

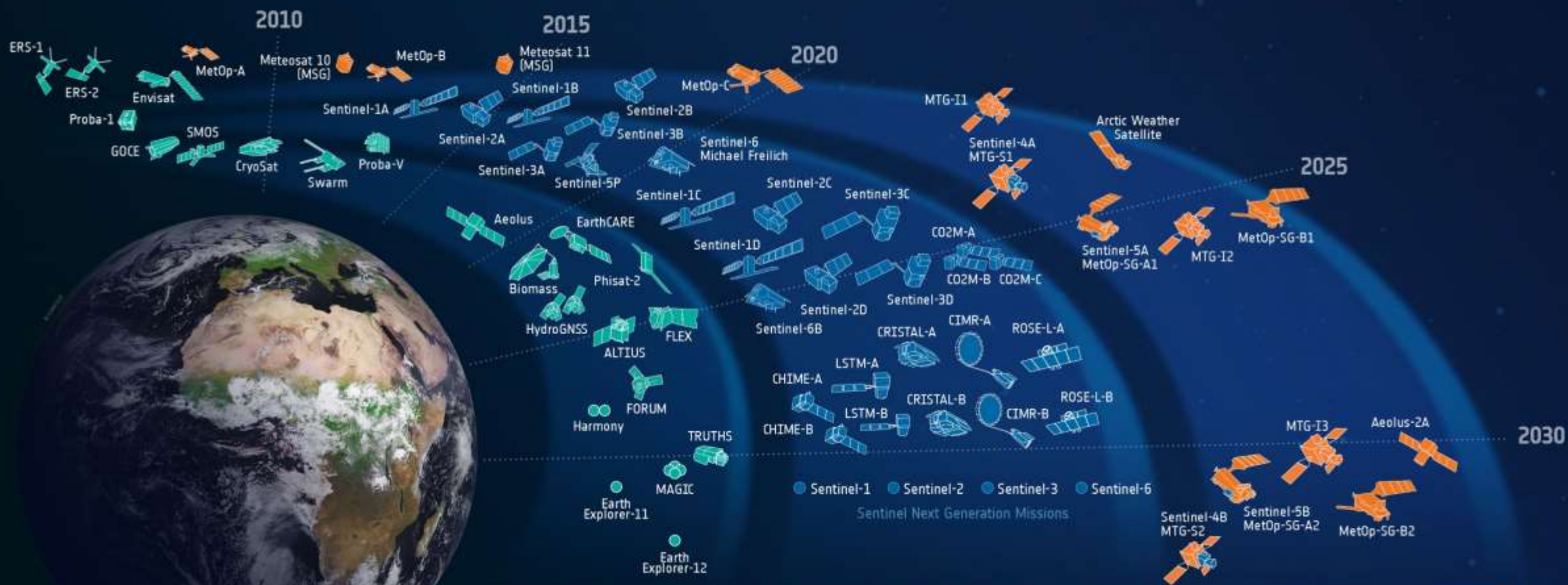


Machine Learning for Earth Observation and Beyond

Xiaoxiang Zhu

Technical University of Munich, Germany

ESA-DEVELOPED EARTH OBSERVATION MISSIONS



Science



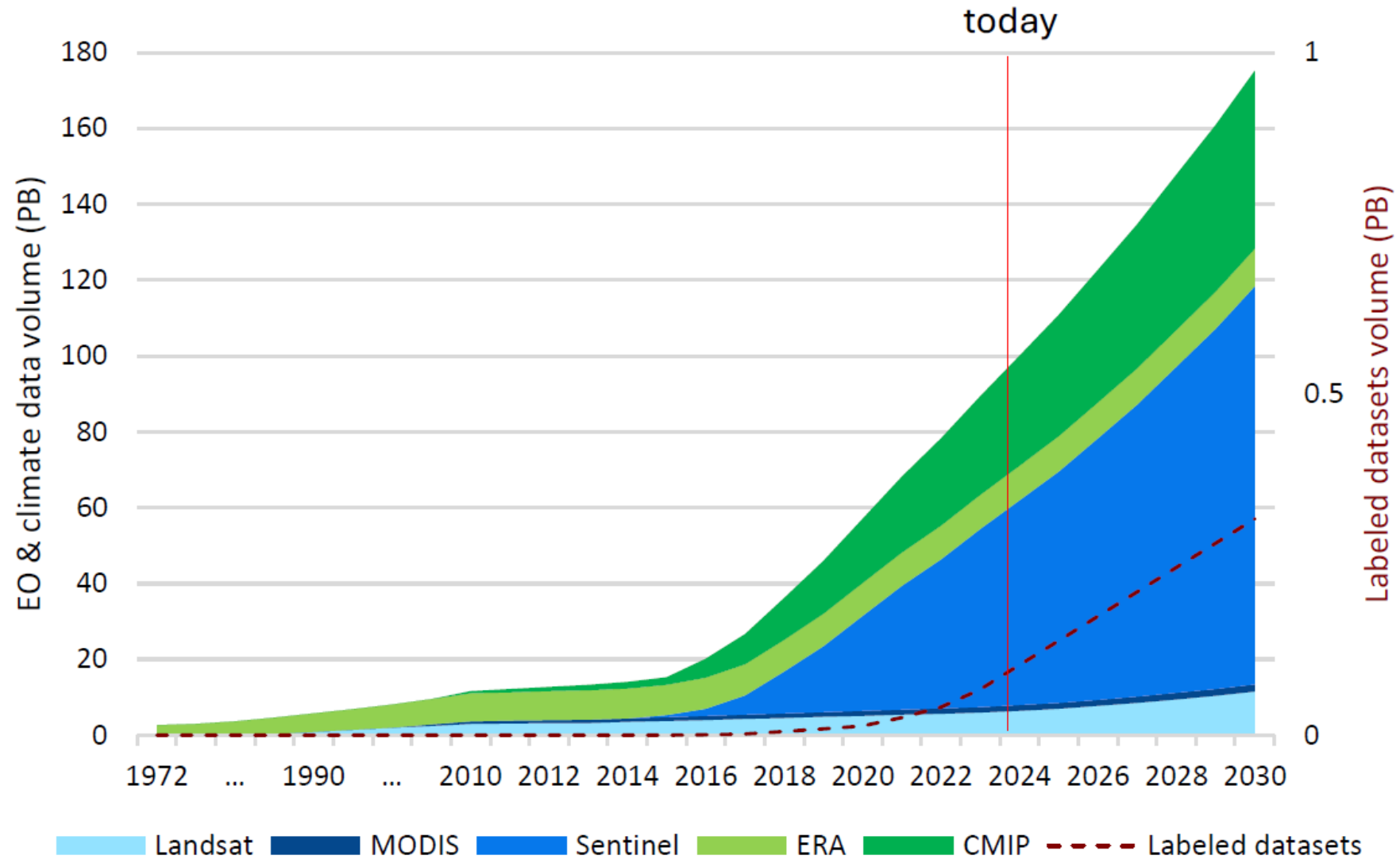
Copernicus

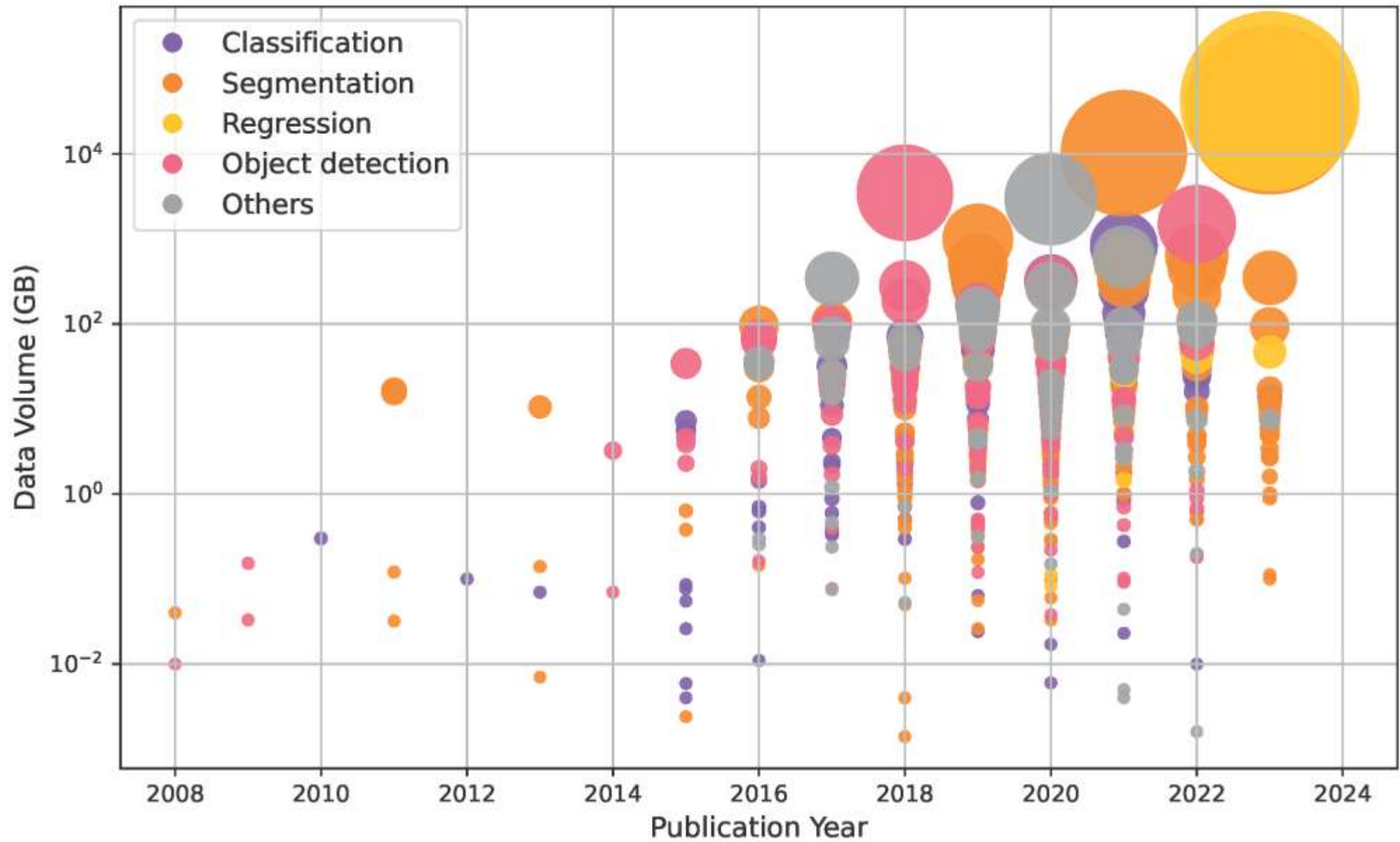


Meteorology



The Open Big Earth Data





AI and Data Science in Earth Observation

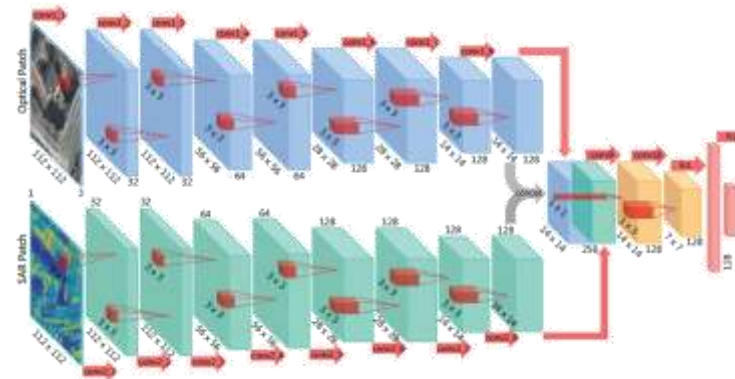
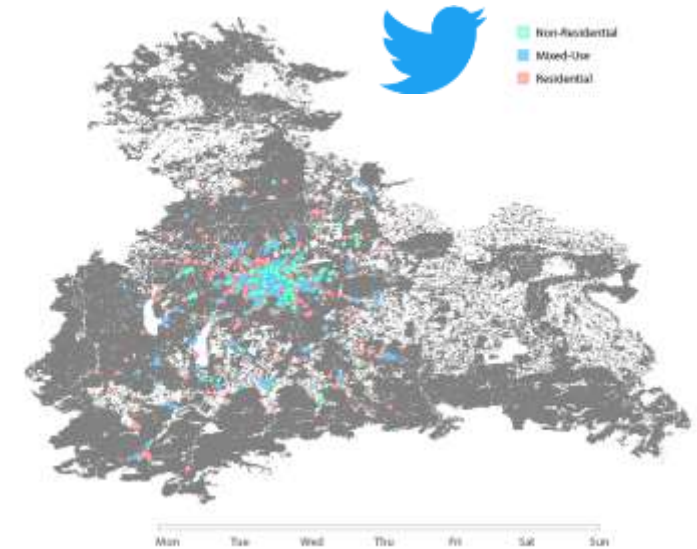
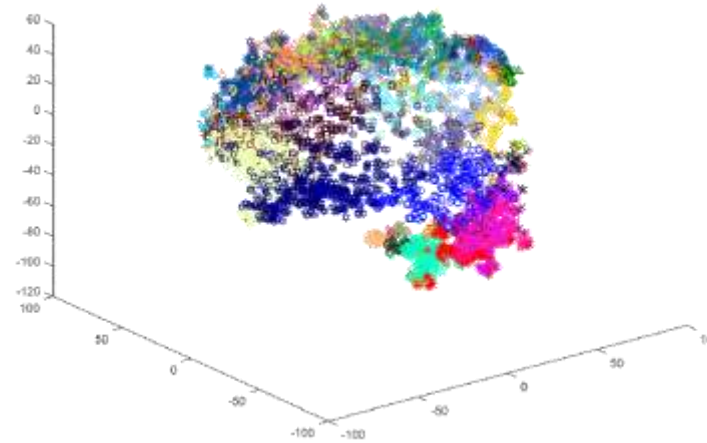
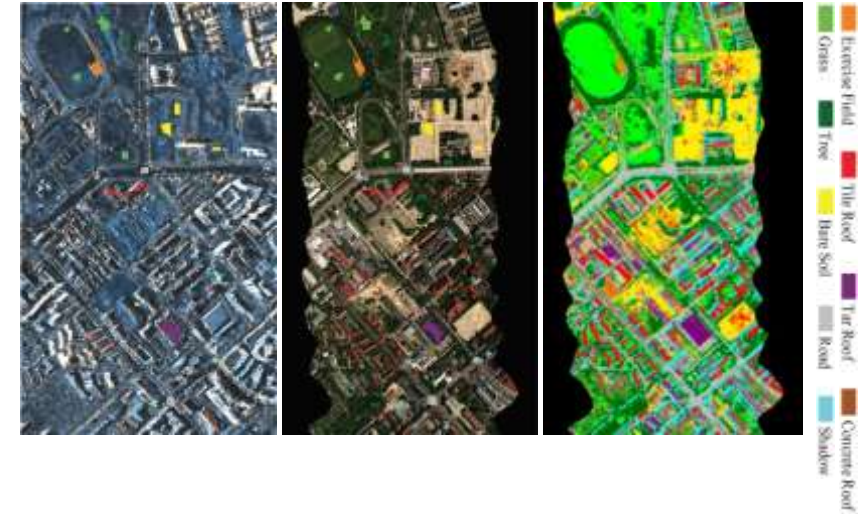
Explorative Signal Processing Methods

Date Fusion

Information Mining

Machine Learning/Deep Learning

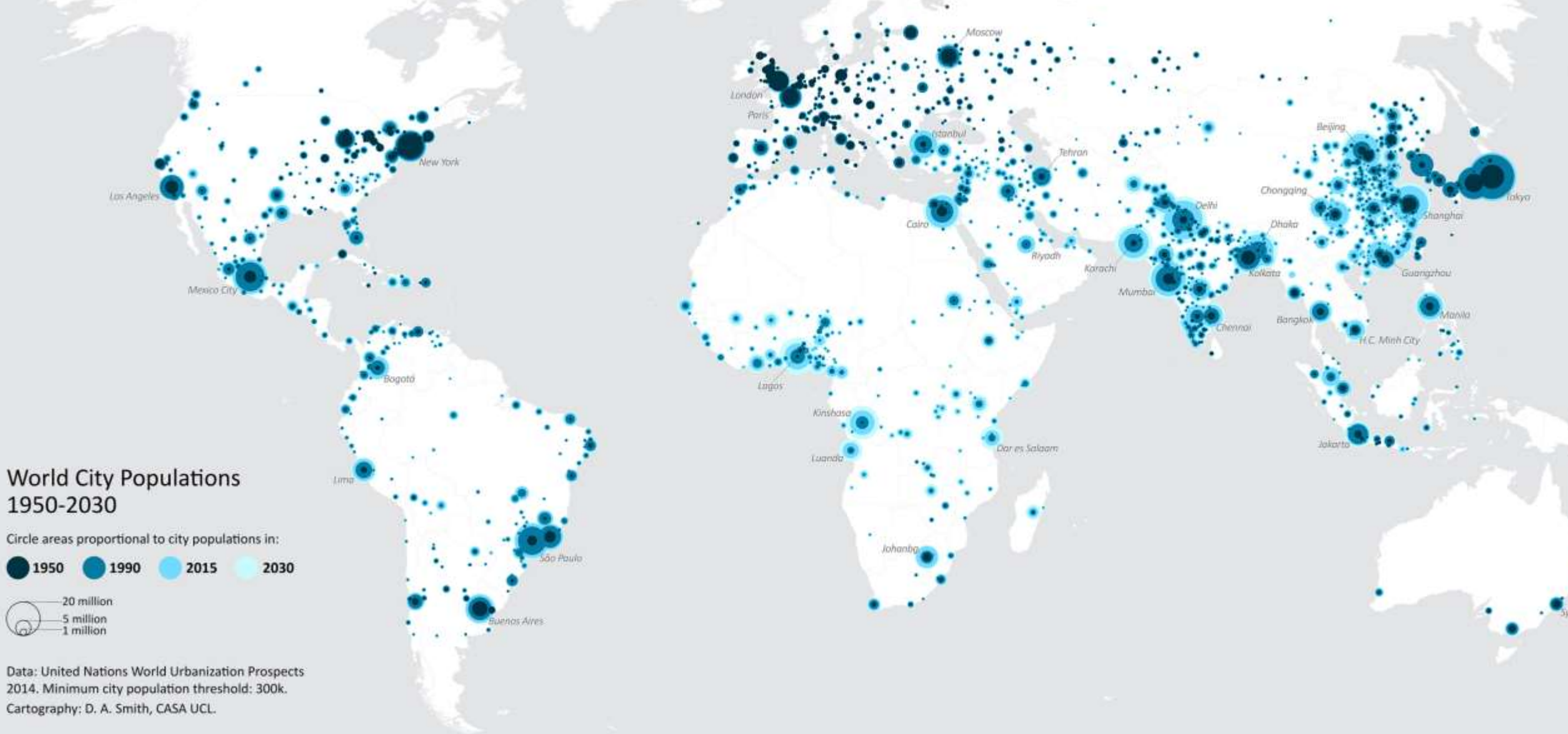
Big Data Management and HPC



What can ML₄Earth offer?

A few motivating examples of tackling societal grand challenges

Urban Growth Happens Mostly in Developing Areas



Open Data for Urban Research?

> **4.1 Billion buildings** in the world (UN, 2019)

604 Million building footprints in OSM < **15%**

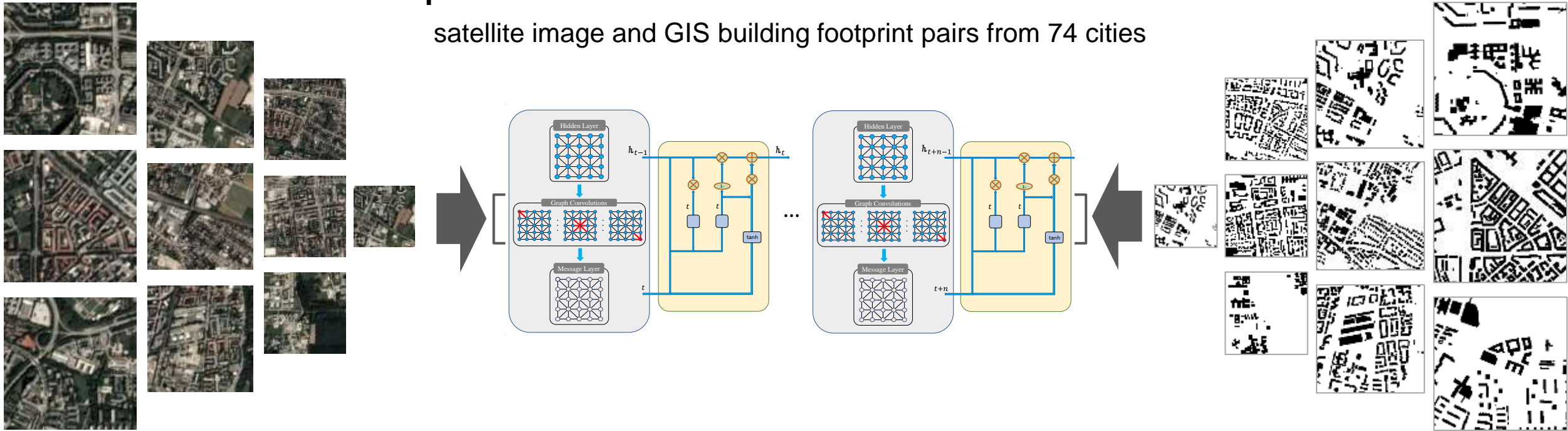
Only **3%** buildings in OSM have **height information < 0.5 %**



Building Footprint Extraction from *NewSpace*-Satellite Images



A Graph Convolutional Recurrent Neural Network trained with satellite image and GIS building footprint pairs from 74 cities



A: Data acquisition

1. Urban detection
2. Image querying

B: Analysis-ready data

3. Radiometric calibration
4. Mosaicing

C: Machine learning pipeline

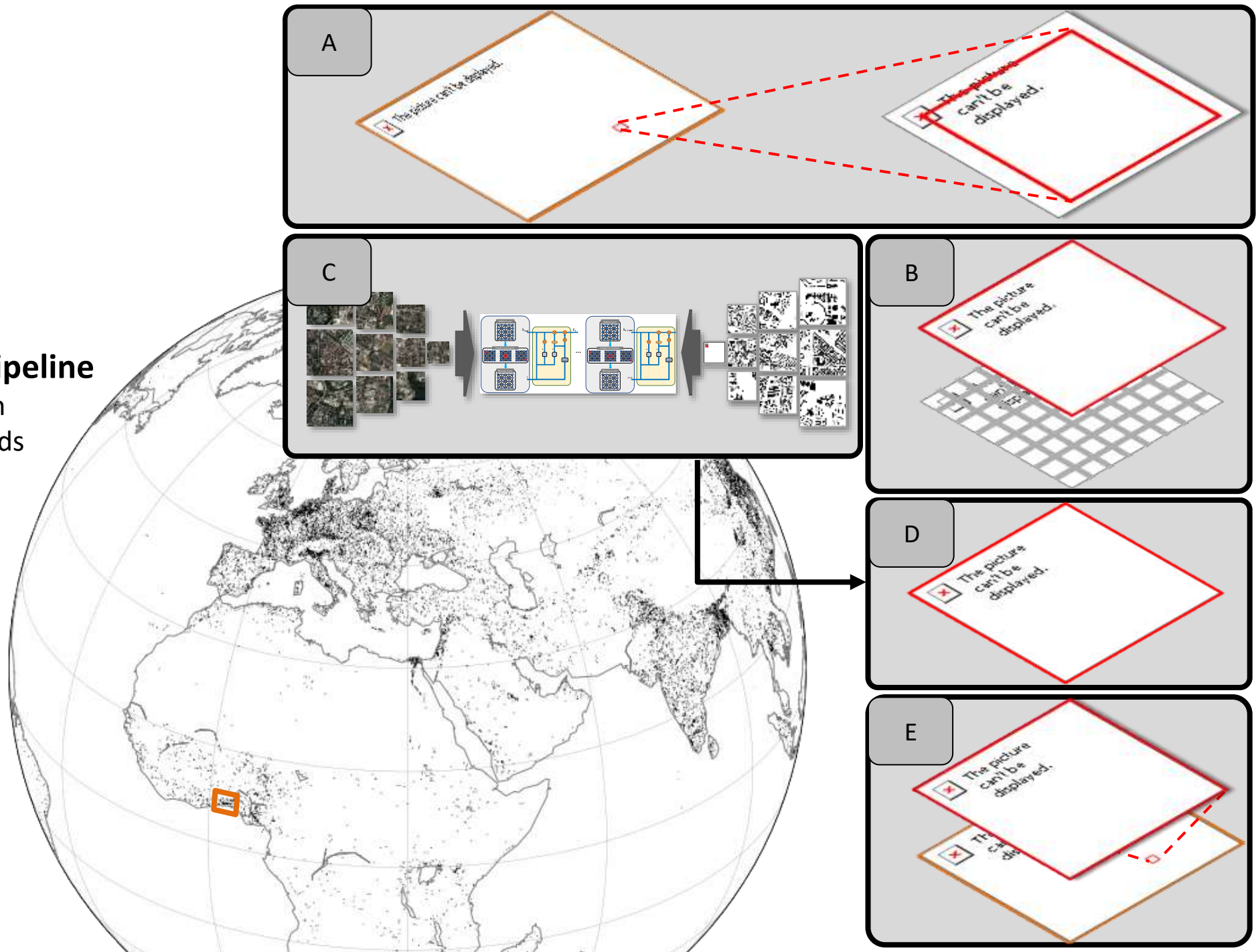
5. Training data preparation
6. Machine learning methods
7. Model training

D: Inferencing

8. Parallel processing
9. Model averaging

E: Post-processing

10. Editing/ Filtering



Data Overview

Sensor: Planet Doves

of images: 779072

of 0.2 degree tiles: 45065



Global Buildings

Red: Google Green: Global OpenBuildingMap(ours) Blue: OpenStreetMap

Two examples

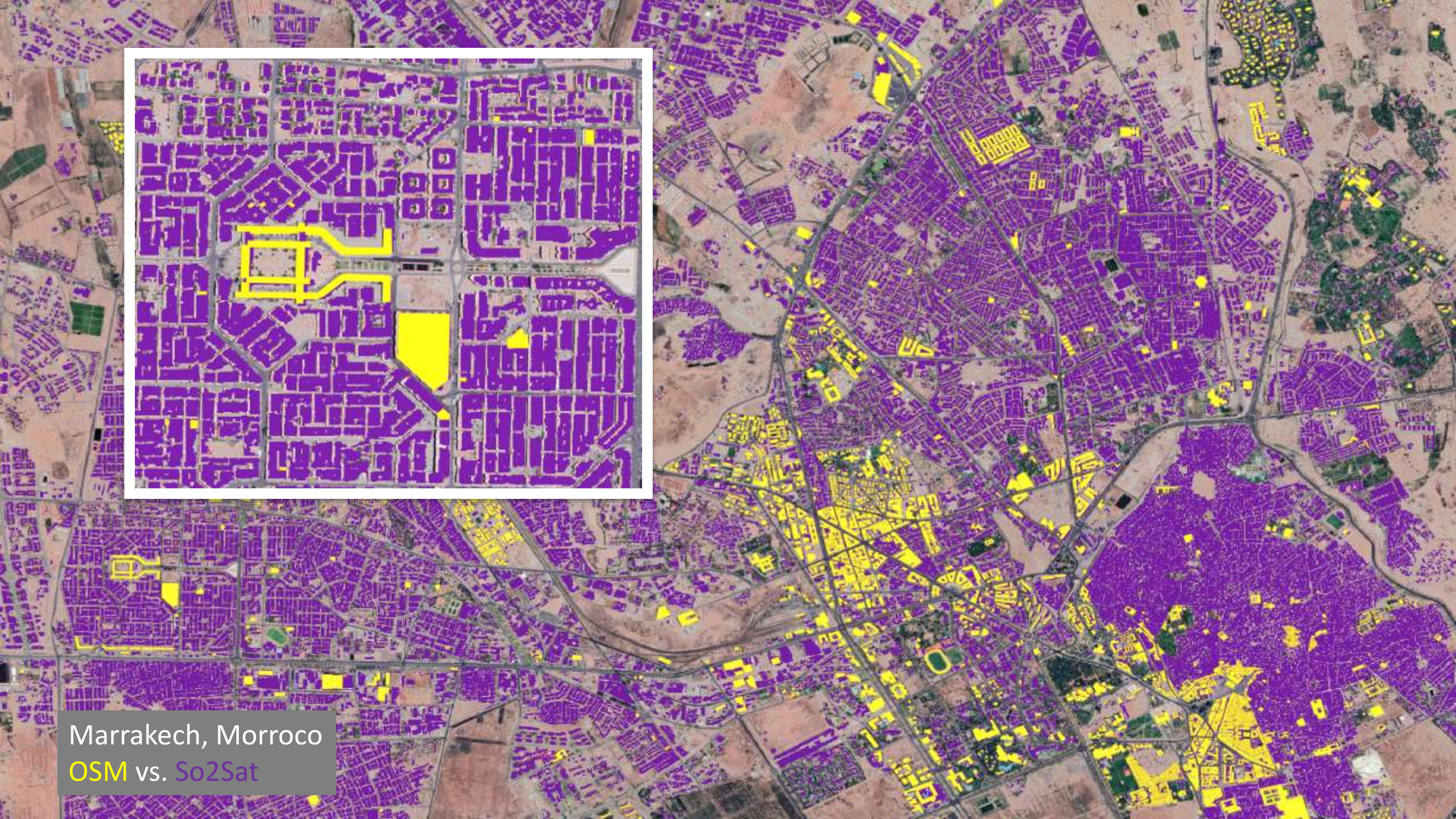
- Marrakech, Morocco
- Cairo, Egypt



Cairo, Egypt

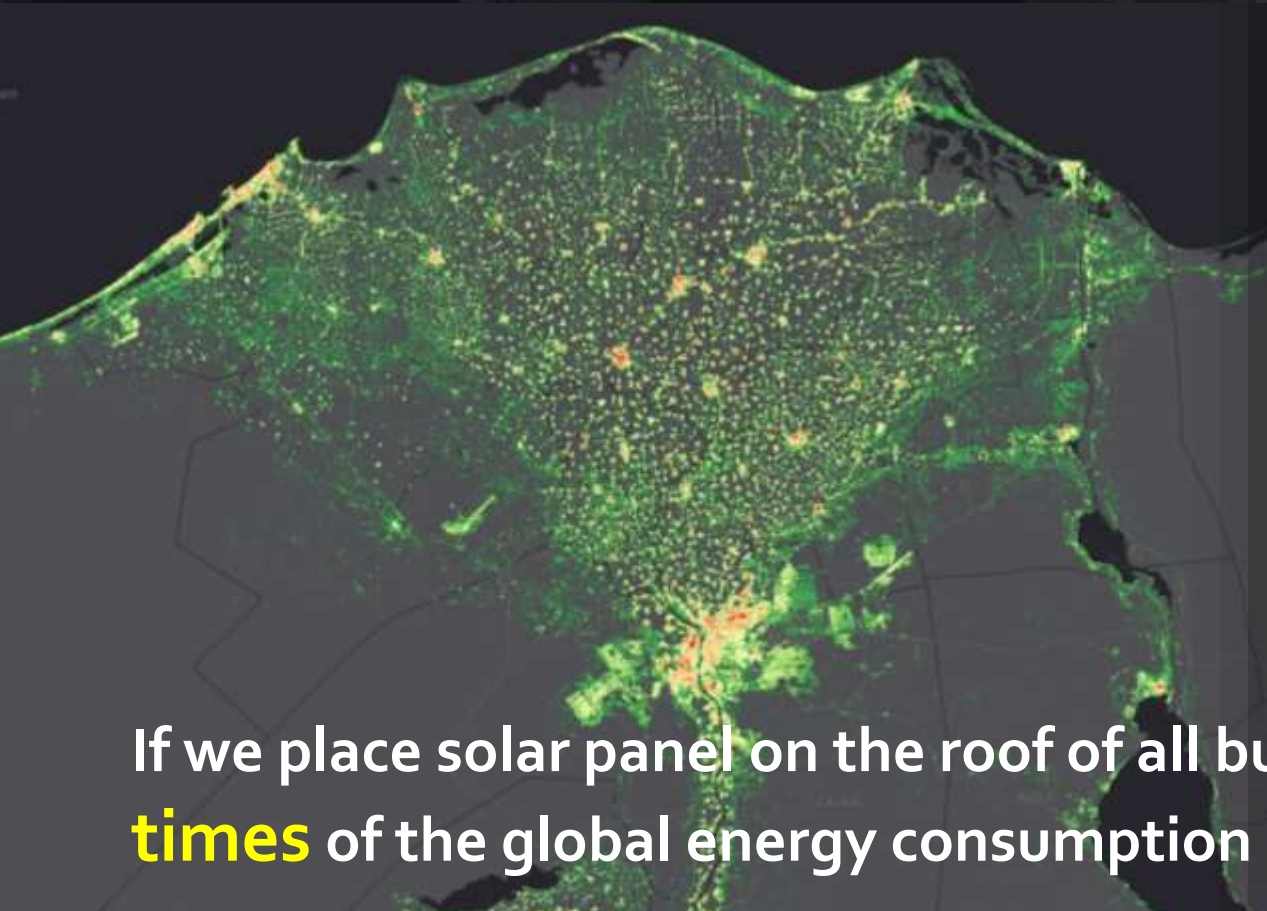
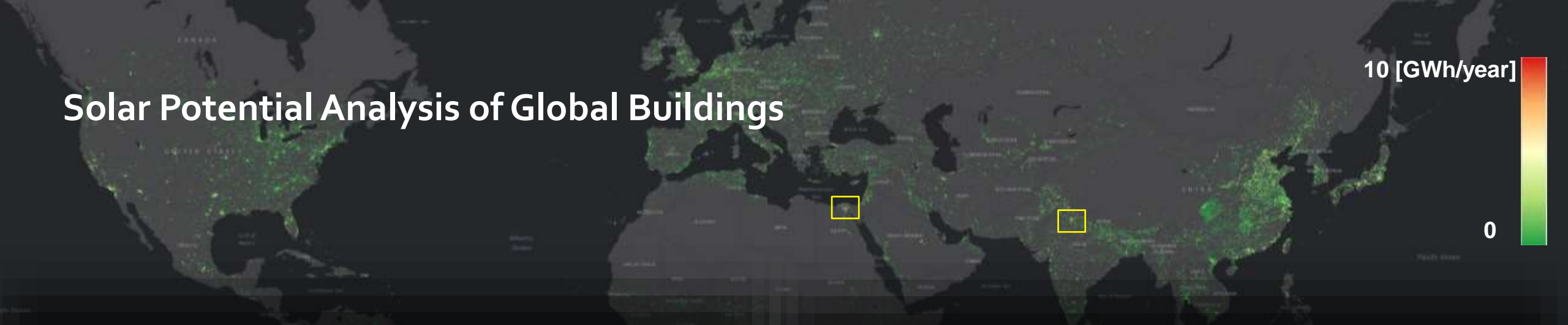
OSM vs. Ours





Marrakech, Morocco
OSM vs. So2Sat

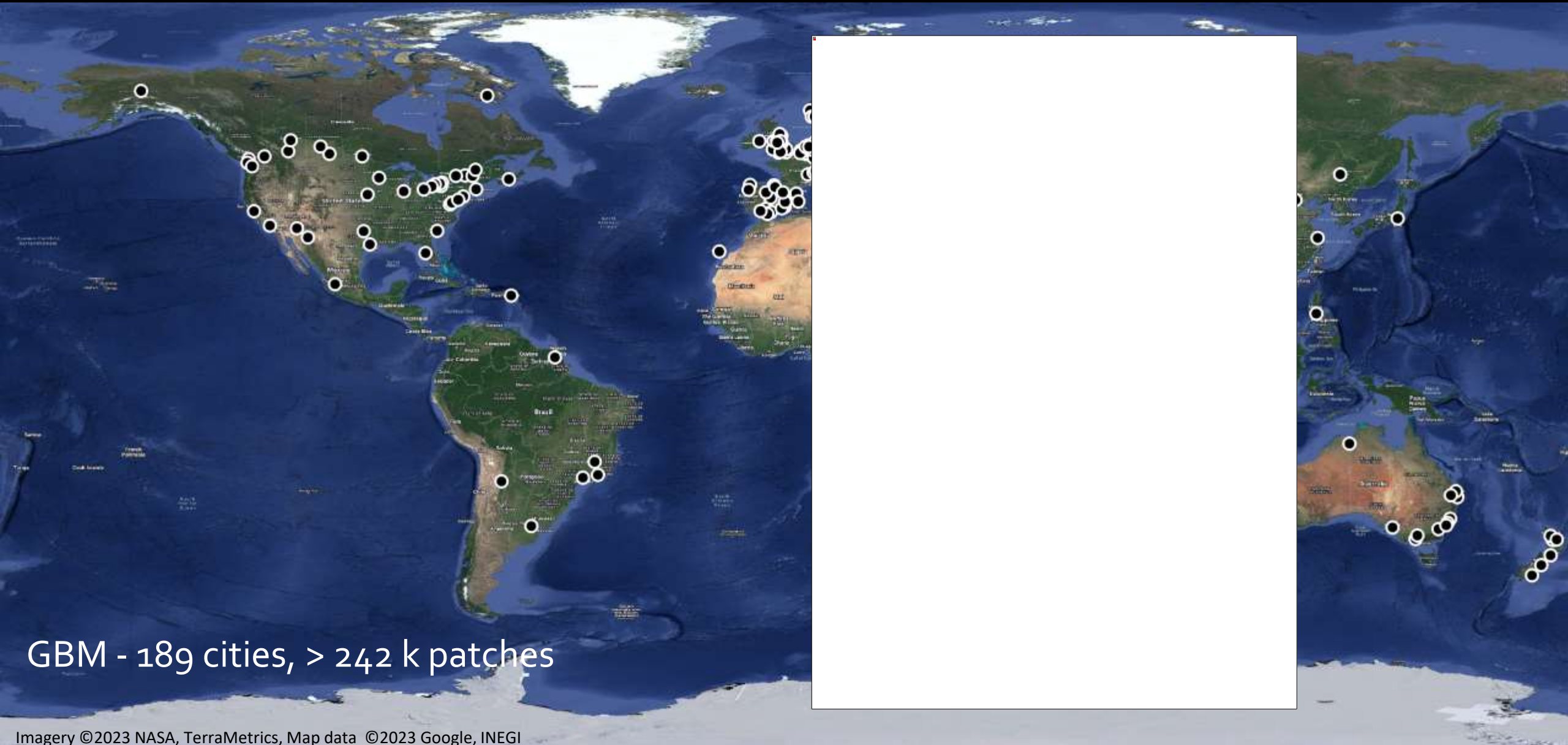
Solar Potential Analysis of Global Buildings



If we place solar panel on the roof of all buildings, it is possible to supply **1.1 to 3.3 times** of the global energy consumption of 2020!

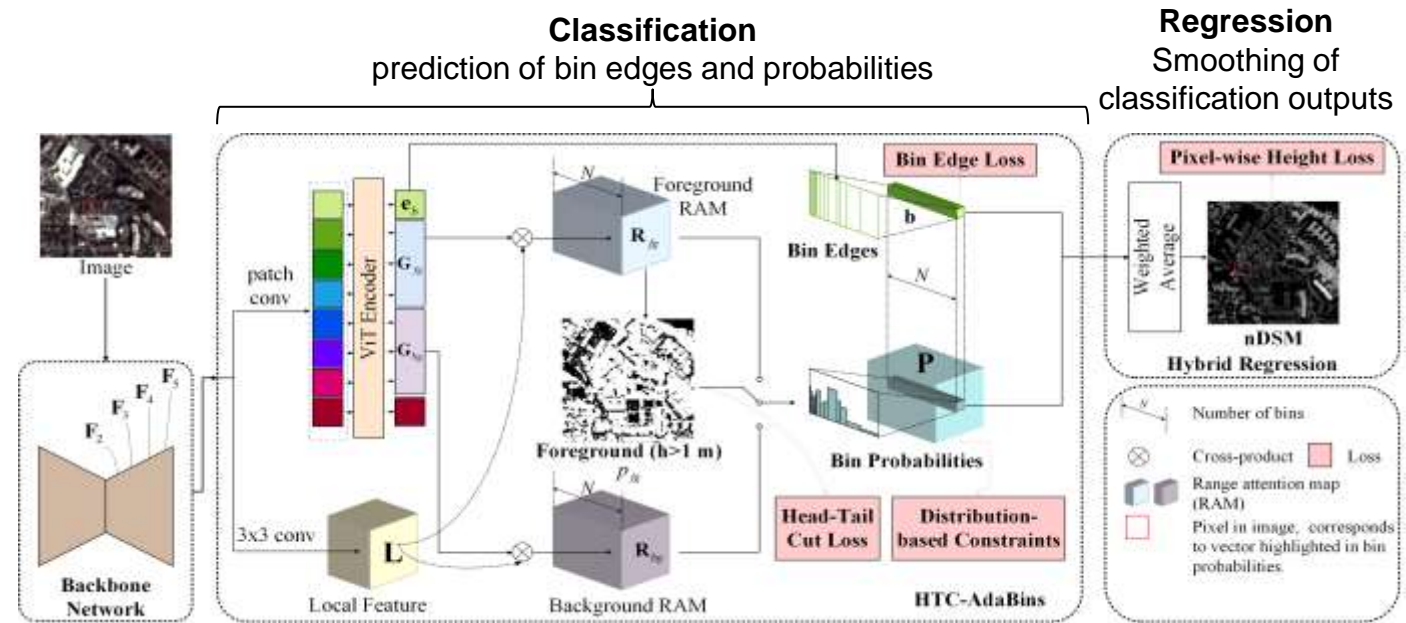
Zhu et al., under review

Global 3D Building Modeling from Single Remote Sensing Images



HTC-DC Net: Monocular Height Estimation from Single Remote Sensing Images

- Ill-posed problem
 - Use of ViTs for enforcement of global consistency
 - Classification-regression paradigm
 - Distribution-based method
- Long-tailed distribution of the height values
 - Head-Tail Cut (HTC) for distinct treatment of background and foreground
- Unconstraint distributions lead to large errors
 - Distribution-based constraints (DC) as regularization



Test Building Height Accuracy of 3.8m on 21 cities across the globe Global inference will be finished by the mid of May!

City Center, Vancouver, Canada



TRISHNA

SGB

LSTM-A

LSTM-B

Monitoring the Antarctic Coastline



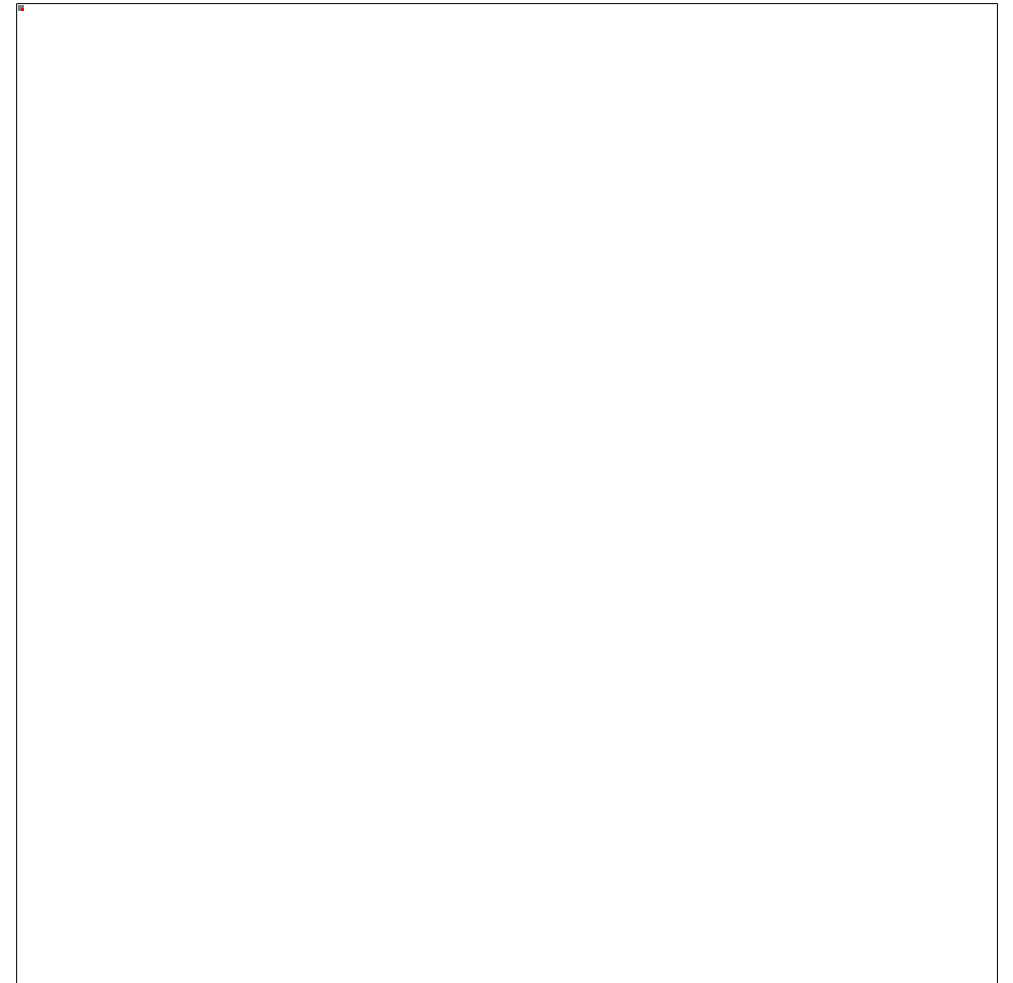
Sentinel 1 GRD @ 40m

Jun 2017 – Dec 2018

Dense Annotations: Land / Water

~ 510,000 km² Training

~ 220,000 km² Validation



HED-UNet: Combined Segmentation and Edge Detection for Monitoring the Antarctic Coastline

- coastline detection as a dual inference task.
- predictions at multiple resolutions
- a hierarchical attention mechanism

Results – Qualitative

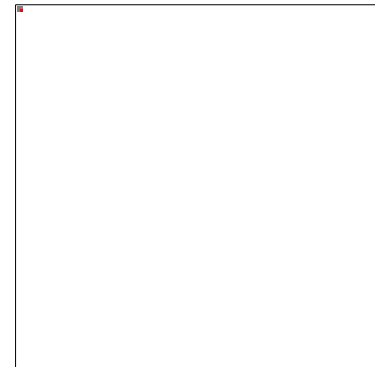
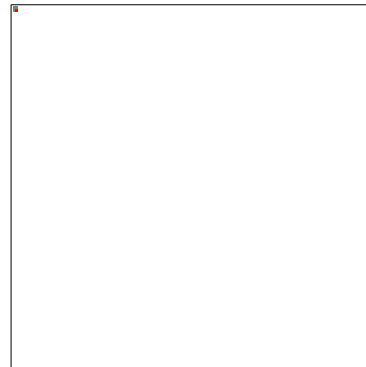
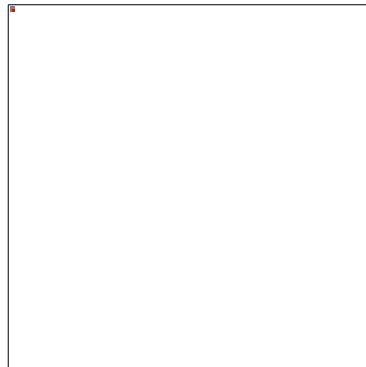
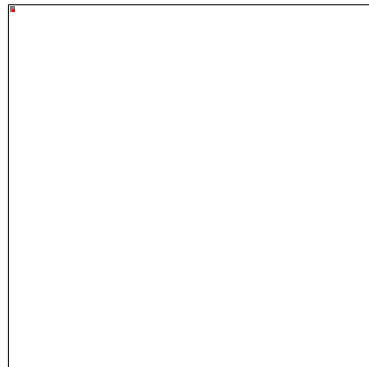
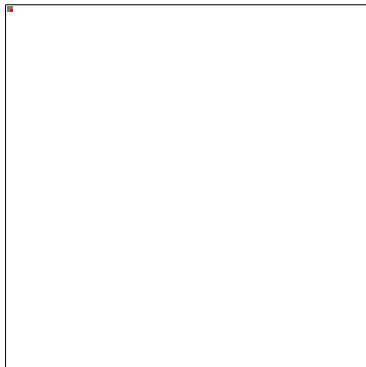
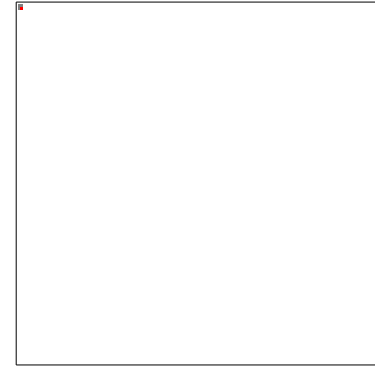
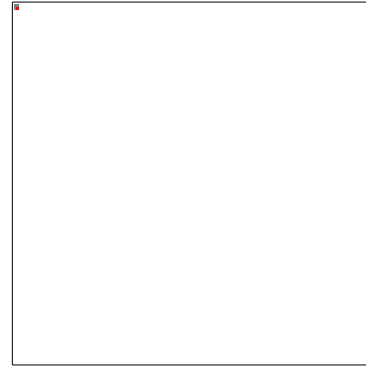
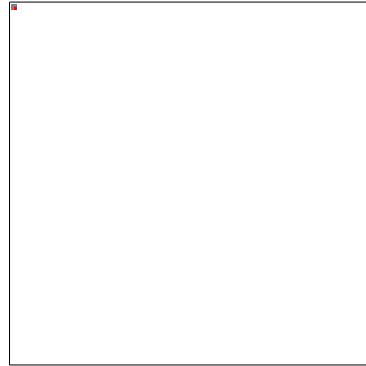
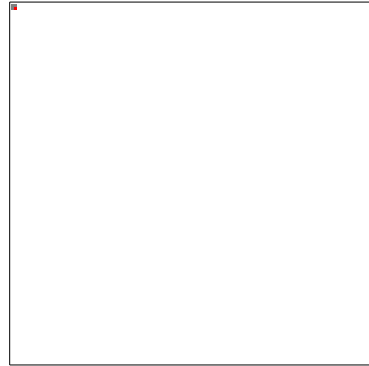
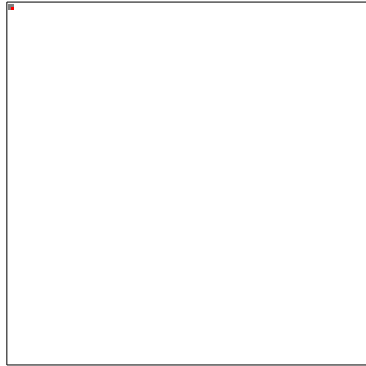
Input

Ground Truth

UNet4

UNet6

HED-UNet6



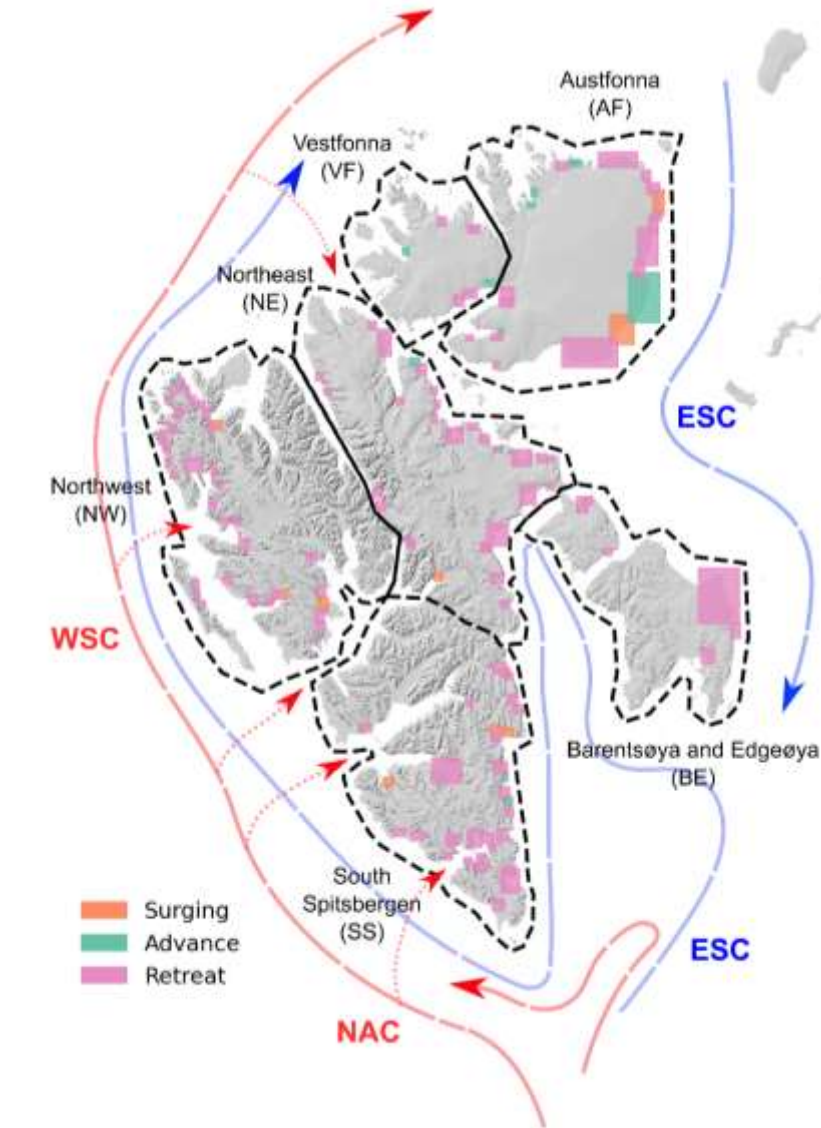
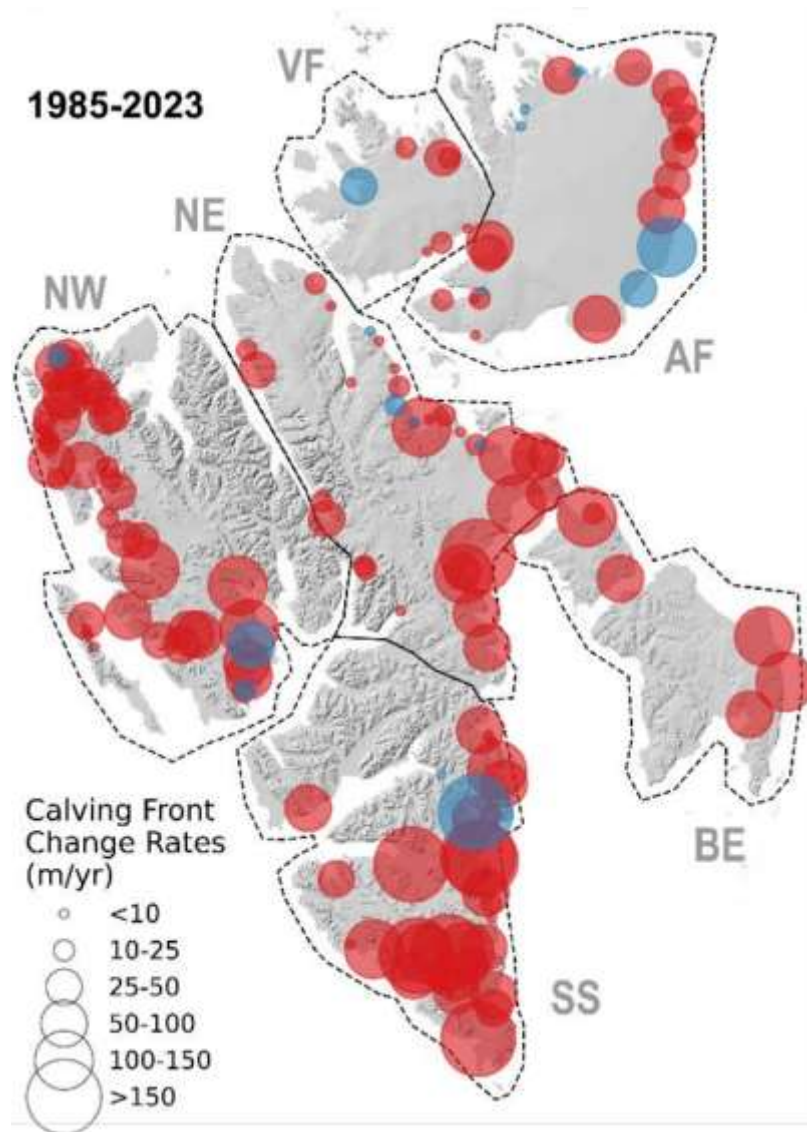
A High-Resolution Calving Front Data Product for Marine-Terminating Glaciers in Svalbard

Features:

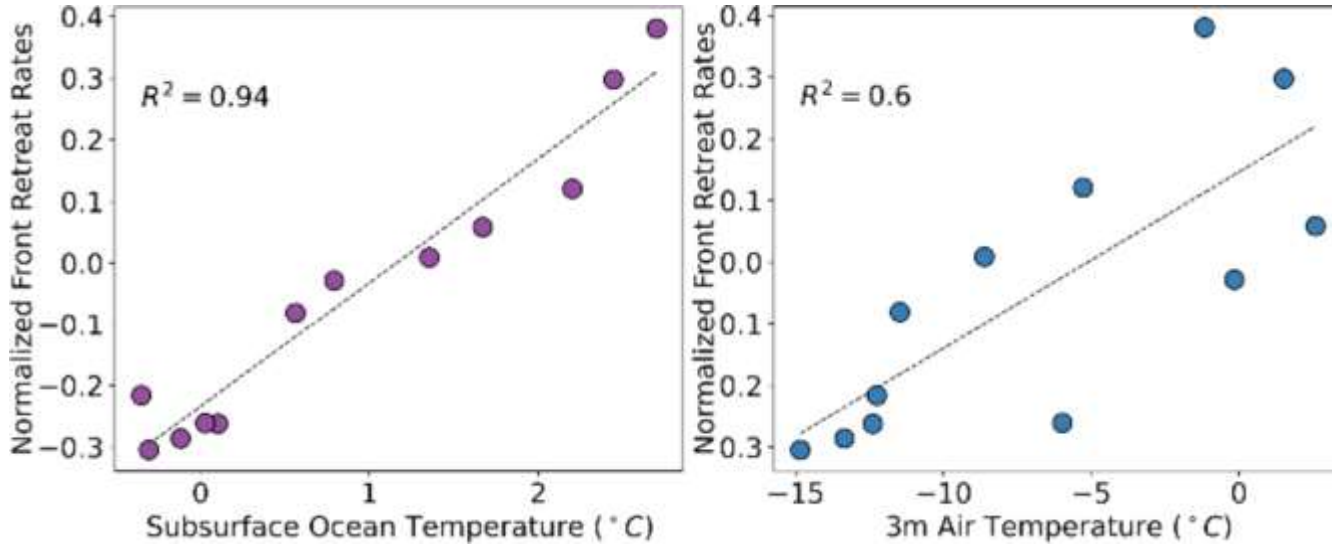
- 149 marine-terminating glaciers in Svalbard
- 124919 glacier calving front positions during the period of 1985-2023



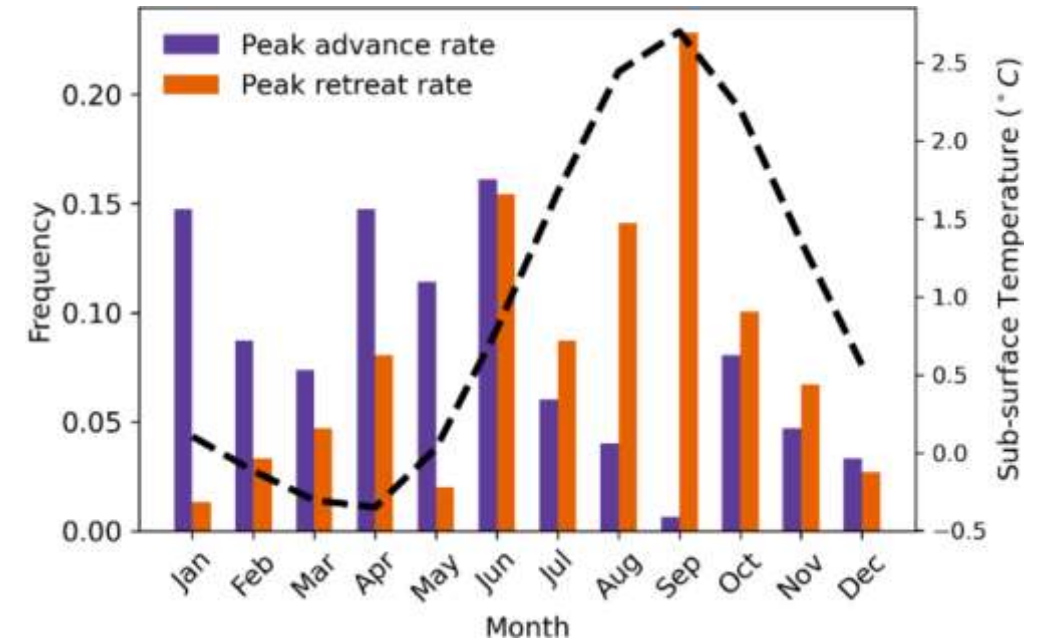
Calving Front Change Rate & Spatial distribution of different glacier types



Correlating with Environmental Factors



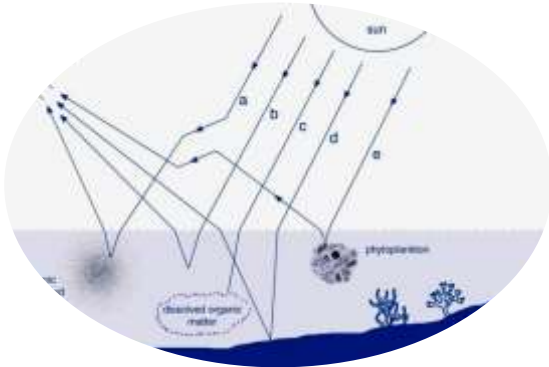
Ocean Temperatures are a much stronger driver than air temperatures



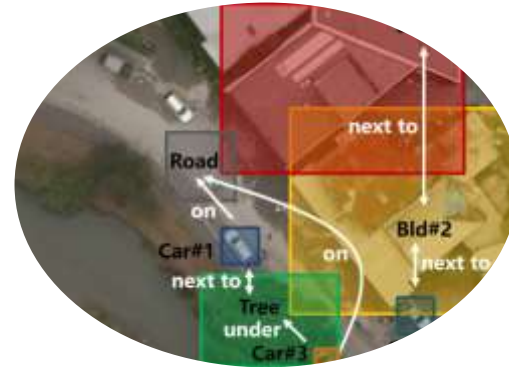
Previously unquantified seasonality

What are the open challenges in ML₄Earth?

Open Issues in AI4EO



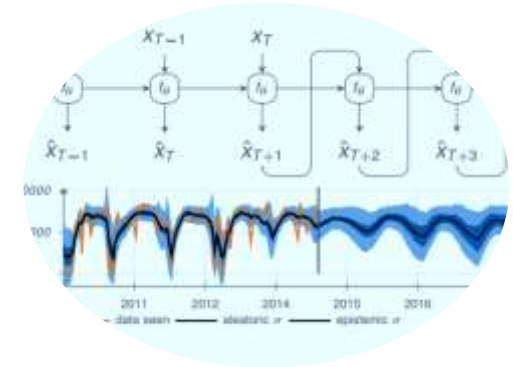
Re-implant **physics, Bayes and domain expertise**



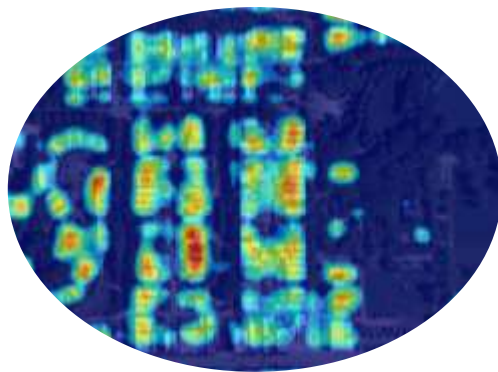
Reasoning



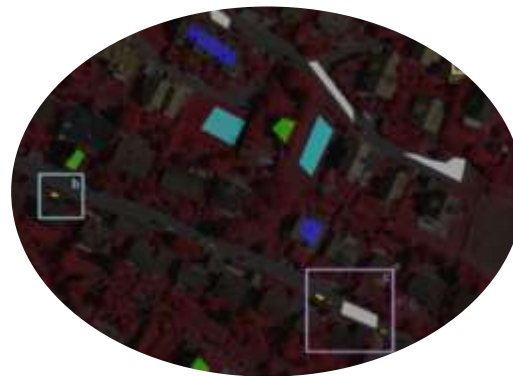
Transferability



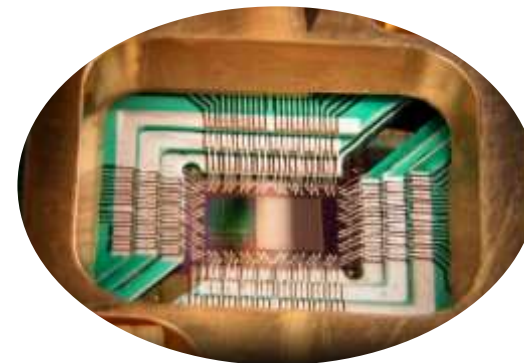
Uncertainty



Explainability



Green AI



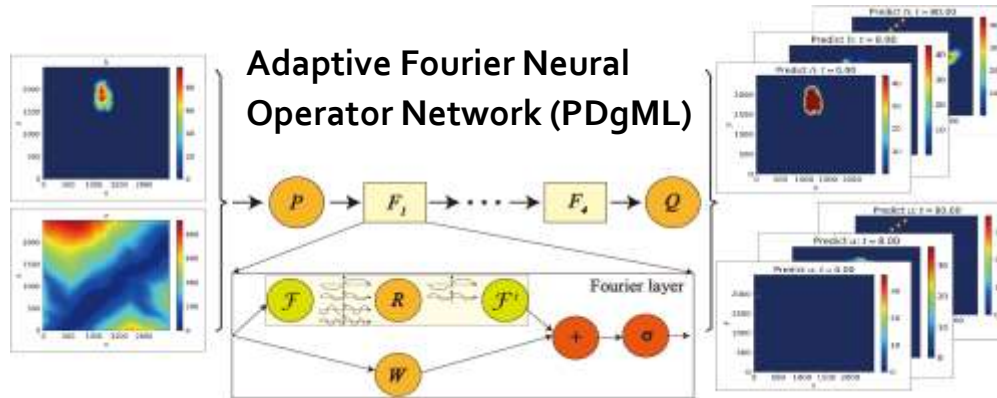
Quantum Machine Learning



Ethics

Physical Data-guided Machine Learning (PDgML) for Urban Flood Modeling and Forecast

Inputs:
 Coordinates X, Y;
 Time domain T;
 Initial conditions (ho);
 Terrain (DEM);
 Rainfall;
 Boundary conditions (Building);



Outputs:
 water height H;
 X-direction velocity of the water U;
 Y-direction velocity of the water V;

Loss function for Urban Flood

a. Data Loss:

$$\mathcal{L}_{data} = \frac{\|u - \hat{u}\|}{\|u\|} \quad u \text{ and } \hat{u} \text{ are simulated and predicted water } (H, U, V)$$

b. Building condition Loss:

$$\mathcal{L}_{BC} = \hat{u}|_{\partial\Omega} - g|_{\partial\Omega} \quad g \text{ represents } H, U, V \text{ equal to } 0 \text{ inside the building}$$

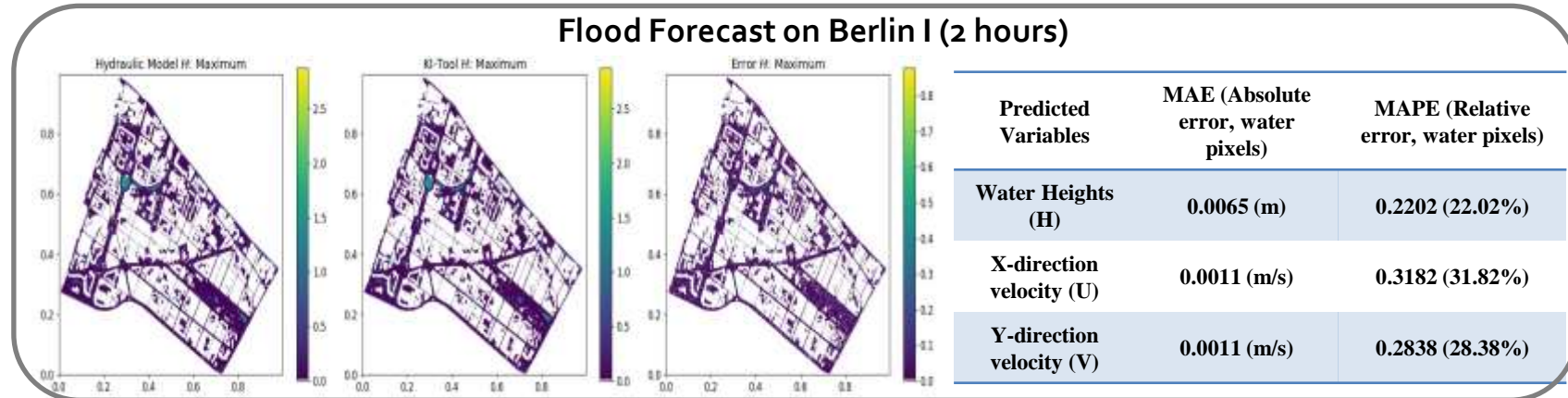
c. Initial condition Loss:

$$\mathcal{L}_{IC} = \hat{u}|_{\Omega, t_0} - g|_{\Omega, t_0}$$

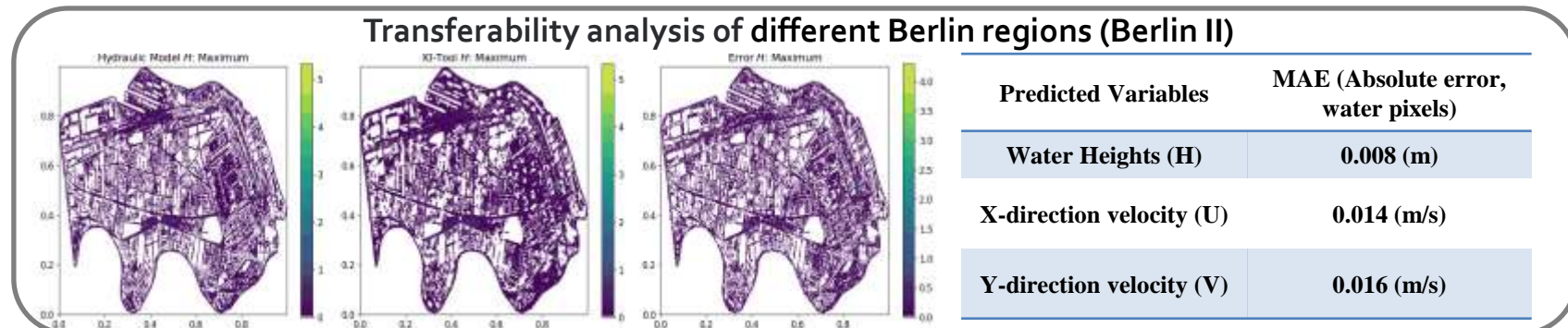


Study Regions

Flood Forecast on Berlin I (2 hours)

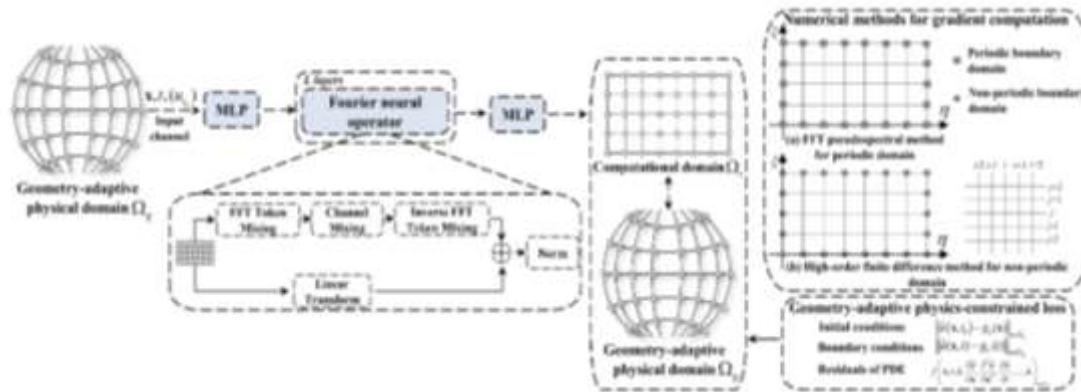


Transferability analysis of different Berlin regions (Berlin II)

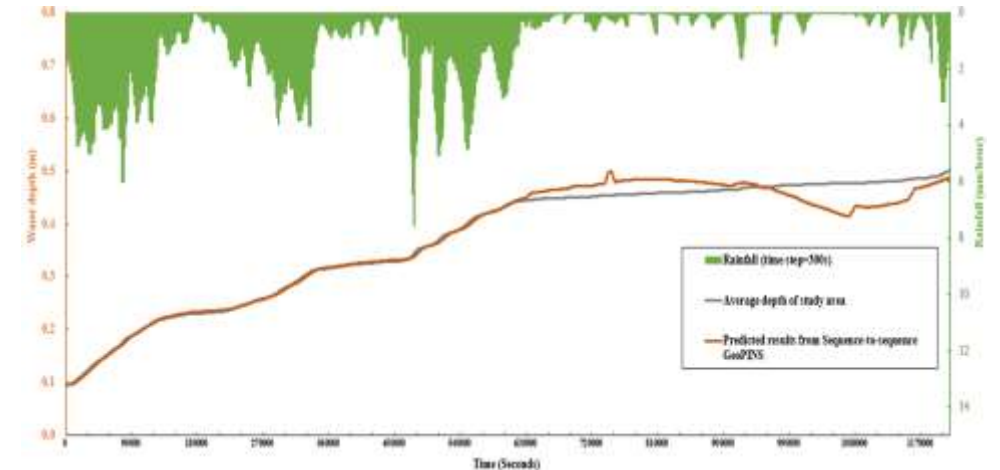


Physics-informed Machine Learning (PiML) for Large-scale Flood Modeling and Forecast

Geometry-adaptive physics-informed neural solver (PiML)



Flood depth validation over a 14-day period
(flood simulation from August 18th to August 30th, 2022)



PiML-based flood forecast achieves an average MAPE of 14.93% for 14-day flood depth predictions. (Pakistan flood prediction in 2022)



Process-based Hydrology in Physics-aware Machine Learning (HydroPML)



MODEL ZOO

More than 100 methods of different physics-aware machine learning



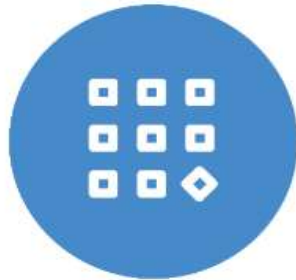
DATASETS

Standard and large-scale datasets for process-based hydrology, such as floods, landslides and other dynamic process datasets



BENCHMARKS

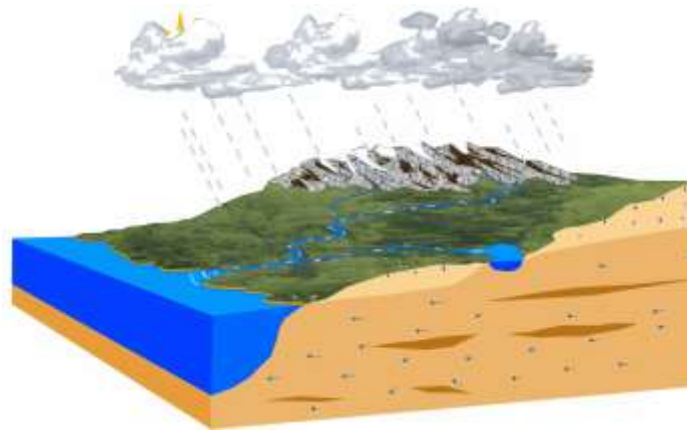
Benchmarking results for different hydrology applications using standard datasets



CASE STUDY

Different application case studies and PaML-based method recommendations

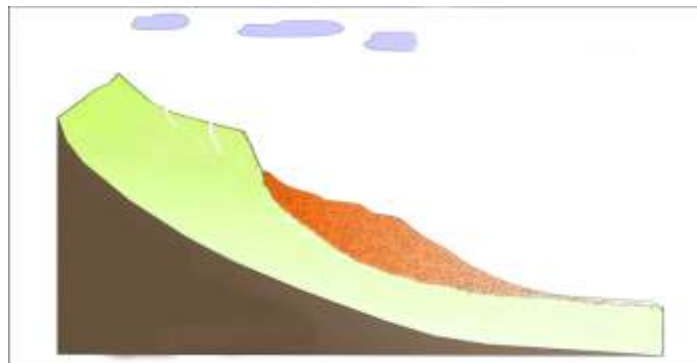
Rainfall-runoff Modeling and Forecast



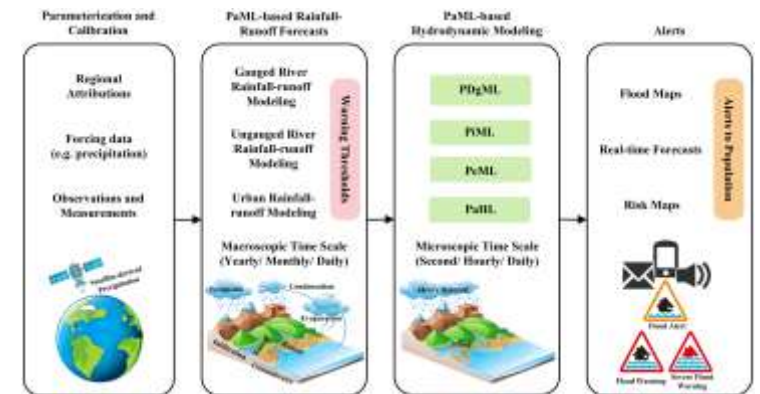
Flood Modeling and Forecast



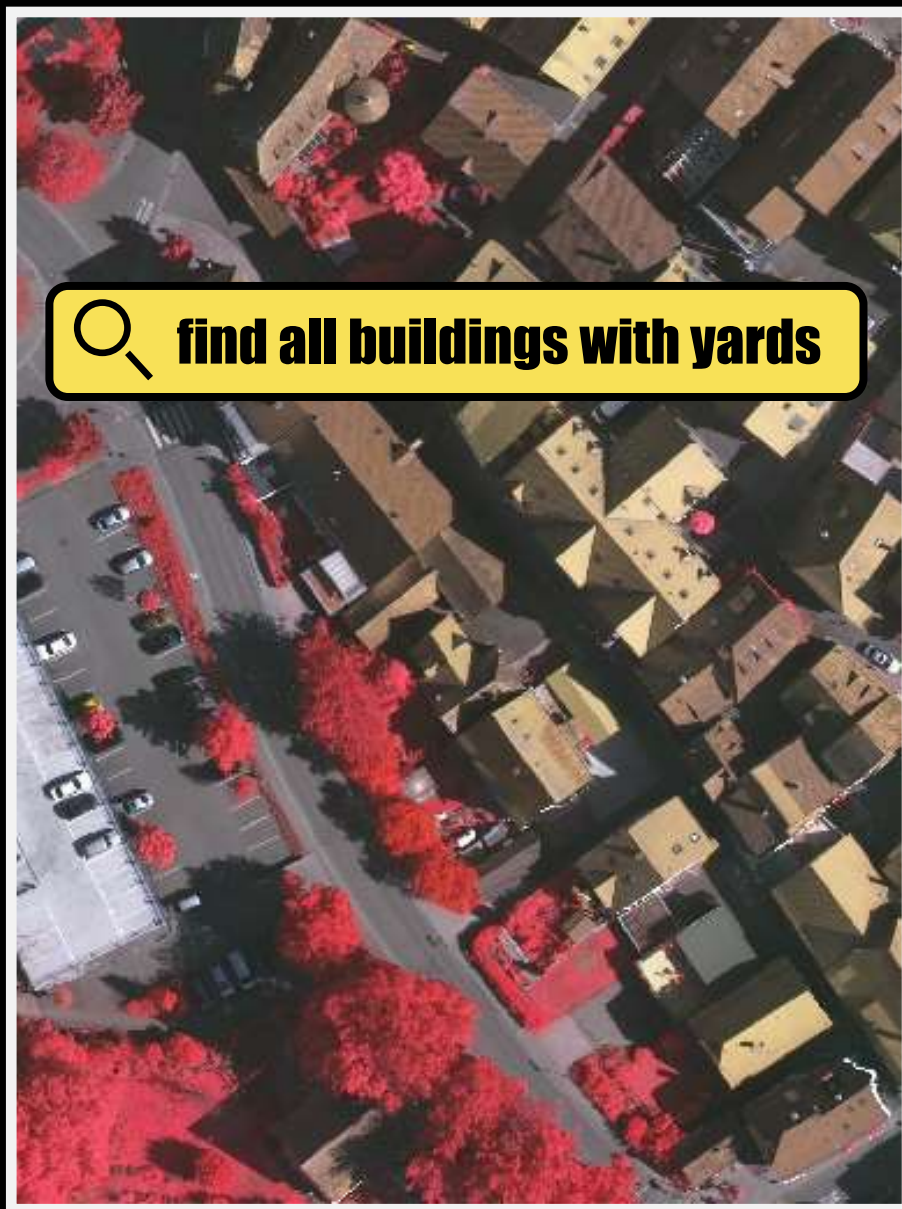
Rainfall-Induced Landslide Modeling and Forecast



Rainfall-runoff-inundation Modeling and Forecast



Human-centered AI analytics of EO data



Human-centered AI analytics of EO data



ChatEarthNet: A Global-Scale Image-Text Dataset Empowering Vision-Language Geo-Foundation Models



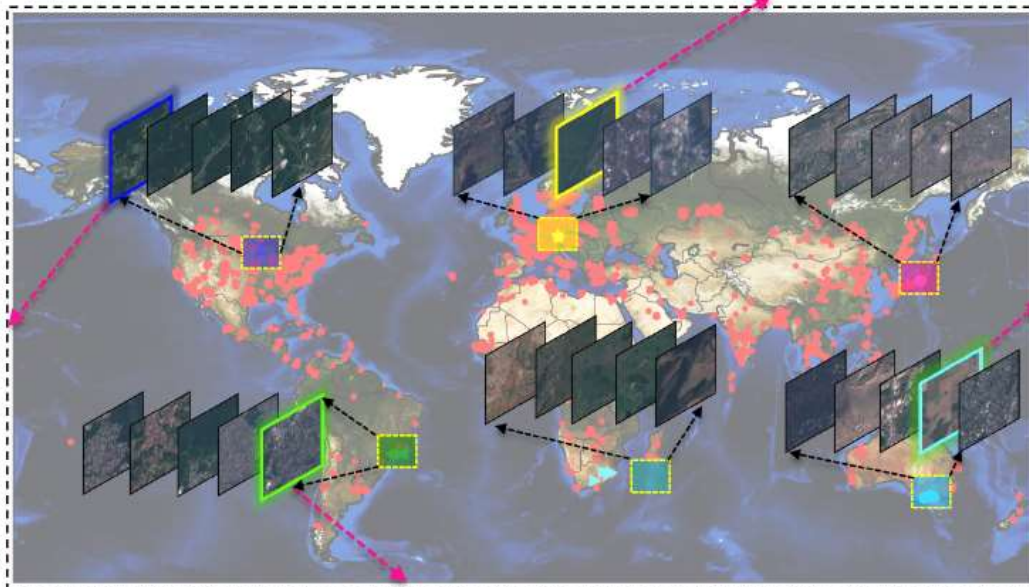
ChatGPT-4V

In this image, the **developed areas** manifest as a branching network, perhaps indicative of transportation routes or urban sprawl, interspersed throughout a substantial presence of **trees**. Small, isolated pockets of **water** are scattered, none large enough to form substantial watercourses or lakes. Sparse instances of **crop** cultivation are noted, with minor patches breaking the dominance of the other vegetation. **Grasslands** are widely distributed, with substantial coverage, suggesting a landscape that balances natural tree coverage with human development and open grassy areas. **The overall theme presents a mixed use of land, where nature and development coexist, but with a clear predominance of tree cover and grassy fields.**



ChatGPT-4V

In this image, **grass** areas predominate, especially in the middle and top left, with substantial coverage observed throughout the image. **Water** bodies are less extensive but have intricately meandering shapes, mainly concentrated towards the middle and top right regions. **Developed areas** have a more concentrated footprint in the bottom right region with geometrically shaped patches throughout. **Tree** coverage is significant but mostly towards the top and bottom left sections, and small **crop** areas are scattered sparsely, with low presence and no distinct shapes or significant coverage. **This image is characterized by the dominance of grass areas and a notable presence of developed areas, particularly in the bottom right quadrant.**



ChatGPT-3.5

The image primarily consists of **crop** fields, which are most dominant across all sections. In the top left, there is a significant expanse of **crop** fields, with a small area of **grass** and **developed land**. Moving to the top right, **crop** fields continue to dominate, followed by a smaller **developed area** and **grassy** patches. In the bottom left, the landscape is mostly covered by **crop** fields, followed by a few **trees** and a small amount of **grass**. The bottom right also exhibits a large area of **crop** fields, accompanied by a small **developed area** and a small portion of **grass**. In the middle section, **crop** fields are again the main feature, with a small number of **trees** and a tiny **developed area**. **Overall, the image depicts a landscape predominantly characterized by crop cultivation, with minor presence of developed areas, trees, and grass.**

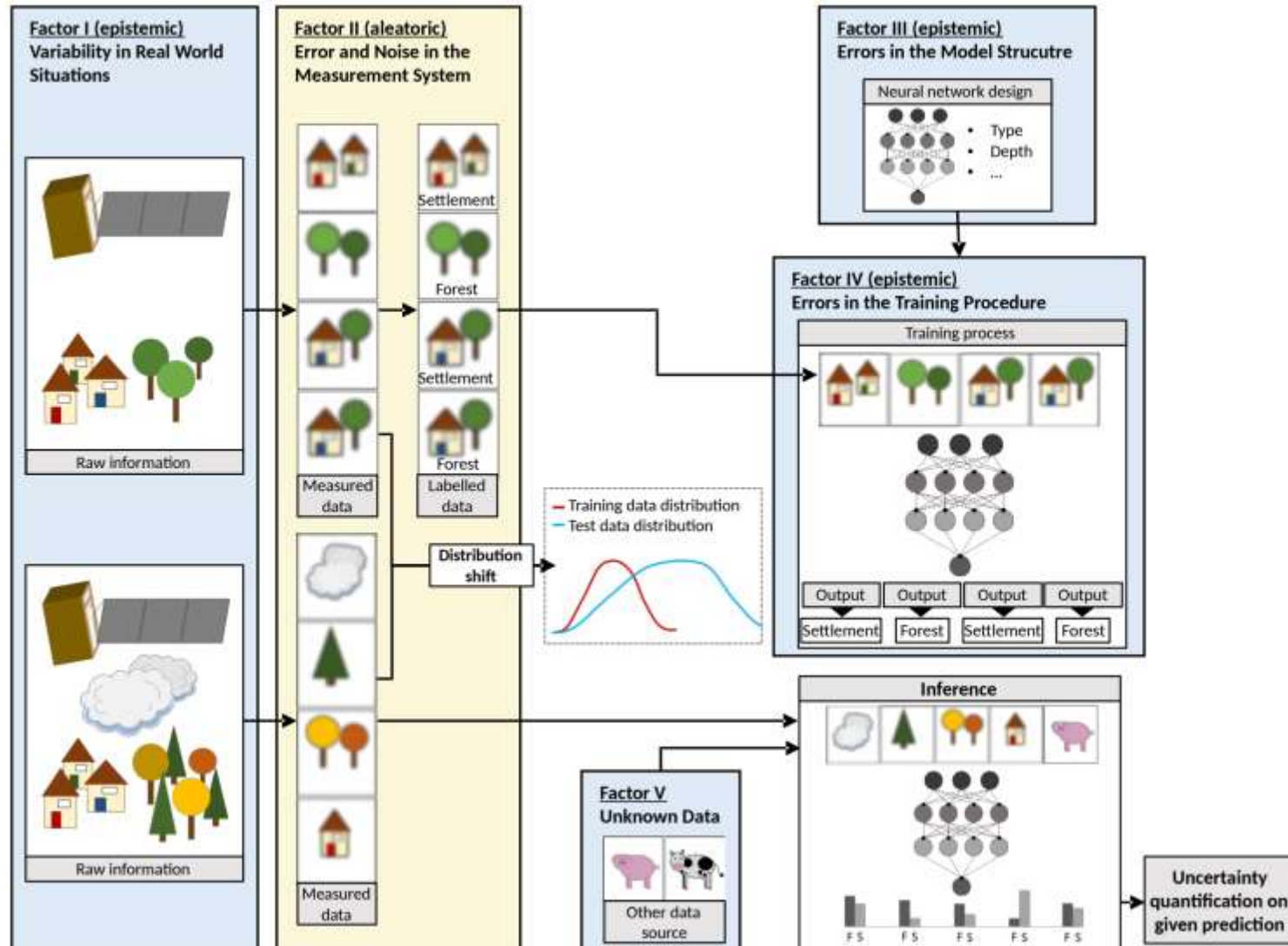


ChatGPT-3.5

The image exhibits varied land cover types across different spatial locations. The dominant land cover type is **developed area**, which occupies an extra large fraction of the overall area. In the top left portion, the land cover primarily comprises **developed area**, followed by a small quantity of **tree**, and extra small quantity of **grass**. Moving to the top right, the **developed area** is still the most prevalent, with an extra large part, accompanied by an extra small part of **water** and an extra small part of **tree**. In the bottom left, the area is mainly covered by **developed area**, followed by a small portion of **tree** and an extra small quantity of **grass**. The bottom right section is also dominated by **developed area**, with a small quantity of **tree** and an extra small quantity of **water**. In the middle, the land cover consists of an extra large quantity of **developed area**, an extra small amount of **tree**, and an extra small portion of **water**. **Overall, the image depicts a landscape dominated by developed areas, with variations in the presence of trees, grass, and water in different spatial locations.**



Uncertainty Quantification



How Certain are the Uncertainty Estimates?

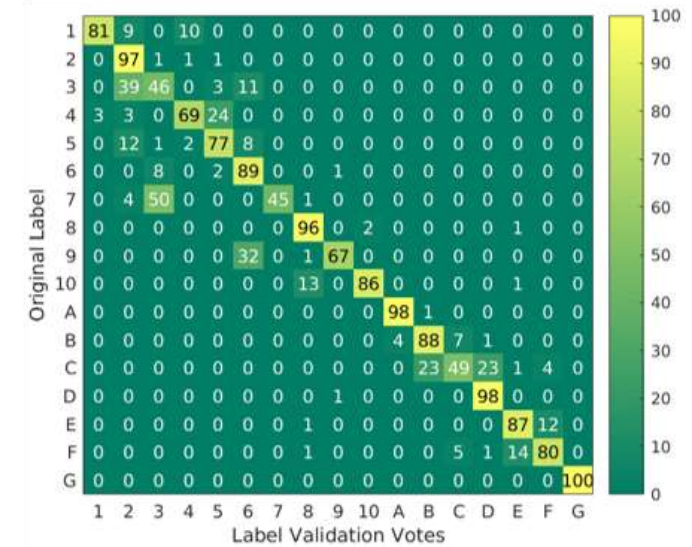
Regression: Tree biomass



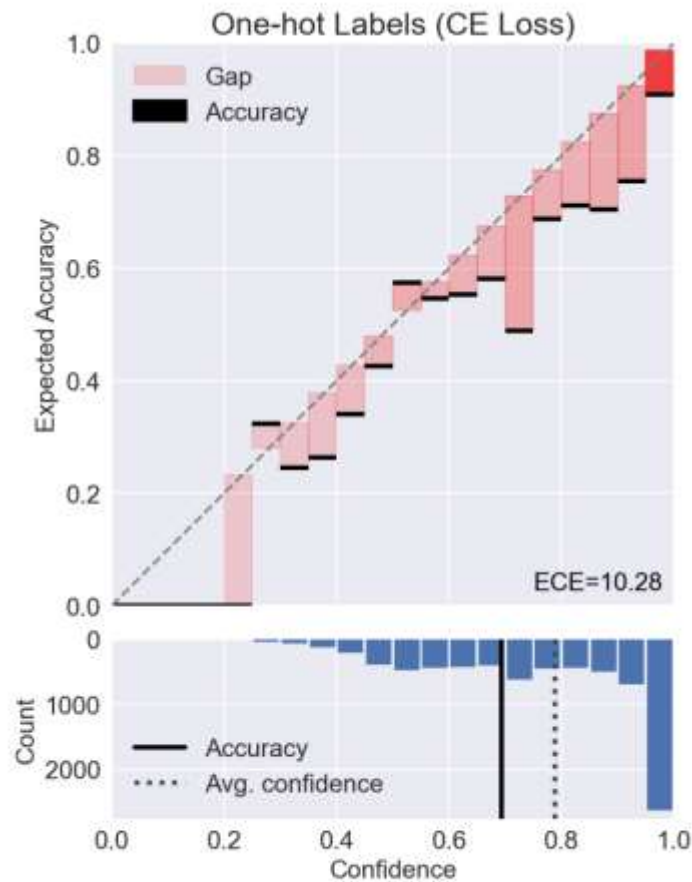
Segmentation: Building footprint



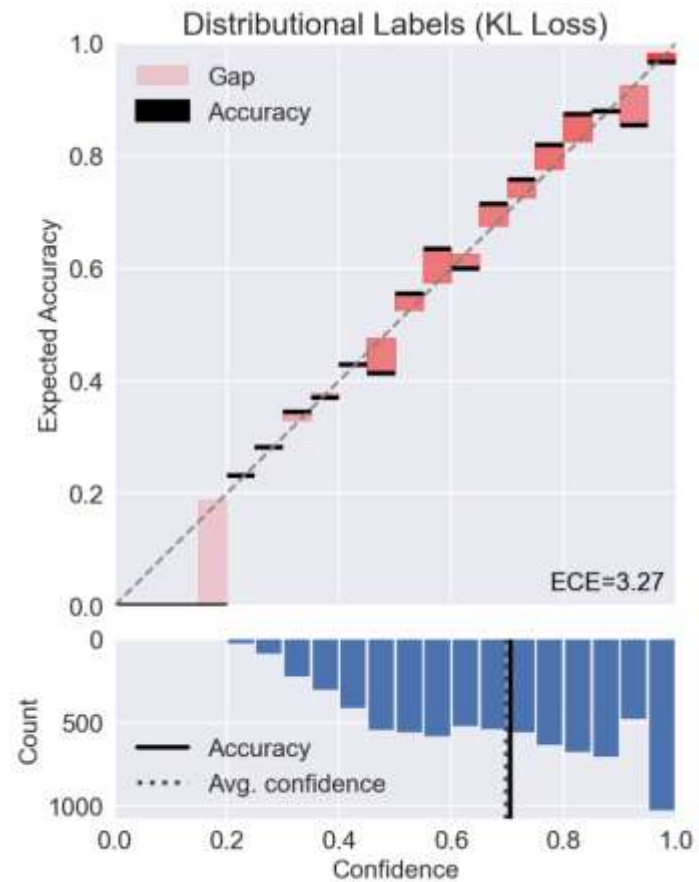
Classification: LCZs



Can Human Uncertainty Improve Model Performance in Earth Observation?



(a) One-Hot Encoding



(b) Label Distribution Encoding

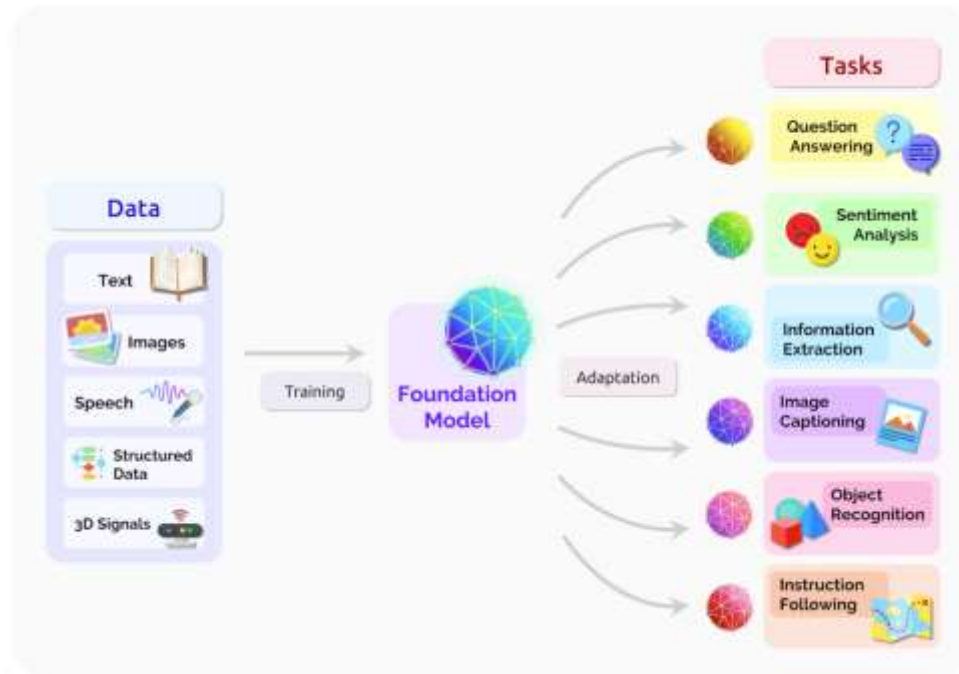
Hey, how about Foundation Models?

What is a foundation model (FM)?

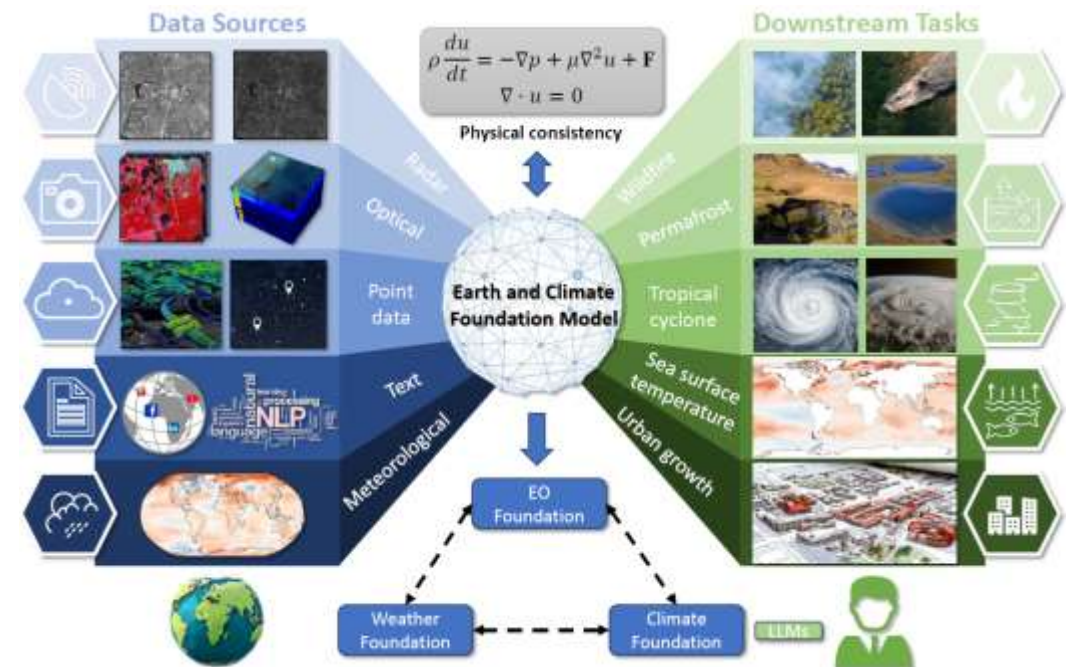
A model that:

- is trained on broad data (generally using self-supervision), and
- can be adapted to a wide range of downstream tasks.

Link to paper:



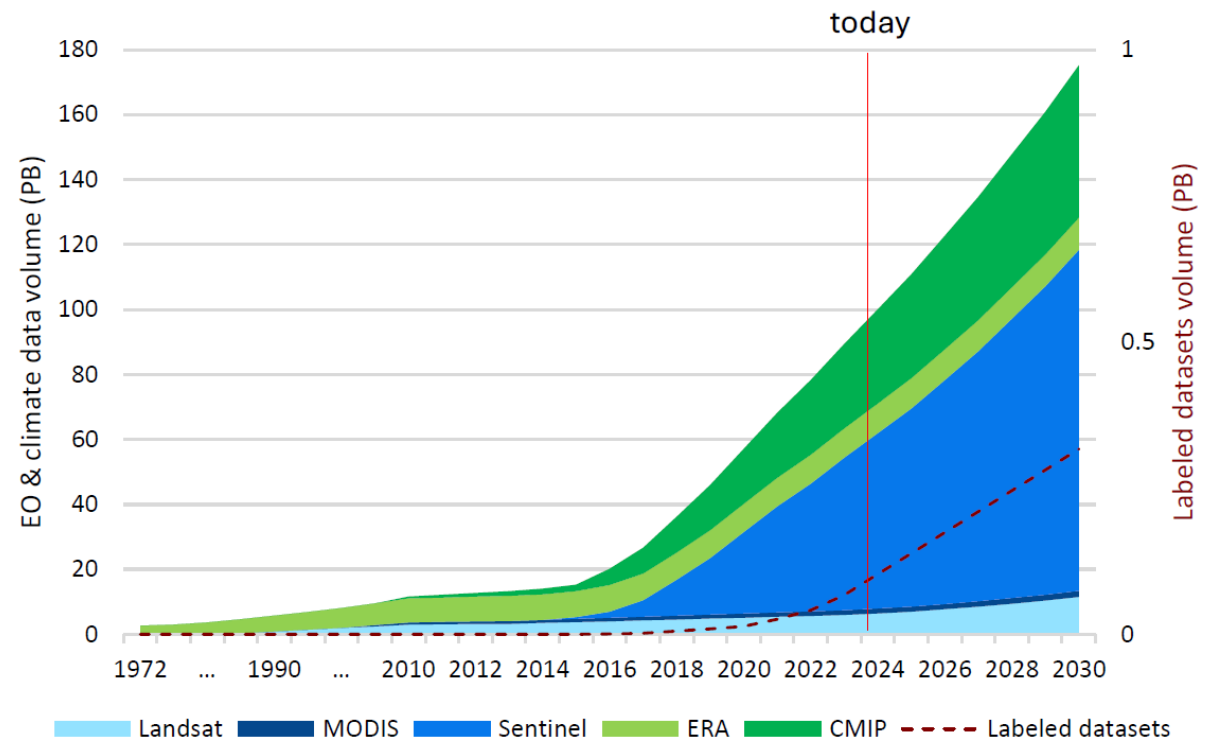
Bommasani, Rishi, et al. "On the opportunities and risks of foundation models." *arXiv* (2021).



Zhu et al., On the Foundations of Earth and Climate Foundation Models, under review.

The potential of Earth and climate FMs

- Unlocking the value of big EO and climate data
- Enhancing label efficiency
- Reducing carbon footprints
- Bridging EO and climate science
- Improving Earth system modeling



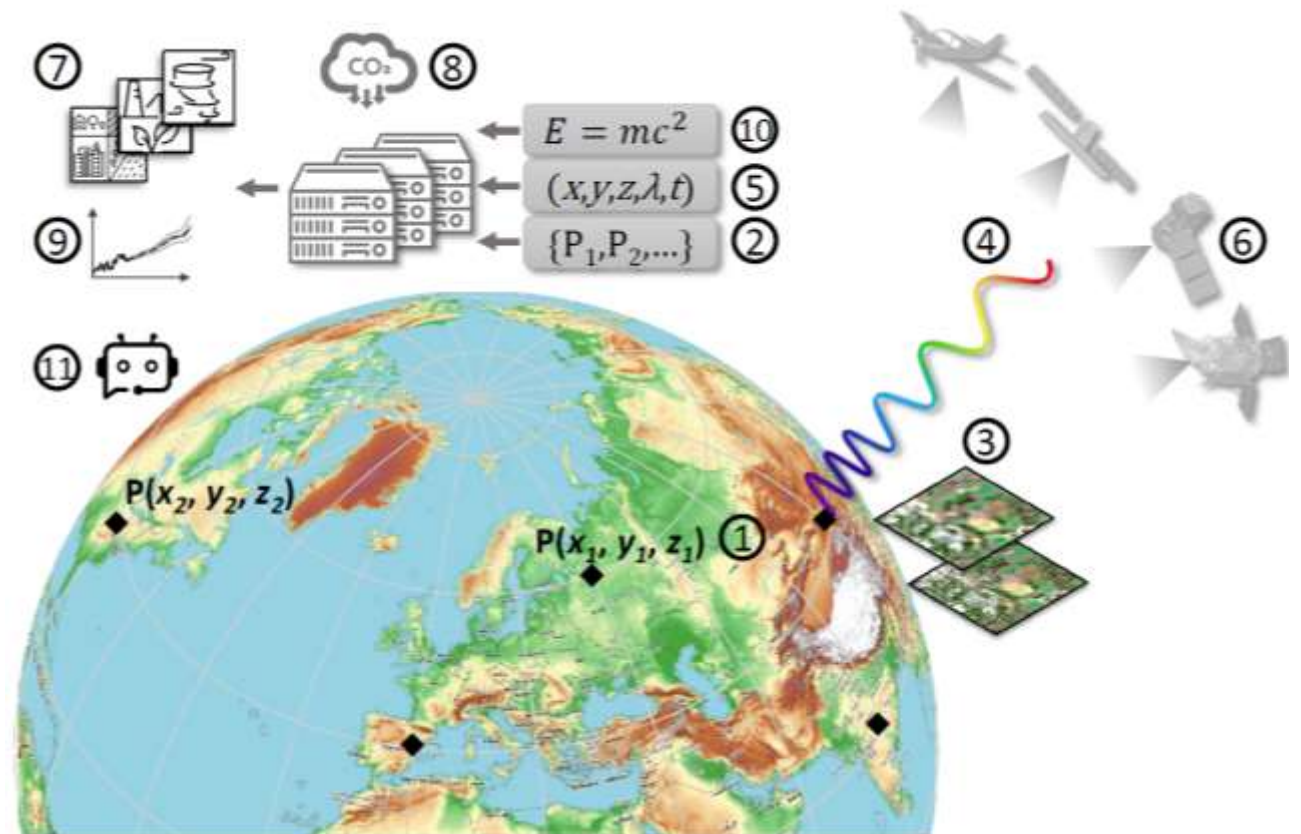
The ideal Earth and climate FM

Must have features:

1. Geolocation embedding
2. Balanced geographical representations
3. Scale awareness
4. Wavelength embedding
5. The time variable
6. Multisensory
7. Task-agnostic
8. Carbon minimized

Highly desirable features:

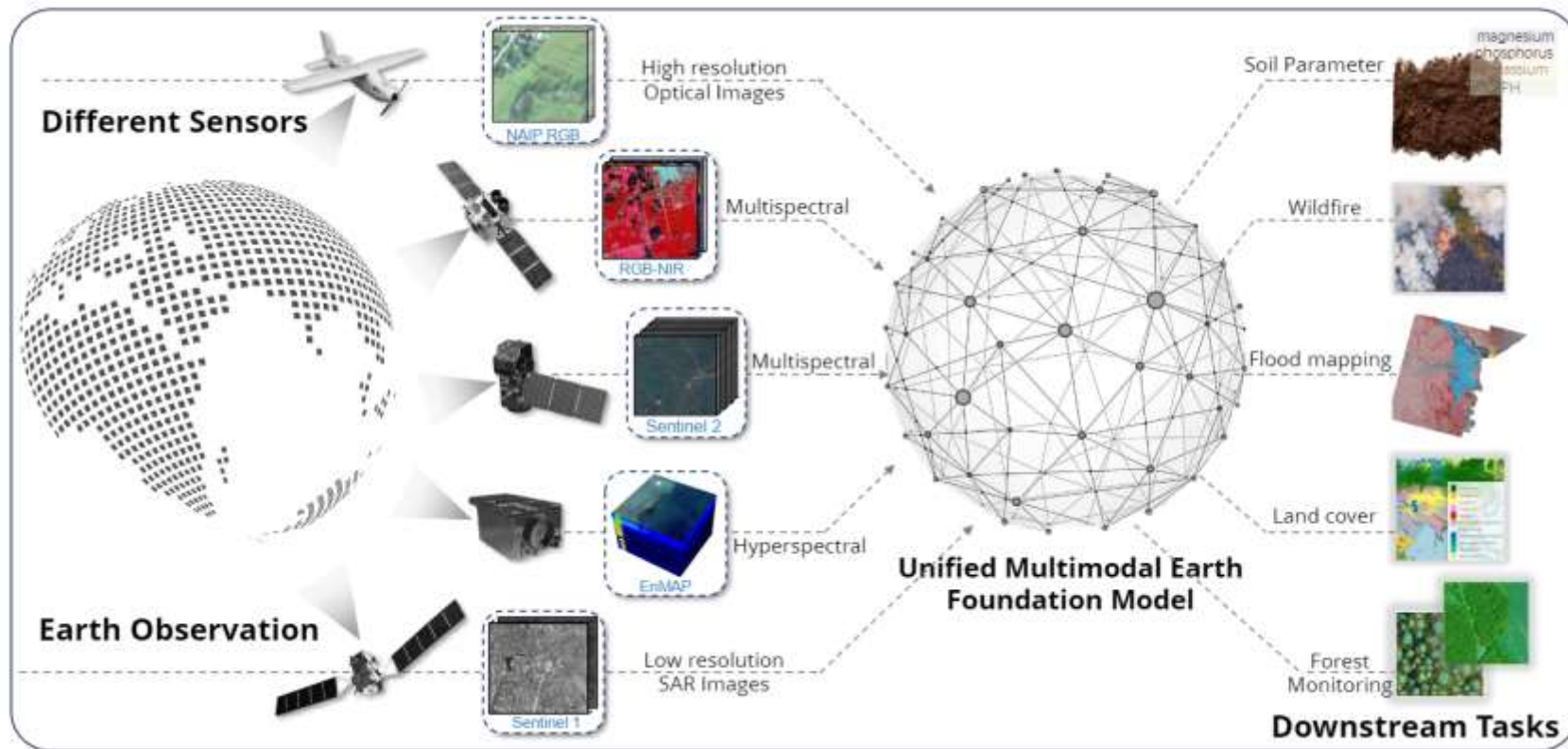
9. Uncertainty quantification
10. Physical consistency
11. AI assistants



Zhu et al., On the Foundations of Earth and Climate Foundation Models, under review.

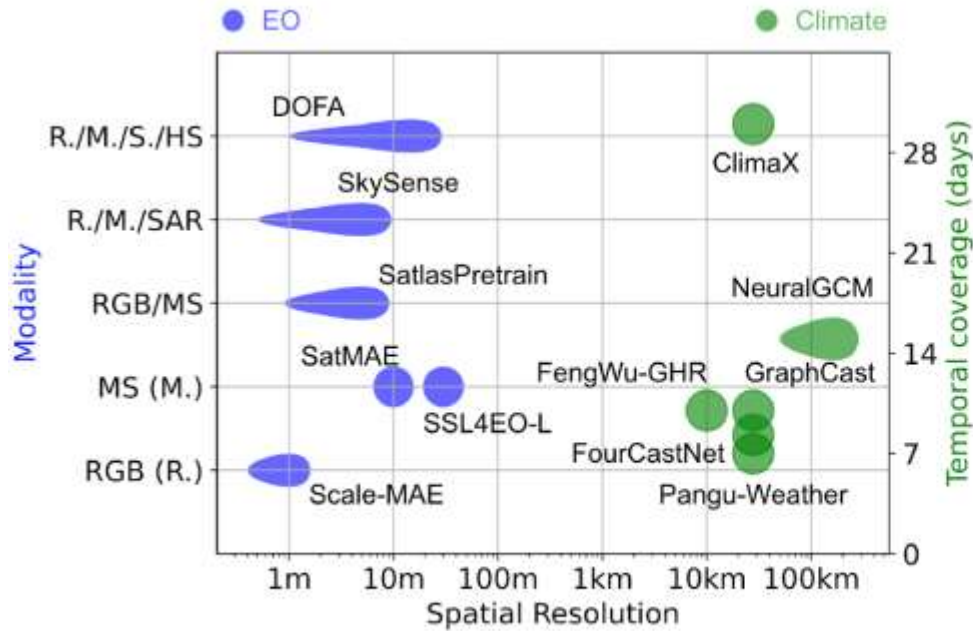
DOFA: Neural Plasticity-Inspired Foundation Model for Observing the Earth Crossing Modalities

- Integrate various data modalities into a single framework adaptively
- Excel across 12 distinct Earth observation tasks, including sensors never seen during pretraining

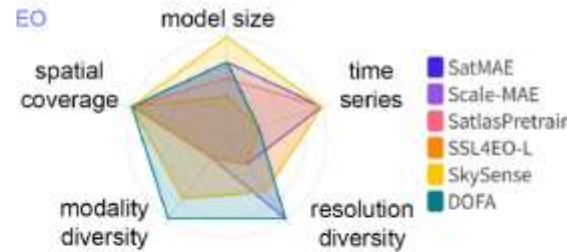


SOTA FMs and gaps

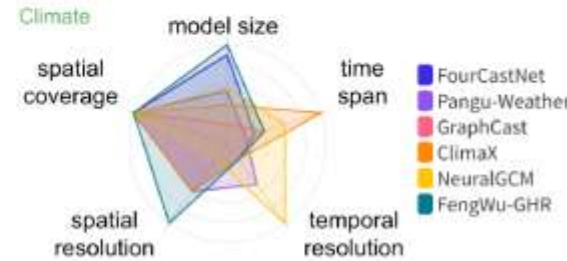
- Most models focus on only one or two of “must-have” features
- Big potential but gaps exist in connecting EO and climate FMs



(a) Comparison between EO and climate spatial resolutions.



(b) EO model comparison.

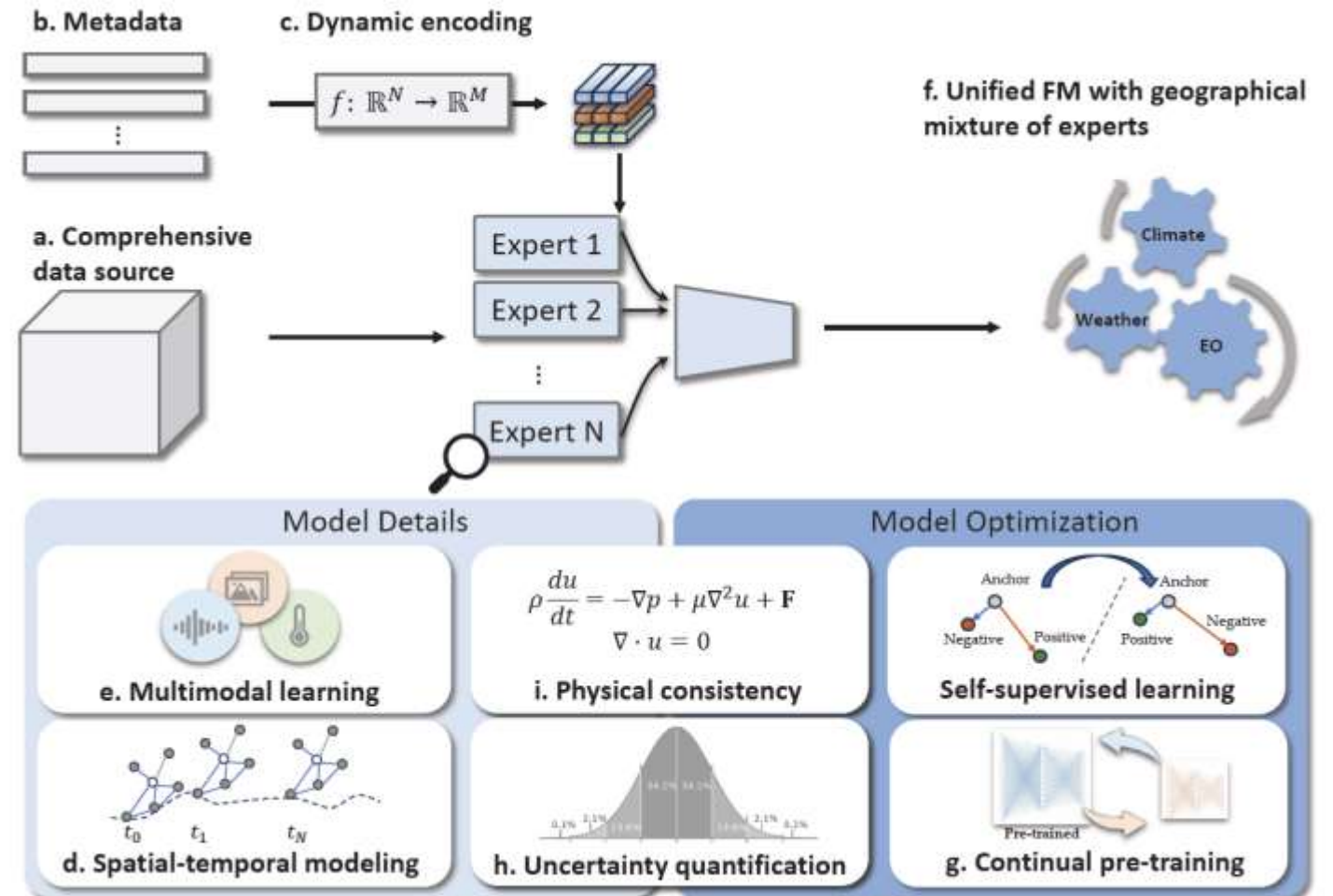


(c) Climate model comparison.

FM	1. Geolocation Embedding	2. Balanced Distribution	3. Scale Awareness	4. Wavelength Embedding	5. The Time variable	6. Multisensory	7. Task-agnostic	8. Carbon-minimized	9. Uncertainty	10. Physical Consistency	11. Language Alignment
SatMAE ⁴⁶	✗	✗	✗	✓	✓	✗	✓	✗	✗	✗	✗
Scale-MAE ⁴⁵	✗	✗	✓	✗	✗	✗	✓	✗	✗	✗	✗
SatlasPretrain ³⁶	✗	✗	✗	✗	✓	✓	✓	✗	✗	✗	✗
SSL4EO-L ³⁵	✗	✓	✗	✗	✗	✓	✓	✗	✗	✗	✗
SkySense ⁴¹	✓	✗	✗	✗	✓	✓	✓	✗	✗	✗	✗
DOFA ⁵²	✗	✗	✗	✓	✗	✓	✓	✓	✗	✗	✗
FourCastNet ⁵⁷	✓	✓	✗	-	✓	✓	✓	✓	✓	✗	✗
Pangu-Weather ⁶⁷	✓	✓	✗	-	✓	✓	✓	✓	✓	✗	✗
GraphCast ⁶⁸	✓	✓	✗	-	✓	✓	✓	✓	✗	✗	✗
ClimaX ⁶³	✓	✓	✗	-	✓	✓	✓	✓	✗	✗	✗
NeuralGCM ⁶⁰	✓	✓	✓	-	✓	✓	✓	✗	✗	✓	✗
FengWu-GHR ⁵⁹	✓	✓	✓	-	✓	✓	✓	✗	✗	✗	✗

The way forward

- a) Comprehensive data source
- b) Metadata integration
- c) Dynamic encoder
- d) Spatial-temporal modeling
- e) Multimodal learning
- f) Geographical mixture of experts
- g) Continual pre-training
- h) Uncertainty quantification
- i) Physical consistency

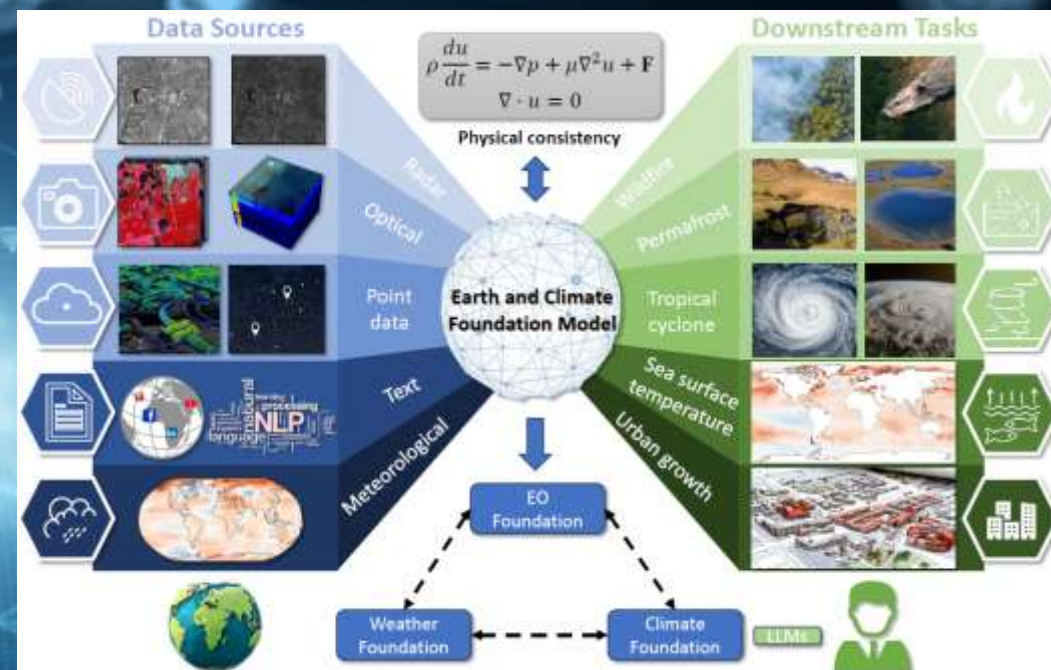


What comes after the Foundation Models?

ML4Earth Foundation Model Workshop

Where: Kolpinghaus in Munich

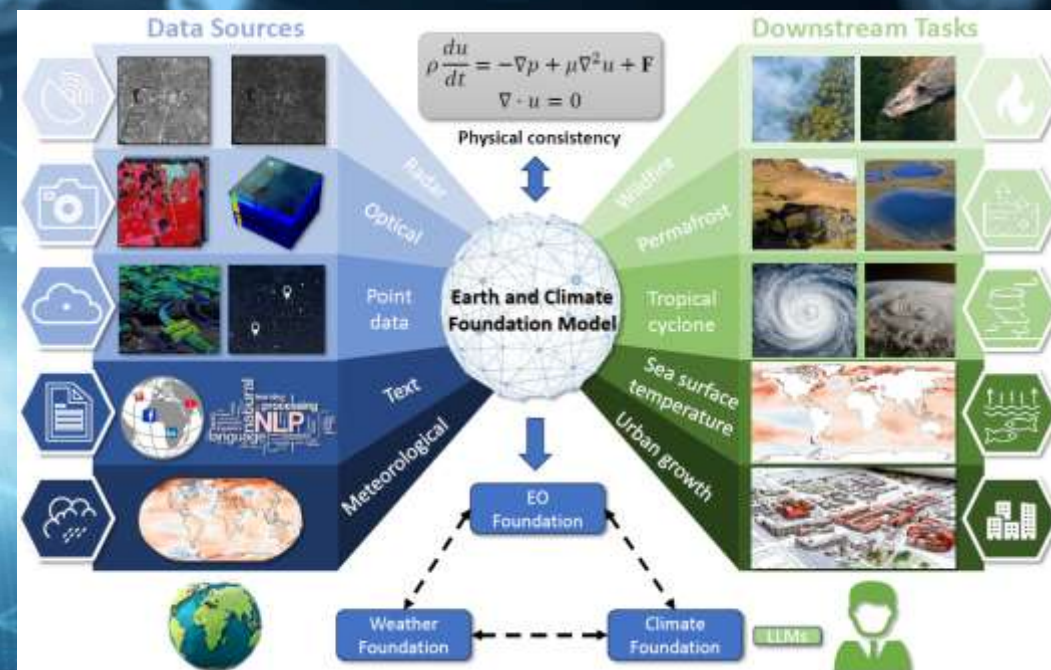
When: September 18, 2024



ML4Earth Foundation Model Hackathon

Where: Technical University of Munich

When: September 19 - 23, 2024



AI4EO Symposium 2024

Where?

TUM Campus Ottobrunn

Lise-Meitner-Straße 9, 85521 Ottobrunn

When?

July 15-16, 2024



Free registration at:



A Few Statements

- Big Earth data and AI4EO offer invaluable geo-information helping tackling societal grand challenges
- Towards large scale and actionable geoinformation retrieval and prediction, fundamental methodological challenges must be addressed, such as uncertainty, physical-consistency, explainability, transferrability etc.
- Earth and climate foundation models have enormous potential, but we are just scratching the surface ...

Contact: xiaoxiang.zhu@tum.de



@xiaoxiang_zhu