Earth Science Information Partners (ESIP) EnviroSensing Cluster
Don Henshaw¹, Corinna Gries², and Fox Peterson³

¹ U.S. Forest Service Pacific Northwest Research Station, ² Center for Limnology, University of Wisconsin, ³ College of Forestry, Oregon State University

EnviroSensing Cluster

Primary objective
Provide sensor network resources for environmental sensor practitioners through a wiki page and regular monthly teleconferences

Sensor and Sensor Data Management Best Practices


Sensor, site, and platform selection

- Selection of sites, science platforms and support systems are interacting planning process
- Communication among PI’s, techs, and information managers
- Data quality and longevity is ultimate goal
- Robust and widely-used core systems and sensors
- Standardize sensor and support hardware, software, designs
- Optimal siting for science objectives can be impeded
- Land ownership/permitting, seasonal weather patterns, logistical access, availability of services (e.g., power sources, communications), operating budget

Data acquisition and transmission

- Manual downloads of sensor data
- May not be sufficient to assure data security
- Does not allow direct control of devices
- Remote data acquisition considerations:
  - Collection frequency and need for immediate access
  - Uni- versus bi-directional transmission methods
  - Bandwidth requirements to transfer the data
  - Line-of-site communication or repeaters
  - Hardware and network protocols
  - Power consumption of the system components
  - Physical and network security requirements
  - Reliability and redundancy
  - Expertise
  - Budget

Sensor data quality assurance and quality control (QA/QC)

- Quality assurance – preventative measures
  - Routine calibration and maintenance
  - Anticipate common repairs and replacement parts
  - Design
  - Access proper installation and protection
  - Sensor redundancy
  - Regular human inspection and evaluation of sensor network
  - Automated alerts in site webcams
- Quality control – checks in near real-time
  - Timestamp integrity (Date/time)
  - Range checks
  - Internal (plausibility) checks
  - Variance checks / Outlier detection
  - Persistence checks
  - Spatial checks / Correlations with nearby sensors

Sensor data archiving

- Archiving strategies
  - Create well documented data snapshots
  - Assign unique, persistent identifiers
  - Maintain data and metadata versioning
  - Store data in text-based formats
- Partner with cross-institution supported archives
  - Federated archive initiatives such as DataONE
  - Community supported, e.g., the LTER NIS
- Best practices
  - Develop an archival data management plan
  - Implement a sound data backup plan
  - Archive raw data (but they do not need to be online)
  - Make data publicly available
  - Assign QC level to published data sets

Primary activities

- Building a sensor and sensor data management best practices guide through community participation
- Monthly teleconferences
- Ongoing maintenance of the ESIP EnviroSensing wiki page and sensor network resource links

Monthly teleconferences

- Monthly discussion forum open to the broader community
- Enlists presentations from sensor research projects and sensor manufacturers and software developers

Research program presentations 2015:

- Monterey Bay Aquarium Research Institute
  - Research product: Smart Open Sensors Consortium
  - http://www.mbari.org/
- Desert Research Institute
  - Research product: AquPerch Data Portal
  - http://www.dri.edu/
- Heat Seek NYC
  - Research product: Heat Seek Temperature Nodes
  - http://heatseeknc.com/

Sensor manufacturer / software developer presentations 2015:

- Aquatic Informatics
  - Software product: Aquarius
  - http://aquaticinformatics.com
- Onset
  - Company products: HOBO Data loggers and HOBOware
  - http://www.onsetcomp.com
- Kisters
  - Software product: WISKI system
  - http://www.kisters.net
- LI-COR
  - Company products: LI-COR Instruments and Eddypro software
  - http://www.licor.com

Sensor data archiving

Table:

<table>
<thead>
<tr>
<th>LTER Data Co-op</th>
<th>Data Archive Initiative (DataONE)</th>
<th>Community supported, e.g., the LTER NIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archer raw data</td>
<td>[Archive raw data] (but they do not need to be online)</td>
<td>Make data publicly available</td>
</tr>
<tr>
<td>Archive data</td>
<td>[Assign QC level to published data sets]</td>
<td>Assign QC level to published data sets</td>
</tr>
</tbody>
</table>

Streaming data management middleware

- "Middleware" software packages and procedures
  - Enable communication and management of data between field sensors and a client such as a database, website or software application
  - Purposes include the collection, archival, analysis, and visualization of data
  - Middleware is often chained together into a scientific workflow to meet multiple functional requirements
  - Considerations
    - Licensing, support, interoperability of components
- Proprietary middleware / software
  - Campbell Scientific – LoggerNet
  - Aquatic Informatics – Aquarius
  - Vista Engineering – Vista Data Vision (VDV)
  - YSI – EcoNet
  - Neokinetics Technology – WQData Live

Open source environments for middleware

- GCE Data Toolbox (MATLAB required)
- CUHIS Hydrologic Information System (HIS)
- DataTurbine Initiative

Image from Campbell et al., Bioscience, 2013.

Aquarius software

Eddypro software

LoggerNet software

Image from Heat Seek NYC, 2013.