Laboratory Assignment

Applied Geospatial Computations (U of H)

Main project: ***Accuracy assessment of SAOCOM-derived InSAR time-series DEMs over Italy***

Target outcome: a conference-quality manuscript suitable for submission to *IEEE T-GRS* or a comparable journal.

Instructor: Prof. Pietro Milillo, pmilillo@uh.edu /Assistant: Amin Tavakkoli: atavakk3@CougarNet.UH.EDU

Note: When communicating with the assistant or point of contact please cc also instructor.

0. Introduction – What this lab is all about

Welcome to your capstone lab for **Applied Geospatial Computations**.

Over the next twelve weeks you will step into the role of a remote-sensing scientist charged with answering one big, practical question:

How good are Digital Elevation Models (DEMs) that we build from SAOCOM L-band InSAR time-series, and what do they add to the elevation products we already have?

To get there, you will:

1. **Build your own SAOCOM DEM**
    - In collaboration with Amin you will be given a SAOCOM-derived digital terrain model which you will have to evaluate. Your Reference point for this dataset will be Dr. Emanuela Bonano at CNR IREA. She is the first author of a paper where the dataset and the SAOCOM constellation is described. Please not that you are required to read all the readings (papers and links provided in this assignment)
2. **Round-up trustworthy reference data**
    - You will download existing Italian elevation layers—TINITALY (10 m), Copernicus DEM (30 m), and high-resolution regional LiDAR tiles—to serve as ground truth.
3. **Compare and quantify accuracy**
    - For every reference you will compute classic statistics (bias, RMSE, MAE, σ) and robust measures (median error, NMAD, LE95).
    - You will repeat the analysis inside each CORINE land-cover class to see where L-band works best and where it may struggle.
4. **Map the blind spots**

- By thresholding the temporal coherence that comes out of your InSAR processor, you will identify areas where SAOCOM simply cannot deliver a height estimate and relate those gaps to land-cover type.

5. **Tell a clear, evidence-based story**
    - Your findings will feed a journal-style paper that shows, with numbers and figures, how SAOCOM time-series DEMs perform across Italy and how they could pave the way for NASA's future **Surface Topography & Vegetation (STV)** mission.

Along the way you will sharpen several core skills:

- **Radar interferometry fundamentals** – turning phase into meters and knowing what can go wrong.
- **Python & Google Earth Engine scripting** – automating large-area downloads, metrics, and plots.
- **Good scientific hygiene** – version-controlled code, reproducible notebooks, and uncertainty budgets to be published on Github.
- **Critical writing** – translating tables of numbers into concise arguments a reviewer can follow.

By the end of the lab you will not only have produced a peer-review-ready manuscript but also a reusable toolbox that future students, and perhaps even space-agency teams, can build on.

Ready? Let's map some heights.

1. How this lab is organized

Item	Deliverable	Weight
Pre-processing notebooks (Python/GEE scripts)	GitHub repo + reproducible README	15 %
Interim memo (2 pages) with early metrics & plots	Week 6	10 %
Final technical report (≈ 20 pages) in journal format	Week 12	55 %
Oral defence (10 min) + slide deck + questions (15-20 min)	Finals week	10 %
Re-usable code packaged as "student's toolbox"		10 %

Besides standard hours of teaching, through the whole lab. experience, you will be coordinating with your instructor, external point of contact and/or lab assistants through quick 30 min meetings that will start with you presenting the progresses you made and challenges you are facing in your project in a 3 min short presentation.

2. Some background: The NASA STV Decadal-Survey Incubation Study

You will be working in a real research context on a **NASA-sponsored international project called the NASA Surface Topography and Vegetation (STV) Decadal Survey Incubation Study**. NASA STV is a concerted R&D effort to mature the innovative sensing technologies and systems needed to map the Earth's surface and vegetation in unprecedented detail, laying the groundwork for a future orbital mission in the early 2030s. It's *not* itself a satellite mission but a coordinated incubation program exploring viable architectures and measurement approaches. Here's a closer look:

Origins & Mission

- **Mandate**: Established under the 2019 ROSES-2019 solicitation (Decadal Survey Incubation Study Teams A.54), alongside a Planetary Boundary Layer team. Work commenced in March 2020, with deliverables due to NASA HQ by early calendar year 2021.
- **Objective**: Refine science and applications requirements, evaluate mature sensing tech (lidar, radar, stereo-photogrammetry), identify gaps, and draft a Science & Applications Traceability Matrix (SATM) connecting Science objectives with measurement requirement.

How does my project fits with NASA STV?

The goal of your project is then to **evaluate the performance and scientific utility of Digital Elevation Models (DEMs) derived from SAOCOM L-band InSAR time-series**, with a focus on their **added value relative to existing global and regional elevation products**. This includes assessing how SAOCOM-based DEMs perform in terms of accuracy, resolution, and reliability across different land cover types and topographic conditions.

The **accuracy results of your analysis will directly feed into [NASA's Science and Applications Traceability Matrix (SATM)]**, a key planning tool that translates high-level scientific goals into quantitative measurement requirements. Your task will be to map the performance characteristics of the SAOCOM-derived DEMs against these SATM requirements, identifying which **NASA STV science goals** can be satisfied by L-band InSAR-based elevation data and under what conditions.

Through this process, you will help define the **role of SAOCOM-like L-band constellation** in the future STV mission architecture and clarify where they complement or surpass existing datasets such as SRTM, TanDEM-X, NASADEM, or Copernicus GLO-30. Your work will contribute to guiding NASA's decisions about sensor selection, orbital design, and data fusion strategies for high-resolution topographic mapping in the coming decade.

STV Focus Areas

The study tackles four key observables, aiming for **contiguous global coverage with repeat seasonal sampling**:

1. **Bare-Earth Land Topography**
2. **Ice & Glacier Surface Elevation**
3. **3D Vegetation Structure** (height, biomass, canopy)
4. **Shallow-Water Bathymetry**

Overall we will assess how well SAOCOM InSAR time-series DEMs over Italy can characterized topography over STV Focus Areas.

3. How InSAR measures topography: a concise refresher

Based on the ESA handbook "InSAR Principles – Guidelines for SAR Interferometry Processing and Interpretation" (2007):

- Single-pass height: $h = (\lambda R_0 \Delta\varphi) / (4\pi B\perp \sin\theta)$. (Paragraph 2.2 Starting at page A-18)
- Coherence budget: temporal, geometric, volumetric, thermal, quantization. (Paragraph 2.6 , 2.7, 3.5, 3.6)
- Time-series InSAR decouples (i) DEM error, (ii) deformation, (iii) atmosphere via N repeat acquisitions and a network inversion (SBAS, PSI, squeeSAR, or TomoSAR).
- L-band advantage (SAOCOM, ALOS-2): longer wavelength (≈ 24 cm) → lower temporal decorrelation over crops & forests.

If these equations and terms do not tells you much you should read the following InSAR tutorial.

3. InSAR time-series DEM generation

3.1 Persistent Scatterers (PSI) approach

Perissin & Rocca (2006, 2008, 2008) demonstrated that combining > 40 ERS/Envisat scenes yields ≤ 1 m vertical precision in Milan.

Height variance (white-noise approximation):

$$\sigma^2_{\Delta h} \approx \left(\frac{\lambda\, R \sin{(\theta)}}{4\pi}\right)^2 \frac{\sigma^2_{\phi}}{N\, \sigma^2_{B}}$$

where R is the target-sensor distance, θ is the incidence angle, λ is the sensor wavelength, σ^2_{ϕ} is the phase noise variance supposed independent of the acquisition and σ^2_{B} isthe variance of the perpendicular baselines of our SAR acquisitions.

3.2 TomoSAR approach

Zhu & Bamler (2010) extended the aperture into the elevation dimension; 4-D TomoSAR simultaneously retrieves elevation + linear velocity of multiple scatterers inside one pixel, using a Wiener-regularised SVD that stabilizes ill-conditioned baselines.

Note: Don't worry if you feel overwhelmed while reading these literature studies. Ultimately, your task will be to assess the accuracy of a DEM that has already been provided to you. While a deep understanding of the underlying algorithms is not essential, gaining some familiarity will help you better contextualize your lab assignment and may prove valuable if you pursue a publication based on your work. **If you found these papers interesting and you would like to deepen your knowledge consider taking my SAR class CIVE 6359.**

4. SAOCOM L-band acquisition plan over Italy

- Constellation: SAOCOM-1A (2018) & SAOCOM-1B (2020) in a 620 km, 97.9° sun-sync orbit.
- Repeat cycle: 16 days; effective sampling ≤ 8 days when both satellites task the same frame.
- Beam modes: StripMap 10 m, ScanSAR 25 m, Polarimetric 10 m; incidence 20–41°.
- Italian campaign: systematic StripMap coverage of the Po valley, Alps and major volcanoes (CNR-IREA & ASI).
- Data access: SAOCOM SLCs via ASI Collaborative Ground Segment (CGS) or CONAE portal.

5. Assignment tasks

5.1 Data assembly

Layer	Resolution	Download hint

SAOCOM SLC stack (≥ 25 scenes, 2018-2025, same beam)	10 m	IREA
TINITALY DEM	10 m	https://tinitaly.pi.ingv.it
Copernicus DEM GLO-30	30 m	https://dataspace.copernicus.eu
Regional LiDAR tiles (Veneto, Lombardia)	1-2 m	Geoportale PCN links provided
CORINE Land-Cover 100 m (CLC 2018)	100 m	GEE ID: COPERNICUS/CORINE/V20/100m

Tip: keep all rasters UTM-32, 33 or 34 N depending on the area.

5.2 DEM generation

Display both SAOCOM DEMs, The reference DEMs and the Corinne Land cover map. What do you notice?

5.3 Accuracy metrics

Aggregate your dataset into a single file and compute for every reference DEM: Bias (ME), RMSE, MAE, σ, plus robust stats: Median error, NMAD (1.4826 × MAD), LE95. Map metrics for the entire AOI and for each CLC land-cover class (at least: urban fabric, arable land, broad-leaf forest, coniferous, pasture).

5.4 Spatial completeness

Derive temporal coherence $\bar{\gamma}$ from the SBAS inversion. Mask pixels with $\bar{\gamma} < 0.3$ → "no-DEM". Quantify percentage area lost per land-cover type and discuss mechanisms (e.g. volumetric decorrelation).

5.5 Deliverables & discussion prompts

- Tables of metrics incl. 95 % confidence intervals (bootstrapped).
- Scatter density plots h_SAOCOM vs h_ref with 1:1 line.
- Map of residuals clipped to ± 5 m. Identify systematic tilts or ramps; propose corrections (e.g. tie-points, GCPs).

6. Glossary (use consistently in your report)

Term	**Meaning**

DEM – Digital Elevation Model	Any elevation of Earth's surface (ambiguous).
DSM – Digital Surface Model	Elevation of natural + built surfaces (tops of trees/buildings).
DTM – Digital Terrain Model	"Bare-earth" elevation (vegetation & structures removed).
NMAD	$1.4826 \times \text{median}(\lvert \text{error} - \text{median}(\text{error}) \rvert)$; robust σ-like metric.
LE95	95th percentile of |error|; non-parametric vertical accuracy.

7. Some Suggestions

1. Start Small, Then Scale

- Begin with a small test area, e.g., a 1 × 1 km tile. Develop and validate your code there.
- Avoid jumping directly into large datasets that take hours to process, it will slow your iteration cycle and make debugging harder.
- Only scale to larger areas once plots, logs, and output metrics look correct at the small scale.

2. Save Intermediate Results

- Save intermediate outputs (e.g., resampled DEM tiles + Land Cover Map).
- These may become invaluable for:
    - Debugging failed steps
    - Visual comparisons
    - Avoiding reprocessing time-consuming steps

3. Maintain Consistent Projections

- Make sure all your data layers (raster and vector) are in the same coordinate system.
- Choose meters-based CRS (e.g., UTM or EPSG:25832) unless you specifically require degrees.
- Do **not blindly trust reprojection functions** in your code, always validate with QGIS or another GIS tool by overlaying layers.

4. Don’t Get Stuck Alone

- Try to debug and isolate the issue yourself first.
- Once you’ve exhausted your ideas, **ask for help early**.
- When reaching out (to your instructor, TA), prepare a 3-minute pitch and be prepared to answer any question:
    - What were you trying to do?
    - What did you expect?
    - What actually happened?
    - What steps have you already tried?

5. Use Version Control

- Set up a GitHub or GitLab repository from Day 1.
- Use branches to experiment.
- Commit your code frequently with clear messages.

6. Document Everything

- Maintain a README.md that describes:
    - Your data sources
    - Processing steps
    - Expected outputs
    - Known limitations
- Use comments inside your code to explain logic and formulas.

8. Modularize Your Code

- Break your code into functions or scripts:
    - Read Data
    - Resample
    - Statistical Analysis
    - Land Cover Accuracy Analysis…
- This will reduce duplication and make testing easier.

9. Log Parameters and Choices

- Store input parameters (e.g., baseline thresholds, coherence filters) in a config file (YAML or JSON).
- Include metadata with each run, such as timestamp, number of acquisitions, and output location.

10. Perform Sanity Checks

- Before interpreting results, always:
  - Plot histograms of elevation differences.
  - Inspect residual maps.
  - Compare random elevation profiles across datasets.
  - Check min/max/mean values for outliers.

11. Automate Where You Can

- Use Snakemake, Makefile, or Python scripts to automate multi-tile or multi-area runs.
- Manual repetition increases error risk, especially under deadlines.

12. Use Scalable Libraries

- For large rasters, use libraries like rioxarray, xarray, or dask to handle tiled/chunked operations.
- Avoid loading entire country-wide rasters into memory unless you're sure it's safe.

13. Write Simple Unit Tests

- Write small test functions to verify:
  - That reprojection keeps shapes aligned
  - That metrics like RMSE or NMAD behave as expected
  - That masks are applied correctly

14. Back Up Your Data

- Keep a copy of:
  - Reference DEMs
  - Resampled rasters
  - Final results (plots, tables, maps)
- Store them in cloud storage or on an external drive.

15. Keep a Research Log

- Use a notebook (physical or digital) to write down:
  - What you did each day
  - What worked, what didn't
  - Questions to ask or investigate

    - This will help you stay organized when writing the final paper.

Final tip: Draft your manuscript as you go—write methods while the code is fresh, drop figures into LaTeX early, and reserve Weeks 11-12 only for polishing & response-to-reviewer rehearsal.

Good luck, and remember: reproducibility > flashiness.