Sustaining the Earth's Watersheds
stewards
Agricultural Research Data System

An Integrated ARS Watershed Data System

USDA

Vision

Research and analyses that inform public debate and support stewardship goals of
individuals
organizations
communities
and the Nation

Mission

To provide access to data and analytical tools to facilitate multidimensional watershed research and conduct analyses to support individual-, organization-, and policy-level decision making

Introduction

- USDA watershed research dates back to 1930's
- Many research watersheds are still operational
- Data have been managed and disseminated independently at each research location
- Data are varied, including GIS maps; satellite images; climatic, runoff, sediment, streamflow, and water quality records; and land use and management.

Needs

- Make data more accessible
- Make data more useful
- Support CEAP research
- Rescue historic data

Objective

- Develop and implement a data system to organize, document, manipulate, and compile climate, water, soil, management, and socio-economic data for assessment of conservation practices and other hydrologic analyses.

System Requirements

- Web-based, server-client architecture
- Meet stakeholder/user needs
- Meet the industry standards
- QA/QC procedures
- Ease-to-maintain, scaleable
- Data access: browser independent, speedy
- "Any-place, any-time" access

Metadata Requirements

- Who, what, when, where, why, and how about every facet of the data

- Document existing and future data to allow users to evaluate the nature and quality of the data

- Metadata search engine for finding the needed datasets

Metadata Standards Considered

- U.S. Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM)

- USDA's geospatial metadata standard
    1. **FGDC full compliance metadata**
    2. **USDA-Service Center Initiative (SCI) minimum compliance metadata**

- National Biological Information Infrastructure (NBII)'s Biological Data Profile of the CSDGM for both geospatial and non-geospatial data

What does the FGDC standard include?

Metadata

Main Sections

Supporting Sections

1. Identification
2. Data Quality
3. Spatial Data Org.
4. Spatial Reference
5. Entity and Attrib.
6. Distribution
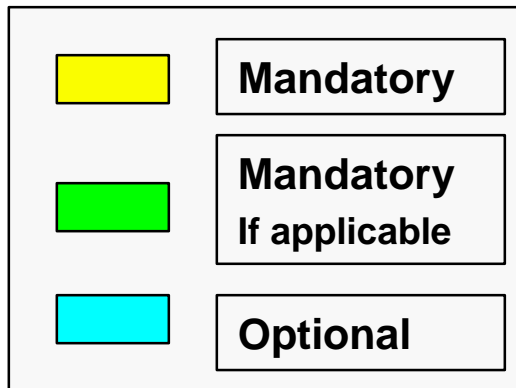7. Metadata Ref.

0. Metadata

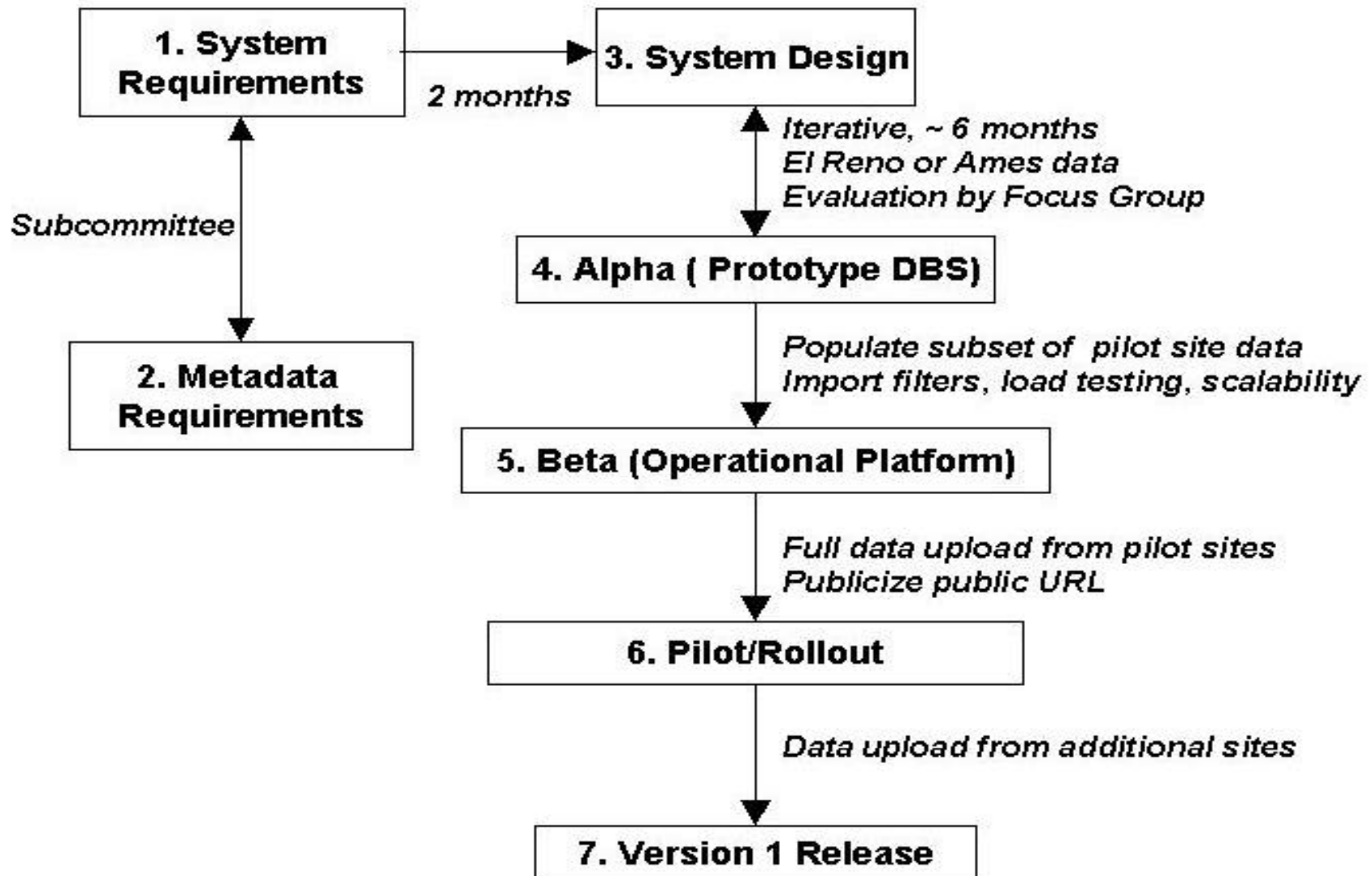8. Citation
9. Time Period
10. Contact

Legend:
- Mandatory (yellow)
- Mandatory If applicable (green)
- Optional (cyan)
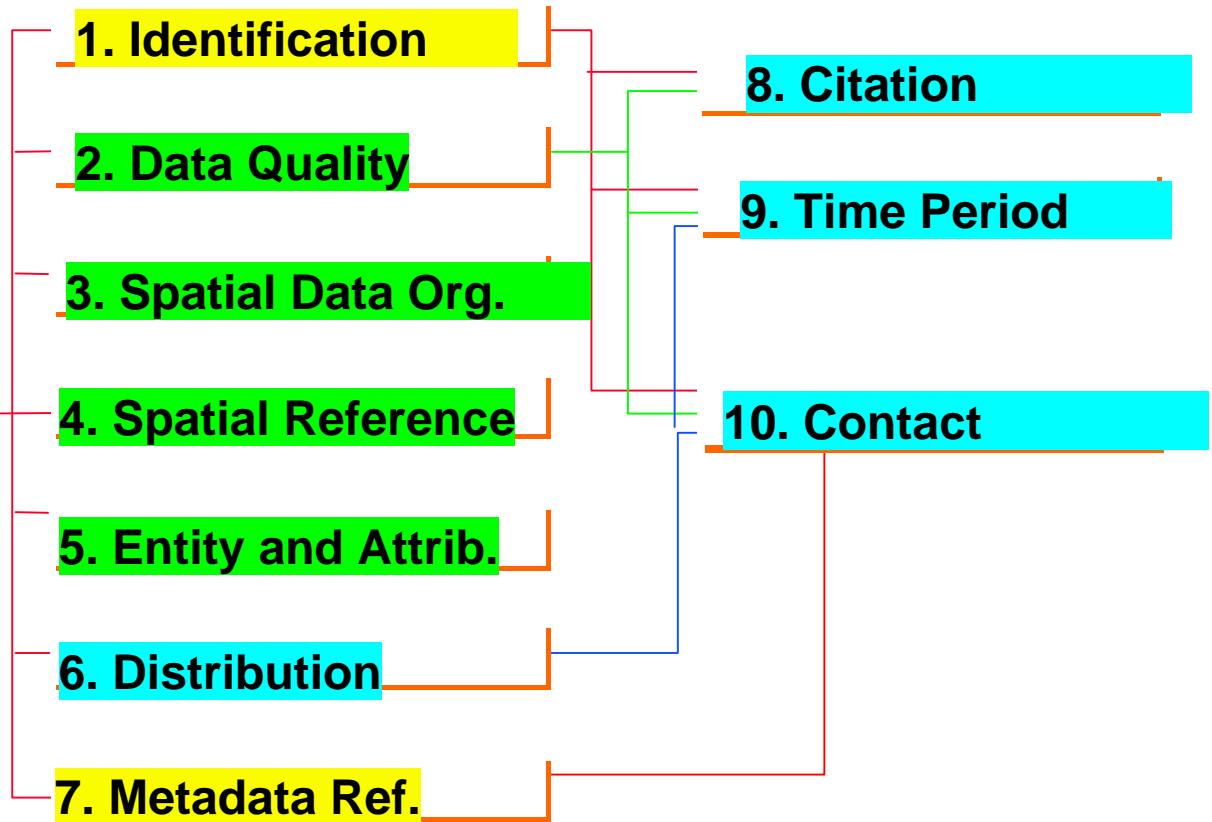
System Development Strategy

1. System Requirements → 3. System Design

2 months

Iterative, ~ 6 months
El Reno or Ames data
Evaluation by Focus Group

Subcommittee

2. Metadata Requirements

4. Alpha (Prototype DBS)

Populate subset of pilot site data
Import filters, load testing, scalability

5. Beta (Operational Platform)

Full data upload from pilot sites
Publicize public URL

6. Pilot/Rollout

Data upload from additional sites

7. Version 1 Release

Implementation Plan – Steiner/Sadler
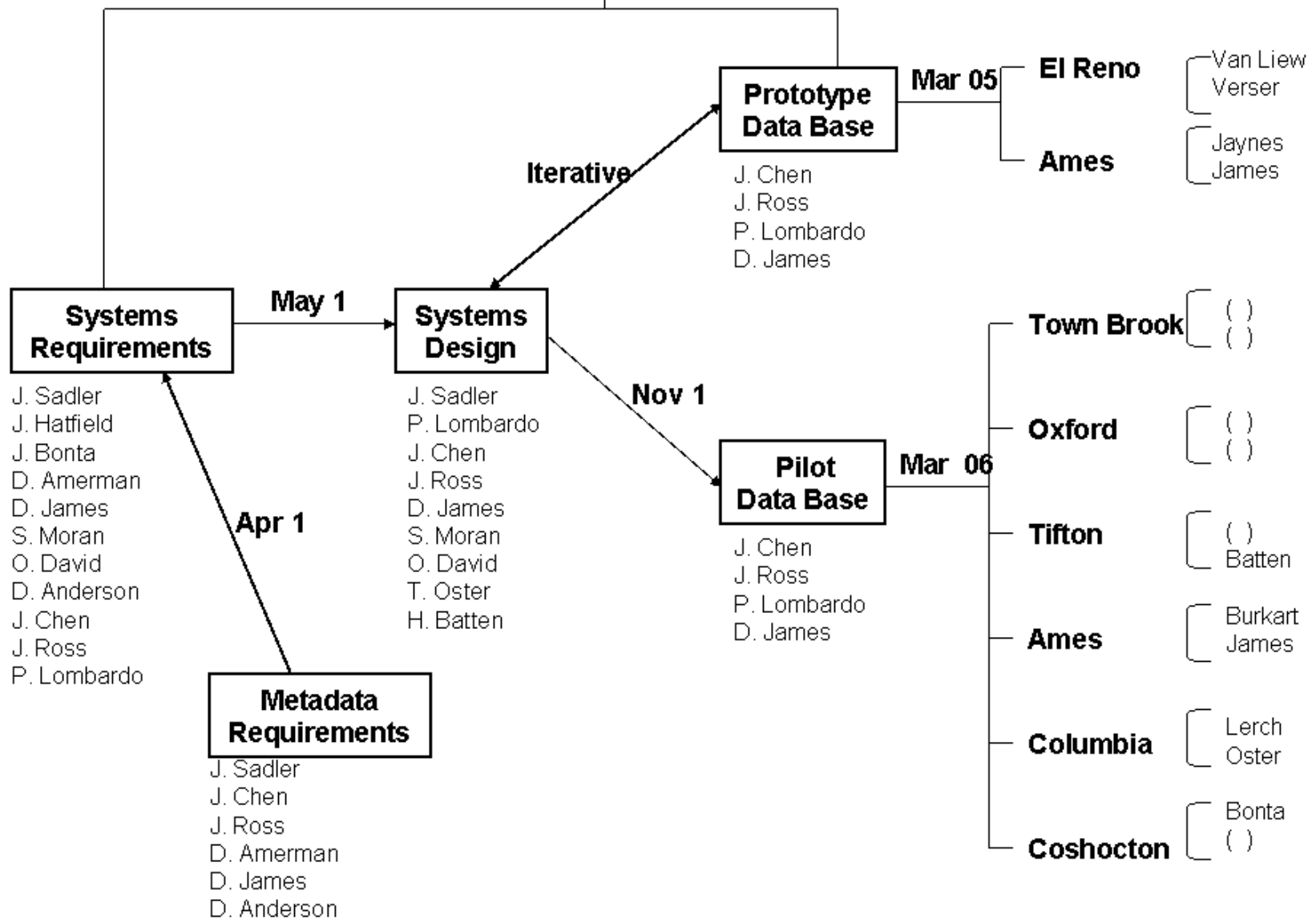
System Architecture

Local Sites

- **Create Metadata**
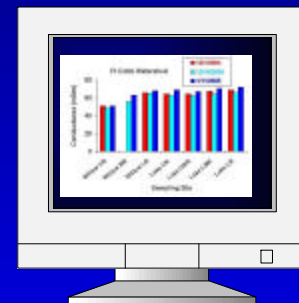- **Create filter for data harvest**
- **Provide URL**

Ft Cobb Watershed

Little Washita River Watershed
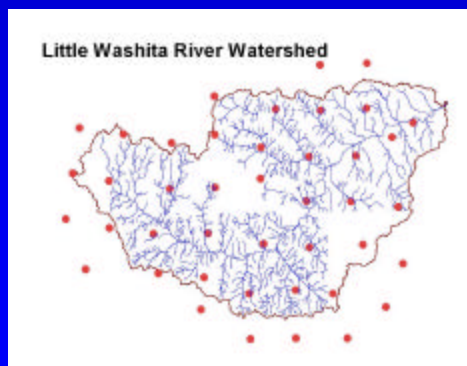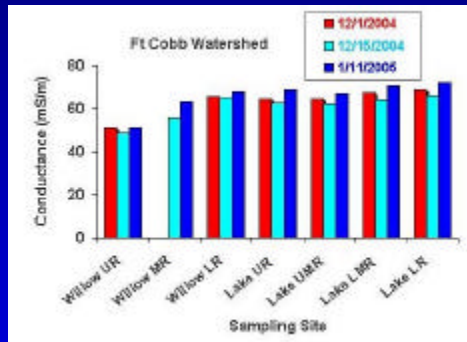
Server

- **Data Management System**
- **Metadata Search Engine**
- **Visualization & Analytical Tools**

Clients

Netscape network

Sustaining the Earth's Watersheds
stewards
Agricultural Research Data System

System Tools

- Utility to create input data for modeling & testing

- Utility to display reports and create downloadable files from query results

- Utility to generate statistics report

- Visualization tools – trends, time series, line/chart comparisons, outliers/abnormality identification

Prototype in El Reno, OK

- Hardware: A dual Xeon processor (2.8GHz CPU), 1 GB of RAM, and a 60GB hard drive

- Software: Mainly running Microsoft 2003 server with the Microsoft IIS and Apache coexisting web servers, and MS SQL/MySQL mixed servers. ASP.NET, PHP, and XHTML scripting languages are used.
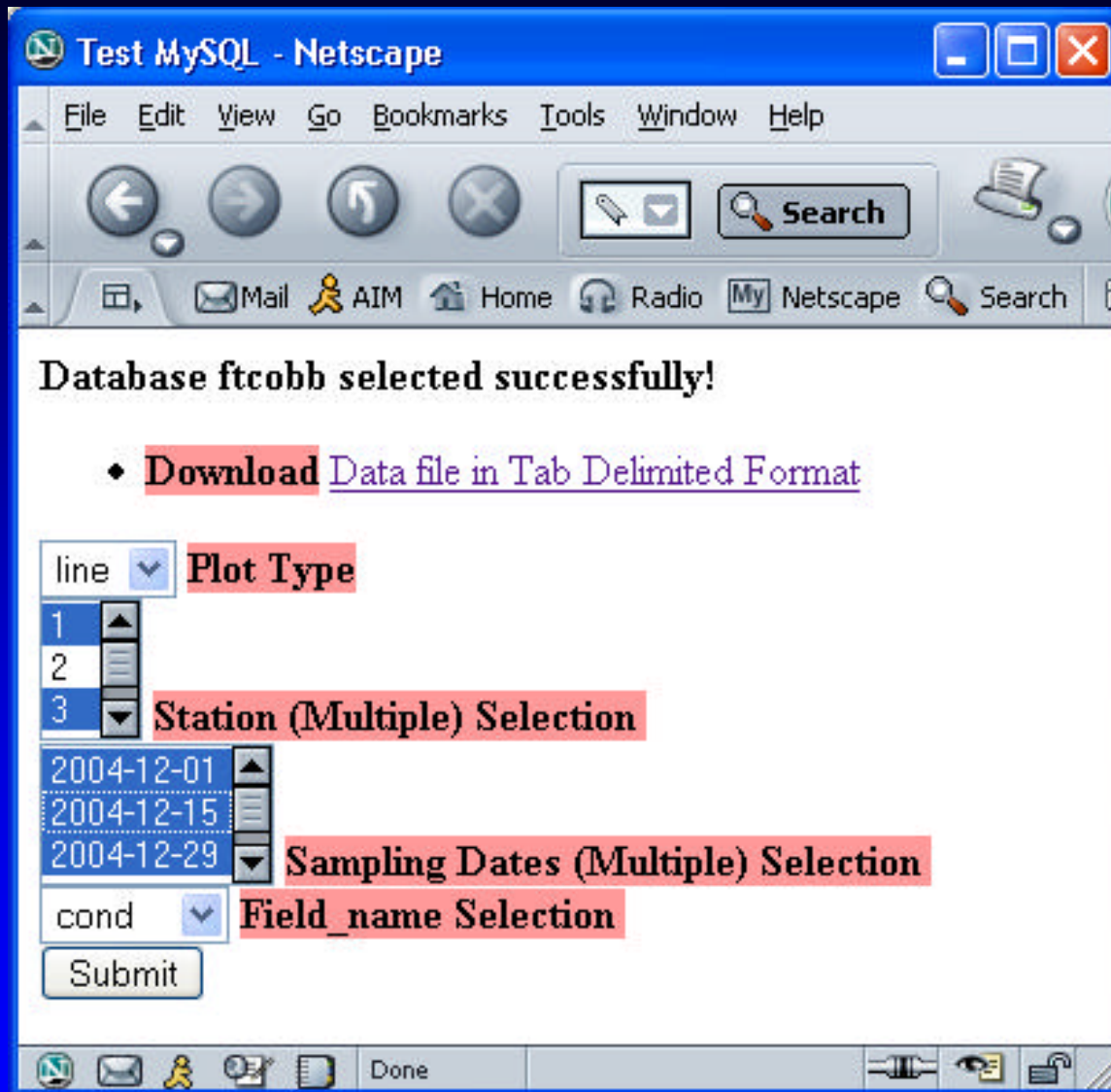
Screen provides
- **Data download**
- **Selection for:**
  - **Plot type**
  - **Station**
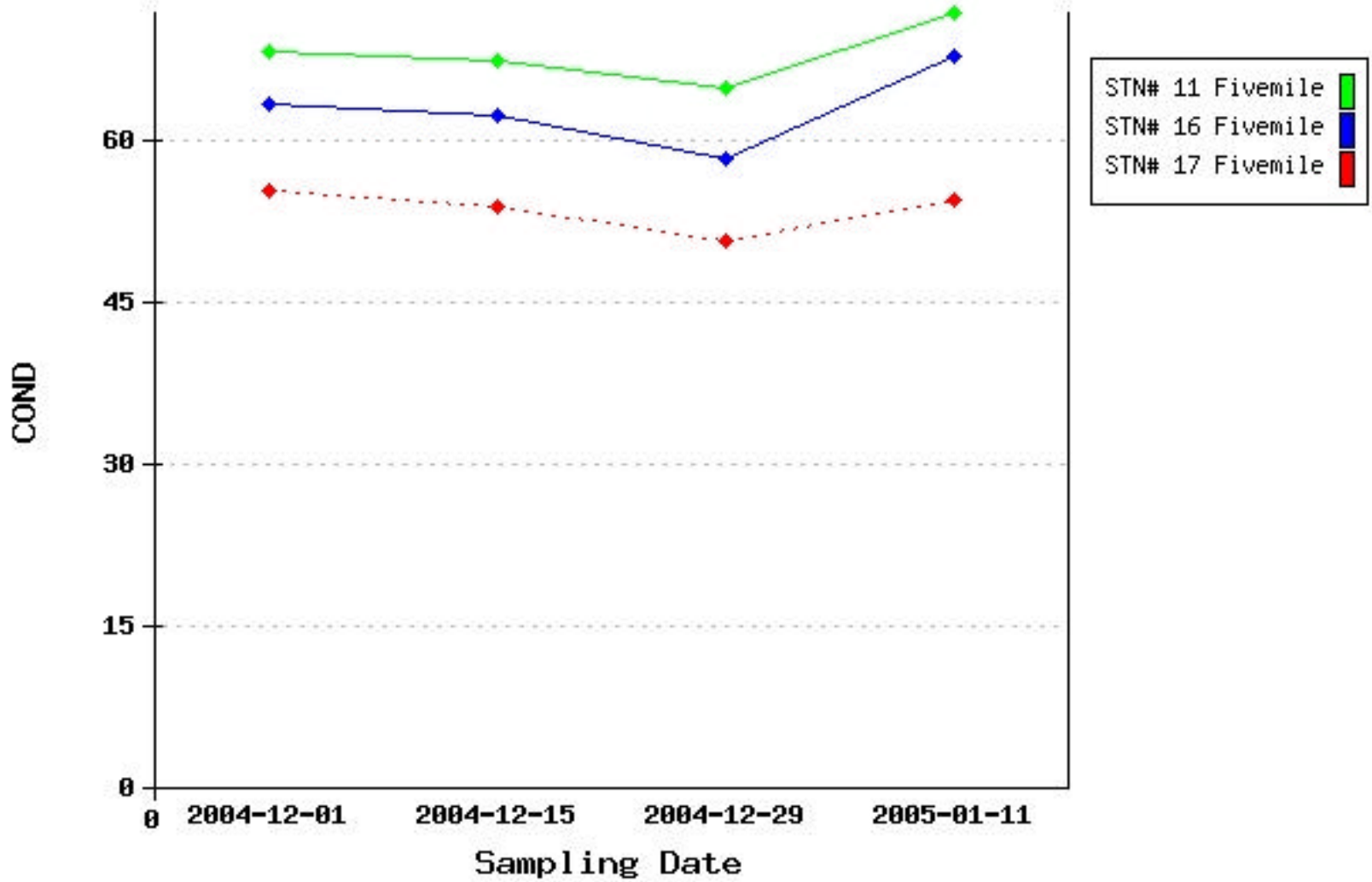 - **Time period**
  - **Parameter**
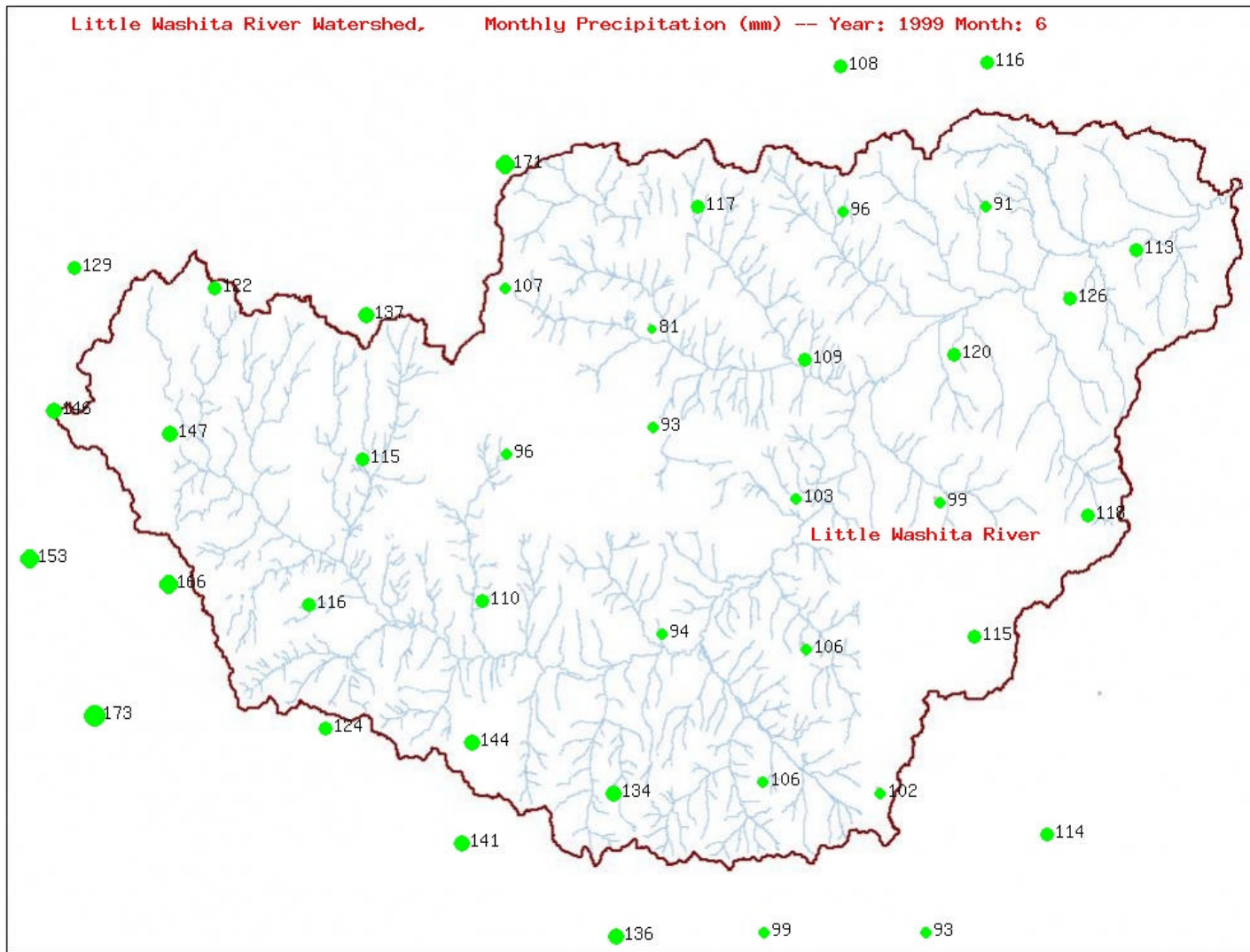
Query Results: Data table for display (downloadable)
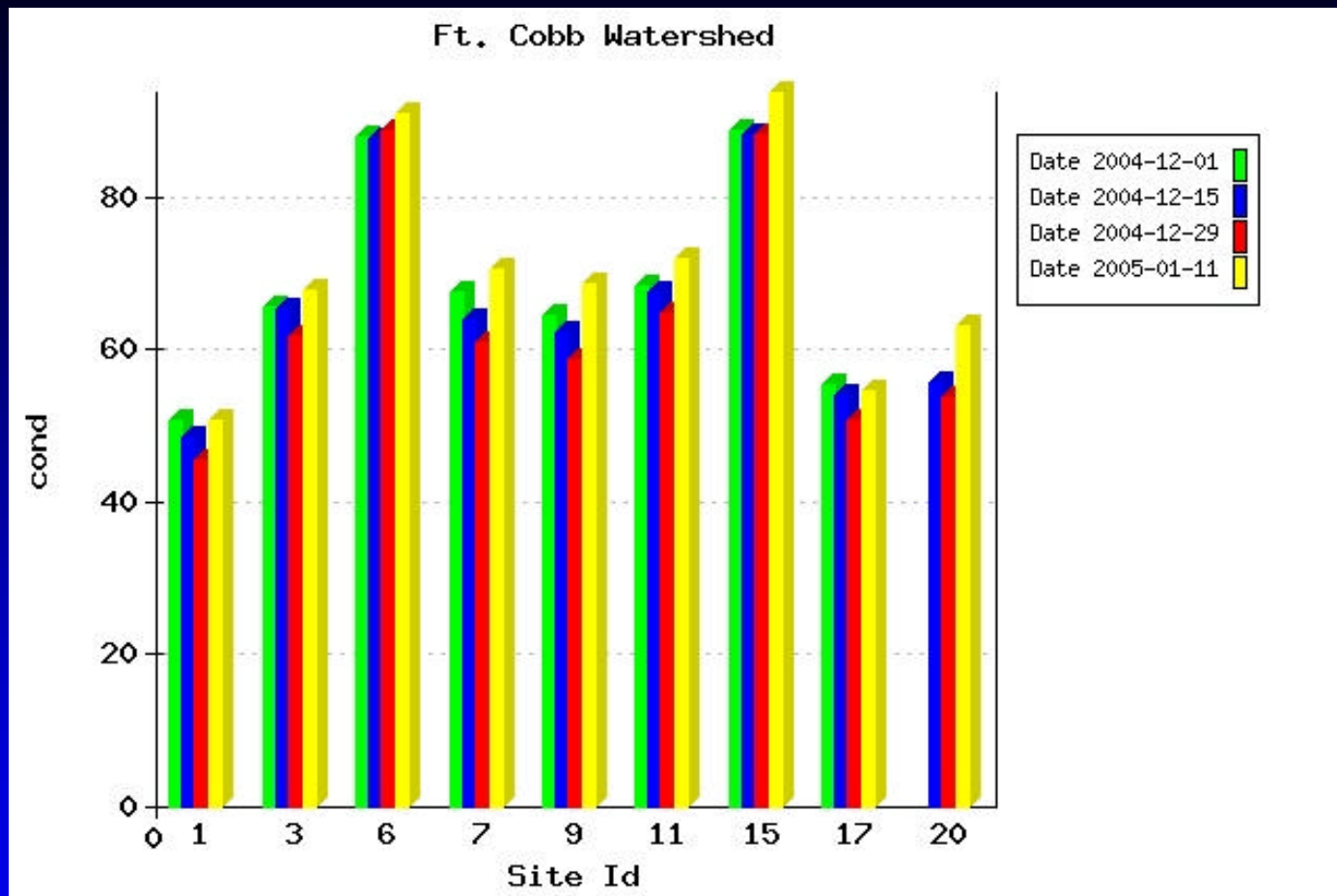
Bar Chart: Comparison of water quality (Conductance) among nine stations during four sampling periods

Ft. Cobb Watershed

Time **series** example

Spatial distribution of monthly precipitation
(symbol sizes represent the relative magnitudes of precipitation)

Summary

The ARS data system will provide

- A friendly environment
- One-stop access and data exploration for ARS scientists and the general public
- Metadata that inform users about the data
- Ability to explore and access the data
- Support to CEAP modeling and analyses
- Increased accessibility, utility, and impact to ARS watershed data.

ARS's commitment - as researchers, staff, research teams, Program Leaders, line managers, and an Agency - to stewardship of data and information from research supported by American taxpayers

Acknowledgements

- Jin-Song Chen, Research Hydrologist
- John Ross, IT Specialist
- E. John Sadler, Co-Chair of Data Team
- Data Team who are supporting this effort