The CUAHSI Community Hydrologic Information System

David Tarboton, David Maidment, Ilya Zaslavsky, Dan Ames, Jon Goodall, Jeffery Horsburgh, Kim Schreuders

CUAHSI HIS

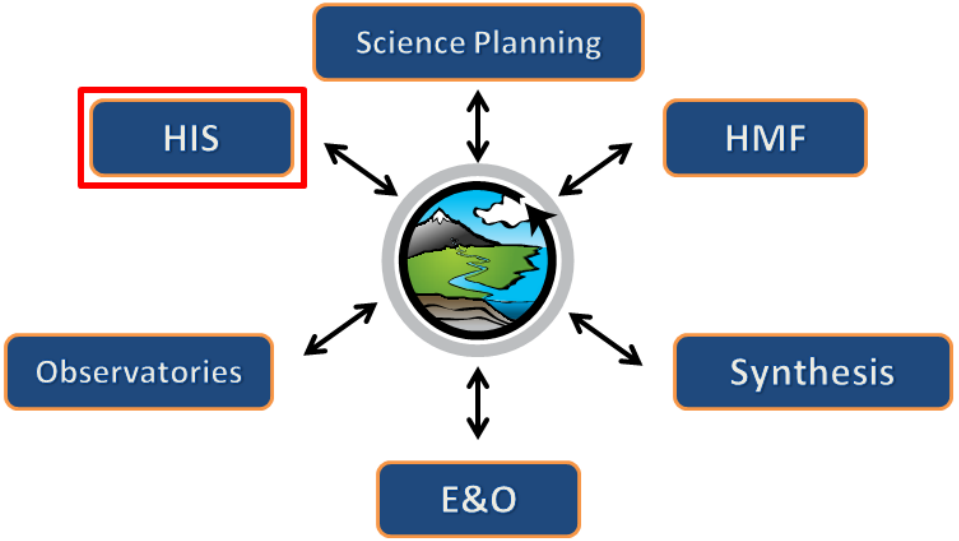
Sharing hydrologic data

http://his.cuahsi.org/

Support
EAR 0622374

What is CUAHSI?

Consortium of Universities for the Advancement of Hydrologic Science, Inc.



- 110 US University members
- 6 affiliate members
- 12 International affiliate members

(as of March 2009)

Infrastructure and services for the advancement of hydrologic science and education in the U.S.

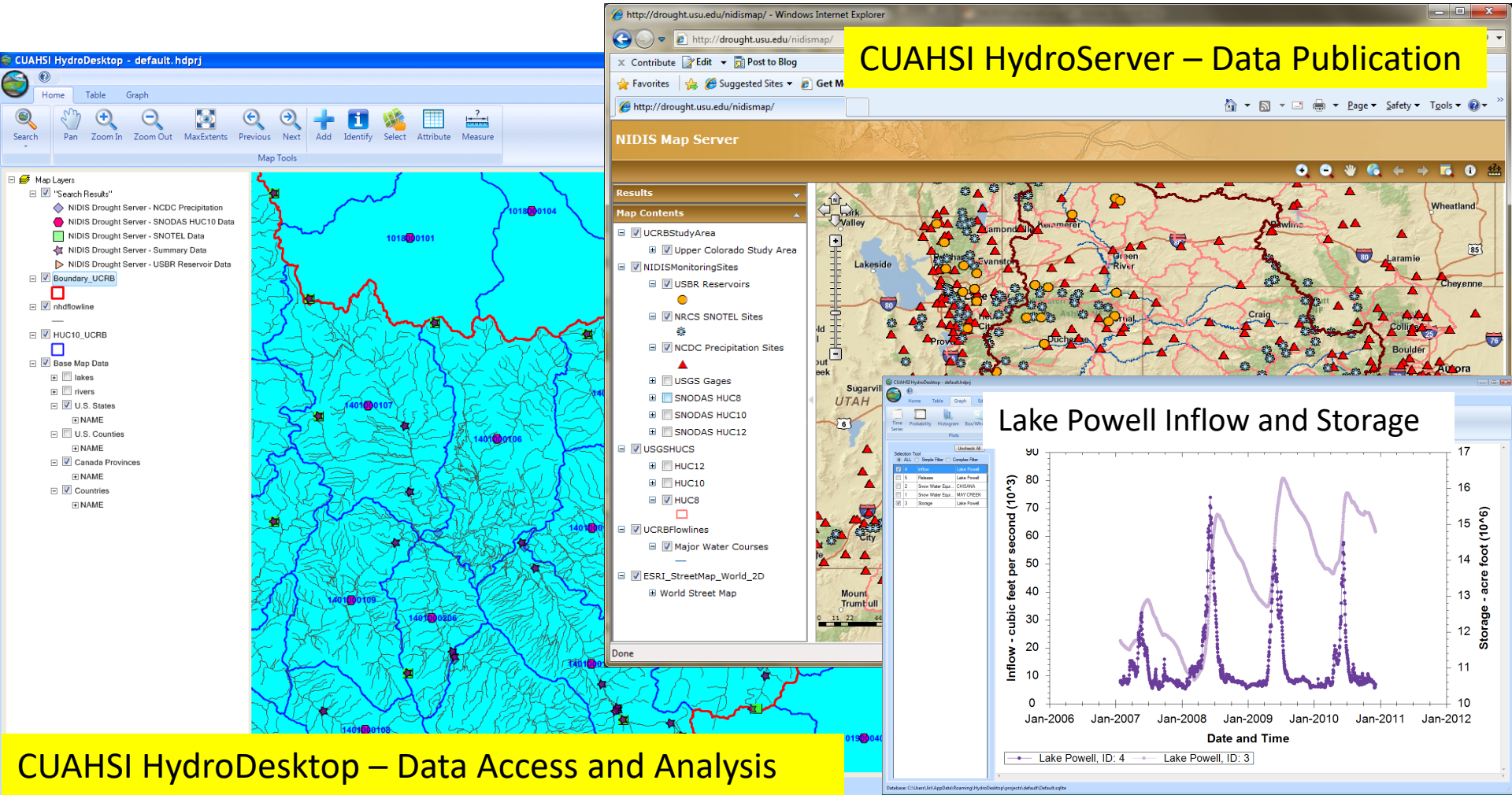
CUAHSI
universities allied for water research

http://www.cuahsi.org/

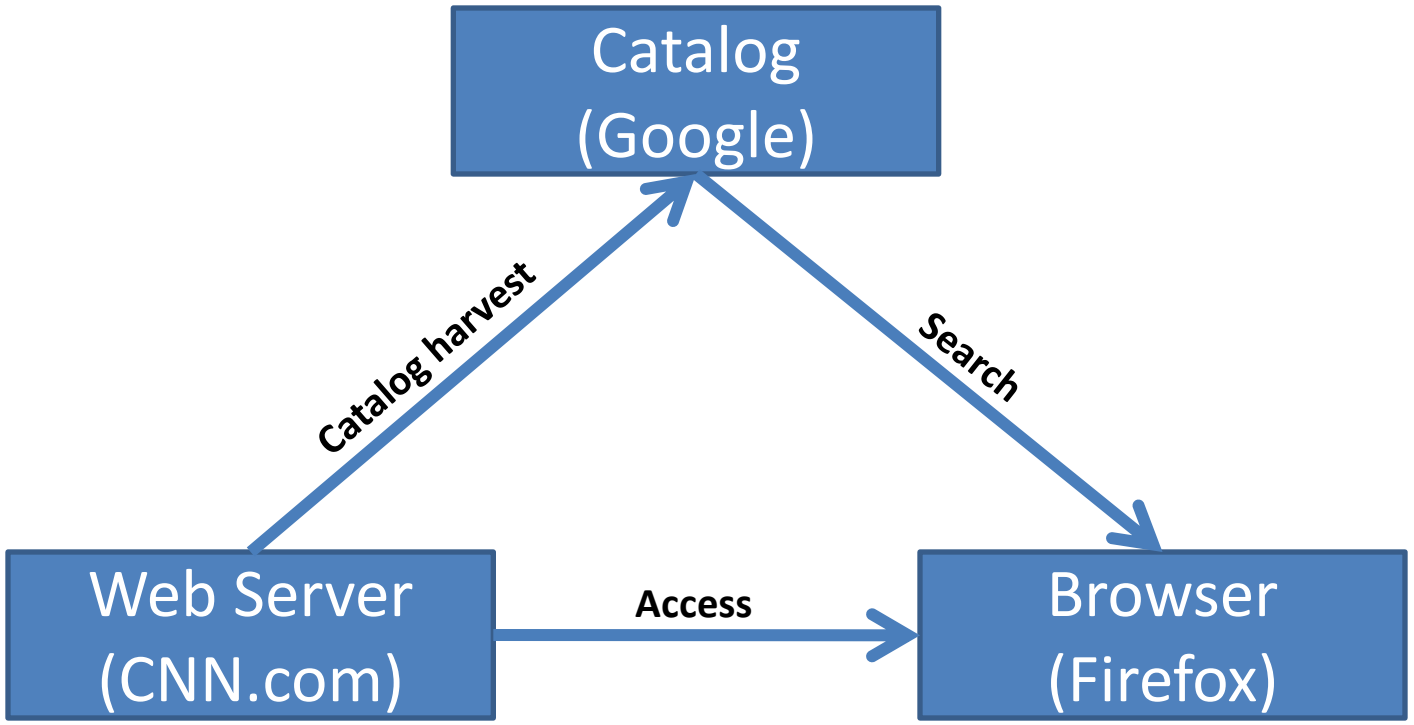
NSF
Support
EAR 0753521

CUAHSI HIS

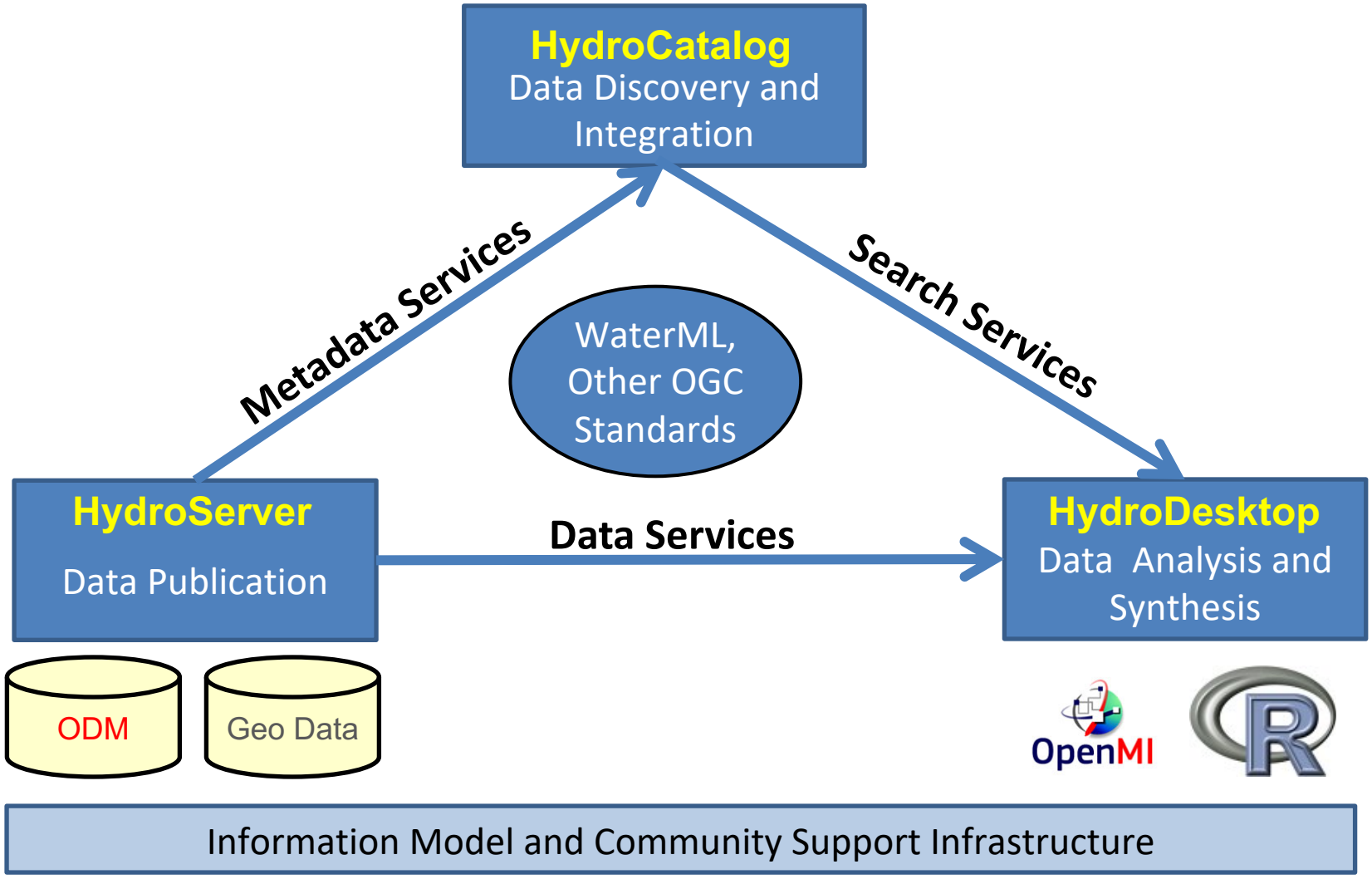
The CUAHSI Hydrologic Information System (HIS) is an internet based system to support the sharing of hydrologic data. It is comprised of hydrologic databases and servers connected through web services as well as software for data publication, discovery and access.



Web Paradigm



CUAHSI Hydrologic Information System Services-Oriented Architecture



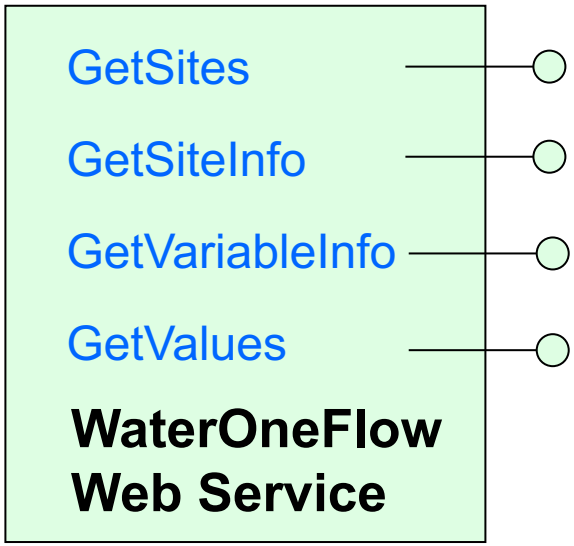
Let’s see some of it

- http://icewater.usu.edu/
- http://hydroserver.codeplex.com
- http://hydrodesktop.codeplex.com

WaterML and WaterOneFlow

WaterML is an XML language for communicating water data
WaterOneFlow is a set of web services based on WaterML

- Set of query functions



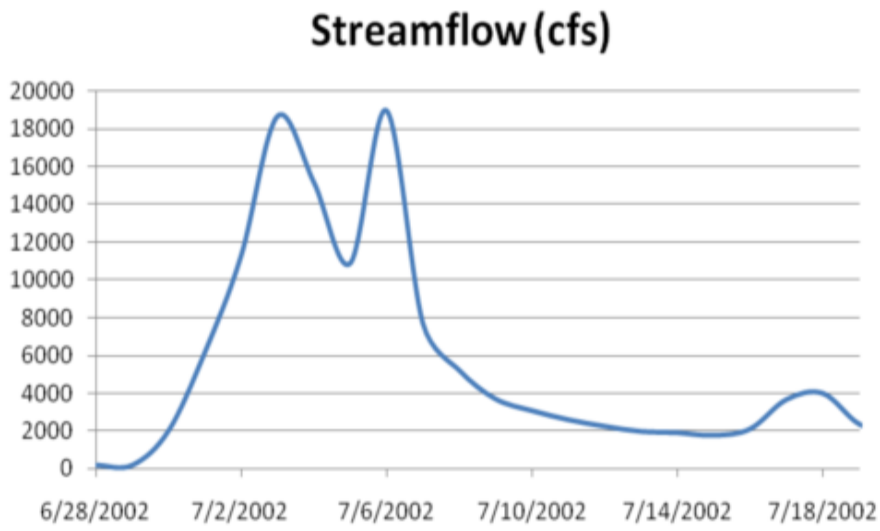
- Returns data in WaterML

```
<timeSeries>
- <sourceInfo xsi:type="SiteInfoType">
    <siteName>Colorado Rv at Austin, TX</siteName>
    <siteCode network="NWIS" siteID="4619631">08158000
  - <geoLocation>
    - <geogLocation xsi:type="LatLonPointType" srs="EPSG
        <latitude>30.24465429</latitude>
        <longitude>-97.694448</longitude>
      </geogLocation>
    </geoLocation>
  </sourceInfo>
- <variable>
    <variableCode vocabulary="NWIS" default="true" variable
    <variableName>Discharge, cubic feet per second</vari
    <units unitsAbbreviation="cfs" unitsCode="35">cubic fee
  </variable>
- <values count="2545">
    <value dateTime="2006-12-31T00:00:00">129</value>
    <value dateTime="2006-12-31T00:15:00">129</value>
    <value dateTime="2006-12-31T00:30:00">129</value>
    <value dateTime="2006-12-31T00:45:00">129</value>
    <value dateTime="2006-12-31T01:00:00">124</value>
    <value dateTime="2006-12-31T01:15:00">129</value>
    <value dateTime="2006-12-31T01:30:00">124</value>
    <value dateTime="2006-12-31T01:45:00">124</value>
    <value dateTime="2006-12-31T02:00:00">124</value>
```

WaterML as a Web Language

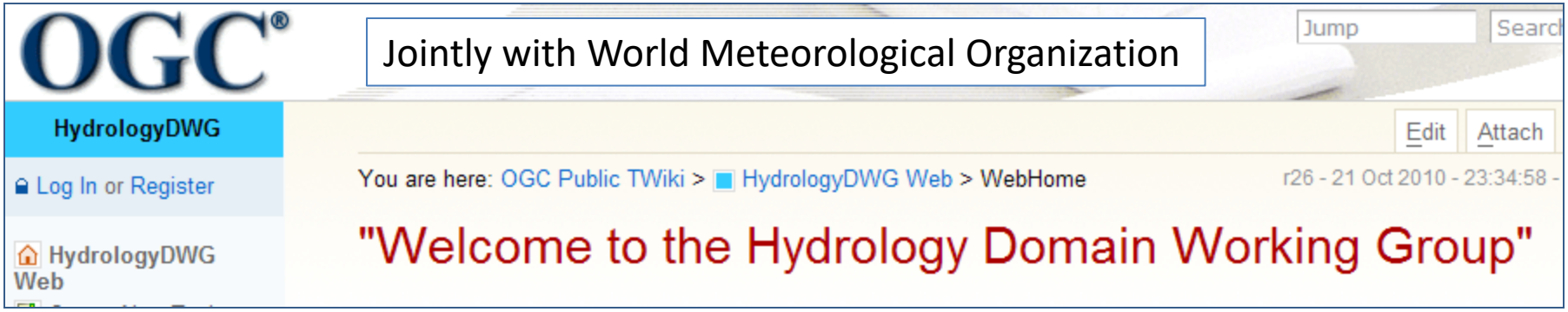
USGS Streamflow data in WaterML language

Discharge of the San Marcos River at Luling, TX
June 28 - July 18, 2002




```
- <values count="21">
    <value qualifiers="A" dateTime="2002-06-28T00:00:00">203</value>
    <value qualifiers="A" dateTime="2002-06-29T00:00:00">195</value>
    <value qualifiers="A" dateTime="2002-06-30T00:00:00">2010</value>
    <value qualifiers="A" dateTime="2002-07-01T00:00:00">6170</value>
    <value qualifiers="A" dateTime="2002-07-02T00:00:00">11300</value>
    <value qualifiers="A" dateTime="2002-07-03T00:00:00">18700</value>
    <value qualifiers="A" dateTime="2002-07-04T00:00:00">15200</value>
    <value qualifiers="A" dateTime="2002-07-05T00:00:00">10900</value>
    <value qualifiers="A" dateTime="2002-07-06T00:00:00">19000</value>
    <value qualifiers="A" dateTime="2002-07-07T00:00:00">7720</value>
    <value qualifiers="A" dateTime="2002-07-08T00:00:00">5230</value>
    <value qualifiers="A" dateTime="2002-07-09T00:00:00">3710</value>
    <value qualifiers="A" dateTime="2002-07-10T00:00:00">3090</value>
    <value qualifiers="A" dateTime="2002-07-11T00:00:00">2610</value>
    <value qualifiers="A" dateTime="2002-07-12T00:00:00">2260</value>
    <value qualifiers="A" dateTime="2002-07-13T00:00:00">1990</value>
    <value qualifiers="A" dateTime="2002-07-14T00:00:00">1920</value>
    <value qualifiers="A" dateTime="2002-07-15T00:00:00">1780</value>
    <value qualifiers="A" dateTime="2002-07-16T00:00:00">2120</value>
    <value qualifiers="A" dateTime="2002-07-17T00:00:00">3680</value>
    <value qualifiers="A" dateTime="2002-07-18T00:00:00">4010</value>
    <qualifier qualifierCode="A" network="USGS" vocabulary="dv_rmk_cd">Approved for publication -- Processing and review completed.</qualifier>
  </values>
```

This is the WaterML GetValues response from NWIS Daily Values



Work Plan

Evolving WaterML into an International Standard

Meetings

- Atlanta OGC TC Meeting - 17 September 2008
- Valencia OGC TC Meeting - 4 December 2008
- Athens OGC TC Meeting - 30 March 2009
- Boston OGC TC Meeting - 22 June 2009
- Darmstadt OGC TC Meeting - 29 September 2009
- Mountain View OGC TC Meeting - 8 December 2009
- Ispra Hydrology DWG Workshop - 15-18 March 2010 - Agenda
- Silver Spring OGC TC Meeting - 15 June 2010
- Toulouse OGC TC Meeting - 22 September 2010
- Toulouse Hydrology DWG Workshop - 21-22 September 2010 - Agenda
- Sydney Hydrology DWG Meeting - 1 December 2010

Meets every 3 months

Teleconferences most weeks

WaterML Version 2 standard being proposed

Vote for adoption 3-6 months later

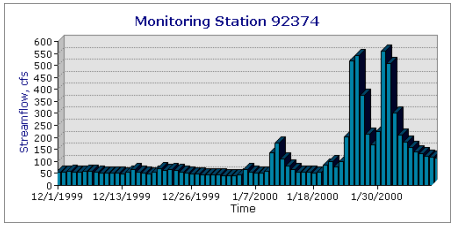
Interoperability Experiments

- GroundwaterInteroperabilityExperiment
- SurfacewaterInteroperabilityExperiment

To be open for public comment April to May 2011
http://www.opengeospatial.org/projects/groups/waterml2.0swg

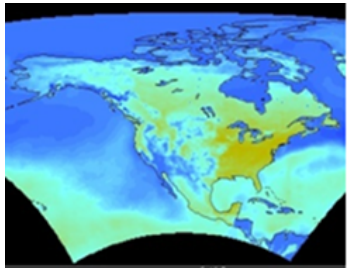
HydroDesktop
Desktop Hydrologic Information System

Observations

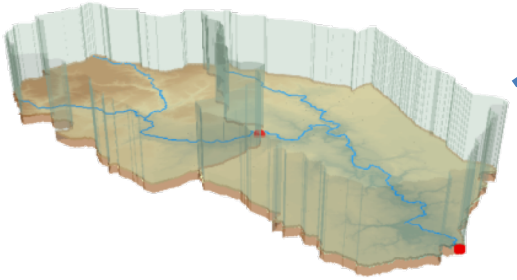


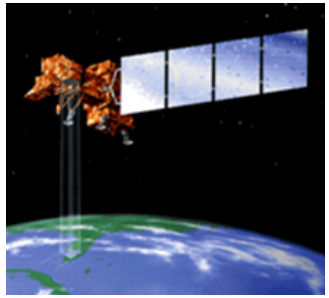
Modeling

Weather and Climate



GIS

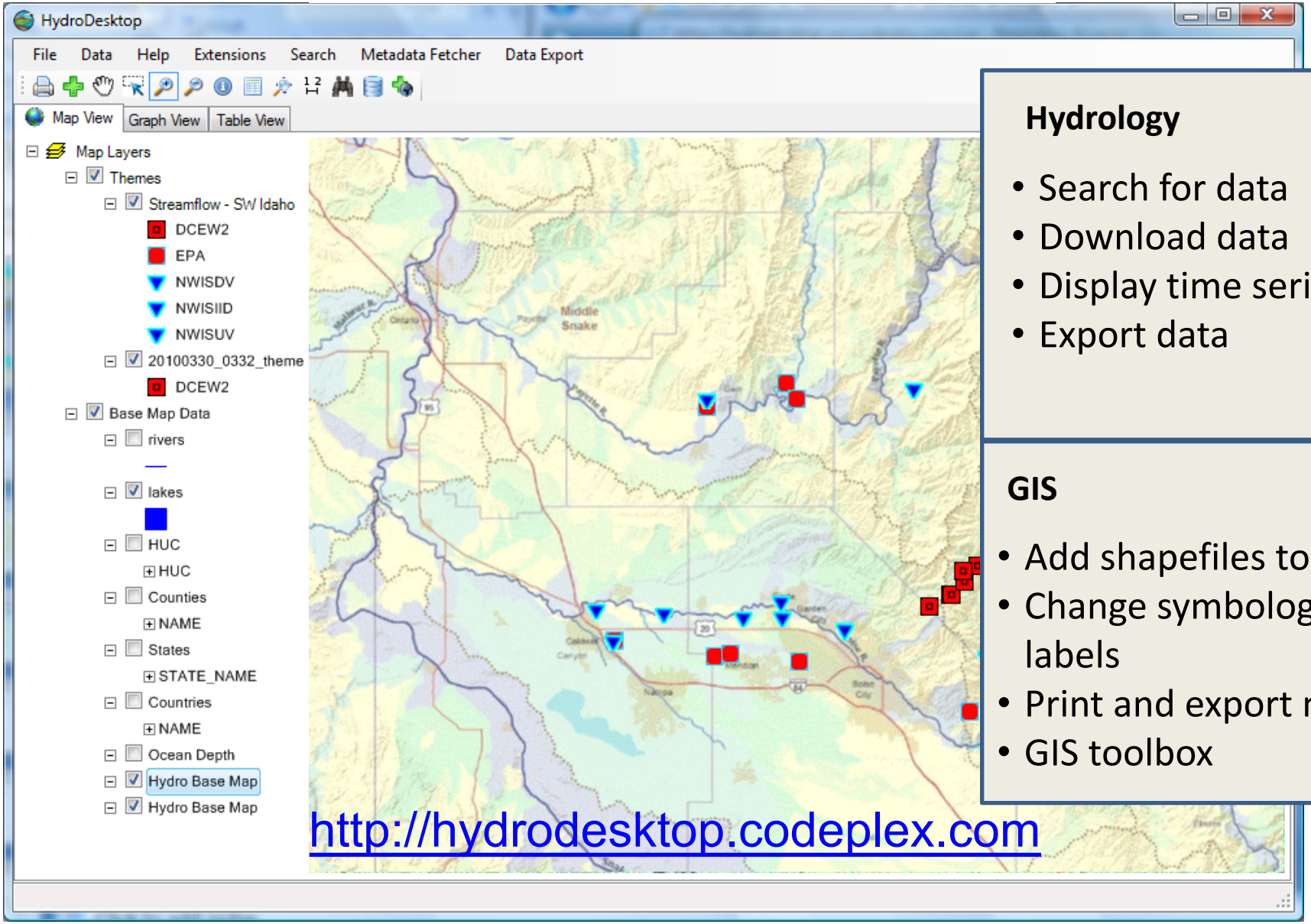




Remote Sensing

HydroDesktop

CUAHSI Open Source Hydrologic Data Tools



Hydrology

- Search for data
- Download data
- Display time series
- Export data

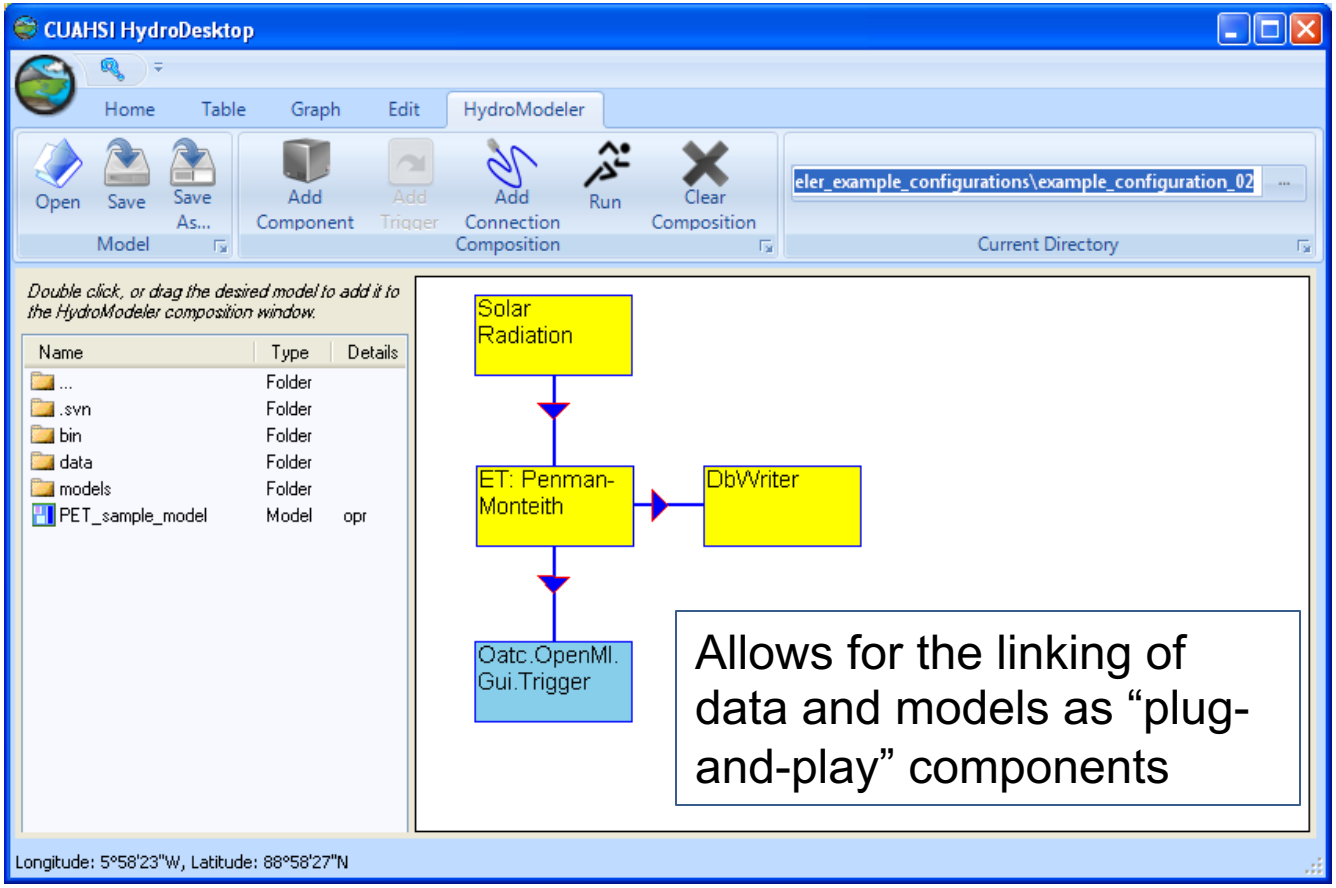
GIS

- Add shapefiles to map
- Change symbology and labels
- Print and export map
- GIS toolbox


HydroDesktop - CUAHSI Hydrologic Information System Desktop Application - Windows Internet Explorer
http://hydrodesktop.codeplex.com
HydroDesktop - CUAH...
Google Calendar
HydroDesktop
CUAHSI Open Source Hydrologic Data Tools
CodePlex
Open Source Community
Search all CodePlex projects
dtarb | Sign Out
Home
Downloads
Documentation
Discussions
Issue Tracker
Source Code
People
License
Project Summary
• 12800 total downloads
• 2000 code commits
• 25 registered developers
Download
Activity

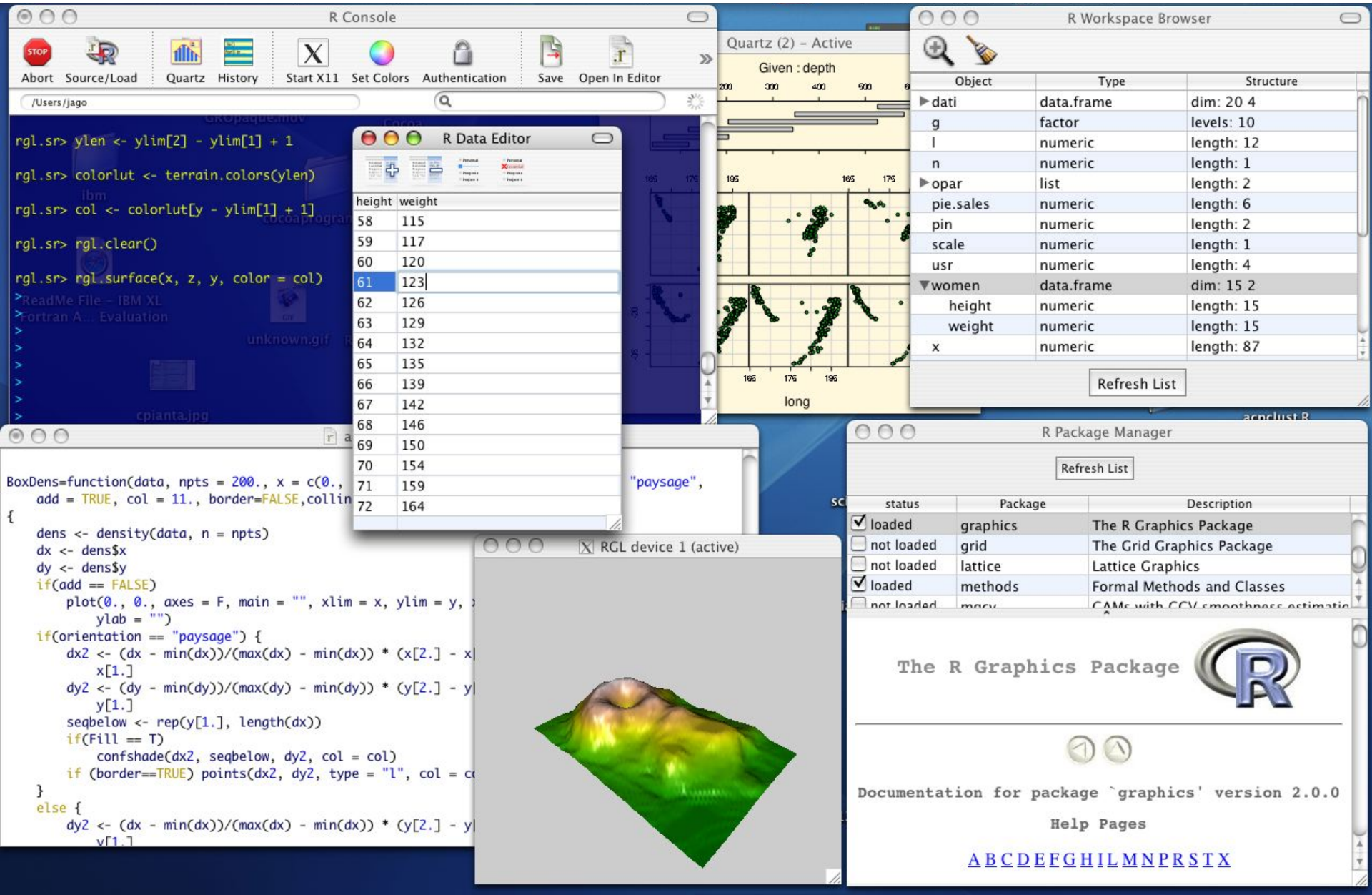
HydroModeler

An integrated modeling environment based on the Open Modeling Interface (OpenMI) standard and embedded within HydroDesktop

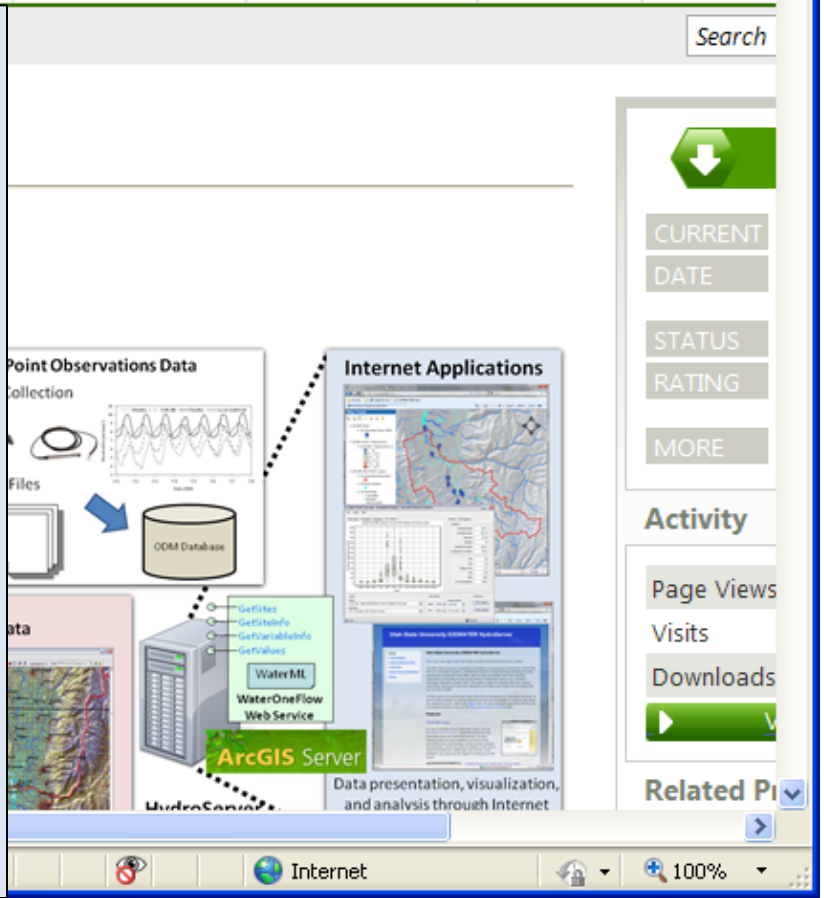


In development at the University of South Carolina by Jon Goodall, Tony Castronova, Mehmet Ercan, Mostafa Elag, and Shirani Fuller

Integration with “R” Statistics Package

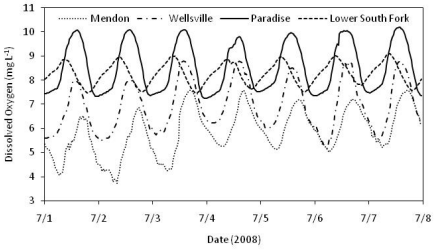


- A platform for publishing space-time hydrologic datasets that:
  - Autonomous with local control of data
  - Part of a distributed system that makes data universally available
- Basis for Experimental Watershed or Observatory data management system
- Standards based approach to data publication
  - Accepted and emerging standards for data storage and transfer (OGC, WaterML)
- Built on established software
  - MS SQL Server, ArcGIS server
- Open Source Community Code Repository
  - Sustainability

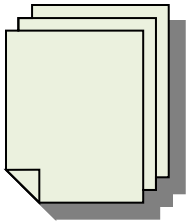


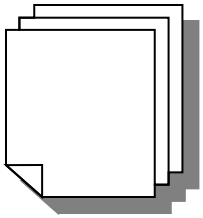
Point Observations Data

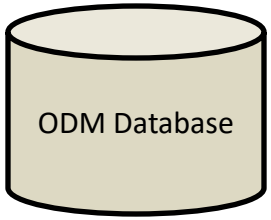
Ongoing Data Collection



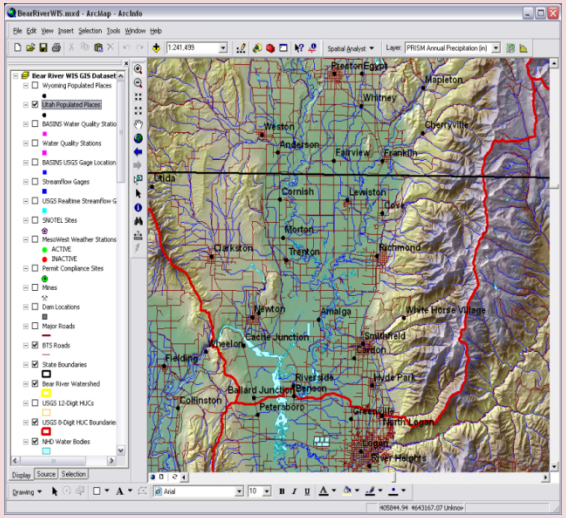
Historical Data Files

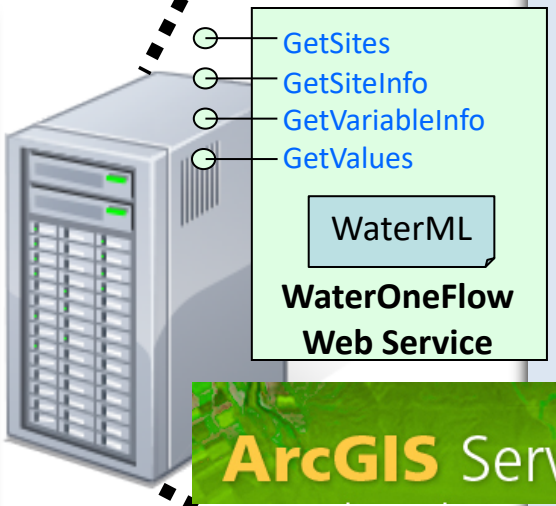






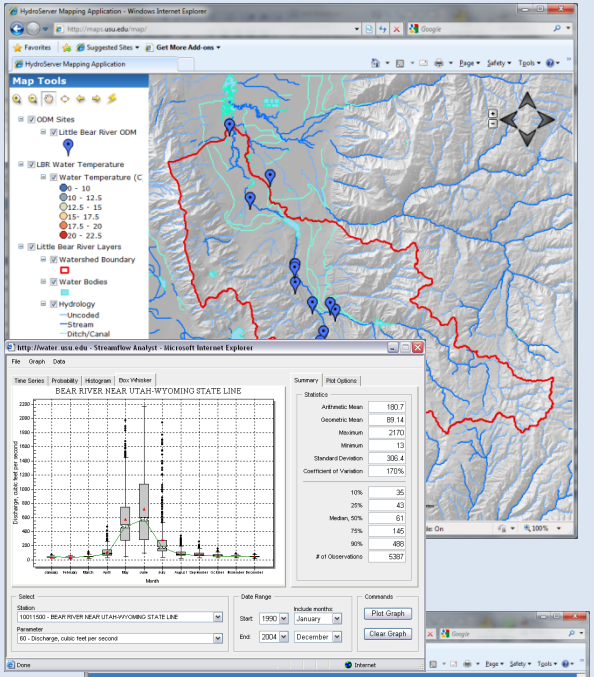
GIS Data

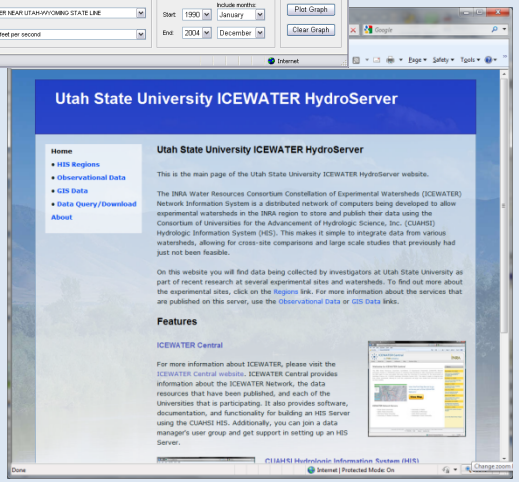




HydroServer

Internet Applications





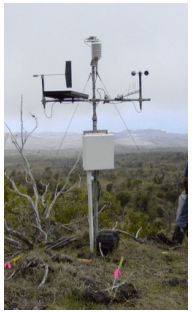
Data presentation, visualization, and analysis through Internet enabled applications

Observation Data Model for hydrologic and environmental measurements

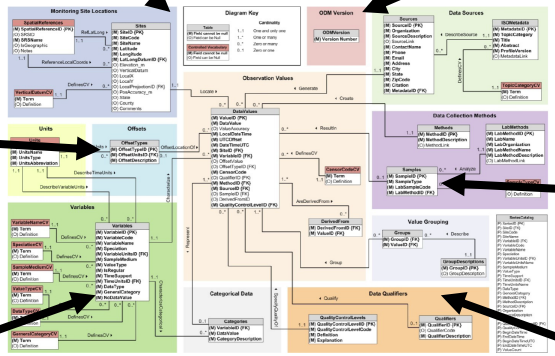
The way that data is organized can enhance or inhibit the analysis that can be done

Streamflow

Groundwater levels



Precipitation & Climate



Soil moisture data

Water Quality

Flux tower data

Why an Observations Data Model

- Provides a common persistence model for observations data

- Syntactic heterogeneity (File types and formats)
- Semantic heterogeneity
  - Language for observation attributes (structural)
  - Language to encode observation attribute values (contextual)
- Publishing and sharing research data
- Metadata to facilitate unambiguous interpretation
- Enhance analysis capability

Scope

- Focus on Hydrologic Observations made at a point
- Exclude Remote sensing or grid data.
- Primarily store raw observations and simple derived information to get data into its most usable form.
- Limit inclusion of extensively synthesized information and model outputs at this stage.

What are the basic attributes to be associated with each single data value and how can these best be organized?

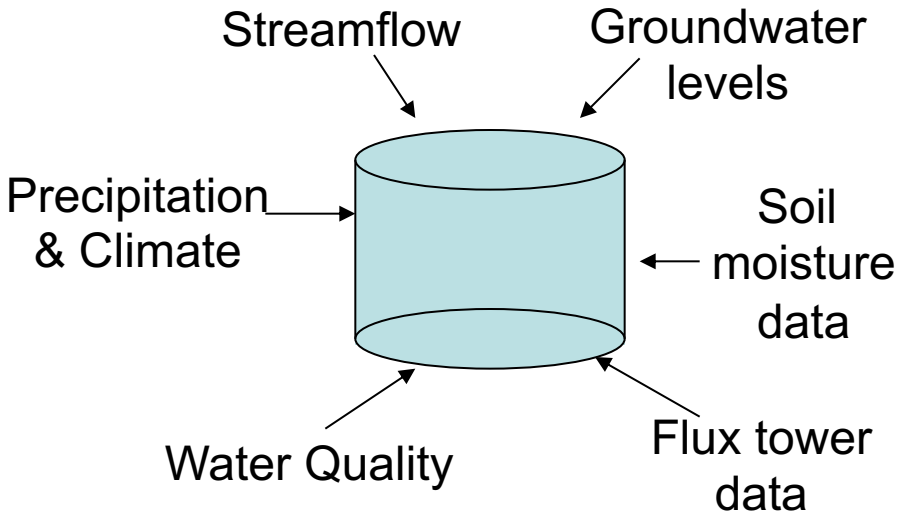
Value
DateTime
Variable
Location
Units
Interval (support)
Accuracy

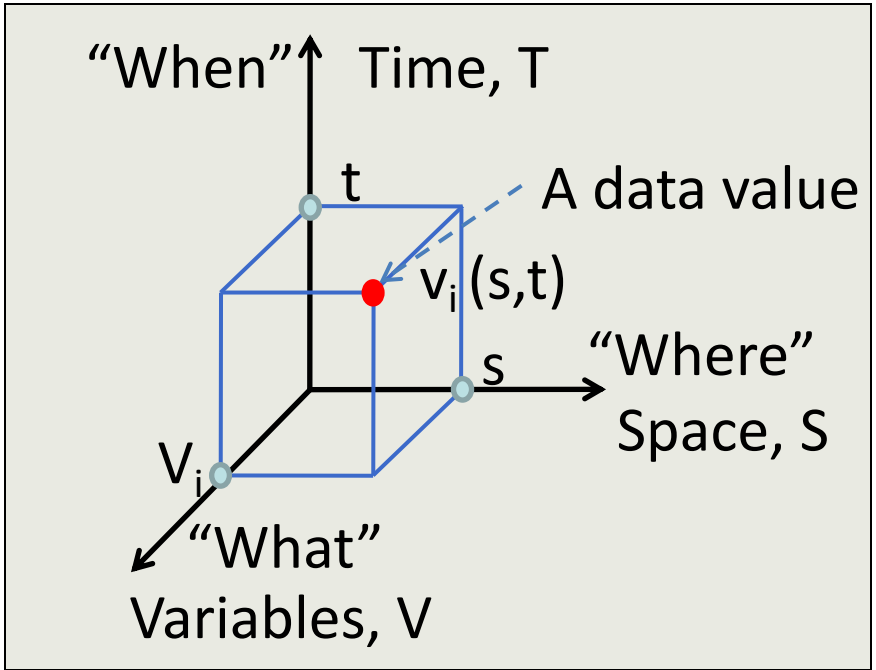
Offset
OffsetType/ Reference Point
Source/Organization
Censoring
Data Qualifying Comments

Method
Quality Control Level
Sample Medium
Value Type
Data Type

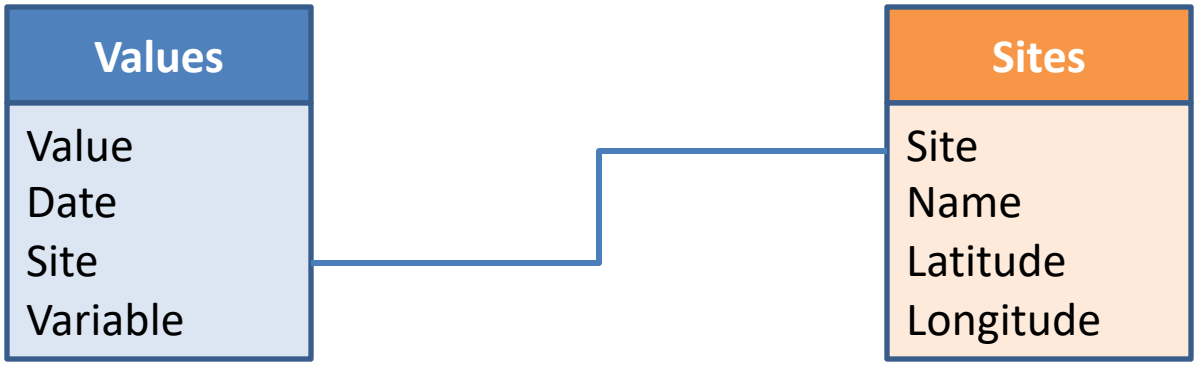
CUAHSI Observations Data Model

- A relational database at the single observation level (atomic model)
- Stores observation data made at points
- Metadata for unambiguous interpretation
- Traceable heritage from raw measurements to usable information
- Standard format for data sharing
- Cross dimension retrieval and analysis





Data Storage – Relational Database



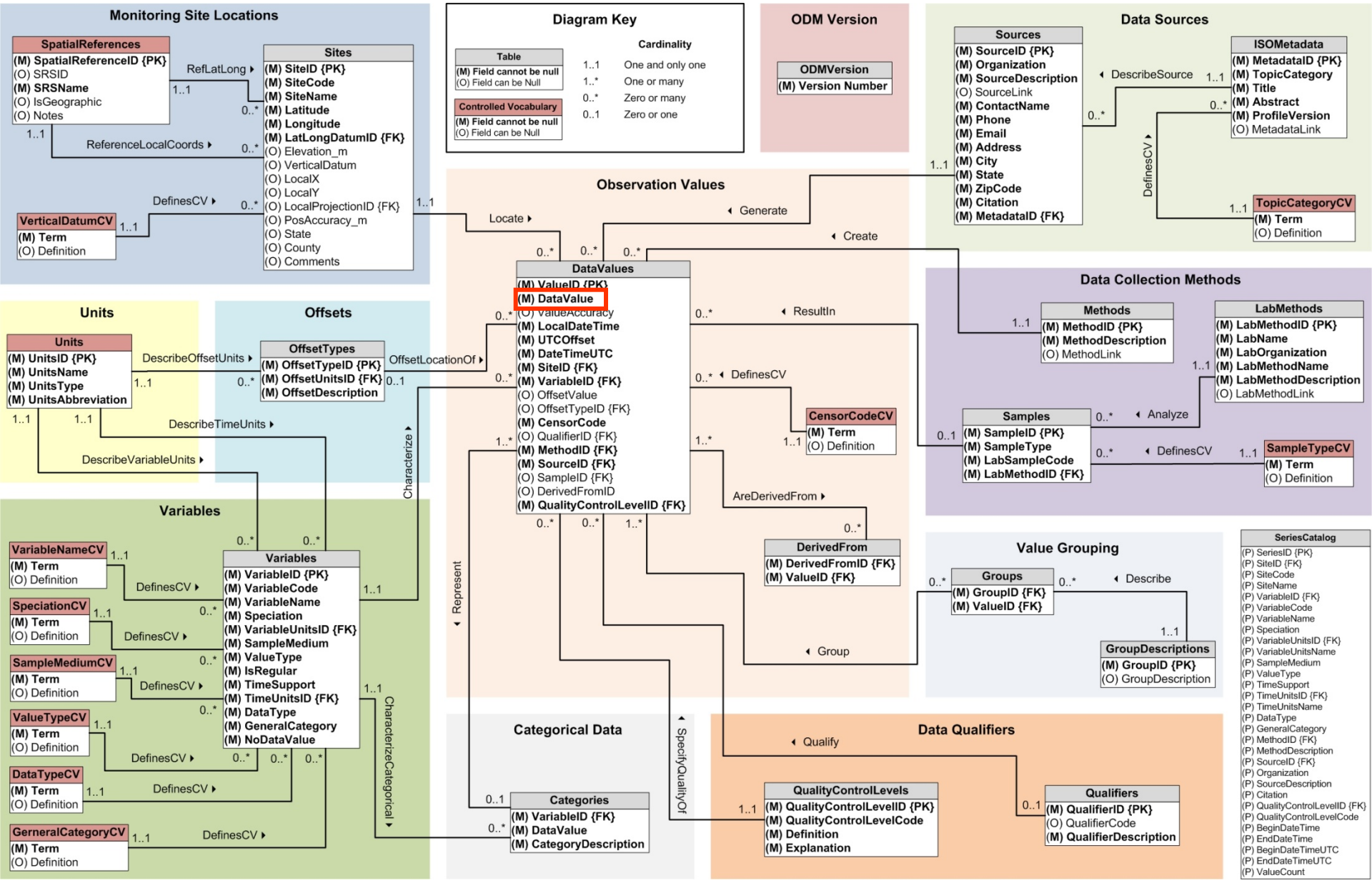
Value	Date	Site	Variable
4.5	3/3/2007	1	Streamflow
4.2	3/4/2007	1	Streamflow
33	3/3/2007	2	Temperature
34	3/4/2007	2	Temperature

Site	Name	Latitude	Longitude
1	Cane Creek	41.1	-103.2
2	Town Lake	40.3	-103.3

Why Use a RDBMS

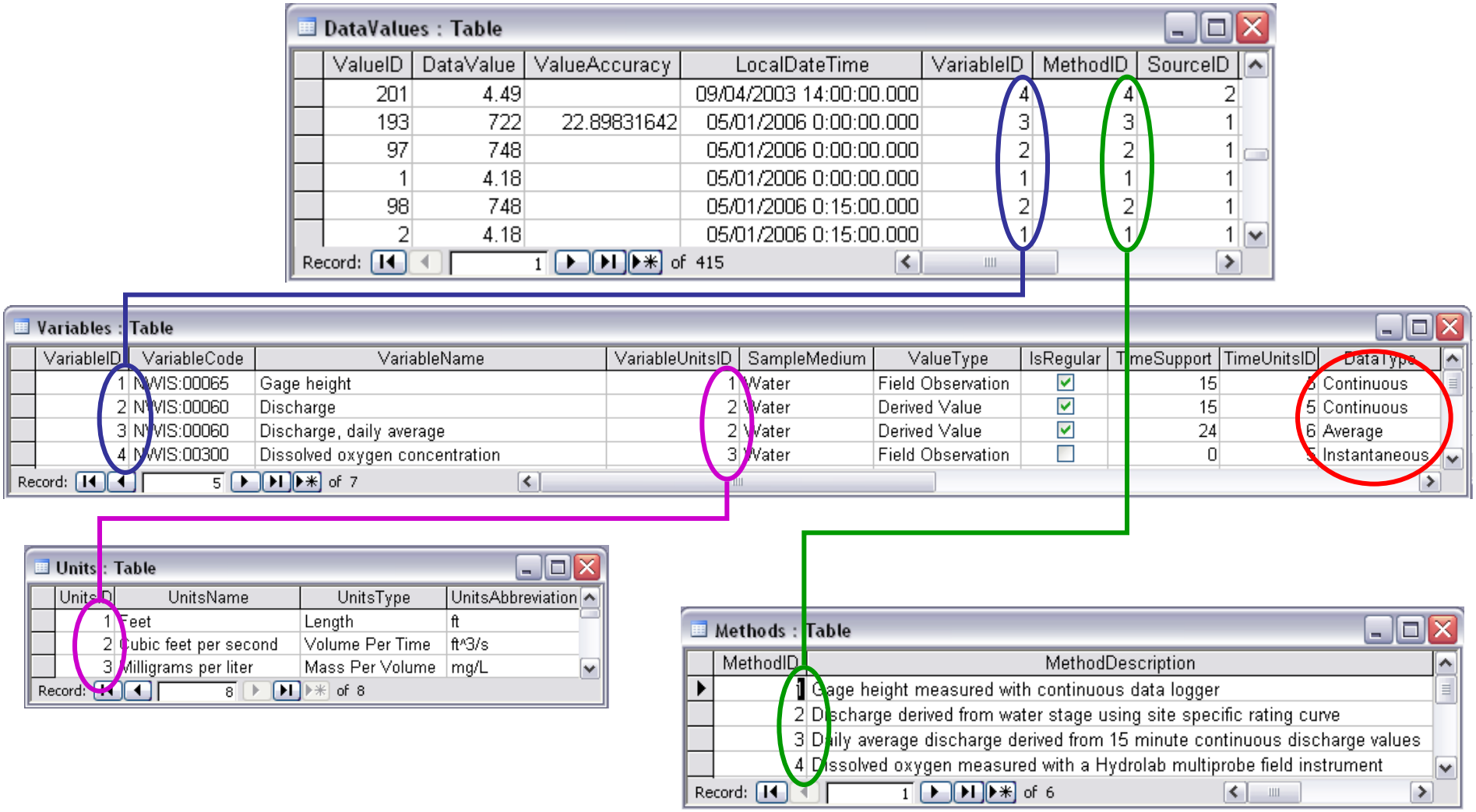
- Mature and stable technology
- Structured Query Language (SQL)
- Sharing of data among multiple applications
  - Data integrity and security
  - Access by multiple users at the same time
  - Tools for backup and recovery
- Reduced application development time

CUAHSI Observations Data Model http://his.cuahsi.org/odmdatabases.html



Horsburgh, J. S., D. G. Tarboton, D. R. Maidment and I. Zaslavsky, (2008), A Relational Model for Environmental and Water Resources Data, *Water Resour. Res.,* 44: W05406, doi:10.1029/2007WR006392.

Discharge, Stage, Concentration and Daily Average Example



DataValues : Table

ValueID	DataValue	ValueAccuracy	LocalDateTime	VariableID	MethodID	SourceID
201	4.49		09/04/2003 14:00:00.000	4	4	2
193	722	22.89831642	05/01/2006 0:00:00.000	3	3	1
97	748		05/01/2006 0:00:00.000	2	2	1
1	4.18		05/01/2006 0:00:00.000	1	1	1
98	748		05/01/2006 0:15:00.000	2	2	1
2	4.18		05/01/2006 0:15:00.000	1	1	1

Record: 1 of 415

Variables : Table

VariableID	VariableCode	VariableName	VariableUnitsID	SampleMedium	ValueType	IsRegular	TimeSupport	TimeUnitsID	DataType
1	NWIS:00065	Gage height	1	Water	Field Observation	☑	15	5	Continuous
2	NWIS:00060	Discharge	2	Water	Derived Value	☑	15	5	Continuous
3	NWIS:00060	Discharge, daily average	2	Water	Derived Value	☑	24	6	Average
4	NWIS:00300	Dissolved oxygen concentration	3	Water	Field Observation	☐	0	5	Instantaneous

Record: 5 of 7

Units : Table

UnitsID	UnitsName	UnitsType	UnitsAbbreviation
1	Feet	Length	ft
2	Cubic feet per second	Volume Per Time	ft^3/s
3	Milligrams per liter	Mass Per Volume	mg/L

Record: 8 of 8

Methods : Table

MethodID	MethodDescription
1	Gage height measured with continuous data logger
2	Discharge derived from water stage using site specific rating curve
3	Daily average discharge derived from 15 minute continuous discharge values
4	Dissolved oxygen measured with a Hydrolab multiprobe field instrument

Record: 1 of 6

Independent of, but can be coupled to Geographic Representation

ODM

Observations Data Model

Sites
SiteID
SiteCode
SiteName
Latitude
Longitude
...

1

OR

CouplingTable
SiteID
HydroID

1

1

1

e.g. Arc Hydro

Feature

HydroPoint
HydroID
HydroCode
FType
Name
JunctionID

Waterbody
HydroID
HydroCode
FType
Name
AreaSqKm
JunctionID

Watershed
HydroID
HydroCode
DrainID
AreaSqKm
JunctionID
NextDownID

ComplexEdgeFeature

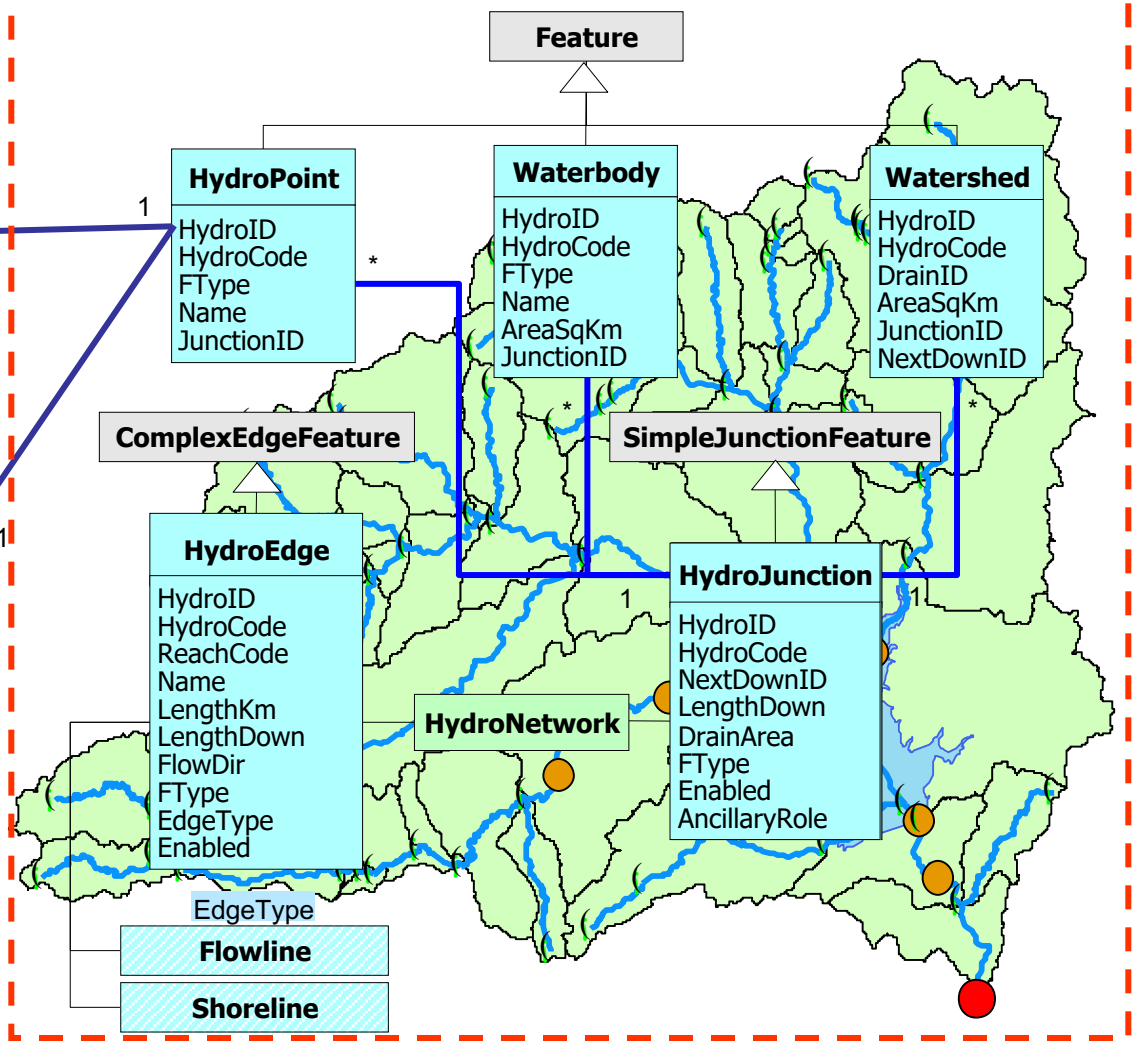
SimpleJunctionFeature

HydroEdge
HydroID
HydroCode
ReachCode
Name
LengthKm
LengthDown
FlowDir
FType
EdgeType
Enabled

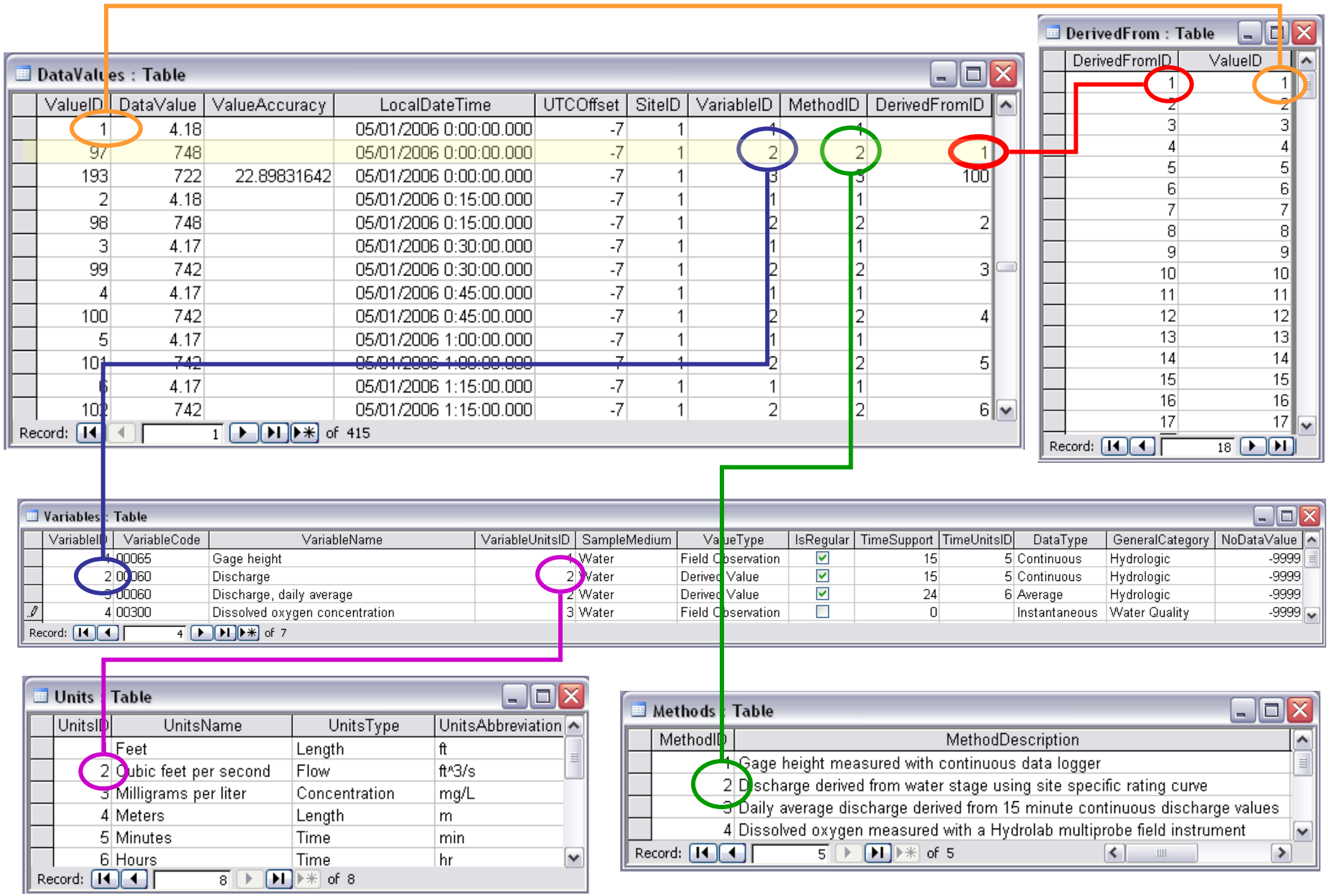
HydroJunction
HydroID
HydroCode
NextDownID
LengthDown
DrainArea
FType
Enabled
AncillaryRole

HydroNetwork

EdgeType
Flowline
Shoreline



Stage and Streamflow Example



DataValues : Table

ValueID	DataValue	ValueAccuracy	LocalDateTime	UTCOffset	SiteID	VariableID	MethodID	DerivedFromID
1	4.18		05/01/2006 0:00:00.000	-7	1	1	1	
97	748		05/01/2006 0:00:00.000	-7	1	2	2	1
193	722	22.89831642	05/01/2006 0:00:00.000	-7	1	3	3	100
2	4.18		05/01/2006 0:15:00.000	-7	1	1	1	
98	748		05/01/2006 0:15:00.000	-7	1	2	2	2
3	4.17		05/01/2006 0:30:00.000	-7	1	1	1	
99	742		05/01/2006 0:30:00.000	-7	1	2	2	3
4	4.17		05/01/2006 0:45:00.000	-7	1	1	1	
100	742		05/01/2006 0:45:00.000	-7	1	2	2	4
5	4.17		05/01/2006 1:00:00.000	-7	1	1	1	
101	742		05/01/2006 1:00:00.000	-7	1	2	2	5
6	4.17		05/01/2006 1:15:00.000	-7	1	1	1	
102	742		05/01/2006 1:15:00.000	-7	1	2	2	6

Record: 1 of 415

DerivedFrom : Table

DerivedFromID	ValueID
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17

Record: 18

Variables : Table

VariableID	VariableCode	VariableName	VariableUnitsID	SampleMedium	ValueType	IsRegular	TimeSupport	TimeUnitsID	DataType	GeneralCategory	NoDataValue
1	00065	Gage height	1	Water	Field Observation	☑	15	5	Continuous	Hydrologic	-9999
2	00060	Discharge	2	Water	Derived Value	☑	15	5	Continuous	Hydrologic	-9999
3	00060	Discharge, daily average	2	Water	Derived Value	☑	24	6	Average	Hydrologic	-9999
4	00300	Dissolved oxygen concentration	3	Water	Field Observation	☐	0		Instantaneous	Water Quality	-9999

Record: 4 of 7

Units : Table

UnitsID	UnitsName	UnitsType	UnitsAbbreviation
1	Feet	Length	ft
2	Cubic feet per second	Flow	ft^3/s
3	Milligrams per liter	Concentration	mg/L
4	Meters	Length	m
5	Minutes	Time	min
6	Hours	Time	hr

Record: 8 of 8

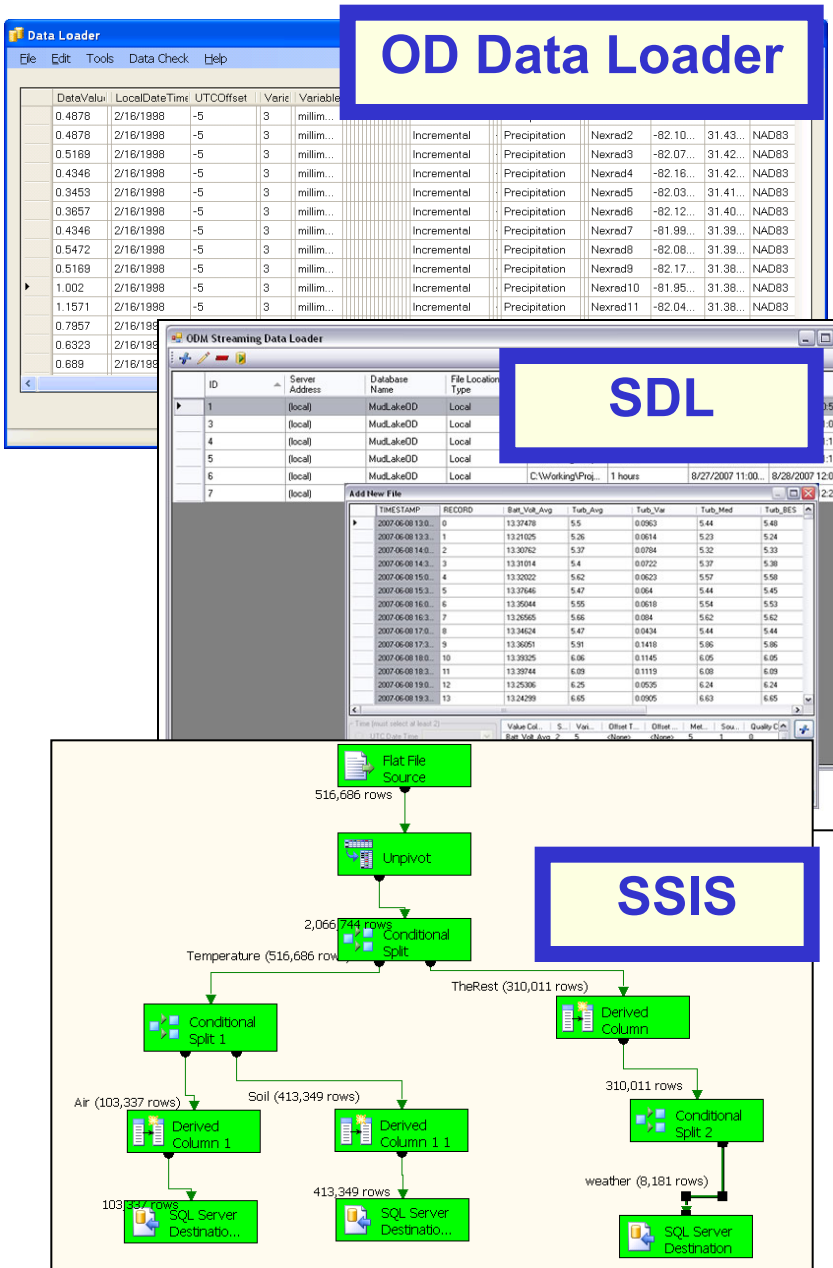
Methods : Table

MethodID	MethodDescription
1	Gage height measured with continuous data logger
2	Discharge derived from water stage using site specific rating curve
3	Daily average discharge derived from 15 minute continuous discharge values
4	Dissolved oxygen measured with a Hydrolab multiprobe field instrument

Record: 5 of 5

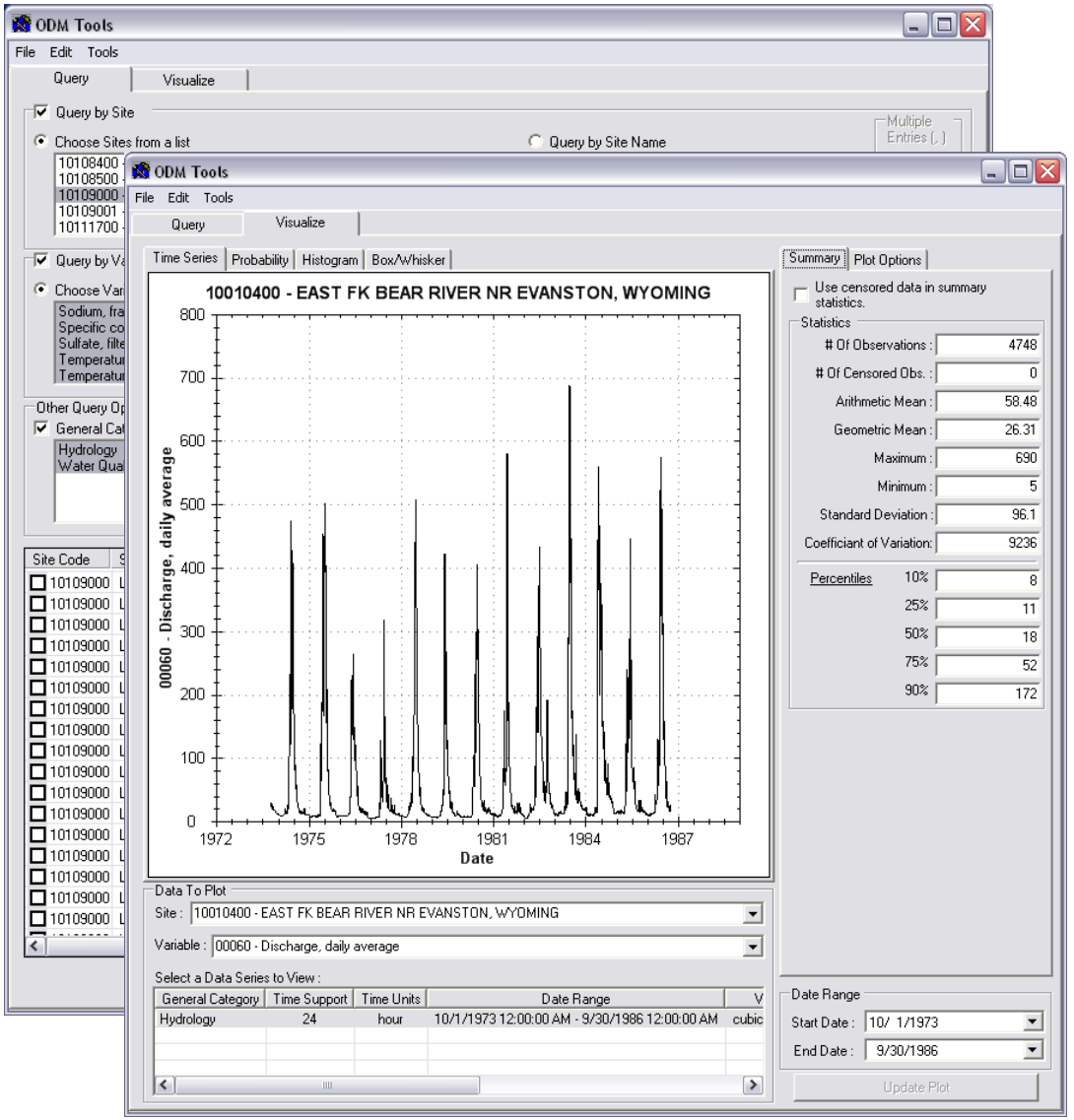
Loading data into ODM

- Interactive OD Data Loader (OD Loader)
  - Loads data from spreadsheets and comma separated tables in simple format
- Scheduled Data Loader (SDL)
  - Loads data from datalogger files on a prescribed schedule.
  - Interactive configuration
- SQL Server Integration Services (SSIS)
  - Microsoft application accompanying SQL Server useful for programming complex loading or data management functions



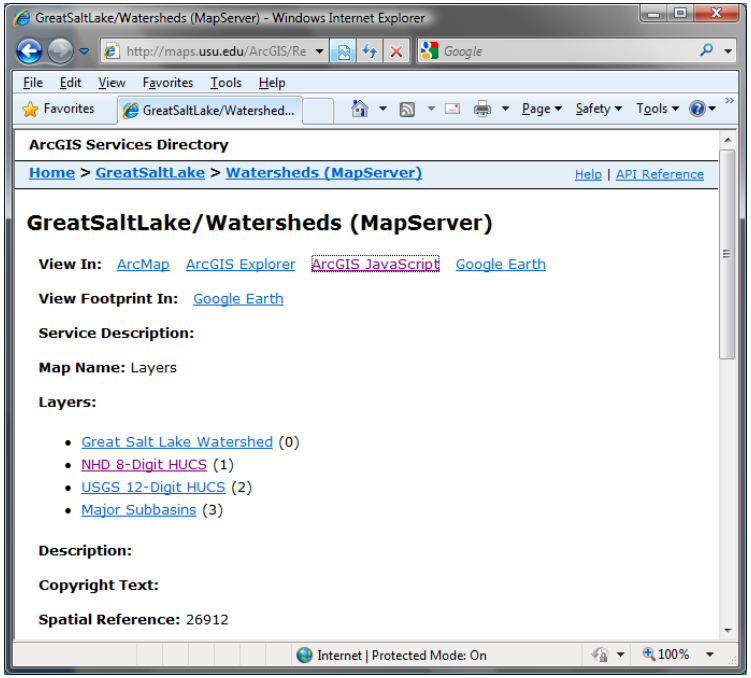
Managing Data Within ODM
ODM Tools

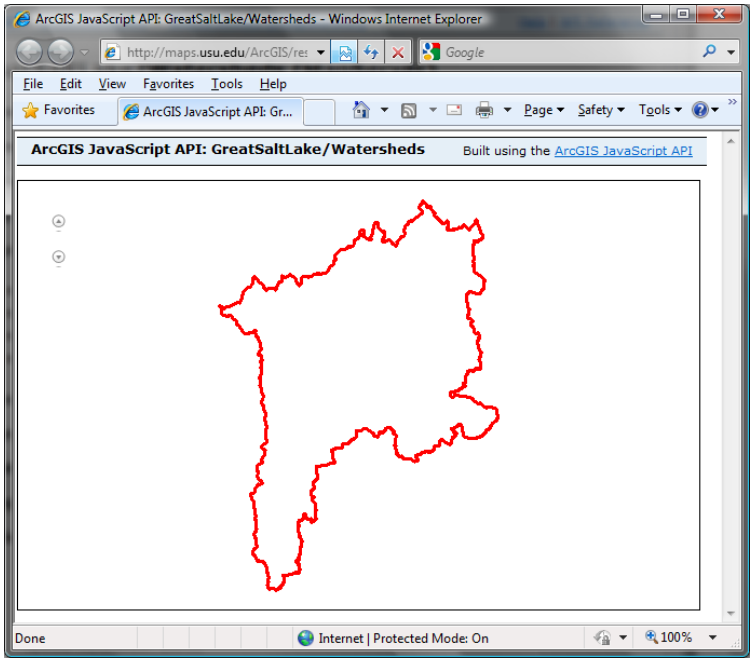
- **Query and export** – export data series and metadata
- **Visualize** – plot and summarize data series
- **Edit** – delete, modify, adjust, interpolate, average, etc.



Publication of Spatial (GIS) Datasets

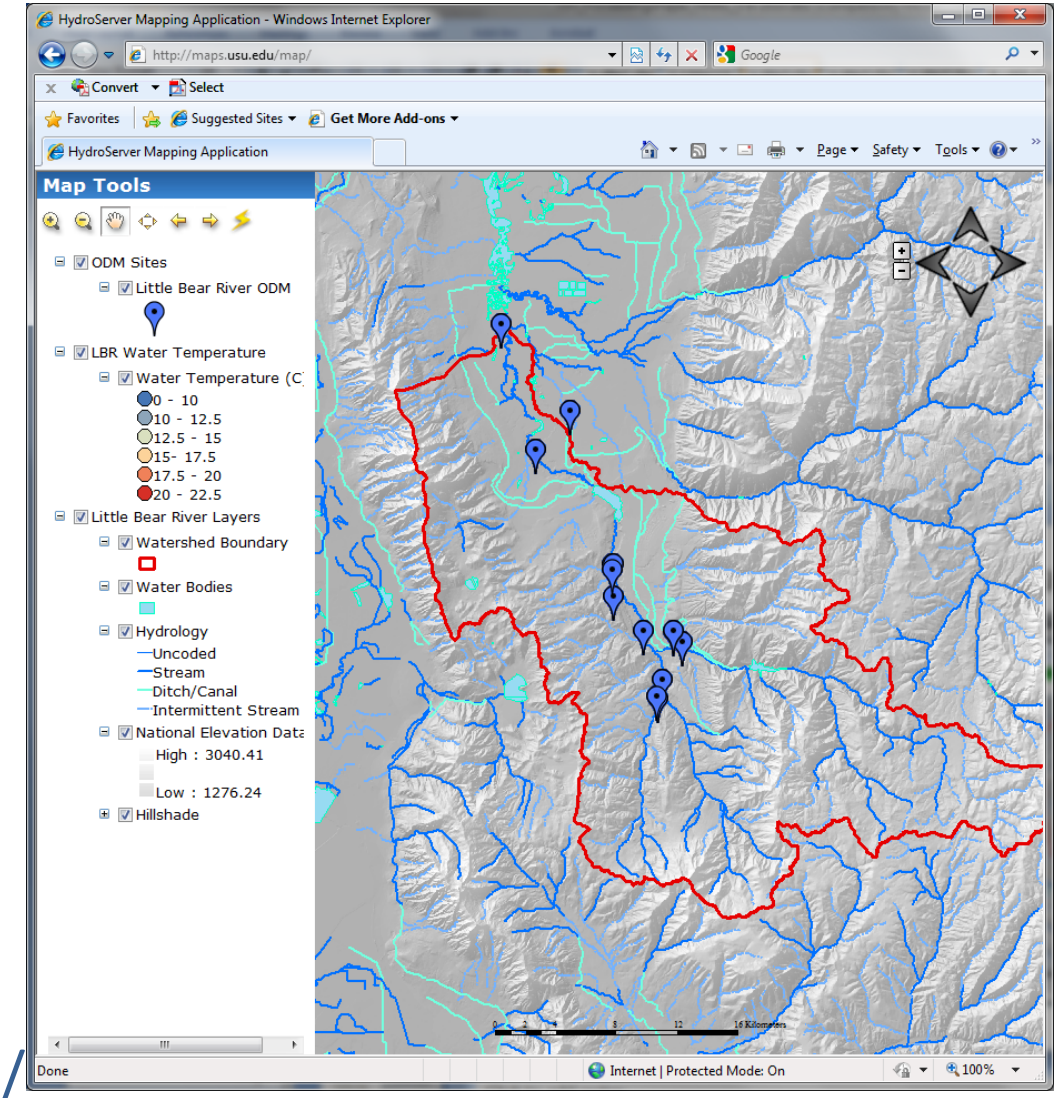
- Publishing spatial datasets using ArcGIS Server
  - Using OGC standards that can be consumed by a number of GIS clients
  - WMS, WFS, WCS





Data Presentation Via a Map Interface

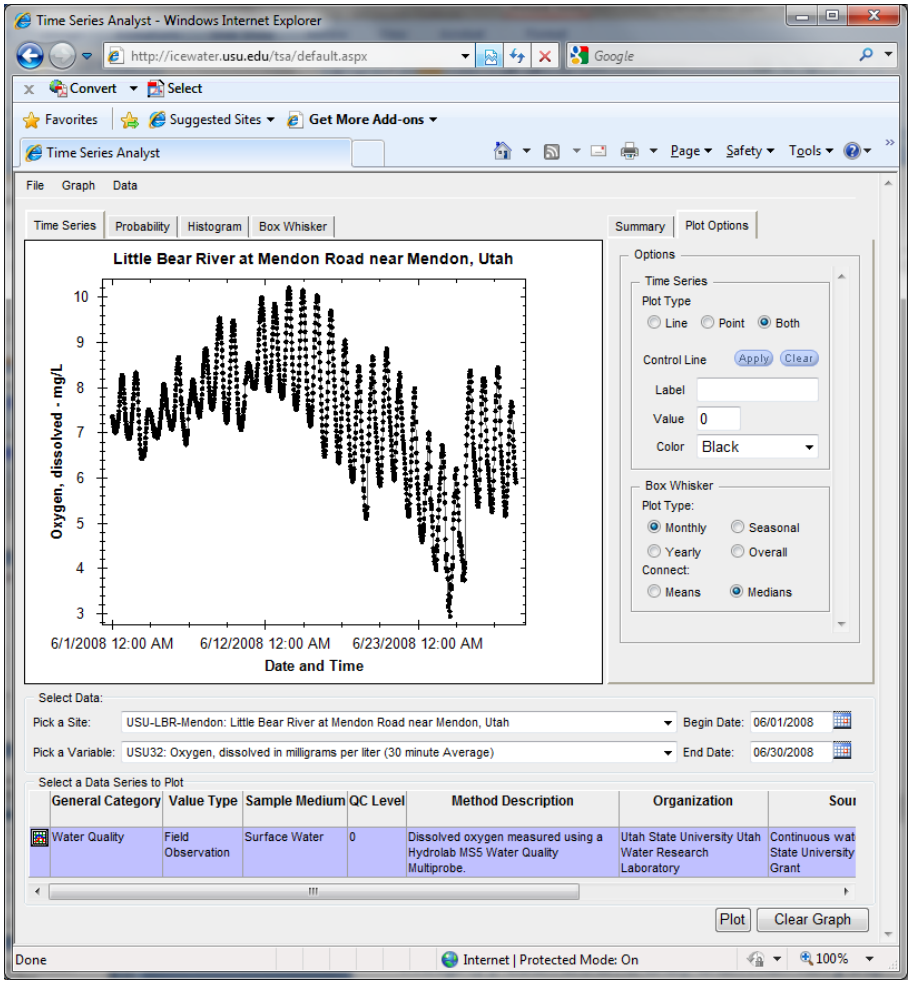
- Internet Map Server built using ArcGIS
- Web browser client
- Combine spatial data and observational data
- Launch data visualization tools
- Based on a “Region”



http://icewater.usu.edu/map/

Data Preview, Visualization, and Analysis
Time Series Analyst

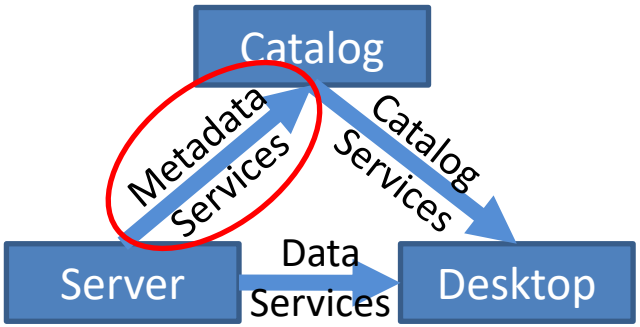
- Web Browser Client
- Multiple ODM Database Support
- Variety of plot types
- Descriptive statistics
- Linked to the map application
- Data preview and download

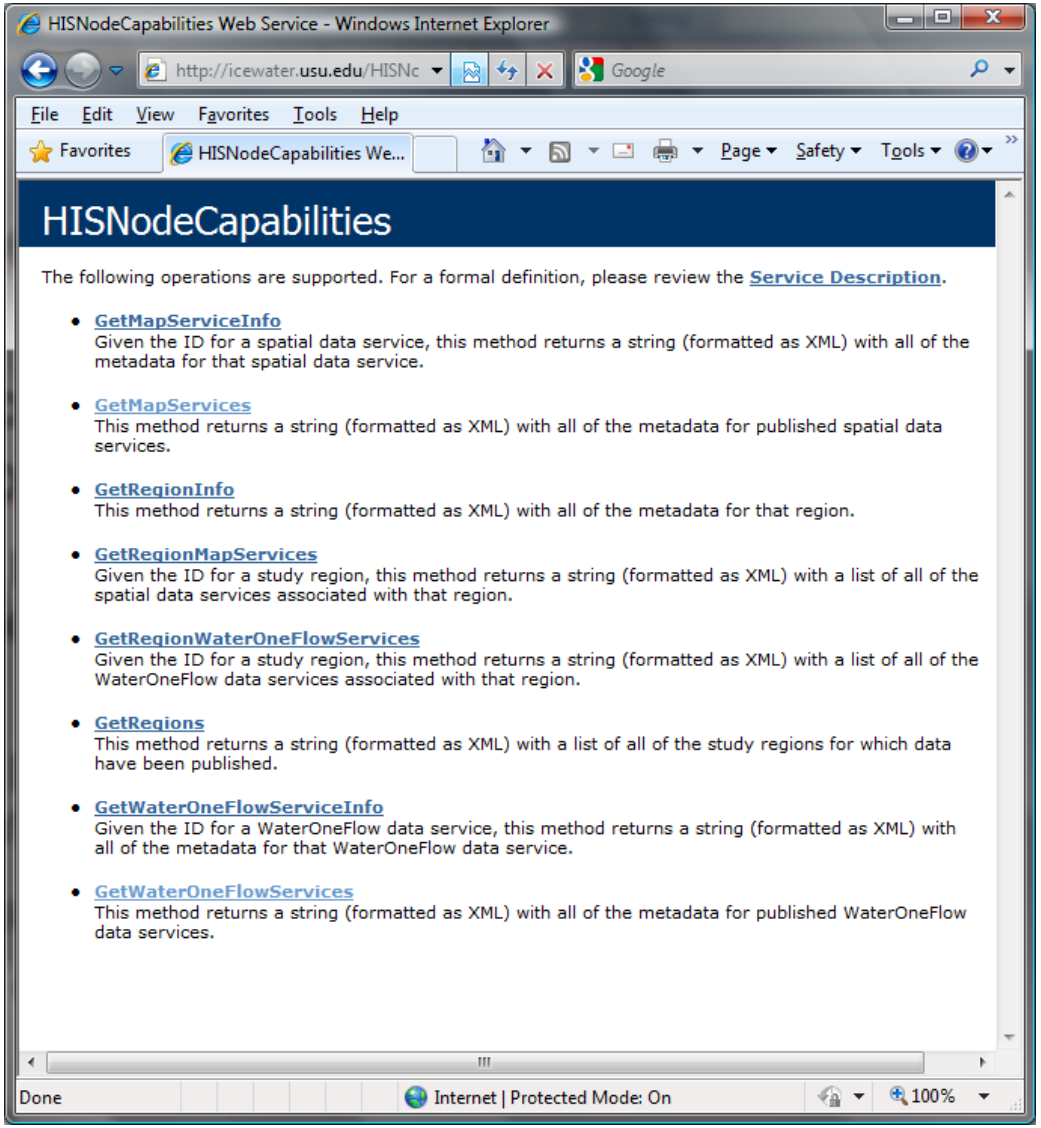


http://icewater.usu.edu/tsa/

HydroServer Capabilities Web Service

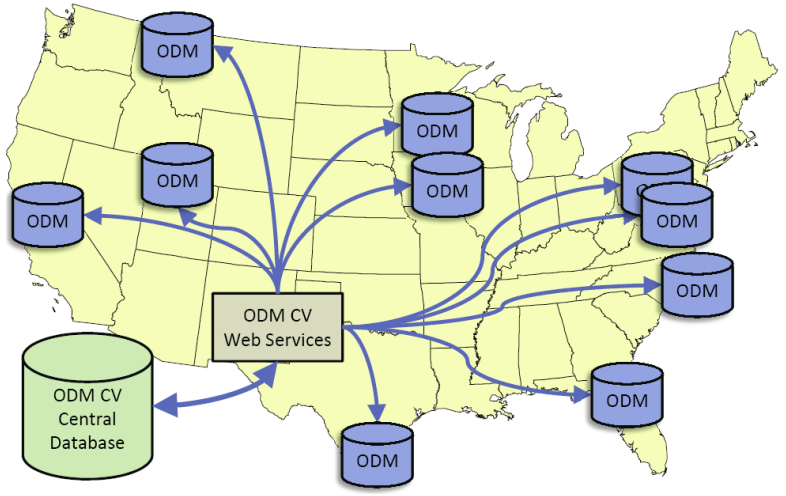
- Publish capabilities of each HydroServer
  - Listing of published observational data services
  - Listing of published spatial data services
- Supports automatic cataloging of available services at HIS Central
- **Makes HydroServers self describing**





Overcoming Semantic Heterogeneity

- ODM Controlled Vocabulary System
  - ODM CV central database
  - Online submission and editing of CV terms
  - Web services for broadcasting CVs



ODM VariableNameCV

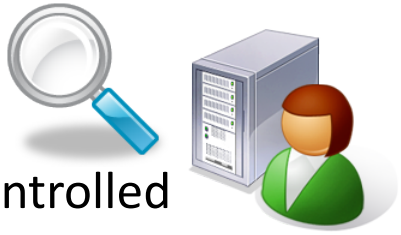
Variable Name	
Investigator 1:	“Temperature, water”
Investigator 2:	“Water Temperature”
Investigator 3:	“Temperature”
Investigator 4:	“Temp.”

Term
…
Sunshine duration
Temperature
Turbidity
…

From Jeff Horsburgh

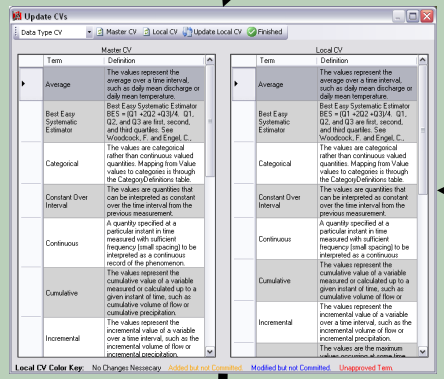
Dynamic controlled vocabulary moderation system

ODM Data Manager

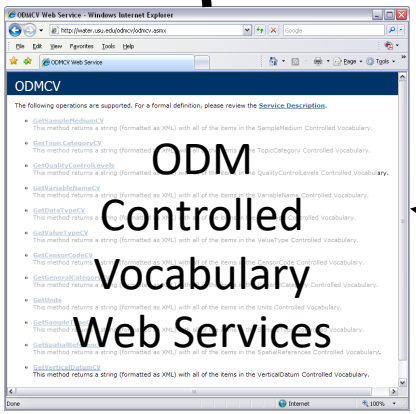


ODM Controlled Vocabulary Moderator

ODM Tools



XML



Local ODM Database

Local Server

Master ODM Controlled Vocabulary

http://his.cuahsi.org/mastercvreg.html

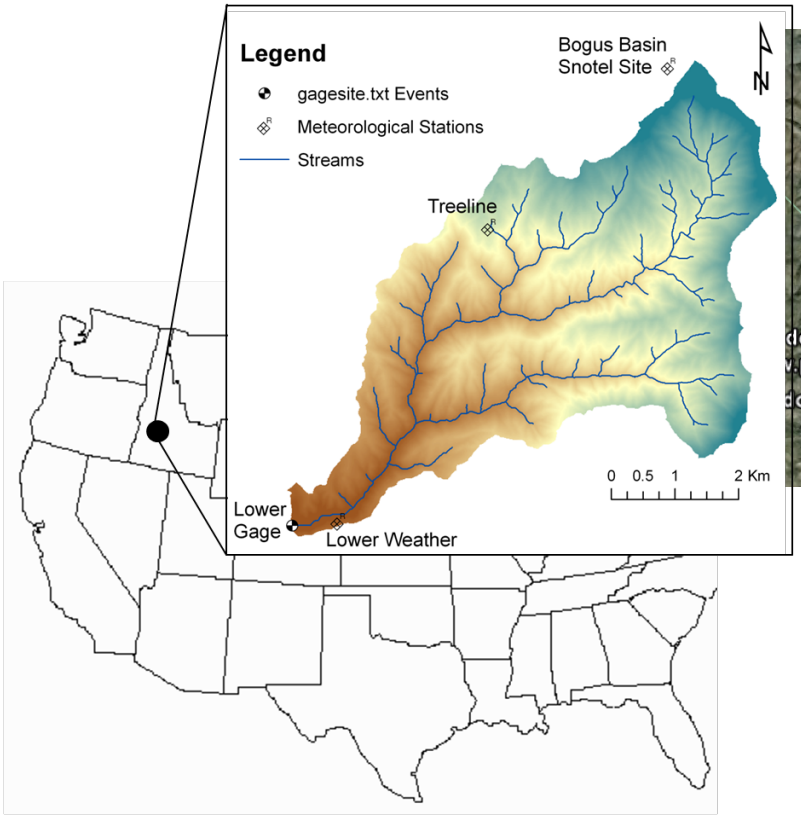
From Jeff Horsburgh

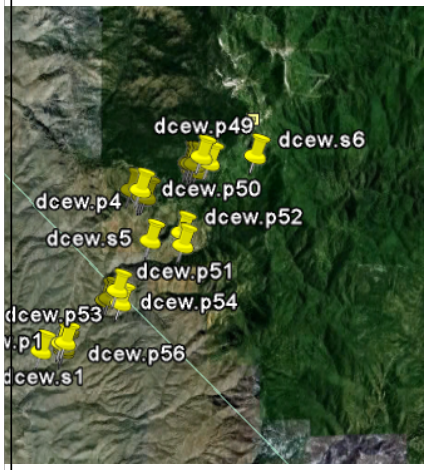
37 Water Data Services on HIS Central from 12 Universities

- University of Maryland, Baltimore County
- Montana State University
- University of Texas at Austin
- University of Iowa
- Utah State University
- University of Florida
- University of New Mexico
- University of Idaho
- Boise State University
- University of Texas at Arlington
- University of California, San Diego
- Idaho State University

Dry Creek Experimental Watershed (DCEW)

(28 km^2 semi-arid steep topography, Boise Front)



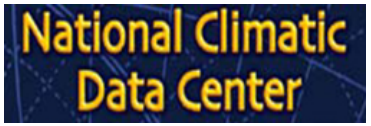


68 Sites
24 Variables
4,700,000+ values

Published by Jim McNamara, Boise State University

Water Agencies and Industry

- USGS, NCDC , Corps of Engineers publishing data using HIS WaterML
- OGC Hydrology Domain Working Group evaluating WaterML as OGC standard
- ESRI using CUAHSI model in ArcGIS.com GIS data collaboration portal
- Kisters WISKI support for WaterML data publication
- Australian Water Resources Information System Water Accounting System has adopted aspects of HIS
- NWS West Gulf River Forecast Center Multi-sensor Precipitation Estimate published from ODM using WaterML



Federal Agency Water Data Services at HISCentral (10/2010)

Network Name	Site Count	Value Count	Earliest Observation	Notes
NWISDV	31,800	304,000,000	01/01/1861	WaterML-compliant GetValues service from NWIS, catalog ingested
EPA	236,000	78,000,000	01/11/1900	SOAP wrapper over WQX services, catalog ingested
NWISUV	11,800	84,500,000	60 DAYS	WaterML-compliant GetValues Service, catalog ingested
NCDC ISH	11,600	3,000,000*	1/1/2005	WaterML-compliant GetValues service from NCDC
NCDC ISD	24,800	18,200,000	1/1/1892	WaterML-compliant GetValues service from NCDC
NWISIID	376,000	86500,000	9/1/1867	SOAP wrapper over NWIS web site, catalog ingested
NWISGW	834,000	8,490,000	1/1/1800	SOAP wrapper over NWIS web site, catalog ingested
RIVERGAGES	1,300	264,000,000	1/1/2000	WaterML compliant REST services from Army Corps of Engineers

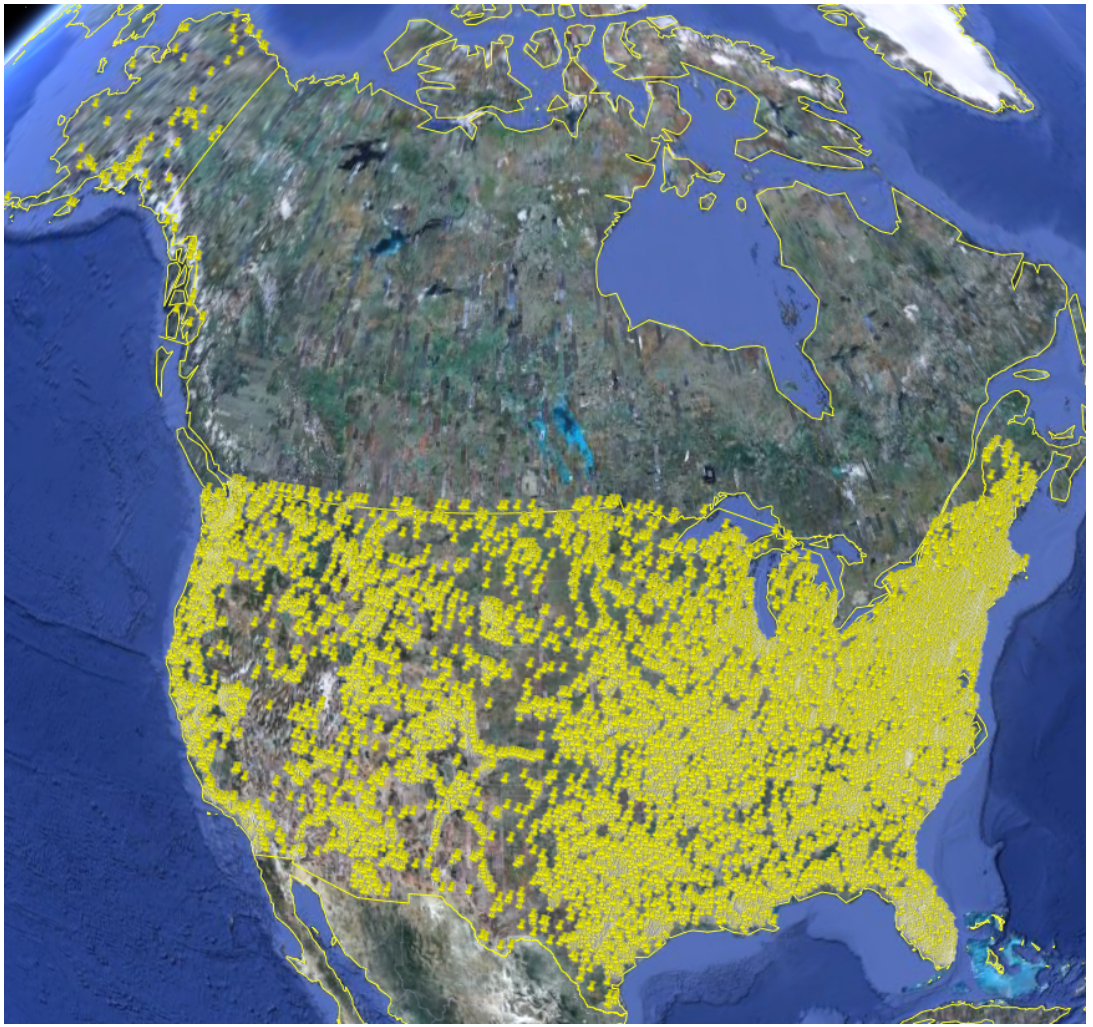
* Estimated

USGS Unit Values Data

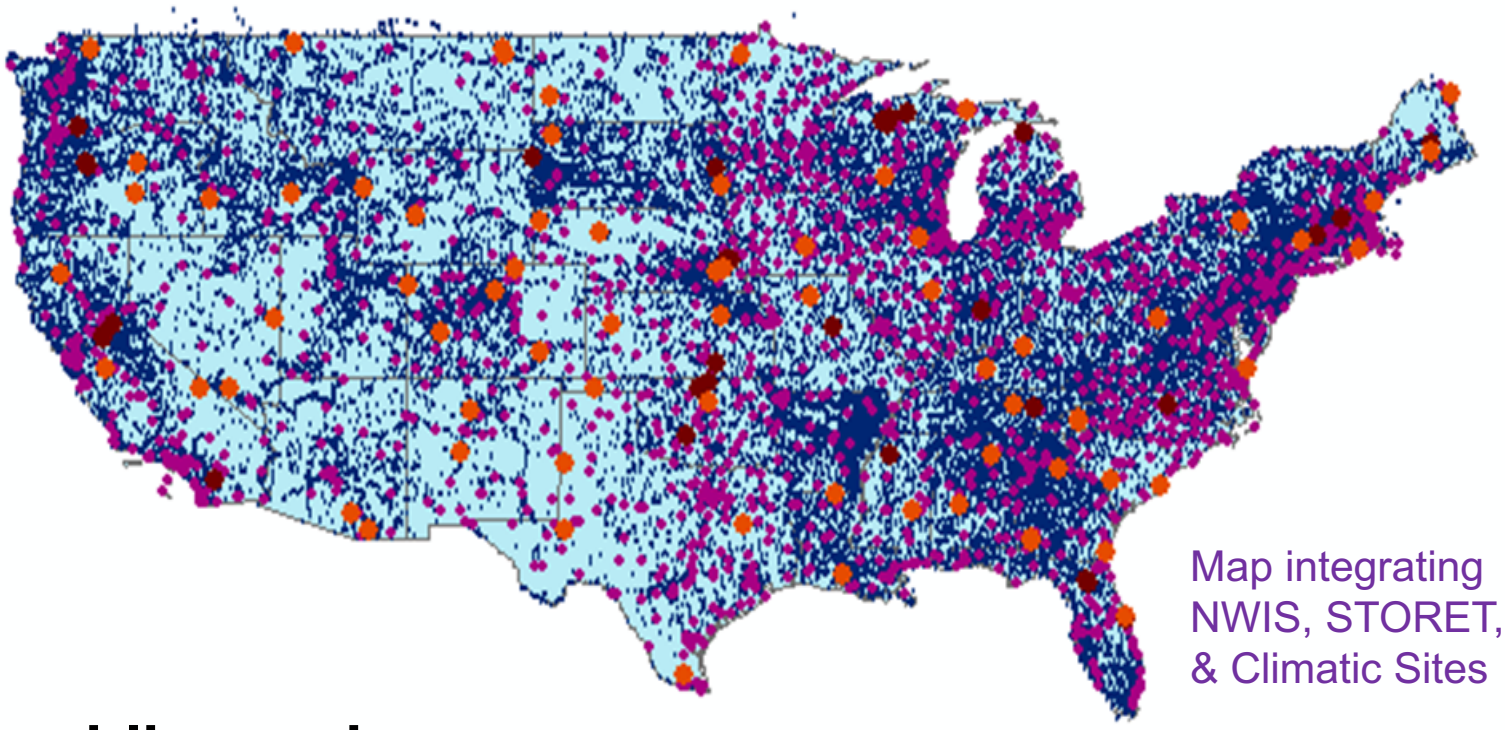
Real time streamflow data over the last 60 days

11188 sites, nationally for the US

Published by USGS National Water Information System



HISCentral Content (11/2010)



Map integrating
NWIS, STORET,
& Climatic Sites

58 public services
18,000+ variables
1.96+ million sites
23.3 million series
Referencing 5.1 billion data values

Available via HISCentral discovery services

Available via GetValues requests

Summary

- **Data Storage** in an *Observations Data Model* (ODM) and publication through HydroServer
- **Data Access** through internet-based *Water Data Services* using a consistent data language, called WaterML from HydroDesktop
- **Data Discovery** through a *National Water Metadata Catalog* and thematic keyword search system at HIS Central
- **Integrated Modeling and Analysis** within HydroDesktop

The combination of these capabilities creates a common window on water observations data for the United States unlike any that has existed before.

- Learn about the CUAHSI-HIS System
- Share your work with information systems and large scale datasets
- Share your use of hydrologic data for teaching
- Interact with other users
- Share your work linking data and modeling
- Show science enabled by HIS
- Hands-on workshops
- Contribute to the future of HIS

For information on presenting or attending see:

http://his.cuahsi.org/conference2011

Contact: David.Tarboton@usu.edu

Thanks! HIS Project Team and Sponsors

- University of Texas at Austin – David Maidment, Tim Whiteaker, James Seppi, Fernando Salas, Jingqi Dong, Harish Sangireddy
- San Diego Supercomputer Center – Ilya Zaslavsky, David Valentine, Tom Whitenack, Matt Rodriguez
- Utah State University – Jeff Horsburgh, Kim Schreuders, Stephanie Reeder, Edward Wai Tsui, Ravichand Vegiraju, Ketan Patil
- University of South Carolina – Jon Goodall, Anthony Castronova
- Idaho State University – Dan Ames, Ted Dunsford, Jiří Kadlec, Yang Cao, Dinesh Grover
- Drexel University/CUNY – Michael Piasecki
- WATERS Network – Testbed Data Managers
- CUAHSI Program Office – Rick Hooper, Yoori Choi, Conrad Matiuk
- ESRI – Dean Djokic, Zichuan Ye

CUAHSI HIS

Sharing hydrologic data

http://his.cuahsi.org/

NSF Support EAR 0622374