Using Oceanographic Mooring and Satellite Datasets to Teach Data Analysis and Scientific Computing Skills for Undergraduate Students

Hilary Palevsky
Boston College
palevsky@bc.edu

Full details & Blob Data Lab materials
My class context

- Data and coding skills focused course in Earth & Environmental Science
- Sophomore undergraduates through 1st year graduate students
- Range of prior programming experience (none – CS major)
- Range of prior experience studying or interpreting environmental data
- Class taught in MATLAB

Key learning goals

- Develop skills and confidence in programming in a scripting language
- Create visualizations of time series and spatial datasets
- Apply basic statistical tools to analyze large datasets
- Become savvy consumers of publicly-available data
- Collaborate productively with groups
Course outline

• Setting your own goals (SYOG)
  • Students set personal goals for the semester, to guide which learning goals they most want to personally prioritize

• Data lab #1 – Historical and future temperature
  • Time-series analysis of monthly mean temperature data (observed & modeled)

• Data lab #2 – Ocean CO₂
  • 4-dimensional data analysis and visualization: Reproduction of Takahashi et al. 2002

• Data lab #3 – The Blob
  • Using raw, messy data from the Ocean Observatories Initiative Station Papa Array to investigate Northeast Pacific heatwaves (aka The Blob)

• Team research projects
  • Students address questions of their own choosing using publicly-available datasets
Learning goals for the Blob Data Lab

• Read in and explore netCDF data files to identify relevant variables and metadata
• Find and download data from online data repositories
• Read documentation to understand how publicly-available data were collected and processed
• Plot and interpret raw data, including identifying and excluding outliers
• Combine and compare data from multiple data sources
• Evaluate the strengths and weaknesses of different data sources that could be used to approach the same question or calculation
Ocean Station Papa: In the midst of The Blob

Satellite data plotted by students at end of the lab
The messy, raw data
Temperature anomalies the students calculate:
Mooring + World Ocean Atlas climatology vs. Satellite SST
Challenges & techniques to address them

Students are often afraid of coding and need to build confidence.
• Students complete data labs in Pair Programming teams
• Low stakes mid-semester Programming Assessment with revisions opportunity
• Explicitly discuss difference between productive & unproductive struggle
• Tap into intrinsic motivation through Setting Your Own Goals reflections

Students are often also resistant to or afraid of group work.
• Include (and pitch) collaboration as explicit skill to develop in course
• Pair Programming team matches informed by individual surveys
• Graded component of Data Labs is a writeup completed individually
Every dataset has a story, and understanding it can guide the choice of suitable analyses; some have labeled this data understanding as *data empathy*. The reason for understanding where the data come from is twofold: first, understanding how the data are generated, their purpose, and generation processes will guide your investigation. Second, understanding the inherent biases in the data gives you a chance to correct them or adjust your results and recommendations.
Where the data come from
Where the data come from
Where the data come from

Newly-funded NSF CAREER project will create 5-minute educational videos on how OOI data are collected for use in undergraduate classrooms.
Interested in forthcoming “data empathy” 5-minute educational videos on how OOI data are collected?

Considering trying out (or adapting) this Blob Data Lab in your own teaching?

Find me after or email me! palevsky@bc.edu

*Full details & Blob Data Lab materials*