Geosciences 215: Earth System Data Science
Spring 2019 Syllabus

Instructor: Hilary Palevsky, hpalevsky@wellesley.edu, Office: Simpson 205

Class meetings: 1:30-4:10 pm Tuesdays, Gray Lot Modulars 302L

Course website: https://sites.google.com/wellesley.edu/geos215-sp-19

Course TA: Lucy Wanzer, lwanzer@wellesley.edu

Course description:

Earth system science increasingly depends on analysis of large data sets from in situ and autonomous observations, satellite remote sensing, and numerical model output. In this course, students will learn to visualize earth system data across broad spatial and temporal scales in maps and time-series plots using MATLAB, and apply statistical tools to analyze trends and variability in time-series data and spatial variability in regional and global datasets. Students will practice these skills in a series of data analysis assignments focused on regional and global climate data, and will apply them in a final research project addressing earth system science questions of their own choosing.

Overarching learning goals for the class:

By the end of this class, you will be able to…

- Visualize earth system data across broad spatial and temporal scales in maps and time-series plots using MATLAB
- Apply statistical tools to analyze trends and variability in time-series data and spatial variability in regional and global datasets
- Describe and critically evaluate methods such as observational compilations, remote sensing, and numerical models used to generate global earth science datasets
- Select and justify appropriate data sources and analytical methods to address scientific questions about earth system processes

Transferrable skills:

In this class, you will work on developing your expertise in this broad set of transferrable skills. You each come in with different sets of prior experience with these skills – therefore the class is designed to allow you to build on your existing strengths and to emphasize those skills you are most interested in developing.

1) Programming in a scripting language, including: creating variables, performing calculations, importing and reshaping arrays of data, creating and using functions, and using help documentation and online searches to troubleshoot code
2) Statistical analyses of data, including: calculating the mean, standard deviation, and linear regression, de-trending, removing outlier data points, and error analysis

3) Visualizing data, including: plotting 2-dimensional data (e.g. variable changing over time), using contours and colormaps to visualize 3-dimensional data, creating maps, and optimizing data visualizations for clarity and accessibility
4) Critical consumption of publicly-available data, including: finding data available through public repositories, reading documentation to understand how data were collected and processed, and assessing the suitability of a dataset to address a particular problem/question
5) Reading scientific papers, including approaches for understanding papers with methods, concepts, and/or technical detail that you don't have previous experience with
6) Research using scientific data, including: formulating a question, framing a hypothesis, selecting methods to test your hypothesis, and interpreting results to address your original question
7) Synthesizing and presenting research findings through written papers and poster presentations
8) Collaborating productively with groups, including: dividing roles across team members, managing all team members' expectations, and successfully producing team-created products.

Assignments:

Class preparation assignments

There will be a number of assignments you will complete in preparation for our Tuesday class meetings. These will include an online MATLAB tutorial, reading scientific papers related to the CO_2 data lab in preparation for class discussions, and preparing for conversations with guests about their data science career paths. Completing these assignments will be critical for your full and effective participation during our class meetings, and each of these assignments will be submitted through a Google Form due before class.

Data labs: Temperature and CO_2

The first two data analysis projects in this class will be multi-part data "labs" where you will conduct guided analysis of datasets that I have selected for you. This will provide an opportunity for you to develop your skills in programming, data analysis, and data visualization, and apply these to questions about climate science that we will pose and discuss together as a class. You work on these labs in pair-programming teams, but will each submit your own lab write-ups explaining and presenting your analysis.

Team research projects

The second half of the semester will be focused on completing team research projects addressing earth system science questions of your own choosing, applying data science approaches to analyze existing publicly-available data sets. We will begin these team projects with an opportunity for everyone in the class to pitch ideas for science questions you want to work on, and personal goals for skills you want to work on in your project, since each team will have lead roles for different aspects of the project (programming lead, science lead, etc.). This will be followed by a project "speed dating" session to help you form teams with common interests and complementary skills. The reason for designing these projects to be completed in teams is that it will allow you to practice working in collaborative teams who each bring

complementary strengths and skills, as happens in real research, enabling you to together complete a more complex and extensive project than you could complete independently. All team deliverable products will be created collaboratively, and accompanied by a joint statement of contributions/lead roles on each section, as well as an individual reflection from each team member.

These team projects will be conducted in two stages: teams of 3-4 people will work on a "Version 1" project focused on scoping the project and research questions, identifying and accessing relevant data sets, and conducting preliminary data processing and analysis to determine a plan for future directions in the "Version 2" portion of the project. Each team will present their Version 1 project to the full class and submit a Version 1 report to me for feedback and assessment. You will then have the opportunity to decide whether to continue with your original project for Version 2, or whether to switch to work with a different team for this second component, which will result in a full written report and a research poster presenting the group's final results and conclusions. The goal of conducting the project in these separate stages is to provide a mid-way milestone to help in guiding your research, and to provide an opportunity to switch directions if you find that a different team's project would be a better fit for your personal goals in the class.

Setting your own goals (SYOGs) and reflections:

This syllabus identifies a set of learning goals and transferrable skills for you to work on that I have set for the course. But equally important are **your** goals for the course. At the beginning of the semester, you will decide which skills you would like to prioritize working on from the set of Transferrable Skills listed earlier in this document. You may also include your own additional skill goals or other personal objectives for the course. The purpose of this is to help you focus on skills that are of personal interest, and use this focus to decide which roles you want to take on in the team research projects. I will ask you to reflect on your progress towards these personal goals (as well as any changes in your personal goals) throughout the semester in individual reflection assignments you will submit along with each lab and research project deliverable. This will include reflecting on:

1) The transferrable skill goals I set for the course –what skills did you find yourself developing by working on this assignment, what effort did you put towards each of these, and how would you self-assess your achievement on each skill involved in this assignment?
2) Your progress thus far towards your own SYOG goals, including any changes in your personal goals for the class,
3) Your individual experience as a member of your group – what were the strengths and weaknesses of the group as a whole, your own personal strengths and weaknesses as a member of the group, and what lessons do you take away about effective group work from this assignment?
4) Anything else you want to share about your experience with this assignment.

Note: I will never reduce your grade because of honesty in anything you reveal in your reflections that I would otherwise have been unaware of.

Assessment:

Grading criteria:

Class preparation assignments – 10%
Temperature data lab – 10%
CO_2 data lab – 15%
Team research project, Version 1 – 20%
Team research project, Version 2 – 25%
Setting your own goals (SYOG) project – 20%

Statement on the Wellesley grading policy:

The College's grading policy http://www.wellesley.edu/registrar/grading/gradingpolicyfaq asks faculty to hold each 100- and 200-level course with 10 or more students to an average of no higher than 3.33. However, it does not require faculty to grade on a "curve" and you will not be competing with your peers for grades in this class – this class is explicitly designed to encourage and reward collaboration and teamwork. There is no arbitrary limit on the number of A's, B's, C's etc., and every student will be assigned the grade they earn and deserve.

Guiding principles:

Recognizing that each student enters this class with a different set of prior experiences and strengths in the skills that this class aims to cover, the assessment scheme is designed to allow different paths through the course focusing on individualized choices of skills to emphasize. Students who receive A's in this class will achieve a satisfactory level in all skill areas, but may select different areas of personal interest to focus on in achieving more complete levels of mastery.

Class policies:

Office hours and email:

I will be available for drop-in meetings during my office hours, which will be scheduled based on a poll of all students in my classes and then announced in class and on the course website. In addition to my scheduled office hours, you are welcome to make an appointment with me at a mutually convenient time. The best ways to set a meeting are to catch me in class or email me. I will respond to your email as soon as I can, usually within 48 hours and often much sooner, but please do not count on more rapid e-mail turn-around time just before deadlines!

Accessibility and accommodations:

It is my goal to create a learning experience that is as accessible as possible. If you anticipate any issues related to the format, materials, or requirements of this course, please meet with me outside of class so we can explore potential options. If you are a student with a disability or condition, either long-term or temporary, and need or think you may need disability-related accommodations, I will be glad to work with you and with Disability Services to accommodate your needs for this course. If you are unsure but suspect you may have an undocumented need for accommodations, you are encouraged to contact Disability Services. They can provide assistance, including screening and referral for assessments. Disability Services can be reached at disabilityservices@wellesley.edu, at 781-283-2434, by scheduling an

appointment online at their website www.Wellesley.edu/disability, or by visiting their offices on the 3rd floor of Clapp Library, rooms 316 and 315.

Course schedule:

This schedule is a fluid document and may change as the semester progresses. Changes will be announced in class and posted to the course website.

Date	Class plan	Assignments
Jan. 29	Introductions, course overview, & logistics + What is earth system data science? + Getting started with GitHub & MATLAB	For class: Install MATLAB & GitHub + complete intro surveys
Feb. 5	Methods of time-series data analysis + Temperature data lab, Part 1	For class: Complete MATLAB Onramp tutorial
Feb. 12	Methods of visualizing data + Temperature data lab, Part 2	For class: Complete SYOG + sign up for SYOG meeting
Feb. 19	****No class – Monday schedule****	DUE Fri. Feb 22: Historical and future temperature data lab
Feb. 26	Analyzing ocean CO_2: Discussion of Takahashi et al., 2002 + Atmosphere & ocean CO_2 data lab, Part 1	For class: Read Takahashi et al. 2002, *Deep-Sea Research II*
Mar. 5	Mapping sparse data: Discussion of Landschützer et al., 2014 + Atmosphere & ocean CO_2 data lab, Part 2	For class: Read Landschützer et al. 2014, *Global Biogeochemical Cycles*
Mar. 12	Finding publicly-available data sources + Group project pitches & "speed dating"	DUE Fri. Mar 15: Atmosphere & ocean CO_2 data lab
Mar. 19	Guide for earth system data consumers, Part 1 + Work on Version 1 group projects	DUE Wed. Mar 20: 1-page Version 1 project prospectus
Mar. 26	****Spring break****	
Apr. 2	Guide for earth system data consumers, Part 2 + Work on Version 1 group projects	
Apr. 9	Class presentations on Version 1 group projects + Pitches and group formation for Version 2 projects	DUE Fri. Apr 12: Group writeup of Version 1 project + individual reflections
Apr. 16	Careers in earth system data science + Work on Version 2 group projects	For class: Questions for guest
Apr. 23	Careers in earth system data science + Work on Version 2 group projects	For class: Questions for guest
Apr. 30	Work on Version 2 group projects	
May 7	Project poster presentations	
Finals week	DUE: Group writeup of Version 2 project + individual final reflections	