Flood Inundation Mapping & Impacts:

Leveraging Google Earth Engine Python API and Spatial Analytics

Mirza Waleed

PhD Candidate @HKBU | GDE Earth Engine | Visiting Student @KAUST

June 2025

waleedgeo.com

Today's agenda

01 Flood Mapping- A brief recap

02 Improving Flood Mapping

03 Why choose Python for your Earth Engine work?

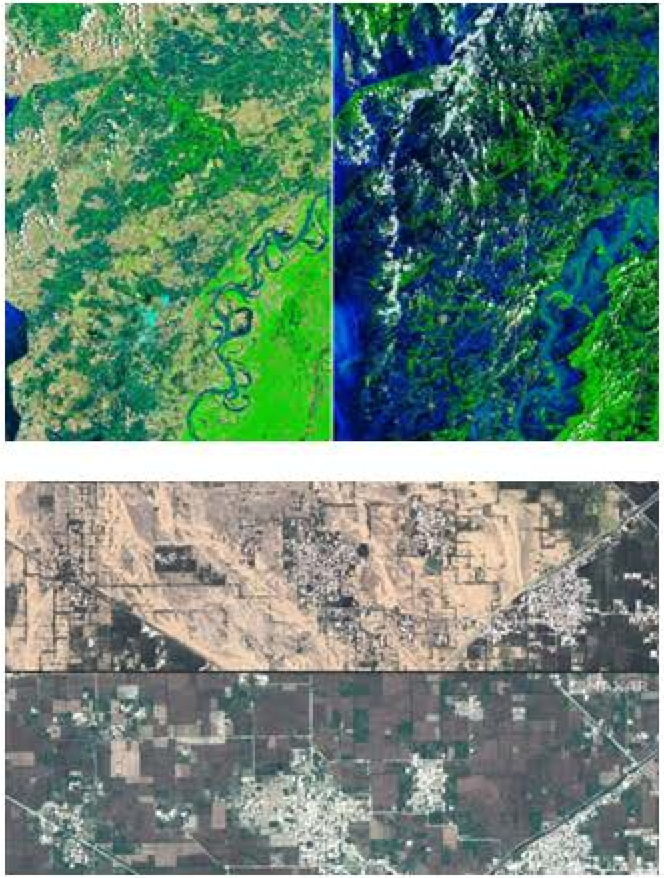
04 **Hands On**: Python - A brief introduction

05 **Hands On**: Earth Engine + Python + `geemap` for flood mapping in colab

Background: Scalable Flood Mapping

- **Global Crisis:** Flooding is the most common natural disaster globally, impacting millions and causing billions in damages annually. Climate change is exacerbating both the frequency and intensity of these events.

- **The Need:** We urgently need rapid, cost-effective, and scalable methods to map flood inundation over large areas to:
  - Aid humanitarian response.
  - Assess damage accurately.
  - Inform recovery efforts.



Satellite images released by Maxar Technologies shows farmland, top, after floodwaters arrived in Gudpur, Punjab province, and in April, bottom.

Google Earth Engine (GEE)? A Paradigm Shift

- **Server-Side Processing:** GEE performs computations on Google's massive infrastructure. This means you can analyze gigapixels of imagery in seconds or minutes, a task that would take days or weeks on a standard desktop.

- **Scalability:** Allows analysis from local basins to entire continents without overwhelming your local system.

My Research:

Addressing Real-World Flood Challenges

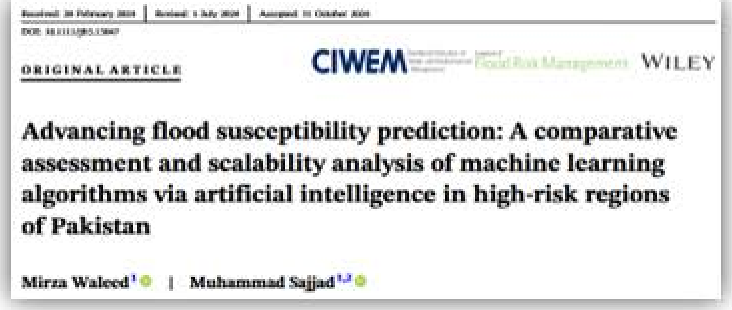
- **Publications:**
  - On the emergence of geospatial cloud-based platforms for disaster risk management: A global scientometric review of google earth engine applications
  - Advancing Flood Susceptibility Prediction: A Comprehensive Assessment of Machine Learning Algorithms for Disaster Risk Management
  - High-resolution flood susceptibility mapping and exposure assessment in Pakistan: An integrated artificial intelligence, machine learning and geospatial framework":

International Journal of Disaster Risk Reduction

ELSEVIER

On the emergence of geospatial cloud-based platforms for disaster risk management: A global scientometric review of google earth engine applications

Mirza Waleed, Muhammad Sajjad



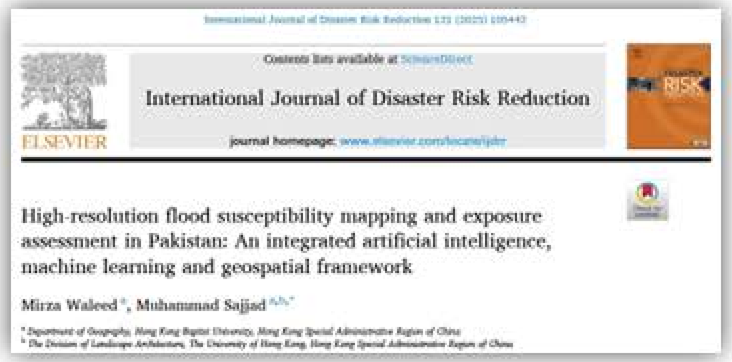
ORIGINAL ARTICLE

CIWEM

WILEY

Advancing flood susceptibility prediction: A comparative assessment and scalability analysis of machine learning algorithms via artificial intelligence in high-risk regions of Pakistan

Mirza Waleed | Muhammad Sajjad



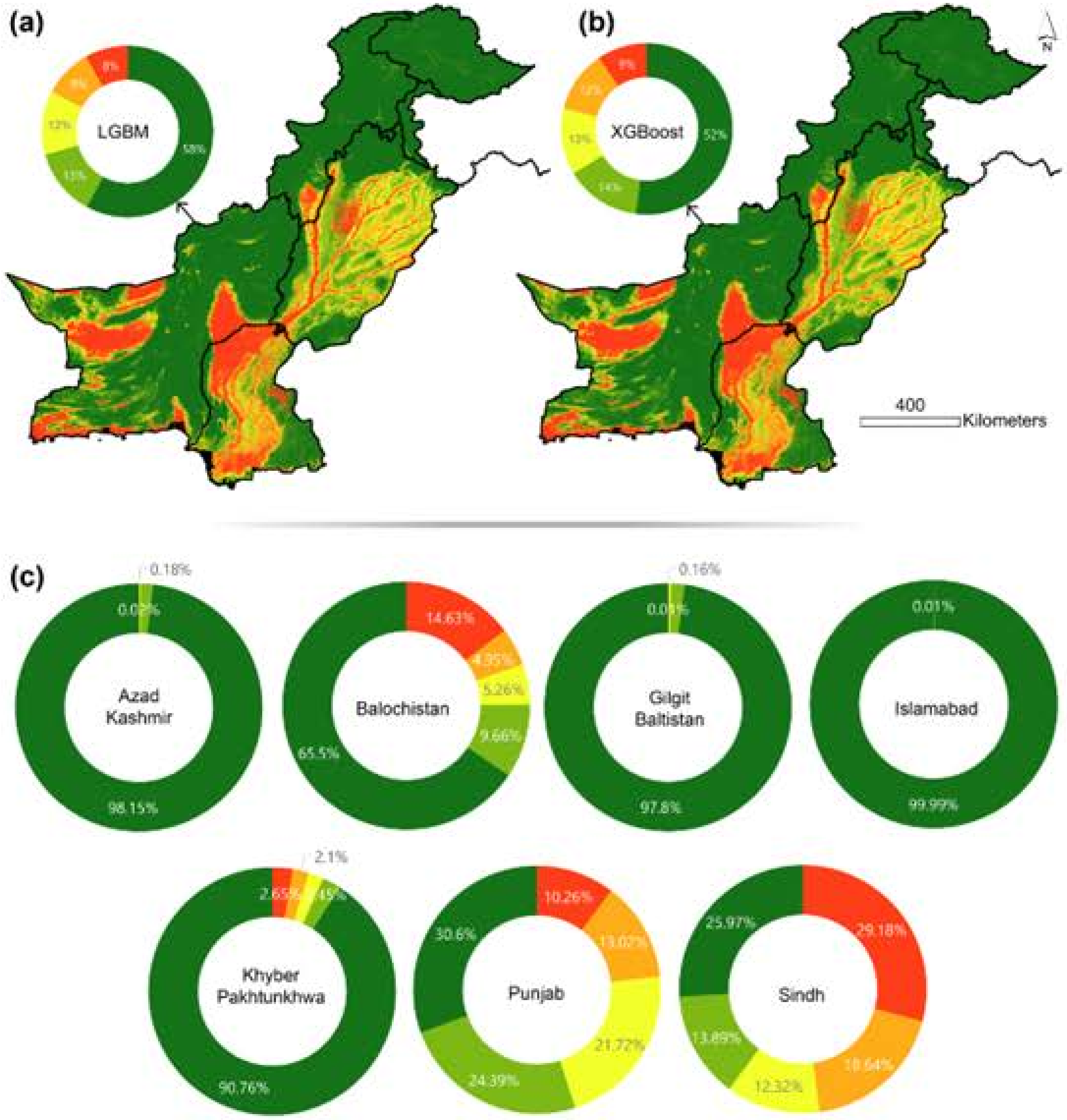
International Journal of Disaster Risk Reduction

ELSEVIER

High-resolution flood susceptibility mapping and exposure assessment in Pakistan: An integrated artificial intelligence, machine learning and geospatial framework

Mirza Waleed, Muhammad Sajjad

Flood Susceptibility in Pakistan



International Journal of Disaster Risk Reduction 121 (2025) 105442

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

High-resolution flood susceptibility mapping and exposure assessment in Pakistan: An integrated artificial intelligence, machine learning and geospatial framework

Mirza Waleed [a], Muhammad Sajjad [a,b,*]

[a] Department of Geography, Hong Kong Baptist University, Hong Kong Special Administrative Region of China
[b] The Division of Landscape Architecture, The University of Hong Kong, Hong Kong Special Administrative Region of China

Map Key

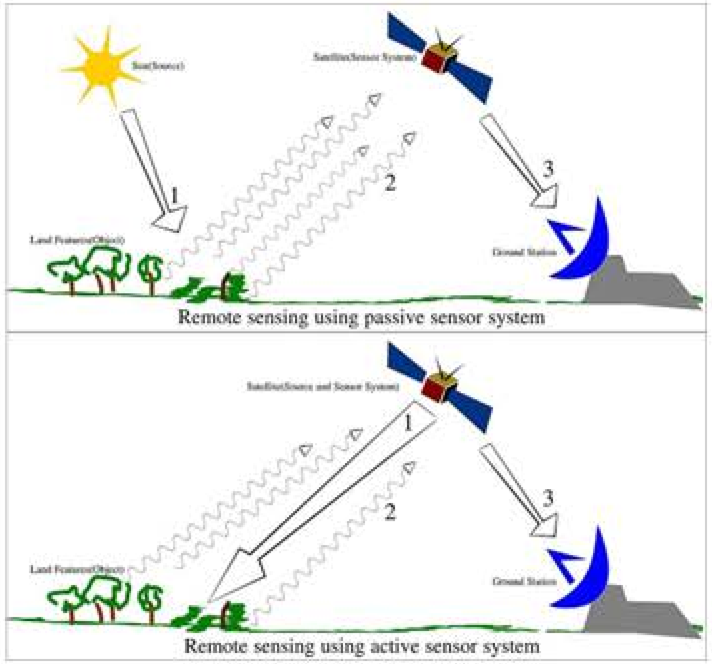
Flood Susceptibility: Very Low Low Moderate High Very High

- Panels (a) & (b): Flood susceptibility maps generated using LGBM and XGBoost models. The donut charts represent the percentage distribution of flood susceptibility classes across the entire study area.
- Panel (c): Provincial flood susceptibility distribution based on LGBM results. Each donut chart shows the proportion of areas under different susceptibility levels, highlighting that Sindh, Punjab, and Balochistan have the highest exposure to high and very high flood susceptibility zones.

Sentinel-1 SAR: The Eye through Clouds

- **SAR Principle:** Sentinel-1 operates on the Synthetic Aperture Radar (SAR) principle. Unlike optical sensors that rely on sunlight and are blocked by clouds or darkness, SAR actively sends out microwave signals and records the energy reflected.

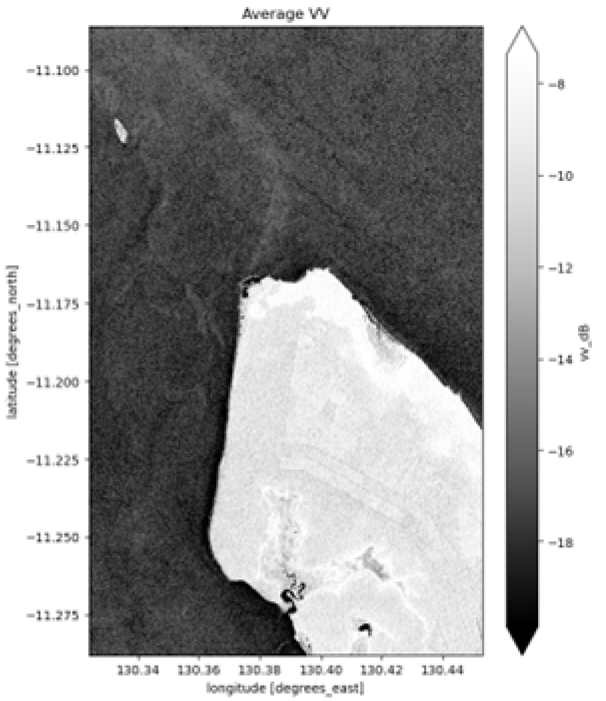
- **All-Weather, Day-and-Night:** This active sensing capability means Sentinel-1 can 'see through' clouds, smoke, and operate regardless of day or night. This is *critical* for flood mapping, as floods often occur during heavy rainfall and under persistent cloud cover.



Sentinel-1 SAR: The Eye through Clouds

- **How it Detects Water:** Smooth surfaces like calm water act like a mirror, reflecting the radar signal away from the sensor, resulting in very low backscatter (dark pixels). Rough surfaces (vegetation, buildings, dry land) scatter the signal in many directions, leading to higher backscatter (brighter pixels). This contrast is key to identifying inundated areas.

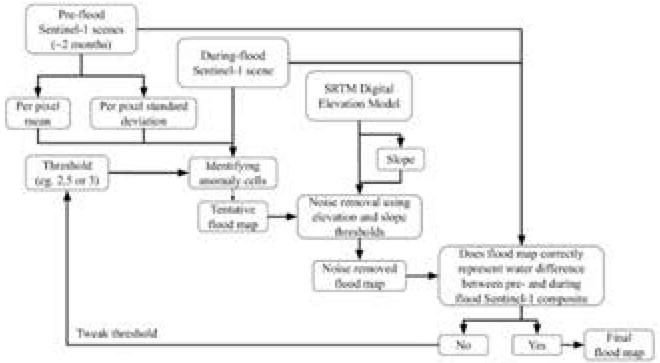
- **Polarizations:** Sentinel-1 provides different polarizations (VV, VH). VV (Vertical send, Vertical receive) is highly sensitive to surface roughness, while VH (Vertical send, Horizontal receive) can provide additional information, especially over vegetated areas.



Fundamentals of Flood Inundation Mapping

- **Core Principle: Change Detection:** The fundamental idea is to compare a "pre-flood" image with a "post-flood" image. We look for areas that were dry before and appear dark (low backscatter, indicating water) after the event.

- **Data Acquisition & Filtering:**
  - Identifying relevant Sentinel-1 GRD (Ground Range Detected) products.
  - Filtering by date (e.g., imagery from 2-3 weeks before the flood, and during/immediately after the flood peak).
  - Filtering by Area of Interest (AOI) and specific orbit *properties (e.g., 'IW' mode for wide swath).
  - Selecting the right polarization bands, typically 'VV' and 'VH'.

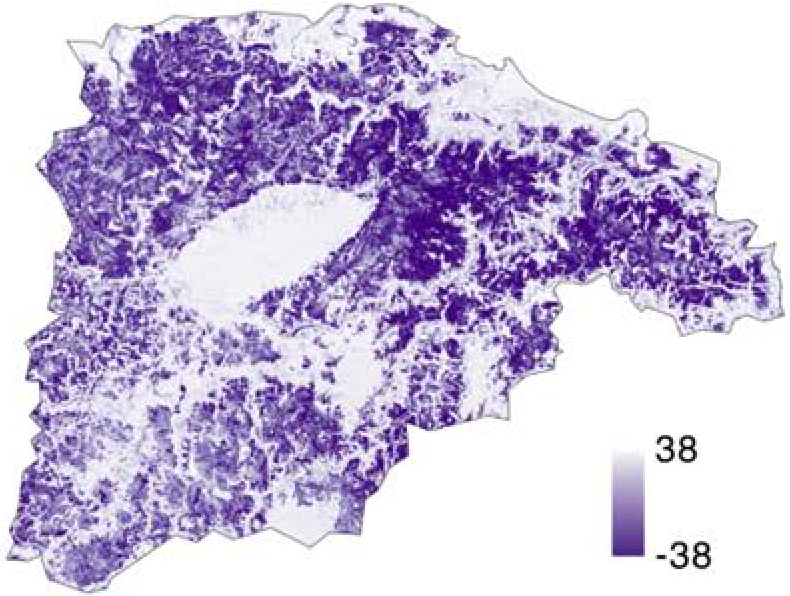


https://doi.org/10.1007/s11069-022-05428-2

Beyond Simple Thresholding: The Power of Z-Scores

- **Limitations of Simple Thresholding:** A single threshold value rarely works perfectly across diverse terrains, land cover types, and different flood conditions. What's 'dark' in one area might be normal terrain elsewhere due to shadowing or specific land cover. This leads to false positives/negatives.

- **Introducing Z-score:** The Z-score (or standard score) is a powerful statistical measure that indicates how many standard deviations an element is from the mean. In our context, it quantifies how anomalous a pixel's backscatter value is compared to its *own historical average*.



Calculation (Z-score)

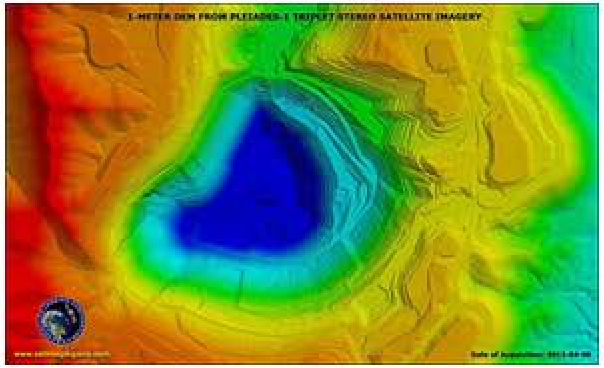
$$Z = (X_i - \mu) / \sigma$$

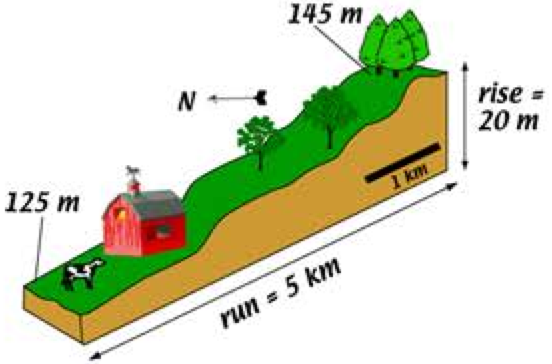
Where:

- X_i = post-flood backscatter value for a pixel.
- μ = mean historical backscatter for that *same pixel* over a defined pre-flood baseline period.
- σ = standard deviation of historical backscatter for that *same pixel*.

Integrating Ancillary Data

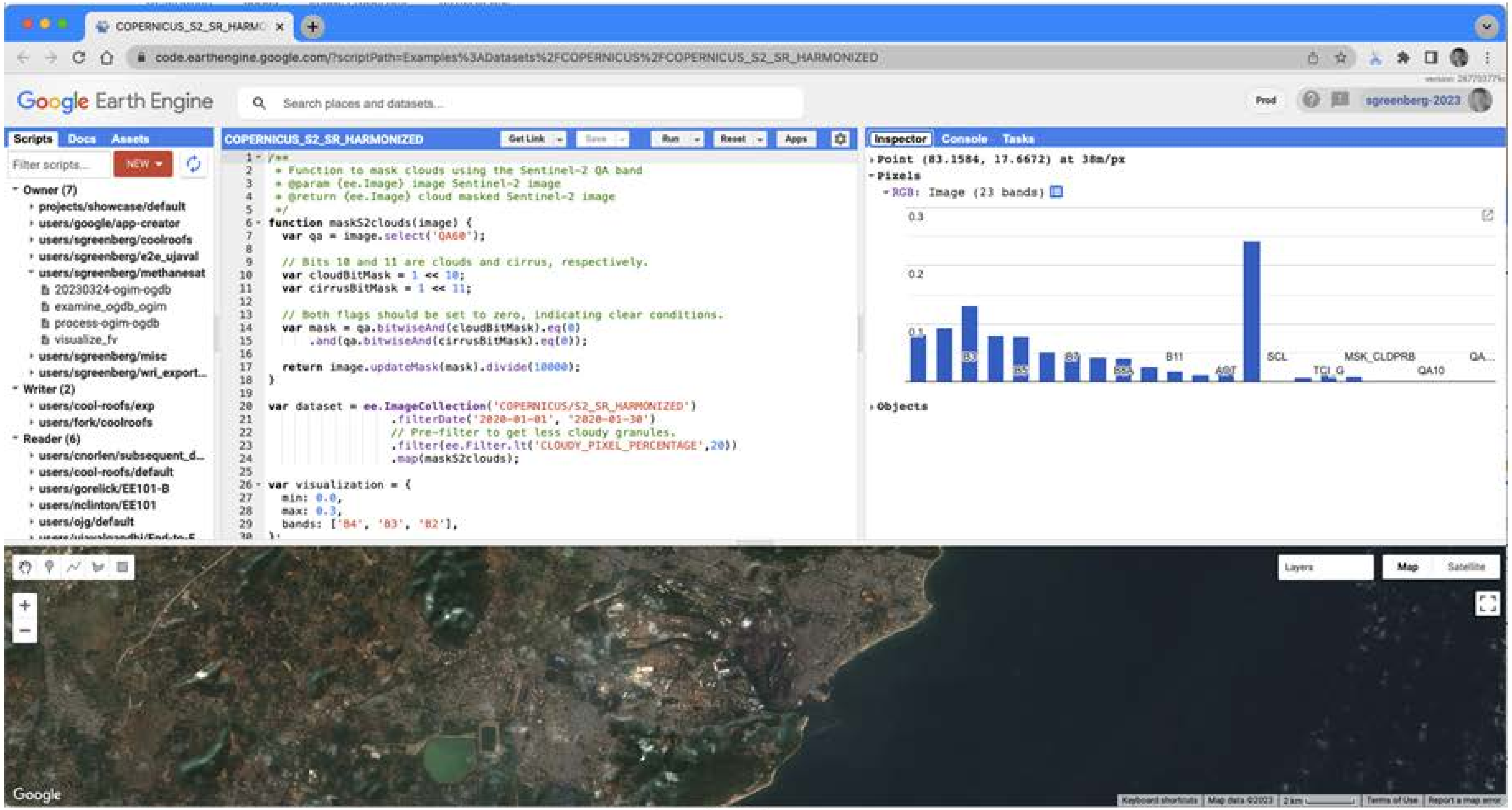
- **JRC Global Surface Water (GSW):**
  - Contains decades of water occurrence statistics. We use it to mask out permanent water bodies (rivers, lakes, oceans). This ensures we map *new* flood inundation, not just existing water.
- **Digital Elevation Models (DEMs):**
  - High-elevation areas (mountains, plateaus) are generally not prone to widespread inundation. We can use DEMs to mask out false positives detected at higher altitudes.
- **Slope:**
  - Steep slopes are typically not inundated globally. Radar shadows on steep terrain can sometimes be misclassified as water. Masking out steeply sloped areas (e.g., >10-15 degrees) helps remove these false positives.





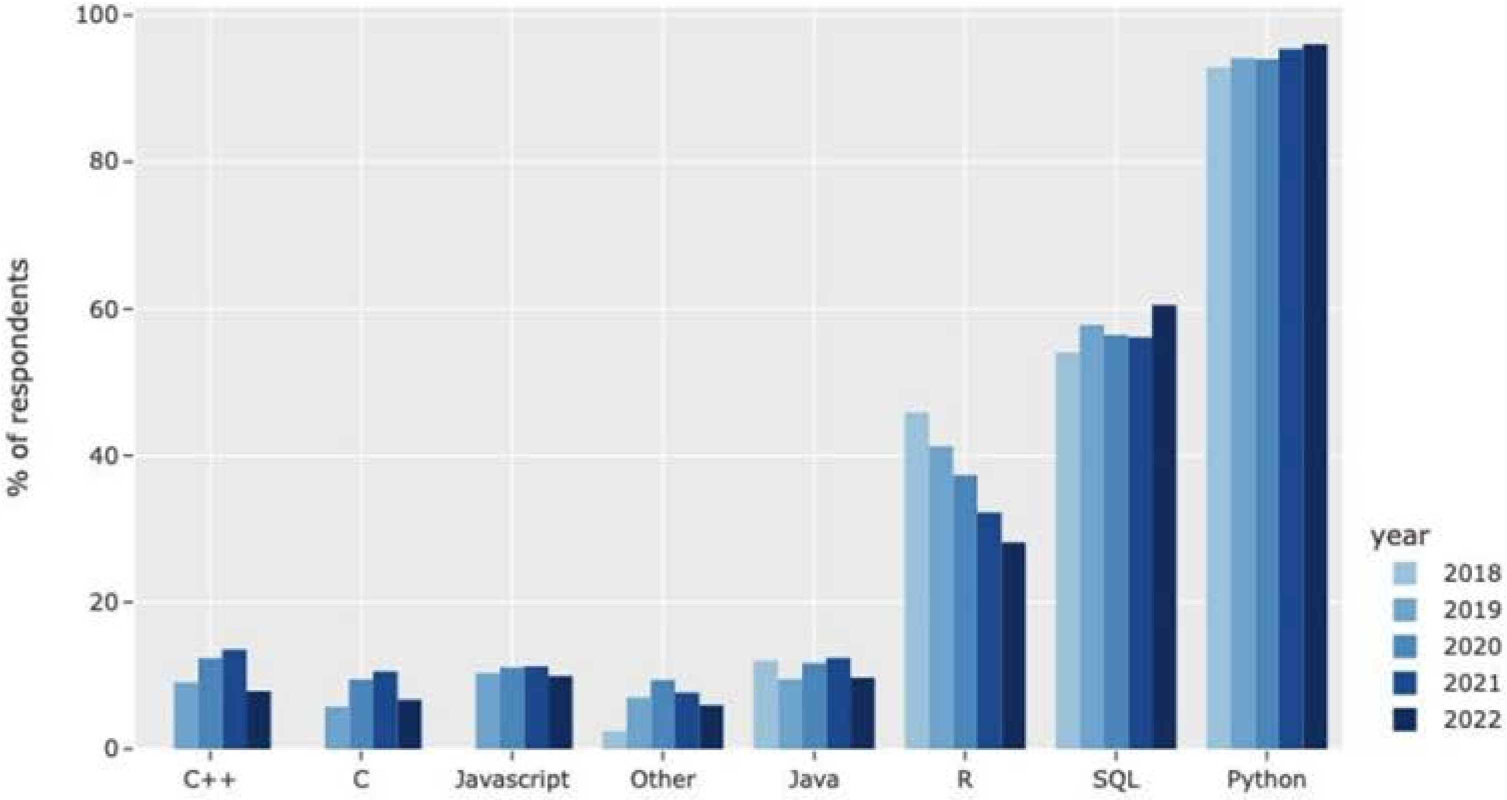
Python in Earth Engine

Integrations in the Earth Engine Code Editor



©waleedgeo.com

Python is the most popular programming language for Machine Learning and Data Science



https://www.kaggle.com/kaggle-survey-2022

A community of Machine Learning and Data …

scikit learn

SciPy

PyData

jupyter

pandas

TensorFlow

PyTorch

… is expanding into geospatial ML and data.

PANGEO

xarray

Zarr

COG

geemap

GeoPandas

Why choose Python + Earth Engine?

01 Automation

Run scripts without user intervention.

02 Libraries

Tensorflow
PyTorch
scikit-learn
Data Analysis
Data Visualization
Physics Libraries

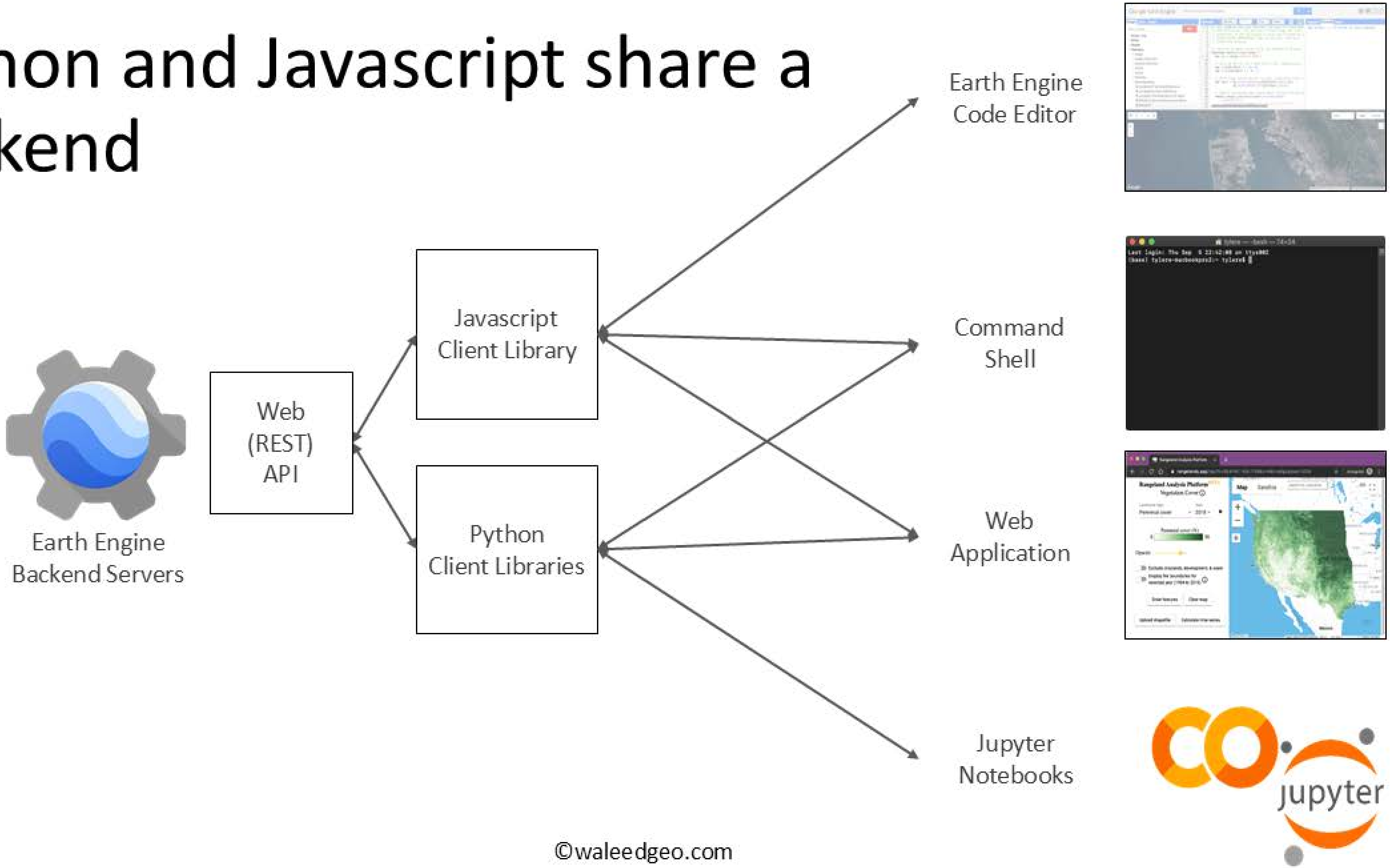
03 Cloud Services

BigQuery
Dataflow / Beam
Vertex AI
Google Cloud Storage

04 Dev Tools

Github
CICD Tools
Testing Frameworks

Python and Javascript share a backend



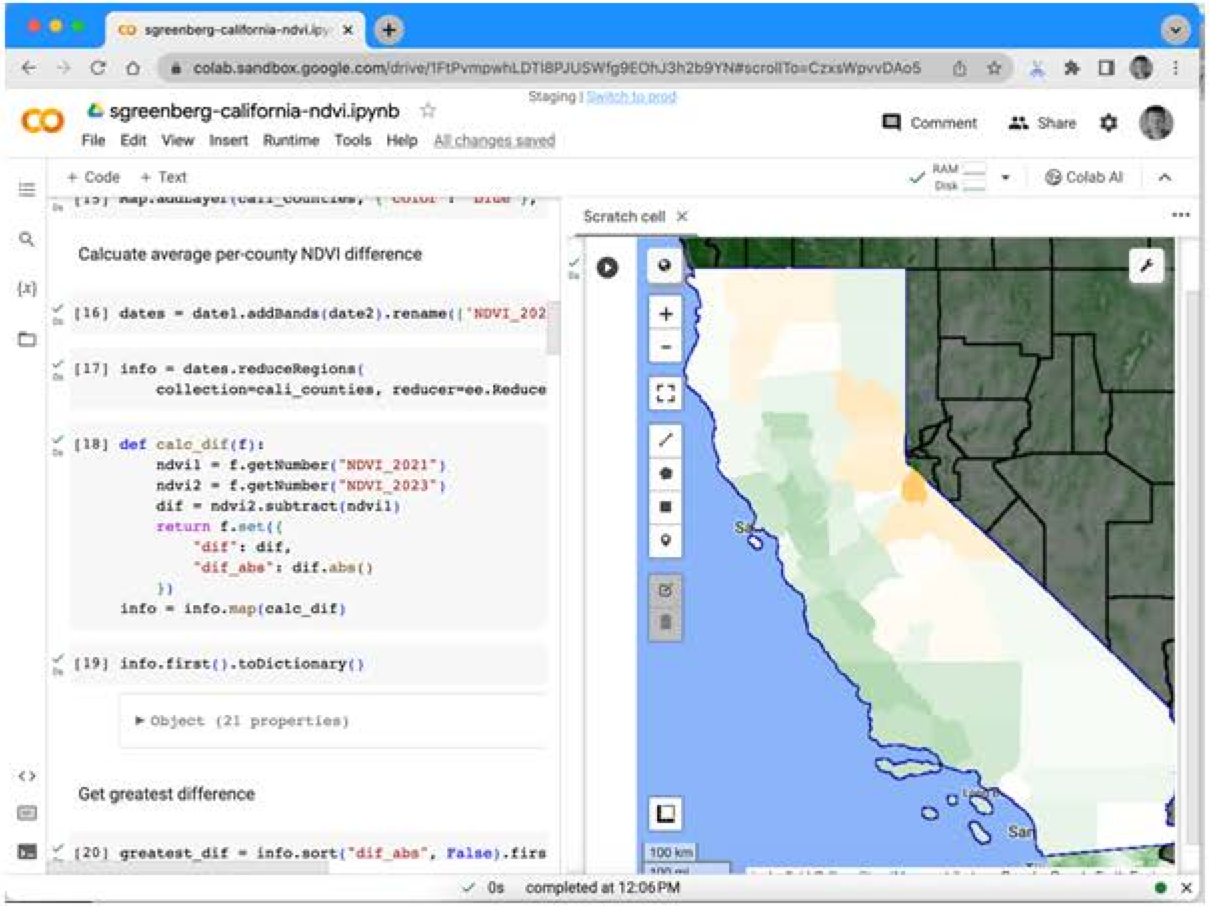
What is a notebook?

Web-based interactive computing platform combining live code, narrative text, visualizations …

Very popular with Data Scientists and ML Engineers

A few UI implementations:

- Jupyter Notebook (project Jupyter)
- JupyterLab (project Jupyter)
- Colaboratory (Google Research)
- Vertex AI Workbench (Google Cloud)
- Kaggle Kernels (Kaggle)



Overview

Google Colab

Zero Setup

Integrated Auth

Hosted on Google's Infrastructure

JupyterLab

Hosted or installed on your computer

Extremely customizable

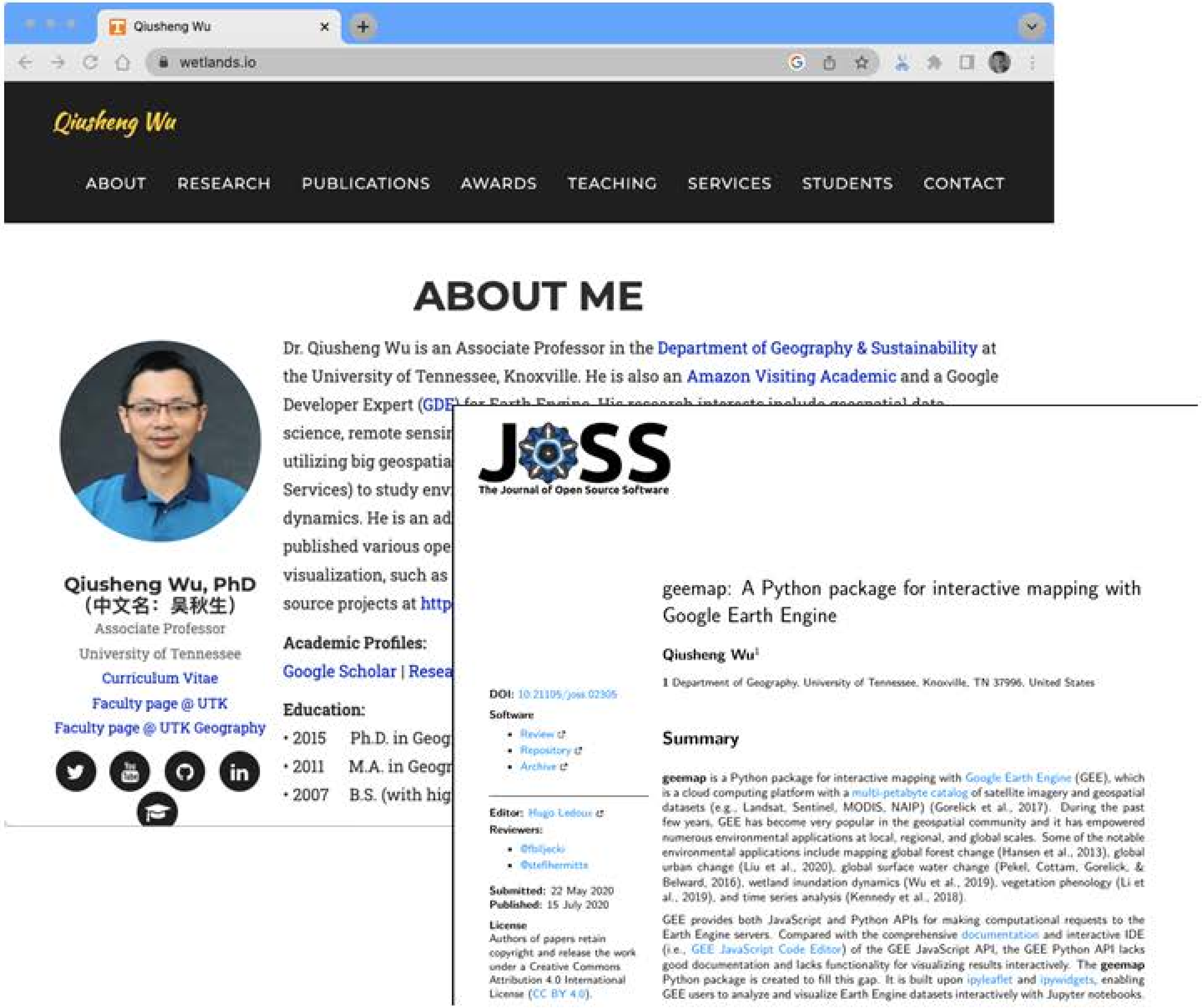
What is geemap?

A Python package for interactive mapping with Google Earth Engine, ipyleaflet and ipywidgets.

Reproduces many Code Editor niceties in a Python notebook.

Uses widgets like folium, ipyleaflet, and plotly

Built by Dr. Qiusheng Wu.



Hands On:
Introduction to Earth Engine in Python via colab

https://github.com/waleedgeo/flood-mapping-workshop