Scalable Geospatial Data Generation Using AlphaEarth Foundations

Luc Houriez[1,2], Sebastian Pilarski[1], Behzad Vahedi[1], Ali Ahmadalipour[1], Teo Honda Scully[1], Nicholas Aflitto[1], David Andre[1], Caroline Jaffe[1], Martha Wedner[1], Rich Mazzola[1], Josh Jeffery[1], Ben Messinger[1], Sage McGinley-Smith[1], Sarah Russell[1]

1. Bellwether at Google X, the Moonshot Factory, 2. Stanford University

Overview

- High-quality labeled geospatial datasets are essential for extracting insights and understanding our planet, however these datasets are often limited to specific geographic regions and are expensive to generate
- We propose and evaluate a method that leverages Google DeepMind's new AlphaEarth Foundations (AEF) to extend existing labeled datasets beyond their original geographic boundaries.
- Using this method, we extend vegetation datasets crucial for wildfire disaster management from the USA into Canada with up to 81% accuracy.
- This work demonstrates that this task can be accomplished using even shallow learning models, such as random forests or logistic regression, despite discussed limitations.

Background

AlphaEarth Foundations (AEF) is a general-purpose, geospatial foundation model-as-data publicly available on Google Earth Engine.

How AEF Works: The model transforms Earth observation data (e.g. Landsat and Sentinel satellites) into a structured, dense latent representation (embedding) (Fig. 1). This output is provided as a dataset of 64-dimensional vectors at a 10-meter resolution (over land), and is updated annually.

Existing Vegetation Type (EVT) is an ecological dataset provided by LANDFIRE (multi government agency program) available in the USA only and used in wildfire management efforts.

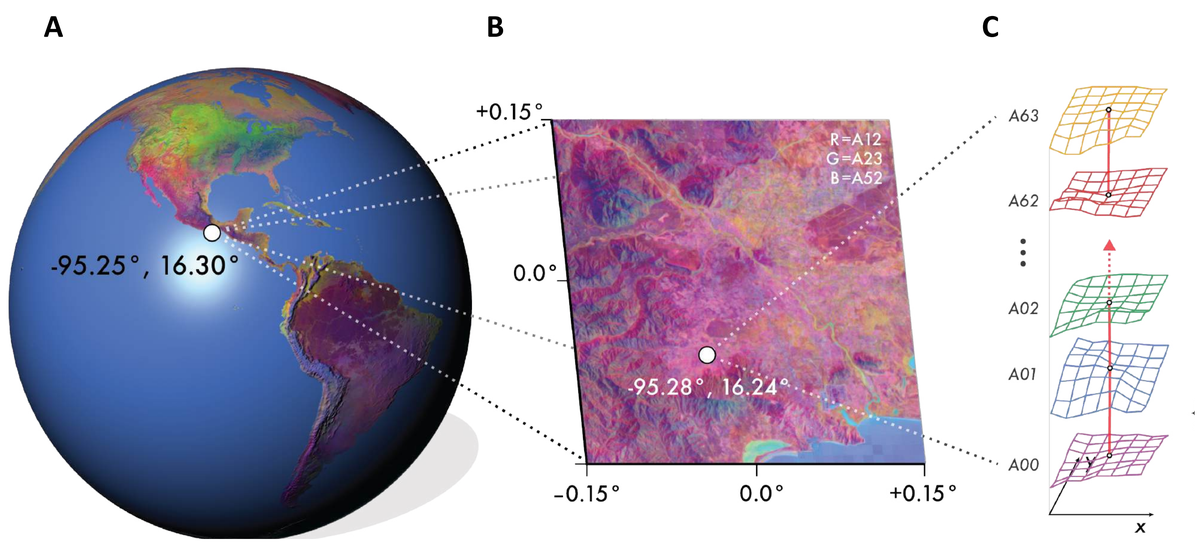

Figure 1. (A) Representation of AEF for the year 2023, note apparent climatic gradients at large scales. (B) AEF produces highly resolved features at 10m2, shown here plotting arbitrary axes in Oaxaca, Mexico. (C) A stack of 64 rasterized AEF bands forms an embedding field, and each individual vector maps to a point on the globe. Figure from Brown et al[1].

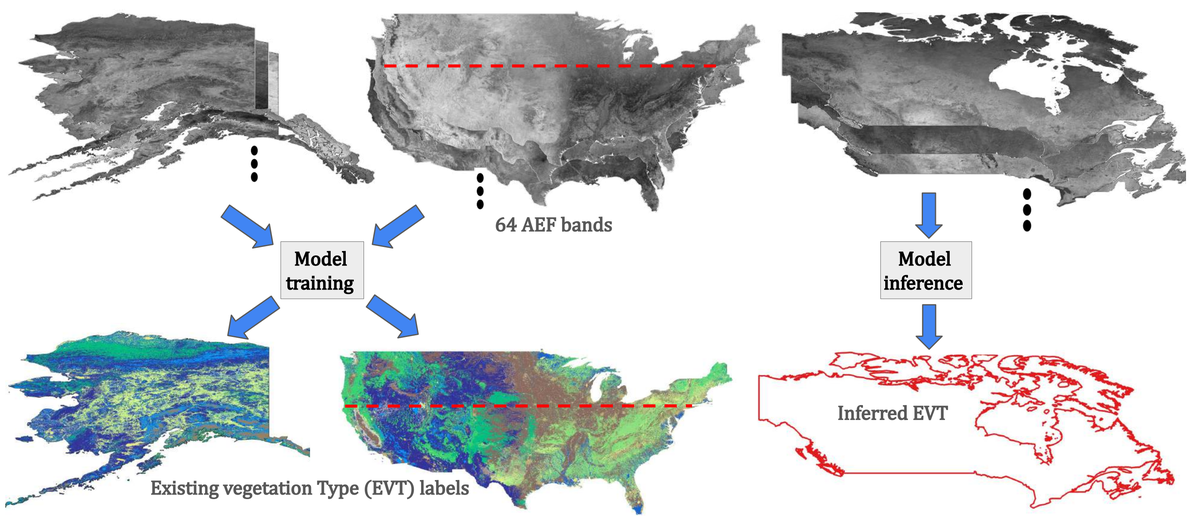

Figure 2. Schematic of model training and inference. The 64 bands of AEF data (input) and EVT data (target) from continental USA north of the red dotted line and Alaska are used to train the model. Running inference on AEF data in Canada provides expected EVT in the previously unlabeled region.

Methods

Models Evaluated: We trained and compared four different classification models:

- Logistic Regression
- Random Forest
- Gradient Boosted Trees
- A U-Net Segmentation Model

Classification Task: We tested this extension at two levels of label granularity:

- **EVTPhys:** 13 broad vegetation classes.
- **EVTGp:** 80 specific vegetation classes (pre-filtered for our region of interest).

Data & Validation:

- Models were **trained** on AEF data from Alaska and the northern continental US.
- Models were **tested** against a "ground truth" EVT test set available in a 90km band along the Southern and Western Canadian border (Fig. 2).

References

[1] Brown, Christopher F., et al. "AlphaEarth Foundations: An embedding field model for accurate and efficient global mapping from sparse label data." arXiv preprint arXiv:2507.22291 (2025).
[2] "Landfire technical documentation. " (2023) Washington DC: US Department of the Interior, US Geological Survey

Results

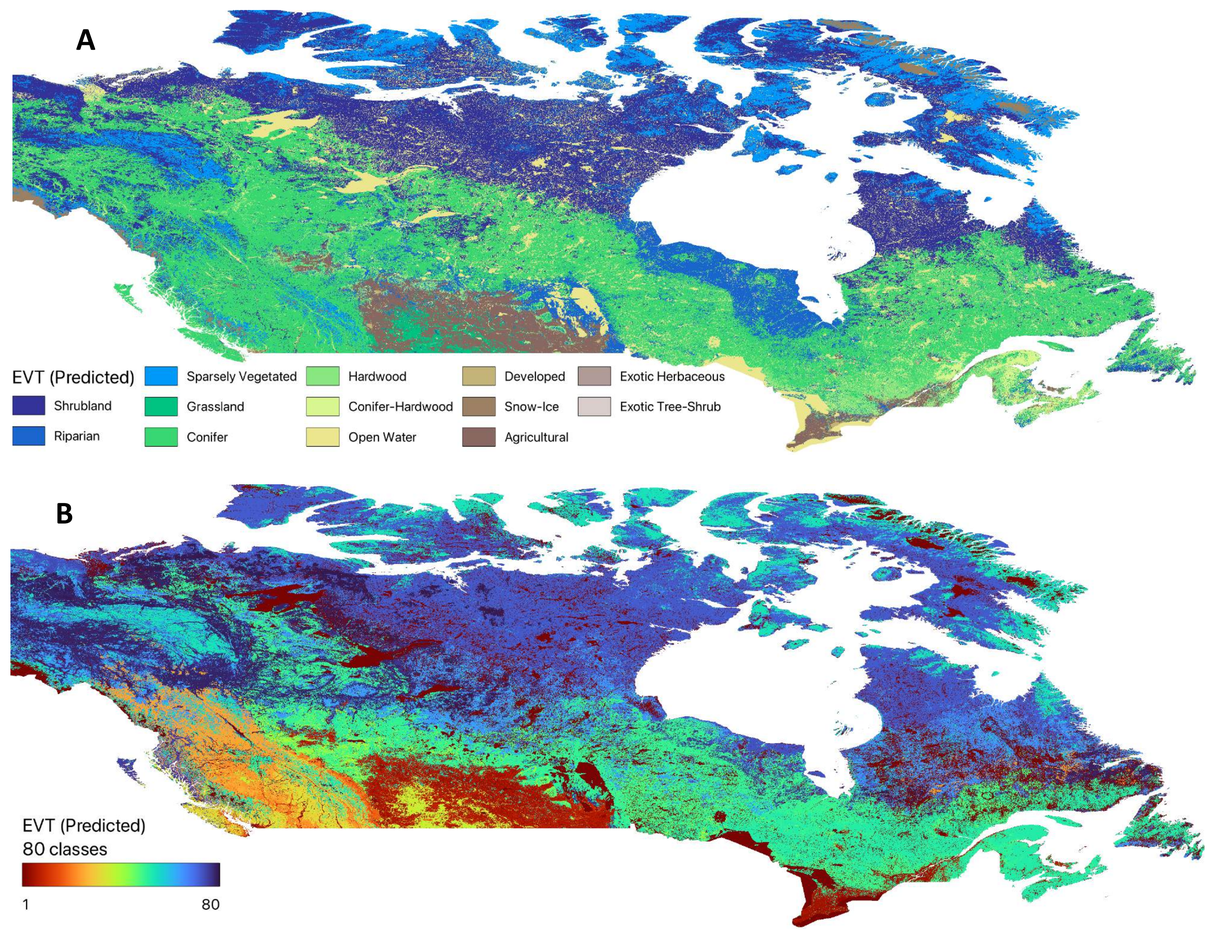

Figure 3: Inference in Canada generated by the segmentation model (A) for EVTPhys and (B) for EVTGP

	Training			Validation			Test		
	ACC	J	F1	ACC	J	F1	ACC	J	F1
Logistic Regression	0.77	0.48	0.60	0.77	0.48	0.59	0.71	0.39	0.51
Random Forest	0.97	0.95	0.97	0.81	0.55	0.67	0.73	0.43	0.55
Gradient Boosted Trees	0.79	0.52	0.65	0.79	0.52	0.64	0.73	0.42	0.54
Segmentation Model	0.79	0.50	0.63	0.79	0.51	0.63	0.73	0.42	0.54

Table 1: Accuracy (ACC), and macro-averaged Jaccard (J) and F1 across data splits for EVTPHYS.

Discussion

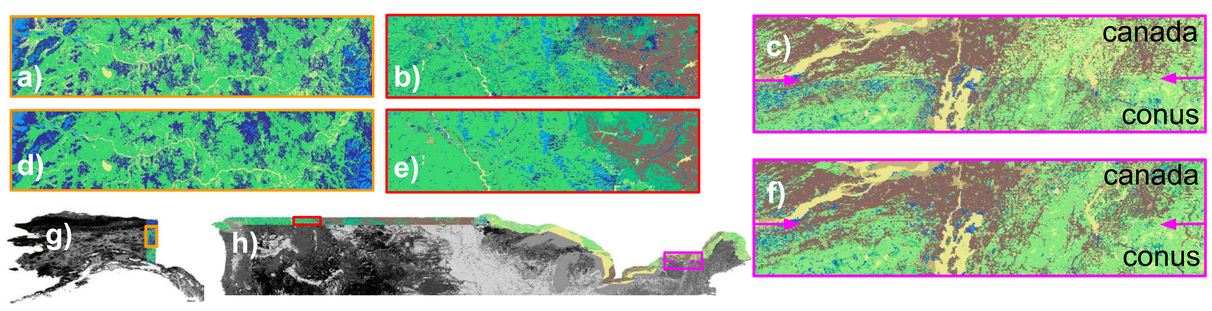

Figure 4: Ground truth EVTPHYS (a–c) compared to gradient boosted trees model inference (d–f) in Canada West (g) and South (h) test regions. Figures (c, f) additionally show land in CONUS across the border which is indicated by the magenta arrows. There, EVT values produced by LANDFIRE seem to exhibit an artificial discontinuity.

Performance is comparable across models, with even simple models exhibiting good metrics in test regions.

Additional testing: Performance varies widely across training regions (Table 2) which may partially be due to discrepancies in LANDFIRE test band data (Fig. 4c).

Limitations:

- As granularity gets finer, performance decreases (Table 3). Notably, AEF targets don't surpass ~40 classes[1] but EVTGP contains 80 classes.
- EVT is the output of a decision tree model[2] which may introduce structural bias.
- Distance (physical and climate) to labeled region conditions performance (table 3)

	Canada South			Canada West			Southern CONUS		
	ACC	J	F1	ACC	J	F1	ACC	J	F1
Logistic Regression	0.67	0.31	0.41	0.82	0.35	0.42	0.59	0.32	0.44
Random Forest	0.69	0.34	0.45	0.83	0.42	0.52	0.68	0.35	0.46
Gradient Boosted Trees	0.69	0.34	0.45	0.83	0.32	0.39	0.64	0.32	0.44
Segmentation Model	0.69	0.34	0.45	0.83	0.37	0.45	0.66	0.36	0.48

Table 2: Model performances for EVTPHYS (13 classes) across 3 distinct test regions. Canada South and Canada West combined comprise the test set in Table 1

	Lat. 41.6 to 38.6			Lat. 38.6 to 35.6			Lat. 35.6 to 33.6		
	ACC	J	F1	ACC	J	F1	ACC	J	F1
Gradient Boosted Trees EVTPHYS (13 classes)	0.76	0.42	0.53	0.69	0.34	0.45	0.55	0.26	0.37
Segmentation Model EVTGP (80 classes)	0.58	0.13	0.19	0.48	0.09	0.14	0.34	0.06	0.09

Table 3: Test results (Accuracy, Jaccard and F1 scores) for different models across distinct latitude bands within CONUS.

Key Takeaways

- We present a flexible and scalable pipeline that leverages AEF embeddings to extend valuable, but limited, geospatial datasets to new regions
- Using the pipeline, we extend vegetation maps from the USA into Canada
- Performance is conditioned by class granularity and distance to labeled data
- For this task, 'simple' models perform on par with U-Net segmentation model