A Jupyter server for the Oregon State University Climate Science Program

In fall 2016, the College of Earth, Ocean, and Atmospheric Sciences (CEOAS) at Oregon State University (OSU) purchased, with funding from Unidata, a Jupyter Notebook server with a primary goal of enabling meaningful climate/geospatial data experiences for undergraduate students. Secondary goals include providing a usable data analysis and visualization platform for the broader CEOAS community, as well as supporting student development of data-serving interfaces.

Using funds from Unidata, we purchased:

- PowerEdge R730 Server with two 14-core Intel Xeon E5-2690 processors (v4 2.6 GHz, 35M Cache), 512 GB system memory (16 x32GB RDIMM), and 24 TB local disk space (6 x 4TB, 7.2K RPM SATA 6Gbps)
- Dell Storage MD1400 external disk array with 72 TB storage (12 x 6TB 7.2K RPM NLSAS 12Gbps)

The server and storage array were delivered in Fall 2016 and operational for the last few weeks of the term. They have been used in two classes so far: ATS 301 Climate Data Analysis (Fall 2016) and ATS 421/521 Climate Modeling (Spring 2017), both taught by Dr. Karen Shell. These classes are part of the relatively new (< 2 years) Climate Science Option of CEOAS' Earth Sciences B.S.

A primary goal of the Climate Science program is to develop data-literate undergrads, who are comfortable working with data and using key tools. Most incoming students have no prior experience with programming. We try to “ease” them into programming, and they build computer skills over multiple classes. Some may later use these skills in their senior research projects. We have standardized on Python as the programming language, with Jupyter Notebook as the interface. The Jupyter Notebook server provides a streamlined system that removes barriers to early learning while also scaling with the needs of students throughout their undergraduate careers.

ATS 301 (Climate Data Analysis) is a sophomore/junior class, first offered in Fall 2016. While we were originally intending for the Jupyter server to be operational before the start of the term, the system came on line two-thirds of the way through the term, unintentionally demonstrating its usefulness.

Python has many advantages. It is free, and there is plenty of documentation available, including an increasing amount of atmospheric science and climate-focused documentation. Jupyter Notebook provides a nice, streamlined, web-based interface for student to run code. They complete each assignment within a single Jupyter Notebook, with code and plots inline (see example below). Unfortunately, while students can install Python for free on their own

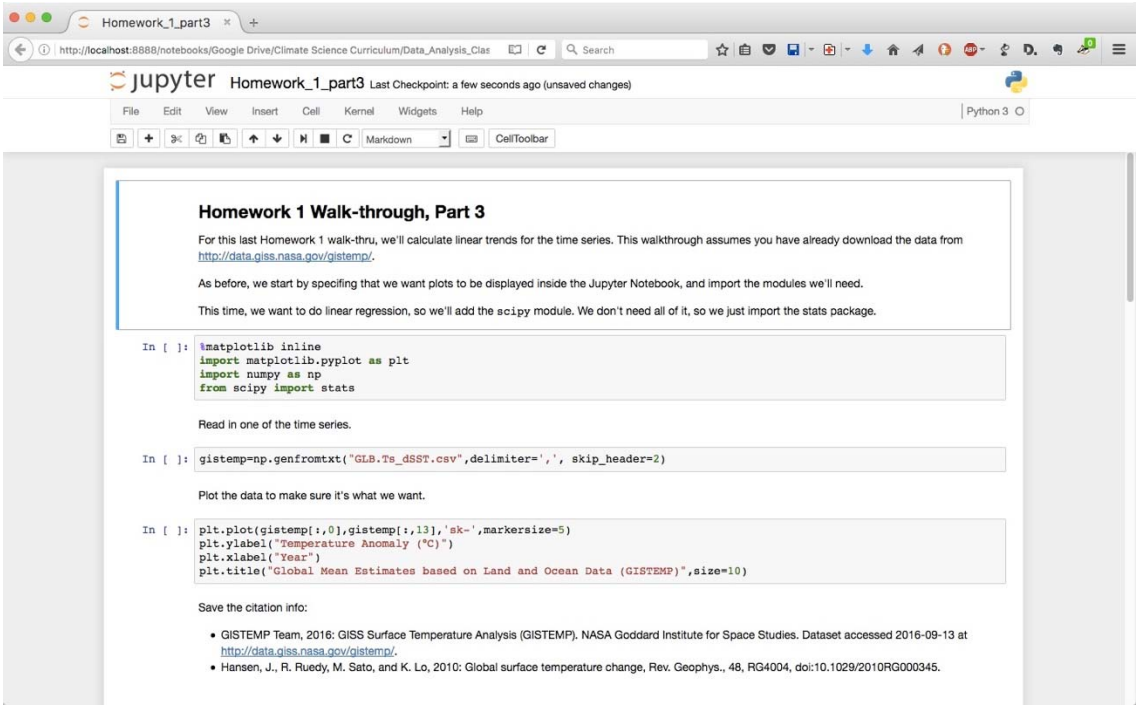
computers, the different versions, combinations of packages, and operating systems make standardization difficult. A notebook that worked on one computer might not on another. Since students were often working on their own computers, they had to move files between computers (e.g., email, USB keys, scp) to share their code or submit assignments. Another problem we encountered was that some of the large datasets actually crashed students' laptops.

Our new Jupyter server fixed these issues. Students just need a web browser (and a VPN connection if off campus) to log in. Everyone is using the same environment. To submit assignments, student simply place their final version in the homework folder; they can easily look at what other students have done; and the instructor can easily troubleshoot notebooks remotely. Additionally, the server is significantly faster. Students appreciated these advantages, such that all of them switched over to the server within a week of its coming online.

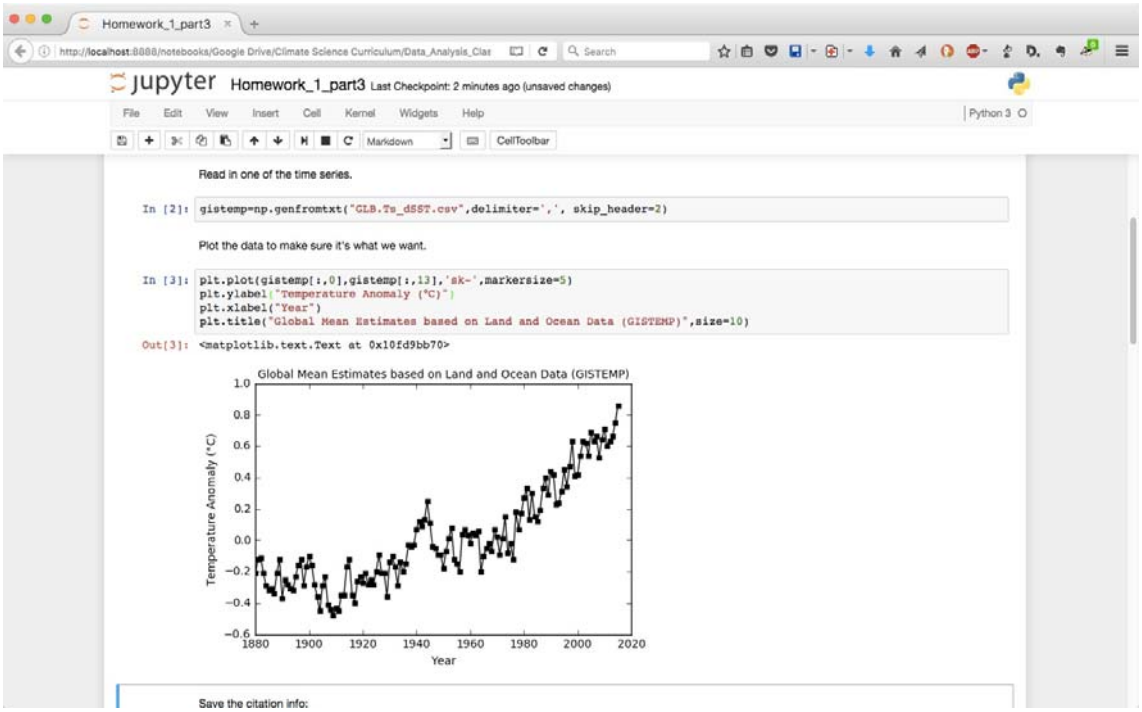
ATS 421/521, a senior and graduate-level climate modeling class, used the server from the very beginning, eliminating the need for installing and configuring Jupyter and Python on various computers. There are still some workflow issues (for example, moving files around using Jupyter requires a couple extra steps, and directory permissions need to be carefully set), but, overall, the use of the Jupyter/Python server has greatly reduced barriers which, in previous classes, have detracted from the science learning objectives, particularly for students who lack a programming background.

These experiences were shared at the 2016 AMS Annual Meeting: "Lessons Learned: Using Python and Jupyter Notebook to Teach Undergraduate Climate Data Analysis." The College published a feature story on new classroom technologies anticipating the purchase of the server < http://ceoas.oregonstate.edu/features/gaze/>.

Students in Climate Data Analysis class. Note the right student's computer monitor showing the Jupyter interface.



Example Jupyter Notebook demonstrating how to read in data from a file and create a plot.



Plot generated when the notebook is evaluated.