Patch CNN on Sentinel-2 data
versus
FC NN trained on AlphaEarth embeddings
(recap for Arjan)

John Toth

Biome-10. Montane Grasslands and Shrublands

Samples	Data	Model	Device	Epochs	Batch	Time	Loss	Accuracy		F1-score	
					size	min		test	train	macro	weighted
Biome-10											
24494	**Sentinel-2**	**CNN Patch**	**cpu**	**500**	**32**	**12**	**0.27**	**89%**	**90%**	**0.68**	**0.89**
24494	Sentinel-2	CNN Patch5	cpu	500	32	15	0.35	88%	88%	0.59	0.87
24878	**AlphaEarth**	**Fully Connected**	**cpu**	**500**	**10**	**18**	**0.02**	**98%**	**99%**	**0.93**	**0.98**

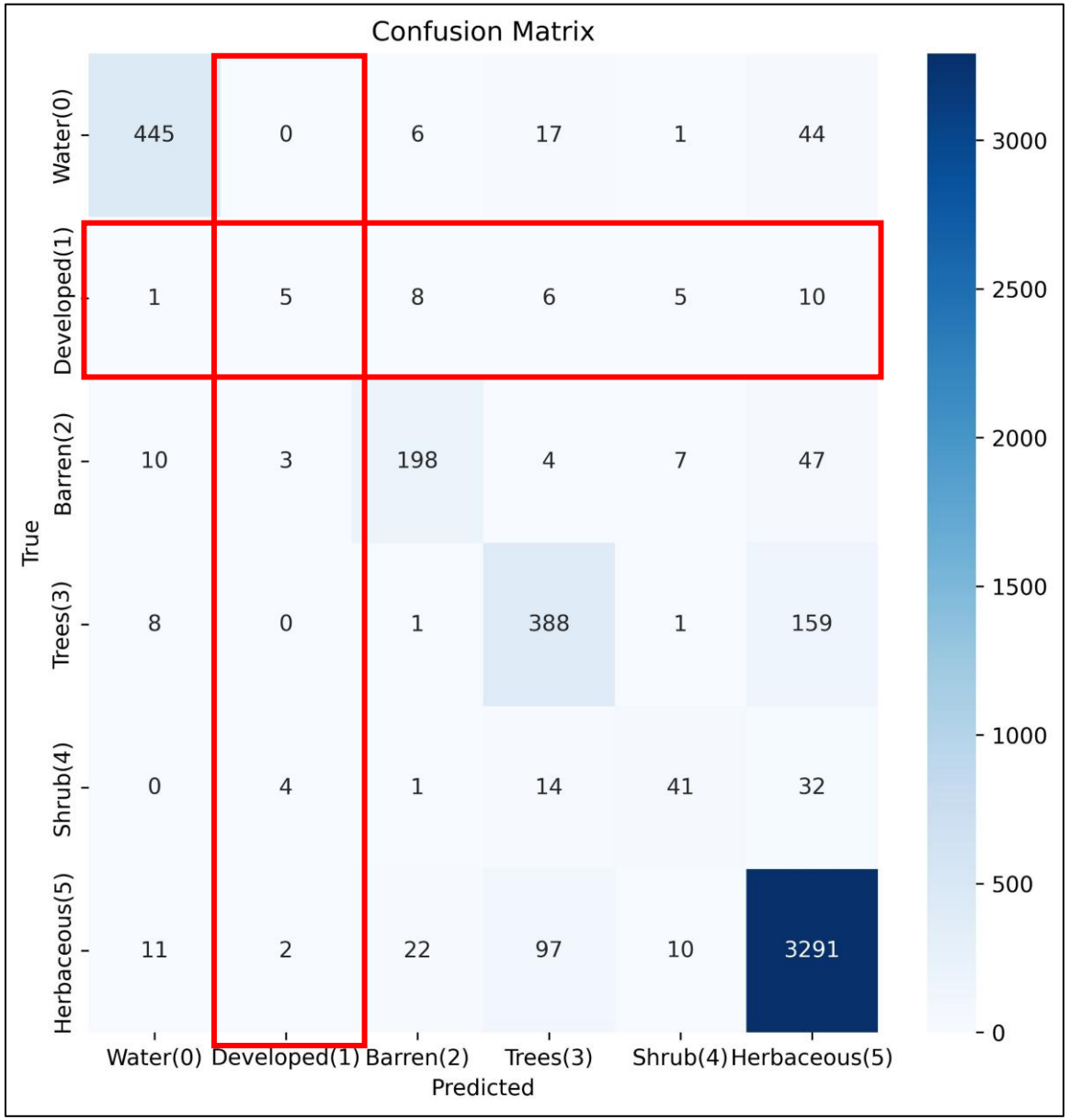
Level 1	Biome 2	Biome 7	Biome 8	Biome 10	Biome 12	Biome 13	All
Water [0]	10.7%	10.2%	2.7%	10.5%	0.6%	1.4%	7.3%
Developed [1]	0.9%	0.9%	3.2%	0.5%	3.9%	2.9%	1.7%
Barren [2]	0.4%	1.5%	2.2%	5.1%	0.8%	25.8%	4.7%
Trees [3]	47.4%	26.2%	7.0%	11.4%	52.5%	4.0%	26.0%
Shrub [4]	1.4%	13.8%	2.5%	1.9%	7.5%	23.1%	10.1%
Herb [5]	39.2%	47.5%	82.4%	70.6%	34.7%	42.9%	50.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Highly imbalanced dataset. 120 ‘developed’ samples out of 24,000 total.

John Toth

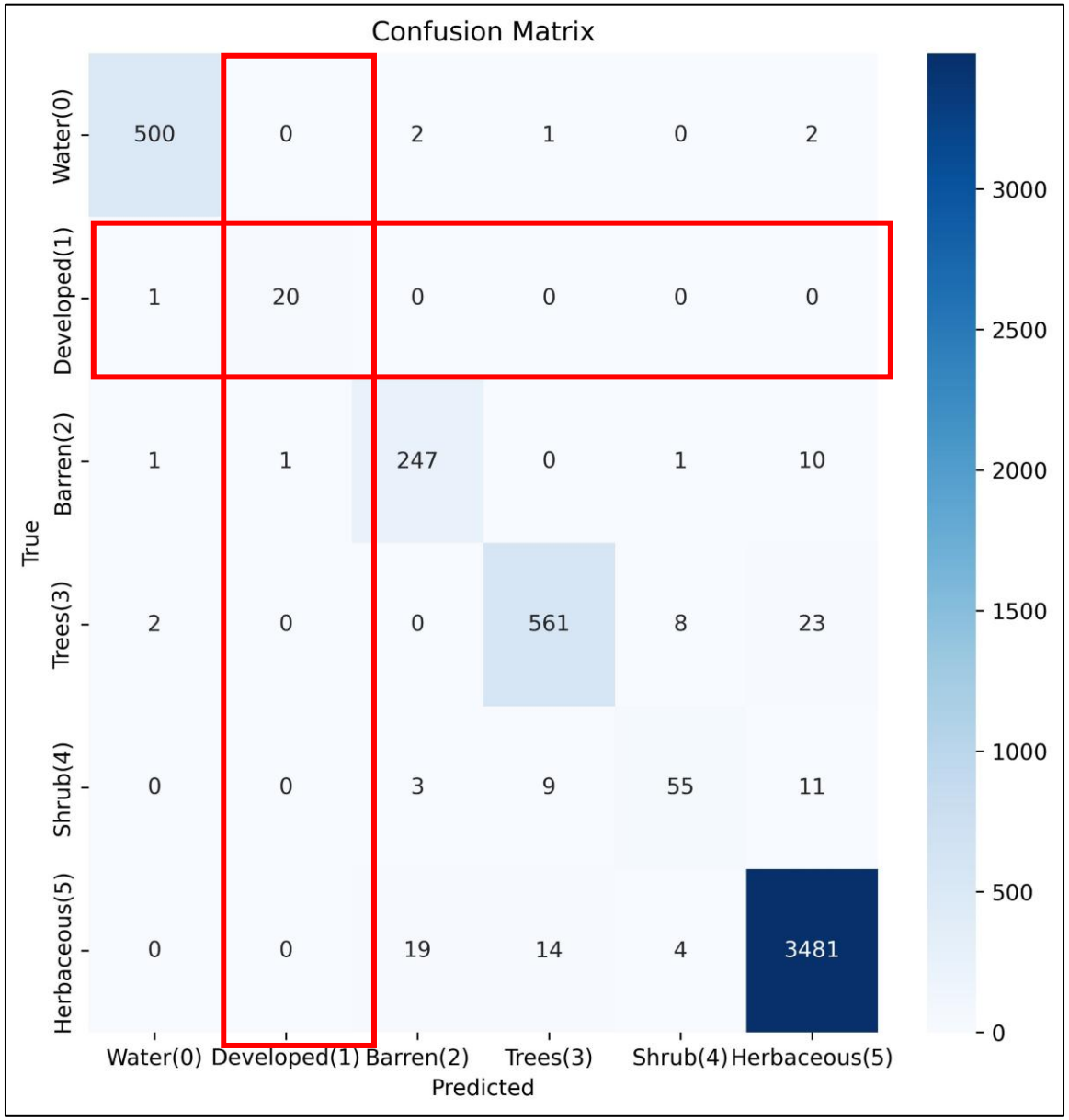
P=0.36, R=0.14, F1-class = 0.2

Sentinel-2



P=0.95, R=0.95, F1-class = 0.95

AlphaEarth

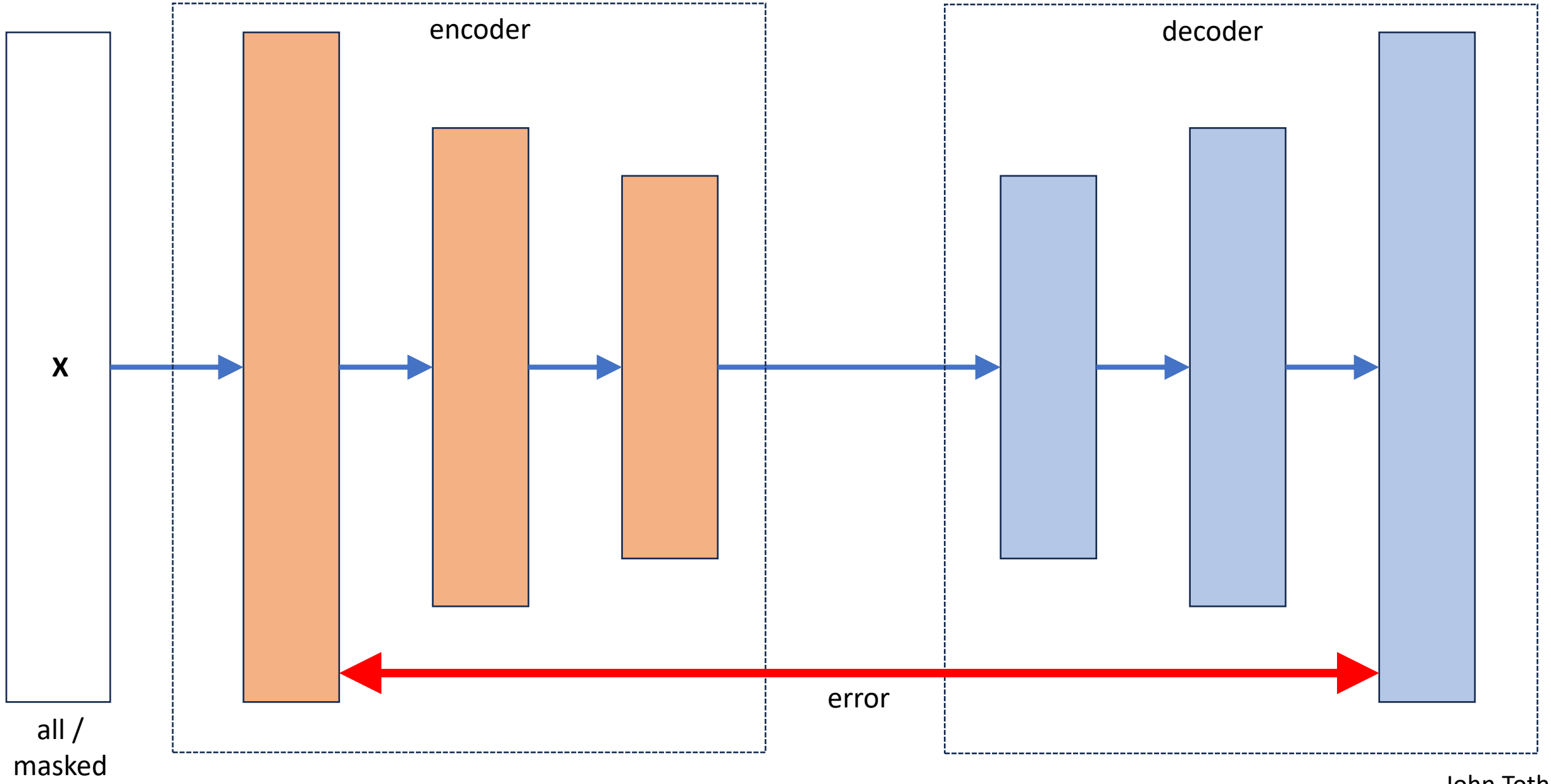


AlphaEarth Geospatial Foundational Model - Inputs

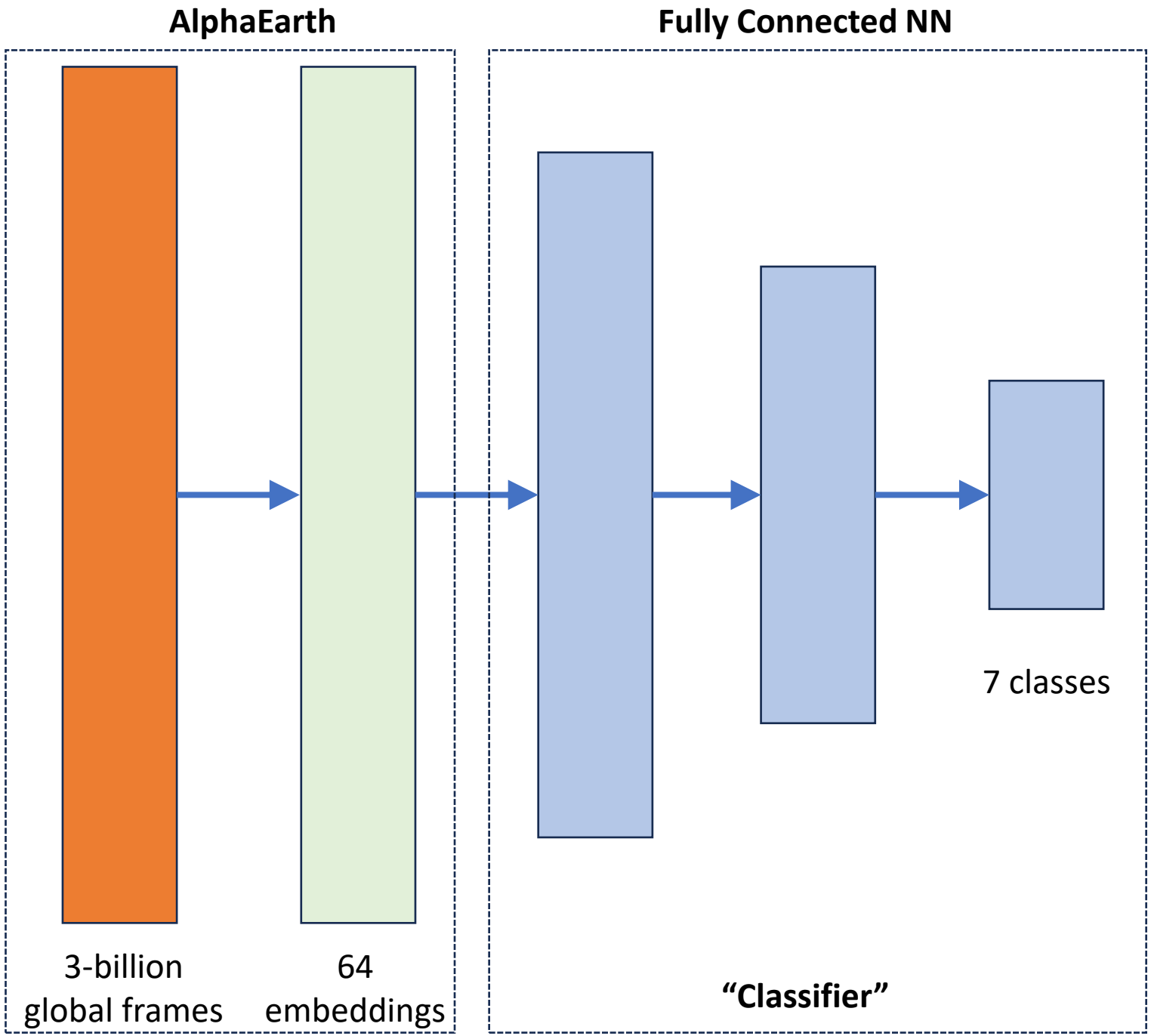
Optical Satellite Imagery - Sentinel-2,Landsat 8, Landsat 9
Radar Data - Sentinel-1, PALSAR2
LiDAR (Light Detection and Ranging) - GEDI (Global Ecosystem Dynamics Investigation): Provides rasterized metrics on canopy height.
Climate & Environmental Reanalysis - ERA5-Land: Monthly aggregated data on atmospheric and land conditions (e.g., temperature, precipitation). GRACE (Gravity Recovery and Climate Experiment): Data on gravity anomalies, which can relate to large-scale water movement.
Elevation Data - GLO-30 (Copernicus DEM)
Textual & Annotated Data - Geotagged Wikipedia: Textual context linked to specific geographic coordinates. NLCD (National Land Cover Database): Used as an annotated data source for land cover classification.
"While the NLCD is specific to the United States, the model learns the fundamental relationships between sensor data and land cover types, which it can then apply globally."

John Toth

Autoencoder – Self Supervised



John Toth



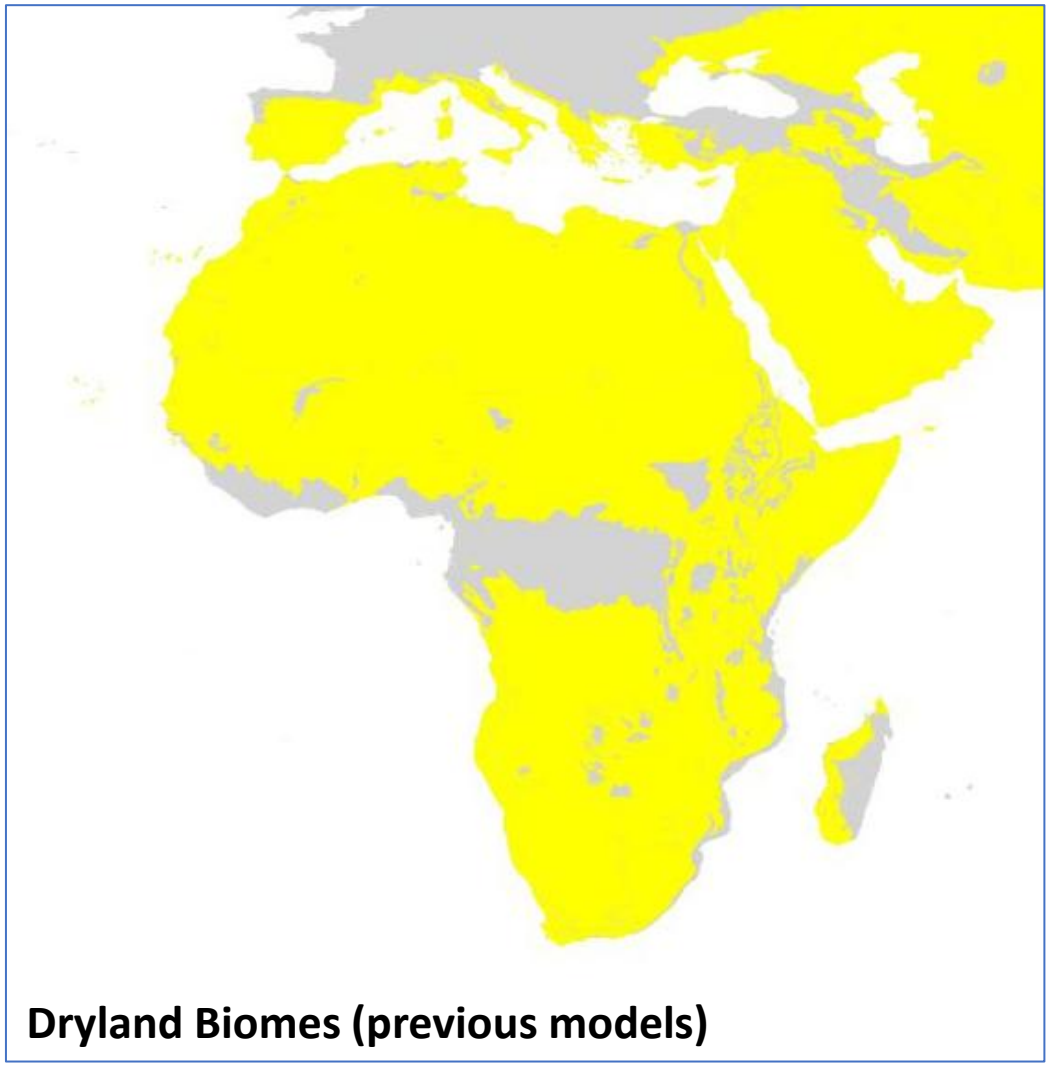
Earth’s surface contains 1,500 Billion 10x10m patches

“AEF was trained on over **3 billion individual image frames** sampled from over **5 million locations globally**. By using this large, diverse sample across different modes and time periods, the model learns a generalized representation of Earth's surface.”

John Toth

All African GLanCE samples

	All Africa	Level1	Level2	Level3
0	Water	26049	26049	26049
1	Developed	11533	11533	11533
2	Barren	4735	2795	4735
3	Trees	73069	2238	73069
4	Shrub	4701	4701	4701
5	Herb (grass)	114050	17168	17168
6	Agriculture		96882	96882
	Total	**234137**	**161366**	**234137**
	All Africa	**Level1**	**Level2**	**Level3**
0	Water	11.1%	16.1%	11.1%
1	Developed	4.9%	7.1%	4.9%
2	Barren	2.0%	1.7%	2.0%
3	Trees	31.2%	1.4%	31.2%
4	Shrub	2.0%	2.9%	2.0%
5	Herb (grass)	48.7%	10.6%	7.3%
6	Agriculture		60.0%	41.4%
	Total	**100.0%**	**100.0%**	**100.0%**



Dryland Biomes (previous models)

John Toth

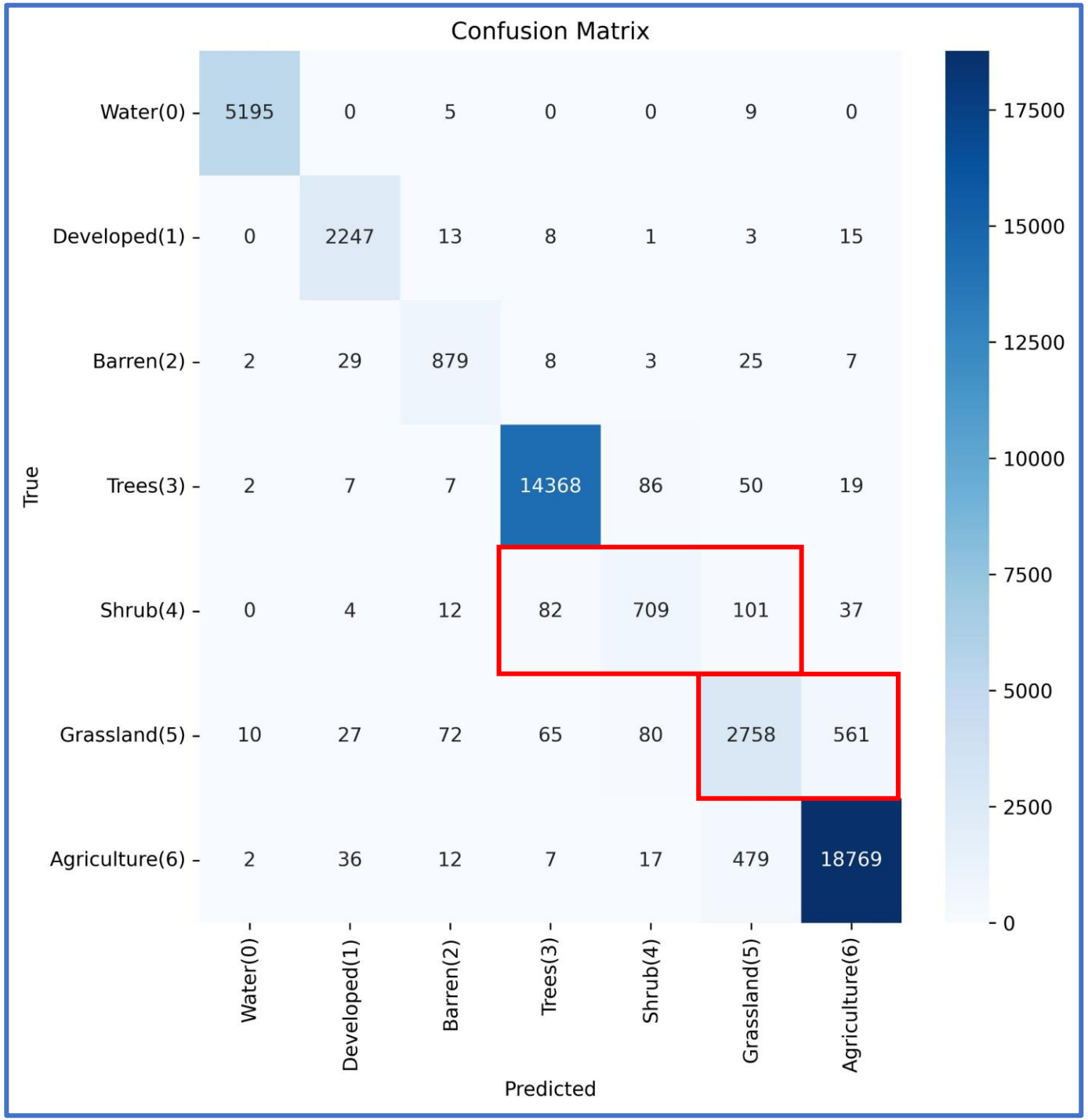
Model Accuracy and F-Score

Region	Samples	Data	Model	Device	Epochs	Batch	Time	Loss	Accuracy		F1-score	
						size	min		test	train	macro	weighted
Biome-2	28353	Sentinel-2	CNN Patch	cpu	500	32	15	0.31	88%	89%	0.62	0.87
	28484	AlphaEarth	Fully Connected	cpu	500	10	21	0.02	97%	99%	0.91	0.97
Biome-7	89297	Sentinel-2	CNN Patch	cpu	500	32	51	0.36	86%	87%	0.77	0.86
	89724	AlphaEarth	Fully Connected	cpu	500	10	67	0.08	96%	97%	0.93	0.96
Biome-8	20381	Sentinel-2	CNN Patch	cpu	500	32	11	0.31	88%	90%	0.58	0.86
	20335	AlphaEarth	Fully Connected	cpu	500	10	15	0.02	97%	99%	0.91	0.97
Biome-10	24494	Sentinel-2	CNN Patch	cpu	500	32	12	0.27	89%	90%	0.68	0.89
	24878	AlphaEarth	Fully Connected	cpu	500	10	18	0.02	98%	99%	0.93	0.98
Biome-12	27006	Sentinel-2	CNN Patch	cpu	500	32	16	0.34	87%	88%	0.81	0.86
	27006	Sentinel-2	CNN Patch	cpu	1000	32	17	0.27	87%	90%	0.80	0.87
	26953	AlphaEarth	Fully Connected	cpu	500	10	20	0.07	96%	98%	0.90	0.96
Biome-13	26628	Sentinel-2	CNN Patch	cpu	500	32	16	0.59	75%	77%	0.68	0.74
	26628	Sentinel-2	CNN Patch	cpu	1000	32	16	0.51	75%	79%	0.71	0.75
	26716	AlphaEarth	Fully Connected	cpu	500	10	19	0.09	93%	97%	0.92	0.93
	26716	AlphaEarth68	Fully Connected	cpu	500	10	19	0.09	93%	96%	0.92	0.93
Biome-All	216159	Sentinel-2	CNN Patch	cpu	500	32	131	0.47	82%	82%	0.76	0.82
	217090	AlphaEarth	Fully Connected	cpu	500	10	182	0.1	96%	96%	0.94	0.96
All_Balanced	74043	AlphaEarth	Fully Connected	cpu	500	10	54	0.09	95%	97%	0.95	0.95
All_Africa	**234137**	**AlphaEarth**	**Fully Connected**	**cpu**	**500**	**10**	**170**	**0.09**	**96%**	**97%**	**0.91**	**0.96**

John Toth

All Africa

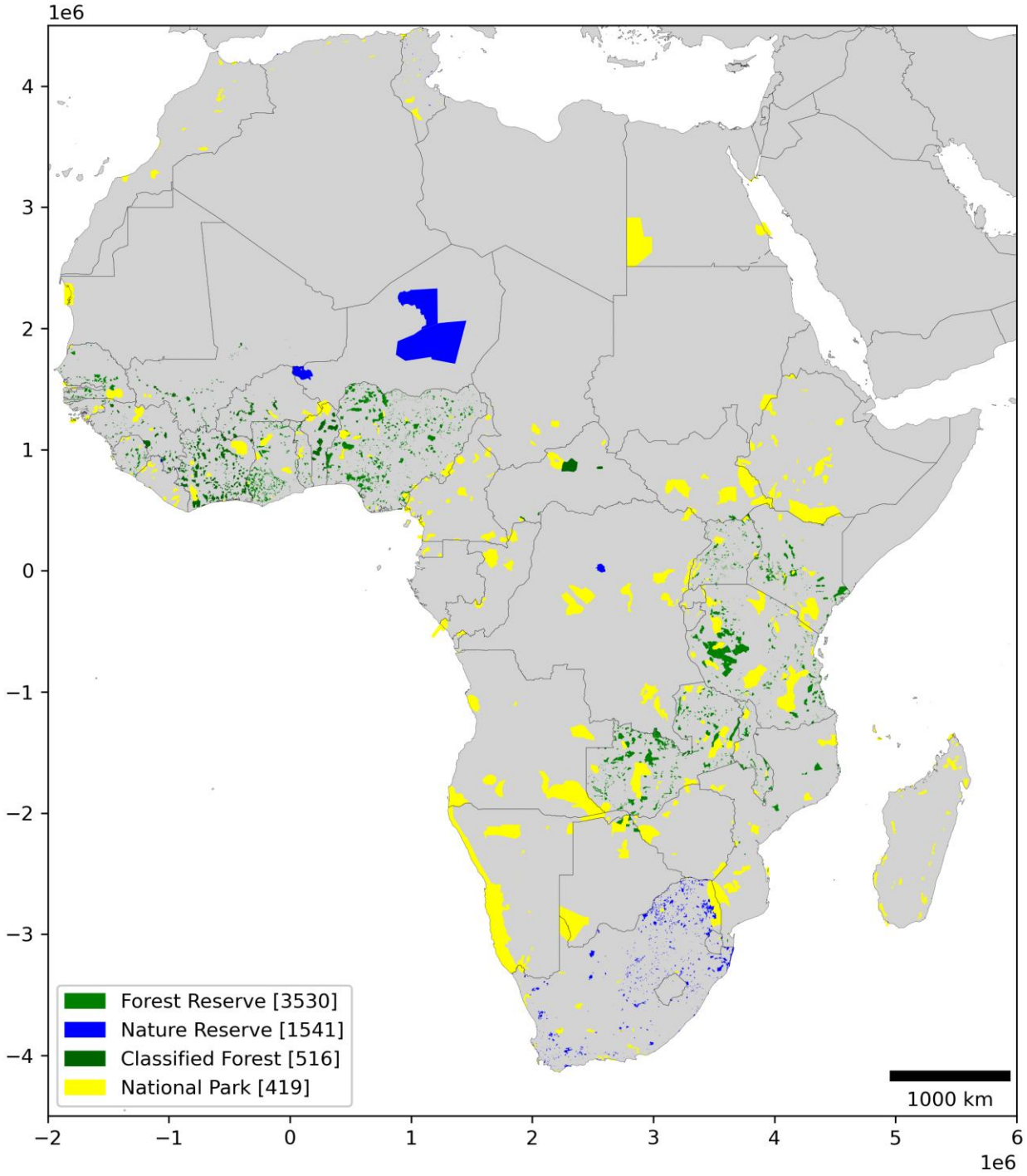
F1_water = 1.0
F1_developed = 0.97
F1_barren = 0.90
F1_trees = 0.99
F1_shrub = 0.77
F1_grassland = 0.79
F1_agriculture= 0.97



John Toth

African Parks

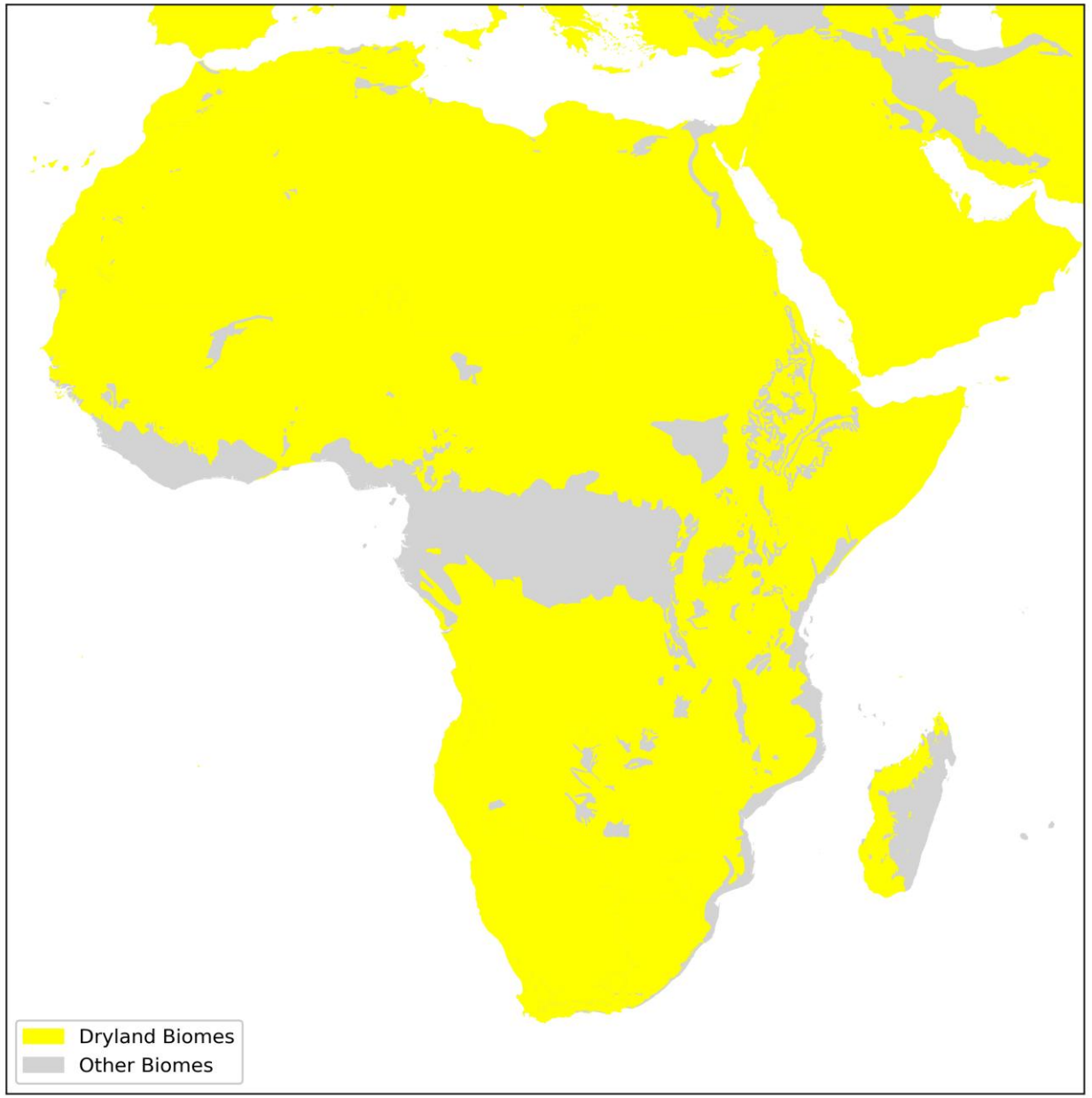
John Toth



World Database on Protected Areas (WDPA). Protected Areas in Africa. The World Database on Protected Areas (WDPA) and World Database on Other Effective Area-based Conservation Measures (WD-OECM) [Online], November 2025, Cambridge, UK: UNEP-WCMC and IUCN. Available at: www.protectedplanet.net.

John Toth

Dryland Biomes

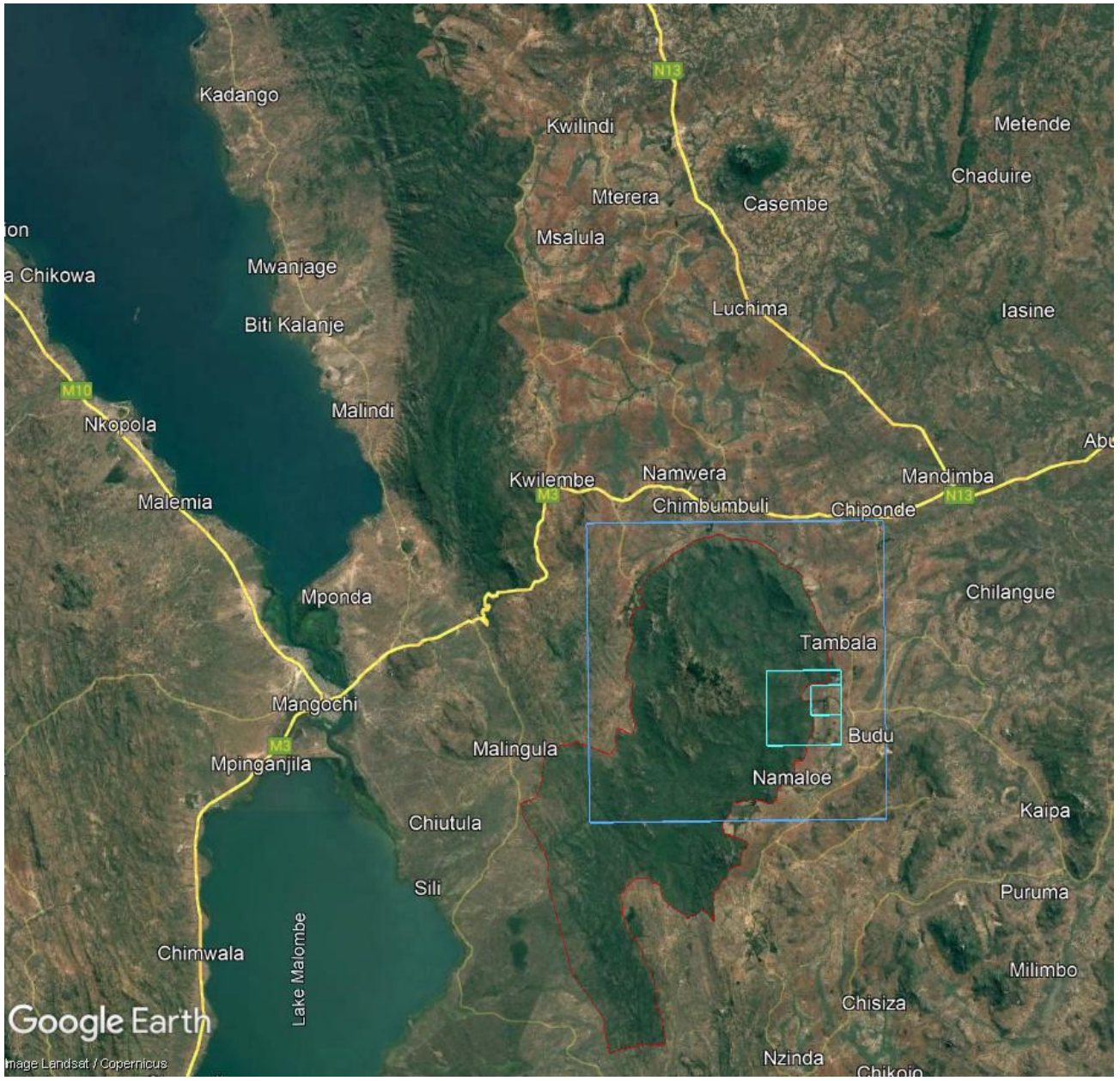


John Toth

Liwonde National Park, Malawai (protectedplanet.net)

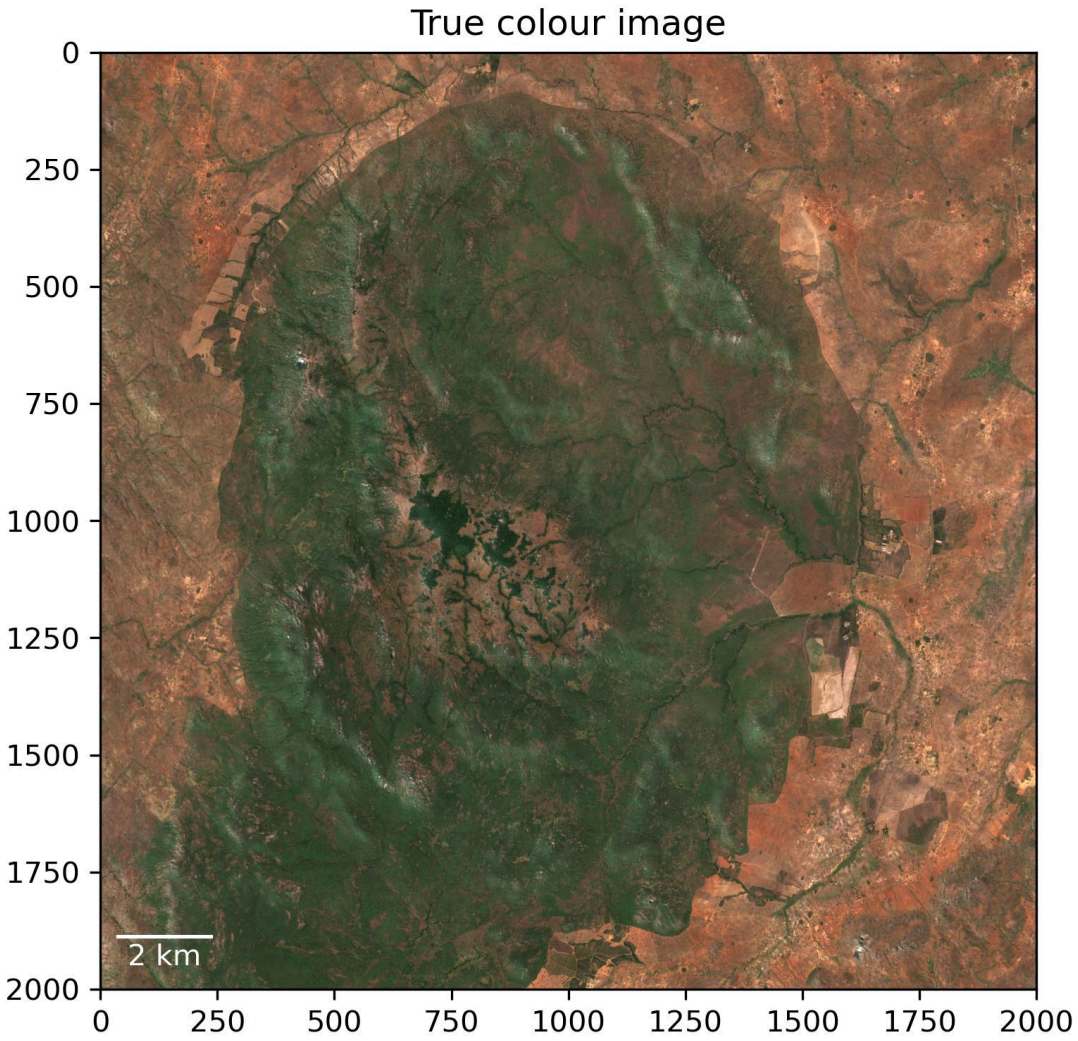
John Toth

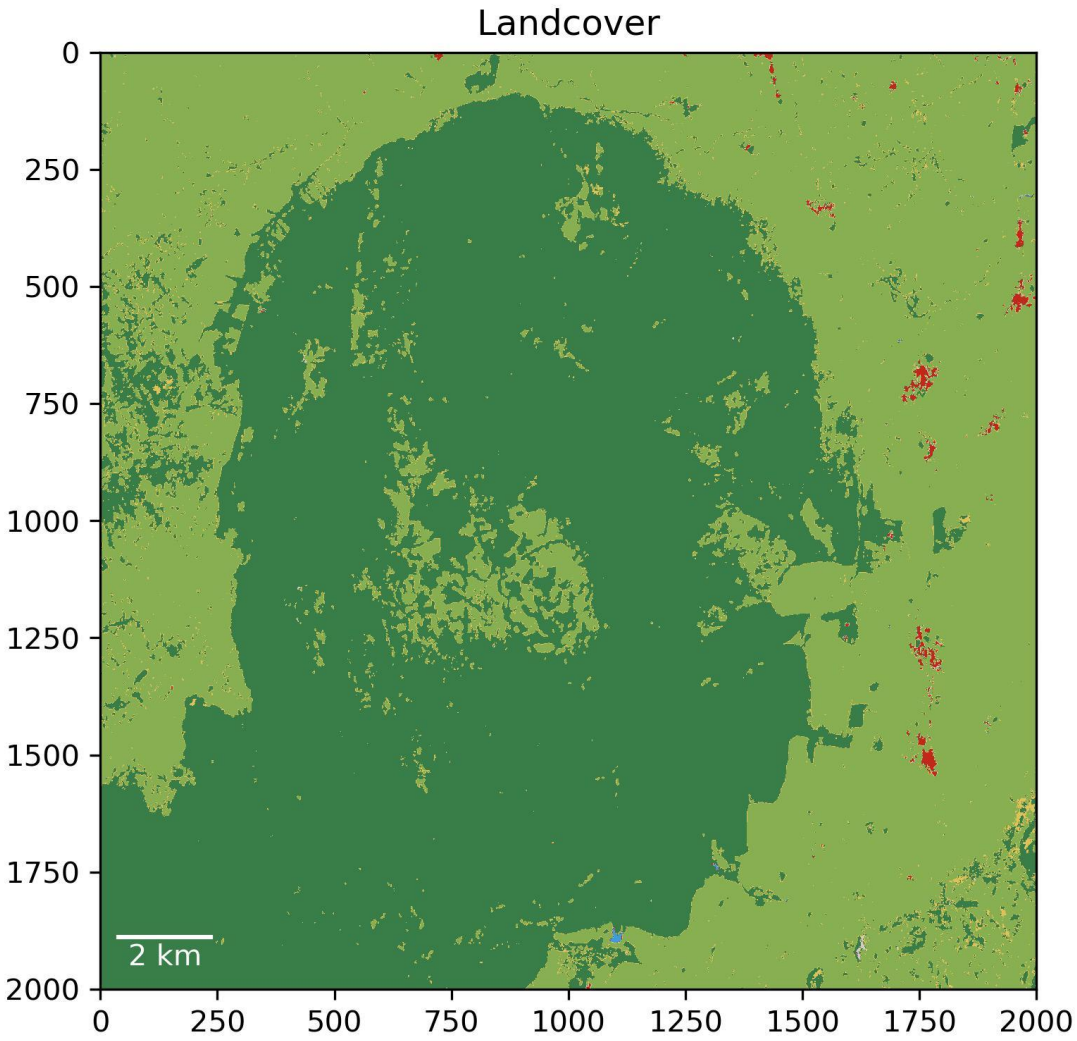
Mangochi Forest Reserve, Malawi



John Toth

Mangochi Forest Reserve, Malawi, 6 classes, 20x20 km

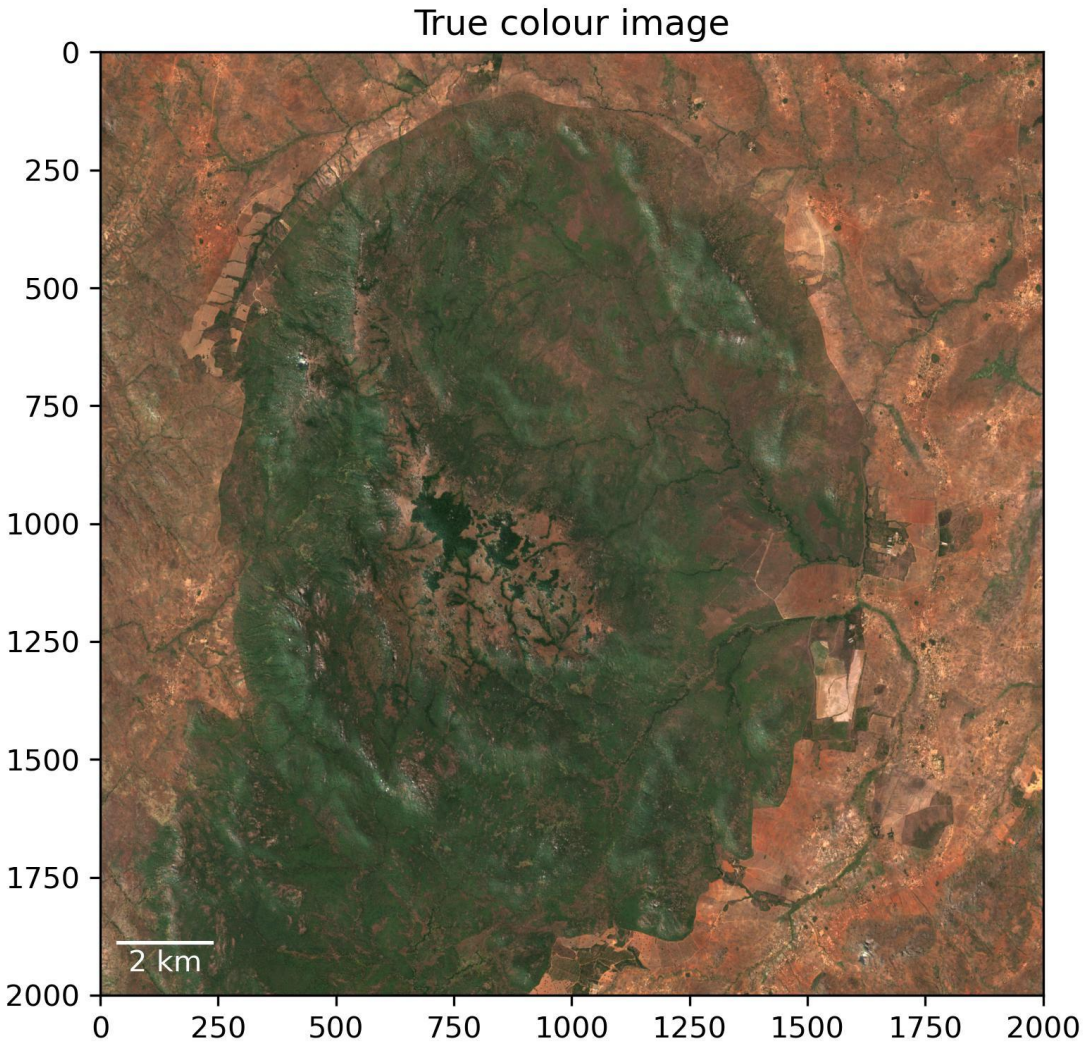


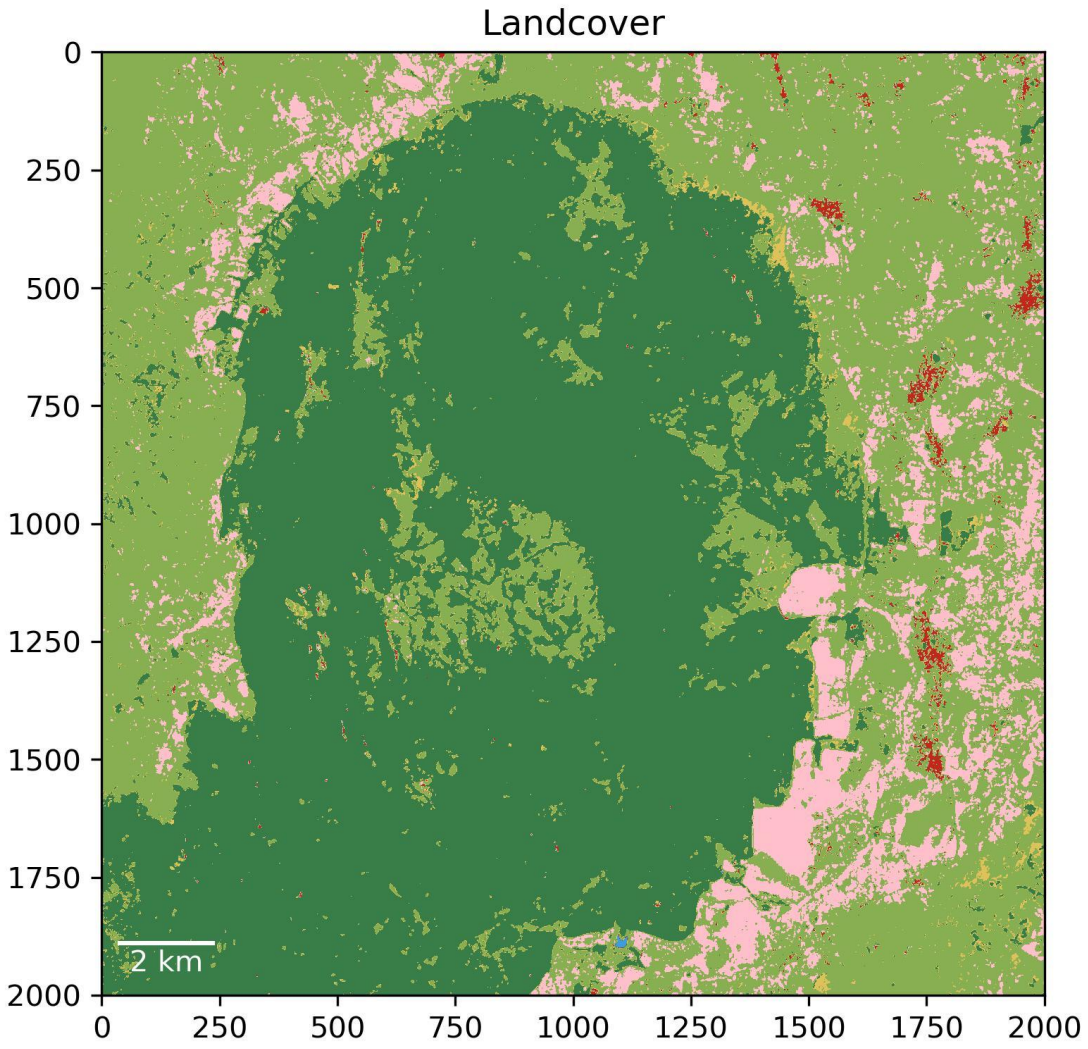


Trained on Dryland Biome Samples

John Toth

Mangochi Forest Reserve, Malawi, 7 classes, 20x20 km

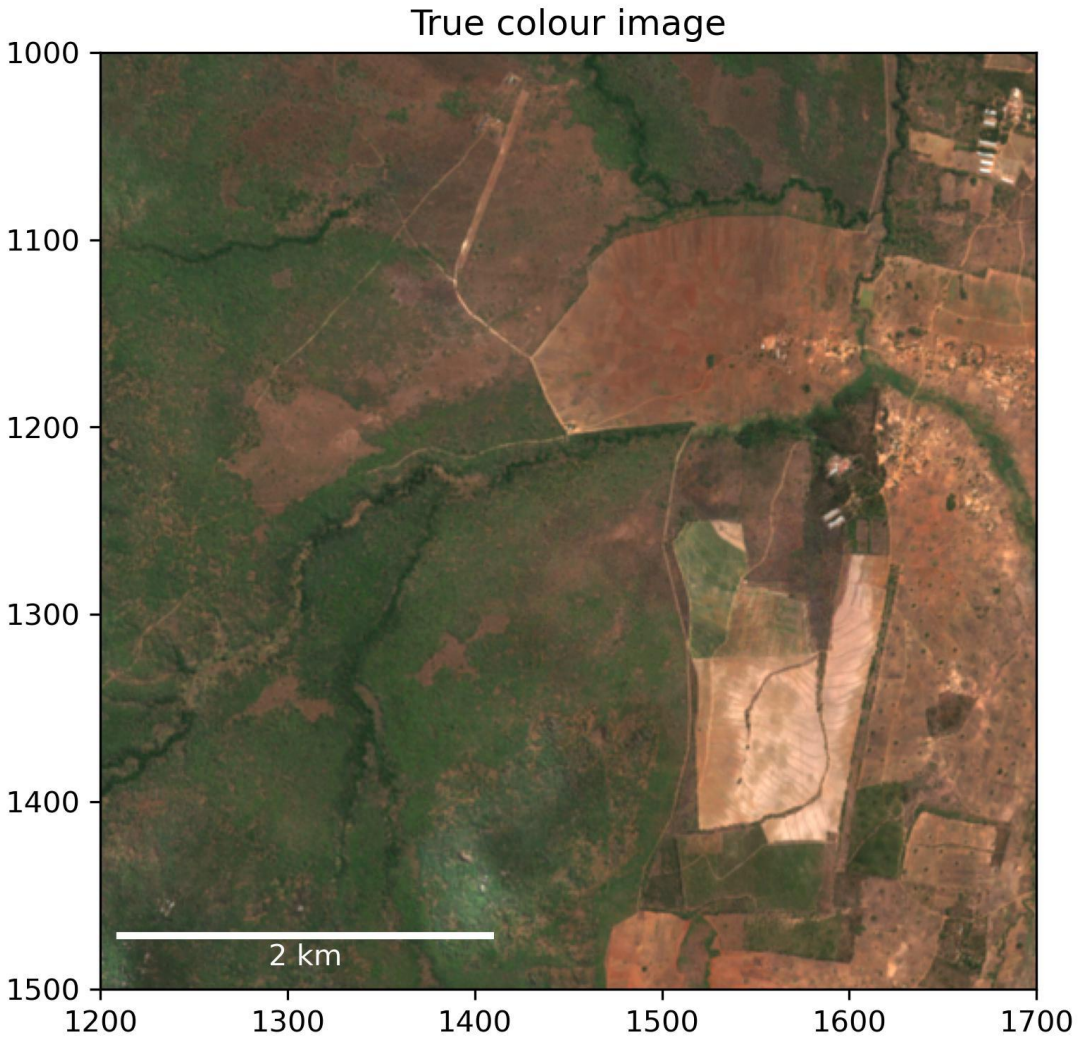


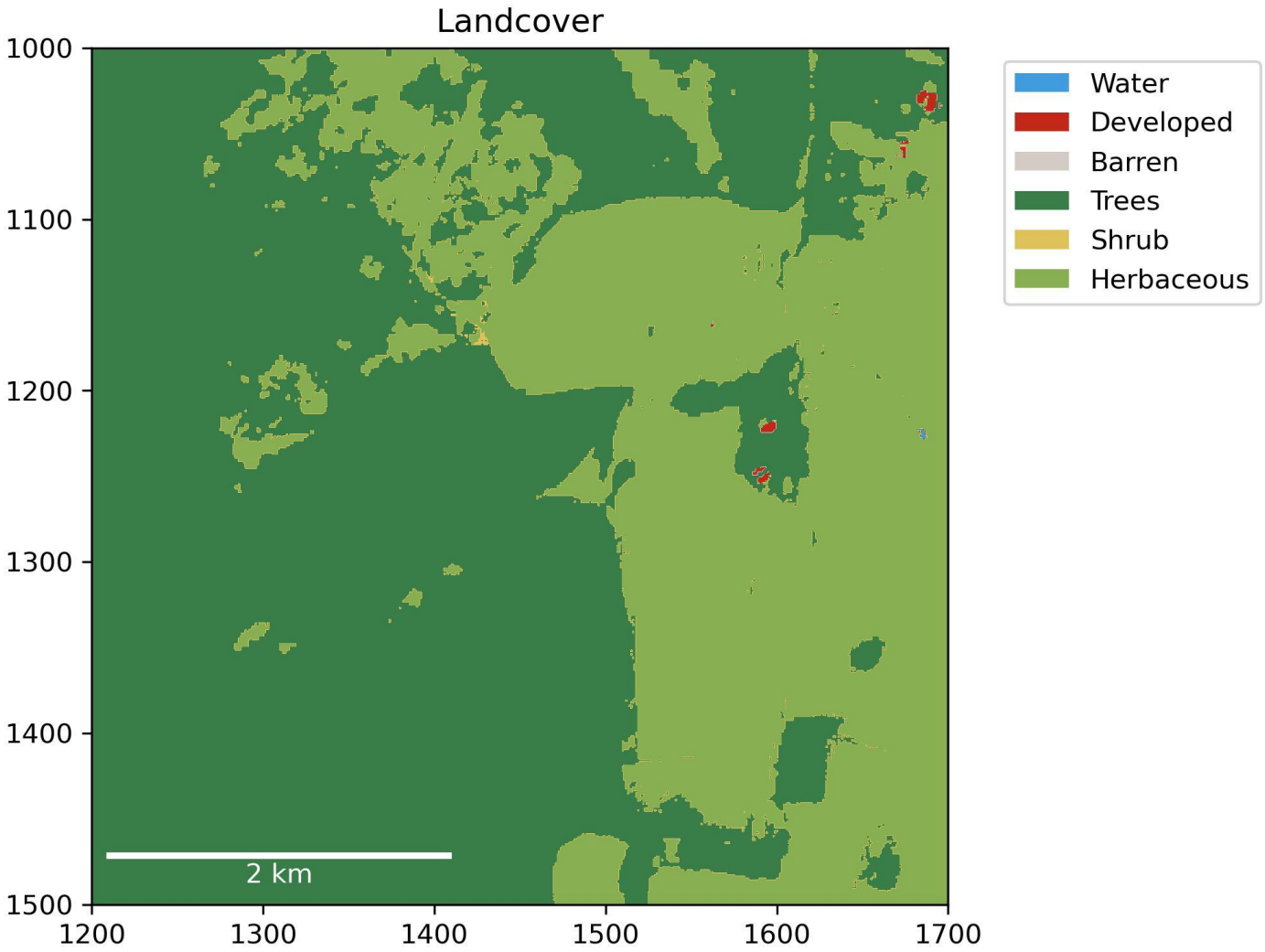


Trained on all African Samples

John Toth

Mangochi Forest Reserve, Malawi, 6 classes, 5x5 km

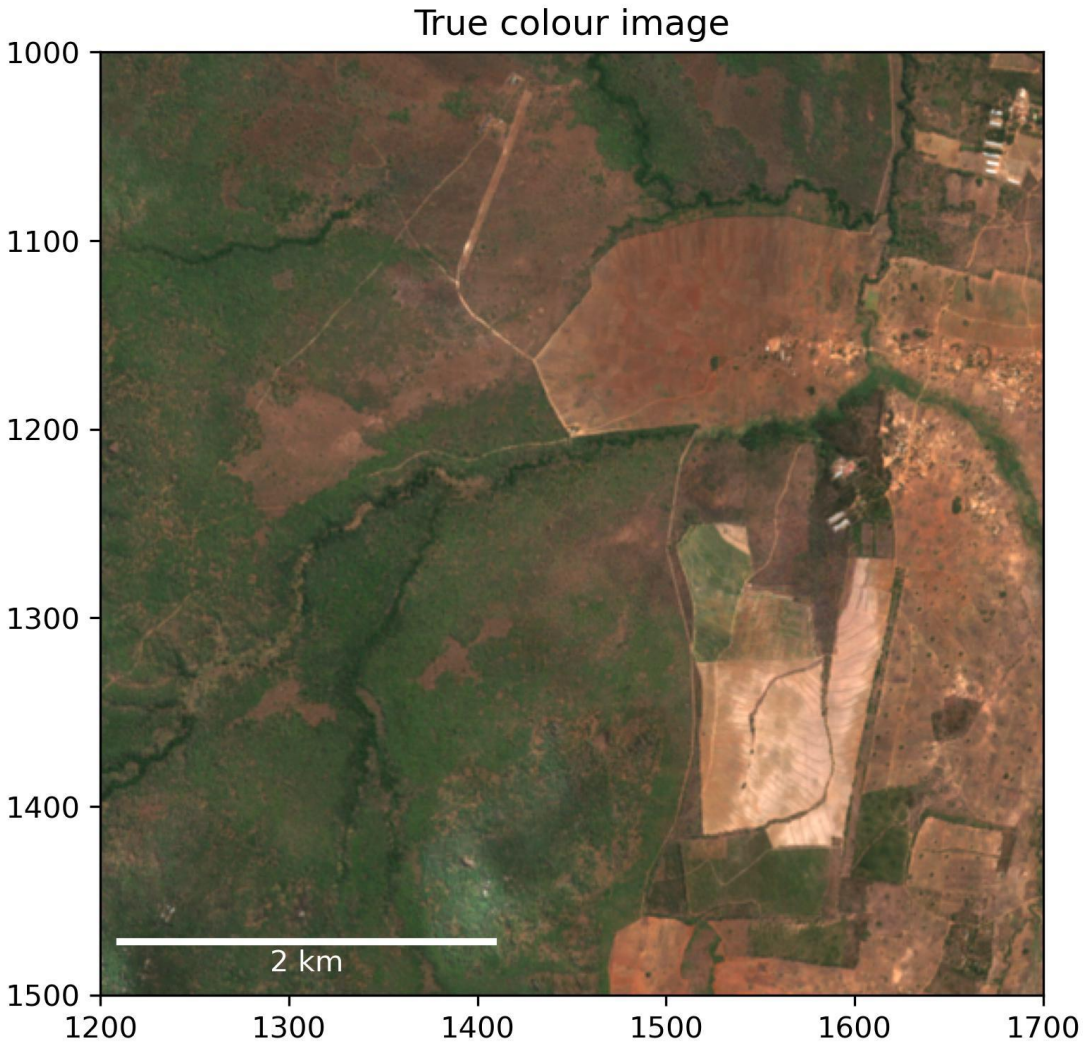


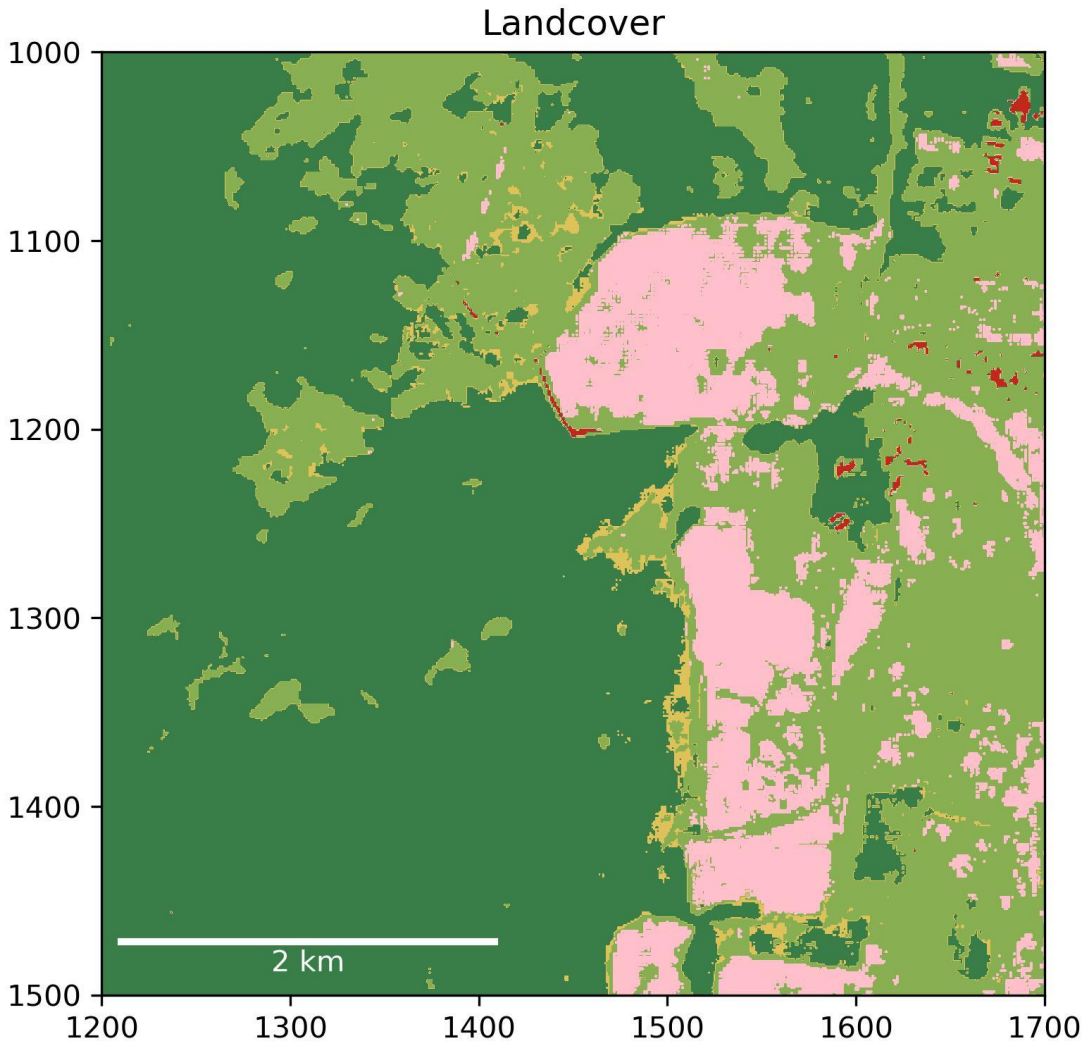


Trained on Dryland Biome Samples

John Toth

Mangochi Forest Reserve, Malawi, 7 classes, 5x5 km

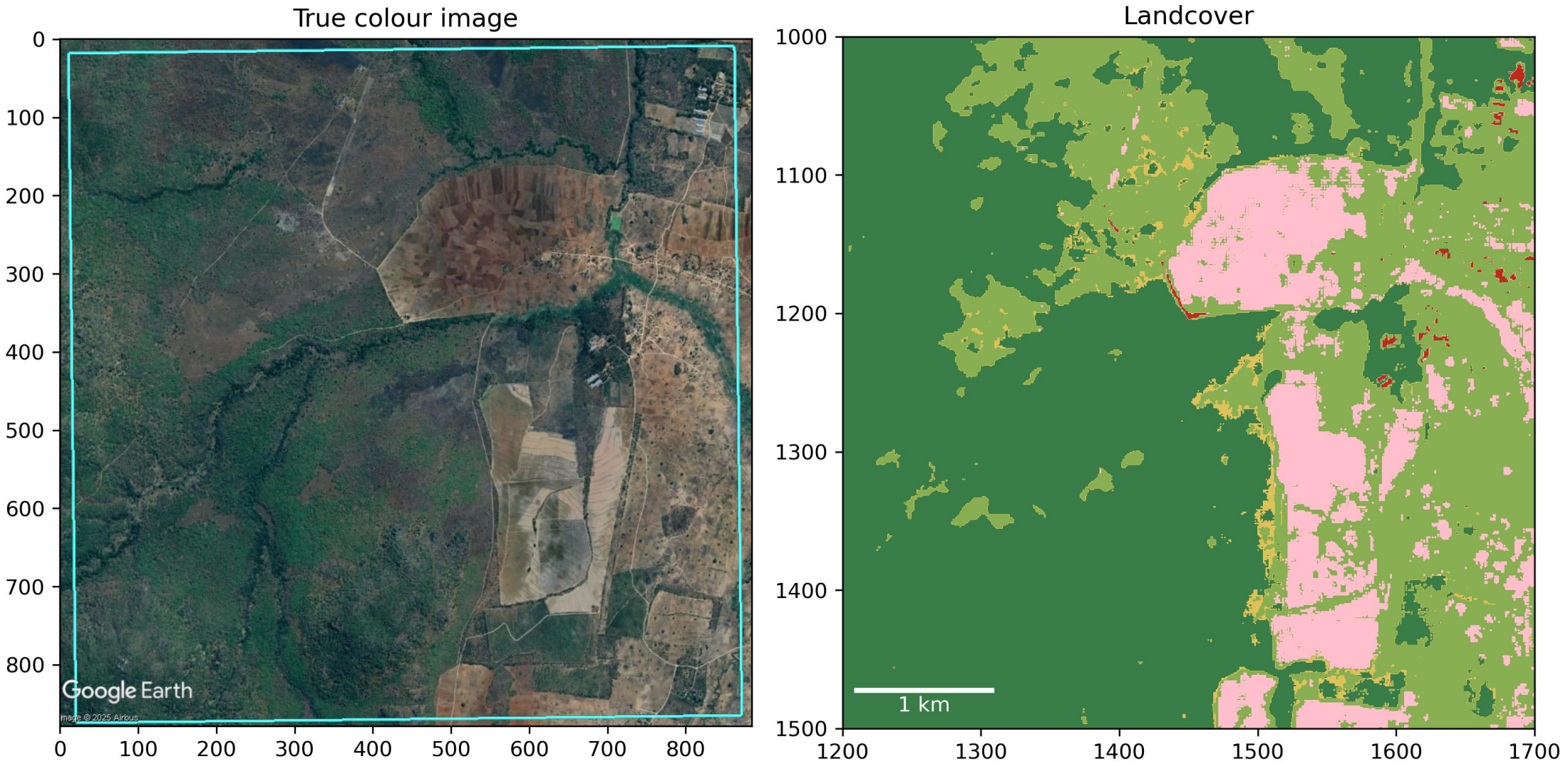




Trained on all African Samples

John Toth

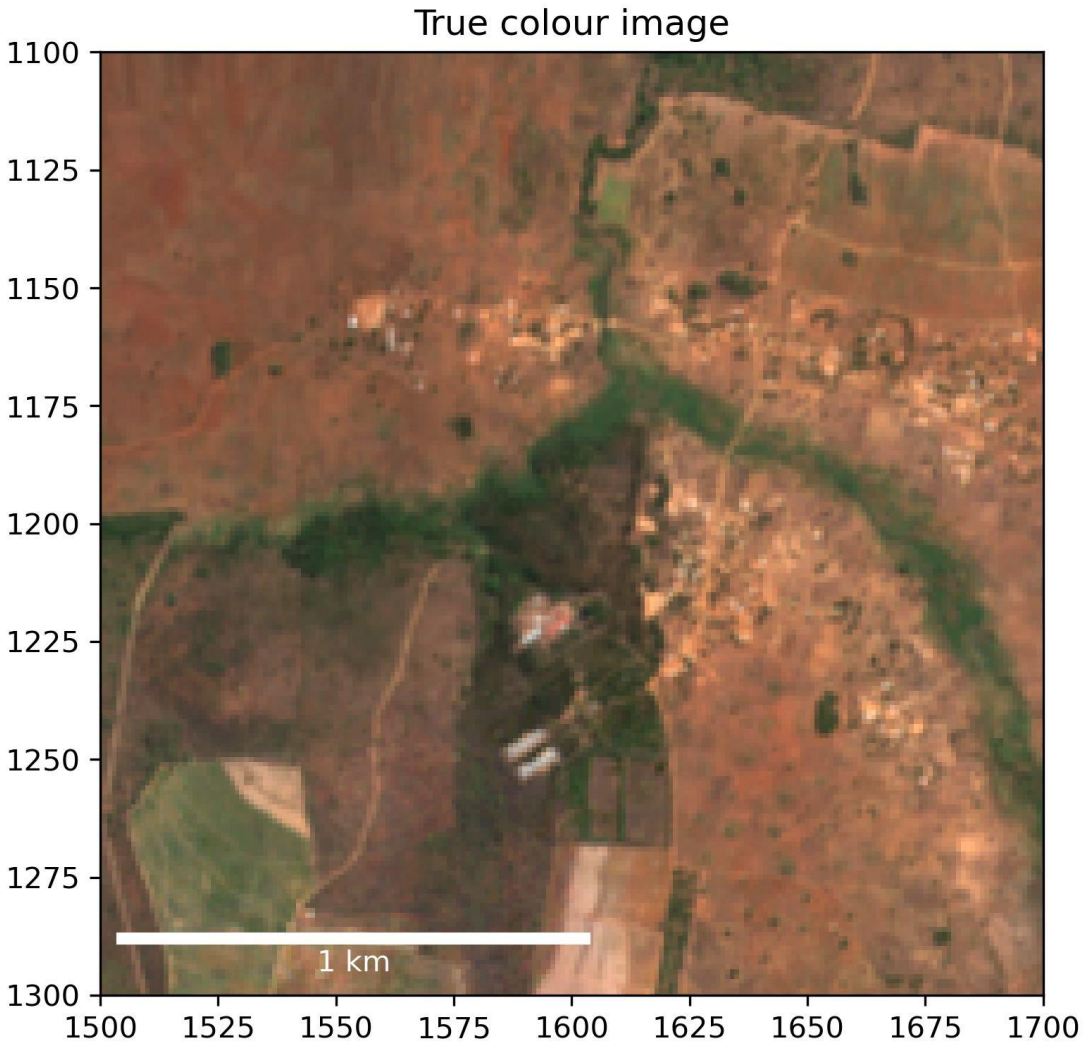
Mangochi Forest Reserve, Malawi, 7 classes, 5x5 km

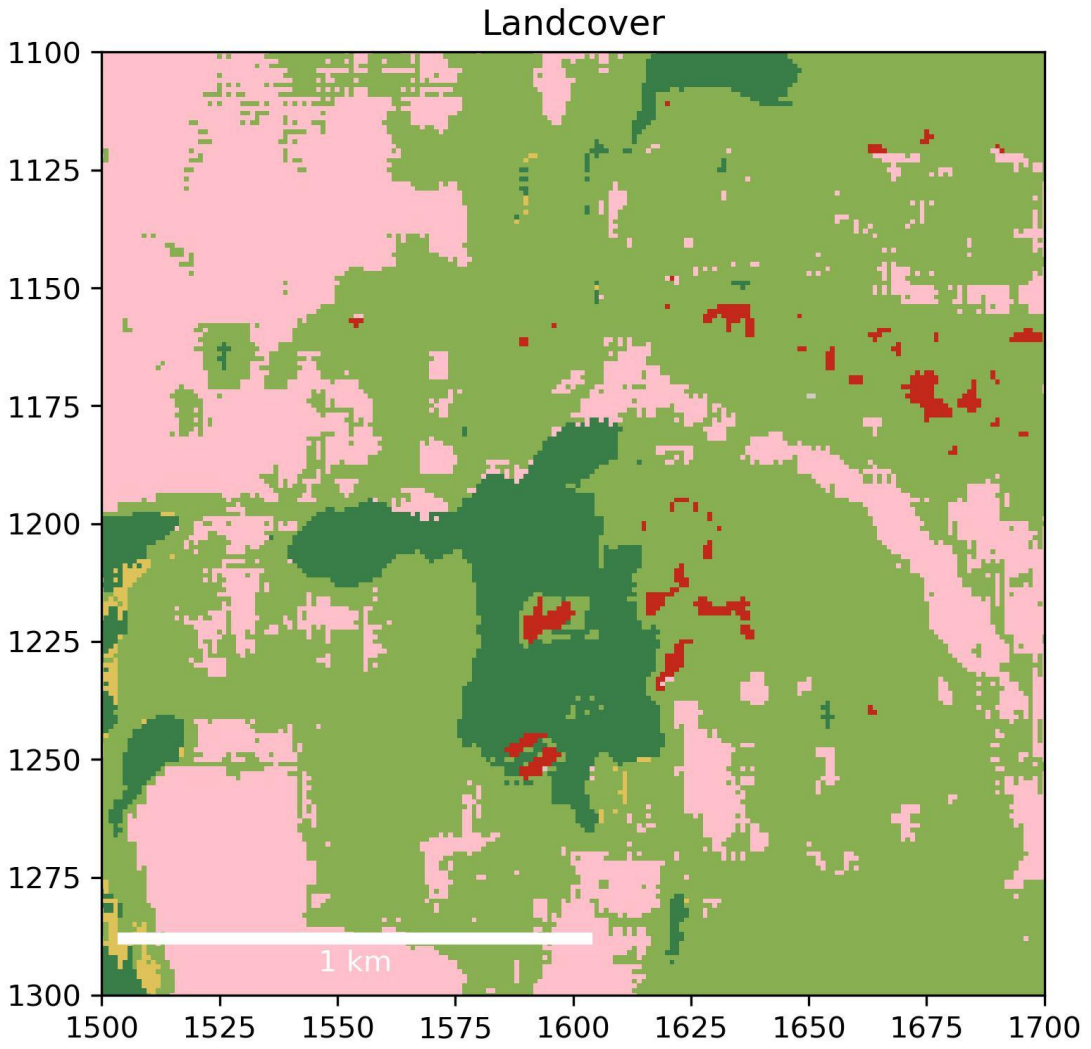


Trained on all African Samples

John Toth

Mangochi Forest Reserve, Malawi, 7 classes, 2x2 km

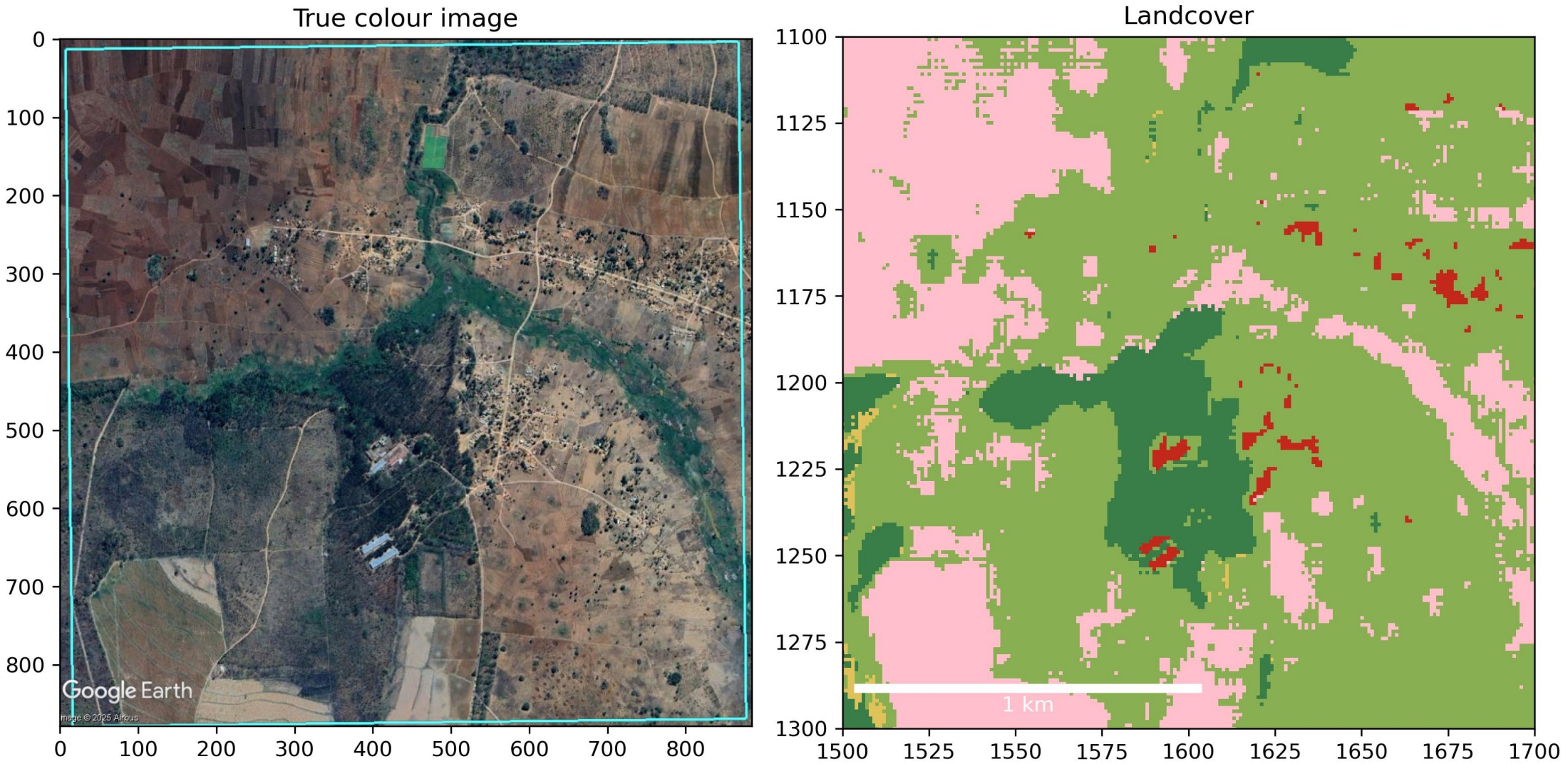




Trained on all African Samples

John Toth

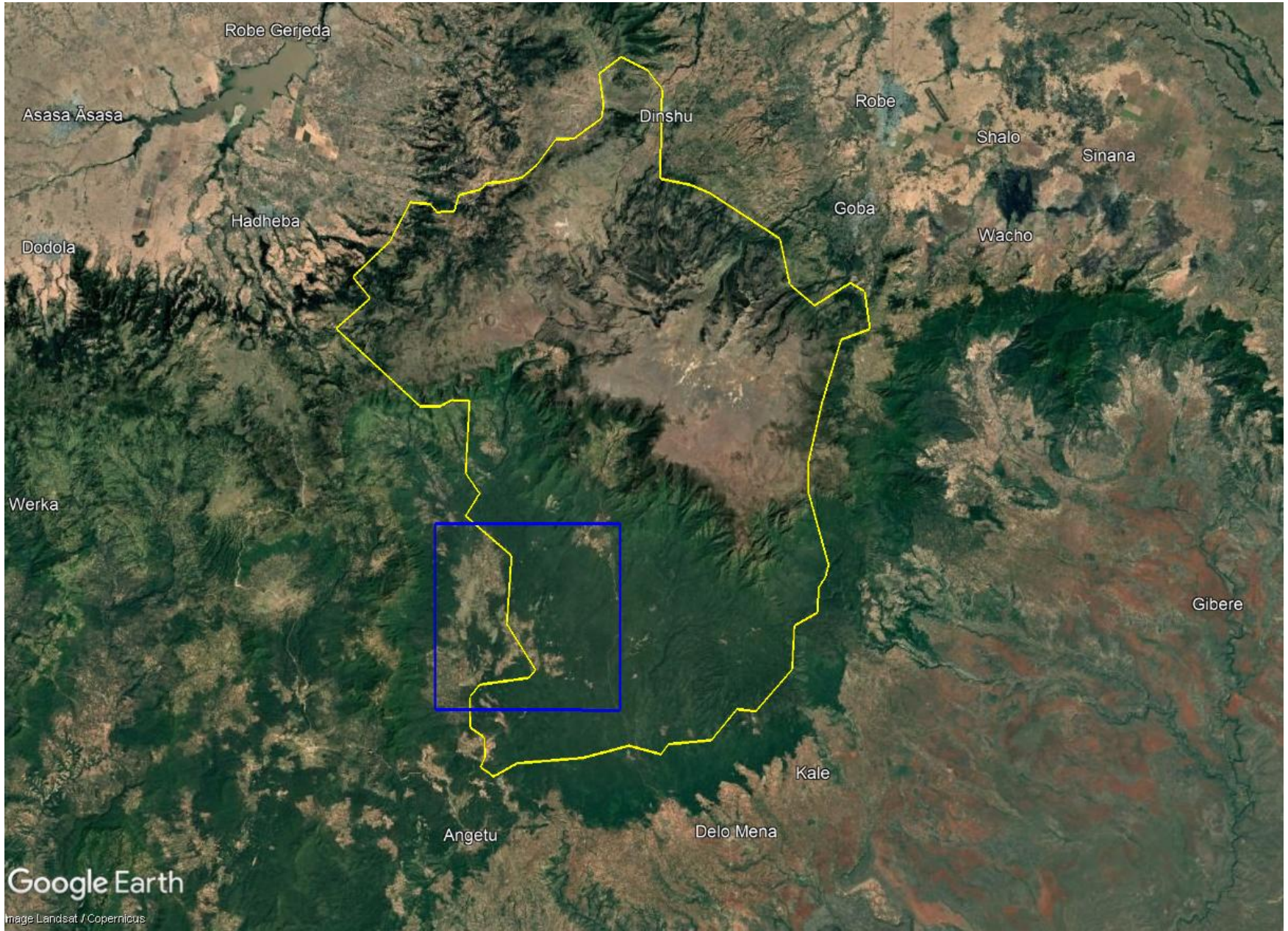
Mangochi Forest Reserve, Malawi, 7 classes, 2x2 km



Trained on all African Samples

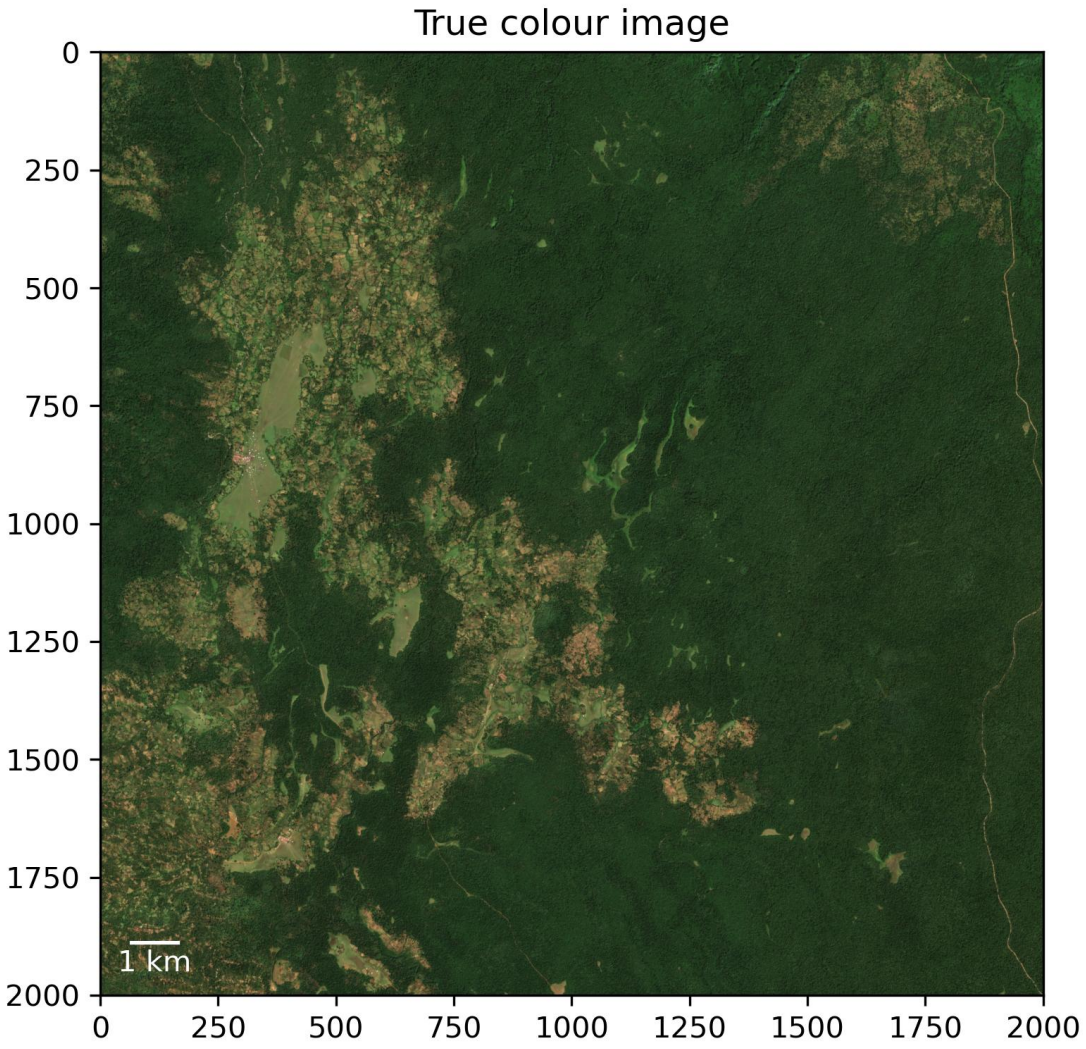
John Toth

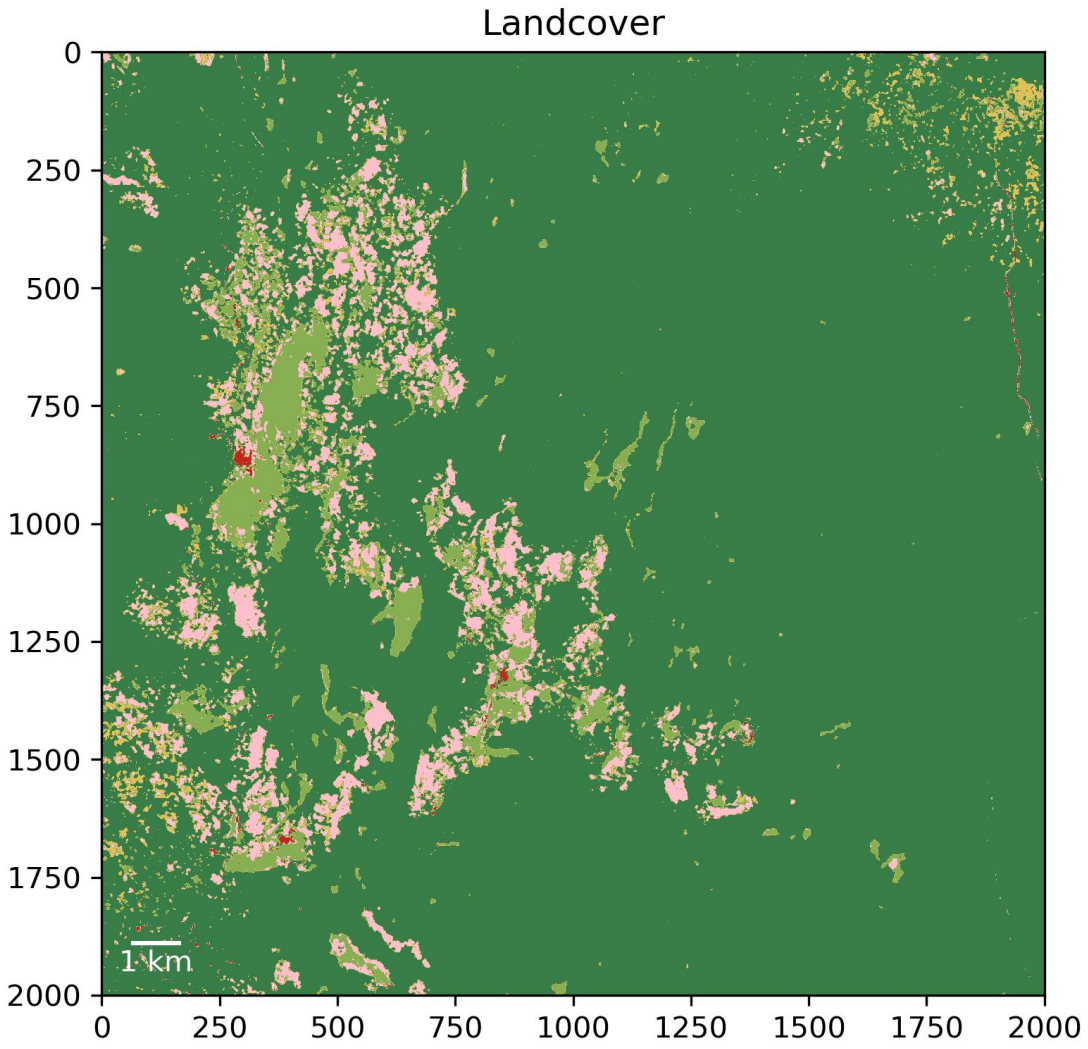
Bale Mountains National Park, Ethiopia



John Toth

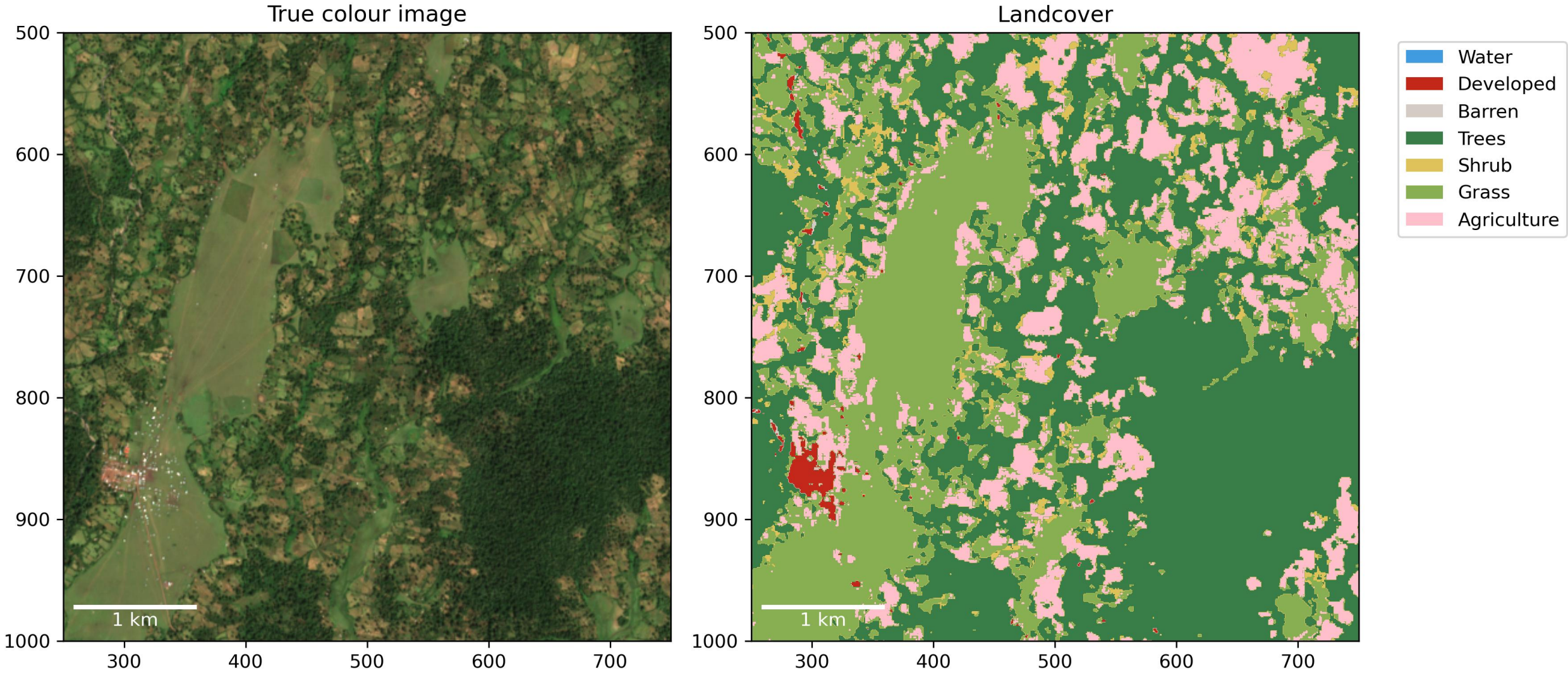
Bale Mountains National Park, Ethiopia, 20x20 km





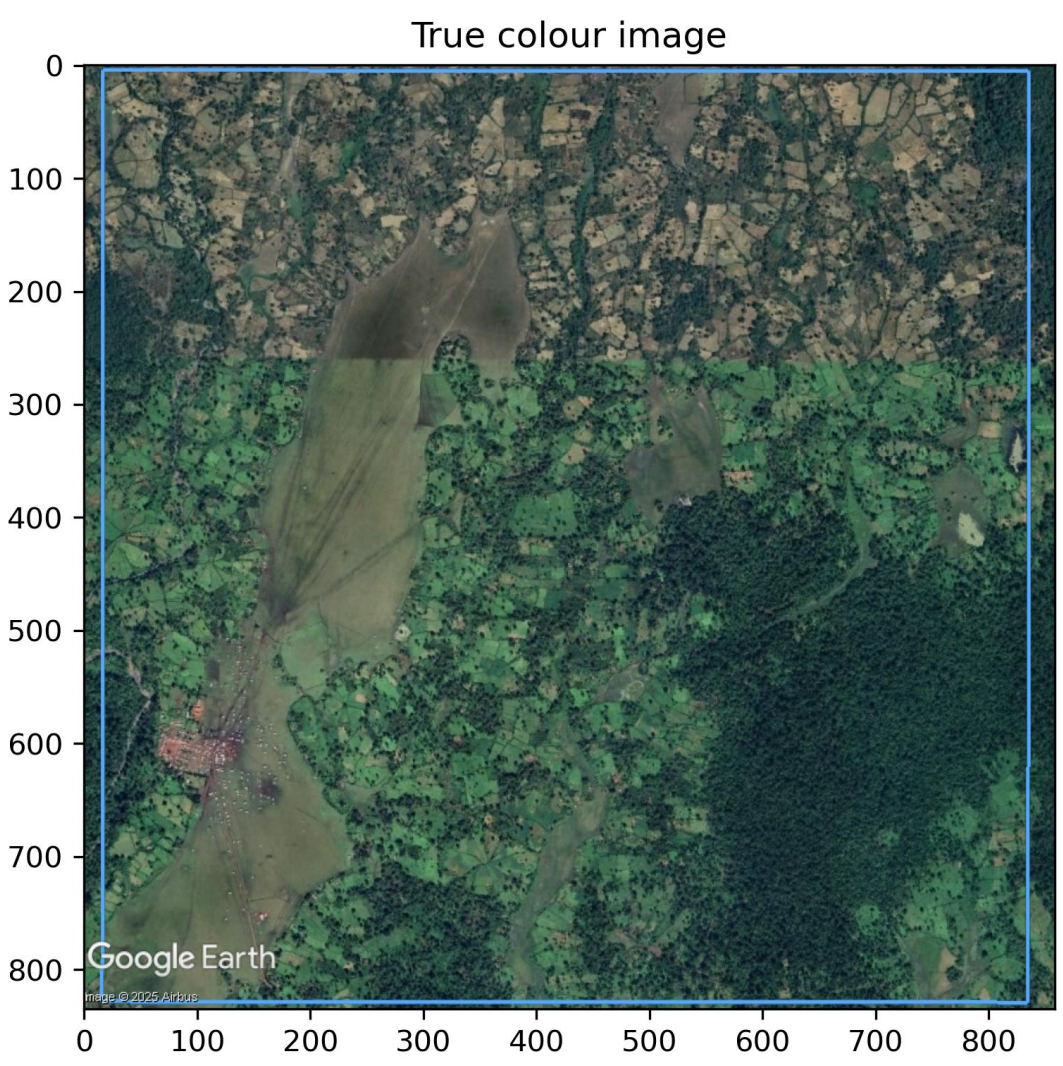
John Toth

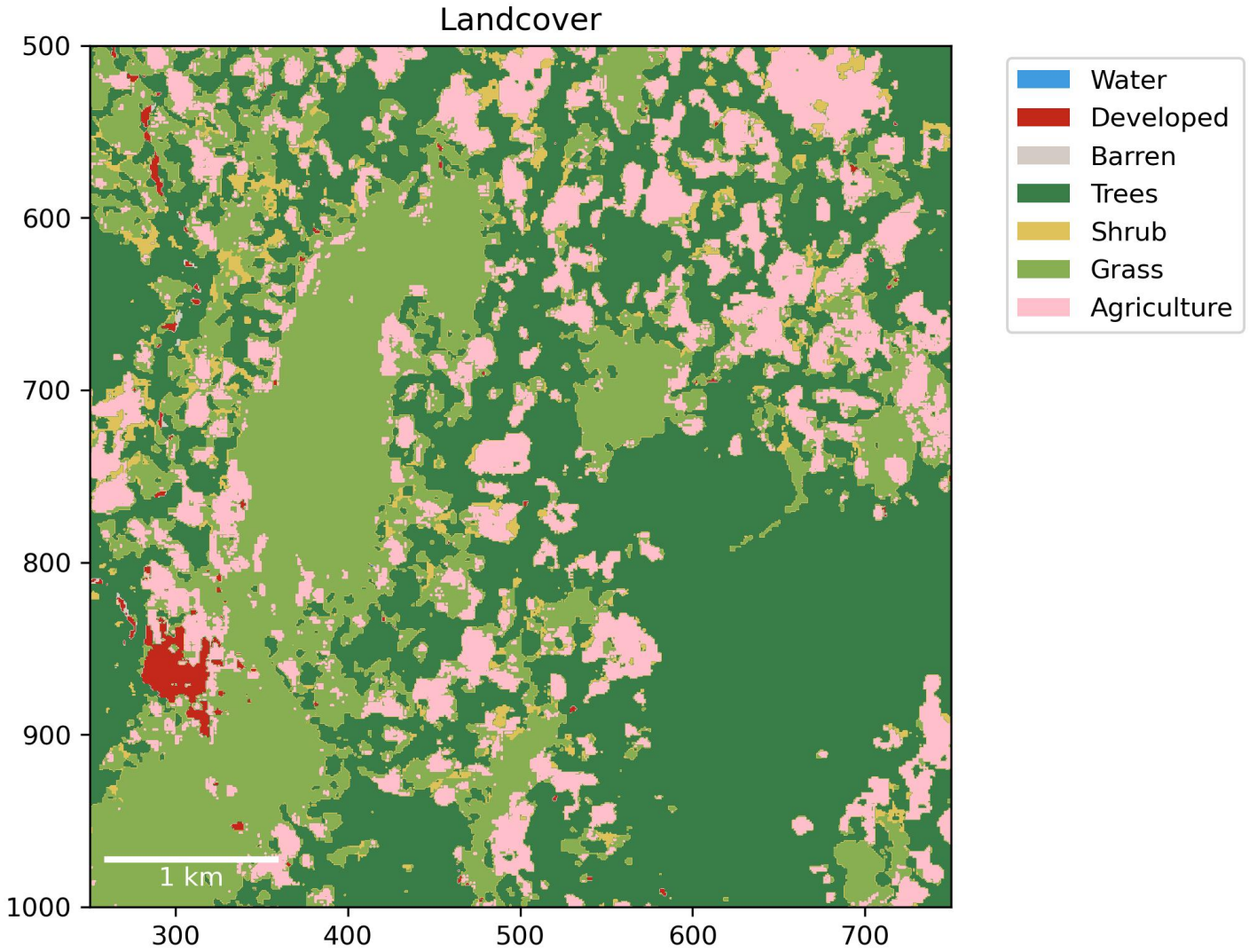
Bale Mountains National Park, Ethiopia, 5x5 km



John Toth

Bale Mountains National Park, Ethiopia, 5x5 km





John Toth