

Cyberinfrastructure

Promoting Collaborative Publication, Interoperability, and Reuse of iUTAH Data and Research Products

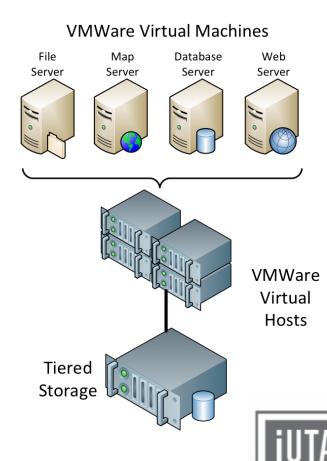
Jeffery S. Horsburgh

Amber Jones and the iUTAH CI Team

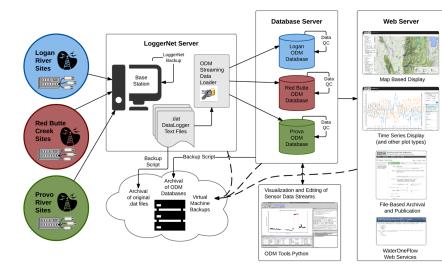
Fall 2016 All Hands Meeting "Focus on Facilities"

Major CI Components

Hardware (Servers and Storage)



Operational Data Management (GAMUT Workflow)



(nowlton Fork P oil nit nowlton Fork RBC soil nitrogen data 0 iUTAH

Modeling and Data Federatio

ШТАН 🎫 / Orga

штан atewide effort

Google-

Twitte

n Facebor

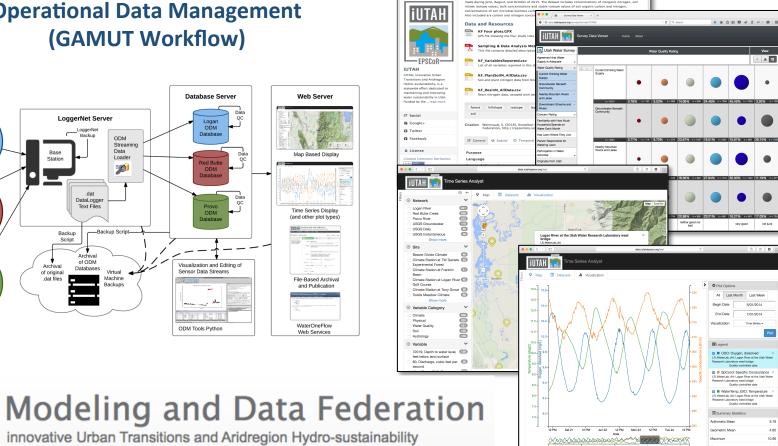
& License

Bari Butte Cre

Provo River

Water Qualit

Variable



Software for Data Sharing,

Publication, and Visualization

innovative Urban Transitions and Aridregion Hydro-sustainability

iUTAH Data Repository

http://repository.iutahepscor.org

- Dataset upload and publication
- Dataset landing page
- Permanent URL and citation
- Limited curation process
- Limitations
 - \circ Not flexible (enough)
 - \odot Hard to maintain
 - No real collaboration capabilities
 No DOI for published resources

Home Development - Da	ta 🗸 About 🗸	Log in Regi		
/ Organizations / il	TAH / Knowlton Fork, RBC, soil Datasets	Organizations Groups		
Knowlton Fork, RBC, soil nitrogen data	Dataset O Activity Stream 🔚 Related			
Followers	Knowlton Fork, RBC, soil nit	rogen data		
0	Agree to data use agreement for data preview/download.			
Organization	This dataset includes measurements of soil nitrogen pools and fluxes f herbaceous) and two landscape positions (upper and lower slopes) in th			
Organization	Butte Creek watershed. Sites are located near the iUtah Knowlton Fork (made during June, August, and October of 2015. The dataset includes	Climate Station, and measurements were		
iutah	nitrate laining June, August, and October 10 2013. The dataset includes concentrations of morganic mitrogen, son nitrate isotope values, bulk concentrations and stable isotope values of soil organic carbon and nitrogen, concentrations of soil microbial biomass carbon and nitrogen, and nitrate leachate from below the rooting zone. Also included are carbon and nitrogen concentrations and isotope values from leaves.			
	Data and Resources			
	KF Four plots.GPX GPS file showing the four study sites where soil nitrogen data is	has been		
	Sampling & Data Analysis Methods This file contains detailed descriptions of field sampling and a	nalytical		
EPSCor	KF_VariablesReported.csv List of all variables reported in this dataset, including definition	ns and		
IUTAH				
iUTAH, innovative Urban Transitions and Aridregion Hydro-sustainability, is a	Soil and plant nitrogen data from Knowlton Fork study plots co	ollected during		
statewide effort dedicated to	KF_ResinN_AllData.csv			
maintaining and improving water sustainability in Utah. Funded by the read more	Resin nitrogen data, assayed with two methods (resin cores an	d free resins),		
	forest hillslope isotope leaching microbial	nitrogen plant-microbe inter		
🖻 Social	301			
Soogle+	Citation: Weintraub, S. (2016), Knowlton Fork, RBC, soil nitrogen			
Twitter	Federation, http://repository.iutahepscor.org/dataset/k	nowiton-fork-rbc-soil-nitrogen-data		
🖬 Facebook	🕼 General 📀 Spatial 💿 Temporal 🗐 Variable & Method	d 🛔 Contact 🌐 Additional		
🔒 License	Purpose Research			
Creative Commons Attribution	Language english			
OPEN DATA	Research Focus Area RFA1			
	Access Information No limitation			

What do we want to do?

- Easily create a digital instance of a dataset or model
- Quickly share it with colleagues (perhaps privately at first)
- Add value through collaboration, annotation, and iteration
- Describe with metadata
- Eventually...share publicly or formally Publish

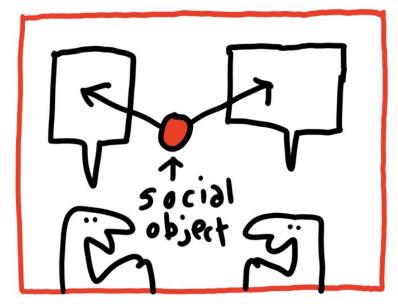


Data and models are "social objects" shared among scientists

Social Objects

"Objects around which social networks form"

Jyri Engeström



Hugh MacLeod http://sharinglab.dk/what-is-a-social-object/



How to fix an Agitator on a Washing Machine

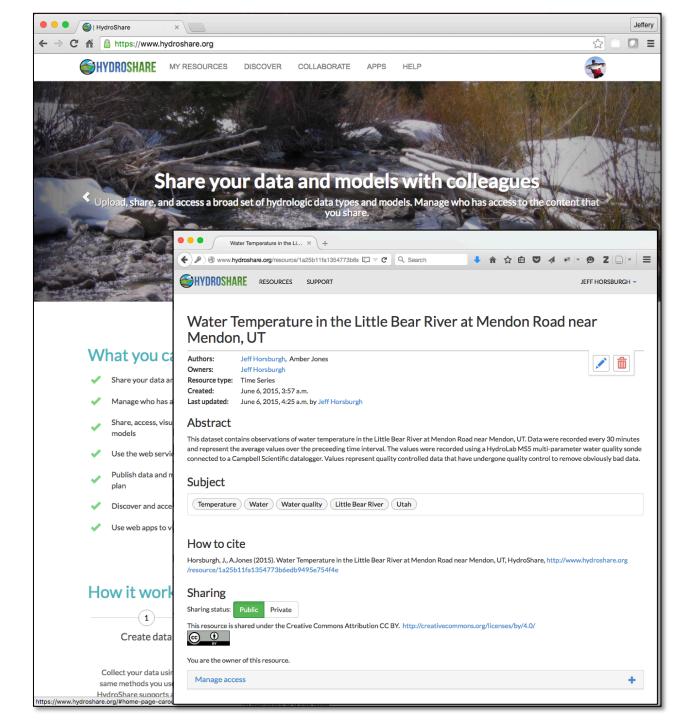
weezie63	
2.000	520,285
+ Add to <\$ Share ···· More	1,341 🗭 94





- Collaborative development project that started about the same time as iUTAH
- <u>Sharing</u> and <u>collaborating</u> around diverse data types used by water scientists

HydroShare's goal: Enable scientists to create **social objects** that add value



HydroShare "Resources"

- <u>**Resource</u>** = primary unit of digital content</u>
 - Create
 - Share
 - Own
 - Access
 - Filter
 - Discover
 - Publish

Resources can be datasets, models, or other digital content

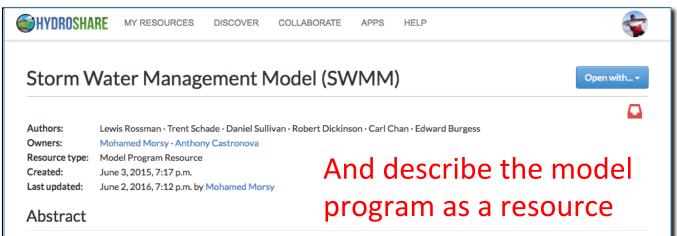
🔴 🔍 🎯 My Resources Hy	vdroShare ×			Je
→ C 🖬 🔒 https://w	ww.hydroshare.org/my-	resources/	Q	
HYDROSHARE MY RE	ESOURCES DISCOVER	COLLABORATE APPS HELP		*
My Resources	🏷 ★	Q. Search	⇔ Owners ⇔ Last m	
Q Filter Owned by me Editable by me	• * •	 Survey of Stormwater Managers in Utah 		, 2016, 6:25
Viewable by meFavorites	= * %	6 🖻 Share and Publish your Data and Models w	vith HydroShare David Tarboton June 28. 11:58 a.	
Labels No labels found.	• * 🗣 📔	🔒 🔄 Cross section survey, Northwest Field Can	al at 200 South Jeffery June 27, Horsburgh p.m.	, 2016, 7:40
	• * •	6 🖻 Water Temperature in the Little Bear Rive Mendon, UT	r at Mendon Road near Jeffery June 21, Horsburgh p.m.	, 2016, 7:14
	• * •	A EC-HMS Version 4.1	David Tarboton June 16, p.m.	, 2016, 8:20
	• * •	🔓 🖄 Logan Digital Elevation Model	David Tarboton June 14, p.m.	, 2016, 1:30
	• * •	🔓 🖄 Logan Specific Catchment Area	David Tarboton June 14, p.m.	, 2016, 1:30
	• * •	🔓 🖄 100yr flood every 3 years	Matthew Turner D.m. P.m.	2016, 11:44
	• * •	6 🖄 Hydrology Domain Cyberinfrastructures:	Successes, Challenges, Jeffery Dec. 17, Horsburgh a.m.	2015, 9:07
	= * >	🔓 🖄 CI-WATER Workshop - Tethys Provo Dam	Break App Data Nathan Swain p.m.	2015, 8:56
	Legend			
CONTACT US		FOLLOW	OPEN SOURCE	
Email us at hydroshare.org		🛩 f 🛗 🗘 in	HydroShare is Open Source. Find us on Githu	њ.

Resource Types in HydroShare

- Generic
- Multidimensional (NetCDF)
- Collections
- Geographic
 - Raster
 - Feature (ESRI Shapefiles)
- Time Series
 - Time Series
 - Referenced Time Series
- Modeling
 - Model Program
 - Model Instance
 - SWAT Model Instance
 - MODFLOW Model Instance
 - Script

Specific Resource types have:

- 1. A specific metadata description
- 2. An expected file format and structure



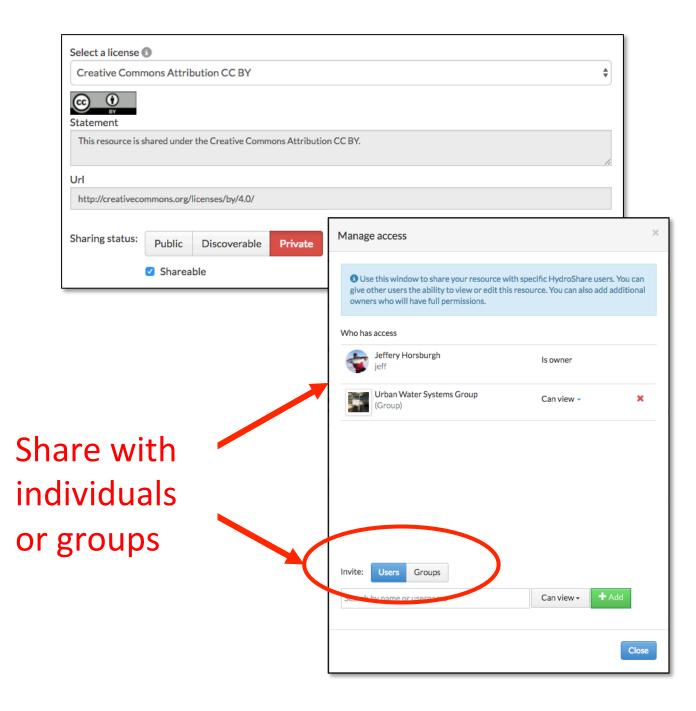
The EPA Storm Water Management Model (SWMM) is a dynamic rainfall-runoff simulation model used for single event or long-term (continuous) simulation of runoff quantity and quality from primarily urban areas. The runoff component of SWMM operates on a collection of subcatchment areas on which rain falls and runoff is generated. The routing portion of SWMM transports this runoff through a conveyance system of pipes, channels, storage/treatment devices, pumps, and regulators. SWMM tracks the quantity and quality of runoff generated within each subcatchment, and the flow rate, flow depth, and quality of water in each pipe and channel during a simulation period comprised of multiple time steps. SWMM was first developed back in 1971 and has undergone several major upgrades since then. The current edition, Version 5, is a complete re-write of the previous release. Running under Windows, EPA SWMM 5 provides an integrated environment for editing drainage area input data, running hydraulic and water quality simulations, and viewing the results in a variety of formats. These include color-coded drainage area maps, time series graphs and tables, profile plots, and statistical frequency analyses.

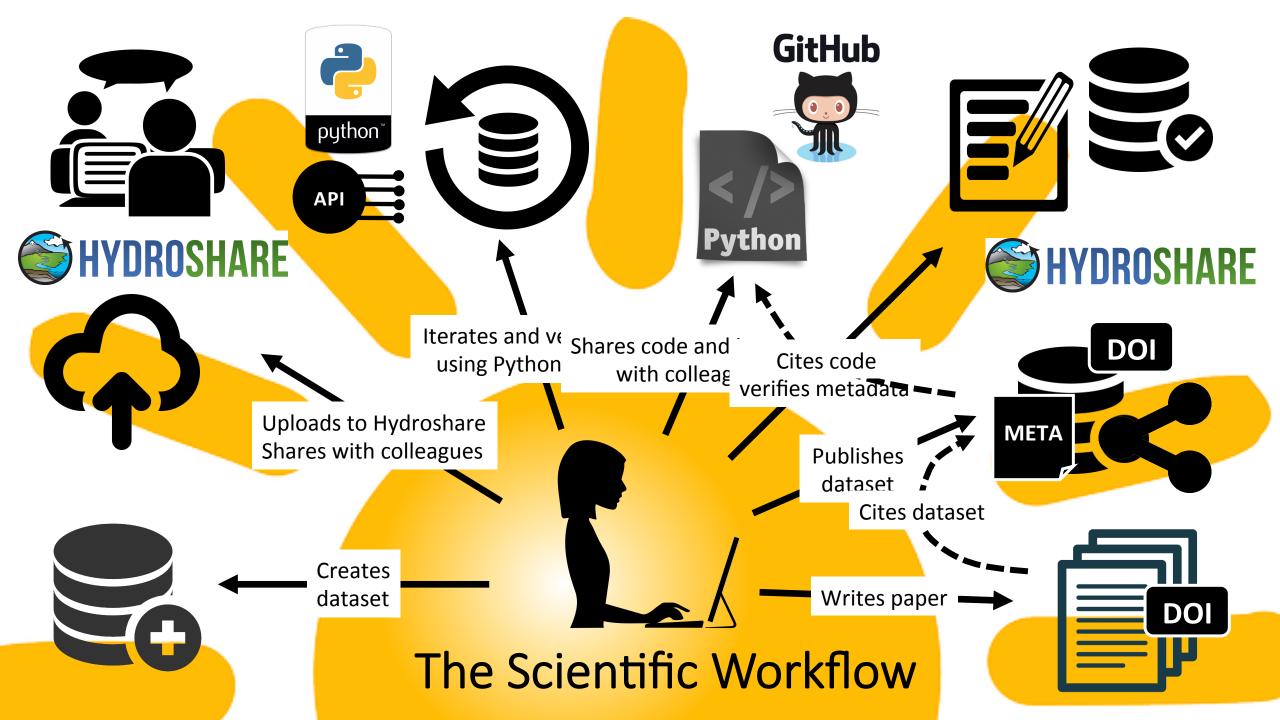
How to cite

Rossman, L., T. Schade, D. Sullivan, R. Dickinson, C. Chan, E. Burgess (2016). Storm Water Management Model (SWMM), Version 5.1.010 with Low Impact Development (LID) Controls, http://www2.epa.gov/water-research/storm-water-management-model-swmm, accessed 6/2/2016,

Resource Sharing in HydroShare

- You control who has access
- 5 Options:
 - <u>Private</u> Only individual users you have given access can view metadata or access content
 - <u>Discoverable</u> Anyone can view the metadata, but only users with permission can access content
 - <u>Public</u> Anyone can view the metadata and access content
 - <u>Published</u> Same as public, but the resource is made immutable and a DOI is assigned
 - <u>Shareable</u> Other users can grant access at their same level of access





iUTAH Resources in HydroShare

GAMUT Raw and QC1 Datasets

uthors:)wners: iesource type: Treated:	IUTAH GAMUT Working Group	C 🖉
ast updated:	iUTAH Data Manager Generic July 8, 2016, 6: 129 p.m. Oct. 28, 2016, 9:42 a.m. by iUTAH Data Manager	
bstract		
reek (RB_KF_B/	ains raw data for all of the variables measured for the iUTAH GAMUT Network aquatic site on the Knowlton Fork tr A). Each file contains a calendar year of data. The file for the current year is updated on a daily basis. The data values s at 15 minute intervals. The file header contains detailed metadata for the site and the variable and method of each	were collected by a
Group, i. G. (201	6). iUTAH GAMUT Network Raw Data at Knowlton Fork Basic Aquatic Site (RB_KF_BA), HydroShare,	
aring status: P	roshare.org/resource/dd1a2fbc2e4422&82a3fbd552&20c32 hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable an specific permission to edit this resource.	
aring status: P	hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable	
aaring status: Provide the second status of the sec	hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable en specific permission to edit this resource.	
EX UNIT OF CONTRACT OF CONTRACTON OF CONTRACTO	hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable en specific permission to edit this resource.	
aring status: Pi u have been give ubject Red Butte Cro Content	hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable en specific permission to edit this resource.	0
O O	hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable en specific permission to edit this resource.	(e) (c)
Definition of the second	hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable en specific permission to edit this resource. eek GAMUT (raw data) (time series) (water quality) tents/IUTAH_GAMUT_RB_KF_BA_RawData_2013.csv tents/IUTAH_GAMUT_RB_KF_BA_RawData_2013.csv tents/IUTAH_GAMUT_RB_KF_BA_RawData_2015.csv	-
Definition of the second	hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ ublic & Shareable en specific permission to edit this resource. eek GAMUT (raw data) (time series) (water quality) tents/IUTAH_GAMUT_RB_KF_BA_RawData_2013.csv tents/IUTAH_GAMUT_RB_KF_BA_RawData_2013.csv	0

GAMUT Discharge Rating Curves

	ARE MY RESOURCES DISCOVER COLLABORATE APPS HELP	
	rge Rating Curve at Red Butte Creek near Foothill D ced Aquatic Site (RB_FD_AA)	Open with
		C
Authors: Owners:	iUTAH GAMUT Working Group iUTAH Data Manager	
esource type: reated:	:: Generic July 19, 2016, 10:03 p.m.	
ast updated:	Oct. 25, 2016, 7:50 p.m. by iUTAH Data Manager	
Abstract		
Bridge (RB_FD_ Rating Curve fil vith each meas lata through 13	ntains a stage-discharge relationship developed for the UTAH CANUT Network aquatic site on Red Butte Cree. AvA) Discharge measurements were collected by a SonKe Neur Tracker. Measured stage and discharge and the le. Information on the site conditions and any issues with discharge measurements are documented in the READ summent. (e.g., output by the FlowTracker instrument) are contained in the azjo directory. This rating curve was u 2/31/2015. New versions of these files may be loaded when new flow measurements are taken. Resulting discher Coperational databases and may be accessed via http://data.iutahepscor.org/ss.	curve are contained in the DME file. Files associated used to generate discharge
How to c	ite	
	016). Discharge Rating Curve at Red Butte Creek near Foothill Drive Advanced Aquatic Site (RB_FD_AA), Hydro ydroshare.org/resource/cf8d84ef37964fa3a10f69ce4b9f9586	Share,
Ge O	s hared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ Public & Shareable iven specific permission to view this resource.	
subject		
Discharge	Flow FlowTracker GAMUT Rating Curve Red Butte Creek Stage-Discharge	
Content		
data/co	ontents/rb-fd-aa-dischargemeasurements-160222.zip	٩
L data/co	ontents/rb-fd-aa-ratingcurve-160222.xlsx	٩
	ontents/rb-fd-aa-ratingcurve-readme-160222.pdf	٢
16.5 KB		U
data/co	ontents/iUTAH_GAMUT_RB_FD_AA_RawData_2016.csv	•

Learn more about the Bagit archive format

Individual Investigator Datasets

	precipitation ions and nitrogen stable isotope	
compos	sition	Open with -
Authors: Owners:	Steven Hall iUTAH Data Manager - Steven Hall	
Resource type:		
Created: Last updated:	Aug. 3, 2016, 11:52 p.m. Sept. 7, 2016, midnight by iUTAH Data Manager	
	Sept. 7, 2010, miningin uy to the transger	
Abstract		
	ons and precipitation amount were measured at 14 sites in the Salt Lake and Cache Valleys from December 2013 to February poradic at several sites. The goal of this study was to identify land use impacts on nitrogen deposition to the iUTAH watershe	
	ploratic at several sites. The goal of this study was to identify land use impacts of hird ogen deposition to the to her water she	
	idings are described in the associated JGR-B manuscript	
How to ci	te	
Hall, S. (2016). i	• • • • • • • • • • • • • • • • • • •	-
http://dx.doi.or	g/10.4211/hs.7ce47150ee344d4e95f2fd7bb4f660ac	
http://dx.doi.or	g/10.4211/hs.7ce47150ee344d4e95f2fd7bb4f660ac	
http://dx.doi.or	y/10.4211/hs.7ce47150ec344d4e95f2fd7bb4f660ac	-
http://dx.doi.org	y/10.4211/hs.7ce47150ec3444e95f2fd7bb4f660ac shared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ Published ven specific permission to view this resource.	-
http://dx.doi.org This resource is Sharing status: F You have been give Subject Logan River	y/10.4211/hs.7ce47150ee34444e95f2fd7bb4f660ac shared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ Published ren specific permission to view this resource. Red Butte ion introgen precipitation chemistry sulfate intrate phosphate chloride fluoric	-
http://dx.doi.or This resource is Dear of the second of th	y/10.4211/hs.7ce47150ee34444e95f2fd7bb4f660ac shared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ Published ren specific permission to view this resource. Red Butte ion introgen precipitation chemistry sulfate intrate phosphate chloride fluoric	-
http://dx.doi.or, This resource is Sharing status: F Subject Logan River (magnesium Content	y/10.4211/hs.7ce4715Dec3444de95f2fd7bb4f660ac shared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ Published ene specific permission to view this resource.	6
http://dx.doi.or This resource is Sharing status: F Subject Logan River magnesium Content	y/10.4211/hs.7ce4715Dec3444dep5f2fd7bb4f660ac shared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/ Published ren specific permission to view this resource. Red Butte (on introgen precipitation chemistry sulfate intrate phosphate chloride fluoric calcium potassium sodium stable isotope 15N introgen isotope Utah	e ()
http://dx.doi.or This resource is Sharing status: F You have been gin Subject Logan River magnesium Content data/cor 1.5 KB data/cor 2.0 KB data/cor 1.5 KB data/cor	y/10.4211/hs.7ce4715Dec3444de95f2fd7bb4f660ac shared under the Creative Commons Attribution CC BY. http://creative.commons.org/licenses/by/4.0/ Published ren specific permission to view this resource. Red Butte ion nitrogen precipitation chemistry sulfate nitrate phosphate chloride fluoric calcium potassium sodium stable isotope 15N nitrogen isotope Utah ntents/160127-jutah-precip-mh4-15n.csv	e () ()

Learn more about the Bagit archive format

Collaborative Groups

- Create a collaborative group
- Share resources with a group
- Manage group membership
- Find groups you are interested in
- Request group membership

HYDROSHARE MY RESOURCES DISCOVER COLLABORATE APPS HELP
Find Groups My Groups
Find Groups
+ Create Group
Freshwater is led by the University of Washington with support of the Mountain to Sea Strategic Research Initiative for advancing freshwater research in the Pacific Northwest and the world. (http://freshwater.uw.edu/) Freshwater researchers create positive change through scientific discovery and technological innovation. This is community resource for education, data and tool sharing for overcoming the global challenges in water quality, resource management, and access.
MEMBERS

stratigraphy, it can also be used in related fields.

tools can be incorporated into your teaching and research. During the first half the workshop,

Collaborative Capabilities of HydroShare

- Dataset/model creation as a "Resource" <u>the Social Objects</u>!
- Resource sharing (public and private)
- Collaborative groups and sharing within groups
- Resource versioning
- Formal <u>P</u>ublication of resources and assignment of DOIs
- Rating and commenting on resources

Moving iUTAH Repository to HydroShare: Next Steps

- Your data collection plans are still valid and your datasets still need to be published
- <u>Modelers</u> there is now a way for you to share your results!
- If you had a dataset in <u>http://repository.iutahepscor.org</u> we have moved it to HydroShare for you
 - You will receive an email from the iUTAH Data Manager with specific instructions for:
 - Creating an account in HydroShare
 - Asserting ownership of your resource
 - Publishing your dataset
- If you have not submitted your data yet
 - Work directly with Amber to get it set up in HydroShare
 - We'll be making some instructional videos



Environ Monit Assess (2015) 187:348 DOI 10.1007/s10661-015-4594-3

Reusability of CI Components

- We have worked hard to ensure that what we have developed for iUTAH is reusable
- All of our source code is in open source code repositories in GitHub
- We have published papers about many of the things we have done



oonents	for a la	a management and public arge-scale, heterogeneous ackman Jones • Jeffery S. Horsburgh • L. Reeder • Maurier Ramírez • aballo			
Exvironmental Modeling & Contents lists availad Environmental Moo journal homepage: www.ed Open source software for visualization at of continuous hydrologic and water qual	delling & Software sevier.com/locate/envsoft	ConMark	anuary 2015/Accepted: 5 May 2015 rs) 2015. This article is published with open accu- is common for hydrology researchers using in situ sensors at high frequencie d durations, and with spatial distribution e data volumes requiring infrastructure fi management, and sharing. The availabili of these data in addressing scientific que t to water availability, water quality, ar sters relies on effective cyberinfrastructure s transformation of raw sensor data in products. It also depends on the ability i o share and access the data in usedb	to and terrestrial site mon meteorologie melt, soil moistu vater quality. We developed for eff monitoring sites t ad software tools we and sharing the s to source and availal of	es for continuous monitoring of com- cal variables, snow accumulation and ure, surface water flow, and surface present the overall workflow we have fectively transferring data from field o ultimate end-users and describe the have deployed for storing, managing, sensor data. These tools are all open ble for others to use.
Jeffery S. Horsburgh ^a , ^a , Stephanie L. Reeder ^b , Ambe ^a Department of Civil and Environmental Engineering, Urb Water Research Laboratory, UL ^b UBM Water Research Laboratory, Ultah State University, Lagar, UT, USA ^c UBM Water Research Laboratory, Ultah State University, Lagar, UT, USA CALL OF 1 N F O Article hatary: Received in privated form 10 March 2015 Nacibal State University, Call State Control of the Consortium of U Concepted 9 April 2015 Nacibal State Control of the Consortium of U Narch 2015 Nacibal State Control of Control of Control of the Consortium of U Stoppart of the Consortium of the Consortium of U Stoppart o	r Spackman Jones ^b , Jacob Meline ^b	🕝 285 commits Branch: master 👻 New p	isualizing time series of environmental observat P 2 branches O releaser pull request Initial repository setup. Fixed unicode decoding error for Changed project structure for Dja Update LICENSE	© Unwatch Viki ↔ Pulse hit Graphs ions. — Edit Create new file Uplead file Idataseries resource.	 ♦ Settings ♦ BSD-3-Clause and artinez and a
Software availability Name of software: ODM Tools Python Developers: Jeffery S, Horsburgh, Stephanie L, Reeder, Amber Spachman Jones, Jacob Meline, and James Patton Contact: JeffLorsburgh@usu.edu War first availabile: 2014 Bardware required: A: personal computer Software required: A: personal computer Software required: A: personal computer Software required: A: porsonal computer Software availability: 2013 Usure code, installers, example ODM databases, and documentation for the ODM Tools Python databases, and documentation for the ODM Tools Python on UCHICODMToolsPython. Cost: Free, Software and source code are released under the New Berkeley Software Distribution (BSD) License, which allows for liberal reuse of the software and code.	1. Introduction Environmental monitoring with <i>in situ</i> environ presents many challenges for data management, grage-scale networks consisting of multiple site personnel. Over the past decade, there has been a in the use of automated data collection in scientif high frequency, extended duration, and spatial dis ollection efforts require cyberinfrastructure to su tate research using sensor data streams. Researd otnoers need tools for data import and storage as we and management. In addition to addressing the sented by managing the sheer quantity of data, mor managers need practices to ensure high data qualing control. In this paper we describe a workflow for scriptee efficiency of continuous, <i>in situ</i> time series dataset etcure and functionality of an open source soft DM Tools Python that implements this workfly hybon enables users to query and export, visualiz	This material is b	Changed project structure for Dja	of agencies and complemen	8732 awarded to Utah
		opinions, finding	s, and conclusions or recommendations express		

Opportunities

- Better tools for working with heterogeneous data
- Software/apps/systems that enhance our "personal cyberinfrastructure"
- Freedom to move data and models from one platform/app/software to another
- Training a next generation of "cyber-savvy" engineers and scientists

An enhanced, cloud connected, and social scientific workflow