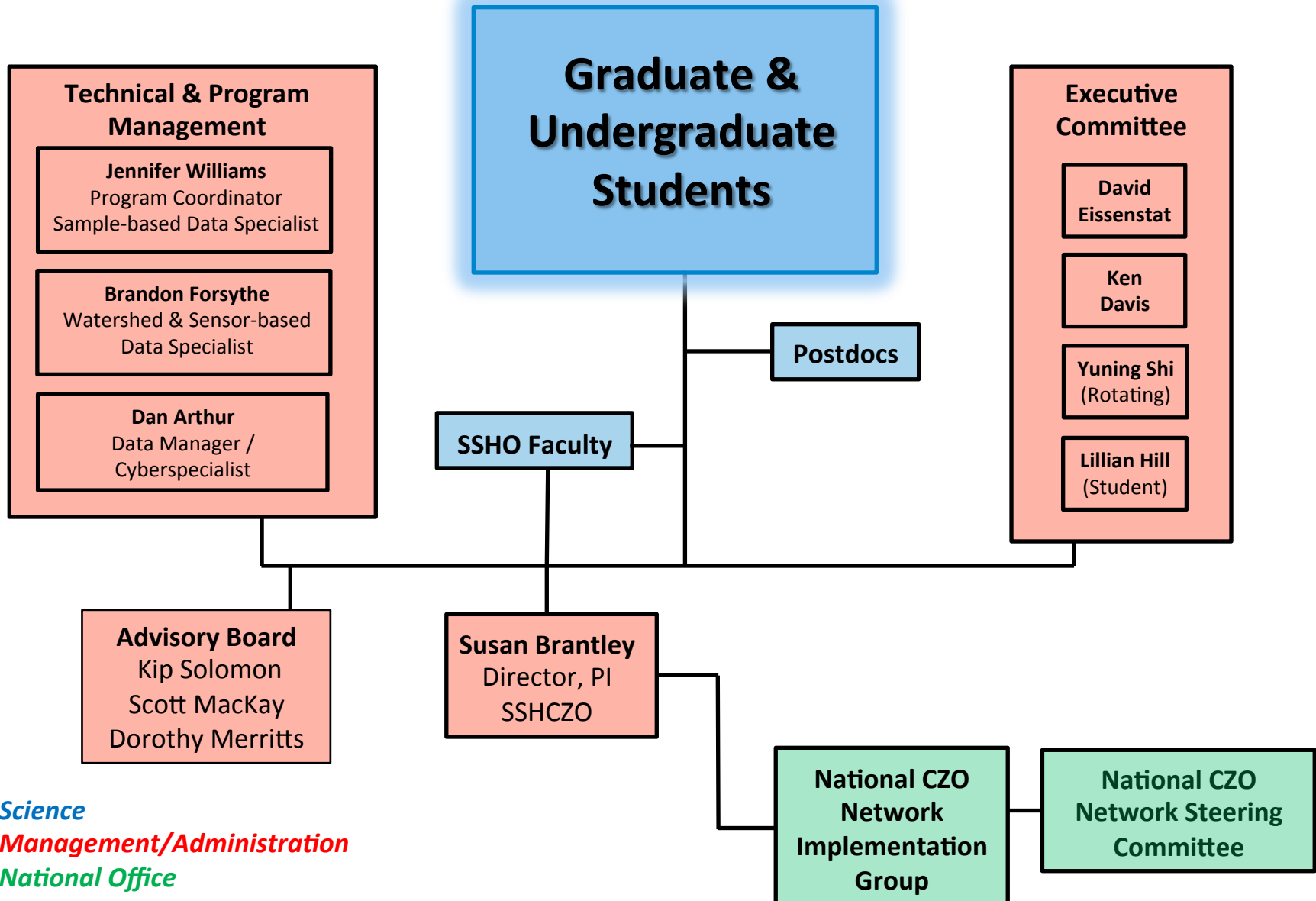


SSHCZO Data Management

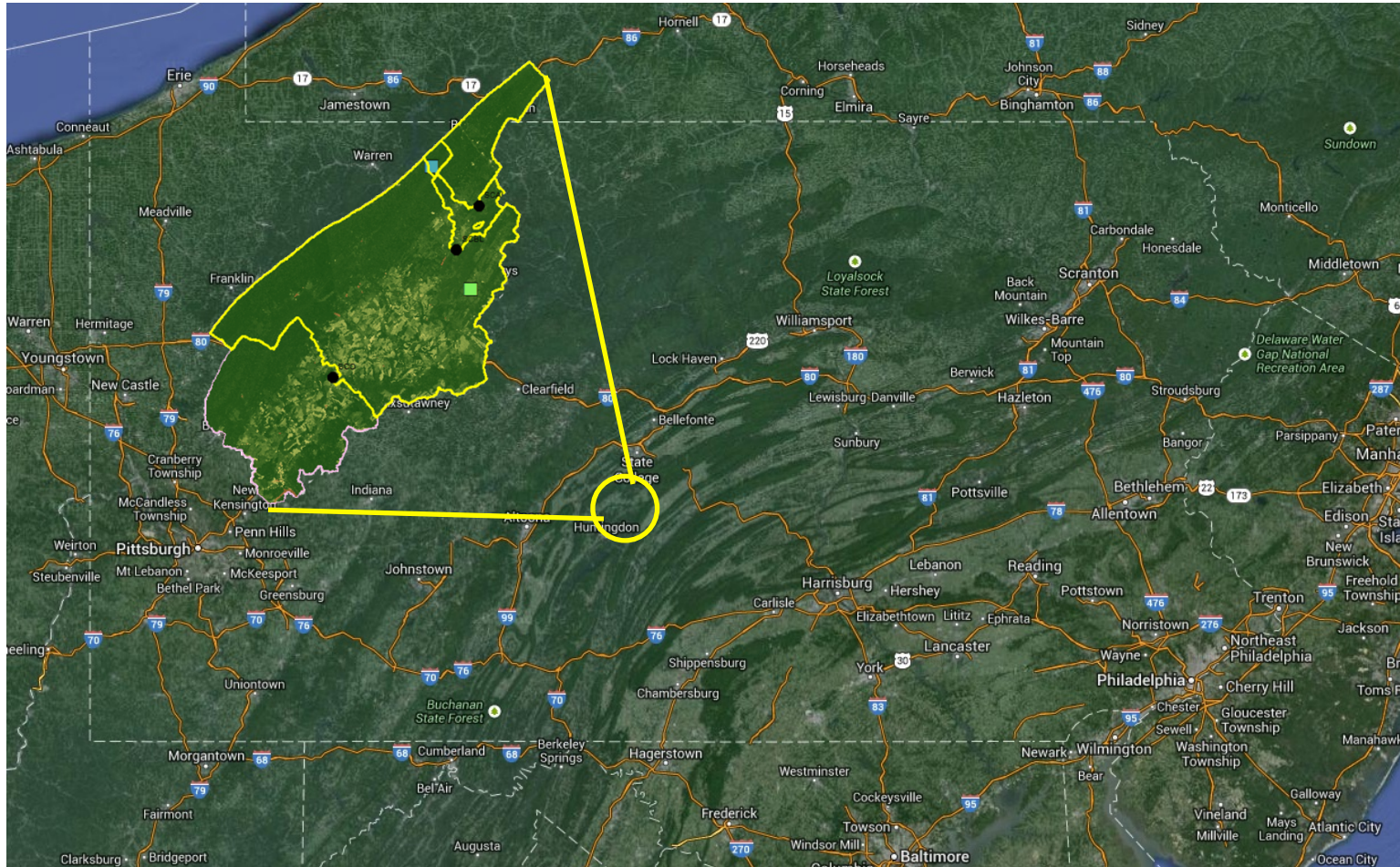
Dan Arthur

October 2015

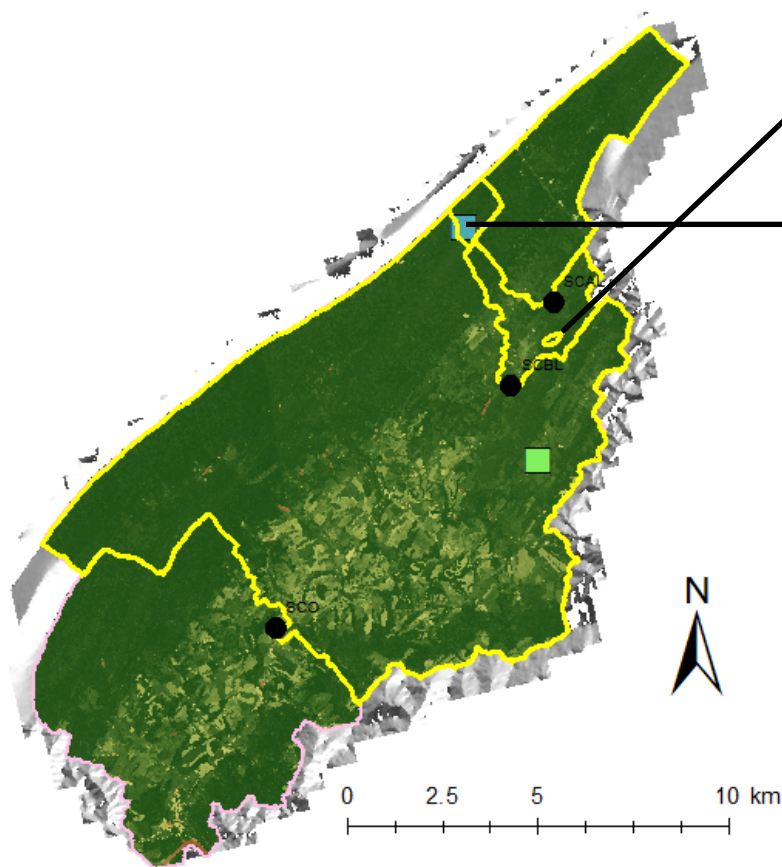
SSHCZO Organization



Shaver's Creek Watershed

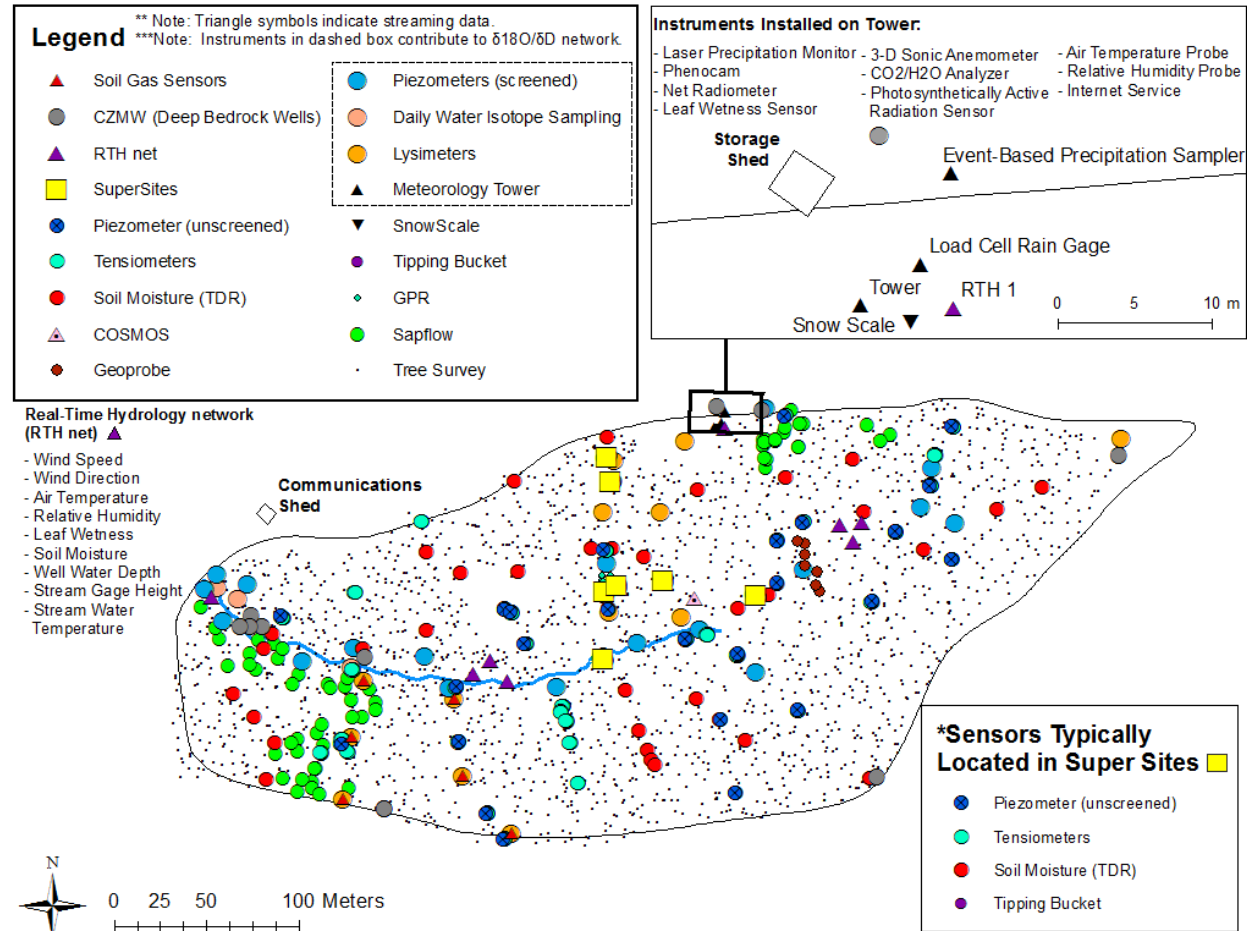


Field Sites

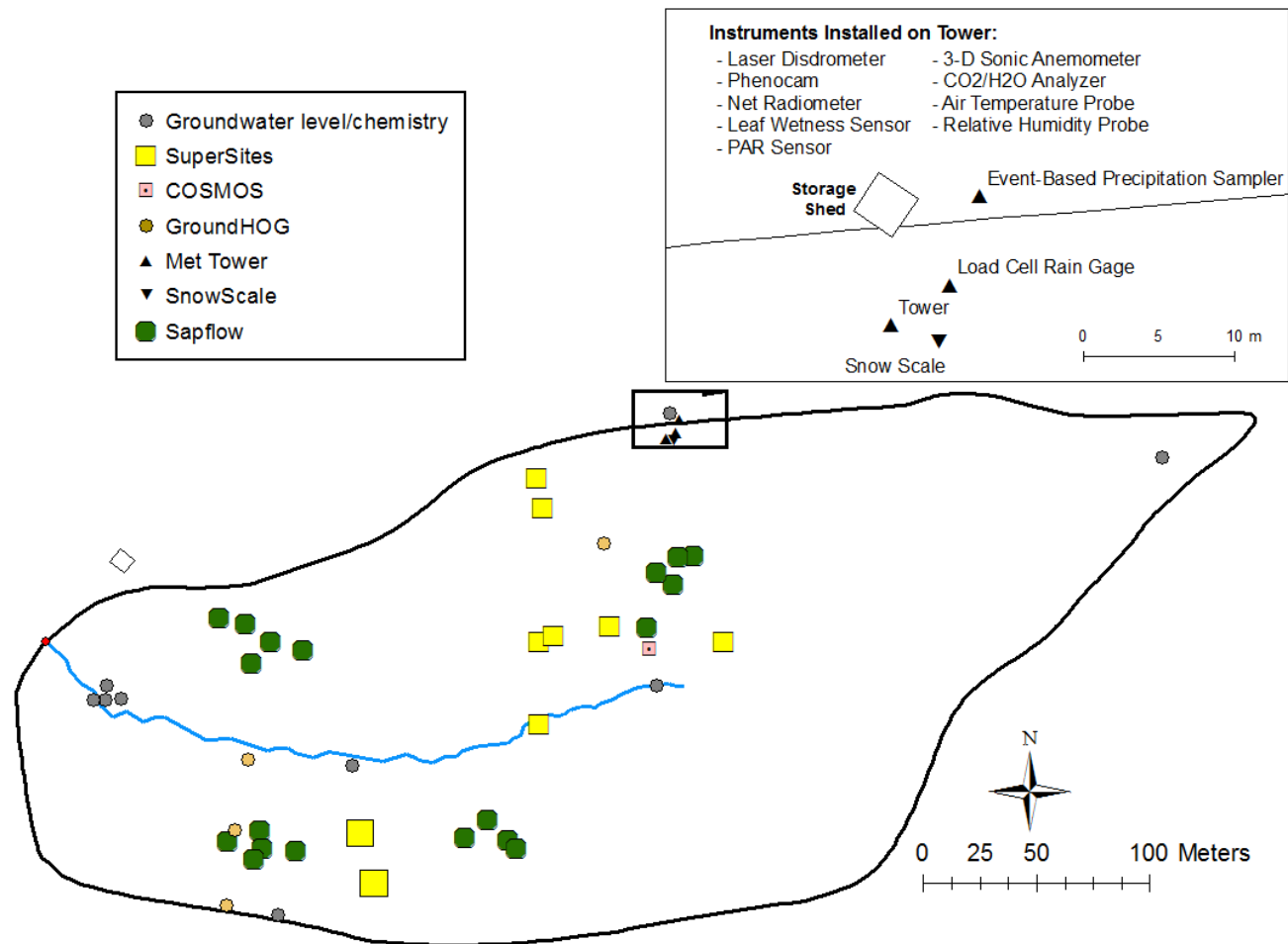


- Shale Hills
 - 0.08 km²
- Sandstone Forested Site (“Garner Run”)
 - 1.43 km²
- Shavers Creek
 - SCAL - Above Lake Perez (16.4km²)
 - SCBL - Below Lake Perez (22.4km²)
 - SCO - Outlet (126.1km²)

Shale Hills I - Instrumentation

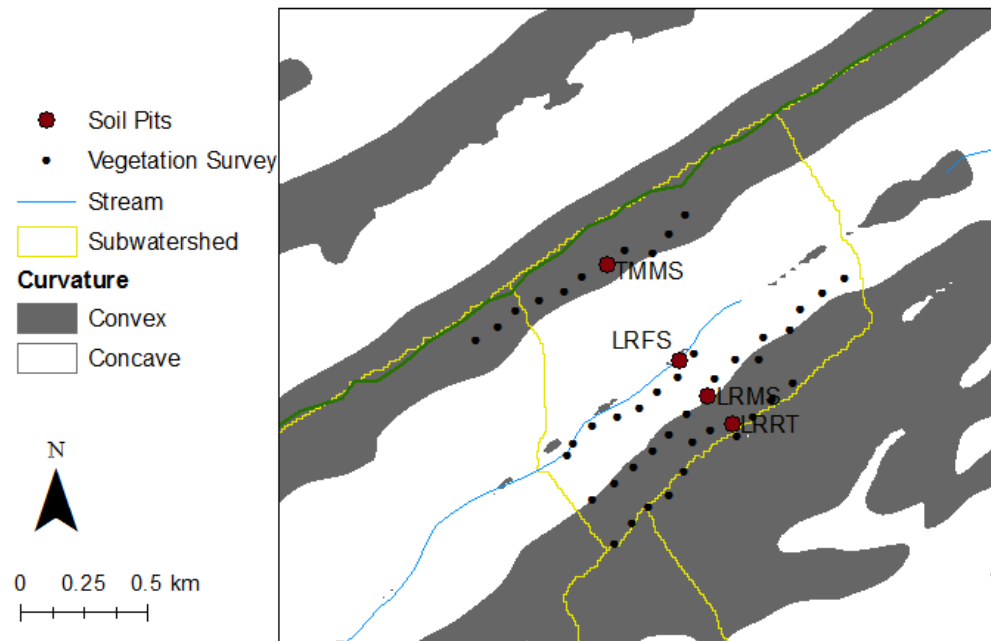


Shale Hills II - Instrumentation



Shale Hills II - Instrumentation

Sandstone Forested Site (“Garner Run”)



SSHCZO Measurement Suite

Measurement	SH	GR	SCAL	SCBL	SCO	AG
3-D Wind Velocity	KD	KD				
4-Component Radiation (upwelling & downwelling; SW & LW)	KD	KD				
Air Temperature	KD	KD				
Chemistry (DOC, TOC, DO, NO3, NH4, K, F, Cl, pH, EC, ORP)	TR	TR	TR	TR	TR	
Chemistry (major cations, Fe, Mn, Al, Si)	SB	SB				
CO2 Concentration	KD	KD				
Dendrometer-Basal Growth	DE	DE				
Digital Imagery / PhenoCam	*	DE				
Discharge	TR	TR	TR	TR	TR	
Groundwater Level	HL	HL				
Groundwater temperature	HL	HL				
H2O Concentration	KD	KD				
Hyporheic zone exchange (streambed water temperature)	TR					
In situ 10Be concentrations in surface boulders	PB	PB				

SH - Shale Hills

GR - Garner Run

SCAL - Shavers Creek - Above Lake

SCBL - Shavers Creek - Below Lake

SCO - Shavers Creek Outlet

AG - Agricultural Site

Blue – Active Data Being Collected

KD - Ken Davis

SB - Sue Brantley

DE - Dave Eissenstat

JK - Jason Kaye

LL - Li Li

TR - Tess Russo

MK - Margot Kaye

PB - Paul Bierman

HL - Henry Lin

* - External

SSHCZO Measurement Suite

Measurement	SH	GR	SCAL	SCBL	SCO	AG
Leaf Area Index	DE					
Leaf Litter	DE	DE				
Precipitation Amount	DA	DA				
Relative Humidity	KD	KD				
Sap Flux	DE	DE				
Soil CO2	JK	JK				
Soil Dielectric	HL	HL				
Soil Moisture - Stevens	HL	HL				
Soil Moisture - TDR	HL	DE				
Soil Moisture - COSMOS	*	LL				
Soil O2	JK	JK				

SH - Shale Hills

GR - Garner Run

SCAL - Shavers Creek - Above Lake

SCBL - Shavers Creek - Below Lake

SCO - Shavers Creek Outlet

AG - Agricultural Site

Blue – Active Data Being Collected

KD - Ken Davis

SB - Sue Brantley

DE - Dave Eissenstat

JK - Jason Kaye

LL - Li Li

TR - Tess Russo

MK - Margot Kaye

PB - Paul Bierman

HL - Henry Lin

* - External

SSHCZO Measurement Suite

Measurement	SH	GR	SCAL	SCBL	SCO	AG
Soil Temperature	HL	HL				
Stage/Level	TR	TR	TR	TR	TR	
Suspended sediment chemistry (major cations, Fe, Mn, Al, Si)	TR					
Tree Survey (Height, Diameter, Species Distribution)	DE	DE				
Turbidity	TR					
Virtual Temperature	KD	KD				
Water chemistry (DOC, TOC, DO, NO3, NH4, K, F, Cl, pH, EC, ORP, major cations, Fe, Mn, Al, Si)	TR	TR				

SH - Shale Hills

GR - Garner Run

SCAL - Shavers Creek - Above Lake

SCBL - Shavers Creek - Below Lake

SCO - Shavers Creek Outlet

AG - Agricultural Site

Blue – Active Data Being Collected

KD - Ken Davis

SB - Sue Brantley

DE - Dave Eissenstat

JK - Jason Kaye

LL - Li Li

TR - Tess Russo

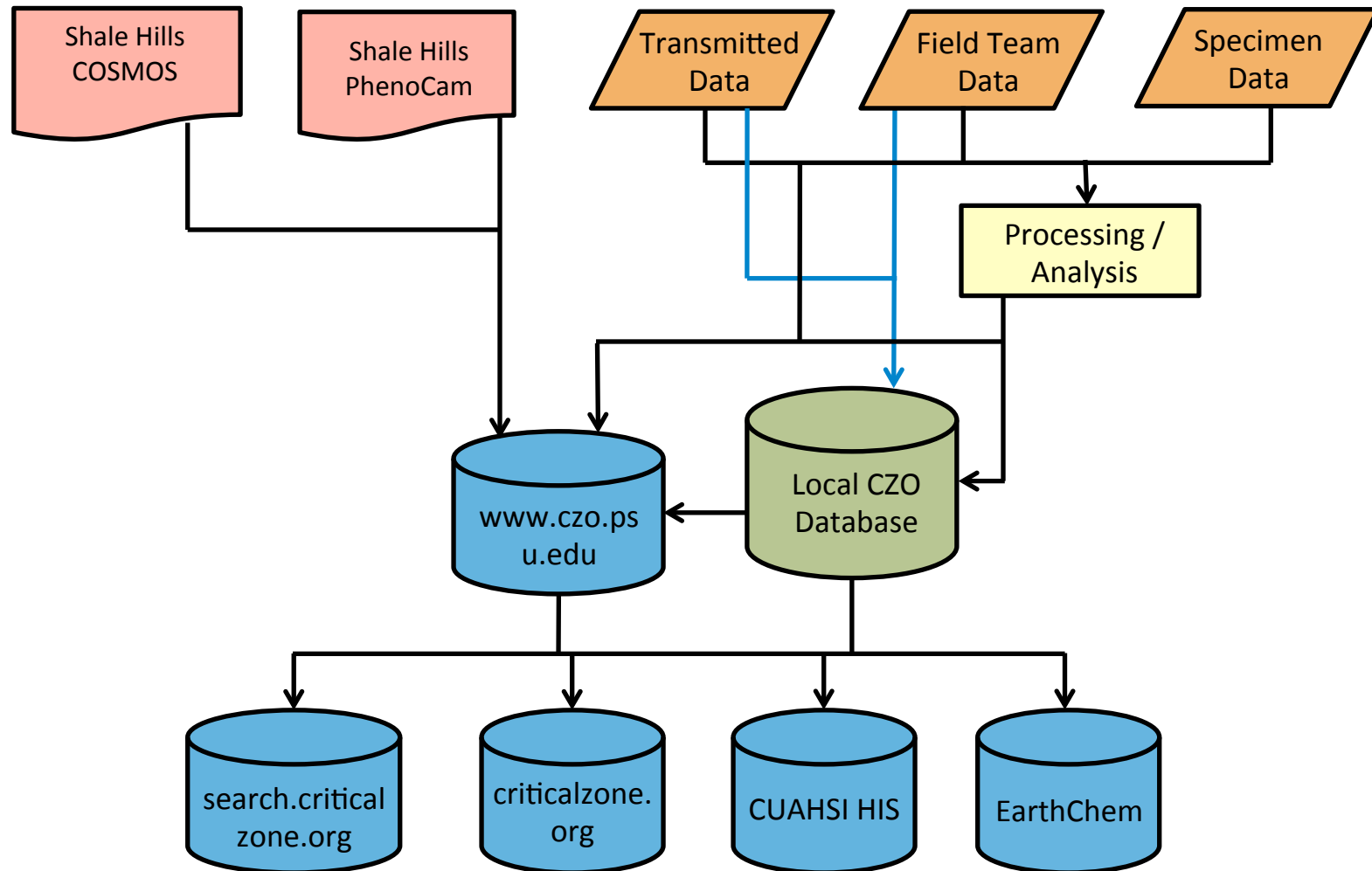
MK - Margot Kaye

PB - Paul Bierman

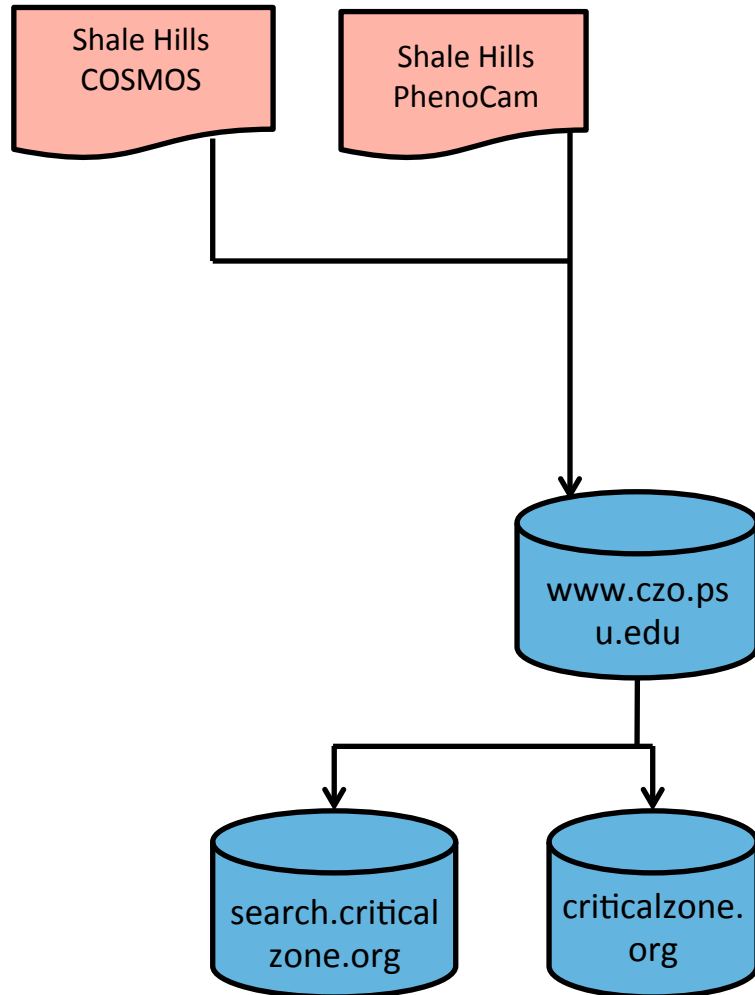
HL - Henry Lin

* - External

Data Flow Process Diagram



Externally-Managed Sources



SH COSMOS:

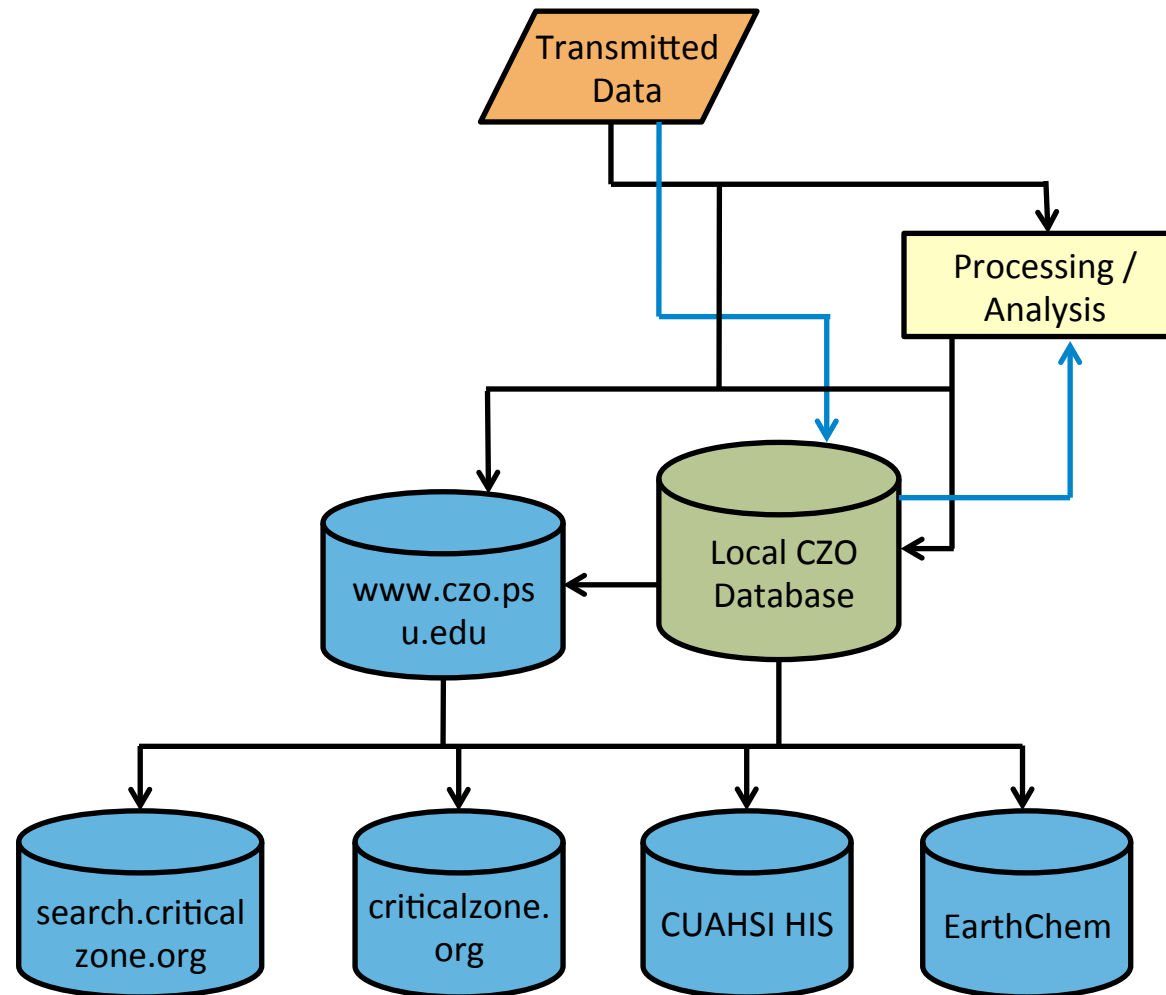
- Soil Moisture Data
- Managed by University of Arizona

SH PhenoCam:

- Photographic Canopy Data
- Managed by Harvard University

- Links for both on www.czo.psu.edu
- Metadata pages on criticalzone.org and search.criticalzone.org

Data Flow Process – Transmitted Data



Data Flow Process – Transmitted Data



Transmitted
Data

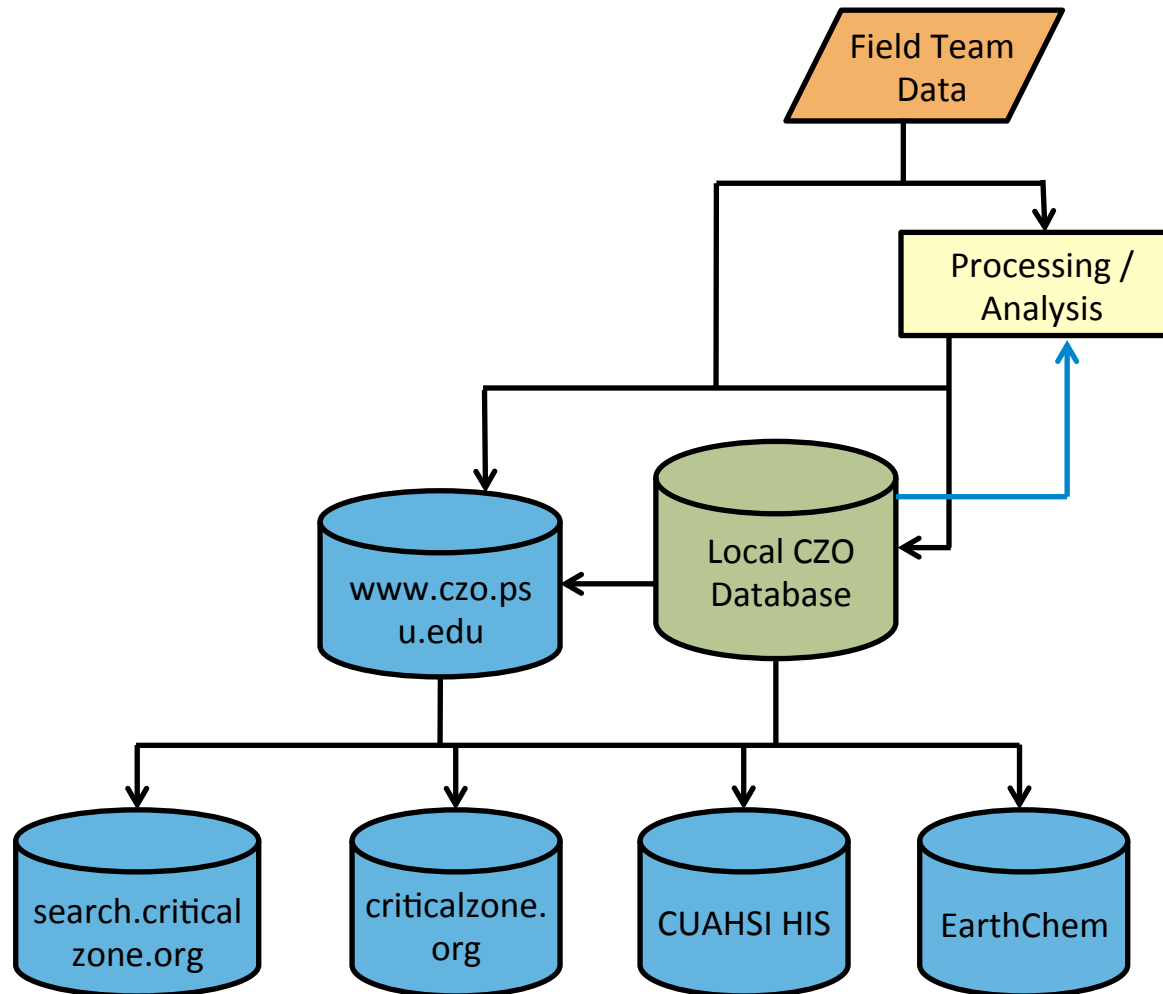
Internet Connection between Shale Hills and PSU campus

LoggerNet -> LNDB -> MS SQL Server 2008 Database (“CZO DB”)

- 10 Hz Raw Eddy Covariance Data
 - Extracted from CZO DB in weekly chunks, processed by grad student (raw & processed files exchanged via box.psu.edu)
- Flux Tower Meteorology
- Precipitation (OTT Pluvio² Weighing Rain Gauge)
- Raw Stream Discharge, Turbidity (weir replacement complete, awaiting recalibration & network infrastructure upgrade completion)

**Currently Shared Live as Raw Data: OTT Precip, 4-Way Solar Radiation. Soon: Turbidity, Discharge*

Data Flow Process – Field Team Data



Data Flow Process – Field Team Data



Field Team
Data

- Grad Student & Postdoc Field Team
- Collect water, gas, etc. samples as needed at Ground HOG, Tower HOG, etc.
- Download data from field dataloggers
- Upload data to internal box.psu.edu space
- Data manager looks over, notifies responsible PI/contact
- Processing completed as necessary, processed files returned to data manager
- Input into CZO DB or otherwise posted to web, external repositories

Data Flow Process – Field Team Data



Field Team
Data

Hydropedological Properties:

- Matric Potential
- Air Temperature
- Electrical Conductivity
- *Dielectric*
- Precipitation
- *Soil Moisture (Stevens Hydra Probes & TDR)*
- Soil Temperature
- Water Table

Sap Flow

Ground HOG:

- Soil Moisture
- Dielectric

Data Flow Process – Field Team Data

Field Team
Data

Hydropedological Properties:

- Matric Potential
- Air Temperature
- Electrical Conductivity
- Dielectric
- Precipitation
- Soil Moisture
- Soil Temperature
- Water Table

Other Data Not Specifically
Collected by Current Field
Team, but Processed Similarly:

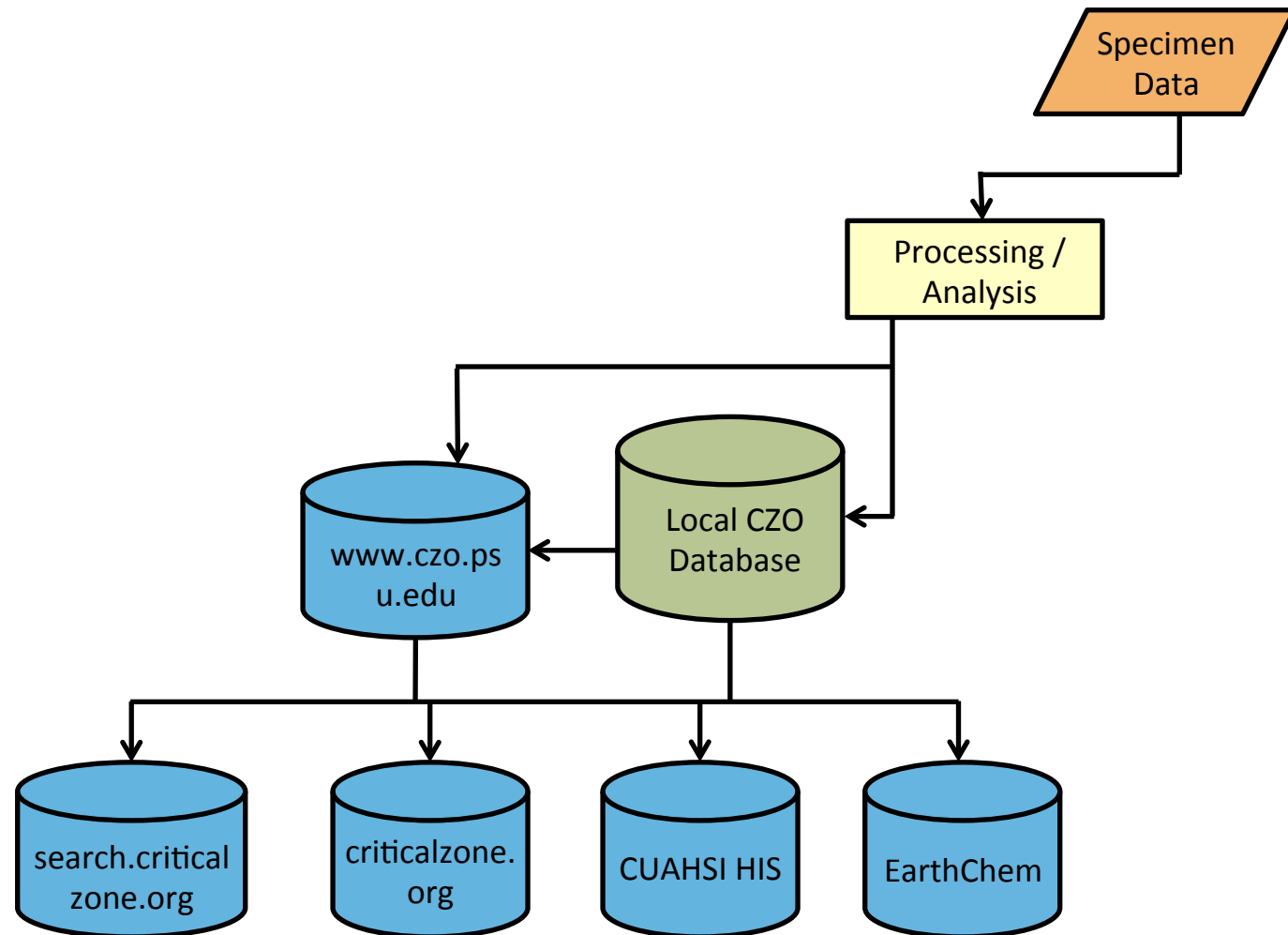
- Tree Survey Updates
- Leaf Area Index
- Litter/Dendroband

Sap Flow

Ground HOG:

- Soil Moisture
- Dielectric

Data Flow Process – Specimen Data



Data Flow Process – Specimen Data



Specimen
Data

Collected Specimens Cataloged (SESAR), Processed, Archived

When Released to Data Manager, Files Linked on Web
(www.czo.psu.edu & criticalzone.org) &/or Imported to CZO DB

DOIs obtained via EarthChem

- Geochemistry
- Aqueous Chemistry
- Soil Gas
- Vegetative Chemistry

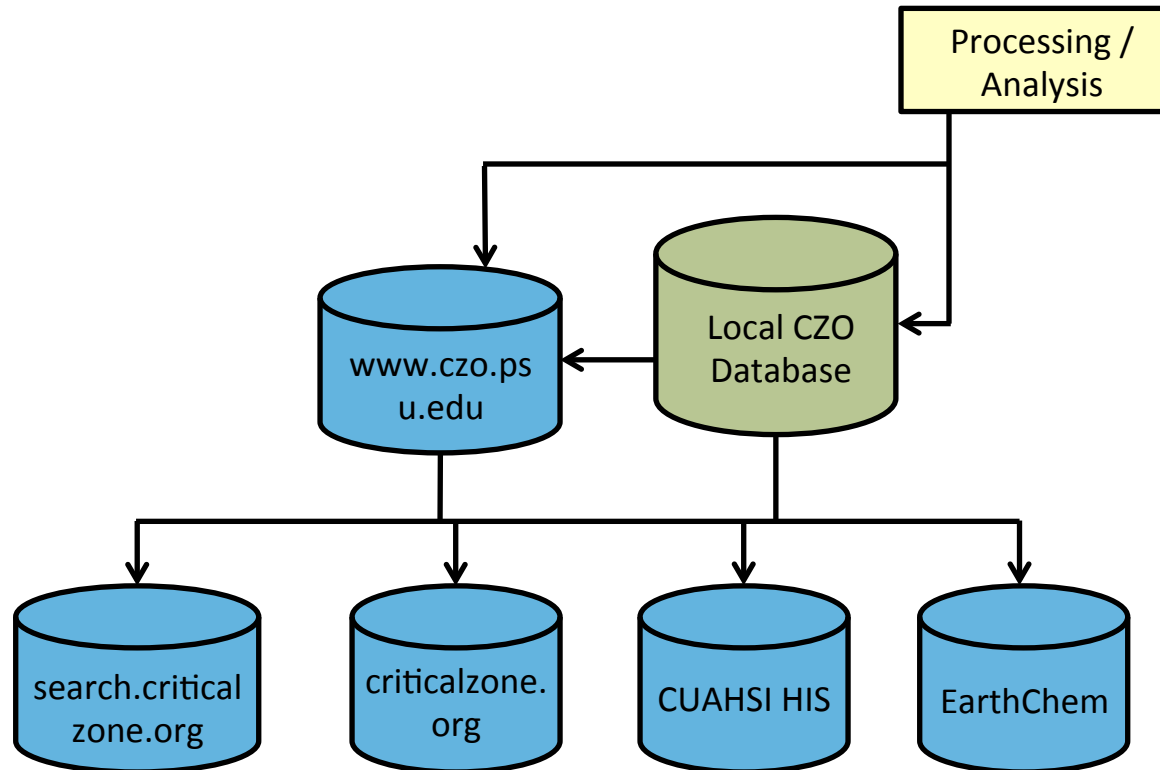
Other Data Types

- Geophysical Data
 - Gamma Counts
 - Seismic Data
 - Ground-Penetrating Radar
 - LIDAR (DEM, Density, Hillshade)
 - GIS/Geospatial Data
-
- Collected from Pls when Available
 - Usually Linked from Web, Not Suitable for CZO DB

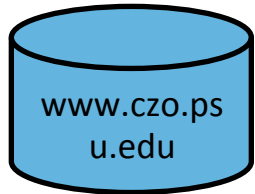
Other Data Acquisition

- Annual “Data Call” Conducted in March
 - Reminders for Data that are
 - Not Transmitted
 - Processed by PIs/grad students/postdocs
 - From Seed Grant PIs
 - E-mail sent, including blank metadata template document with instructions, with follow-up reminders

Data Flow Process – Processing & Access



Data Flow Process – Processing & Access



PSU CZO Web Site

HTML Pages Organized by 3 Broad Data Classifications:

- Time Series
 - Hydropedology, Land/Surface Flux, Discharge, etc.
- Geospatial
 - GIS Database, LIDAR, Tree Survey, etc.
- Geochemical/Geophysical
 - Aqueous Chemistry, Seismic Refraction, etc.

Links to:

- CSV/XLS Files
- Proprietary Files
- PHP Scripts linked to CZO DB

Data Flow Process – Processing & Access



criticalzone.org/shale-hills/data

Dataset Pages Generated by Expression Engine Content Management System (CMS)

Content Maintained by Data Manager

Metadata for each dataset

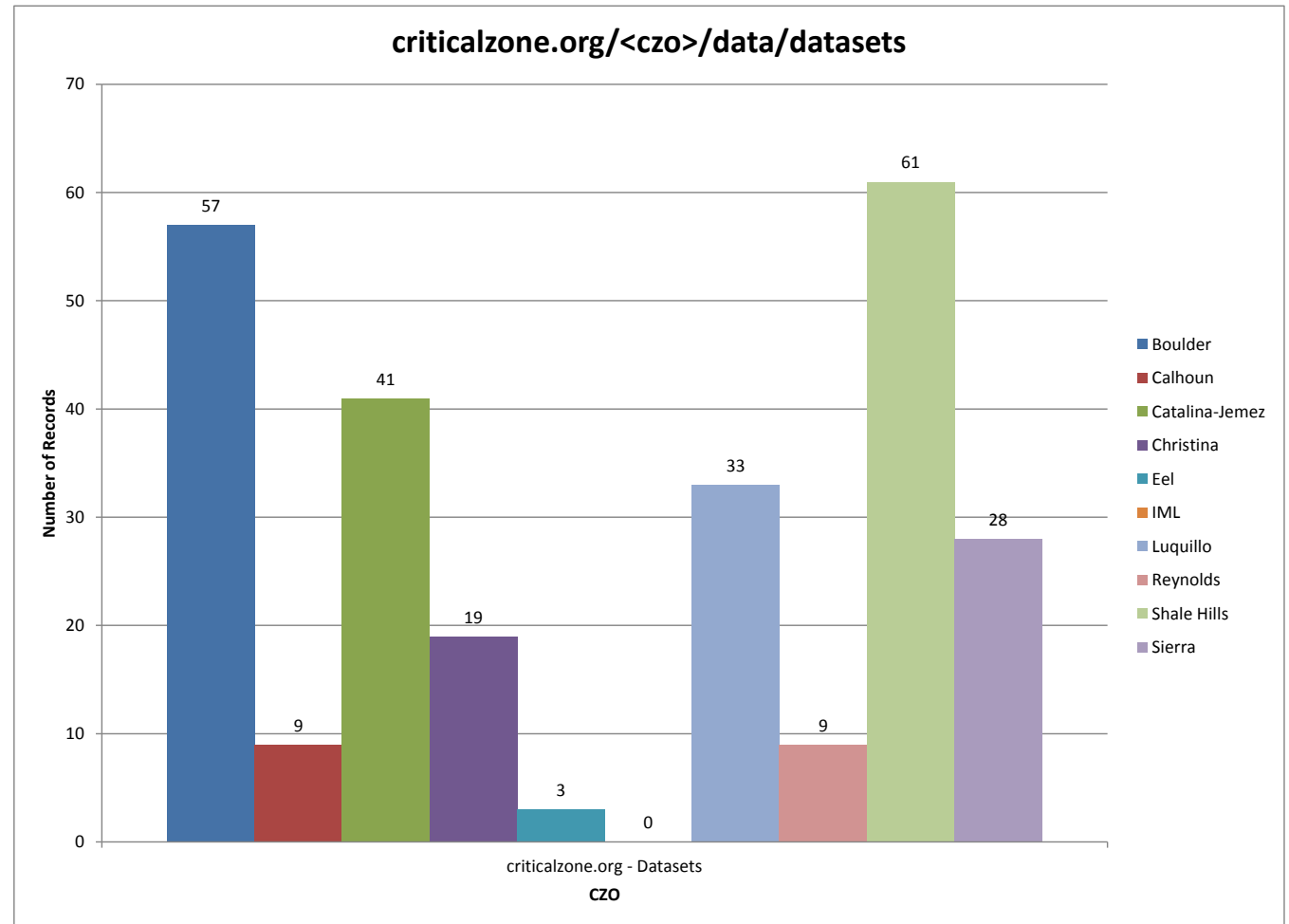
Link(s) to dataset components (e.g., yearly subsets)

- Direct Links to Files on www.czo.psu.edu

or

- Links to PHP scripts on www.czo.psu.edu

Data Flow Process – Processing & Access



61 Datasets with 269 Components (subclassifications; e.g., year, subtype)

*Correct as of 21 Oct 2015

Data Flow Process – Processing & Access

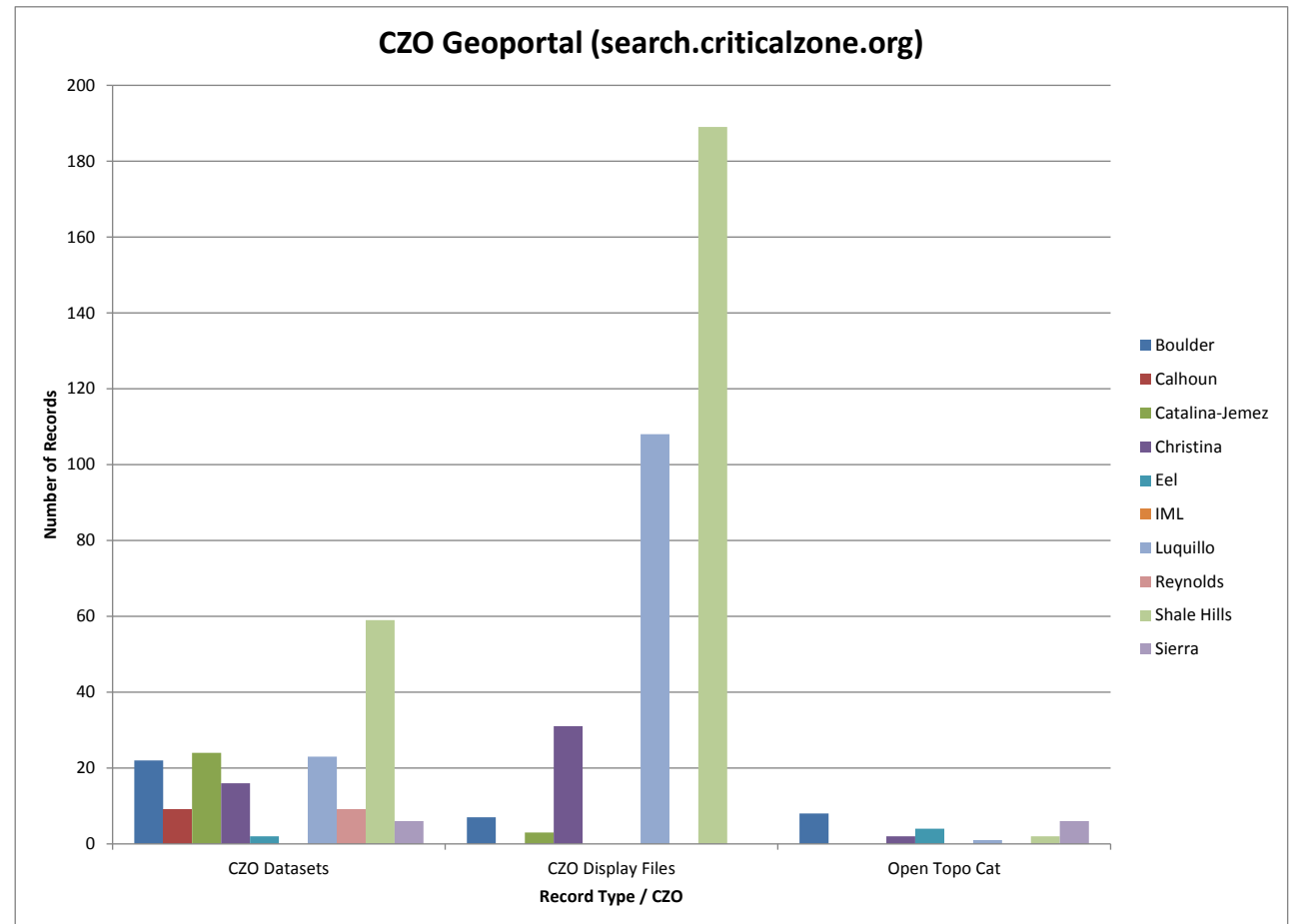


SDSC CZO Geoportal

Data Manager Creates
CZO Display Files v1 &
Initiates Harvest

Data: Primarily Time
Series, Some
Geochemical

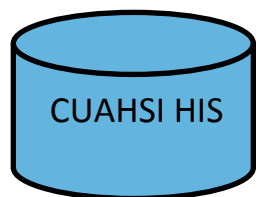
Converted from Source
Files or CZO DB Tables



SSHCZO: 59 Datasets, 189 Display Files, 2 Open Topography LIDAR Catalog Entries

*Correct as of 21 Oct 2015

Data Flow Process – Processing & Access

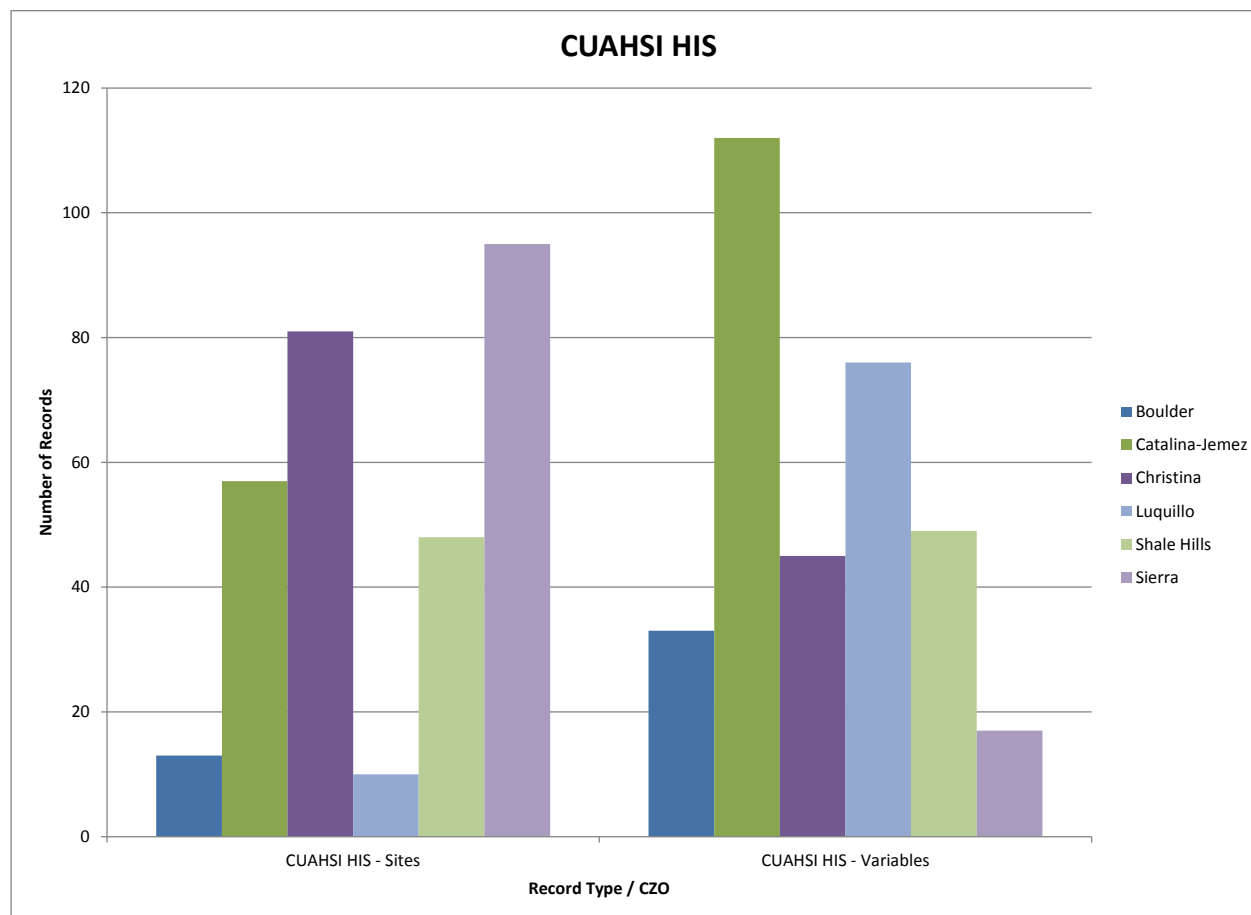


CUAHSI HIS

Receives metadata from
SDSC Geoportal

- Time Series
- Aqueous Chemistry

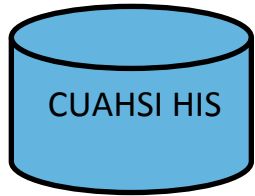
Clients: HydroDesktop,
Web Client



SSHCZO: 48 sites, 49 variables

*Correct as of 21 Oct 2015

Data Flow Process – Processing & Access

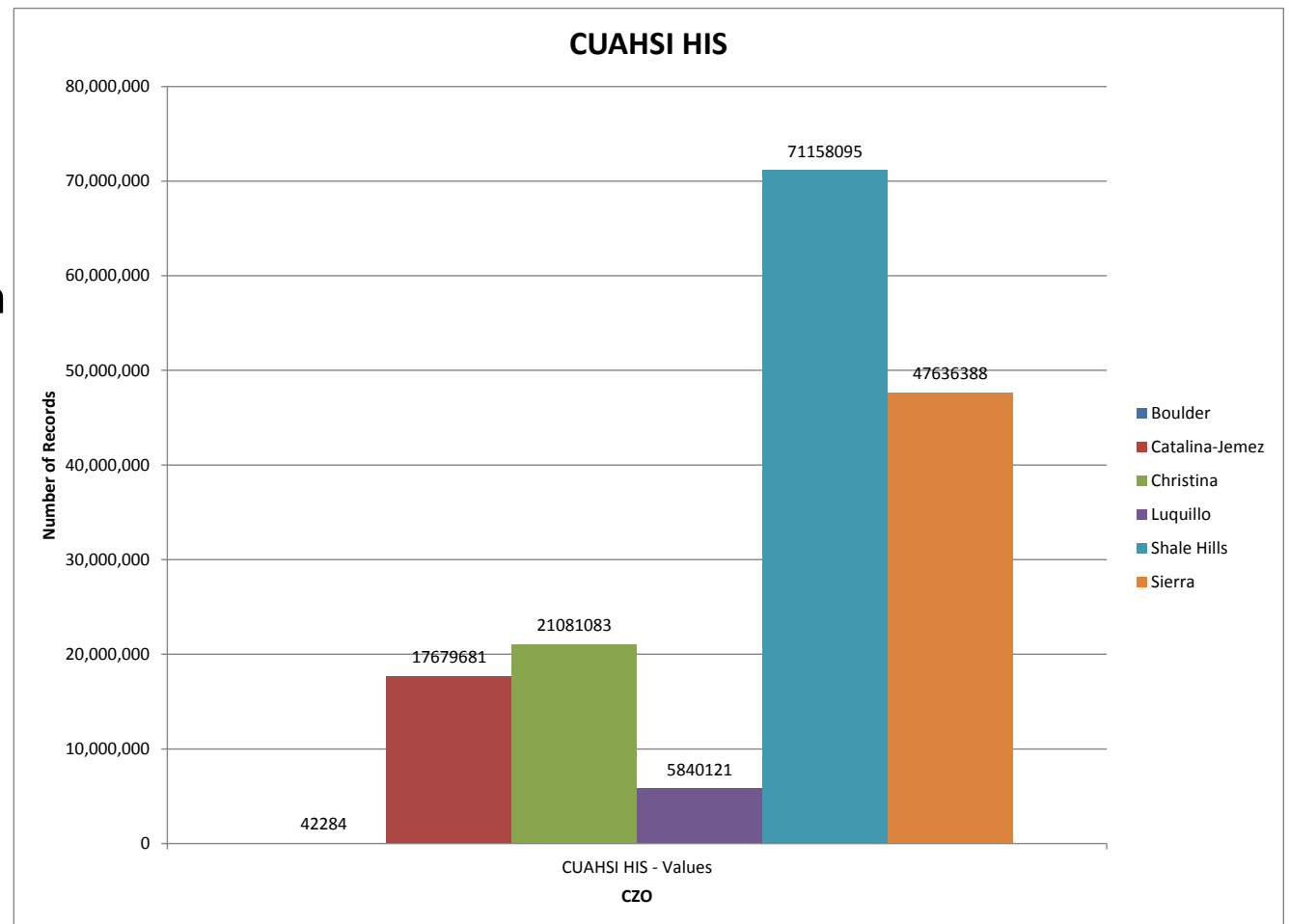


CUAHSI HIS

Receives metadata from
SDSC Geoportal

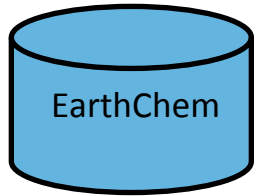
- Time Series
- Aqueous Chemistry

Clients: HydroDesktop,
Web Client



*Correct as of 21 Oct 2015

Data Flow Process – Processing & Access

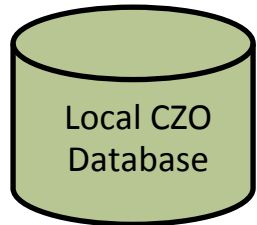


EarthChem

DOI Source – 30 for SSHCZO, including:

- Aqueous Chemistry
- Tree Survey

Data Flow Process – Processing & Access



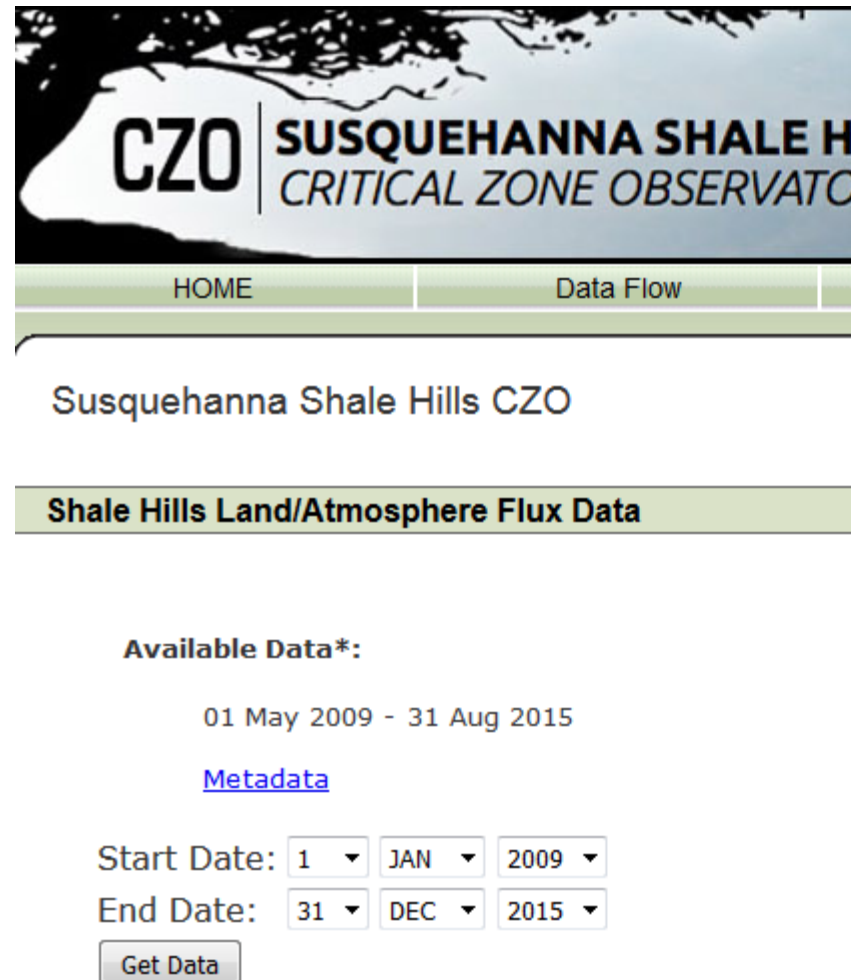
Local CZO DB

Microsoft SQL Server 2008

Administered by PSU ITS and
CZO Data Manager

Structure: Relatively Flat with
Few Table Relationships

- Facilitates easier
access, flexible tools



The screenshot shows the web interface for the CZO (Critical Zone Observatory) Susquehanna Shale Hills. The header features the CZO logo and the text 'SUSQUEHANNA SHALE HILLS CRITICAL ZONE OBSERVATORY'. Below the header is a navigation bar with 'HOME' and 'Data Flow' links. The main content area is titled 'Susquehanna Shale Hills CZO' and 'Shale Hills Land/Atmosphere Flux Data'. Under 'Available Data*', it shows the date range '01 May 2009 - 31 Aug 2015' and a link to 'Metadata'. At the bottom, there are date selection dropdowns for 'Start Date' (1, JAN, 2009) and 'End Date' (31, DEC, 2015), followed by a 'Get Data' button.

CZO | SUSQUEHANNA SHALE HILLS
CRITICAL ZONE OBSERVATORY

HOME Data Flow

Susquehanna Shale Hills CZO

Shale Hills Land/Atmosphere Flux Data

Available Data*:

01 May 2009 - 31 Aug 2015

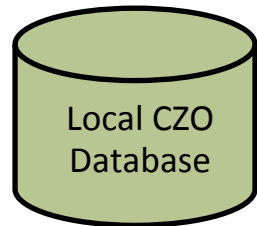
[Metadata](#)

Start Date: 1 JAN 2009

End Date: 31 DEC 2015

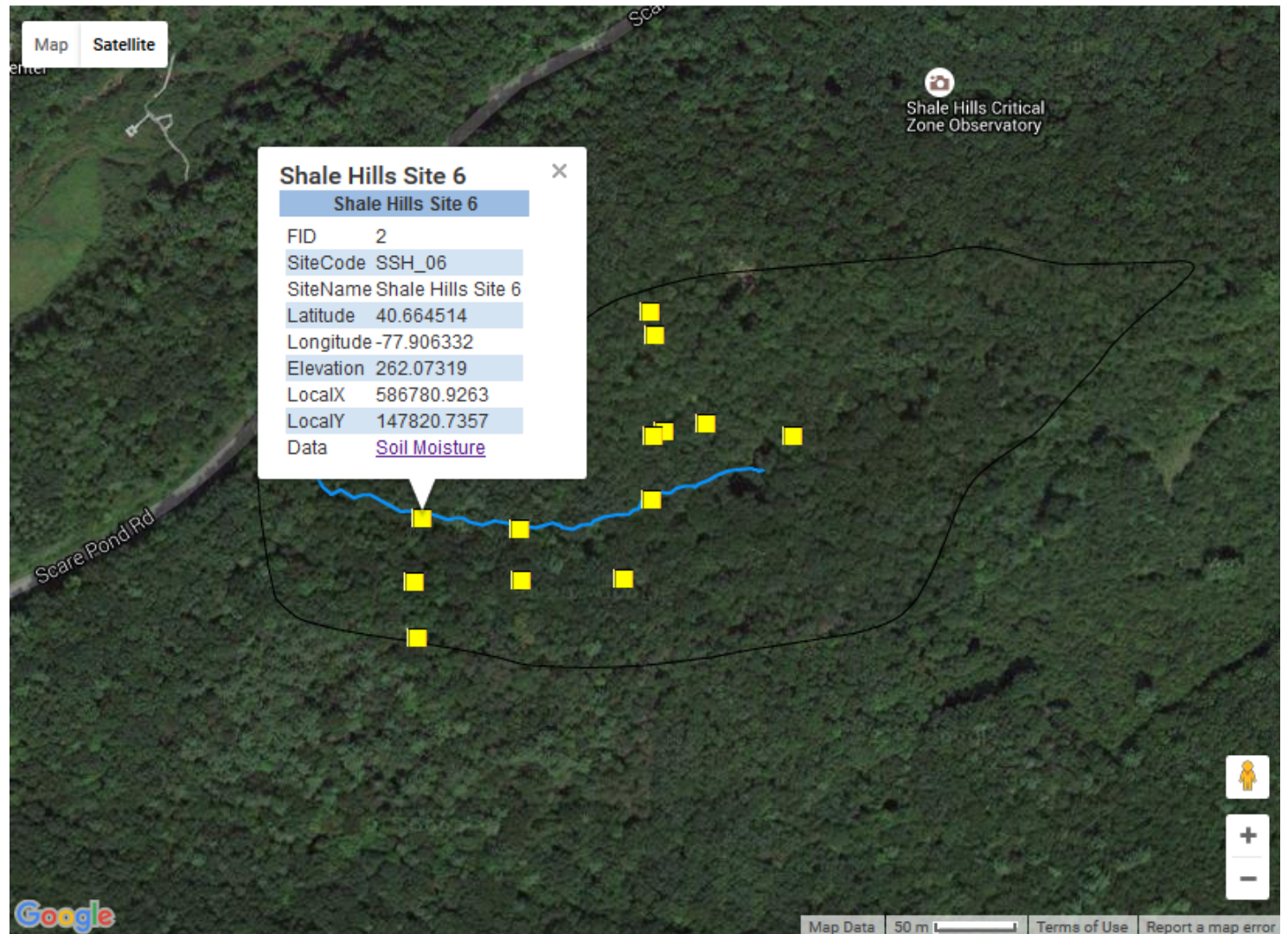
Get Data

Data Flow Process – Processing & Access



Map of Hydropedological Instrumentation Sites - Shale Hills - Soil Moisture

Click a site location for information and links to download data:



CZO Instrument Map

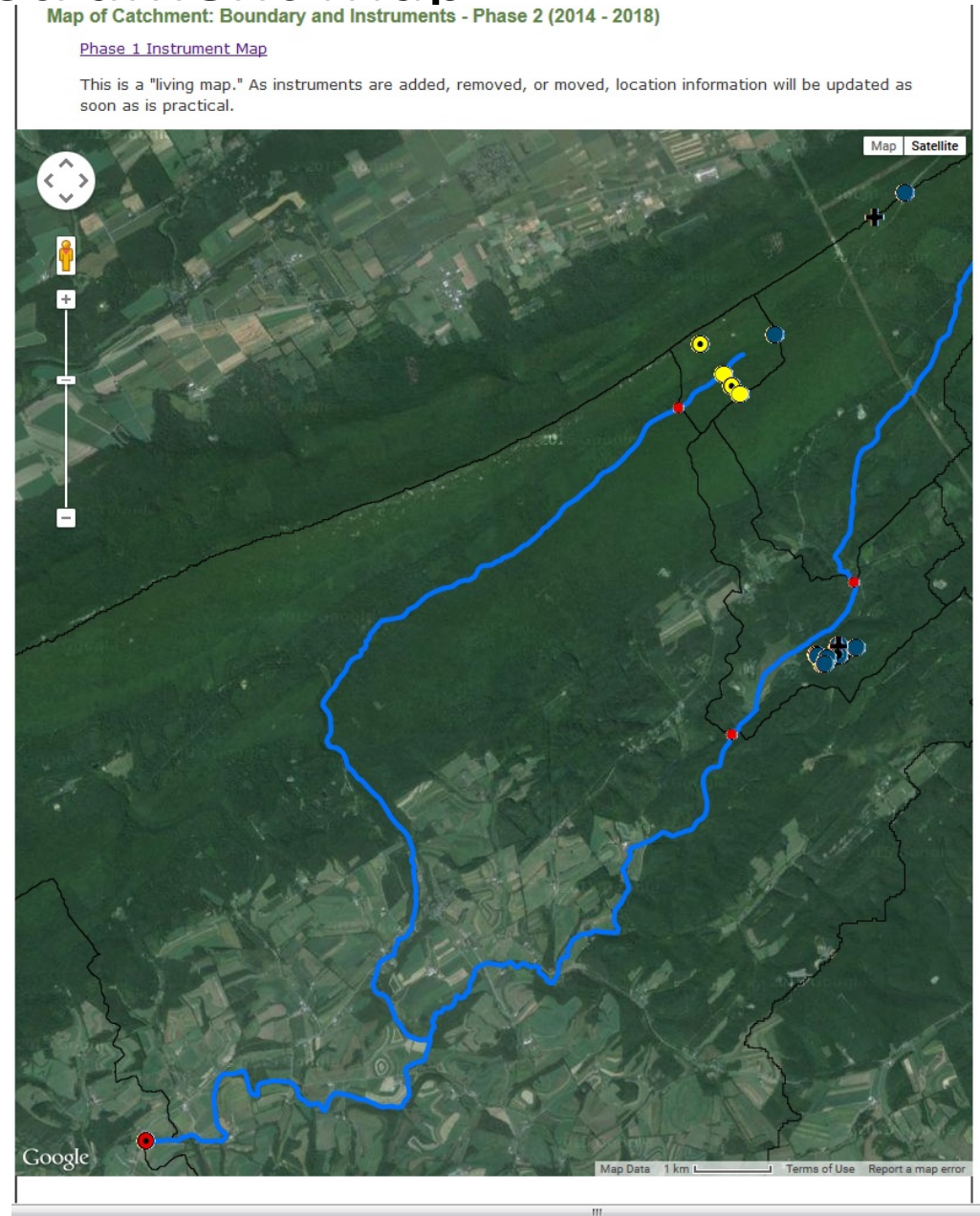
http://www.czo.psu.edu/data_overview.html

Clickable, customizable Google Map

- Instrument layers generated from ArcGIS geodatabase via KMZ
- Selectability & control via JavaScript

Legend:

	GroundHOG - Soil Moisture		GroundHOG - SM, Sap, Gas
	Stream Site		Stream Site S::CAN
	Well (Current/Planned)		Flux Tower
	Meteorology		
	Waterway		Subwatershed Boundary



Datasets – www.czo.psu.edu & criticalzone.org

†Measurements presumed finished or not ongoing

<u>Garner Run</u>	<u>Public</u>	<u>Embargo</u>
COSMOS Neutron Counts & Soil Moisture		2015
†Fiber-Optic Distributed Temperature Sensor		2014
Soil Geochemistry		2014
† Vegetation Transects		2014
 <u>Shale Hills</u>	 <u>Public</u>	 <u>Embargo</u>
†Watershed Reanalysis	1990 – 2010	
Stream Flow/Discharge Reanalysis	2008 – 2015	
†Stream Flow	2006 – 2012	
†Groundwater	2009 – 2012	
Precipitation (hourly and daily)	2006 – 2013	
Precipitation (daily and 10-minute, raw)	2009 – 2015	
†Real-Time Hydrology Network (RTH2 & 3) Soil Moisture	2009 – 2012	
Land-Atmosphere (“Surface”) Fluxes	2009 – 2015 (31 Aug)	
†12 Hydropedologic Properties	2007 – 2012	2013
TDR Soil Moisture	2006 – 2015	
†LPM Disdrometer	2009 – 2014	
Sap Flow	2009 – 2012	2013 – 2014

Datasets – www.czo.psu.edu & criticalzone.org

†Measurements presumed finished or not ongoing

<u>Shale Hills</u>	<u>Public</u>	<u>Embargo</u>
†Flux Tower Meteorology	2009 – 2013	
COSMOS (Shale Hills; links to external site)	2011 – 2015	
†Bedrock Elevation	2010	
†Terrestrial Laser Mapping	2010	
Tree Survey	2008	2010 – 2012
†Leaf Area Index	2010	
†LIDAR Vegetation Survey	2010	
Litter / Dendroband		2011 – 2013
Above-Ground Tree Biomass Data		2008 – 2012
†LAI Movie	2010	
PHENOCAM (links to external site)	2012 – 2015	
National GIS (DEM, Digital OrthoPhoto, Geology (USGS), Land Cover, Soils, Watershed Boundary)	unknown	
Soil Gas & Porewater Concentration		2013 – 2014
†Gamma Logs	2012	
†Wireline Logs	2006 – 2008	
†Solute Tracer Tests	2011	

Datasets – www.czo.psu.edu & criticalzone.org

†Measurements presumed finished or not ongoing

<u>Shale Hills</u>	<u>Public</u>	<u>Embargo</u>
†Seismic Refraction & Surface Wave Analyses		2013
Soil Gas Concentration, pCO ₂ & CO ₂ Flux	2008 – 2010	
Soil Geochemistry (link to CZChemDB)		2014
†Aqueous Geochemistry (ground, soil pore, and stream water)	2006 – 2011	
†Stable Isotope Hydrology (precip, stream, ground, soil water)	2008 – 2011	
†Tree Tracer Study	2012	
Root Length Density		2013
†Tree Isotope Chemistry		2011 – 2014
†Tree Throw / Tree Fall Data		2014
†Ground Penetrating Radar	2008	
Model Input Soil Parameters	2014	
<u>Shavers Creek Watershed</u>	<u>Public</u>	<u>Embargo</u>
Discharge and Water Chemistry		2014
Surface and Groundwater Chemistry		2014
†LIDAR (leaf-off and leaf-on)	2010	
†LIDAR-Derived (DEM, Density, and Hillshade)	2010	
†Vegetation Analysis	2010	

Datasets – www.czo.psu.edu & criticalzone.org

†Measurements presumed finished or not ongoing

<u>Shavers Creek Watershed</u>	<u>Public</u>	<u>Embargo</u>
Geodatabase (GIS)	2014	
National GIS (DEM, Geology (USGS), Land Cover, Soils, Watershed Bdy.)	unknown	
<u>Shale Transect</u>	<u>Public</u>	<u>Embargo</u>
†Meteorology	2010 – 2014	

Additional Activities

- In Development: Geotagged/Metadata-Linked Photo Database
- Processing Assistance
 - Hydropedological Properties
 - MATLAB scripts
 - Flux Tower Eddy Covariance
 - FORTRAN code for processing of 10 Hz tower data to EC, including QC pass
- Assuming Responsibility for All Precipitation Instrumentation and Measurement