Data Lab #3: The Blob

Part 2: Combining data from multiple sources

In Part 2 of this lab, you will create two different time series of ocean temperature anomalies at Ocean Station Papa in the northeast Pacific Ocean. These will use the two datasets you acquired and extracted in Part 1: one based on the Ocean Observatories Initiative mooring data and one from the satellite sea surface temperature data. These data provide information about the timing and strength of the marine heatwave known at The Blob.

In addition to providing scientific information about The Blob that you will interpret in your writeup, a key skill goal for Part 2 of this lab is for you to gain practice combining data from multiple data sources.

In this lab, we will be interpreting ocean temperature data from three different sources:

1) Data collected from a sensor deployed by the Ocean Observatories Initiative
2) Satellite observations of sea surface temperature compiled by the National Oceanic and Atmospheric Administration
3) A gridded interpolation of a compilation of many data points, known as the World Ocean Atlas (similar to the gridded data in the LDEO ocean pCO_2 dataset we used in Data Lab #2)

In Part 1, you began working with data from the first two of these sources. In Part 2, you will bring all three of these sources together.

In order to simplify the lab, I have compiled the World Ocean Atlas data (the one additional dataset you hadn't worked with yet) and written the code to extract it from the original netCDF files into a MATLAB [structure array](#) for you to work with.

To get the World Ocean Atlas data, you will need to download the files from: https://tinyurl.com/NEPacificWOAdata. As in Part 1, you will need to put the folder with these files them in your MATLAB path to work with them – but don't put them inside your GitHub repository because they take up too much space!

Writing the MATLAB code to bring these three datasets together:

- Starter code with step-by-step directions for Part 2 is in GitHub classroom at: https://classroom.github.com/a/r3D4xR6Z
- The analysis in Part 2 will build on the analysis from Part 1, so you will need both repositories in your path to complete the analysis.
- As for Part 1 of this lab, even if you are working with a partner, you will each have your own GitHub repository and write the code yourself.

Questions to consider in interpreting your analysis for the lab writeup:

- What are the similarities/differences between the OOI-based and satellite-based sources ocean temperature anomaly data? Are there reasons you might expect differences based on how each type of data was collected?

- What are the strengths and weaknesses of each of these data types in enabling analysis of marine heatwaves such as The Blob? Based on this, why might we want to use more than one source of data to make the same type of calculation in addressing research questions about The Blob or similar phenomena?

Extension options:

For this lab you are only required to complete one extension (though you are of course welcome to do more if you desire).

Reading scientific papers: In order to understand the background for this lab and write the introduction, you will all need to explain the phenomenon of ocean heatwaves and the Blob, but as an extension you can gain a deeper understanding of this background by reading on or more relevant papers from the literature and including a brief summary of the paper's findings and relevance to your analysis.

Good places to start from to find relevant papers include the links provided in the Part 1 description and slides (note that the science news articles are based on original scientific papers that you could find and read), but you're welcome to read any relevant paper based on your interests, and I'm happy to help you if you have trouble finding a paper of interest.

Critical consumption of publicly-available data: For your methods section, you will describe the sources of all the types of data used here, but as an extension you can do more extensive background reading on the OOI and/or satellite SST data sources used in this lab. This would comprise reviewing online documentation to understand the methods behind how the data were originally collected, as well as how they were processed prior to being provided publicly for analysis.

Visualizing data: In Part 1, you made one figure showing a regional map of the satellite SST data. Make a series of SSTanom maps from different times within the OOI mooring time series record to show how the Ocean Station Papa record fits into a broader regional context, and show the times you selected on the time series plot.