Hydroinformatics (USU CEE 6930)

Class Time: Tuesday / Thursday 3:00 - 4:15 PM

Class Locations:

Utah State University: Distance Education 105
University of Utah: Park Bldg 302
Brigham Young University: 368F Clyde Building

Instructors:

Jeff Horsburgh
Utah State University
Office: UWRL 201
Office Hours: T/TH 4:15 – 5:15 PM in ENGR 223, or by appointment
E-mail: jeff.horsburgh@usu.edu
Phone: 435-797-2946

David Rosenberg
Utah State University
Offices: ENGR 213 / UWRL 205
Office Hours: T/TH 4:15 to 5:30 PM in ENGR 213, or by appointment
E-mail: david.rosenberg@usu.edu
Phone: 435-797-8689

Dan Ames
Brigham Young University
Office: 242J Clyde Building
Office Hours:
E-mail: dan.ames@byu.edu
Phone: 208-406-8055

Steven Burian
University of Utah
Offices: MCE 2044
Office Hours: M/W 10:00 – 11:00 AM, or by appointment
E-mail: steve.burian@utah.edu
Phone: 801-585-5721

Class Canvas Website: https://usu.instructure.com/courses/127332

Course Offering

This course will be simultaneously offered across the CI-WATER (http://ci-water.org/) partner campuses (USU, UofU, and BYU). Each university has designated an instructor to manage things on their end, be present during class time, help deliver course content, and evaluate plus support work by students on their campus. Support for development of this course was provided by National Science Foundation Grant No. EPS 1135482. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Course Overview

Hydroinformatics is the study, design, development, and deployment of hardware and software systems for hydrologic data collection, distribution, interpretation, and analysis to aid in the understanding and management of water in the natural and built environment. This class will introduce students to fundamental and advanced hydroinformatics concepts and procedures including automated data collection networks, relational databases and data management software, metadata and semantics, data storage file formats and standards, data transformations and automation of data manipulation tasks to support modeling and analysis, web based data distribution and access using web services, and integrated networks of hydro-climate data.

The course has an open, project-type format where students will work individually or in a small group over the semester to discover, organize and manage data for a hydrology or water resources problem of their interest. Projects may include designing appropriate data models and automating data loading, manipulation, and transformations in support of data intensive analyses or modeling. Class time will include lectures focused on learning and developing data management, transformation, and task automation skills, class discussions, code writing exercises to solve data manipulation tasks, demonstration of software and data systems, and student presentations of their project work. The course will better prepare students to work in data-intensive research and project work environments and emphasize development of reproducible processes for managing and transforming data in ways that others can easily and completely reproduce on their own to support analyses and modeling. Additionally, this course will better prepare students to work across multiple software platforms and systems used in data management.

Learning Objectives

Upon successfully completing this course, students will be able to:

Data and the Data Life Cycle:

1. Describe the data life cycle
2. Determine the dimensionality of a dataset, including the scale triplet of support, spacing extent for both space and time
3. Generate metadata and describe datasets to support data sharing
4. Discover and access data from major data sources

Databases and Data Models:

5. Store, retrieve, and use data from important data models used in Hydrology such as ArcHydro, NetCDF, and the Observations Data Model (ODM)
6. Develop data models to represent, organize, and store data
7. Design and use relational databases to organize, store, and manipulate data
8. Query, aggregate, and pivot data using Structured Query Language (SQL), Excel, R, and other software systems

Visualization, Transformations, Analysis, and Modeling:

9. Create reproducible data visualizations
10. Write and execute computer code to automate difficult and repetitive data related tasks
11. Manipulate data and transform it across file systems, flat files, databases, programming languages, etc.
12. Retrieve and use data from Web services
13. Organize data in a variety of platforms and systems common in hydrology and engineering
14. Prepare data to support hydrologic, water resources, and/or water quality modeling

Pre-requisites

Required:

- Graduate student standing
- Proficiency in Microsoft Excel and Geographic Information Systems (GIS)
- Familiarity and ability to write simple programs in any programming language such as C, C++, C#, Fortran,Visual Basic, R, Matlab, Python, or Java.

Recommended:

- Concurrent or prior enrolment in a database class, Hydrologic Modeling, Integrated River Basins/Watershed Planning and Management (CEE 6490), GIS in Water Resources (CEE6440)

Texts

Required readings will be posted on the class website or distributed in class.

Some potential references, but not required:

- Kumar, P., (2005), Hydroinformatics: Data Integrative Approaches in Computation, Analysis, and Modeling, CRC Press, 552 p.
- Grayson, R. and G. Blöschl, ed. (2000), Spatial Patterns in Catchment Hydrology: Observations and Modelling, Cambridge University Press, Cambridge, 432 p, full PDF text available at http://www.catchment.crc.org.au/special_publications1.html.

Class Schedule (subject to change - updated regularly on class webpage)

Date	Topic	Lecturer	Course Work
Aug. 28	Data and data problems in the information age	Horsburgh	
Aug. 30	Data management and the data life cycle	Horsburgh	
Sep. 4	Metadata	Horsburgh	
Sep. 6	**CI-WATER Symposium (No Class)**		**Project Topic Due**
Sep. 11	Data models	Horsburgh	
Sep. 13	Data model design	Horsburgh	**ILO 1 Due**
Sep. 18	Storing and manipulating data using Google Fusion Tables	Ames	
Sep. 20	Data Visualization and Mapping in Microsoft Excel	Ames	**ILO 2 Due**
Sep. 25	Database implementation	Horsburgh	
Sep. 27	Using SQL to query a relational database	Horsburgh	**ILO 3 Due**
Oct. 2	Aggregating and pivoting data – SQL and Excel	Rosenberg	
Oct. 4	Moving data across platforms – database to R and analyzing data	Horsburgh	**ILO 4 Due**
Oct. 9	Programming for automation of common data management tasks, part 1	Ames	
Oct. 11	Programming for automation of common data management tasks, part 2	Ames	**ILO 5 Due**
Oct. 16	**Student project interim oral presentations**		**Oral Presentation Due**
Oct. 18	**Friday Class Schedule (Fall Break – No Class)**		
Oct. 23	Service Oriented Architecture for Sharing Hydrologic Data: The CUAHSI HIS	Ames	
Oct. 25	Using HydroDesktop to search, discover, view, and analyze data in the HIS network	Ames	**ILO 6 Due**
Oct. 30	HydroServer Lite: A quick way to get data online and join the CUAHSI HIS	Ames	

Nov. 1	Customizing HydroServer Lite to meet the needs of your organization	Ames	**ILO 7 Due**
Nov. 6	Water resources modeling example	Rosenberg	
Nov. 8	Data model and data preparation for water resources modeling example	Rosenberg	**ILO 8 Due**
Nov. 13	Introduction to the OpenMI Modeling Environmnent in HydroDesktop	Ames	
Nov. 15	Using OpenMI and HIS data to perform hydrologic modeling	Ames	
Nov. 20	Visualizing high-dimensional data	Rosenberg	**ILO 9 Due**
Nov. 22	**Thanksgiving Holiday (No Class)**		
Nov. 27	Course wrap-up and review	Horsburgh	
Nov. 29	**Final Project Oral Presentations (USU)**		**Project Reports Due**
Dec. 4	**Final Project Oral Presentations (UofU)**		**Peer Review Due**
Dec. 6	**Final Project Oral Presentations (BYU)**		
Dec. 13	**1:30 PM Final Oral Presentations Continued**		

Description of Required Course Work

Individual Learning Opportunities (ILOs)

Students will complete a series individual learning opportunities (ILOs). Each ILO will pose a problem related to data organization, management, or transformation and will require use of software tools and/or computer programming/automation to solve. The list of ILOs follows, but is subject to change:

- ILO-1. Metadata and the data lifecycle
- ILO-2. Data model design
- ILO-3. Google Fusion Tables and visualization and mapping in Excel
- ILO-4. Database implementation and loading data
- ILO-5. Querying, visualizing, transformation, and analysis
- ILO-6. Automation of data management tasks
- ILO-7. Accessing data using web services
- ILO-8. Deploying an HIS HydroServer to share your data
- ILO-9. Preparing data for input to a model

Students will submit the answer to each ILO in the format of a 1-page engineering report or briefing paper. Each report should be formatted with standard 12-point font, 1-inch margins, and be fully self-contained to include an introduction to the problem, methods used, and results obtained so that a technically-versed reader not familiar with the problem statement can understand the rationale for, methods used, and results of the work presented. The restriction to 1 page is to help develop clear and succinct writing skills. Planners, managers, politicians, and decision makers rely on the work of engineers and hydroinformatics experts but are very busy people who rarely understand computer code or the underlying technologies. References, figures, tables, listing of code, and more detailed explanations can be included in appendices that do not count towards the 1-page limit. Students not satisfied with their performance on an ILO report may revise and resubmit it up to 1 week after the instructors evaluate and return the ILO report.

Semester Project

A large part of the course work will involve a semester-long project. Individuals or groups of up to 3 students will choose a hydrology or water resources data management problem to study for the semester. As part of the project, students will identify the data, data dimensionality, and lifecycle of the data required to solve the problem and develop metadata to describe and a data model to organize the data. Subsequently, students will implement the data model in a data management system and undertake querying, pivoting, and transformation activities to move the data into a model and/or analysis that uses the data. Students will report and make recommendations based on their results. Semester project work will comprise the following components:

1. **Topic**: A title and short paragraph describing the data management problem to be explored (not graded).
2. **Interim oral presentation**: A brief, in-class presentation that gives background on the data management problem you have selected for your semester project. Described the data, including dimensionality, lifecycle, metadata, and data model and/or formats that you will use to organize the data.
3. **Final Project Report**: Students will write a final project report providing general information on their data management project, as well as documentation of and results of their work. Final reports should synthesize all the work done on the project for the semester and allow readers to access and reproduce the methods developed and used. ***Please note that final project reports will be made publicly available on the course website***.
4. **Peer review**: Students will be assigned to provide feedback to another project report. Students will assess whether the work they review is reproducible and identify ways the study and write-up could be improved.
5. **Final oral presentation**: Your final project oral presentation should expand upon the first interim presentation to include details about software used, code written, databases developed, data transformations performed, modeling and/or analysis undertaken, and key results. You will deliver your final presentation in class, and it will be broadcast to all three campuses.

Class Participation

We expect students to read assigned readings ahead of time and come to class and share their impressions of the reading(s) or ask questions on points they did not understand. At times, we may discuss readings in a seminar format. During lectures or discussions, the instructors will ask many questions, and, if needed, call on you individually to ensure everyone participates.

Grading of and Expectations for Submitted Work

Approximate Grading:

Individual Learning Opportunities		50%
Semester Project		40%
Interim oral presentation	10%	
Final project report	15%	
Peer review	5%	
Final oral presentation	10%	
Class Participation		10%

90 to 100% -- at least some sort of A
80 to 90% -- at least some sort of B
70 to 80% -- at least some sort of C
< 70% -- most probably some sort of F

There is no curve. All submitted work (that meets grading rubric standards) will earn an "A".

All submitted work will be graded for technical correctness, organization, clarity, presentation, and other criteria according to the Grading Rubric available on the class website for the item. We will ask to meet with students who submit low quality work to discuss improvement strategies.

Submitted work must be:

- Original, typed with 1" margins in a standard 11-point font, printed, and stapled.
- Have a title page with title, student name(s), date, email address(es), class, and instructor.
- Handed in at the **beginning** of class on the due date listed on the class web page.
- Turned in with the self-assessment portion of the Grading Rubric completed.
- Turned in with the Group (and self) Rating Form completed (for group work items).
- Turned in with duplicate electronic copy and the self-assessment via Canvas.
- For group work, only one paper and Canvas copy need be submitted.
- **We do not accept late assignments.** They will be **graded as zero**. In extenuating circumstances (birth/death in the immediate family; grave illness with doctor's note), contact the instructors **prior** to the due date and make alternative arrangements.

Academic Integrity: We expect each student to uphold academic integrity. See http://www.usu.edu/policies/PDF/Acad-Integrity.pdf. For example, USU, the CEE department, and the instructors take plagiarism seriously, and we will hold offending parties to the full extent of the USU Code. When in doubt, acknowledge sources, cite references, and properly quote material that is not your own.

Electronic Policies

1. Class Webpage: We will post all class materials to the class webpage(https://usu.instructure.com/courses/127332) including readings, lecture materials, and descriptions of and grading rubrics for all course work.
2. Canvas: Submit electronic versions of all work and self evaluations to Canvas (for grading and archiving purposes).
3. Project Reports: Please note that your project reports will be made publicly available.
4. Email: **Include "Hydroinformatics" in the subject line** of all email so that we can timely attend and respond to emails. Unless you request that we don't, we may forward email questions and answers to the entire class.

Expectations of Students

- Be on-time to class and ready to learn / participate when class starts.
- Read assigned readings ahead of time and come to class prepared to share your impression(s) of the reading(s) and/or ask questions on points you do not understand.
- Turn off or keep silent all electronic devices that may distract the instructors or other students. We will ask students using phones, pagers, PDAs, music players, etc. to leave class.
- Contribute to class discussions while being respectful of and listening to others' points of view.
- Turn in all work on time in the required format.

- Bring questions and concerns forward during class, office hours, or by chat or email.
- Put in approximately 2 – 5 hours outside of class for each 1 hour of in-class time.

Expectations of the Instructors

- Start class on time.
- Respect the value of student's time.
- Call equally on all students for class participation.
- Learn student names by some point through the semester.
- Facilitate an environment of inclusivity and non-discrimination.
- Respond to email within 30 hours when we are not traveling out of town.
- Return graded work within 1 week from when work is submitted.

Disability/Special Accommodations

Please talk to one of the instructors immediately if you require disability or other special accommodations.

Additional Resources for Students

- Partial list of data management problems for semester study (see the last page of syllabus)
- CUAHSI Hydrologic Information System Website: http://his.cuahsi.org
- Class web page: https://usu.instructure.com/courses/127332
- Dr. Rosenberg's web page: http://www.engr.usu.edu/cee/faculty/derosenberg/
- Dr. Horsburgh's web page: http://jeffh.usu.edu
- Dr. Burian's web page: http://www.civil.utah.edu/~burian/

Please direct further questions or concerns about the syllabus or the course to one of the instructors by email, in person, or phone.

Importance of IDEA Course Objectives

USU uses the IDEA system for students to provide course evaluations at the end of the semester. As instructors, we have rated the relevance of the 12 IDEA Objectives to what we would like to accomplish with this course. Our teaching materials and methods will be focused most on those objectives rated as "Essential" or "Important."

IDEA Course Objective	Minor/No Importance	Important	Essential
1. Gaining factual knowledge (terminology, classifications, methods, trends)	X		
2. Learning fundamental principles, generalizations, and theories	X		
3. Learning to apply course material (to improve thinking, problem solving, and decisions)			X
4. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course			X
5. Acquiring skills in working with others as a member of a team		X	
6. Developing creative capacities (writing, inventing, designing, performing in art, music, drama, etc.)		X	
7. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)	X		
8. Developing skill in expressing myself orally or in writing		X	
9. Learning how to find and use resources for answering questions or solving problems			X
10. Developing a clearer understanding of, and commitment to, personal values	X		
11. Learning to analyze and critically evaluate ideas, arguments, and points of view		X	
12. Acquiring an interest in learning more by asking my own questions and seeking answers			X

Some Potential Data Management Problems to use for Semester-Long Project

- Search out, download, and input required data into a Water Evaluation and Planning (WEAP) system model or other watershed hydrology or water quality model and evaluate changes in land use, water use, or climate.
- Develop an empirical model for the prediction of Logan River streamflow (peak, annual volume, spring runoff volume) based on available precipitation and snow data. An analysis that combines information from SNOTEL, NCDC precipitation, NOHRSC SNODAS, DEM, Land Use/Land Cover, USGS Streamflow, NWS CBRFC, etc.
- Assess the potential for water shortage due to drought through the use of drought indices that combine information from streamflow, reservoir storage, snow storage, precipitation and other appropriate data resources. (Build around NIDIS drought project http://drought.usu.edu, http://extension.usu.edu/drought/htm/drought, http://www.drought.gov/)
- Assess and quantify the impacts of water use on inflows to the Great Salt Lake (or one of the contributing basins). Required integration of data on Natural Hydrology as well as diversions, consumptive uses, transfers, etc.
- Assess the impacts of retreating glaciers on Water Resources in parts of the world where streamflow is glacier fed, e.g. northern Rocky Mountains, Glacier National Park and Wyoming, Himalaya's, South America. Could combine information from remote sensing on glacier extent and properties, climate forecasts (ICPP and various downscaling activities) and draw upon application of UEB snowmelt model for glacier melt evaluation.
- Assess the potential for water shortage in a major river basin due to increasing demand and climate change (e.g., the Colorado River basin). Combine information from streamflow, storage, land use/land cover. Climate change and downscaling (ICPP etc). Could draw upon Barnett and Pierce Lake Mead going dry work.
- Assess the trends and patterns in streamflow in parts of the world and associations with climate, (e.g., shifts in volume and timing of spring runoff in association with climate change, ENSO, PDO etc.). Requires combination of data from multiple sources to do the analysis.
- And others!