

# Breakout Workshop: Scientific computing and the Open Source software revolution: Opportunities for ocean colour science

Joaquín E. Chaves (NASA GSFC/SSAI), Erdem M. Karaköylü (NASA GSFC/SAIC), Joel P. Scott (NASA GSFC/SAIC)

Tuesday, April 9

# Agenda:



- 2:00--2:05: Intro, motivation and goals (co-chair: **J. Chaves**, NASA GSFC/SSAI):Open Source software revolution: Resources and opportunities for ocean color science research, data distribution, education and mentoring
- 2:05--2:25 Python for multi-year GOCI ocean color products analysis: sharing the advantages and issues (Dr. Myung-Sook Park, KIOST)
- 2:25--2:30 QA and interactive discussion.
- 2:30--2:50 Leveraging Python & Jupyter Notebook to validate satellite Ocean Color retrievals via SeaBASS data (co-chair: **J. Scott**, NASA GSFC/SAIC)
- 2:50--2:55 **QA and interactive discussion.**
- 2:55--3:00 Coffee break
- 3:00–3:20 Use of Python and SeaDAS for Ocean Color Processing, Analysis and Visualization (**Bruce Monger**, Cornell University)
- 3:20--3:25 QA and interactive discussion.
- 3:25--3:45 The importance of open science principles for expanding and diversifying the user base of marine satellite data (**Dr. Hayley Evers-King**, Plymouth Marine Lab)
- 3:45–3:50 **QA and interactive discussion.**
- 3:50--4:10 Probabilistic Programming Workflow in Python (co-chair **E. Karaköylü**, NASA GSFC/SAIC)
- 4:10--4:15 **QA and interactive discussion.**
- 4:15--4:30 Summary and group discussion (Co-chairs, others)

# Agenda:



## Key questions:

- What recent shifts among Ocean Color community have there been towards open source tools?
- How and why is the community using these tools?
- How are teaching, mentoring, and research deliverables being changed by open source?
- Are these tools promoting reproducibility, code preservation, accessibility, diversity & inclusion?



# What is Open Source software?

"Open source software is software with source code that anyone can inspect, modify, and enhance"



Breakout Workshop: Scientific computing and the Open Source software revolution:

nttps://opensource.com/resource es/what-open-source



## Some early history:

- Historically (1950-60s) software was free, profits were on hardware.
- AT&T's commercialization of UNIX ensues the OS movement:
- Richard Stallman, then at MIT is credited as the founder of the open source software movement in the early 1970s.
- GNU, BSD & MIT licenses for UNIX-like systems
- Linux



https://blog.timescale.com/opensource-demise-of-proprietarysoftware-a49f73f54165







python

javascript

2020



#### Projections of future traffic for major programming languages

Breakout Workshop: Scientific computing and the Open Source https://stackoverflow.blog/2017/09/06/incrediblesoftware revolution:





Breakout Workshop: Scientific computing and the Open Source software revolution:

https://github.com/elegant-scipy/elegantscipy/blob/master/figures/python-inastronomy.png



Ø		Help	Donate	Login	Register	
Find, ir w	istall and publish Pyth ith the Python Packag	on packa e Index	iges			
Search	projects	Q.				
	Or <u>browse projects</u>					
175,025 pro	ojects 1,276,989 releases 1,821,889 f	les 318,182 u	sers			
Package Index	The Python Package Index (PyPI) is a reposit programming language. PyPI helps you find and install software developed about installing packages. Package authors use PyPI to distribute their softwa for PyPI.	ory of software fo and shared by the Py re. Learn how to pac	r the Pytho thon comm kage your Py	n unity. Learr thon code	1	







## Open Source and diversity in scientific computing:





ABOUT V PROJECTS V PROGRAMS V SPONSOR V BLOG

DONATE



#### NumFOCUS Diversity Statement

NumFOCUS welcomes and encourages participation in our community by people of all backgrounds and identities. We are committed to promoting and sustaining a culture that values mutual respect, tolerance, and learning, and we work together as a community to help each other live out these values.

We have created this diversity statement because we believe that a diverse community is stronger, more vibrant, and produces better software and better science. A diverse community where people treat each other with respect has more potential contributors, more sources for ideas, and fewer shared assumptions that might hinder development or research.

Although we have phrased the formal diversity statement generically to make it all-inclusive, we recognize that there are specific identities that are impacted by systemic discrimination and marginalization. We welcome all people to participate in the NumFOCUS community regardless of their identity or background.

Breakout Workshop: Scientific computing and the Open Source software revolution:

https://numfocus.org/co de-of-conduct



## ...Let's hear it from the speakers



Back up slides for discussion at the end...



### PYPL PopularitY of Programming Language

https://www.r-bloggers.com/pypl-language-rankings-python-ranks-1-r-at-7-in-popularity/

Repositories as research deliverables and complement to traditional papers:



## The Journal of Open Source Software



The Journal of Open Source Software is an affiliate of the Open Source Inititative.





https://joss.theoj.org/











http://pyviz.org/

Objects from nearly every other plotting library can be used with Panel, including specific support for all those listed here plus anything that can generate HTML, PNG, or SVG. HoloViews also supports Plotly for 3D visualizations.

pillow



The EarthML Tutorial offers a general-purpose overview of the concepts and tools involved, and the Topics section shows examples of how these tools may be used to perform machine learning and related tasks in the Earth sciences, such as :



Please feel free to report issues or contribute code .

Installation

Step 1: Install a Miniconda (or Anaconda ) environment 🦷

Any Linux or Mac OS X computer with a modern web browser (preferably Googie Chrome) should be suitable. Windows machines can also run most of the functionality, but the EarthML team is currently (11/2018) investigating some Windows-specific issues. 5GBB of RAM is required for some of the examples, but most will run fine in 4GB.



Space Flight Center



#### http://earthml.pyviz.org/index.html

Exponential growth in Machine Learning and Artificial Intelligence has been driven by open source:





# **O** PyTorch

## Der Link



https://link.springer.com/book/10.1007/978-3-319-65633-5



#### Earth Observation Open Science and Innovation pp 165-218 | Cite as

#### Machine Learning Applications for Earth Observation

Authors

Authors and affiliations

David J. Lary 🖂 , Gebreab K. Zewdie, Xun Liu, Daji Wu, Estelle Levetin, Rebecca J. Allee, Nabin Malakar, Annette Walker,

Hamse Mussa, Antonio Mannino, Dirk Aurin

Open Access Chapter First Online: 24 January 2018



Part of the ISSI Scientific Report Series book series (ISSI, volume 15)

#### Abstract

Machine learning has found many applications in remote sensing. These applications range from retrieval algorithms to bias correction, from code acceleration to detection of disease in crops, from classification of pelagic habitats to rock type classification. As a broad subfield of artificial intelligence, machine learning is concerned with algorithms and techniques that allow computers to "learn" by example. The major focus of machine learning is to extract information from data automatically by computational and statistical methods. Over the last decade there has been considerable progress in developing a machine learning methodology for a variety of Earth Science applications involving trace gases, retrievals, aerosol products, land surface products, vegetation indices, and most recently, ocean applications. In this chapter, we will review some examples of how machine learning has already been useful for remote sensing and some likely future applications. Open Source in education, mentoring, and knowledge distribution:



Repositories as research deliverables and complement to traditional papers:



	Pull requests Issue	es Marketplace I	Explore	. +
MATINE Ocean Optics UMaine MISC Lab O University of Maine, USA	http://misclab.umeoce	. 🖂 emmanuel.bos	s@maine	Report abu
Find a repository Type: All - La	nguage: All -			
Python package to unpack and calibrate binary of ACS	data from a WetLabs		Top language	ges I MATLAB 🔵 JavaScript
InLineAnalysis Process optical in-line data measured continuou Inline instrument acs vessel eco	sly on research vessel flowthrough lisst		bio-argo People	instrument 0 >
MATLAB ★ 2 4 MIT Updated on Nov 8, 20	18		This organiza members. Yo who's a part	tion has no public u must be a member to see of this organization.
FloatExplorer Explore Bio-Argo profiles with ease. flask dashboard bio-argo JavaScript dt GPL-3.0 Updated on Nov 1, 2018	3			
FloatProcessor Process Bio-Argo profiles with ease. processing calibration bio-argo • Python ★1   GPL-3.0 Updated on Oct 31,	2018	·		
and FOR				