Large Scale Masked Autoencoding for Reducing Label Requirements on SAR Data

M. Allen, F. Dorr, J. A. Gallego-Meija, L. Martinez-Ferrer, A. Jungbluth, F. Kalaitzis, R. Ramos-Pollan

FDLEUROPE.ORG

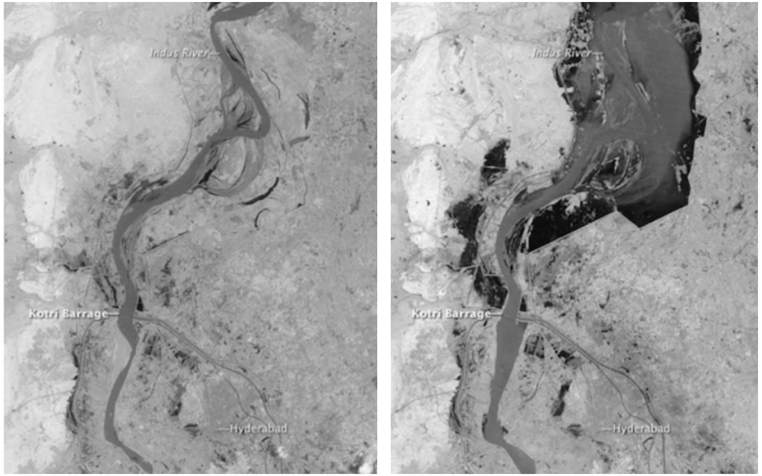
esa Φ-lab | FDL 2023 EUROPE | Google Cloud | NVIDIA | SCAN AI | Centre for Satellite Data in Environmental Science | PASTEUR iSi | AIRBUS | The Danish Meteorological Institute | D-ORBIT | CATAPULT Satellite Applications | UNIVERSITY OF OXFORD | TRILLIUM EUROPE

Global Challenges & Earth Observation

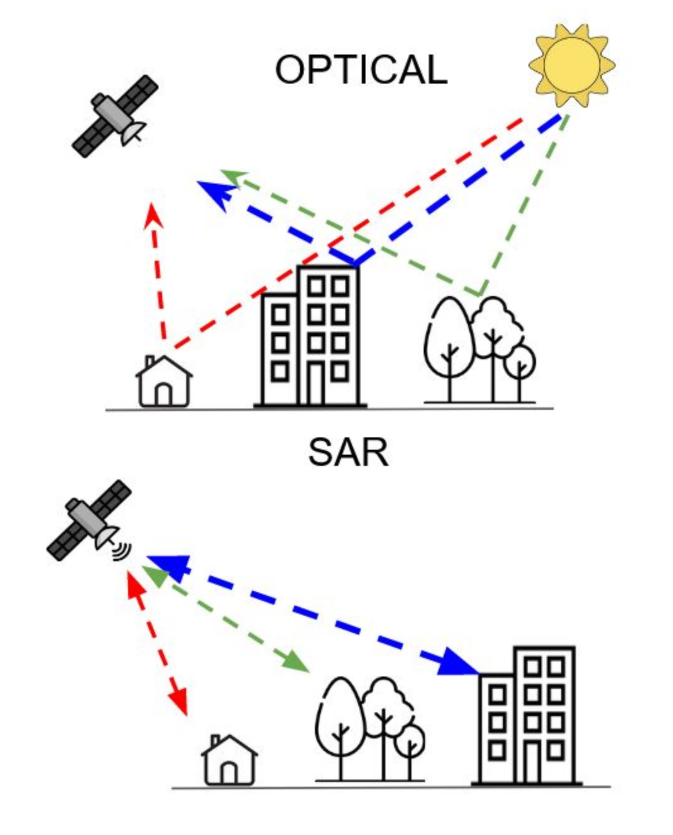
Causes and consequences of climate change evolve rapidly at global scale

Satellite data is crucial for monitoring and mitigation

Effect of interventions is limited by reliance on optical data

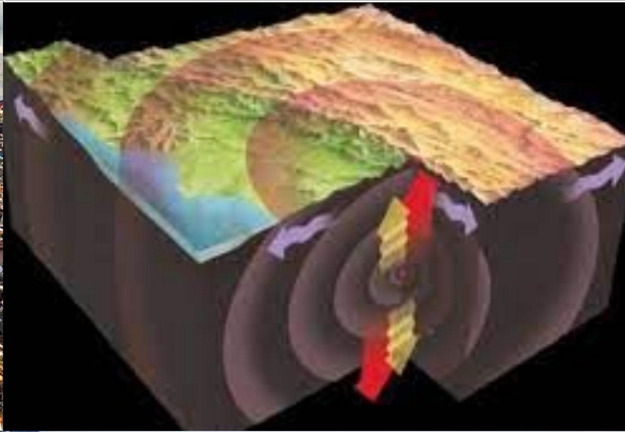


All-weather, day night coverage



Rapid Revisit Times

Detecting Small Scale Changes



SAR - Complexities

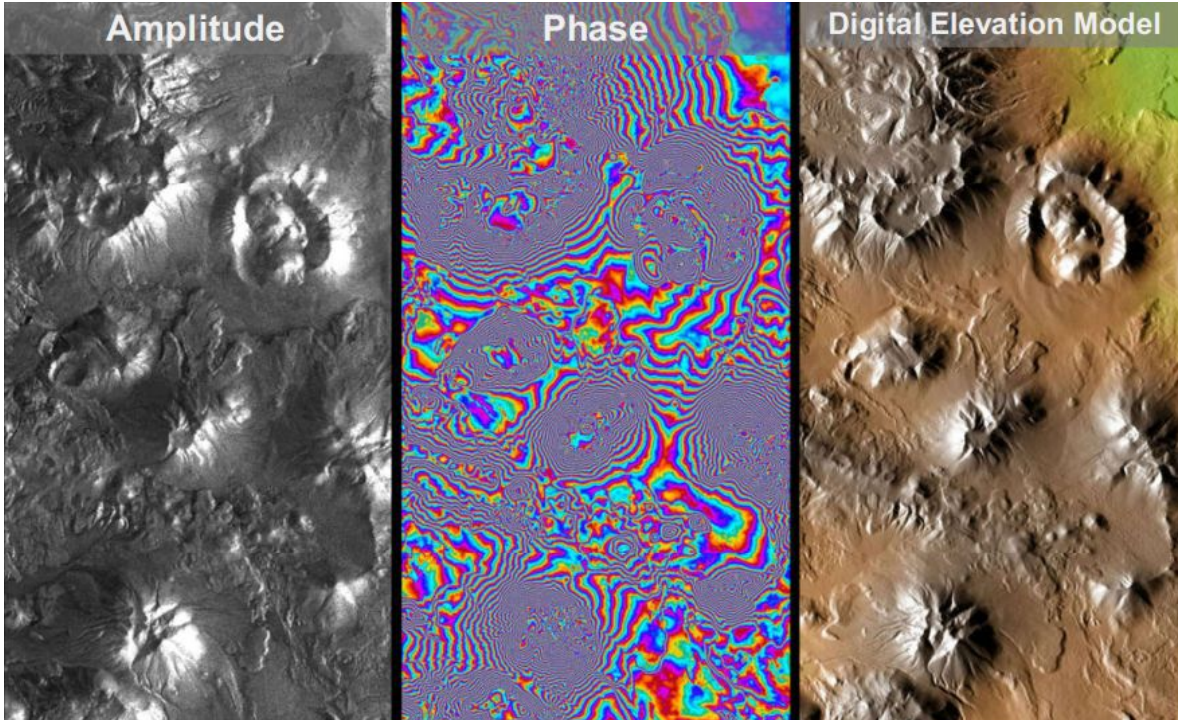
Preprocessing SAR data requires expert knowledge

Different ‘modalities’ of SAR data (amplitude, coherence, interferometry)

Difficult to interpret and label

Large, self-supervised models would be ideal

Off-the-shelf pretrained RGB models are not appropriate

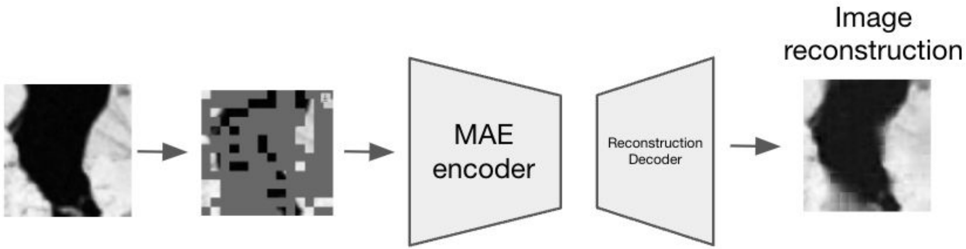


Model

Masked Autoencoder (He 2021) with 2 minor modifications

Increased channels to 12
VV, VH polarisations
VV-VH log difference
4 seasonal averages

Decreased patch to image size ratio
Factor of 2 - remote sensing imagery less correlated at distant pixels
~Confirmed by linear probe



Input Datasets

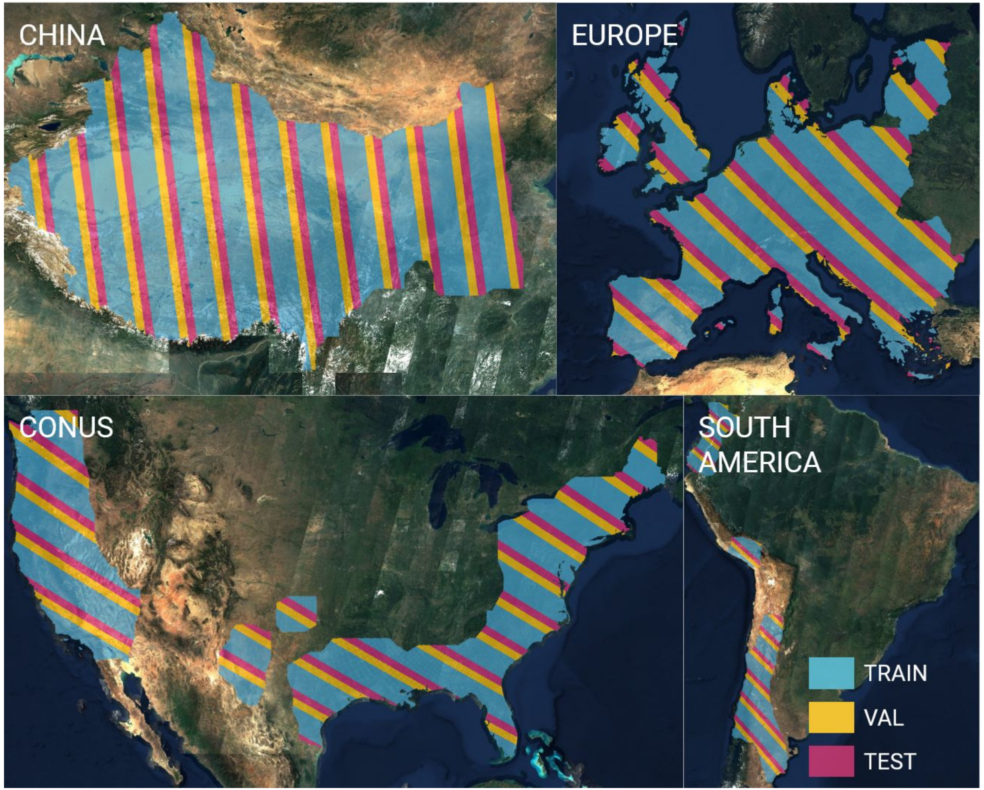
Diverse terrain - approx 10% Earth surface

Split by geographic banding (60:20:20)

Pretrain - CONUS, Europe, China

Test - Europe, South America

Sentinel 1 GRD polarimetry
448x448 tiles - 10m resolution

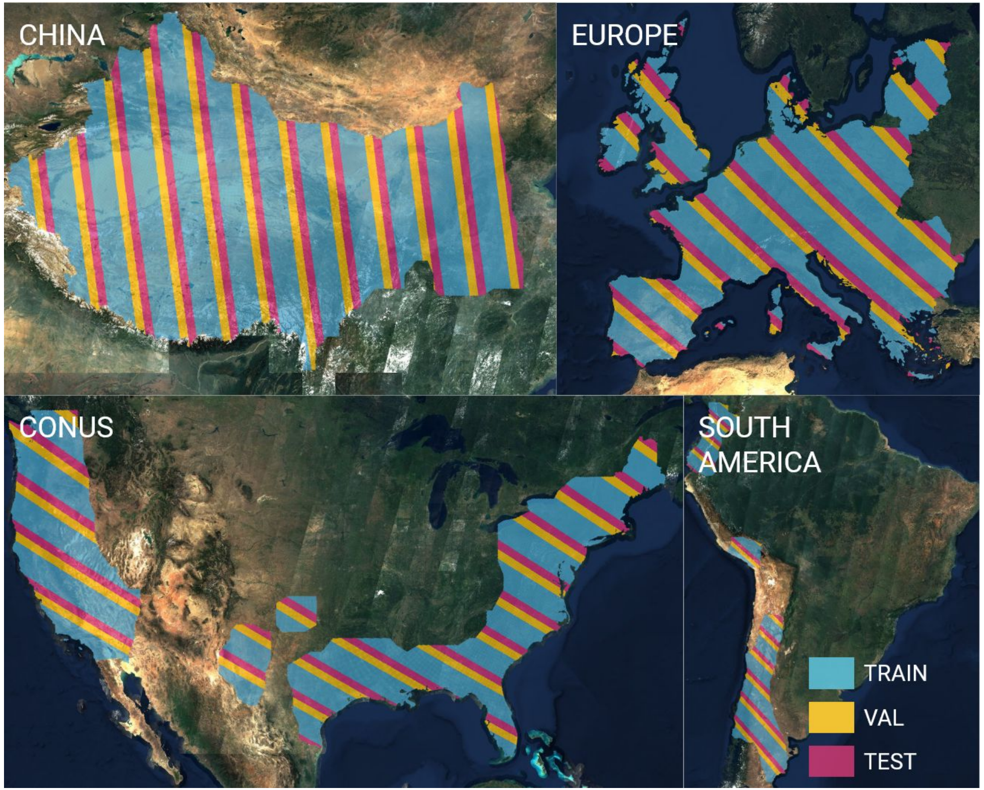


Labelled Datasets

MODIS Vegetation Continuous Fields
Percentage vegetation cover (regression per-tile)
Regression head - 1D Conv then 3 FC layers

ESA World Cover
Land cover classification (semantic segmentation)
~10m/pixel
11 classes
Segmentation head - SETR-PUP (Zheng 2021)

Data Ablations - 0.1, 1, 10, 100% labels



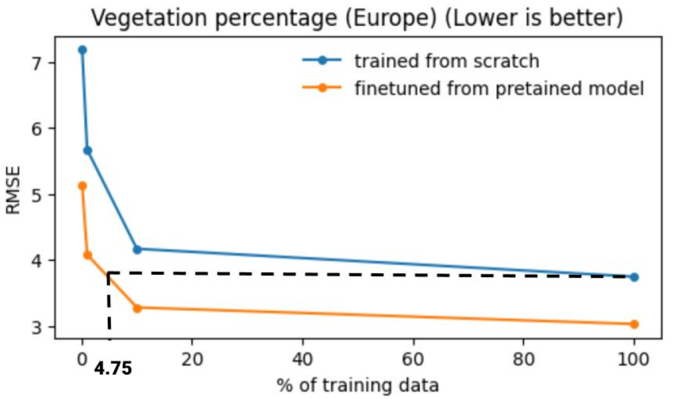
Results - MODIS Vegetation

Consistent **increase in performance for both regions** for all **data percentages**

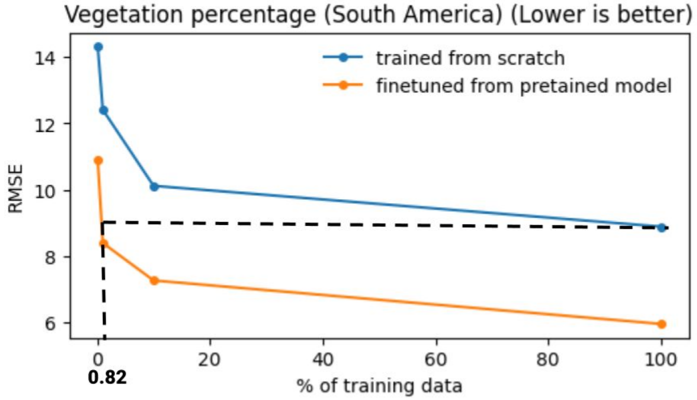
Large reduction in label requirements (10x) for fixed performance

The **effect is greater on the region outside the pretraining set**

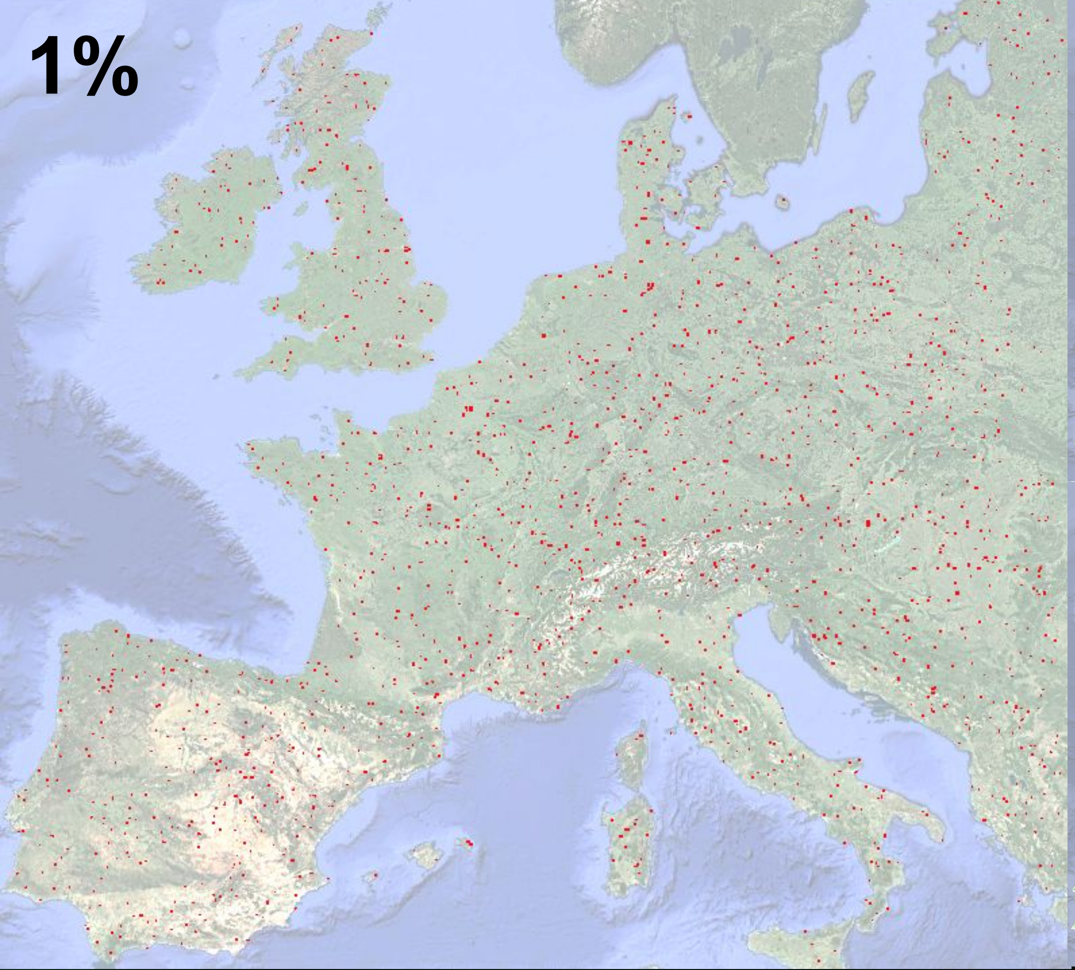
Test region (Europe) IN pretraining set

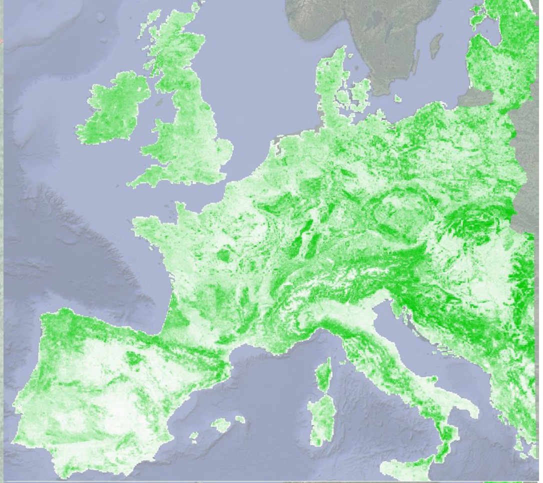


Test region (S. America) OUTSIDE pretraining set



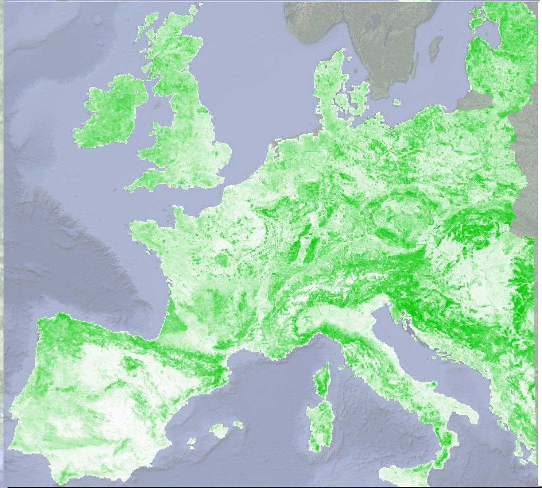
1%





MODIS Vegetation prediction

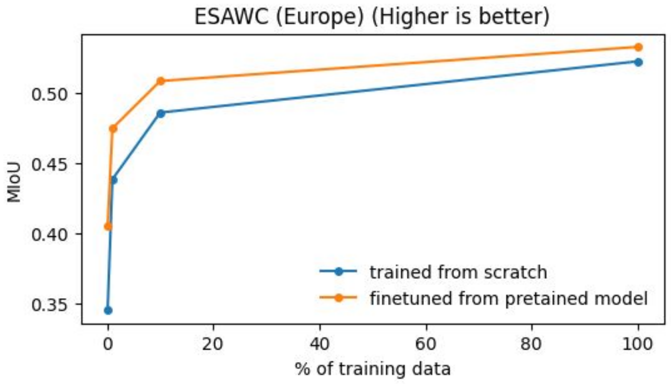
Model trained on **1% of the data**



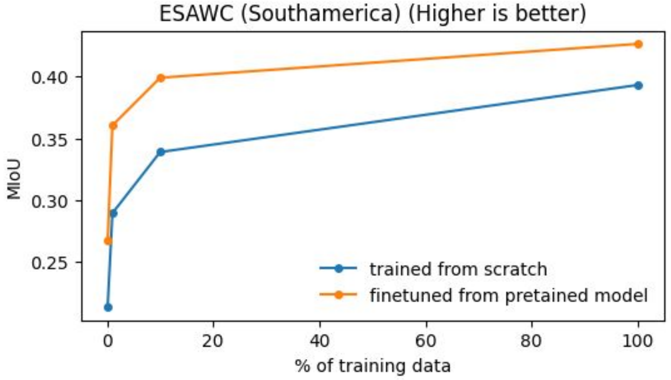
Labels

Results - ESA WC

Test region (Europe) IN pretraining set



Test region (S. America) OUTSIDE pretraining set

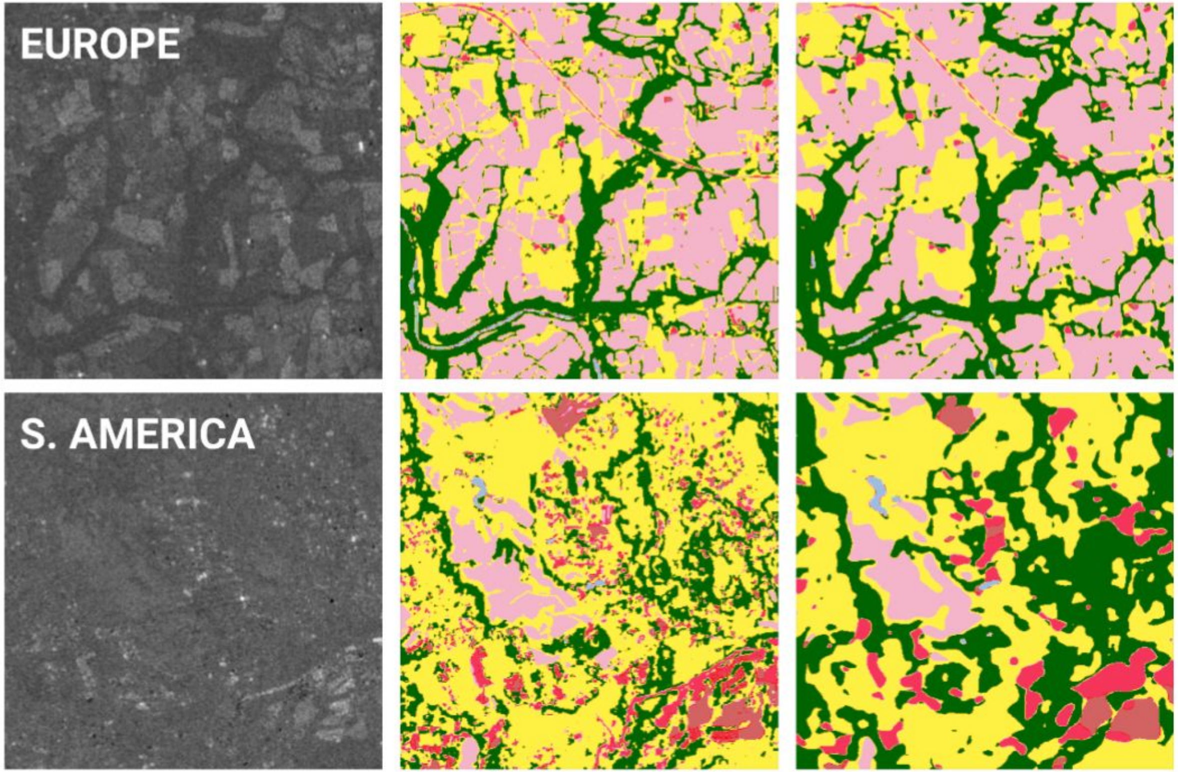


Same **increase in performance for both regions** for all **data percentages**

Again - **effect is greater on the region outside the pretraining set**

Smaller reduction (~60% EUR, ~90% SA)

SAR | **LABELS** | **PREDICTIONS (1%)**



- Tree cover
- Shrubland
- Grassland
- Cropland
- Built-up
- Bare / sparse vegetation
- Snow and ice
- Permanent water bodies
- Herbaceous wetland
- Mangroves
- Moss and lichen

Conclusions

Consistent performance increases for land cover segmentation & vegetation prediction

This effect was greater when the downstream region was **outside** the pretraining set

Promising early results for scaling to a larger model with more data

Needs more in-depth comparison to other models - direct comparison precluded by implementation differences so far

See other work @NeurIPS 2023 with same authors