HydroWatch
Open Source Informatics for Sensor Observatories

by
Collin Bode

P.I.’s: Inez Fung, William Dietrich, Todd Dawson, Jim Bishop, Mary Power, Ron Cohen, Kari Kaufman
Collaborators: Ginger Ogle, Rohit Salve, Kevin Simonin
Students: Hyojin Kim, Percy Link, Jasper Oshun, Daniella Rempe

Integrative Biology and Earth & Planetary Sciences,
UC Berkeley

Funding by W.M. Keck Foundation &
National Center for Earth-surface Dynamics
Outline

1. Sensor Observatory Concept
2. Hydrowatch project
3. Study Site
4. Networking
5. Sensors
6. Informatics (CUAHSI ODM)
Sensor Observatory

“Mapping, Tracing, and Wireless Sensor technologies may do for ecology in the 21st century what DNA sequencing did for genetics in the 20th century.”

– Beth Burnside, molecular biologist & retired Vice Chancellor for Research at UCB

**Definition:**
a distributed array of sensors connected by a network that is aggregated into a single dataset.

- Spatially distributed
- High temporal resolution
- Real-time access

Sensor Observatory

Field Station Sensor Data Workflow

Physical and Networking Components

Sensors

Physical and Networking Components

Informatics

UC Berkeley Campus Colocation Facility

Web Server

Field Stations Data Server

Filter for data out of range, flag and warn responsible parties

Convert raw data into proper units

Parse Logger files and load into database

Raync to Field Station Data Server, rename files as appropriate

Connect to logger over serial or ethernet and pull data

Data Logger

Data Logger ASCII file output

Local Caching Server

Field Stations Data Server

Web Pages

Wrapper for web services

Educators

Researchers

General Public
HydroWatch Project

**Vision:** to exploit new technology to have highly resolved mapping of water as it transits from canopy to ground to stream and returns to the atmosphere in a remote setting.

**PI:** Inez Fung, Atmospheric Scientist, Climate modeler

**Co-PI’s:**
- Bill Dietrich, Geomorphology
- Todd Dawson, Tree physiology
- Mary Power, Ecology
- Jim Bishop, Aquatic chemistry
- Ron Cohen, Atmospheric chemistry
- Kari Kaufman, Statistics
HydroWatch Project: Current Research Questions

- What controls the depth of weathering front into bedrock under ridge and valley topography? **Daniella Rempe**
- In seasonally dry environments, do vegetation communities and their response to climate change depend on rock moisture? **Jasper Oshun**
- How does the spatial and temporal structure of transpiration across hillslopes depend on environmental conditions? **Percy Link**
- What controls the chemical evolution of runoff through the critical zone and what does that reveal about runoff processes? **Hyojin Kim**
HydroWatch Project: Data Collection

1. Sensor Observatory
   2. Geospatial Data
      - 2 LiDAR flights 2004, 2009
      - Bare-earth & Vegetative Canopy DEMs
3. Automated water collection: lab chemical & isotope analysis
   - Four 24-position Isco water samplers: 3 wells, 1 stream
   - Networked for remote sampling adjustment
   - Modified bottle collectors to isolate water
4. Regular Manual Sampling and Monitoring
   - Tracing isotopes through the system: vegetation, soil, lysimeter water, well, stream
   - Well level hand measurements
   - Collection of Isco water bottle samples
   - Rain water collection
5. Targeted Field Campaigns
   - Picarro CRDS analyzer in situ monitoring of $d^{18}O$ and $dD$ in water & vapor
   - Conservative tracer experiments during dry season
6. Geophysics: mapping fractured rock moisture
   - Electrical Resistivity Tomography (ERT)
   - Seismic refraction
   - Neutron logging in wells
7. Sensor Development
   - Electrical resistance Probes (ERP) for soil moisture (Salve, 2011)
   - CO2 motes in wells
Study Site:

Angelo Coast Range Reserve

http://angelo.berkeley.edu
Study Site:

Angelo Coast Range Reserve

http://angelo.berkeley.edu
Study Site: Angelo Reserve

- Reserve 30 km²
- South Fork Eel River 143 km²
- Elder Creek 17 km²
Study Site: Rivendell
Study Site: Rivendell
Study Site: Rivendell, planar view

Catchment area: 3,500 m$^2$
Sensor Observatory: Networking

Field Station Sensor Data Workflow

- Physical and Networking Components
- Fixed Wireless
- Data Logger
- Local Caching Server
- IST Point of Presence
- Internet VPN tunnel
- UC Berkeley Campus Colocation Facility
- Field Station Data Server
- Web Server
- Educators
- Researchers
- General Public

Workflow Functions

- Connect to logger over serial or ethernet and pull data
- Parse Logger files and load into database
- Convert raw data into proper units
- Filter for data out of range, flag and warn responsible parties
- Wrapper for web services
- Web Pages
## Networking: Radio Frequencies

### Low frequencies
- better transmission
- bad interference
- lower data transfer
- low power

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>530kHz – 1.7MHz</td>
<td>AM Radio</td>
</tr>
<tr>
<td>88 – 108MHz</td>
<td>FM Radio</td>
</tr>
<tr>
<td>3 - 30MHz</td>
<td>VHF – ship radios</td>
</tr>
<tr>
<td>170 – 698MHz</td>
<td>TV broadcast</td>
</tr>
<tr>
<td>30 – 300MHz</td>
<td>UHF – animal tracking</td>
</tr>
<tr>
<td>700MHz</td>
<td>4G – AT&amp;T, Verizon</td>
</tr>
<tr>
<td>800MHz</td>
<td>Cell phones</td>
</tr>
<tr>
<td>900MHz</td>
<td>Unlicensed</td>
</tr>
<tr>
<td>1.5GHz</td>
<td>GPS</td>
</tr>
<tr>
<td>1.7GHz</td>
<td>4G – Sprint, T-Mobile, USA, MetroPCS</td>
</tr>
<tr>
<td>2.4GHz</td>
<td>WiFi, Unlicensed</td>
</tr>
<tr>
<td>4.8GHz</td>
<td>Point-to-point, Unlicensed</td>
</tr>
<tr>
<td>12.0GHz</td>
<td>Satellite (Rappaport, 2002)</td>
</tr>
</tbody>
</table>

### High frequencies
- poor transmission
- less interference
- higher data transfer
- high power

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>700MHz</td>
<td>4G – AT&amp;T, Verizon</td>
</tr>
<tr>
<td>800MHz</td>
<td>Cell phones</td>
</tr>
<tr>
<td>900MHz</td>
<td>Unlicensed</td>
</tr>
<tr>
<td>1.5GHz</td>
<td>GPS</td>
</tr>
<tr>
<td>1.7GHz</td>
<td>4G – Sprint, T-Mobile, USA, MetroPCS</td>
</tr>
<tr>
<td>2.4GHz</td>
<td>WiFi, Unlicensed</td>
</tr>
<tr>
<td>4.8GHz</td>
<td>Point-to-point, Unlicensed</td>
</tr>
<tr>
<td>12.0GHz</td>
<td>Satellite (Rappaport, 2002)</td>
</tr>
</tbody>
</table>
Networking: Radios

Backhaul: High Power P2P Radios (5-15 watts)
  - Ubiquiti Rocket
  - Motorola Canopy
  - Eion VIP110-24

Mesh Radios
  - Tranzeo Mesh
  - Memsic Eko Motes

Medium Power Radios (1-5 watts)
  - Campbell Scientific RF450
  - Custom Wifi radios

Low Power Radios (milliwatts)
  - Cellular modems
  - Telemetry radios
Networking: Rivendell

4 Wireless networks

**Network Architectures**

1. **Backbone**
   1. Heirarchical tree structure
   2. Eion VIP110-24
   3. 2.4ghz, 5 watts
   4. VINES protocol

2. **Campbell PakBus**
   - Manually defined relays
   - 900mhz, 1watt
   - Pakbus protocol

3. **Motes**
   - Mesh structure
   - 2.4ghz, extremely low power
   - Single base station to Backbone
Networking: Trees

Trees in Rivendell

- Treebeard (65m)
- Ingrid (44m)
- Ursula (44m)
- Ilean (55m)
- Flatcap (50m)
- Flatcap (50m)
- Ursula (44m)

Legend:
- Red line: Boundary
- Black dots: Trees
- Blue dots: Locations

Elder Creek

Brandywine (30m)
Networking: Power
Networks: Wireless Backbone

Collin Bode, configuring at 35 meters

Peter Steel, Reserve Steward, Setting up radio and antenna
Networks: REB Construction

- Redwood height: 40 meters
- Solar System total weight: 80 kilos
- Arms: very sore
Sensor Observatory: Sensors

Field Station Sensor Data Workflow

Physical and Networking Components:
- Fixed Wireless
- Sensor
- Data Logger

Sensors

Workflow Functions:
1. Connect to logger over serial or ethernet and pull data
2. Ramyc to Field Station Data Server, rename files as appropriate
3. Parse Logger files and load into database
4. Convert raw data into proper units
5. Filter for data out of range, flag and warn responsible parties

Data Logger ASCII file output

Local Caching Server

Field Stations Data Server

Web Server

UC Berkeley Campus Colocation Facility

Field Station Clearinghouse Database

Educators

Researchers

General Public

Web Pages
Sensors: Rivendell
<table>
<thead>
<tr>
<th>Value Measured</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barometric pressure</td>
<td>7</td>
</tr>
<tr>
<td>CO2 Sensor</td>
<td>2</td>
</tr>
<tr>
<td>Hail, intensity</td>
<td>2</td>
</tr>
<tr>
<td>Light, net radiometer</td>
<td>1</td>
</tr>
<tr>
<td>Light, shortwave</td>
<td>12</td>
</tr>
<tr>
<td>Light, shortwave, offline</td>
<td>30</td>
</tr>
<tr>
<td>Logger, battery voltage</td>
<td>15</td>
</tr>
<tr>
<td>Logger, temperature</td>
<td>15</td>
</tr>
<tr>
<td>Precipitation, tipping bucket</td>
<td>9</td>
</tr>
<tr>
<td>Relative humidity</td>
<td>15</td>
</tr>
<tr>
<td>Sap flow</td>
<td>106</td>
</tr>
<tr>
<td>Snow pack height, sonic</td>
<td>1</td>
</tr>
<tr>
<td>Soil Hygrometer/Psychrometer</td>
<td>54</td>
</tr>
<tr>
<td>Soil moisture, Electrical resistance</td>
<td>192</td>
</tr>
<tr>
<td>Soil moisture, sm200</td>
<td>30</td>
</tr>
<tr>
<td>Soil moisture, time domain reflectometer</td>
<td>100</td>
</tr>
<tr>
<td>Temperature, air</td>
<td>14</td>
</tr>
<tr>
<td>Temperature, air, offline</td>
<td>30</td>
</tr>
<tr>
<td>Temperature, water</td>
<td>3</td>
</tr>
<tr>
<td>Water content reflectometers</td>
<td>15</td>
</tr>
<tr>
<td>Water samples, isco</td>
<td>4</td>
</tr>
<tr>
<td>Water turbidity</td>
<td>1</td>
</tr>
<tr>
<td>Well and stream water level</td>
<td>23</td>
</tr>
<tr>
<td>Well water level, manual</td>
<td>12</td>
</tr>
<tr>
<td>Wind, direction</td>
<td>6</td>
</tr>
<tr>
<td>Wind, speed</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>706</strong></td>
</tr>
</tbody>
</table>
Sensors: Rivendell
Sensors: Rivendell
Sensors: Wells
Sensor Observatory: Informatics

Sensor Database Design Goals
- Standardization
- Automation
- System Maintenance
- Quality Control
Informatics: Hydrologic Information Systems (HIS)
HIS: HydroServer

Point Observations Data

Ongoing Data Collection

Historical Data Files → ODM Database

Internet Applications

GIS Data

GetSites
GetSiteInfo
GetVariableInfo
GetValues

WaterML
WaterOneFlow Web Service

ArcGIS Server

Data presentation, visualization, and analysis through Internet enabled applications

HydroServer
HIS: Observations Data Model (ODM)

Microsoft SQL Server Database Schema
Informatics: Berkeley Sensor Database

- Ginger Ogle, Database Programmer
- LAMP (Linux, Apache, MySQL, Perl)
- SQLserver → MySQL
- Core ODM schema only

159.7 million records as of April 18, 2013
Online since Oct, 2008
Informatics:

http://sensor.berkeley.edu

Sensor Database
- Number of users: 30
- Data storage, retrieval
- Data quick visualization
- Maintenance
**Monitoring Collection:** A conceptual grouping of stations, usually being used by a specific research project or organization. *Access Control*

**Station:** One datalogger with sensor devices attached, such as a weather station. The Station is also used to define the output data file which is pulled from the logger.

**Method:** A way to measure a parameter. Usually a particular instrument make and model. *Already exists in ODM.*

**Device:** A specific instance of a method that is installed. A device can be decommissioned, or moved to another location.

**Datastream:** The "data flow" of one type of data from a device of a particular method at a particular location.
Informatics: Datasreams

Datasream: The "data flow" of one type of data from a device of a particular method at a particular location.
Informatics: Quality Control

Project Requirements:
- Store all versions of data (raw, converted, derived, corrected)
- Sanity check on incoming data
- Allow users to flag data
- Show flags & annotations on all data

### NASA/EOS Data Quality Levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Definition</th>
<th>Data Flag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Instrument Data</td>
<td>P</td>
<td>Passed sanity check</td>
</tr>
<tr>
<td>1A</td>
<td>Instrument Data + Metadata</td>
<td>U</td>
<td>Unchecked value, no sanity check</td>
</tr>
<tr>
<td>1B</td>
<td>Instrument Data in Sensor Units + Metadata</td>
<td>NV</td>
<td>No value recorded (NaN)</td>
</tr>
<tr>
<td>2</td>
<td>Geophysical Units</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Geophysical Units + Space/Time Uniformly Scaled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Derived Data, from Multiple Variables or from a Model</td>
<td>VE</td>
<td>Value exceeds device minimum range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>Invalid Data: equipment malfunction</td>
</tr>
</tbody>
</table>

### Data Flags

- **P**: Passed sanity check
- **U**: Unchecked value, no sanity check
- **NV**: No value recorded (NaN)
- **VB**: Value below device minimum range
- **VE**: Value exceeds device minimum range
- **X**: Invalid Data: equipment malfunction
**Informatics: Researcher Use**

http://sensor.berkeley.edu

<table>
<thead>
<tr>
<th>Type of Access</th>
<th>Requirements</th>
<th>Data Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web Access</td>
<td>Web browser &amp; valid project login</td>
<td>Graph &amp; CSV download</td>
</tr>
<tr>
<td>Direct MySQL Access</td>
<td>Knowledge of SQL</td>
<td>SQL output</td>
</tr>
<tr>
<td>Matlab</td>
<td>Matlab MySQL connector or Mathworks Database Toolbox</td>
<td>Matlab matrix format</td>
</tr>
<tr>
<td>Bulk Data Dumps</td>
<td>Valid project login</td>
<td>Zipped, CSV, weekly refresh</td>
</tr>
</tbody>
</table>

---

**Query the Sensor Database**

Use this form to query data in the sensor database. If there is more than one level of data (year, month, day), create a table for each data point or only the most recent version of the data will be included in the results. Please email [sensor@berkeley.edu](mailto:sensor@berkeley.edu) if you need access to earlier versions of the data.

See [Sensors Database](http://sensor.berkeley.edu) for current info about data sources, collection times and ranges.

Data Range is required, and you must select at least one Station and one Variable.

---

**Query Sensor Data**

[Download](http://sensor.berkeley.edu) the results. Data-file name: .txt, file size: 178 KB.

```
SELECT * FROM Data WHERE Station = 'value' AND Date = 'value';
```

http://sensor.berkeley.edu

---

**Note:** If your query included other stations or variables that are not shown below, they either no data was recorded for that station during the time period requested, or the variable you chose is not measured at that station.
Incident ID: 1
Title: Bear damage
Station Name(s): L1_1, L1_2, L2_1, L2_2, L3_1, L3_2, L4_1
Datastream(s): resistance probes, rain collector, wells
Data Flagged: ERPs: 290,304 values
Well 6: 3,024 values
Well 7: 3,024 values
StartTime: 2009-09-24 (+/- 1 months)
EndTime: 2009-10-24 (+/- 1 days)
Reported By: Bill Dietrich (2009-10-26 00:00:00)

Description:
Daniella, Jasper and I visited the site this past weekend. Unfortunately a bear did some damage recently. It systematically knocked over and chewed on everyone of the plastic bag covered resistance probes. They were each snapped at the soil boundary. I haven't had a chance to check to see when this might have happened according to the online record. The bear also knocked over the rainfall collector (James says he has found teeth marks in it before). It broke the extended plastic tube on well 7 and the wooden stake holding up the plastic tarp. and it knocked off the cap of well 6. We didn't detect other damage-- but there could be some. I encourage you all to check your favorite on line device. We put the cap back on 6, propped up well 7 tube and put back a stake. and reset up the rain collector.
Lessons Learned:

**Data Principle:** Each set of sensors requires an individual who is responsible for their functioning.
- Receive email alerts when data is out of bounds
- Check sensor database graphs for unusual behavior
- Write incident reports when something goes wrong
- Get credit on publications using their data

**Wireless Motes:** Wonderful idea. Excellent potential. Difficult to realize.

**Power:** real-time systems require a lot of power. Planning for sufficient power for your climate and deployment is critical to its success.

**Maintenance:** the infrastructure takes a lot of maintenance. Large deployments need an informaticist.