Environmental Modelling and Software 111 (2019) 121–133

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Development of a participatory Green Infrastructure design, visualization and evaluation system in a cloud supported jupyter notebook computing environment

Lorne Leonard[a,*], Brian Miles[b], Bardia Heidari[c], Laurence Lin[d], Anthony M. Castronova[e], Barbara Minsker[f], Jong Lee[g], Charles Scaife[d], Lawrence E. Band[h]

[a] Department of Civil and Environmental Engineering, Institutes of Energy and the Environment, The Pennsylvania State University, University Park, PA 16802, USA
[b] Brian Miles, CGI Group Inc., Lafayette, LA, USA
[c] Department of Civil and Environmental Engineering, University of Illinois at Urbana Champaign, Urbana, IL, USA
[d] Department of Environmental Sciences, University of Virginia, Virginia, USA
[e] Consortium for the Advancement of Hydrological Science, Cambridge, MA, USA
[f] Department of Civil and Environmental Engineering, Southern Methodist University, Dallas, TX, USA
[g] National Center for Supercomputing Applications, The University of Illinois Champaign, Urbana, IL USA
[h] Department of Environmental Sciences, Department of Engineering Systems and Environment, University of Virginia, Virginia, USA

A B S T R A C T

Land use planners, landscape architects, and water resource managers are using Green Infrastructure (GI) designs in urban environments to promote ecosystem services including mitigation of storm water flooding and water quality degradation. An expanded set of urban sustainability goals also includes increasing carbon sequestration, songbird habitat, reducing urban heat island effects, and improvement of landscape aesthetics. GI is conceptualized to improve water and ecosystem quality by reducing storm water runoff at the source, but when properly designed, may also benefit these expanded goals. With the increasing use of GI in urban contexts, there is an emerging need to facilitate participatory design and scenario evaluation to enable better communication between GI designers and groups impacted by these designs. Major barriers to this type of public participation is the complexity of both parameterizing, operating, visualizing and interpreting results of complex ecohydrological models at various watershed scales that are sufficient to address diverse ecosystem service goals. This paper demonstrates a set of workflows to facilitate rapid and repeatable creation of GI landscape designs which are incorporated into complex models using web applications and services. For this project, we use the RHESSys (Regional Hydro-Ecologic Simulation System) ecohydrologic model to evaluate participatory GI landscape designs generated by stakeholders and decision makers, but note that the workflow could be adapted to a set of other watershed models.

1. Introduction

A growing paradigm in urban environmental management is the adaptation and enhancement of ecosystem services with Green Infrastructure (GI) to mitigate the effects of urbanization on stormwater flooding and water quality degradation, urban heat islands, air quality and other adverse impacts. As hydrological and ecosystem models evolve to explicitly represent the influence of fine scale landscape form on the cycling and export of water, carbon and nutrients, public perceptions and preferences need to be included in the management process by which water resource managers and landscape planners design and implement new urban design and infrastructure. Often the public is not contemplating the important interaction between the water cycle and ecosystem processes, with the design form (architecture and engineered) and management of their own property or neighborhood. Therefore, environmental modeling software should be designed to address the strongly integrated set of ecosystem services affected by urban form and infrastructure, and engage individuals and communities about their own role in improving water quality and mitigating extreme events at the parcel to watershed scales. A major management barrier is the complexity of installing, compiling, and consuming datasets with watershed models using sufficient spatial resolution and process representation of the fine scale ecohydrological interactions of landscape design and structure. The conventional process by many hydrological modelers is to prepare data as an offline process from multiple sources, copy the processed datasets to the compute environment for modeling, then process the watershed model results offline with calibration and visualization. These processes are cumbersome

* Corresponding author.
E-mail addresses: lorne.leonard.psu@gmail.com, lnl3@psu.edu (L. Leonard), selimnairb@gmail.com (B. Miles), hdrhrtm2@illlinois.edu (B. Heidari), hl8vq@virginia.edu (L. Lin), acastronova@cuahsi.org (A.M. Castronova), minsker@smu.edu (B. Minsker), jonglee1@illinois.edu (J. Lee), cis9zg@virginia.edu (C. Scaife), leb3t@virginia.edu (L.E. Band).

https://doi.org/10.1016/j.envsoft.2018.10.003
Received 30 July 2017; Received in revised form 2 September 2018; Accepted 8 October 2018
Available online 13 October 2018
1364-8152/

Software availability

Name GI Designer, RHESSys notebooks.
Software Required Internet browser with HTML 5 support. Later versions are recommended.
Program Languages Python, JavaScript, HTML, C.
Availability and cost The notebooks are available from GitHub: https://github.com/leonard-psu/RHESSys-Jupyter-notebooks to use in any Jupyter compute environment. To use the Jupyter notebooks with HydroShare, an account is required that is free. The notebooks are available by using HydroShare's discover search engine. The Green Infrastructure (GI) notebook presented in this manuscript is available here: Leonard, L., L. Band, L. Lin, B. Miles (2018). Green Infrastructure Designer with RHESSys Workflow, HydroShare, https://doi.org/10.4211/hs.3f7680cf83dc426e858d5b48cb95a565.
The GI web application is here https://hydroterremodels.psu.edu/RHESSYS/GI_Designer/GI.html. EcoHydroLib is available from https://github.com/selimnairb/EcohydroLib and RHESSysWorkflows from https://github.com/selimnairb/RHESSysWorkflows.

and make it difficult for other professionals and the public to participate in improving the modeling process and critiquing watershed model results. The research presented in this manuscript is a step towards improving these processes, by using workflows to make data and model processes more efficient to evaluate watershed model results and engage other professionals and the public with GI landscape designs to evaluate their impact to ecohydrological interactions.

The premise of this paper has **two components**, conceptualized to facilitate participatory design and evaluation of GI planning. The participatory design assumes that stormwater engineers, community groups and potentially individual residents would collaborate in the design of preferred GI, with rapid evaluation of different design effectiveness for reaching stormwater restoration goals. The **first** component is to engage and connect storm water engineers and property owners with rapid charrettes of applying GI landscape designs on their property using web-based visualizations. Then, connecting the GI design with agency regional plans as an integrated system within the local watershed to be used as inputs for ecohydrological modeling to rapidly evaluate the design efficacy of mitigating impacts in both the built and natural environments, with the potential to iterate this process to balance desirable and effective modification of the landscape utilizing GI.

Here, we demonstrate using a spatially distributed ecohydrological model, the Regional Hydro-Ecologic Simulation System (Tague and Band, 2004) (RHESSys), although the GI data workflows serve as a template for other models to adapt and consume. The **second** component is to simplify and improve the modeling setup to consume fine resolution land information and GI in RHESSys models by using multiple workflows and Jupyter notebooks to share data, code, simulations, and visualizations. Some of these workflows were provided by the HydroTerre expert system, to access Essential Terrestrial Variables from national datasets within the United States of America (Leonard and Duffy, 2016a, 2014, 2013). We demonstrate using the GI web application within a Jupyter notebook to setup a RHESSys model at the Dead Run watershed located in Baltimore, Maryland, USA. The RHESSys notebooks serve as a prototype for beginner and expert users to modify their region of interest and data sources. These notebooks are available to the community within the HydroShare collaborative environment (CUAHSI, 2017; D. G. Tarboton et al., 2014) and as resources via GitHub for other Jupyter hosted environments.

- We developed a new GI web application to allow web-users to design GI with a plan and street view perspective. Users can assign custom GI attributes for their modeling requirements and estimate costs to share with other professionals and the public.
- New GI-model-data workflows (GI-RHESSys-HydroTerre) were created to process the GI landscape designs from the GI web application into fine spatial resolution model parameter files. These new workflows integrate the RHESSys ecohydrological model and the HydroTerre expert system to integrate GI data bundles with local custom spatial datasets (e.g. terrain, soils, land cover) supplied by the user or existing HydroTerre national spatial data (United States).
- We demonstrate using a new GI-RHESSys Jupyter notebook that uses the GI web application to create GI landscape designs. These designs are processed using the GI-RHESSys-HydroTerre data workflows and then processed within the GI-RHESSys Jupyter notebook. The GI-RHESSys Jupyter notebook uses the existing EcoHydroLib and RHESSys workflows to prepare data and execute the RHESSys ecohydrological model using the user's GI landscape designs. These notebooks serve to facilitate participatory use and evaluation of GI planning.

This article is structured with Section 2 summarizing background information and the inspiration to combine GI with RHESSys and workflows using a Jupyter notebook. Section 3 discusses the back-end web application design using multiple workflows. Section 4, describes and demonstrates the GI web application for modelers to create GI landscape designs. Finally, Section 5 demonstrates using the web application within Jupyter notebooks to develop a new RHESSys ecohydrological model.

2. Background

2.1. The role and requirements for Green Infrastructure (GI)

The motivation for GI in urban areas is to manage and protect the water cycle without relying solely on conventional engineered solutions (i.e. dams, levees, detention ponds). Here we operationally define GI to include various forms of soil, vegetation, and water storage treatments, including tree canopy, rain gardens, bioswales, permeable pavement that promote processes mimicking natural infiltration, transpiration, subsurface storage, transport and processing of reduce and slow precipitation-runoff. Land use planners, landscape architects, and water resource managers are using GI solutions to reduce storm water flow from impervious surfaces (e.g. roofs, roads, parking lots) in urbanized watersheds, mitigate downstream water quality problems and urban heat islands, while creating amenities for residents (Jefferson et al., 2017; Miles and Band, 2015a; Miles, 2014; Poff et al., 1997; Walsh et al., 2005). To treat non-point source storm water in older established urbanized areas (i.e. Baltimore Maryland, USA), federal, state, and municipal agencies are working with residential and commercial owners to retrofit impervious surfaces. For example, rather than have roofs drain rainwater directly through gutters into roads and storm sewers, the rainwater is stored using rain barrels and tanks to be slowly released between storms or routed to infiltrate into the ground using pervious surfaces, including rain gardens, bioswales, and wetlands before the water is directed to roads and storm sewers. Many municipalities have also set goals to increase tree canopy cover to promote multiple ecosystem services (Pincetl et al., 2013). In the examples used in this paper, we emphasize landscape designed GI (e.g. swale drains, rain gardens) and increased tree canopy coverage.

The goals of federal and local agencies are to improve water quality and reduce costs from flooding events and large engineering projects by mimicking the pre-development hydrology by treating storm water

The following are our primary research contributions:

runoff close to or at the origin (United States Environmental Protection Agency, 2017). Common goals in stormwater treatment include reduction in peak flows, increased baseflow, and reduction in nutrient

and sediment loading. Therefore, these agencies are interested in GI landscape design solutions that will address coupled ecosystem and hydrologic cycling, ranging from the building site to the watershed scale, meeting regulations and policies from a variety of government and non-government organizations. Additional sustainability goals expressed by municipalities include a range of other ecosystem services benefiting from increased tree cover. To better understand how successful GI landscape design solutions can be at reducing storm water runoff at a watershed scale, potential designs need to be integrated so their combined effects can be evaluated. Doing so suggest the use of a distributed ecohydrologic model that couples water, carbon and nutrient cycling along hydrologic flowpaths to scale up GI designs at the residential scale (datasets with resolutions on the order of meters) to target watersheds (up to ~10 square kilometers) to evaluate the impact of reducing storm water overland flow and improve ecohydrology.

2.2. RHESSys and workflows

For this project, we used the Regional HydroEcological Simulation System (RHESSys) (Tague and Band, 2004) to evaluate GI landscape designs for three major reasons. First, RHESSys has been developed to simulate hydrological and ecological processes at patches (1–10-m resolution), hillslopes (10–100-m resolution), watershed, and basin scales. It integrates the above ground canopy processes, e.g. evapotranspiration and net canopy photosynthesis, and below ground biological processes, e.g. root function, litter and soil organic matter decomposition, nitrification, and denitrification (Running and Coughlan, 1988), soil water distribution and routing (Beven and Kirby, 1979; Wigmosta et al., 1994), and topo-climate extrapolation (Running et al., 1987). The explicitly modeled ecohydrological interaction between vegetation and water provides the basis for landscape design in GI. While RHESSys was initially developed for research and prediction in unmanaged and rural watersheds, it has been adapted for the study and forecast of ecohydrology of managed and urban watersheds (Bell et al., 2017; Martin et al., 2017; Mittman et al., 2012). Second, RHESSys simulates spatially distributed and coupled water, carbon and nitrogen dynamics using a landscape hierarchical structure over nested patch, hillslope, and watershed scales. This feature makes it possible to evaluate and monitor GI performance at fine spatial scale and their cumulative and emergent effect at the watershed landscape. Third, instead of lumping at coarse spatial scale as other watershed models (e.g. Soil and Water Assessment Tool, SWAT, Hydrological Simulation Program Fortran, HSPF), water and solutes are spatially distributed within hillslopes and watersheds in RHESSys. This feature allows for simulating the distribution/routing of rain/storm water from impervious surfaces (e.g. roofs and roads), and the cycling of carbon and nutrients into designed GI in the landscape.

RHESSys and other hydrological models (i.e. PIHM, SWAT, VIC) require modelers to prepare (i.e. re-project) and transform (i.e. generate model input files) a diverse range of geospatial datasets from multiple agencies (i.e. USGS, NOAA, NASA, EPA, USDA) or from local sources. These processes are time consuming, require a broad range of expertise and are often difficult to reproduce. Given the prospect of petabyte datasets, automation is essential to process the diverse range of datasets for hydrological modeling (Leonard and Duffy, 2013). To prepare RHESSys models that incorporate GI designs, the EcohydroLib (Miles and Band, 2017a, 2015b) and RHESSys workflows (Miles and Band, 2017b) simplify the process of generating RHESSys input from federal agencies or local datasets and generating model parameter files. Fur-

2.3. HydroShare's interactive jupyter notebook computing environment

HydroShare (https://www.hydroshare.org/) is an online, collaborative infrastructure to share hydrologic data, model, and applications (Crawley et al., 2017; Horsburgh et al., 2015; Morsy et al., 2017; D. Tarboton et al., 2014; D. G. Tarboton et al., 2014). Hydroshare manages data and model files as fully referenced “resources” which can be searched and cited. A resource is defined as “a set of digital content that includes the science data files or model files and their corresponding metadata information” (DCMI, 2018; HydroShare, 2018; OAI, 2018). HydroShare enables users to share hydrological models by using the Jupyter (http://jupyter.org/) Notebook Application, a server-client application that allows web users to edit and run notebook documents (IPython) within a web browser on a local or remote compute environment (Jupyter Team, 2017a,b; Perez and Granger, 2007). This paper demonstrates using Jupyter notebooks to prepare a RHESSys hydrological model with GI landscape designs supported with data and model workflows, using HydroShare's Jupyter cloud compute environment supported by Consortium of Universities for the Advancement of Hydrologic Science, Inc CUASHI.

Jupyter notebook documents are ideal to introduce new modelers, other professionals (i.e. landscape planners) and stakeholders to all the steps with documentation and samples included in a single document. In this case, we use the Jupyter notebook to prepare and execute a RHESSys hydrological model on a remote cloud compute environment, and include design and visualization of new GI. We have developed a collection of RHESSys Jupyter notebooks prepared for different user skill levels (i.e. beginner to expert) and data requirements (i.e. gage location based, custom datasets, and with national datasets) that allows users to go through all the steps. The notebooks can be accessed from GitHub (L. Leonard and Band, 2017) or by searching for RHESSys resources via HydroShare. Users interested in learning how to use RHESSys do not need to install RHESSys and dependencies on their own compute environment. Using the HydroShare cloud environment, RHESSys notebooks are ready to be executed to create an ecohydrological model. Hence, new users spend more time on learning RHESSys capabilities, while experienced modelers copy and edit an existing notebook for their study location to devote more time on analysis and refinement of model results. Furthermore, after calibrating and refining ecohydrological model results, the Jupyter notebook is easy to share with other users by downloading the notebook in different file formats or by creating a HydroShare resource to share with the community for critique. For example, residents, stakeholders, and other professionals (i.e. landscape architects) can critique the GI design, and evaluate ecohydrological performance in a collaborative web application, before requesting GI design changes to meet the decision makers objectives.

Examples of hydrologic research using HydroShare's Jupyter notebook infrastructure include TauDEM, with a notebook that processes terrain datasets (Castronova, 2017) and an introduction to Landlab to build numerical landscape models (Bandaragoda, 2016) (Bandaragoda et al., 2017). Furthermore, work by (Heidari et al., 2018) is using the Storm Water Management Model (SWMM) to study trees and rain gardens at the Dead Run watershed with the GI web application described in this article (Heidari et al., 2018). Outside of HydroShare, water related research using Jupyter notebooks include the evaluation of analytics solutions for steady interface flow where the aquifer extends below the sea (Bakker et al., 2017) and end-to-end workflows for assessing sea surface temperature, salinity and water levels, predicted by coastal ocean models (Subramanian et al., 2015).

thermore, these Python tool suites enable modelers to incorporate their own custom and often high-resolution datasets. To encourage outreach, provenance, and reproducibility, this research incorporates RHESSys and workflows using an interactive Jupyter notebook computing environment with cloud resources for modelers to evaluate GI designs.

2.4. *Constraints*

Incorporation of vegetation biodiversity (e.g. tree species) at fine spatial scale has been challenging for watershed models including RHESSys, which requires detailed information regarding vegetation

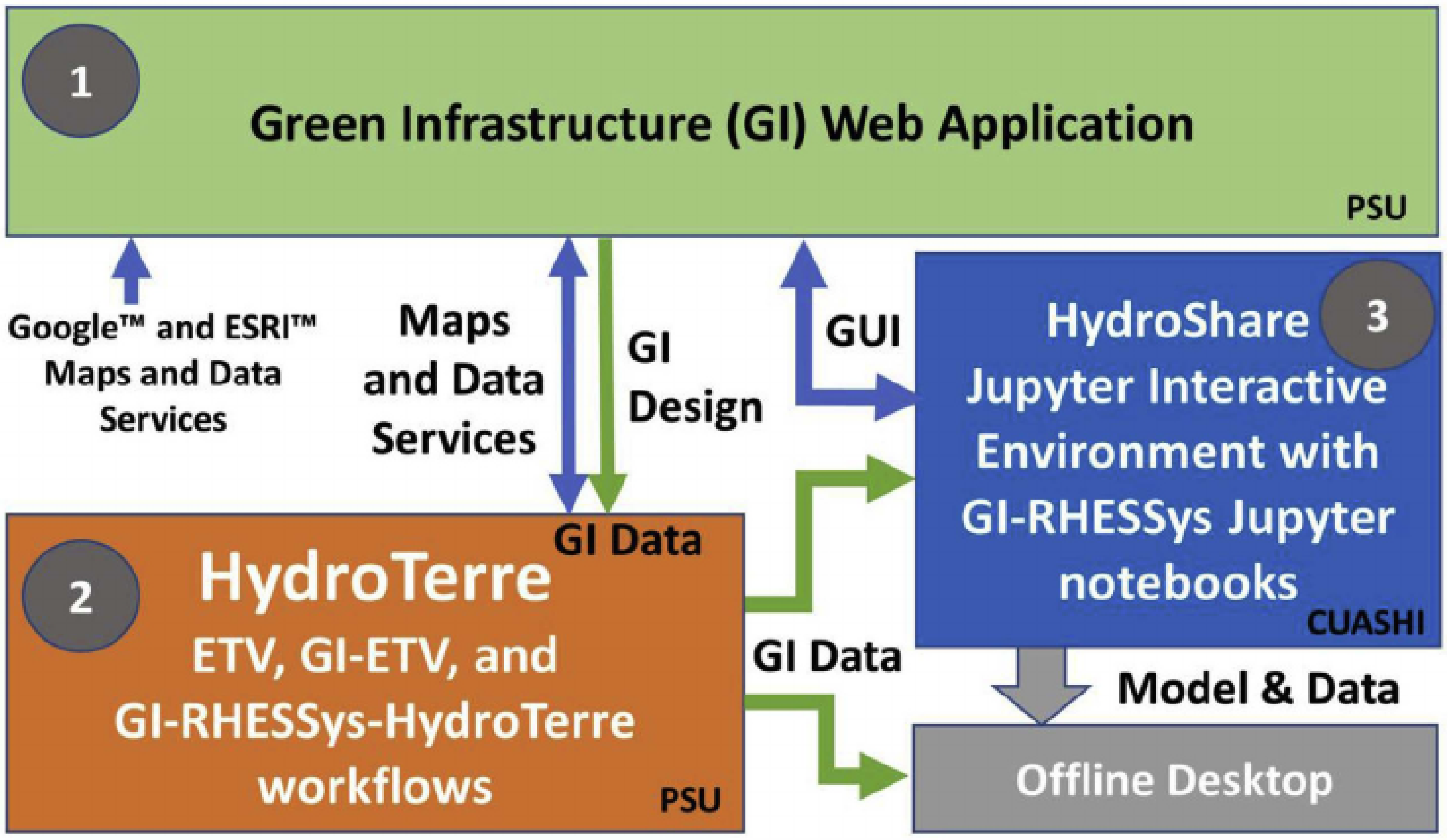


Fig. 1. There are three main components to using Green Infrastructure (GI) landscape designs within GI-RHESSys Jupyter notebook environment. (1) The GI web application that enables users to create GI designs online at street scales. (2) HydroTerre services (Essential Terrestrial Variables (ETV)) that merge GI designs with national and/or custom datasets. Referred to as GI-RHESSys-HydroTerre data workflows in this manuscript to explain data origin. (3) HydroShare's Jupyter compute environment that enables users to incorporate GI designs with RHESSys workflows and setup a RHESSys hydrological model. In this manuscript, we refer to these workflows (both data and model) and compute environment as GI-RHESSys Jupyter notebook. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

species (or plant functional type, PFT) composition and density at high spatial resolution. This type of information is difficult to acquire even with satellite imagery. In current RHESSys applications, classifying vegetation as grass, deciduous and evergreen is a common practice. Efforts have been made to further incorporate more diverse vegetation information in RHESSys based on PFT water use traits (Lin et al. in review). This feature will allow RHESSys to better simulate GI that are customized with vegetation species and management purposes appropriate for different climates and designs (Pataki et al., 2013).

Additionally, this article focuses on the software development to create GI designs using web applications and workflows that are embedded within cloud distributed compute environments using Jupyter notebooks that are easy to edit and share with the water resource and ecosystem management community as new data and model resources. In this presentation we do not validate data inputs, for example the appropriateness of placing GI by users, or inputs such as sewer infrastructure. However, spatial restrictions in GI placement can be spatially represented within the core GIS layers. Furthermore, we do not demonstrate calibrating and validating GI-RHESSys model results used in the Jupyter notebooks, instead we leave these issues as an exercise for notebook modeler users or assume a previously developed and calibrated model in the case where users will design and test new GI.

3. GI back-end workflows

This section describes the software architecture building blocks supporting the GI web application using multiple workflows required by the GI-RHESSys Jupyter notebook to evaluate GI landscape designs with RHESSys. The GI web application consumes multiple software and data services (Section 3.1) to enable interactive hydrological modeling compute environments. How the GI design web application is embedded into HydroShare's Jupyter interactive environment is described in Section 3.2, followed by how users' GI designs are transformed by the

provides tools for users to design GI in both two-dimension (2D) and 2.5 dimension (2.5D) environments. The second component is HydroTerre's services (http://hydroterre.psu.edu) hosted at PSU to place users' GI designs within both national and custom datasets. Referred to as GI-RHESSys-HydroTerre data workflows in this manuscript to explain data origin. The last component is HydroShare's Jupyter compute environment supported by Consortium of Universities for the Advancement of Hydrologic Science, Inc CUASHI (CUAHSI, 2017). In this manuscript, we refer to these workflows (both data and model) and compute environment as GI-RHESSys Jupyter notebook.

There are two GI web application versions. The first is a standalone application for users to download GI-RHESSys-HydroTerre data bundles for their modeling goals using a personal desktop computer. This version is available from PSU at (web link [1]). Users create a GI design, supported by map and data services from Google Maps™ (Google, 2017), ESRI™ (ESRI, 2017) and HydroTerre (Leonard and Duffy, 2016b, 2013, 2014) When the user is ready to evaluate their GI design, the design is transformed by HydroTerre data workflows to create GI data bundles. These data bundles can be downloaded for offline use.

The second version of the GI web application is an embedded web application for cloud-based modeling using Jupyter notebooks (web link[2]). Users interact with the GI web application as she or he would with browser software (i.e. Chrome, Firefox etc.) inside the notebook. The web user follows the same steps as he or she would with the standalone application. However, rather than downloading the data bundle for personal desktop use, the GI data bundle is accessed by a GI-RHESSys Jupyter notebook to create a RHESSys hydrological model using cloud resources.

3.2. *HydroShare jupyter interactive compute environment*

One method within HydroShare to share data and models is to store Jupyter notebooks within a resource. Jupyter notebooks are documents produced by the Jupyter Notebook Application containing computer

GI-RHESSys-HydroTerre data workflows (Section 3.3) for RHESSys model workflows and hydrological modeling (Section 3.4) to evaluate GI landscape designs hydrological impact.

3.1. Overview of main components

There are three main components to use GI landscape designs within the HydroShare Jupyter environment with RHESSys as summarized in Fig. 1. The first component is the GI web applications currently hosted at Penn State University (PSU). The GI web application

source code (e.g. R, Python, JavaScript, and HTML) and rich text elements (e.g. Tables, Figures) that are executable documents for hydrological data and model analysis (Jupyter Team, 2017a,b). For example, the GI-RHESSys Jupyter notebook is available from HydroShare as a resource here (Leonard et al., 2017). Fig. 2 summarizes how the GI web application interacts with the Jupyter notebook environment.

[1] https://hydroterremodels.psu.edu/RHESSYS/GI_Designer/GI.html.
[2] https://hydroterremodels.psu.edu/RHESSYS/GI_Designer/GI_HS.html.

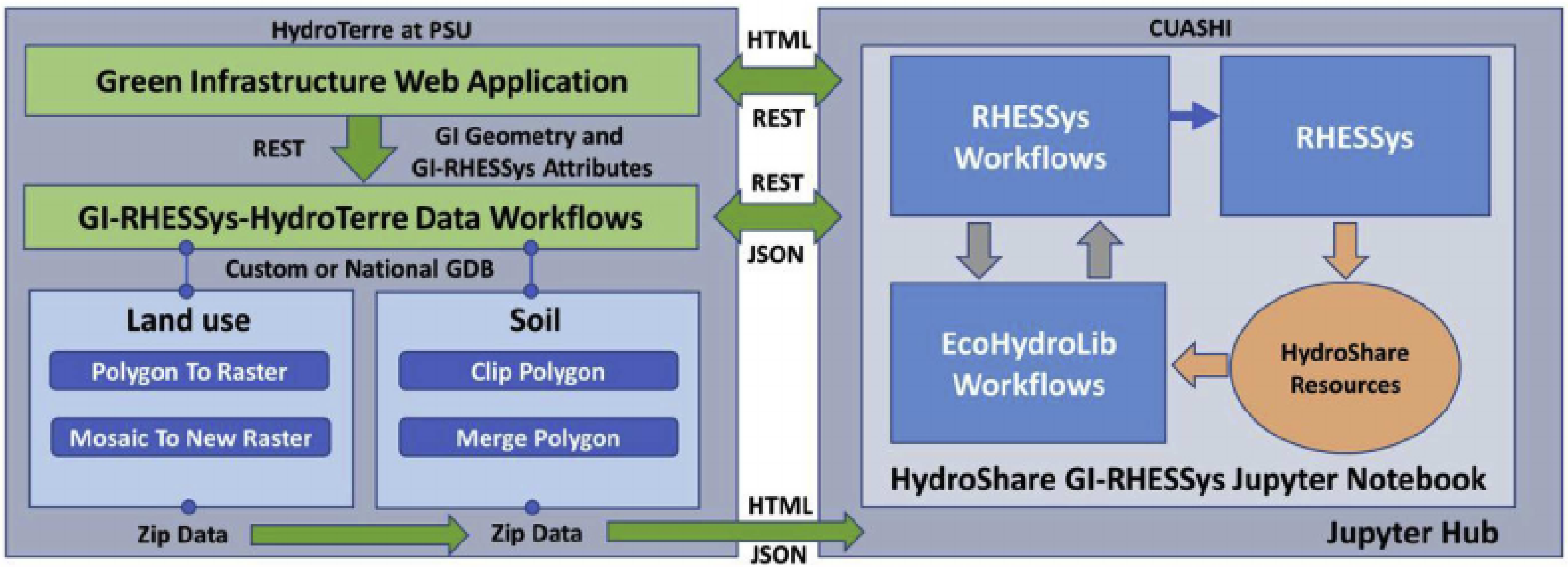


Fig. 2. The Green Infrastructure (GI) web application uses RESTful web services to submit GI geometry and RHESSys attributes to GI-RHESSys-HydroTerre data workflows. These Land use and Soil workflows combine custom or National Geodatabases (GDB) with the GI web application user's GI designs. The workflow results are bundled and shared with the HydroShare GI-RHESSys Jupyter notebook. The notebook uses EchoHydroLib and RHESSys workflows to process GIS datasets, prepare RHESSys input files, and execute the RHESSys hydrological model. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

When users open the GI notebook via HydroShare, the user is accessing a cloud-based compute environment. Dedicated compute nodes use OpenStack (2017), an open-source cloud computing framework, with JupyterHub (Jupyter Team, 2017a,b) to launch Jupyter notebooks. The purpose of JupyterHub is to manage authentication and spawn single-user servers on demand for many users on a compute environment (Jupyter Team, 2017a,b). Each user gets a complete Jupyter notebook server to execute their data and modeling processes. Effectively, the user has their own interactive Linux based compute environment instance for hydrological modeling using a web browser.

Embedded inside the GI-RHESSys Jupyter notebook (maintained by CUASHI), the GI web application (hosted at PSU) is accessible using an HTML IFrame object and JavaScript code. The user creates GI designs as explained in Sections 4 and 5 within the embedded GI web application. Using Representational State Transfer RESTful web services, the GI design is processed using GI-RHESSys-HydroTerre workflows at PSU (Section 3.3) with either National or custom data products. HTML web links pointing to the location of the GI-RHESSys-HydroTerre data bundles are shared with the Jupyter notebook. Within the GI-RHESSys Jupyter notebook, the Python based EcoHydroLib (Miles and Band, 2017a) and RHESSys workflows (Miles and Band, 2017b) consume the GI data bundles, data from federal agencies (i.e. USGS and USDA), and with custom data as HydroShare resources (i.e. Dead Run) to generate a RHESSys hydrological model using the user's GI landscape design. Generating a RHESSys hydrological model within the Jupyter notebook is explained in Section 3.4.

3.3. GI-RHESSys-hydroterre data workflows

Once the user is ready to evaluate their GI landscape design using the GI-RHESSys Jupyter notebook, the design geometry and RHESSys attributes are sent using RESTful services to the GI-RHESSys-HydroTerre data workflows. The GI web application provides the GI geometry. The main steps within the GI-RHESSys-HydroTerre land use data workflow include (1) converting GI design polygon geometry to a raster dataset; (2) assign user defined land use values to the new raster dataset; (3) merge the designed raster dataset with either custom land cover datasets or with a watershed scale clipped National Landcover dataset.

The second GI-RHESSys-HydroTerre data workflow adjusts soil types outlined by the user with polygon shapes. Within the GI web application, users assign nine RHESSys soil properties (Table 1) to the soil geometry. The main steps within the GI-RHESSys-HydroTerre soil data workflow include (1) erase GI design polygon geometry from custom soil dataset or watershed scale clipped national soil datasets (i.e. (SSURGO, 2011; STATSGO, 2011)); (2) merge the erased dataset with the user defined soil dataset.

With both the land use and soil workflows, the resultant datasets are zipped and a message is sent to the GI web application using a HTML callback, indicating where the data bundle zip files are located. These messages are kept as global variables within the GI web application for access by the GI-RHESSys Jupyter notebook environment and RHESSys workflows.

Table 1
The Soil and Land use GI-RHESSys-HydroTerre data workflows require the following attributes. The Soils workflow requires parameters for the pedotransfer functions. Both workflows require geometry descriptions and RHESSys unique parameter identifications.

GI-RHESSys-HydroTerre Soil Data Workflow Input Requirements		
Name	Type	Description
Value	int	RHESSys Soil Identification
Ksat	double	Saturated Hydraulic Conductivity
PctClay	double	Clay Percentage
PctSilt	double	Silt Percentage

HydroTerre data workflows within HydroTerre compute resources. The GI geometry (point and polygon) is converted to HydroTerre's projection coordinate system (Albers Conical Equal Area). Table 1 summarizes the input variables required to execute the GI-RHESSys-HydroTerre data workflows.

There are two GI-RHESSys-HydroTerre data workflows that transform the user's GI design into new datasets within the watershed boundary extent (Fig. 2). The first workflow modifies land use datasets delineated by the user with polygon boundaries. Within the GI web application, users assign pre-defined RHESSys land use classifications (i.e. vegetation type and surface imperviousness) to the land use

PctSand	double	Sand Percentage
Porosity	double	Porosity
FieldCap	double	Field Capacity
AvlWatCap	double	Plant available water capacity
DrnWatCont	double	Drainage water capacity
Geometry	polygon	Shape of user defined GI

GI-RHESSys-HydroTerre Land Use Data Workflow Input Requirements

Name	Type	Description
Value	int	RHESSys Land Use Identification
Geometry	polygon	Shape of User Defined GI

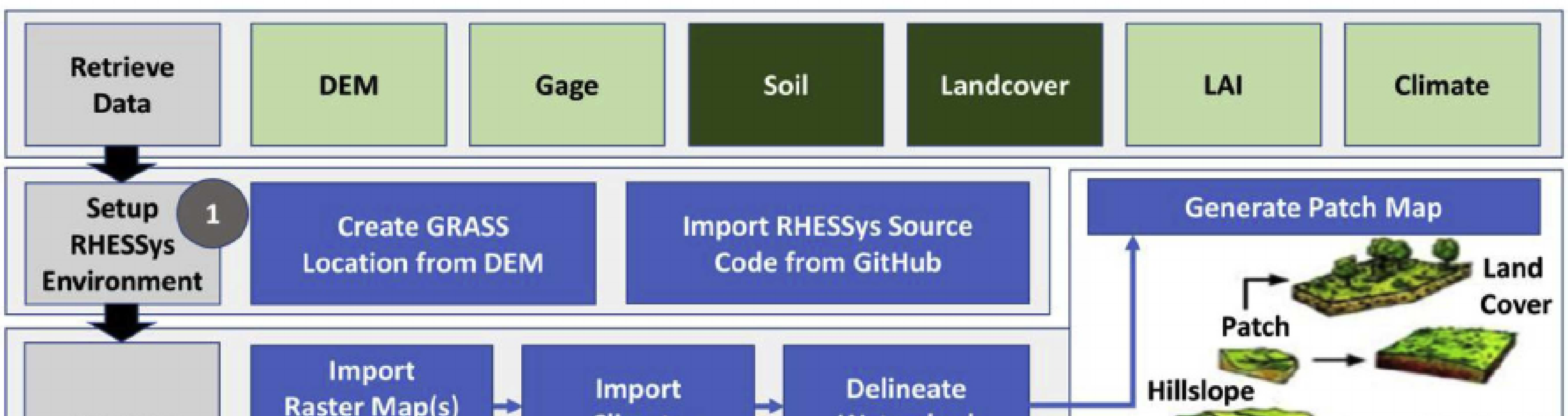
3.4. EcoHydroLib and RHESSys model workflows

The last components to evaluate GI designs with RHESSys hydrological models within a GI-RHESSys Jupyter notebook are to provide base model data required for the initial RHESSys model setup not generated by the GI-RHESSys-HydroTerre data workflows and prepare the data for RHESSys. In Fig. 3, the dark green polygons indicate data sets provided by the GI-RHESSys-HydroTerre data workflows, while the light green polygons designate the missing datasets (elevation, vegetation leaf area index, soils, and climate) required by RHESSys. Leaf Area Index (LAI) is a primary ecosystem variable providing the surface area exchanging water, carbon and energy with the atmosphere, and needs to be assigned appropriate to vegetation type and structure. Within the GI-RHESSys Jupyter notebooks, the user can assign HydroShare resources to these missing datasets, upload their own datasets, or they can use EcoHydroLib workflow scripts to retrieve data from federal agencies (i.e. USDA, USGS, NASA). To use the EcoHydroLib workflow scripts, the user needs to define the bounding box or extent of the catchment that he or she is studying. Alternatively, the user can specify a National Hydrography Datasets (NHD) streamflow gage to select upstream reaches to determine the catchment extent.

After the user has retrieved RHESSys input datasets in the GI-RHESSys Jupyter notebook, she or he needs to prepare the RHESSys model within the HydroShare compute environment. RHESSys uses a set of parameter files. The worldfile describes the hierarchical structure of the watershed (basin-hillslope-patch-vegetation layer) containing the full set of initial state variables and parameters (terrain, soils, vegetation). The flowtable stores a flow connectivity network linking patches within hillslopes to the stream network. The RHESSysWorkflows Python tool suite has been designed to simplify these steps required to build the model parameter files with three general categories. The first category is to setup the RHESSys environment by importing the RHESSys source code via GitHub (RHESSys, 2017) and compiling the latest stable release of RHESSys on the Jupyter Linux compute environment. Furthermore, RHESSys requires that all spatial data be stored in a Geographic Resources Analysis Support System (GRASS) (Neteler et al., 2012) Geographic Information System (GIS) mapset and this is achieved by creating a GRASS location using the Digital Elevation Model (DEM).

The second category processes the watershed data using the GRASS mapset by importing and registering raster datasets (i.e. landcover, vegetation type and LAI) using EcoHydroLib workflows. Due to the variety of climate data formats, the user is responsible for supplying climate data, either by uploading data, using an existing Hydroshare resource, or using the North American Land Data Assimilation System (NLDAS) (NLDAS, 2011) climate data supplied from the GI-RHESSys -HydroTerre data workflows (L. N. Leonard and Band, 2017). If the user has specified a streamflow gage, the DelineateWatershed tool delineates the watershed. Delineation is required so the watershed stream network be extracted, and the watershed be hierarchically partitioned (Fig. 3 illustrates hierarchy) into hillslopes draining into each stream reach and ecosystem patches using the GeneratePatchMap tool (Miles and Band, 2017c). After generating patches, soil properties are assigned to each patch using the GenerateSoilTextureMap function. The RegisterLandcoverReclassRules tool generates reclassification rules required to assign landcover properties to patches with the GenerateLandcoverMaps software package (Miles and Band, 2017c).

The last category generates the RHESSys input parameter files (world and flowtable) based on the GRASS GIS mapset. The first step is to use the GenerateWorldTemplate package to create the initial state of the world (the full domain, both spatially and temporally, within the watershed dynamics) by assigning default values from data stored in the GRASS datasets. Next, the CreateWorldfile tool creates the world file using the template and fills out all required parameters for each hillslope and patch based on registered spatial data. The CreateFlowtable function creates a flow table to describe connectivity between the watershed patches and whether to include roads and rooftops (optional) with the watershed connectivity description. The RunLAIRead tool initializes vegetation carbon stores in the worldfile and the remaining step is to create a Temporal Event Controller (TEC) file to change RHESSys land use parameters at certain times, for example to simulate fire, harvest, urban and road developments, and to control model output generation. After finishing all these steps, it is now possible to modify the parameter sets using the user's GI landscape



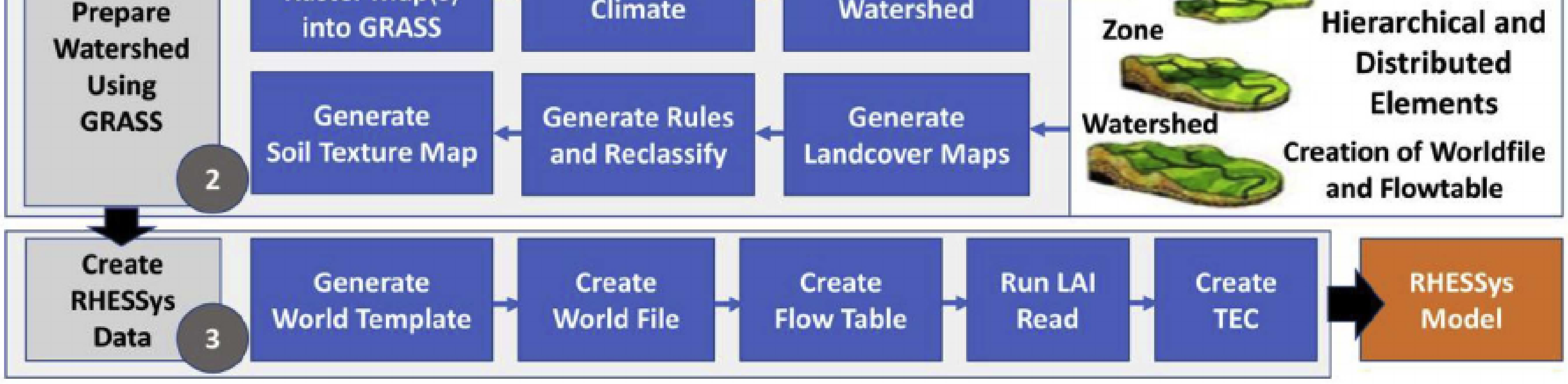


Fig. 3. To implement a RHESSys hydrological model using the GI-RHESSys Jupyter notebook, datasets must be retrieved first. The dark green polygons represent data provided by using the GI web application and GI-RHESSys-HydroTerre data workflows. The light green polygons represent data provided by the EcoHydroLib workflow scripts, user supplied datasets, or using HydroShare resources. (1) The RHESSysWorkflows Python tool suite sets up the RHESSys environment, (2) prepares the watershed datasets using GRASS, and (3) creates the required RHESSys input files. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

design from the GI web application tool. The next section focuses on the GI web application interface and user steps to create a GI landscape design.

4. GI web application

Section 3 explained the back-end components utilized by both the GI web application and GI-RHESSys Jupyter notebook using multiple data and model workflows. Section 4 focuses on the GI web application by explaining the Graphical User Interface (GUI) and visualization components for users to create GI designs that are processed by the back-end workflows and to be used by GI-RHESSys Jupyter notebooks. Section 4.1 summarizes the main components of the GI web application. Section 4.2 explains the GUI layout and users' steps to create GI designs. Section 4.3 explains the technical details of how the GI web application exchanges data with the GI-RHESSys Jupyter notebook demonstrated in Section 5.

4.1. GI web application components

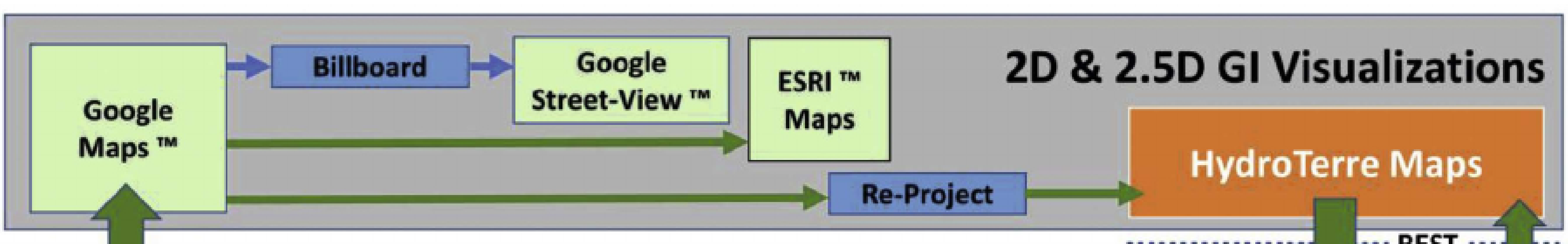
The components to create GI features using the GI web application is summarized in Fig. 4. There are three key components to the GI web application. The first component is the GI material dictionaries that allow users to add, delete, and edit GI features. The second component uses these GI features in 2D and 2.5D visualizations and are used to assist users in placing GI and understand the visual impact of their landscape design. The last element is GI tools that aid users to create or edit their own GI features, share GI designs, and estimate costs.

Fig. 4, Section 3 (blue dashed line) explained the components for generating data required to evaluate GI designs using multiple workflows that are retrieved within HydroShares GI-RHESSys Jupyter notebook interactive compute environment. To initiate the GI-RHESSys-HydroTerre data workflows, the GI web application is required to create the data structures shown in Table 1. These data structures are retrieved sfrom the GI material dictionaries, defined by users via the web interface. Table 2 summarizes the tree and surface GI type attributes stored in these material dictionaries. Attributes highlighted in green are identical to those in Table 1. Attributes highlighted in pink are used to estimate the costs for planting GI trees (small, medium, or large with estimated labor costs) and GI surface materials such as soil are calculated by volume (depth is an attribute, and area is calculated with GI geometry) and estimated labor costs. These costs are defined by the user and are expected to be location dependent. Attributes highlighted in yellow are for object descriptions with text and icons within the GI visualization tools. The remaining attributes in Table 2 are for creating unique keys within the material dictionaries. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The remaining data type required in Table 1 (see Section 3) is GI geometry. There are two GI geometry types; point and polygon. Point geometry represents the center positions of trees and shrubs. As users add GI point materials to the map visualizations, icons marked with sizes (S = Small, M = Medium, L = Large) are placed on the maps using latitude and longitude positions. Furthermore, the tree attribute (highlighted in orange) is used to create circle geometry to represent the crown shape. Buffering is necessary to assign land use area changes. Polygon geometry is used to delineate surface GI types such as soils, taller grasses, lawn, rain gardens, and impervious surfaces.

To add, delete, and edit GI features, users rely on the 2D and 2.5D visualizations to place GI in suitable locations. Google Maps™ is the primary map for adding features at the map view. As users add tree GI features, billboards (Décoret et al., 2003) that are called 2.5D geometry are placed within the Google Street View™. Billboard sizes are dependent on the user's current tree size preference (Small, Medium, Large). Users are able to edit or delete GI features in either Google™ visualization tool. We note that at this time perspective visualization is limited to the area visible within Google Street View, but ongoing work will incorporate point cloud generated views. Beyond the street view. To provide additional datasets to aid users in placing GI, for example terrain topography, ESRI Basemaps™ services are also available. The remaining map visualization component is a GI-RHESSys-HydroTerre data workflow that provides national Essential Terrestrial Variables (ETV) and custom datasets from HydroTerre. The Web Mercator projection and coordinate system used by Google Maps™ differs from the HydroTerre systems, therefore GI features are re-projected to Albers Conical Equal Area.

The remaining web application design categories are GI tools dedicated to enriching GI features and estimate costs to implement the



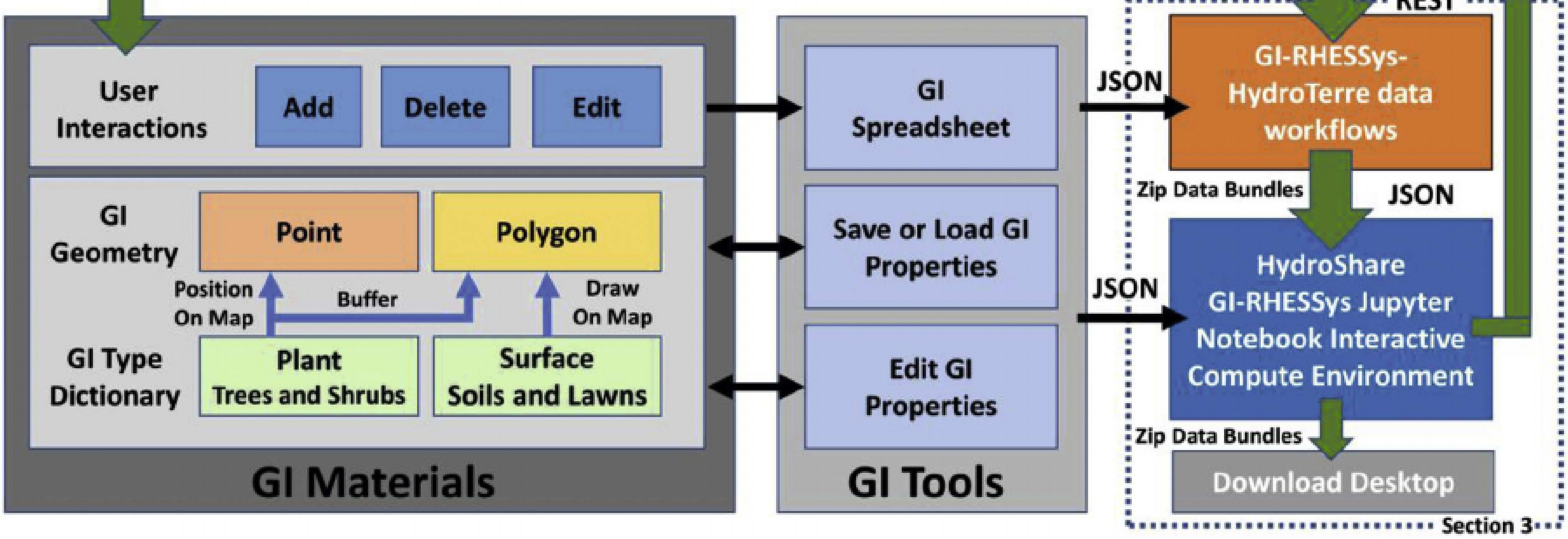


Fig. 4. The GI Web application components. Users add, delete, and edit GI materials derived from GI dictionaries that are used to store RHESSys attributes and GI geometry. Map visualizations are used to place and display the location of GI geometry. GI tools provide users the ability to estimate costs, edit and create custom GI. Restful services and Java Script Object Notation (JSON) are used to share data attributes to the GI-RHESSys-HydroTerre data workflows and connect with HydroShare's GI-RHESSys Jupyter notebook interactive compute environment.

user's GI landscape design. As the user adds, edits, or removes GI features from the map views, the GI inventory are coordinated with the "GI Spreadsheet" view. For example, when the user replaces soil types in front of a home owner's property to retain and slow surface water movement, the volume of material is calculated, and the total cost of labor and materials is added to the "GI Spreadsheet". All GI features, such as tree sizes and material costs, can be customized by the GI web application user to meet his or her local region differences. For example, material costs in city environments are often higher than outer urban environments. Another example is when labor costs are zero, as the home owner is doing the project themselves, or by volunteers. The "Edit GI Property" tool enables web users to add their own GI features that meet the GI attributes described in Table 2. Once the user has finished her or his GI design, the design can be saved to the user's desktop for sharing or to load back to the GI web application at another time. Otherwise, users' GI designs are not stored and are lost when users close the browser.

4.2. Using the web application to create GI landscape designs

The GI web application available here (web link [3]) as an independent service, or here (web link [4]) embedded within the GI-RHESSys Jupyter notebook is shown in (Fig. 5). The overarching strategy in designing the GI web application GUI layout was to split the browser screen into two large panels to aid and simplify user creation of GI designs. The left panel is designated to place GI in a 2D map environment and the right panel is dedicated to visualizing GI in a 2.5D environment with supporting GI tools.

To show how to create GI with the GI web application, the Dead-Run catchment, located in Baltimore County, Maryland, USA, is demonstrated using custom three-meter high-resolution datasets. The watershed scale GI design has been prepared and uploaded as geodatabase layers. Further information about these designs is available from (Miles and Band, 2015a). Other locations within the CONUS are supported, although at lower resolutions of 10 or 30 m constrained by national dataset properties unless additional local information is available. We note that high resolution lidar terrain data and land cover are currently available or being developed by many municipalities, and the Chesapeake Bay Consortium has developed these data for the full Chesapeake Bay Watershed (http://chesapeakeconservancy.org). In the Dead Run example, the user specifies the Dead-Run watershed location (that is supported with a pre-defined GI designed database) by picking depth with the camera perspective and object occlusion.

As the user adds GI features, quantities with estimated costs are automatically added to the "GI Spreadsheet" located in Fig. 5I and shown in Fig. 6A. For users to modify the predefined GI properties such as costs that are kept in the GI dictionaries as discussed in Section 3.1, the "Edit GI Properties" tool located in Fig. 5J and displayed in Fig. 6B is available to modify the existing GI or to create new custom GI. To share existing and custom GI properties with other users, or to save GI designs to user's local desktop for future use, the "Save or Load GI" tool located in Fig. 5K and presented in Fig. 6C is offered to users. Once users have finished their GI landscape design, the GI-RHESSys-HydroTerre workflows described in Section 3, creates the GI datasets, when users interact with the tools found by clicking the right panel tab at Fig. 5L. The next section focuses on how the GI web application connects these GI datasets with the GI-RHESSys Jupyter notebook.

4.3. Connecting GI designs with the GI-RHESSys jupyter notebook environment

Once the user has completed a GI landscape design, he or she can download the resultant data bundles from GI-RHESSys-HydroTerre data workflows to their desktop by clicking on the download buttons shown in Fig. 6D. Within the GI-RHESSys Jupyter notebook environment, the user has the same capability, however the process is cumbersome. Recall, the HydroShare GI-RHESSys Jupyter notebook environment is a web application providing access to a remote Linux cloud compute environment. When users click on the download buttons, the browser will direct the zip file to the user's own compute environment, not the remote Linux environment. Hence, users would need to upload the zip files to the HydroShare GI-RHESSys Jupyter notebook environment requiring not only extra steps but be comfortable with the Linux compute environment.

To avoid these extra steps, browser window event listeners were added to the GI-RHESSys Jupyter notebook to receive messages from the GI web application summarized in Table 3. There are three general groups to post messages from the GI web application. The first group contains the GI material dictionaries (with default and any user custom GI) as described in Section 4.1, accordingly the GI attributes (Table 2) can be analyzed within the GI-RHESSys Jupyter notebook. The second group provides access to the GI geometry designed by the user. Both the first and second groups enable access to GI data useful for models that do not use spatial GIS datasets. For example, the EPA's SWMM uses routing, through a system of digitized drainage lines or pipes (Gironas the location marked at (Fig. 5A) and the web visualizations will zoom to this location. Users can also use the Google Pegman™ (Fig. 5B) to change locations with all maps and the Google StreetView™ will zoom to the Pegman location. To pan (hand icon) the Google Map™, add GI point features (balloon icon), or create GI polygon surfaces (shape icon), the user picks from the toolbar located at (Fig. 5C). Users specify GI properties such as size and depth at Fig. 5D that are assigned to GI materials chosen by the menu located at Fig. 5E. For example, to place a small Honey Locust tree, the user clicks on small tree size (Fig. 5D), specifies the tree from the menu list (Fig. 5E), clicks on the shape icon (Fig. 5C) and the mouse cursor in the left panel changes from the hand icon to a cross icon. Now where the user clicks on the map (left panel), a tree icon marked with S for small is placed on the map (Fig. 5F). Depending on the camera view in Google StreetView™ (right panel) the tree billboard will appear immediately (Fig. 5G), otherwise the user can rotate the camera view to see the tree in a perspective view. To create surface GI, users follow similar steps by specifying feature (soil) depth (Fig. 5D), selecting the shape icon (Fig. 5C) and clicking on the map multiple times to define the polygon vertices (Fig. 5H). However, surface GI are not added to the Google StreetView™ due to the limitation of

[3] https://hydroterremodels.psu.edu/RHESSYS/GI_Designer/GI.html.
[4] https://hydroterremodels.psu.edu/RHESSYS/GI_Designer/GI_HS.html.

et al., 2010) and not derived from the elevation model that models such as RHESSys use. Research by (Heidari et al., 2018) use these GI JavaScript Object Notation (JSON) data structures with SWMM and Jupyter notebooks to analyze trees and rain gardens described elsewhere. The last group provides Uniform Resource Locator (URL) locations of the workflow results. By using wget (Free Software Foundation, 2015), the data bundle results are downloaded within the GI-RHESSys Jupyter notebook environment. Section 5 demonstrates these GI capabilities by executing a RHESSys model at Dead-Run watershed with the GI-RHESSys Jupyter notebook compute environment.

5. Demonstration of using landscape designs within GI-RHESSys jupyter notebook

The previous sections have discussed the GI web application for users to develop GI landscape designs and incorporate these designs with datasets required by RHESSys using multiple workflows. This section focuses on demonstrating the GI web application within the GI-RHESSys Jupyter notebook compute environment to develop a RHESSys model using the GI landscape design at Dead Run watershed from Section 4. Section 5.1 provides an overview of how users access the GI-RHESSys Jupyter notebook and the last section demonstrates using the GI-RHESSys Jupyter notebook to develop a RHESSys

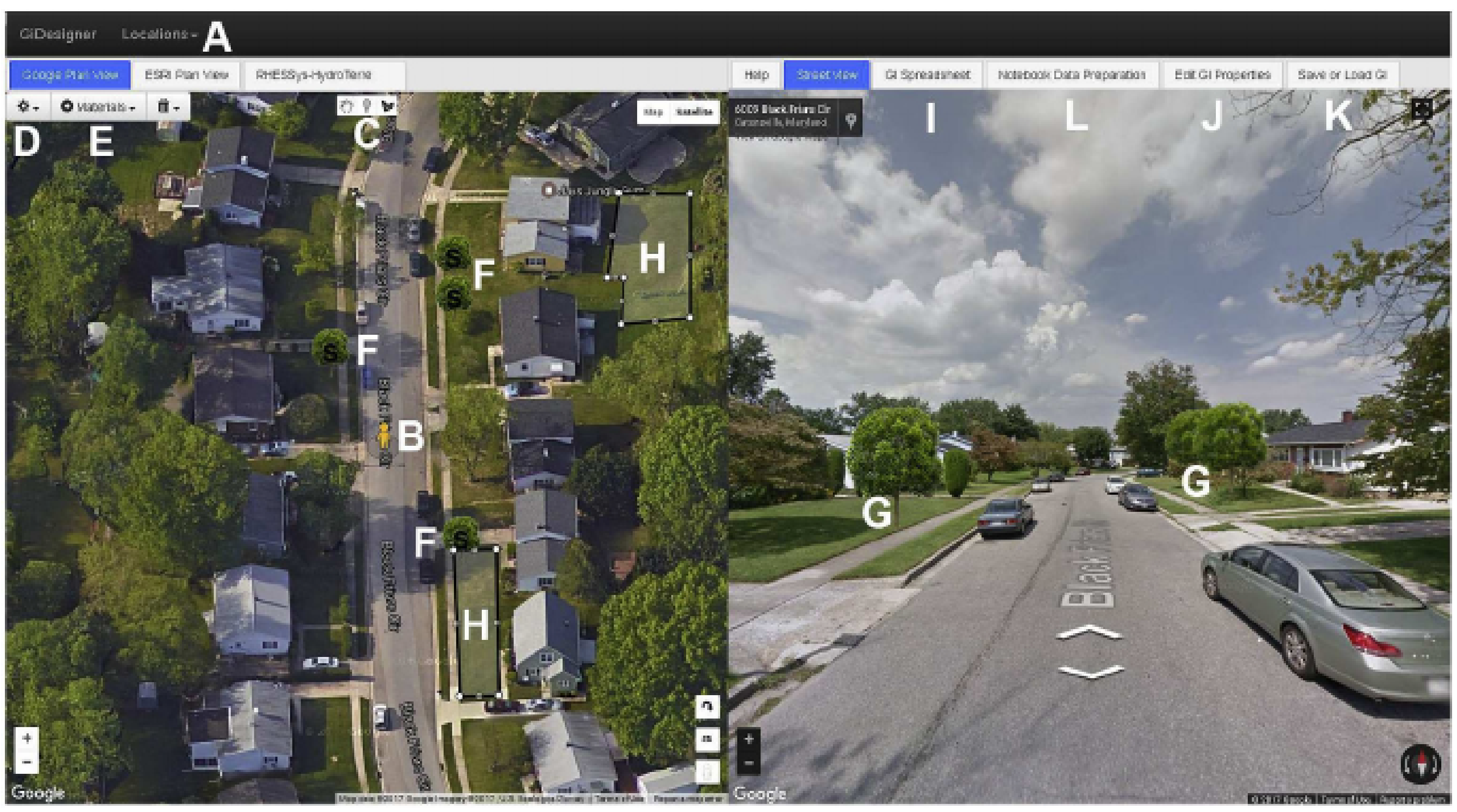


Fig. 5. The GI web application Graphical User Interface (GUI). The left panel is used to place GI materials. The right panel is for displaying 2.5D plant materials and tools for estimating costs, editing, and creating new GI materials. (A) To change the map location to custom datasets. (B) Pegman location to change map and street-view locations. (C) To pan and create GI features on map. (D) GUI location to specify GI sizes. (E) GUI location to select GI materials. (F) User placed small trees on map. (G) User placed small trees in street-view. (H) User placed surface materials on map. (I) GUI location to load GI Spreadsheet. (J) GUI location to edit and create new GI materials. (K) GUI location to save or load GI user designs. (L) GUI location to execute GI-RHESSys-HydroTerre data workflows and retrieve GI data bundles.

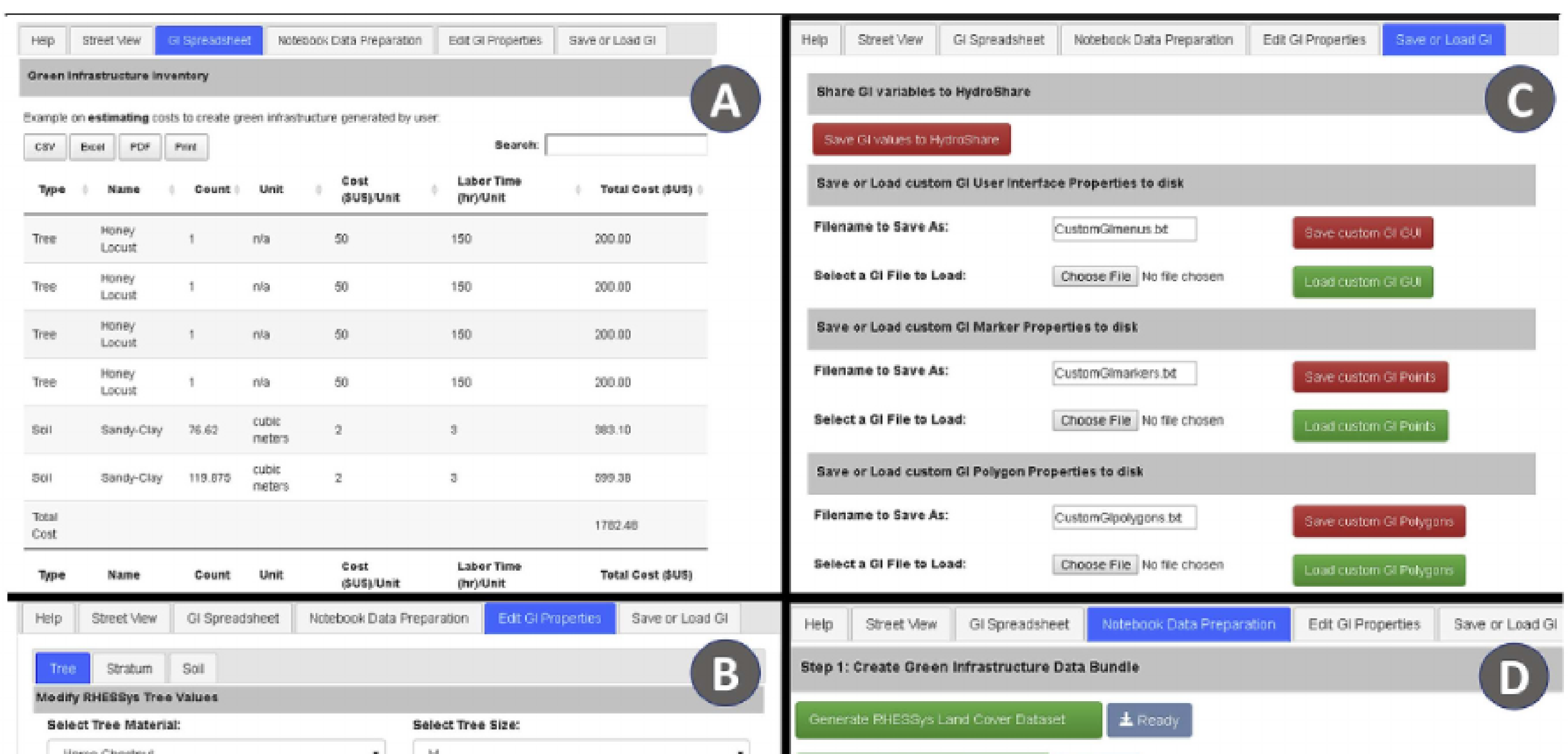


Fig. 6. GI web application tools. (A) Spreadsheet with estimated GI costs from user landscape design. (B) Tool to add new and edit existing GI features. (C) Tools to save and load GI user landscape designs and materials. (D) Tools to execute GI-RHESSys-HydroTerre data workflows and retrieve GI data bundles.

hydrological model at Dead Run watershed.

5.1. *Accessing GI-RHESSys jupyter notebooks*

There are two ways to access GI-RHESSys Jupyter notebooks dependent on the user's confidence with computational environments. The first method is to download the notebooks directly from HydroShare's GitHub repository (web link [5]), where official notebooks supported by HydroShare are located with dependencies (i.e. libraries). Additionally, users can access the latest developer versions of GI-RHESSys Jupyter notebooks using GitHub from here (L. Leonard and Band, 2017).

The second method is to visit the HydroShare website at https://www.hydroshare.org/. Users are required to have an account with HydroShare to use these resources. Sign up is free. Once the user has logged into HydroShare, click on "discover" at the top menu bar. A keyword search of "RHESSys" will return a list of resources available from HydroShare. Find and click on the resource titled "Green Infrastructure Designer with RHESSys Workflow". A web page document appears describing details when the resource was created and last updated, an abstract, subject keywords, how to cite the resource, and a Zipped BagIt archive (Boyko et al., 2014) containing the GI-RHESSys Jupyter notebook. Users can download the GI-RHESSys Jupyter notebook to their desktop using the Zipped BagIt archive if desired.

The main advantage of using the second method to access the GI-RHESSys Jupyter notebook is the ability to use the "Open with" button (located top right of the web page document) to open the notebook. Depending on the user's credentials, the "Open with" button will provide a list of different compute environments. To use the GI-RHESSys Jupyter notebook, click on a JupyterHub. Assuming the user has valid credentials, a new web page is loaded displaying a Jupyter notebook within the JupyterHub compute environment. A welcome page appears

[5] https://github.com/hydroshare/hydroshare-jupyterhub/tree/master/notebooks.

Table 2
Plant and Surface GI attributes storied in GI dictionaries. Orange cells indicate tree attributes. Yellow cells are for interface web tools. Pink cells are used to estimate costs. Green cells are for GI-RHESSys-HydroTerre data workflows.

Plant GI Attributes		
Name	**Type**	**Description**
Type	string	Species Name
Plant_id	string	Unique Name used in GI dictionary
Radius	double	Tree or shrub crown radius size
Pretty_name	string	Name used in tooltips and graphics
Tree_Icon	string	Graphic for Google Street View™
Tree_map_icon	string	Graphic used to represent species on maps
Size	string	Tree or Shrub Sizes (Small, Medium, Large)
Cost	double	Estimated GI Cost in USD
Labor_cost	double	Estimated Labor Cost in USD
Landuse	int	RHESSys Land use Identification
Surface GI Attributes		
Name	**Type**	**Description**
Surface_name	string	Unique Name used in GI dictionary
Pretty_name	string	Name used in tooltips and graphics
Depth_Size	double	Soil depths (Small, Medium, Large)
Depth_value	double	Value assign Soil Depth
Cost	double	Estimated GI Cost in USD
Labor_cost	double	Estimated Labor Cost in USD
Surface_id	int	RHESSys Surface Identification
Ksat	double	Saturated Hydraulic Conductivity
Clay	double	Clay Percentage
Silt	double	Silt Percentage
Sand	double	Sand Percentage
Porosity	double	Porosity
FieldCap	double	Field Capacity
AvlWatCap	double	Plant available water capacity
DrnWatCont	double	Drainage water capacity

(Fig. 7 Step 5) for the RHESSys model. Assuming there were no errors reported, the RHESSys model is ready to be executed in the GI-RHESSys Jupyter notebook compute environment as shown in Fig. 7 Step 6. There is documentation on what the command options are for the "RunModel" tool that prepares and executes the RHESSys model (Fig. 7A).

The RHESSys model using the "JupyterHub" compute environment took 45 min to finish. The main reason the model compute time took a long time for a short simulation period (the default value of 6 months defined in the RHESSysWorkflow Python object which users can adjust) is the high-resolution three-meter datasets used for defining the watershed properties and the resources allocated at the cloud site. The last GI-RHESSys Jupyter notebook code cell generates visualizations for RHESSys variables as shown in Fig. 7B.

After executing all the notebook steps, the user has generated a GI landscape design, prepared data for the watershed location, and created a RHESSys hydrological model. Next steps for using the GI-RHESSys Jupyter notebook include the development of workflows for calibrating the model parameters (Fig. 7C) with the "RunModel" tool, trying different datasets from local or federal agencies, or datasets from other HydroShare resources. The GI-RHESSys Jupyter notebook is easy to edit or clone for other research goals and a collection of different RHESSys Jupyter notebooks are available here (L. Leonard and Band, 2017) with different data inputs and user skill level requirements. After users have evaluated the GI-RHESSys Jupyter notebook for the first time, the learning curve to compare GI designs has been significantly reduced. The authors encourage users to focus on creating ideal GI landscape designs with quantitative analysis with the RHESSys ecohydrological model and share their data and model results within the HydroShare community as resources and public outreach.

Table 3
The GI web application posts the following data structures as messages to the GI-RHESSys Jupyter notebook application. Event Listeners in the notebook are used to download zipped data bundles (highlighted in pink), GI dictionaries (highlighted in green), and the user's GI geometry landscape design (highlighted in orange).

Key	Type	Description
soil_download_url	URL	Location to download GI-RHESSys soil workflow data bundle.

landuse_download_url	URL	Location to download GI-RHESSys land use workflow data bundle.
hs_tree_dictionary	JSON	Dictionary of plant (i.e. trees or shrubs) GI attributes
hs_stratum_dictionary	JSON	Dictionary of surface (i.e.lawns) GI attributes
hs_soil_dictionary	JSON	Dictionary of soil GI attributes
hs_point_markers	JSON	List of point geometry locations (i.e. Tree and Shrubs)
hs_polygon_markers	JSON	List of polygon geometry locations (i.e. Soils and Lawns)

with instructions on how to open and use notebooks.

For users that downloaded GI-RHESSys Jupyter notebooks from GitHub or used the Zipped BagIt method, it is feasible to upload Jupyter notebooks within HydroShare's JupyterHub environment manually by using the Jupyter tree (click on Jupyter icon) to upload files in the users' compute workspace. Users using their own JupyterHub environment will require dependencies such as the EcoHydroLib and RHESSys workflows python tools (Section 3) be installed. Instructions for installation are here (Miles and Band, 2017c). The HydroShare JupyterHub compute environment has all the RHESSys dependencies preinstalled for all users. The next section discusses how to open and use the GI-RHESSys Jupyter notebook.

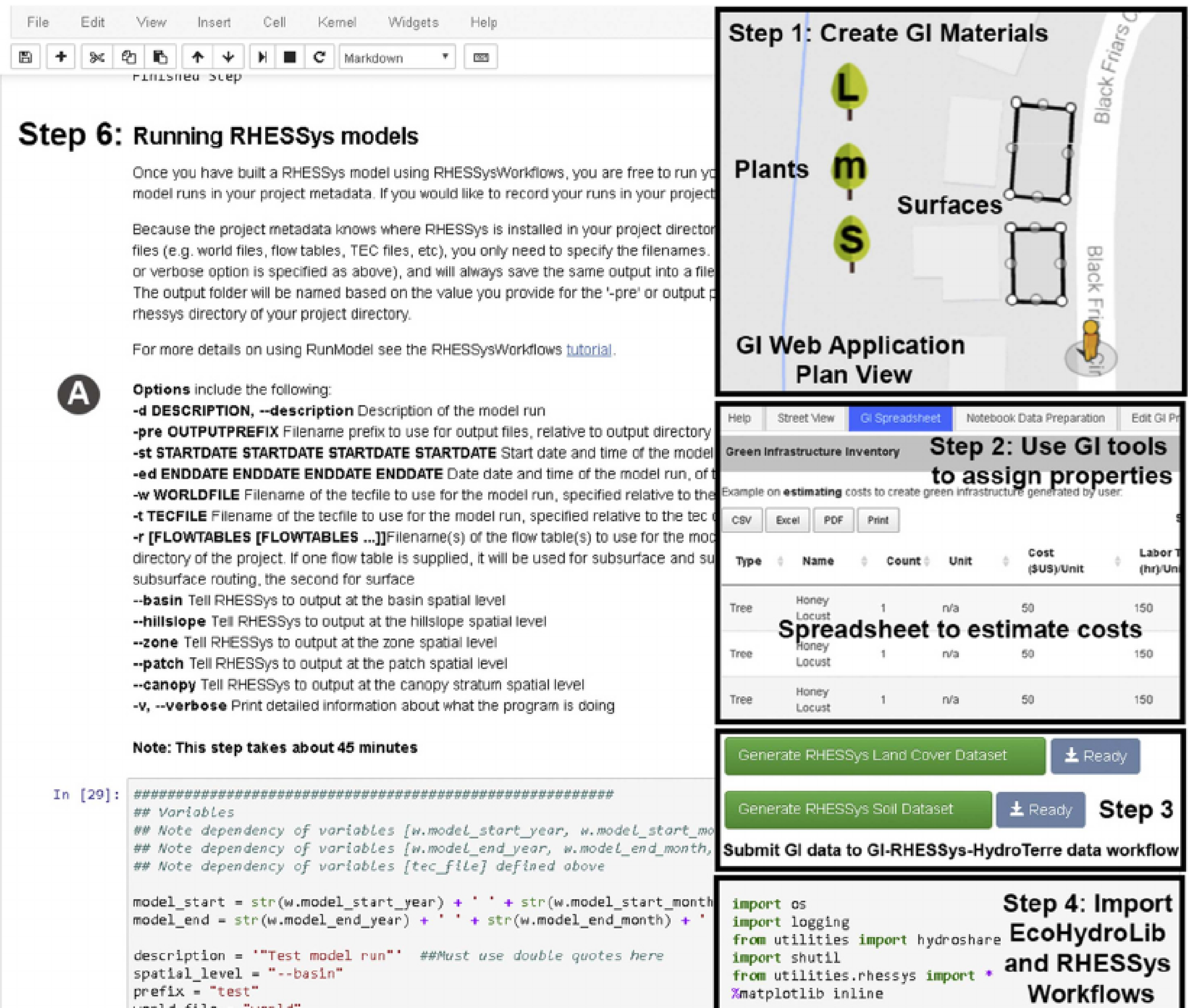
5.2. Using a jupyter notebook to create a RHESSys model with GI designs

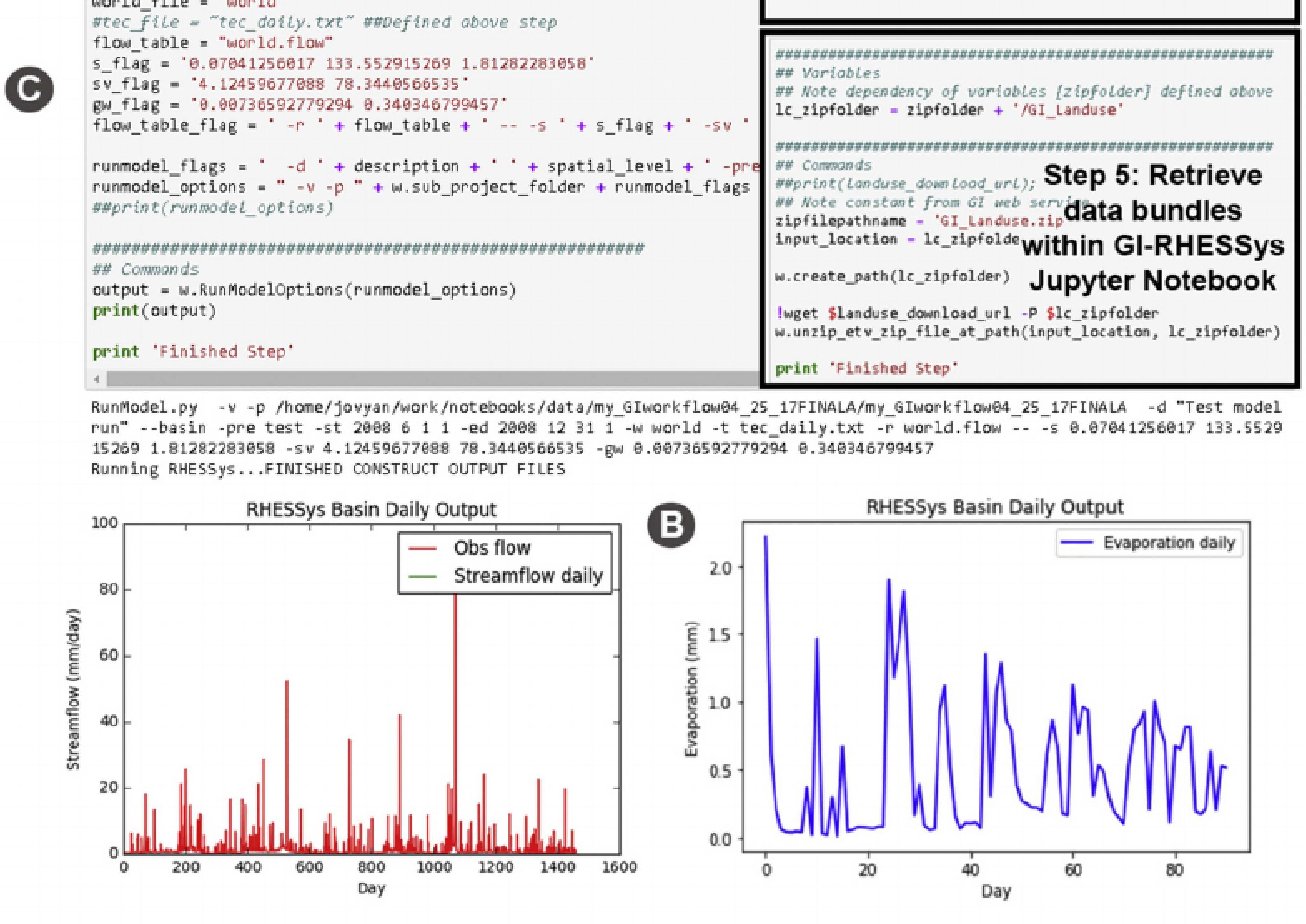
The supplementary material provides the steps in opening and using the GI-RHESSys Jupyter notebook to create GI landscape designs and a RHESSys model. Briefly, the user creates a GI landscape design (Fig. 7 Steps 1–3) using the GI web application explained in Sections 4.2 and 4.3. Next, the user is ready to evaluate their GI landscape design in the GI-RHESSys Jupyter notebook environment. Users need to import the EchoHydroLib and RHESSys workflows (Section 3.4) into the GI-RHESSys Jupyter notebook (Fig. 7 Step 4) and prepare the GI data

6. Conclusion

The provision and improvement of ecosystem services in urban areas depends on sound designs that integrate an understanding and forecast of the interactions of hydrologic, ecological, climate and social processes. Successful approaches to environmental restoration of the built environment are also increasingly seen as requiring some level of participation and input from residents and other stakeholders. Community participation in the use of advanced environmental models used to assess the benefits of different urban restoration designs requires both visualization and access or input to computational systems that are usually not available to non-specialists. Our approach demonstrates how GI web applications can rapidly visualize GI landscape designs at the site scale for participatory planning between stormwater engineers and property owners in urban environments with pre-defined and custom materials, and automate the translation of user developed landscape designs to parameterize complex environmental models.

We demonstrated these capabilities at the Dead Run watershed in Baltimore, Maryland, USA, a study site of the Baltimore Long Term Ecological Research Site (http://beslter.org/), combining GI landscape designs at site locations with watershed scale designs. User GI landscape designs are then transformed and integrated with custom and national datasets using GI-RHESSys-HydroTerre data workflows. The





(caption on next page)

Fig. 7. Steps 1 and 2 summarize the main steps to prepare GI materials for the GI-RHESSys Jupyter notebook. See Sections 4.1, 4.2 and the supplementary material for further details. Step 3 submits the user GI design to the GI-RHESSys-HydroTerre data workflows. See Section 4.3 for details about how the connection with the data bundles operates with the GI-RHESSys Jupyter notebook. See Section 3.3 about workflow operations and processes in the backend. Step 4 imports the EchoHydroLib and RHESSys workflows (Section 3.4) into the GI-RHESSys Jupyter notebook. Step 5 retrieves the GI data bundles to be accessible by the GI-RHESSys Jupyter notebook as explain in Section 4.3. Step 6, after the users have prepared the RHESSys input datasets using the RHESSys workflows (Section 3.4), RHESSys is ready to be executed using HydroShare's JupyterHub compute environment. (A) Example documentation explaining the RHESSys workflow tools. (B) Visualization results from the RHESSys hydrological model using GI landscape designs from the embedded GI web application within the GI-RHESSys Jupyter notebook. (C) RHESSys model input variables and values for calibration.

data workflows generate data bundles and JSON data structures for hydrological modeling. Next, we demonstrated the GI web application embedded inside a GI-RHESSys Jupyter notebook hosted at HydroShare's cloud compute environment that consumes these data structures to create a RHESSys ecohydrological model with GI landscape designs. The documented RHESSys Jupyter notebooks serve as templates for beginner users to quickly learn how to prepare model inputs using EcoHydroLib and RHESSys workflows. Furthermore, these notebooks are easy to clone and edit, and the embedded workflows make the process for both beginner and expert users to rapidly create ecohydrological models at other watershed locations straightforward by using cloud compute resources to generate the RHESSys ecohydrological models.

7. Future directions

The next step with this research is to improve the visualization and data acquisition capabilities of the GI web application. This includes workflows for users to upload sewer infrastructure Computer Aided Designs and GI planning designs, improving the GI material dictionaries

- Awards 1148453 and 1148090, HydroShare: NSF Office of Cyberinfrastructure
- Award DEB 1027188, Long-Term Ecological Research program, Baltimore Ecosystem Study.
- Award DBI-1639145: SocioEconomic Synthesis Center, support for Water Science Software Institute.
- Award no. 1239678 EAGER: Collaborative Research: Interoperability Testbed-Assessing a Layered Architecture for Integration of Existing Capabilities
- Award no. 0940841 DataNet Federation Consortium.
- Award no. 1148090 Collaborative Research: SI2-SSI: An Interactive Software Infrastructure for Sustaining Collaborative Innovation in the Hydrologic Sciences

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2018.10.003.

References

with three dimensional trees and integrating Google StreetViews™ with
LiDAR (Light Detection and Ranging) captured topography and street infrastructure to create photo realistic GI landscape designs combined with simulated dynamics (i.e. vegetation growth and extreme events from climate and hydrological model results). Furthermore, new workflows and tools are necessary to validate GI landscape designs, such as placing trees in suitable locations and quantifying whether the GI landscape designs meet local and state regulations using automation and hydro-informatics, and placement relative to utilities and other infrastructure.

We see a need from local agencies in multiple cities to use these web applications and services to interactively evaluate GI landscape designs with the public and other professions for immediate feedback on GI policies and implementations. However, spatially distributed and continuous-time models like RHESSys are computationally expensive and time consuming. Our vision is that the RHESSys Jupyter notebooks using compute resources such as HydroShare's JupyterHub will serve as templates and resources for agencies to access model results similar to a reference library. As the data and models improve, agencies access the latest improved model results with services such as the GI web application, to evaluate proposed GI landscape designs and policies that are shared with other agencies and the community.

Credits

A number of scientists, engineers and other professionals associated with the Baltimore Ecosystem Study, and with stormwater utilities in Baltimore City and County, Portland, OR, Durham, NC, Chicago and Phoenix provided important feedback on conceptual design and initial versions of this software.

The research and development were a synthesis of information and methods developed by a set of NSF grants:

- Award no. 1331813 Collaborative Research: CyberSEES: Type 2: A New Framework for Crowd-Sourced Green Infrastructure Design, NSF CISE

Bakker, M., Miller, A.D., Morgan, L.K., Werner, A.D., 2017. Evaluation of analytic solutions for steady interface flow where the aquifer extends below the sea. J. Hydrol. 1–5. https://doi.org/10.1016/j.jhydrol.2017.04.009.

Bandaragoda, C., 2016. How to run a Landlab model in HydroShare. [WWW Document]. HydroShare. http://www.hydroshare.org/resource/25040a158eac4730b31eb5ebcc3a7339.

Bandaragoda, C., Castronova, A.M., Phuong, J., Strauch, R., Istanbulluoglu, E., Nudurupati, S., Tarboton, D.G., Yin, D., Wang, S., Barnhart, K., Tucker, G., Hutton, E., Hobley, D.E.J., Gasparini, N.M., Adams, J., 2017. Reproducible earth-surface modeling with Landlab on HydroShare. Hydro. https://doi.org/10.4211/hs.38002ee4bf594901a29055bdf20b13da.

Bell, C.D., Tague, C.L., McMillan, S.K., 2017. A model of hydrology and water quality for stormwater control measures. Environ. Model. Software 95, 29–47. https://doi.org/10.1016/j.envsoft.2017.05.007.

Beven, K.J., Kirby, M.J., 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol. Sci. Bull. 24, 43–69. https://doi.org/10.1080/02626667909491834.

Boyko, A., Kunze, J., Littman, L., Vargas, B., 2014. The BagIt file packaging format. V0.97) [WWW Document]. Netw. Work. Gr. Internet-Draft. https://tools.ietf.org/html/draft-kunze-bagit-14, Accessed date: 7 January 2017.

Castronova, A.M., 2017. JupyterHub terrain processing. HydroShare [WWW Document]. https://www.hydroshare.org/resource/68313c40c7ff48cd9e065ceada3c8426/.

Crawley, S., Ames, D., Li, Z., Tarboton, D., 2017. HydroShare GIS: visualizing spatial data in the cloud. Open Water J 4.

CUAHSI, 2017. Consortium of Universities for the Advancement of Hydrologic Science. [WWW Document]. https://www.cuahsi.org/, Accessed date: 7 January 2017.

DCMI, 2018. Dublin core metadata initiative. [WWW Document]. http://dublincore.org/.

Décoret, X., Durand, F., Sillion, F.X., Dorsey, J., 2003. Billboard clouds for extreme model simplification. ACM Trans. Graph. 22, 689–696. https://doi.org/10.1145/882262.882326.

ESRI, 2017. ESRI Basemaps. [WWW Document]. http://www.esri.com/data/basemaps, Accessed date: 1 July 2017.

Free Software Foundation, 2015. GNU Wget.

Gironas, J., Roesner, L.A., Rossman, L.A., Davis, J., 2010. A new applications manual for the storm water management model (SWMM). Environ. Model. Software 25, 813–814. https://doi.org/10.1016/j.envsoft.2009.11.009.

Google, 2017. Google maps APIs. [WWW Document]. https://developers.google.com/maps/, Accessed date: 1 July 2017.

Heidari, B., Leonard, L., Band, L.E., Minsker, B., 2018. IDEAS for GI_Deadrun. HydroShare WWW Document. http://www.hydroshare.org/resource/75dfc27015f1467388d712b562657835, Accessed date: 15 April 2018.

Horsburgh, J.S., Morsy, M.M., Castronova, A.M., Goodall, J.L., Gan, T., Yi, H., Stealey, M.J., Tarboton, D.G., 2015. Hydroshare: sharing diverse environmental data types and models as social objects with application to the hydrology domain. JAWRA J.

Am. Water Resour. Assoc n/a–n/a. https://doi.org/10.1111/1752-1688.12363.

HydroShare, 2018. Resource types. WWW Document]. CUAHSI. https://help.hydroshare.org/hydroshare-resource-types/.

Jefferson, A.J., Bhaskar, A.S., Hopkins, K.G., Fanelli, R., Avellaneda, P.M., McMillan, S.K., 2017. Stormwater management network effectiveness and implications for urban watershed function: a critical review. Hydrol. Process. 31, 4056–4080. https://doi.org/10.1002/hyp.11347.

Jupyter Team, 2017a. JupyterHub. WWW Document. https://jupyterhub.readthedocs.io/en/latest/, Accessed date: 7 January 2017.

Jupyter Team, 2017b. Jupyter. WWW Document. https://jupyter.org, Accessed date: 7 January 2017.

Leonard, L., Band, L.E., 2017. RHESSys-Jupyter-notebooks. WWW Document]. GitHub. https://github.com/leonard-psu/RHESSys-Jupyter-notebooks, Accessed date: 7 January 2017.

Leonard, L., Band, L.E., Lin, L., 2017. Green infrastructure designer with RHESSys workflow. HydroShare WWW Document. https://doi.org/10.4211/hs.https://doi.org/10.4211/hs.3f7680cf83dc426e858d5b48cb95a565.

Leonard, L., Duffy, C., 2016a. Visualization workflows for level-12 HUC scales: towards an expert system for watershed analysis in a distributed computing environment. Environ. Model. Software 78, 163–178. https://doi.org/10.1016/j.envsoft.2016.01.001.

Leonard, L., Duffy, C., 2016b. Visualization workflows for level-12 HUC scales: towards an expert system for watershed analysis in a distributed computing environment. Environ. Model. Software 78, 163–178. https://doi.org/10.1016/j.envsoft.2016.01.001.

Leonard, L., Duffy, C.J., 2014. Automating data-model workflows at a level 12 HUC scale: watershed modeling in a distributed computing environment. Environ. Model. Software 61, 174–190. https://doi.org/10.1016/j.envsoft.2014.07.015.

Leonard, L., Duffy, C.J., 2013. Essential Terrestrial Variable data workflows for distributed water resources modeling. Environ. Model. Software 50, 85–96. https://doi.org/10.1016/j.envsoft.2013.09.003.

Leonard, L.N., Band, L.E., 2017. Sample RHESSys workflow. WWW Document]. GitHub. https://github.com/hydroshare/hydroshare-jupyterhub/blob/master/notebooks/help/rhessys.ipynb, Accessed date: 7 January 2017.

028.

Neteler, M., Bowman, M.H., Landa, M., Metz, M., 2012. Environmental Modelling & Software GRASS GIS : a multi-purpose open source GIS. Environ. Model. Software 31, 124–130. https://doi.org/10.1016/j.envsoft.2011.11.014.

NLDAS, 2011. North American land data assimilation system. WWW Document]. http://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php.

OAI, 2018. Open archives initiative. WWW Document]. https://www.openarchives.org/organization/.

OpenStack, 2017. OpenStack. WWW Document]. https://www.openstack.org/software/, Accessed date: 7 January 2017.

Pataki, D.E., McCarthy, H.R., Gillespie, T., Jenerette, G.D., Pincetl, S., 2013. A trait-based ecology of the Los Angeles urban forest. Ecosphere 4 art72. https://doi.org/10.1890/ES13-00017.1.

Perez, F., Granger, B.E., 2007. IPython: a system for interactive scientific computing. Comput. Sci. Eng. 9, 21–29. https://doi.org/10.1109/MCSE.2007.53.

Pincetl, S., Gillespie, T., Pataki, D.E., Saatchi, S., Saphores, J.-D., 2013. Urban tree planting programs, function or fashion? Los Angeles and urban tree planting campaigns. Geojournal 78, 475–493. https://doi.org/10.1007/s10708-012-9446-x.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The natural flow regime. Bioscience 47, 769–784. https://doi.org/10.2307/1313099.

RHESSys, 2017. RHESSys : the regional hydro-ecologic simulation system. [WWW Document]. GitHub. https://github.com/RHESSys/RHESSys, Accessed date: 7 January 2017.

Running, S.W., Coughlan, J.C., 1988. A general model of forest ecosystem processes for regional applications I. Hydrologic balance, canopy gas exchange and primary production processes. Ecol. Model. 42, 125–154. https://doi.org/10.1016/0304-3800(88)90112-3.

Running, S.W., Nemani, R.R., Hungerford, R.D., 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. Can. J. For. Res. 17, 472–483. https://doi.org/10.1139/x87-081.

SSURGO, 2011. Soil survey geographic database. WWW Document. http://soils.usda.gov/survey/geography/ssurgo, Accessed date: 8 January 2010.

Martin, K.L., Hwang, T., Vose, J.M., Coulston, J.W., Wear, D.N., Miles, B., Band, L.E., 2017. Watershed impacts of climate and land use changes depend on magnitude and land use context. Ecohydrology 10 e1870. https://doi.org/10.1002/eco.1870.

Miles, B., Band, L.E., 2017a. EcoHydroLib. WWW Document]. GitHub. https://github.com/selimnairb/EcohydroLib, Accessed date: 7 January 2017.

Miles, B., Band, L.E., 2017b. RHESSysWorkflows. WWW Document]. GitHub. https://github.com/selimnairb/RHESSysWorkflows, Accessed date: 7 January 2017.

Miles, B., Band, L.E., 2017c. RHESSysWorkflows installation instructions. WWW Document]. GitHub. https://github.com/selimnairb/RHESSysWorkflows#installation-instructions, Accessed date: 7 January 2017.

Miles, B., Band, L.E., 2015a. Green infrastructure stormwater management at the watershed scale: urban variable source area and watershed capacitance. Hydrol. Process. 29, 2268–2274. https://doi.org/10.1002/hyp.10448.

Miles, B., Band, L.E., 2015b. Ecohydrology models without borders? In: Environmental Software Systems. Infrastructures, Services and Applications: 11th IFIP WG 5.11 International Symposium, ISESS 2015, Melbourne, VIC, Australia, March 25-27, 2015. Proceedings, pp. 311–320. https://doi.org/10.1007/978-3-319-15994-2_31.

Miles, B.C., 2014. Small-scale Residential Stormwater Management in Urbanized Watersheds: a Geoinformatics-driven Ecohydrology Modeling Approach. The University of North Carolina at Chapel Hill.

Mittman, T., Band, L.E., Hwang, T., Smith, M.L., 2012. Distributed hydrologic modeling in the suburban landscape: assessing parameter transferability from gauged reference Catchments1. JAWRA J. Am. Water Resour. Assoc. 48, 546–557. https://doi.org/10.1111/j.1752-1688.2011.00636.x.

Morsy, M.M., Goodall, J.L., Castronova, A.M., Dash, P., Merwade, V., Sadler, J.M., Rajib, M.A., Horsburgh, J.S., Tarboton, D.G., 2017. Design of a metadata framework for environmental models with an example hydrologic application in HydroShare. Environ. Model. Software 93, 13–28. https://doi.org/10.1016/j.envsoft.2017.02.

STATSGO, 2011. State soil geographic database. WWW Document]. http://soils.usda.gov/survey/geography/statsgo/, Accessed date: 8 January 2010.

Subramanian, V., Signell, R., Fernandes, F.P.A., Hernandez, D., 2015. An end-to-end workflow for assessing sea surface temperature, salinity and water level predicted by coastal ocean models. In: OCEANS 2015. MTS/IEEE Washington. IEEE, pp. 1–5. https://doi.org/10.23919/OCEANS.2015.7404614.

Tague, C.L., Band, L.E., 2004. RHESSys : Regional Hydro-Ecologic Simulation System - an object- oriented approach to spatially distributed modeling of carbon, water, and nutrient cycling. Earth Interact. 8 (19) 42 p. https://journals.ametsoc.org/doi/abs/10.1175/1087-3562%282004%298%3C1%3ARRHSSO%3E2.0.CO%3B2.

Tarboton, D., Horsburgh, J.S., Idaszak, Ray, Heard, J., Valentine, D., Couch, A., Ames, D.A.N., Goodall, J.L., Band, L., Merwade, V., Arrigo, J., Hooper, R., 2014. A resource centric approach for advancing collaboration through hydrologic data and model sharing. In: 11th International Conference on Hydroinformatics, HIC 2014, New York City, USA.

Tarboton, D.G., Idaszak, R., Horsburgh, J.S., Heard, J., Ames, D., Goodball, J.L., Merwade, V., Couch, A., Arrigo, J., Hooper, R., Valentine, D., Maidment, D.R., 2014. HydroShare: advancing collaboration through hydrologic data and model sharing. In: International Environmental Modelling and Software Society (IEMSs) 7th International Congress on Environmental Modelling and Software.

United States Environmental Protection Agency, 2017. Green infrastructure. WWW Document]. http://water.epa.gov/polwaste/green/, Accessed date: 7 January 2017.

Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M., Morgan, R.P., 2005. The urban stream syndrome: current knowledge and the search for a cure. J. North Am. Benthol. Soc. 24, 706–723. https://doi.org/10.1899/04-028.1.

Wigmosta, M.S., Vail, L.W., Lettenmaier, D.P., 1994. A distributed hydrology-vegetation model for complex terrain. Water Resour. 30.