh elements using *Data Model Loader* toolset. (b) Classifications corresponding to geospatial data layers are either assigned to the nodes or the centroid of the mesh elements. (c) Topology of the stream segments is defined by establishing relations between a stream reach and its neighbors.

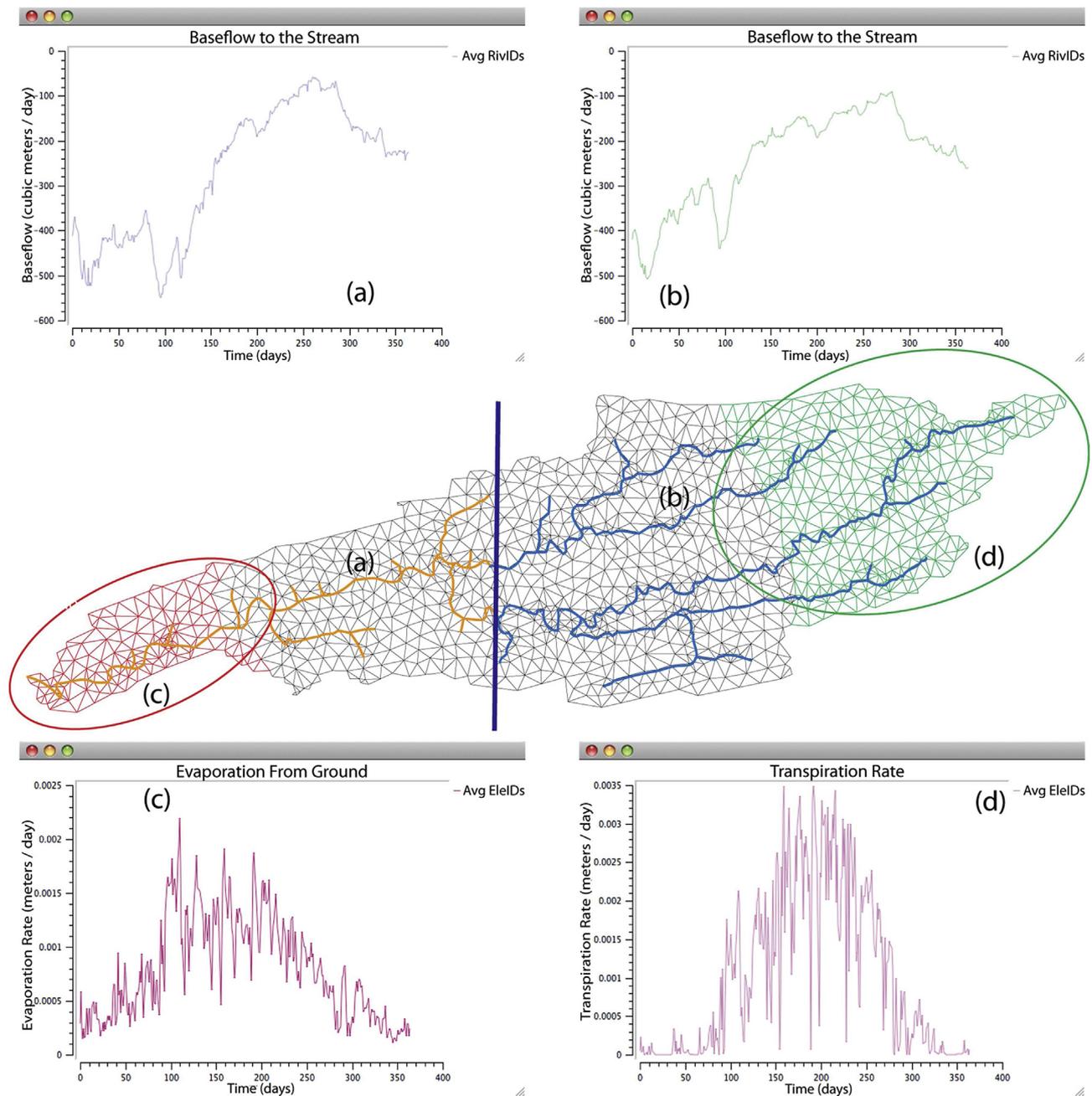


Fig. 10. An example of ‘Time Series’ analysis plot showing fluxes from a yearlong PIHM simulation in Mahantango Creek watershed. (a) Net groundwater contribution to the stream segments from the left half of the watershed (left of vertical line). (b) Net groundwater contribution to the stream segments from the right half of the watershed (right of vertical line). Negative sign in the baseflow rates imply flow to the river. (c) Rate of ground evaporation from a downstream region (encircled) near the outlet of the watershed. (d) Rate of transpiration from the upland headwater region (encircled) of the watershed.

architecture), and tightly coupled GIS interface to the PIHM. PIHMgis incorporates multiple vector and raster geospatial data layers to decompose the model domain at appropriate spatial resolution, while ensuring a desirable balance between accuracy of data representation and number of triangular mesh elements. The framework allows sophisticated distributed hydrologic modeling capabilities within an open-source GIS. The core of the framework is an object oriented shared data model, which provides flexibility to describe both GIS and hydrologic model objects and can incorporate new data descriptors (Kumar et al., 2010). The shared geodatabase based architecture (Fig. 3) used in the development facilitates minimum data redundancy, reduced storage

requirements, and improves hydrogeologic data retrievability from within different components of the PIHMgis framework.

The framework also establishes a formal strategy for integration of distributed hydrologic modeling, analysis, and visualization system, supported by national data products. Organization of tools in a procedural framework streamlines distributed hydrologic modeling operations, while also allowing use of PIHMgis components such as *Raster Processing* (as in *TauDEM* and *ArchHydro*) as standalone softwares. The framework can be used to set up a new model simulation in any watershed in U.S., using national geodatabases, in matter of minutes. Since the framework is open source, it can be modified and new scientific tools may be

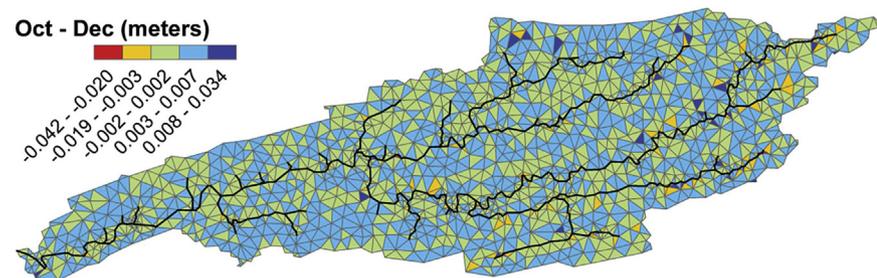
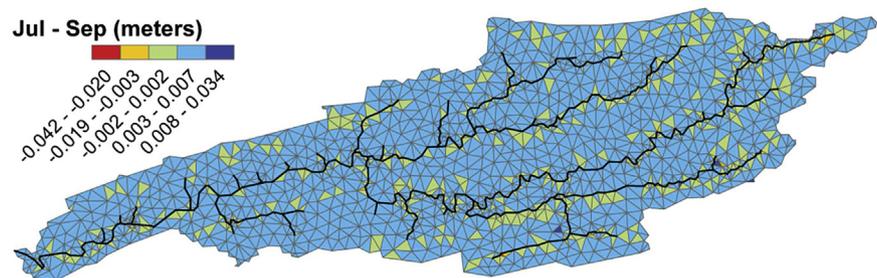
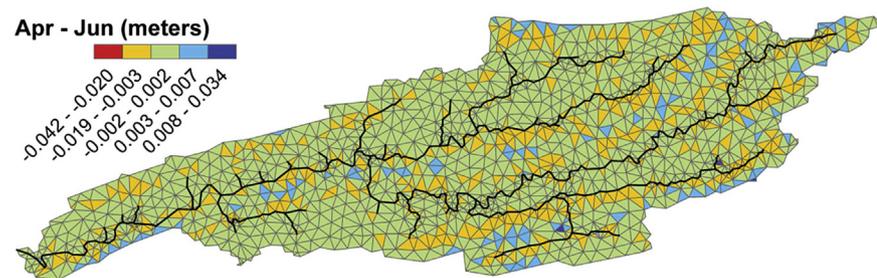
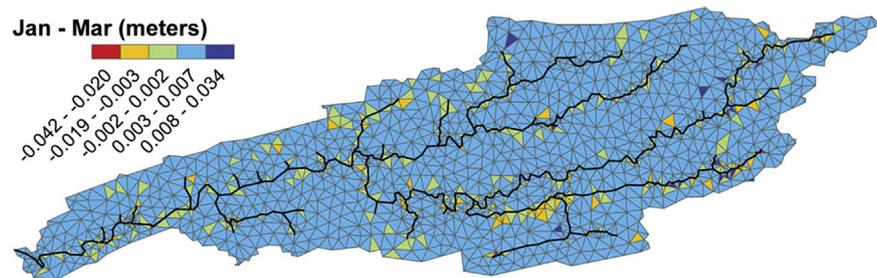


Fig. 11. An example of 'Spatial' analysis plot from a yearlong PIHM simulation in Mahantango Creek watershed. The plots show spatial and seasonal variability in simulated groundwater recharge across the watershed.

Table 4

Efficacy in prototyping a model simulation at Mahantago Creek watershed. The simulation was performed on a 2.53 GHz Intel Core 2 Duo, 4 GB DDR3 Memory, running Mac OS X Version 10.6.6 using PIHMgis Version 2.3. Computation wall times are reported in min:s.

a. Raster processing		d. Data model loader	
Modules	Time (m:s)	Modules	Time (m:s)
Fill pits	01:56	Mesh file	00:01
Flow direction	00:15	Attribute file	00:02
Stream grid	00:03	River file	00:02
Link grid	00:04	Parameters file	00:01
Stream polyline	00:02	Soil file	00:02
Catchment grid	00:04	Geology file	00:02
Catchment polygon	00:53	Land cover file	00:01
b. Vector processing		e. Run PIHM	
Modules	Time (mm:ss)	Modules	Time (m:s)
Dissolve	00:02	PIHM ^a	147:00
Polygon to polyline	00:01		
Simplify polyline	00:02		
Split line	00:01		
Vector merge	00:01		
c. Domain decomposition		f. Analysis	
Modules	Time (m:s)	Modules	Time (m:s)
Shape topology	00:02	Time series plot	09:05
Run triangle	00:01	Spatial plot	06:21
Generate Mesh	00:01		

^a A prototype simulation of one-year time period.

integrated. As the framework is also supported on the Unix/OSX or Linux platforms, it can be used under high performance computing environment to support parallel hydrologic model simulation or for running parallel instances of the serial model.

In future, the level of coupling between QGIS and PIHM in PIHMgis can be extended to embedded coupling, thus providing opportunity to modify parameters and states during the simulation. The salient characteristics needed to facilitate such coupling are already in place in terms of: (a) a shared data model support for GIS and hydrologic model; (b) use of the same programming language for the development of GIS tools and hydrologic model; (c) a data structure design that makes it easier to query the data model and model control parameters; (d) an option of stopping the model simulation at any time and restarting after reinitializing the state variables and parameters; and (e) the ability to launch different applications using library loaders and linkers. The algorithms implemented in PIHMgis can also be ported to a web-based cloud environment to make these tools available as web-services. These cloud-based modeling tools may query consolidated national geospatial databases (such as one on the HydroTerre server, www.hydroterre.psu.edu) to further cement cloud-based model-data integration. The strategy would provide educators, scientists, engineers and managers cross-platform access of integrated web-based modeling, analysis, and visualization tools in rapidly evolving post desktop (PC) environment.

Acknowledgments

This study was supported by grant from the National Science Foundation (EAR 0725019 Shale Hills-Susquehanna Critical Zone Observatory), the Duke University start-up grant, and the Environmental Protection Agency (Chesapeake Bay Program Office, CB96305901). We would like to thank Massimiliano Zappa and two other anonymous reviewers for their constructive comments that greatly improved this manuscript.

References

Abbott, M., Bathurst, J., Cunge, J., O'Connell, P., Rasmussen, J., 1986a. An introduction to the European hydrological system – système hydrologique européen, "SHE",

- 1: History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.* 87, 45–59.
- Abbott, M., Bathurst, J., Cunge, J., O'Connell, P., Rasmussen, J., 1986b. An introduction to the European hydrological system – système hydrologique européen. "SHE", 2: structure of a physically-based, distributed modelling system. *J. Hydrol.* 87, 61–77.
- Abel, D.J., Kilby, P.J., 1994. The systems integration problem. *Int. J. Geogr. Inform. Syst.* 8 (1), 1–12.
- Bhatt, G., 2012. A Distributed Hydrologic Modeling System: Framework for Discovery and Management of Waters Resources. PhD dissertation. Civil and Environmental Engineering, Pennsylvania State University, 372 pp.
- Boyle, D.P., Gupta, H.V., Sorooshian, S., 2003. Multicriteria calibration of hydrologic models. In: Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. (Eds.), *Calibration of Watershed Models*, Water Sci. Appl. vol. 6. AGU, Washington, D.C., pp. 185–196.
- Brimicombe, A., 2003. GIS, Environmental Modeling and Engineering, first ed. Taylor & Francis, London. pp. 320.
- Chen, X., Kumar, M., McGlynn, B.L., 2014. Variations in streamflow response to large hurricane-season storms in a Southeastern US watershed. *J. Hydrometeorol.* <http://dx.doi.org/10.1175/JHM-D-14-0044>.
- Clark, M.J., 1998. Putting water in its place: a perspective on GIS in hydrology and water management. *Hydrol. Process.* 12 (6), 823–834.
- Cohen, S.D., Hindmarsh, A.C., 1994. CVODE User Guide. Technical Report UCRL-MA-118618. Lawrence Livermore National Laboratory.
- Coroza, O., 1997. Enhancing runoff modeling with GIS. *Landsc. Urban Plan.* 38, 13–23.
- Deckmyn, J., Paniconi, C., Giacomelli, A., Kleinfeldt, S., 1997. A modular approach to integrated environmental modeling systems incorporating GIS and visualization functionality, CRS4-TECH-REP-96/26. In: *Proc. of an International Workshop on New Development in Geographic Information Systems*. ERIM International, Ann Arbor, Milano, Italy, pp. 85–95.
- Douglas, D.H., Peucker, T.K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Can. Cartogr.* 10 (2), 112–122.
- DeVantier, B.A., Feldman, A.D., 1993. Review of GIS applications in hydrologic modeling. *J. Water Resour. Plan. Manag.* 119, 246–261.
- DHI: MIKE SHE technical reference, version 2005, 2005. DHI Water and Environment, Danish Hydraulic Institute, Denmark.
- Duffy, C.J., Leonard, L., Bhatt, G., Yu, X., Giles, C.L., 2011. Watershed reanalysis: towards a national strategy for model-data integration. *e-Science Workshops 2011*: pp. 61–65.
- Engel, B.A., Srinivisan, R., Arnold, J., Rewerts, C., Brown, S.J., 1993. Nonpoint source (NPS) pollution modeling using models integrated with geographic information systems (GIS). *Water Sci. Technol.* 28 (3–5), 685–690.
- Entekhabi, D., Eagleson, P.S., 1989. Land surface hydrology parameterization for atmospheric general circulation models including subgrid scale spatial variability. *J. Clim.* 2 (8), 816–831.
- Freeze, R.A., Harlan, R.L., 1969. Blueprint for a physically-based digitally simulated, hydrologic response model. *J. Hydrol.* 9, 237–258.
- Goodchild, M.F., 1992. Geographical information science. *Int. J. Geogr. Inform. Syst.* 6, 31–45.
- Goodchild, M., Haining, R., Wise, S., 12 others, 1992. Integrating GIS and spatial data analysis: problem and possibilities. *Int. J. Geogr. Inform. Syst.* 6, 407–423.
- Haverkamp, S., Fohrer, N., Frede, H.G., 2005. Assessment of the effect of land use patterns on hydrologic landscape functions: a comprehensive GIS-based tool to minimize model uncertainty resulting from spatial aggregation. *Hydrol. Process.* 19, 715–727. <http://dx.doi.org/10.1002/hyp.5626>.
- Hellweger, F.L., Maidment, D.R., 1999. Definition and connection of hydrologic elements using geographic data. *J. Hydrol. Eng.* 4 (1), 10–18.
- Hindmarsh, A.C., Brown, P.N., Grant, K.E., Lee, S.L., Serban, R., Shumaker, D.E., Woodward, C.S., 2005. SUNDIALS: suite of nonlinear and differential/algebraic equation solvers. *ACM Trans. Math. Softw.* 31 (3), 363–396.
- Huang, B., Jiang, B., 2002. AVTOP: a full integration of TOPMODEL into GIS. *Environ. Model. Softw.* 17, 261–268.
- Jenson, S.K., Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogram. Eng. Remote Sens.* 54 (11), 1593–1600.
- Kinerson, R.S., Kittle, J.L., Duda, P.B., 1999. BASINS: Better Assessment Science Integrating Point and Nonpoint Sources, Decision Support Systems for Risk-based Management of Contaminated Sites, pp. 1–24.
- Kirchner, J.W., 2006. Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.* 42 (3).
- Kollet, S.J., Maxwell, R.M., 2006. Integrated surface-groundwater flow modeling: a free-surface overland flow boundary condition in a parallel groundwater flow model. *Adv. Water Resour.* 29 (7), 945–958.
- Kopp, S.M., 1996. Linking GIS and hydrological models: where have we been, where are we going? In: Kover, K., Nachtenebel, H.P. (Eds.), *HydroGIS 96: Application of Geographic Information Systems in Hydrology and Water Resources*, IAHS Publ. vol. 235, pp. 13–21.
- Kouwen, N., 1988. WATFLOOD: a micro-computer based flood forecasting system based on real-time weather radar. *Can. Water Resour. J.* 13 (1), 62–77.
- Kumar, M., 2009. Toward a Hydrologic Modeling System. PhD dissertation, Civil and Environmental Engineering, Pennsylvania State University. <https://etd.libraries.psu.edu/paper/9870/5576>.

- Kumar, M., Duffy, C.J., Salvage, K.M., 2009a. A second-order accurate, finite volume-based, integrated hydrologic modeling (FIHM) framework for simulation of surface and subsurface flow. *Vadose Zone J.* 8, 873–890. <http://dx.doi.org/10.2136/vzj2009.0014>.
- Kumar, M., Bhatt, G., Duffy, C.J., 2009b. An efficient domain decomposition framework for accurate representation of geodata in distributed hydrologic models. *Int. J. Geogr. Inform. Sci.* 23 (12), 1569–1596.
- Kumar, M., Bhatt, G., Duffy, C.J., 2010. An object-oriented shared data model for GIS and distributed hydrologic models. *Int. J. Geogr. Inform. Sci.* 24 (7), 1061–1079.
- Kumar, M., Marks, D., Dozier, J., Reba, M., Winstrol, A., 2013. Evaluation of distributed hydrologic impacts of temperature-index and energy-based snow model simulations. *Adv. Water Resour.* 56, 77–89.
- Lahlou, M., Shoemaker, L., Choudry, S., Elmer, R., Hu, A., Manguerra, H., Parker, A., 1998. Better Assessment Science Integrating Point and Nonpoint Sources: BASINS 2.0 User's Manual. EPA-823-B-98-006. U.S. Environmental Protection Agency, Office of Water, Washington, DC, USA.
- Leon, L.F., 2009. MapWindow interface for SWAT (MWSWAT). In: Soil and Water Assessment Tool (SWAT) Global Application. WASWAC Special Publ, vol. 4.
- Leonard, L., Duffy, C.J., 2013. Essential terrestrial variables data workflow for distributed water resources modeling. *Environ. Model. Softw.* 50, 85–96.
- Leveque, R.J., 2002. Finite Volume Methods for Hyperbolic Problems. Cambridge University Press, 558 pp.
- Li, S., Duffy, C., 2011. Fully coupled approach to modeling shallow water flow, sediment transport, and bed evolution in rivers. *Water Resour. Res.* 47, W03508.
- Lu, J., Sun, G., McNulty, S.G., Comerford, N.B., 2009. Sensitivity of pine flatwoods hydrology to climate change and forest management in Florida, USA. *Wetlands* 29 (3), 826–836.
- Liu, Z., 2005. ArcTOP: a distributed hydrologic modeling system of tight coupling TOPKAPI with GIS. *Hydrology* 25 (4), 18–22.
- Luzio Di, M., Srinivasan, R., Arnold, J.G., Neitsch, S.L., 2002. Arcview Interface for SWAT2000, User's Guide. U.S. Department of Agriculture, Agriculture Research Service, Temple, Texas.
- Maidment, D.R., 1993. GIS and hydrological modeling. In: Goodchild, M.F., Parks, B., Steyaert, L. (Eds.), *Environmental Modeling with GIS*. Oxford University Press, New York, pp. 147–167.
- Maidment, D.R., 1996. GIS and hydrologic modeling: an assessment of progress. In: Proceedings of the Third International Conference on Integrating GIS and Environmental Modeling. Santa Fe, New Mexico.
- Maidment, D.R., 2002. Arc Hydro: GIS for Water Resources. ESRI Press, Redlands, CA, 203 pp.
- McDonald, M.G., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference groundwater flow model. U.S. Geological Survey Techniques of Water-Resources Investigations Book 6, Ch. A1.
- McDonnell, R.A., 1996. Including the spatial dimension: using geographical information systems in hydrology. *Prog. Phys. Geogr.* 20, 159–177.
- Meixner, T., Gupta, H., Bastidas, L., Bales, R., 2003. Estimating parameters and structures of a hydrologic model using multiple criteria. In: Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. (Eds.), *Calibration of Watershed Models*, Water Sci. Appl, vol. 6. AGU, Washington, D. C., pp. 213–228.
- Moore, I.D., 1996. Hydrologic modeling and GIS. In: Goodchild, M.F., Parks, B.O., Steyaert, L.T. (Eds.), *GIS and Environmental Modeling: Progress and Research Issues*, pp. 143–148. Fort Collins, CO.
- Nelson, E.J., 1997. WMS v5.0 Reference Manual, Environmental Modeling Research Laboratory. Brigham Young University, Provo, Utah, p. 462.
- NLDAS Vegetation Parameters. UMD Vegetation parameters tables, <http://ldas.gsfc.nasa.gov/nldas/NLDASmapveg.php> (accessed 2013).
- Nyerges, T.L., 1987. GIS research issues identified during a cartographic standards process. In: Proceedings, of International Symposium on Geographic Information Systems.
- Nyerges, T.L., 1993. Understanding the scope of GIS: its relationship to environmental modeling. In: Goodchild, M.F., Parks, B.O., Steyaert, T. (Eds.), *Environmental Modeling with GIS*. Oxford University Press, New York, pp. 75–93.
- O'Callaghan, J.F., Mark, D.M., 1984. The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process.* 28, 328–344.
- Olivera, F., Valenzuela, M., Srinivasan, R., Choi, J., Cho, H., Koka, S., Agrawal, A., 2006. ArcGIS-SWAT: a geodata model and GIS interface for SWAT. *J. Am. Water Resour. Assoc.* 42 (2), 295–309.
- Panday, S., Huyakorn, P.S., 2008. MODFLOW SURFACT: a state-of-the-art use of vadose zone flow and transport equations and numerical techniques for environmental evaluations. *Vadose Zone J.* 7, 610–631.
- Pitman, A.J., Henderson-Sellers, A., Yang, Z.L., 1990. Sensitivity of regional climates to localized precipitation in global models. *Nature* 346, 734–737.
- Qu, Y., 2004. An Integrated Hydrologic Model for Multi-process Simulation using Semi-discrete Finite Volume Approach. PhD dissertation. Pennsylvania State University.
- Qu, Y., Duffy, C.J., 2007. A semidiscrete finite volume formulation for multiprocess watershed simulation. *Water Resour. Res.* 43, W08419. <http://dx.doi.org/10.1029/2006WR005752>.
- Refsgaard, J.C., 1997. Parameterization, calibration and validation of distributed hydrological models. *J. Hydrol.* 198 (1–4), 69–97.
- Ruppert, J., 1995. A Delaunay refinement algorithm for quality 2-dimensional mesh generation. *J. Algorithm.* 18 (3), 548–585.
- Shewchuk, J.R., 1996. Triangle: engineering a 2D quality mesh generator and Delaunay triangulator, in applied computational geometry: towards geometric engineering. In: *Lecture Notes in Computer Science*, vol. 1148. Springer-Verlag, Berlin, pp. 203–222.
- Shi, Y., Davis, K.J., Duffy, C.J., Yu, X., 2013. Development of a coupled land surface hydrologic model and evaluation at a critical zone observatory. *J. Hydro-meteorol.* 14, 1401–1420.
- Shrestha, R.R., Rode, M., 2008. Multi-objective calibration and fuzzy preference selection of a distributed hydrological model. *Environ. Model. Softw.* 23 (12), 1384–1395.
- Singh, V.P., Fiorentino, M., 1996. *Geographical Information Systems in Hydrology*. Dordrecht, Netherlands.
- Smith, P.N., Maidment, D.R., 1995. Hydrologic Data Development System. CRWR Online Report 95-1. Center for Research in Water Resource, The University of Texas at Austin, Austin, TX.
- Storck, P., Bowling, L., Wetherbee, P., Lettenmaier, D., 1998. Application of a GIS-based distributed hydrology model for prediction of forest harvest effects on peak streamflow in the Pacific Northwest. *Hydrol. Process.* 12 (6), 889–904.
- Sui, D.Z., Maggio, R.C., 1999. Integrating GIS with hydrologic modeling: practices, problems and prospects. *Comput. Environ. Urban Syst.* 23, 33–51.
- SUNDIALS. <http://www.llnl.gov/casc/sundials/>, (accessed Sept. 2006).
- TauDEM. <http://hydrology.neng.usu.edu/taudem/> (accessed Sept., 2006).
- VanderKwaak, J.E., 1999. Numerical Simulation of Flow and Chemical Transport in Integrated Surface-subsurface Hydrologic System. Doctorate thesis. Department of Earth Sciences, University of Waterloo, Ontario, Canada.
- Viviroli, D., Zappa, M., Gurtz, J., Weingartner, R., 2009. An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools. *Environ. Model. Softw.* 24 (10), 1209–1222.
- Wang, R., Kumar, M., Marks, D., 2013. Anomalous trend in soil evaporation in semi-arid, snow-dominated watersheds. *Adv. Water Resour.* 57, 32–40.
- Wosten, J.H.M., Lilly, A., Names, A., LeBas, C., 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90, 169–185.
- Yeh, G.T., Huang, G.B., 2003. A Numerical Model to Simulate Water Flow in Watershed System of 1-D Stream-river Network, 2-D Overland Regime, and 3-D Subsurface Media (WASH123D: Version 1.5). Technical Report. Department of Civil and Environmental Engineering, University of Central Florida, Orlando, Florida.
- Yu, X., Bhatt, G., Duffy, C., Shi, Y., 2013. Parameterization for distributed watershed modeling using national data and evolutionary algorithm. *Comput. Geosci.* 58, 80–90.
- Yu, X., Lamačová, A., Duffy, C., Krám, P., Hruška, J., White, T., Bhatt, G., 2014. Modeling long term water yield effects of forest management in a Norway spruce forest. *Hydrol. Sci. J.* <http://dx.doi.org/10.1080/02626667.2014.897406>.