

Introducing Students to Scientific Python for Atmospheric Science

Sean Arms, Julien Chastang, Maxwell Grover, Jon Thielen,
Matthew Wilson, and Douglas Dirks

Python Workshop

What: Roughly 140 students gathered for a hands-on course focused on providing an introduction to Python for the atmospheric sciences.

When: 12 January 2020

Where: Boston, Massachusetts

<https://doi.org/10.1175/BAMS-D-20-0069.1>

Corresponding author: Sean Arms, sarms@ucar.edu

In final form 24 April 2020

©2020 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

AFFILIATIONS: **Arms, Chastang, and Dirks**—Unidata, UCAR, Boulder, Colorado; **Grover**—University of Illinois at Urbana–Champaign, Champaign, Illinois; **Thielen**—Iowa State University, Ames, Iowa; **Wilson**—University of Nebraska–Lincoln, Lincoln, Nebraska

Within the geoscience community, enthusiasm for the scientific Python ecosystem continues to grow (Nagpal and Gabrani 2019). As software tools created with Python continue to emerge and mature, researchers and educators are able to leverage their knowledge of the language to tackle problems from processing Earth observation data (Raspaud et al. 2018) to weather analysis (Blumberg et al. 2017). The success and growth of the American Meteorological Society's Symposium on Advances in Modeling and Analysis Using Python—which debuted at the January 2010 AMS Annual Meeting (Lin 2012)—is a good indicator of the continuing adoption of Python tools in the AMS community.

Enthusiasm for the Python scientific ecosystem was especially apparent among students at the 2020 AMS Annual Meeting, with undergraduate and graduate students playing significant roles in organizing and conducting Python-focused events. Of special interest was a large workshop conducted as part of the 19th Annual AMS Student Conference, which brought together roughly 140 students for a 90-min hands-on course focused on providing an introduction to Python for the atmospheric sciences (Fig. 1).

One of us (Grover), a graduate student at the University of Illinois at Urbana–Champaign, originally proposed the idea of a hands-on workshop to the AMS Student Conference organizing committee in May 2019. At the time, he was a summer intern at the Unidata Program Center at the University Corporation for Atmospheric Research (UCAR) in Boulder, Colorado. Because it is integral to Unidata's mission to help students build skill in accessing and using geoscience data, and because Unidata staff members regularly lead workshops on a variety of topics including Python, the original idea was for members of Unidata's Python development team to teach during the Student Conference session. As planning progressed, however, the idea arose that students could serve as instructors for the session, with Unidata staff providing organizational and technical support. It helped that the three students who volunteered to teach (Grover, Thielen, and Wilson) had all been Unidata summer interns and had good working relationships with the Unidata developers involved in the project (Arms and Chastang).

With the assistance of the Student Conference organizers, a no-cost workshop session was scheduled for the afternoon of Sunday, 12 January 2020, and advertised with the following synopsis:

Are you looking to make the transition to Python but unsure of where to start? Do you already know Python but want to see atmospheric science specific applications? Are you looking for data? If so, then please join us for a hands-on Python Users' Workshop where beginners and experts alike will learn skills that will enhance your ability to find, analyze, and explore data, all while doing so in the cloud. Specialized local software installations are not necessary. All



Fig. 1. Roughly 140 students participated in the Python Workshop, held as part of the 19th Annual AMS Student Convention at the 2020 AMS Annual Meeting.

you need is a laptop or tablet (a keyboard may be helpful) and a GitHub login to participate. In-person follow-up help will be available all week, as well as remotely once our time together as a community in Boston comes to an end.

The response to this was tremendous; well over 100 students registered to participate in advance of the workshop, and roughly 140 ended up participating.

Workshop format

Because we hoped the workshop would be a hands-on experience, where attendees were able to follow along with what the instructors were doing in real time, we wanted to avoid any stumbling blocks relating to computers or software configuration. Our solution was to preconfigure all of the software used during the workshop and provide it to students via cloud-hosted Jupyter notebooks (Zonca and Sinkovits 2018; Wilkins-Diehr et al. 2016). In practice, students were asked to connect to a Jupyter notebook environment via a “JupyterHub” provided by Unidata as part of the Unidata Science Gateway project, which relies on resources from the National Science Foundation (NSF) Jetstream project (Stewart et al. 2015). Once logged in to the JupyterHub, participants could access preconfigured PyAOS (Python for the Atmospheric and Oceanic Sciences) computing environments that allowed them to work with the same code examples demonstrated by the workshop presenters, potentially experimenting with the presented material by making code alterations. This technique allowed the instructors to better illustrate the concepts they were teaching and improved the students’ ability to absorb the material.

The three workshop sections dealt with topics ranging from a general introduction to relevant scientific Python techniques to tackling specific tasks in atmospheric science using (among other tools) Unidata’s MetPy library (May et al. 2016) for data analysis and Siphon (May et al. 2015) for remote data access. Jupyter notebooks used during the session are publicly available on GitHub (Chastang et al. 2020).

First up was Matthew Wilson, currently a graduate student at the University of Nebraska–Lincoln and a Unidata summer intern in 2017. He presented “Numpy, Matplotlib, and Cartopy Basics,” including a discussion of arrays, a cornerstone of scientific programming, and the Python library for array-oriented numerical computation, NumPy (Fig. 2). To demonstrate, Matthew led students through an exercise plotting tornado observations on a map with the Matplotlib and Cartopy libraries.

Next was Jon Thielen, currently a graduate student at Iowa State University and a Unidata summer intern in 2018. His topic was “Isentropic Analysis with xarray and MetPy,” and his examples helped students understand how to use the Python xarray library to tackle a geoscience-specific application using multidimensional arrays. He also covered unit manipulation and remote data access, and concluded with a demonstration of isentropic analysis.



Fig. 2. Matthew Wilson, currently a graduate student at the University of Nebraska–Lincoln, presented “Numpy, Matplotlib, and Cartopy Basics.”

Third was Max Grover, currently a graduate student at the University of Illinois at Urbana-Champaign and a Unidata summer intern in 2019. His “Declarative Plotting Tutorial” covered remote data access followed by the introduction of a GEMPAK-like plotting interface for plotting of satellite, model forecast data, and surface observation data using the MetPy library.

Participant reactions

Student participants were not asked to provide formal feedback on the session, but several did provide comments after the workshop. Katurah McCants, an undergraduate meteorology major at Metro State University of Denver, said “I really enjoyed attending the workshop this year and hope to take one of the longer workshops at next year’s conference. Unidata creates an environment where individuals who are novice to programming, like myself, feel at ease when learning difficult new concepts. The workshop was a great crash course in using Python in a variety of applications for meteorologists.”

“My experience with the Unidata workshop was extremely rewarding! I cannot say enough positive outcomes about attending the AMS Unidata Python Workshop,” said Ayesha Wilkinson, currently pursuing a master’s degree in Atmospheric Sciences and Meteorology at Howard University. “During the workshop, the Unidata team preloaded the workshop onto a workbook so it was extremely easy to pay attention and follow. They gave us their website which included: tips, tutorials, videos, and Q&A sections. After the workshop, I was able to go to their Unidata booth and ask more questions! I was so thankful.”

“The short course was well-organized and given by a group of experts, who were well equipped to handle questions from attendees,” said Brian Carroll, an Atmospheric Physics PhD candidate at the University of Maryland, Baltimore County (UMBC). “The multitude of versions, distributions, packages, and functions available to Python users can be overwhelming for beginners, and this workshop started by clarifying these first steps then moved on to obtaining and plotting datasets. With little prior knowledge, I learned how to access and plot near-real-time and archived meteorological datasets (e.g. satellite imagery and soundings) in Python.”

Mikael Peter Hiestand, a Ph.D. candidate at The Pennsylvania State University, commented that the course “did a good job of covering Python basics for those new to the language and then provided a good walk through of Python packages for meteorological applications. The only way it could have been improved would have been to have the code ready for student download at the end of the workshop.”

“Unidata’s AMS Python Workshop helped me to realize how easy and fast a Jupyter notebook can be in helping you to visualize meteorological data,” said Kyra Bryant, a doctoral student and graduate research assistant at Tennessee State University. “I knew I could display results using Python, but I didn’t realize how simple it could be! In only an hour and a half, I walked out equipped with knowledge of a powerful meteorological application tool.”

Looking ahead

Beyond the explicit benefits for the student community of holding this workshop, a secondary motivation was to explore the scalability of the Science Gateway JupyterHub configuration. While similar configurations of Jetstream resources had been used successfully for smaller workshops, this was the first “live” test of the Unidata gateway supporting more than 100 participants simultaneously. The smooth running of this workshop gives us confidence that the Unidata Science Gateway has the potential to support numerous students without technical issues.

Given the general level of student enthusiasm exhibited before, during, and after this workshop, Unidata hopes to be able to hold a similar event next year. In initial conversations, the 2021 AMS Student Conference organizing team has been supportive of the idea. We hope to

once again engage students as instructors for such a session, and we are interested in hearing from anyone who would like to participate.

Acknowledgments. We thank the 2020 AMS Student Conference organizers for their support in making this workshop possible. We also thank Jeremy Fischer and the Jetstream Team at Indiana University for their assistance with Unidata's projects in the Jetstream cloud. The Unidata Program is funded primarily by the National Science Foundation under Award 1901712.

References

- Blumberg, W. G., K. T. Halbert, T. A. Supinie, P. T. Marsh, R. L. Thompson, and J. A. Hart, 2017: SHARPy: An open-source sounding analysis toolkit for the atmospheric sciences. *Bull. Amer. Meteor. Soc.*, **98**, 1625–1636, <https://doi.org/10.1175/BAMS-D-15-00309.1>.
- Chastang, J., S. Arms, M. Grover, J. Thielen, and M. Wilson, 2020: Unidata AMS 2020 Student Conference PyAOS workshop. GitHub repository, <https://github.com/Unidata/pyaos-ams-2020>.
- Lin, J. W., 2012: Why Python is the next wave in Earth sciences computing. *Bull. Amer. Meteor. Soc.*, **93**, 1823–1824, <https://doi.org/10.1175/BAMS-D-12-00148.1>.
- May, R., S. Arms, J. Leeman, and J. Chastang, 2015: Siphon. Unidata, UCAR/NCAR, <https://doi.org/10.5065/D6CN72NW>.
- , ———, P. Marsh, E. Bruning, and J. Leeman, 2016: MetPy. Unidata, UCAR/NCAR, <https://doi.org/10.5065/D6WW7G29>.
- Nagpal, A., and G. Gabrani, 2019: Python for data analytics, scientific and technical applications. *2019 Amity Int. Conf. on Artificial Intelligence*, Dubai, United Arab Emirates, IEEE, 140–145, <https://doi.org/10.1109/AICAL.2019.8701341>.
- Raspauld, M., and Coauthors, 2018: PyTroll: An open-source, community-driven Python framework to process earth observation satellite data. *Bull. Amer. Meteor. Soc.*, **99**, 1329–1336, <https://doi.org/10.1175/BAMS-D-17-0277.1>.
- Stewart, C. A., and Coauthors, 2015: Jetstream: A self-provisioned, scalable science and engineering cloud environment. *Proc. 2015 XSEDE Conf.: Scientific Advancements Enabled by Enhanced Cyberinfrastructure (XSEDE '15)*, St. Louis, MO, Association for Computing Machinery, Article 29, <https://doi.org/10.1145/2792745.2792774>.
- Wilkins-Diehr, N., S. Sanielevici, J. Alameda, J. Cazes, L. Crosby, M. Pierce, and R. Roskies, 2016: An overview of the XSEDE extended collaborative support program. *ISUM 2015: High Performance Computer Applications*, I. Gitler and J. Klapp, Eds., Communications in Computer and Information Science Series, Vol. 595, Springer, 3–13, https://doi.org/10.1007/978-3-319-32243-8_1.
- Zonca, A., and R. S. Sinkovits, 2018: Deploying Jupyter Notebooks at scale on XSEDE resources for science Gateways and workshops. *PEARC '18: Proc. Practice and Experience on Advanced Research Computing, Pittsburgh, PA, Association for Computing Machinery, Article 8*, <https://doi.org/10.1145/3219104.3219122>.