POLITECNICO MILANO 1863 | DIPARTIMENTO DI INGEGNERIA CIVILE E AMBIENTALE

AI for Good Global Summit

8 – 11 July 2025
Geneva, Switzerland

palexpo

The role of GeoAI and Foundational Models in shaping an AI-driven future for all

Foundation Models Powering Sustainability and the SDGs

09.07.2025 | **Maria Antonia Brovelli**, Politecnico di Milano,

UN GGIM Academic Network, UN Open GIS, ISDE, ISPRS

Maria Antonia Brovelli, Mohanad Yousef Ahmad Diab, Seyederfan Eshghollahi, Polychronis Kolokousis, Julia Anna Leonardi, Seyed Mohammad Moein Peyghambar, Fatemeh Soleimaniansomarin, Alberto Vavassori, Qiongjie Xu

Content

LangRS: A zero-shot segmentation using GroundingDINO and SAM

Mohanad Diab, Research collaborator at GEOLab

Dr. Polychronis Kolokoussis, Visiting Faculty at GEOLab

Moein Peyghambar Zadeh, Master Student in Geoinformatics Engineering

Maria A Brovelli, Professor, Head of GEOLab

01

LangRS

LangRS is a **Python** package that enables **zero-shot segmentation** of aerial/satellite images using **foundation models;** Grounding DINO for object detection and Segment Anything Model (SAM) for mask generation .

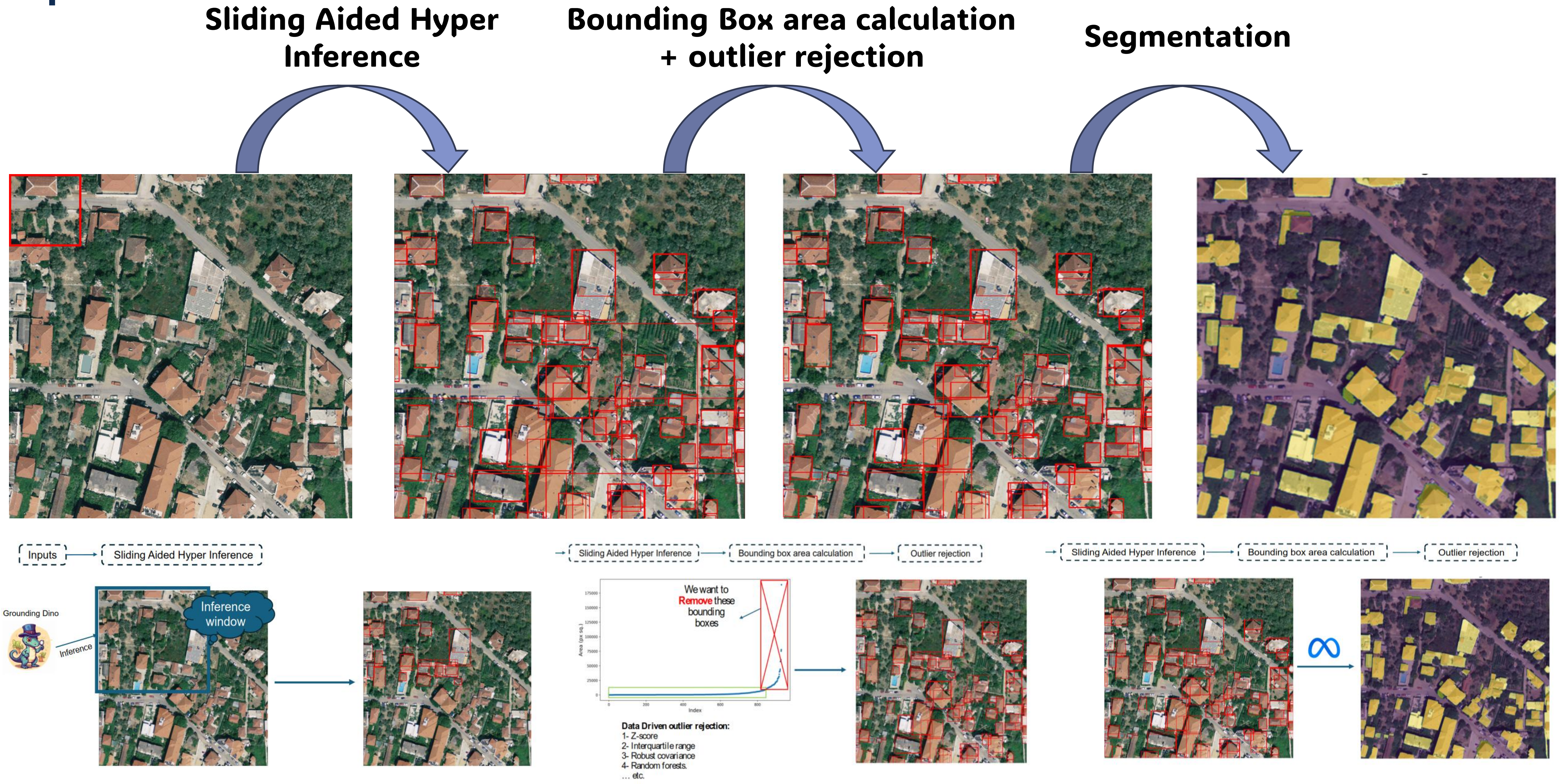
- **Built on:**
  The Segment-Geospatial (SAMGeo) framework, crafted to process geospatial imagery efficiently .

- **Why it matters:**
  Enables optimized, fast, accurate segmentation of features like buildings, trees, cars...etc. without training a new AI model.

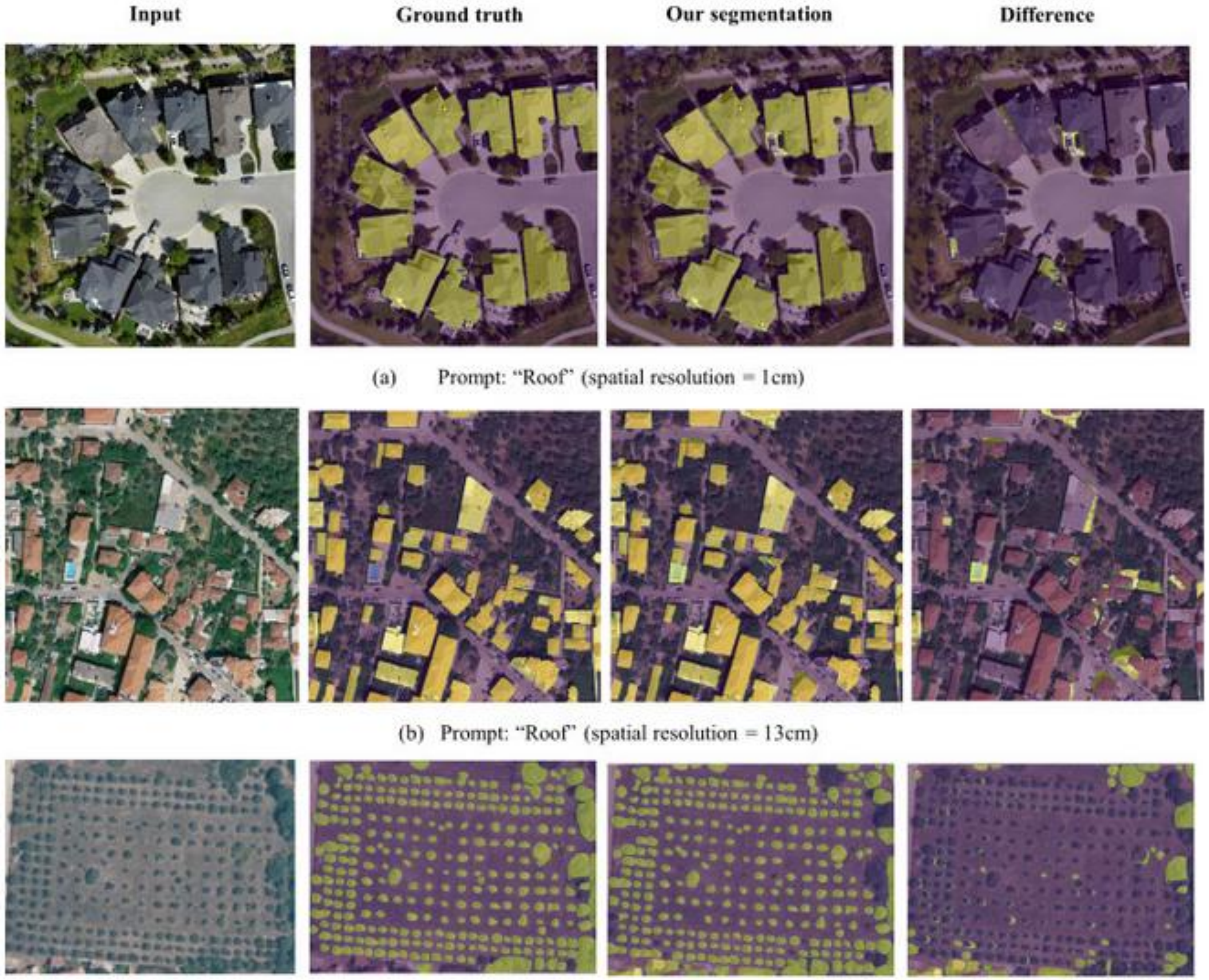
Grounding DINO

SAM

Pipeline


Sliding Aided Hyper Inference
Bounding Box area calculation + outlier rejection
Segmentation
Inputs
Sliding Aided Hyper Inference
Grounding Dino
Inference
Inference window
Sliding Aided Hyper Inference
Bounding box area calculation
Outlier rejection
We want to Remove these bounding boxes
Area (px sq.)
Index
Data Driven outlier rejection:
1- Z-score
2- Interquartile range
3- Robust covariance
4- Random forests.
… etc.
Sliding Aided Hyper Inference
Bounding box area calculation
Outlier rejection

Results



(a) Prompt: "Roof" (spatial resolution = 1cm)

(b) Prompt: "Roof" (spatial resolution = 13cm)

(c) Prompt: "Tree" (spatial resolution = 25cm)

(a) Prompt: "White cars"

(b) Prompt: "Blue cars"

Model metrics for different models trained on different training datasets.

Area	Platform	Prompt	Resolution (cm)	Accuracy	TPR	TNR
1	UAV	Roof	1	98.9	98	2
2	UAV	Blue car	5	99.95	97	3
2	UAV	White car	5	98.1	96.7	3.3
3	Airborne	Roof	13	96.1	97	3
3	Airborne	Car	13	99.4	84.9	15.1
4	Airborne	Tree	25	91	84.4	15.6
5	Satellite	Tree	60	80	75	25

Where to find more pieces of information?

- **Peer-reviewed publication** @ Artificial Intelligence in Geosciences
- **Oral presentation** @ Esa-Nasa International Workshop on AI Foundation Model for EO (2025)
  Video presentation? Slides?
- **Hands-on workshop** @ AI for Good GeoAI seminar series (2025)
- **Python Package** with 50+ starts on Github

LangRS paper

LangRS Github

LangRS hands-on

Adapting a Pretrained Foundation Model for Secret Runway Detection using Sentinel-2 Imagery

Fatemeh Soleimanian, Master Student at GEOLab

Dr Polychronis Kolokoussis, Visiting Faculty at GEOlab

Maria A Brovelli, Professor, Head of GEOLab

02

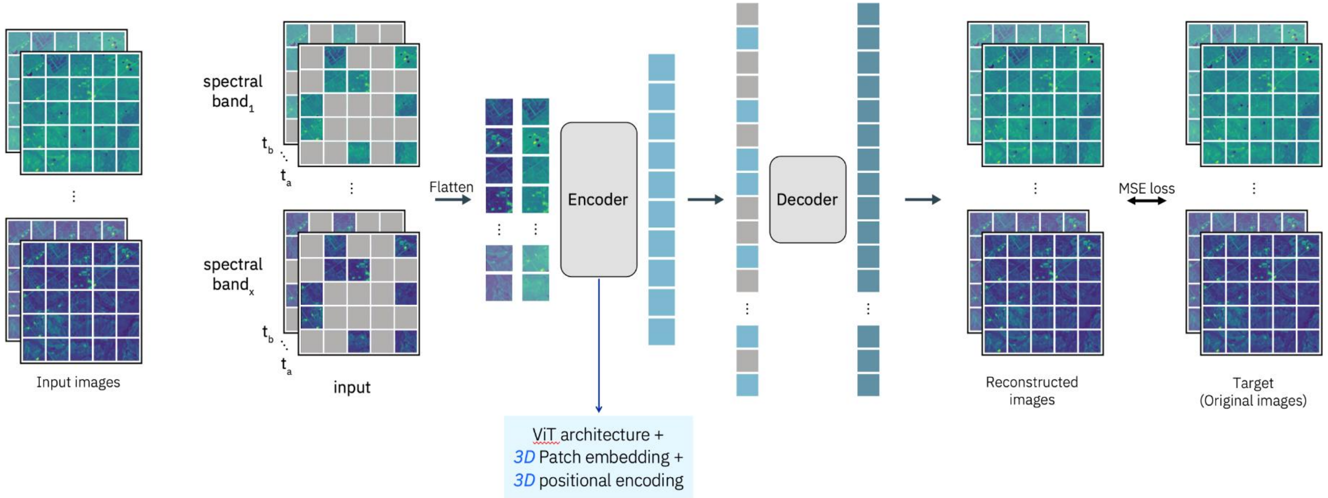
Illegal runways

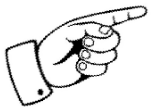
Source: https://plataforma.brasil.mapbiomas.org/pistas-de-pouso

Illegal runways hidden deep in the Amazon threaten biodiversity and fuel environmental crime. This study explores how GeoAI, powered by the Prithvi foundation model, can detect these clandestine sites on Sentinel-2 satellite imagery efficiently, accurately, and with a small number of labeled data.

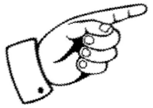
Technical Goals:

- Fine-tune a geospatial foundation model to detect runways from Sentinel-2 imagery with limited labelled data (~100 images).
- Examine the impact of different parameters of the model on the model performance and achieve its optimal performance by appropriate hyperparameter tuning.





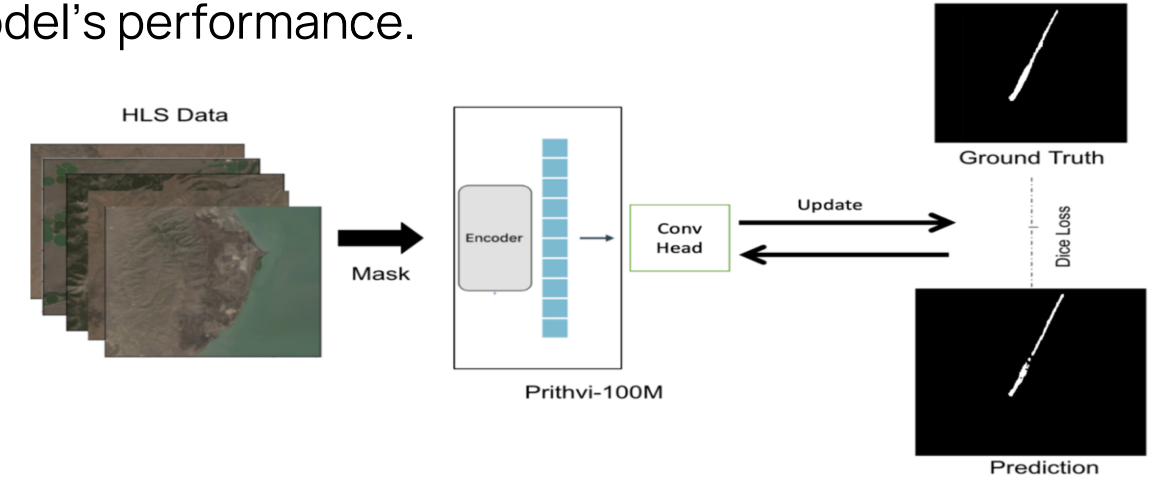
Selected Foundation Model: **IBM-NASA Prithvi-EO-1.0-100M**



Ground truth: Masks created based on runway contour lines, using GEO-SAM plugin in QGIS.
The contour lines are available on ITU's AI for good challenge on "ZINDI" website:
https://zindi.africa/competitions/geoai-amazon-basin-secret-runway-detection-challenge

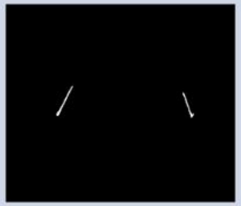
Results

Evaluation: Both numerical metrics (**IoU**, **mIoU**, **Accuracy**) and visual inspection on unseen datasets were used to assess the model's performance.

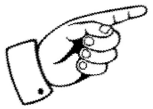


Numerical results of the 5 best experiments

Parameters	Exp8	Exp15	Exp16	Exp22	Exp23
Evaluation method	Iteration-based	Iteration-based	Iteration-based	Iteration-based	Iteration-based
Patch size	16	16	8	16	16
Loss function	Dice	Cross Entropy	Cross Entropy	Dice	Cross Entropy
number of training Dataset	67	67	67	40	40
mIoU	**79.25**	**78.5**	**79.12**	**80.14**	**79.66**
Accuracy	82.37	84.51	84.13	84.87	88.09

Image	Ground truth	Experiment 8
		
		

Validation: For validation, we run the obtained weight of the training process, on new unseen data that haven't been used in training loop.



- The Prithvi FM, fine-tuned with limited labelled data, achieved strong segmentation performance on a quite difficult RS task.
- The results highlight the Prithvi model 's efficiency, particularly valuable in scenarios with limited, imbalanced datasets.
- Optimal hyperparameter tuning significantly improved accuracy.

Foundation Models for Land Cover Understanding: NASA Prithvi FM and MOLCA

Qiongjie Xu,
PhD Student

Seyederfan Eshghollahi,
Master Student in Geoinformatics Engineering

Maria A Brovelli,
Professor, Head of GEOLab

03

GHRLC Maps

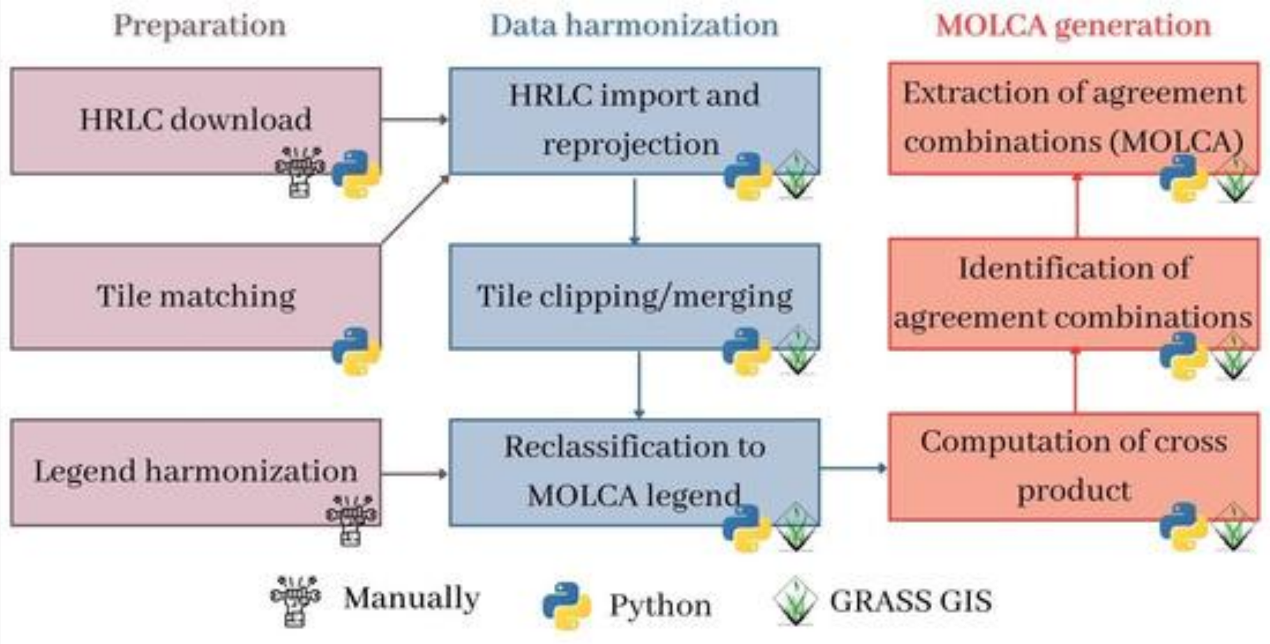
✓ Global high-resolution land cover maps are not homogeneous in terms of spatial and temporal resolution, nor in their classification legends. The mismatch in legends is mainly due to differences in their intended field applications. To provide a benchmarking tool, we have created the Map of Land Cover Agreement, currently available for three major regions of the world.

Dataset name	Provider	Spatial resolution	Temporal resolution	Land cover focus
Dynamic World (DW)	World Resources Institute Google	10 m	Near real-time from 2015-06-27 to present	General (9 classes)
ESRI LULC	ImpactObservatory, Microsoft, and Esri	10 m	Annual (2017 - 2023)	General (10 classes)
Forest/Non-Forest (FNF)	JAXA-EORC	25 m	Annual (2007 - 2010)	Forest (3 classes)
FROM-GLC10	Tsinghua University	10 m	2017	General (10 classes)
FROM-GLC30	Tsinghua University	30 m	2010, 2015, 2017	Genera l (10 classes)
Global Forest Cover (GFC)	JRC	10 m	2020	Forest (1 class)
GHS-BUILT-S R2023A	JRC	10 m	2018	Built-Up surface proportion (continuous 0–100%)
GHS-BUILT-S1 2016	JRC	20 m	2016	Built-Up (2 classes)
Global 30m impervious-surface dynamic dataset (GISD30)	CAS	30 m	Interval-encoded (per pixel): before 1985, 1985-1990, 1990-1995, 1995-2000, 2000–2005, 2005–2010, 2010-2015, 2015-2020	Built-Up (1 classes)
GlobeLand30 (GL30)	NGCC	30 m	2000, 2010, 2019	General (10 classes)
Global Land Cover Estimation (GLanCE) v001	BU/EE/NASA ES/USGS EROS	30 m	Annual (2001 - 2019)	General (7 classes)
Global 30m land-cover (dynamics monitoring) product with a fine classification system (GLC-FCS30/D)	CAS	30 m	Every 5years (1985, 1990, 1995) and Annual (2000 - 2022)	General (35 classes)
Global Surface Water (GSW)	JRC	30 m	Annual (1984-2021)	Water (3 classes)
Global Urban Footprint (GUF)	DLR	12 m	2011	Built-Up (2 classes)
Global annual wetlanddataset at 30mwith a fine classficiationsystem (GWL-FCS30/D)	CAS	30 m	Annual (2000 - 2022)	Wetland (8 classes)
MapBiomas	SEEG and Climate Observatory	30 m	Annual (1985 - 2023)	General (6 classes)
Tree canopy cover	Hansen/UMD/Google/USGS/NASA	30 m	2000	Tree canopy cover proportion (continuous 0–100%)
World Settlement Footprint (WSF)	DLR	10 m	2015, 2019	Built-Up (2 classes)
WorldCover	ESA	10 m	2020, 2021	General (11 classes)

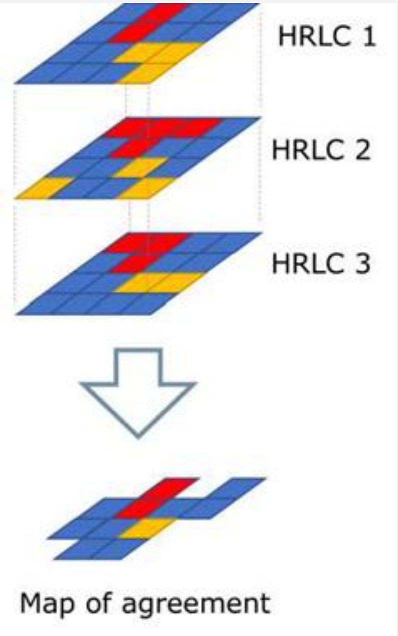
Map of Land Cover Agreement (MOLCA)

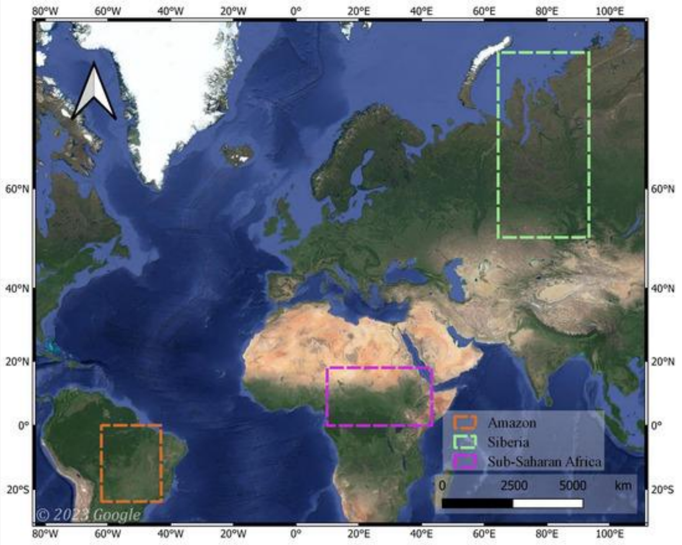
- Created by integrating multiple existing high-resolution land cover datasets through a consensus-based approach
- *Spatial coverage*: Sub-Saharan Africa, the Amazon, and Siberia
- *Spatial Resolution*: 10 m
- *9 Land cover classes:* Shrubland, Grassland, Cropland, Built-up, Bareland, Permanent ice and snow, Water, Wetland, Forest
- OA = 96%

⬇️ Schema of the MOLCA generation procedure

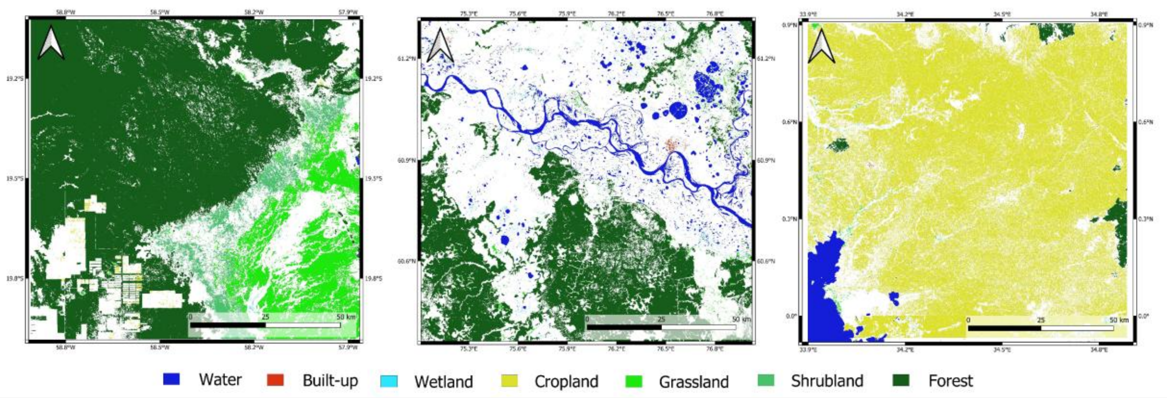


⬇️ Illustration of extraction of the map of agreement





⬆️ Regions where MOLCA data are produced



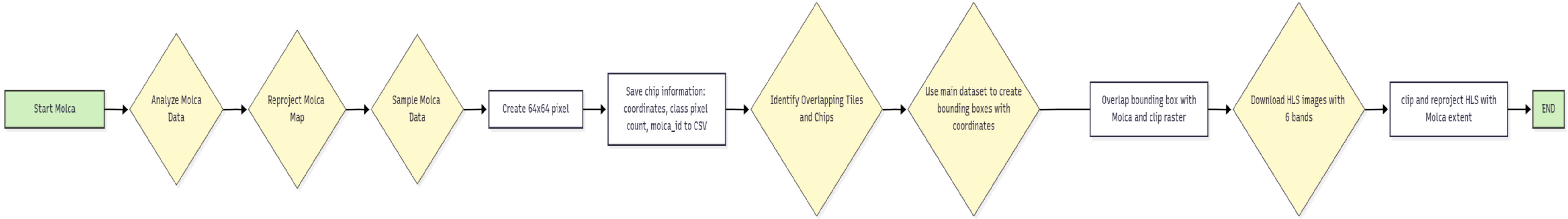
⬆️ Example of MOLCA tile in the Amazon region (**left**), in Siberian region (**center**), and African region (**right**).

Reference: https://doi.org/10.3390/rs15153774

Preprocessing and Integration of MOLCA and HLS Data for Land Cover Analysis

Class	Number of chips
shrubland	90642
Grassland	337547
Cropland	339001
Wetland	927
Bare land	2966468
Built-up	515
Water	2184682
Permanent Ice and snow	32118
Forest	6320288

- The MOLCA dataset includes over 2000 large TIFF files from three regions.
- It is reprojected from WGS84 to UTM to match HLS data and ensure consistent chip extraction.
- 64x64 pixel chips are sampled to minimize null pixels and reduce class overlap.
- Special thresholds are applied for underrepresented classes like Built-up (Class 13).
- Overlapping chips from tile redundancies are removed using a greedy algorithm that retains the most saturated samples.
- An HLS integration pipeline downloads, stacks, and clips imagery to match the Molca chips.
- Cloud coverage is managed by selecting cloud-free images or applying cloud masks when necessary.



Exploring Geospatial Foundation Models in urban downstream applications and with hyperspectral data

Julia Leonardi,
PhD Student

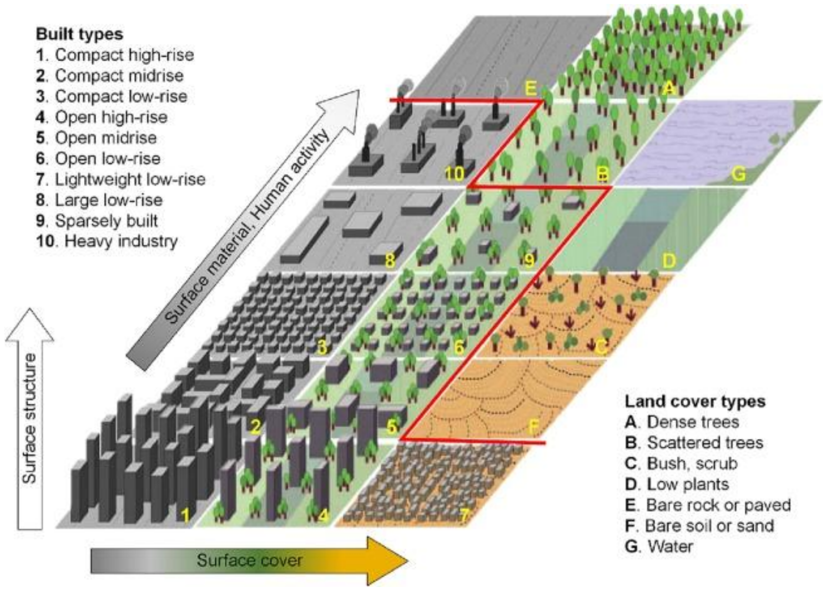
Alberto Vavassori,
Assistant Professor

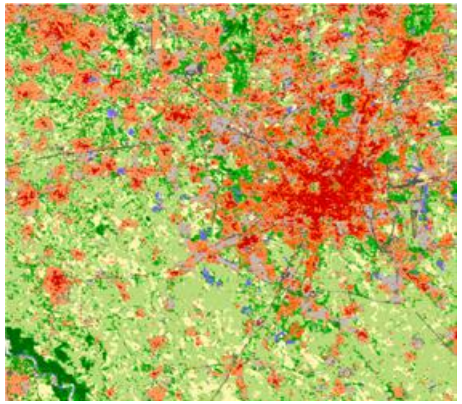
Maria A. Brovelli,
Professor, Head of
GEOLab

04

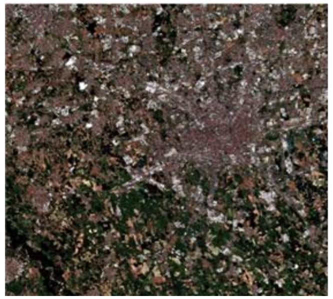
Using Geospatial Foundation Models for Local Climate Zones (LCZ) Mapping

- An experimental activity building on earlier work with the Random Forest algorithm, testing how well the **Prithvi-1.0 Geospatial Foundation Model (GFM)** could map **Local Climate Zones (LCZs)** in Milan
- Aimed to explore how GFMs can support real-world tasks like **urban environment monitoring**
- A key challenge was the **limited compatibility of the available LCZ dataset**, which made fine-tuning difficult
- With a better-fitting dataset, **the model showed** noticeably improved performance and **real potential**
- Future work could involve building a more suitable LCZ dataset, **testing other GFMs**, and exploring **hyperspectral data**

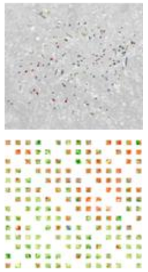




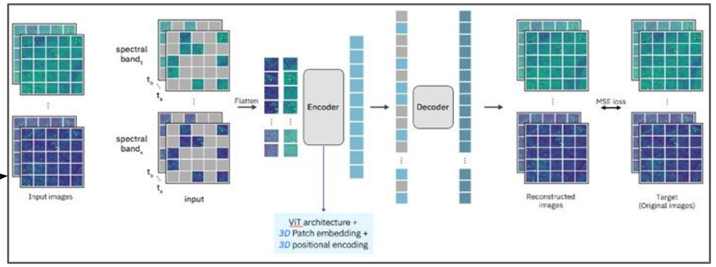
Classified S-2 scene



S-2 scene of Milan



Training labels

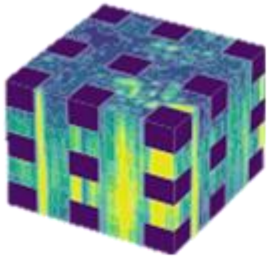


Prithvi-1.0 (100M)

Enhancing Geospatial Foundation Models with Hyperspectral Satellite Imagery – an ongoing review

- An ongoing effort focused on compiling a **comprehensive and exhaustive review** of what **feature extraction** methods for **hyperspectral images** can be **integrated into a GFM** architecture.

- Identifying **gaps in spectral dimension processing** for **efficient feature extraction** for **large-scale** model **pre-training**



Identify gaps in spectral integration → Explore effective HSI integration methods → Apply these methods to enhance GFMs → Benchmark performance on a range of key tasks

Future work on Geospatial Foundation Models and the Hyperspectral modality

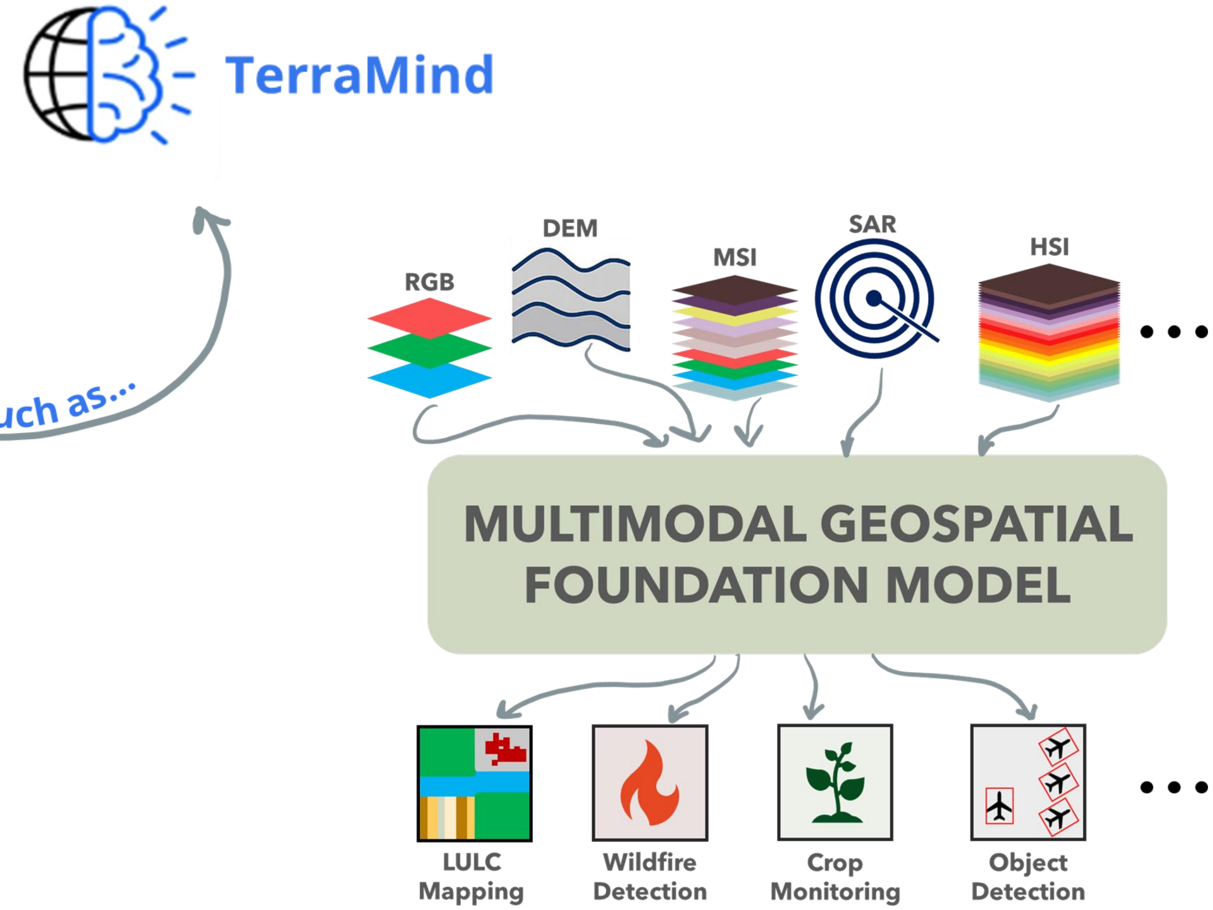
Not Only HSI – Toward Truly Multimodal Geospatial Models

The ultimate goal is a foundation model that **natively handles diverse geospatial data**.

🎯 **Next Steps**

- Investigate **sensor-agnostic** architectures
- Explore **cross-modal pretraining strategies** and work on existing models
- Benchmark models across varied tasks

HSI is just one piece of a larger multimodal puzzle. Our interests lie in GFMs that are not only robust to new tasks and regions but also capable of reasoning across sensor types and other input modalities.



Conclusions

05

What's ahead?

- Pushing forward the initiatives already in motion
- **Course on Geospatial Foundation Models** for 20 students from Politecnico di Milano, enrolled in Computer Science and Geoinformatics Engineering programs
- **2 PhD students** (one in the field of Computer Science Two PhD students — one in Computer Science and one in Geoinformatics — will collaborate on supervised, contrastive, and self-supervised learning approaches, with a strong focus on **continual learning for real-time updates and streaming geospatial data, including Earth Observation.** The call is expected to open in **November/December**.
- **Welcoming collaboration opportunities in this field.** Our journey is ongoing — the best is yet to come.

Thanks for your attention

Maria Antonia Brovelli
maria.brovelli@polimi.it