UNDERGRADUATE RESEARCH PORTFOLIO REPORT

Cloud-Based Sentinel-1 SAR Analysis of the 2022 Pakistan Floods

Rajanpur District / Mithankot Case Study

Sentinel-1 SAR • Google Earth Engine • Python • Remote Sensing

Prepared as an undergraduate technical research portfolio project
for prospective graduate research and research-assistantship evaluation

Revised academic edition • September 2026

Subhan Nadeem

Abstract

This undergraduate research project investigates the use of Copernicus Sentinel-1 Synthetic Aperture Radar (SAR) imagery and the Google Earth Engine (GEE) Python API to examine flood-related surface-water expansion in the Rajanpur–Mithankot region of Punjab, Pakistan, during the 2022 monsoon disaster. The analysis uses a rectangular study area defined by coordinates [70.1, 28.8, 70.6, 29.4] (EPSG:4326), with a June 2022 pre-monsoon reference period and an August 25–September 5, 2022 peak-flood period. Sentinel-1 Ground Range Detected imagery was filtered to Interferometric Wide (IW) mode and VV polarization, composited using the median, clipped to the study region, and rendered as comparable grayscale backscatter images. The original project records a flood-extent output of 422.88 km^2 and visually demonstrates a marked increase in low-backscatter areas during the peak-flood period. The work is positioned as a proof-of-concept and research portfolio study rather than a validated operational flood product. In particular, the retained code excerpt documents image acquisition and visualization but does not preserve the full thresholding, permanent-water separation, pixel-area reduction, or independent validation steps that would be required to reproduce and statistically verify the 422.88 km^2 estimate. The revised report therefore emphasizes methodological transparency, appropriate scientific language, relevant remote-sensing literature, and explicit limitations while preserving the original project output.

Keywords: Sentinel-1; synthetic aperture radar; flood mapping; Google Earth Engine; Pakistan floods; Rajanpur; remote sensing; geospatial analysis

Editorial scope note: Project-specific methods, dates, coordinates, images, code excerpts, and the 422.88 km^2 result are taken from the supplied report. Peer-reviewed and official literature has been added in this revision to strengthen scientific context. No undocumented algorithm, validation result, or accuracy statistic has been invented.

Contents

1. Introduction and Research Objective
2. Literature and Technical Background
3. Study Area, Data, and Computing Environment
4. Methodology
5. Implementation and Reproducibility
6. Results
7. Discussion
8. Limitations and Validity
9. Future Research Extension
10. Conclusion

References

Appendix A. Original Python Extraction/Visualization Excerpt

1. Introduction and Research Objective

The 2022 monsoon floods in Pakistan created a major humanitarian and economic disaster. The national Post-Disaster Needs Assessment reported that approximately 33 million people were affected and estimated combined damages and economic losses above US$30 billion (Government of Pakistan et al., 2022). Rapid spatial information was therefore important for understanding the evolving flood footprint across the Indus Basin.

Optical Earth-observation imagery is valuable for flood mapping, but persistent monsoon cloud cover can obstruct the land surface during the period when observations are most urgently needed. Synthetic Aperture Radar (SAR) provides an important alternative because it is an active microwave system and can acquire imagery during day or night and under most cloud conditions. Sentinel-1, in particular, has become a widely used source for rapid flood delineation because of its C-band SAR instrument, systematic acquisitions, open data policy, and suitability for cloud-independent disaster monitoring (European Space Agency, 2016; Schumann & Moller, 2015).

This project was developed as an undergraduate remote-sensing and cloud-geospatial analysis exercise focused on the Rajanpur–Mithankot area of Punjab. Its purpose was not to propose a new operational flood-mapping algorithm, but to demonstrate the ability to define a study region, access and process satellite data through GEE, construct comparable pre- and post-event SAR composites, generate quantitative and visual outputs, and document the workflow in Python and version control.

Research objective

The primary objective is to examine whether Sentinel-1 VV SAR imagery processed in a cloud environment can provide a useful pre-event versus peak-event representation of flood-related surface-water expansion in the defined Rajanpur study region during the 2022 Pakistan monsoon floods.

Research questions

- RQ1. How does the spatial pattern of Sentinel-1 VV backscatter differ between the June 2022 baseline and the August–September 2022 peak-flood period within the selected study region?
- RQ2. Can a lightweight GEE/Python workflow provide reproducible access, compositing, clipping, and visualization of Sentinel-1 data for an undergraduate flood-analysis case study?
- RQ3. What flood-extent value was produced by the original project workflow, and what methodological qualifications are necessary when interpreting that value?

2. Literature and Technical Background

SAR flood mapping is based on the interaction between microwave energy and the Earth's surface. Calm open water commonly produces low radar backscatter because much of the incident energy is reflected away from the sensor in a specular direction. Rough soil, crops, and built surfaces tend to return more energy through diffuse or complex scattering. This contrast can make open water appear dark relative to many surrounding land covers in SAR imagery. However, the relationship is not universally simple: wind-roughened water, flooded vegetation, urban double-bounce scattering, radar shadow, smooth bare surfaces, and acquisition geometry can complicate interpretation (Schumann & Moller, 2015).

Flood-mapping studies therefore commonly extend beyond visual interpretation. Published Sentinel-1 workflows have used automatic thresholding, change detection, terrain screening, permanent-water masks, morphological filtering, fuzzy logic, and validation against independent reference products. Twele et al.

(2016) demonstrated a fully automated Sentinel-1 flood-mapping chain for near-real-time applications. Tiwari et al. (2020) and Moharrami et al. (2021) applied Otsu-based thresholding in GEE and evaluated the resulting flood maps against independent data. Chen and Zhao (2022) further showed how thresholding can be refined to reduce false classification from terrain shadows and other low-backscatter surfaces.

Pakistan has also been studied directly using Sentinel-1. Earlier work on Pakistan demonstrated that thresholding and contextual refinement can map large flood events while accounting for terrain, permanent water, and flood-like low-backscatter surfaces. For the 2022 disaster, Roth et al. (2023) evaluated Sentinel-1-based flood monitoring across Pakistan, while Chen et al. (2024) presented a nationwide automated analysis based on thresholding and fuzzy logic. These studies provide an important benchmark: they confirm the scientific relevance of Sentinel-1 for the 2022 Pakistan floods, while also showing that operational-quality mapping requires more extensive preprocessing and validation than the compact undergraduate workflow retained in the present project.

Why VV polarization and IW mode?

The supplied project uses VV polarization and Interferometric Wide (IW) swath mode. IW is Sentinel-1's principal land-acquisition mode, providing a wide swath suitable for regional analysis. VV backscatter can provide clear discrimination between open water and many non-water surfaces, although published research often evaluates multiple polarizations or additional contextual information to improve robustness (European Space Agency, 2016; Schumann & Moller, 2015).

3. Study Area, Data, and Computing Environment

The analysis focuses on a rectangular area surrounding Rajanpur District and Mithankot in southern Punjab near the Indus River. The original project uses a bounding box rather than an administrative district polygon; accordingly, results in this report should be interpreted as applying to the defined analysis rectangle, not automatically to the full administrative area of Rajanpur District.

Parameter	Specification
Study focus	Rajanpur District / Mithankot area, Punjab, Pakistan
Analysis geometry	Rectangle [70.1, 28.8, 70.6, 29.4] in EPSG:4326
Pre-flood baseline	1 June 2022 – 30 June 2022
Peak-flood period	25 August 2022 – 5 September 2022
Satellite dataset	Copernicus Sentinel-1 Ground Range Detected (GRD) SAR
Acquisition mode	Interferometric Wide (IW)
Polarization used	VV
Cloud processing	Google Earth Engine Python API in Google Colab
Visualization	GEE image rendering; Folium/Matplotlib referenced in

Parameter	Specification
	original report; Figma used for dashboard presentation
Reported project output	422.88 km^2 flood extent (original project result; not independently revalidated in this revision)

Table 1. Study design and computational environment reconstructed from the original project materials.

4. Methodology

The documented workflow is a cloud-native pre/post SAR comparison. The implementation filters Sentinel-1 imagery by study geometry and date, restricts acquisition mode to IW, selects the VV band, creates a median composite, clips the result to the study region, and renders a grayscale image using a nominal display range of −25 to 0 dB. The same structure was used for the June baseline and the August–September flood period.

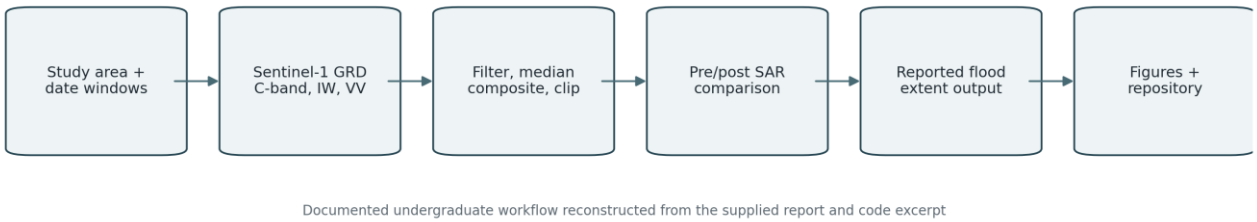


Figure 1. Documented workflow based on the supplied report and retained code excerpt.

4.1 SAR interpretation principle

Open standing water often exhibits low C-band SAR backscatter because of specular reflection away from the sensor. In contrast, rougher terrain, vegetation, and built surfaces tend to return stronger signals. The project therefore interprets the expansion of coherent dark, low-backscatter regions between the baseline and flood-period composites as evidence consistent with increased surface-water coverage. This physical interpretation is well established, but dark pixels are not synonymous with floodwater in all environments; radar shadow and other smooth surfaces can generate similarly low returns (Schumann & Moller, 2015).

4.2 Temporal compositing

For each date window, the image collection is reduced using the median. Median compositing can suppress isolated noise and provide a representative surface response across multiple acquisitions, but it also changes the interpretation from a single-date flood snapshot to a multi-date composite. The result therefore represents conditions across the selected period rather than one exact acquisition time.

4.3 Methodological transparency

The original narrative states that a mathematical threshold was used to isolate inundated pixels, and it reports a flood area of 422.88 km^2. However, the code excerpt preserved in the report does not show the threshold value, water-mask expression, permanent-water exclusion, pixel-area calculation, or spatial reduction used to derive that number. For academic transparency, this revised report does not reconstruct

those missing operations from assumption. Instead, it preserves 422.88 km^2 as the reported output of the original project and explicitly identifies the missing reproducibility details as a limitation.

5. Implementation and Reproducibility

The project was implemented in Google Colaboratory using the Earth Engine Python API. This architecture allowed satellite imagery to remain server-side while the notebook issued filtering, compositing, clipping, and rendering operations. Cloud processing is appropriate for an undergraduate project because it reduces local storage and compute requirements while exposing the researcher to reproducible geospatial scripting (Gorelick et al., 2017).

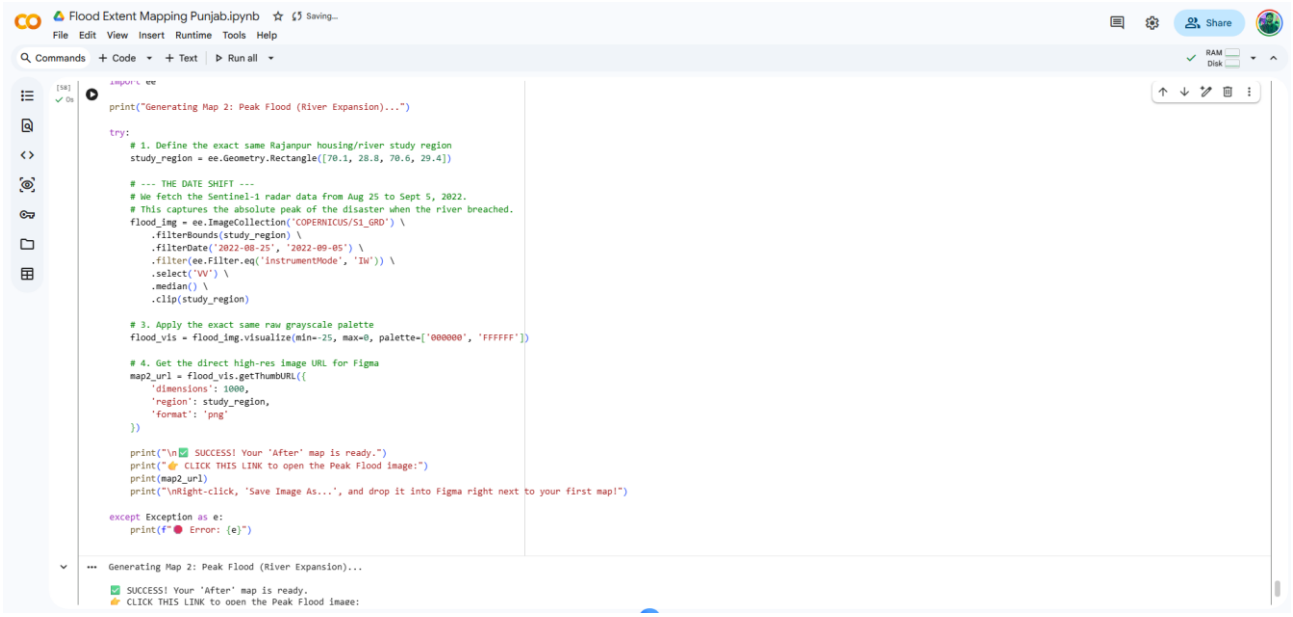


Figure 2. Original Google Colab notebook screenshot showing the peak-flood Sentinel-1 extraction and visualization workflow.

The retained notebook screenshot confirms the key implementation sequence: defining the study rectangle, filtering COPERNICUS/S1_GRD by date and geometry, restricting instrumentMode to IW, selecting VV, applying a median reducer, clipping the result, and generating a visualization thumbnail. This is consistent with the core source-code excerpt reproduced in Appendix A.

5.1 Reproducibility status

Component	Status	Evidence / qualification
Study coordinates	Documented	Explicit rectangle in report and code excerpt.
Baseline/flood date windows	Documented	Explicit date windows are retained.
Sentinel-1 mode and polarization	Documented	IW and VV are explicitly selected.
Composite operation	Documented	Median compositing is shown in code.
Visualization range	Documented	−25 to 0 dB grayscale rendering is

Component	Status	Evidence / qualification
		shown.
Flood classification threshold	Not preserved	Narrative mentions thresholding, but the threshold value and expression are absent from the retained code.
Permanent-water / terrain masking	Not documented	No mask is visible in the retained excerpt.
Pixel-area reduction to 422.88 km^2	Not preserved	The numerical result is reported, but the area-calculation code is not shown.
Independent accuracy assessment	Not performed in supplied report	No reference dataset, confusion matrix, or uncertainty analysis is provided.

Table 2. Reproducibility status of the supplied undergraduate workflow.

6. Results

The original project produced two principal outputs: (1) comparable pre-monsoon and peak-flood Sentinel-1 VV composites, and (2) a reported flood-extent metric of 422.88 km^2 within the defined analysis region. The strongest directly reproducible evidence in the supplied material is the change in spatial backscatter pattern between the two composite images.

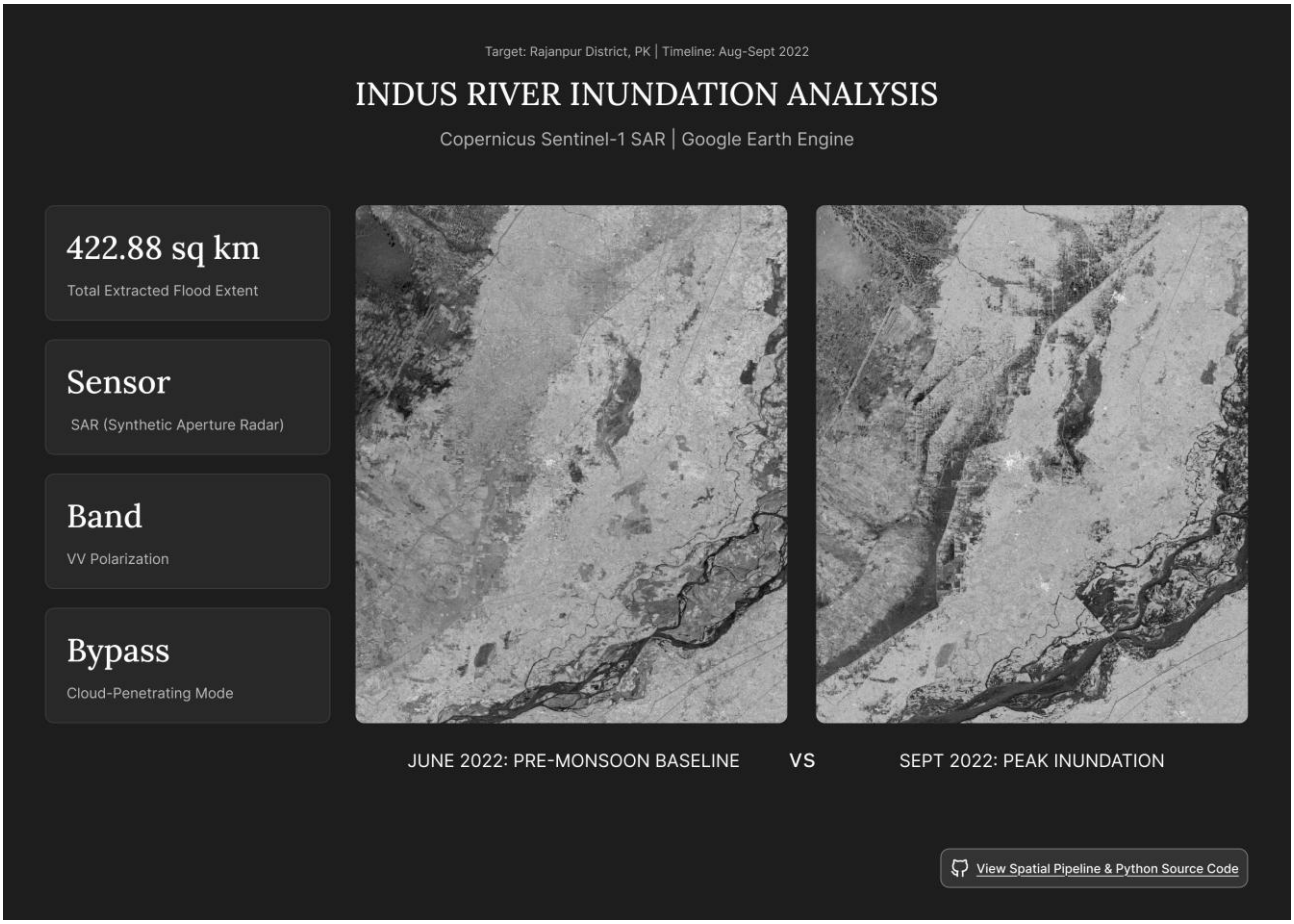


Figure 3. Original project dashboard comparing the June 2022 baseline with the August–September 2022 peak-flood composite. The 422.88 km² value is retained as the project-reported extent.

6.1 Pre-monsoon baseline



Figure 4. June 2022 Sentinel-1 VV median composite for the study region (original project output).

The June composite shows the normal river network as relatively dark features embedded within a more heterogeneous agricultural and settled landscape. The image provides a reference against which the later low-backscatter expansion can be visually assessed.

6.2 Peak-flood period



Figure 5. August 25–September 5, 2022 Sentinel-1 VV median composite for the study region (original project output).

The peak-period composite contains substantially larger contiguous zones of low backscatter than the June baseline. In the context of the 2022 flood event and the Indus floodplain setting, this pattern is consistent with widespread surface-water expansion. The visual evidence supports the qualitative conclusion that inundation increased markedly within the analysis region during the selected peak period.

6.3 Reported flood-extent metric

Metric	Original project output
Flood extent	422.88 km^2

Table 3. Quantitative result reported by the original project.

The 422.88 km^2 value should be interpreted as a project-reported estimate rather than a statistically validated ground-truth measurement. The supplied document does not retain enough classification and area-reduction code to independently regenerate the value or quantify its uncertainty.

7. Discussion

The project demonstrates the central practical advantage of SAR for monsoon flood analysis: it permits surface observation when optical imagery may be obstructed by cloud. The visible increase in low-backscatter area between the June and August–September composites is physically plausible for a major flood event and is consistent with the broader scientific record of severe inundation across Pakistan in 2022. National-scale Sentinel-1 analyses published after the event likewise documented extensive flooding and demonstrated the value of automated SAR-based monitoring (Roth et al., 2023; Chen et al., 2024).

However, similarity to broader flood patterns does not validate the exact local area estimate. Operational and publication-grade flood products typically include explicit classification logic, permanent-water separation, terrain screening, acquisition-geometry controls, post-processing, and independent validation. The distinction is important: this undergraduate project succeeds as a proof-of-concept in cloud geospatial processing and event visualization, but its retained documentation is not sufficient to claim a fully validated flood classification product.

The use of a rectangular bounding box is also analytically important. Because the region is not clipped to an administrative boundary, the 422.88 km^2 result cannot be interpreted automatically as the amount of Rajanpur District flooded. It applies to the project's defined analysis rectangle. This clarification improves geographic precision and prevents overgeneralization.

From a research-training perspective, the project nonetheless demonstrates several transferable skills: formulating a remote-sensing problem, selecting appropriate satellite data, scripting server-side geospatial operations, reasoning about SAR backscatter, comparing multi-temporal imagery, generating research figures, documenting code, and recognizing the difference between visual evidence and validated classification. These are relevant foundations for graduate work in remote sensing, GIS, environmental data science, disaster analytics, and cloud geospatial computing.

8. Limitations and Validity

Classification reproducibility: The retained code does not show the threshold or Boolean flood-mask operation used to produce the numerical area output.

Area calculation: The pixel-area and spatial-reduction code used to produce 422.88 km^2 is not included in the supplied excerpt.

Permanent water: No explicit removal of the normal Indus River channel or other permanent water bodies is documented.

Acquisition geometry: The retained code does not constrain orbit direction, relative orbit, or incidence angle; these factors can affect backscatter comparability.

Speckle and contextual filtering: No speckle filter, connected-pixel filter, terrain mask, or flood-lookalike screening is documented in the excerpt.

Validation: The project does not provide independent reference data, a confusion matrix, precision/recall, intersection-over-union, or an uncertainty interval.

Temporal interpretation: Median compositing over multi-day windows provides a period summary rather than a single-date flood snapshot.

Spatial scope: The analysis geometry is a bounding box around the Rajanpur/Mithankot area rather than the Rajanpur administrative boundary.

Polarization scope: The analysis uses VV only; other polarizations or multi-channel information could improve discrimination in some land-cover conditions.

9. Future Research Extension

A graduate-level extension should preserve the accessibility of the current workflow while adding explicit classification and validation. A defensible next version could: (1) match pre- and post-event acquisitions by orbit geometry; (2) apply an adaptive or literature-supported water threshold; (3) remove permanent water using a historical surface-water product; (4) mask steep terrain and low-backscatter shadows where appropriate; (5) remove isolated pixels through connected-component or morphological filtering; (6) compute flood area from the final binary mask using ee.Image.pixelArea(); and (7) validate the map against an independent flood product or cloud-free optical reference. These steps are consistent with established Sentinel-1 flood-mapping practice (Twele et al., 2016; Tiwari et al., 2020; Moharrami et al., 2021; Chen & Zhao, 2022).

Sensitivity analysis should also be added. Rather than reporting a flood area to two decimal places without uncertainty, the researcher should evaluate how the result changes across plausible thresholds and masking choices. This would convert the present proof-of-concept into a stronger research study while preserving the same cloud-based toolchain.

10. Conclusion

This project demonstrates an undergraduate-level application of Sentinel-1 SAR, Google Earth Engine, and Python to the 2022 Pakistan floods. Within a defined Rajanpur–Mithankot analysis rectangle, the workflow produces comparable June and August–September VV composites and shows a pronounced increase in low-backscatter regions consistent with flood-related surface-water expansion. The original project reports an inundation estimate of 422.88 km^2.

The revised report deliberately separates what is directly documented from what would be required for publication-grade verification. The supplied code reliably demonstrates data access, filtering, compositing, clipping, and visualization, but it does not retain the complete classification and area-reduction steps needed to reproduce the 422.88 km^2 value. Accordingly, the result is presented as a project output rather than an independently validated measurement.

As a research portfolio artifact, the project provides credible evidence of initiative and technical capability in remote sensing, cloud geospatial analysis, scientific programming, and research communication. Its most valuable next step is not stronger promotional language, but a more explicit and validated flood-mask workflow. That extension would create a clear path from an undergraduate proof-of-concept toward a graduate-level research study.

Research artifact and version-control evidence

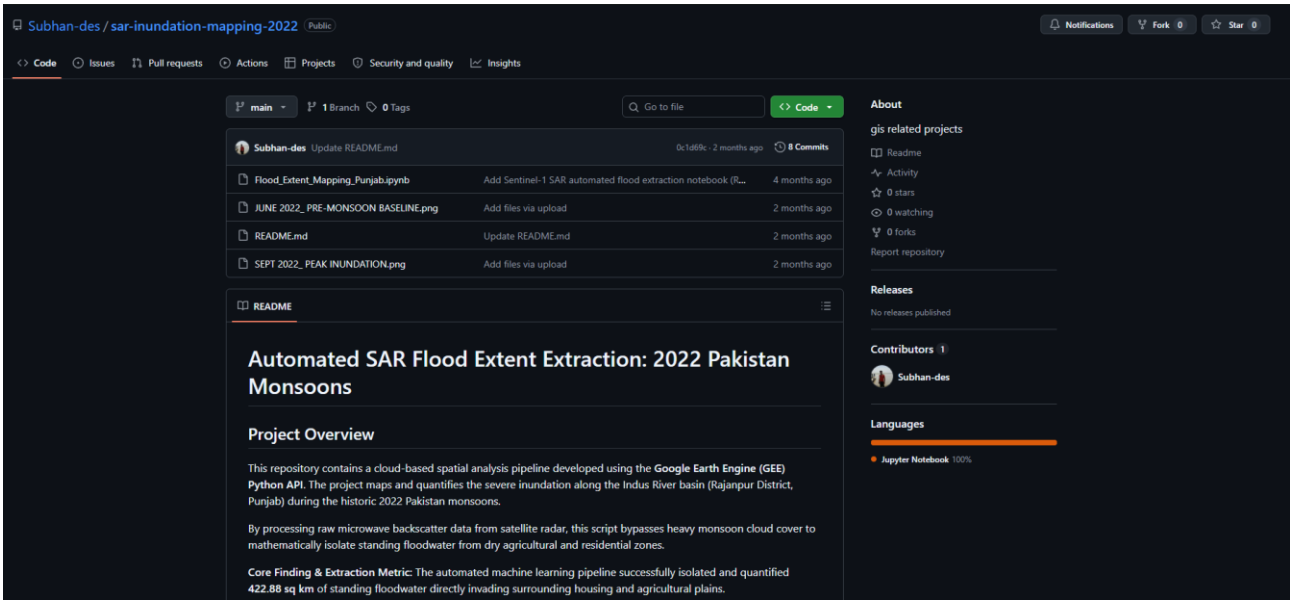


Figure 6. Original repository screenshot documenting notebook and image assets associated with the project.

References

Chen, F., Zhang, M., Zhao, H., Guan, W., & Yang, A. (2024). Pakistan's 2022 floods: Spatial distribution, causes and future trends from Sentinel-1 SAR observations. Remote Sensing of Environment, 304, 114055. https://doi.org/10.1016/j.rse.2024.114055

Chen, Z., & Zhao, S. (2022). Automatic monitoring of surface water dynamics using Sentinel-1 and Sentinel-2 data with Google Earth Engine. International Journal of Applied Earth Observation and Geoinformation, 113, 103010. https://doi.org/10.1016/j.jag.2022.103010

European Space Agency. (2016). Sentinel-1 product definition (Issue 2/7). European Space Agency.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Government of Pakistan, Asian Development Bank, European Union, United Nations Development Programme, & World Bank. (2022). Pakistan floods 2022: Post-disaster needs assessment.

Moharrami, M., Javanbakht, M., & Attarchi, S. (2021). Automatic flood detection using Sentinel-1 images on the Google Earth Engine. Environmental Monitoring and Assessment, 193, 248. https://doi.org/10.1007/s10661-021-09037-7

Roth, F., Bauer-Marschallinger, B., Tupas, M. E., Reimer, C., Salamon, P., & Wagner, W. (2023). Sentinel-1-based analysis of the severe flood over Pakistan 2022. Natural Hazards and Earth System Sciences, 23, 3305–3317. https://doi.org/10.5194/nhess-23-3305-2023

Schumann, G. J.-P., & Moller, D. K. (2015). Microwave remote sensing of flood inundation. Physics and Chemistry of the Earth, Parts A/B/C, 83–84, 84–95. https://doi.org/10.1016/j.pce.2015.05.002

Tiwari, V., Kumar, V., Matin, M. A., Thapa, A., Ellenburg, W. L., Gupta, N., & Thapa, S. (2020). Flood inundation mapping—Kerala 2018: Harnessing the power of SAR, automatic threshold detection method and Google Earth Engine. PLOS ONE, 15(8), e0237324. https://doi.org/10.1371/journal.pone.0237324

Twele, A., Cao, W., Plank, S., & Martinis, S. (2016). Sentinel-1-based flood mapping: A fully automated processing chain. International Journal of Remote Sensing, 37(13), 2990–3004. https://doi.org/10.1080/01431161.2016.1192304

Appendix A. Original Python Extraction/Visualization Excerpt

The following code is retained from the supplied report. It documents the image-extraction and visualization stage of the project. It should not be interpreted as a complete reproduction of the flood-mask or area-calculation logic used to obtain the reported 422.88 km^2 value.

```
import ee

def extract_sar_imagery():
    print("Initializing Automated SAR Extraction Pipeline...")
    try:
        study_region = ee.Geometry.Rectangle([70.1, 28.8, 70.6, 29.4])

        flood_img = (ee.ImageCollection('COPERNICUS/S1_GRD')
            .filterBounds(study_region)
            .filterDate('2022-08-25', '2022-09-05')
            .filter(ee.Filter.eq('instrumentMode', 'IW'))
            .select('VV')
            .median()
            .clip(study_region))

        flood_vis = flood_img.visualize(
            min=-25,
            max=0,
            palette=['000000', 'FFFFFF']
        )

        map_url = flood_vis.getThumbURL({
            'dimensions': 1200,
            'region': study_region,
            'format': 'png'
        })

        print(f"Extraction Successful. Asset URL: {map_url}")
    except Exception as e:
        print(f"Pipeline Error: {e}")

# The same extraction structure was used for the June 1–30 baseline.
```

End of report