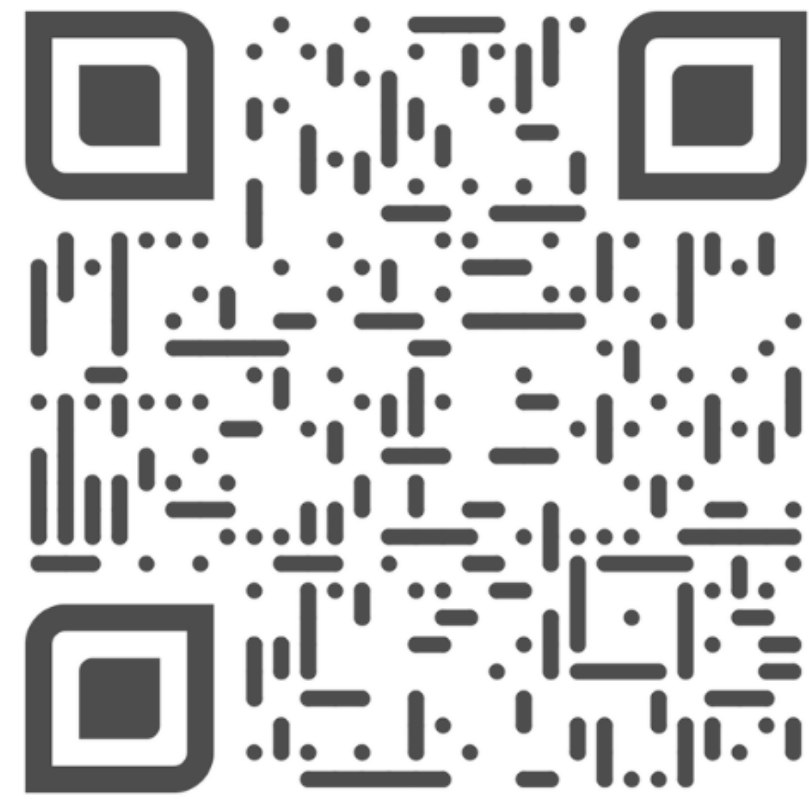
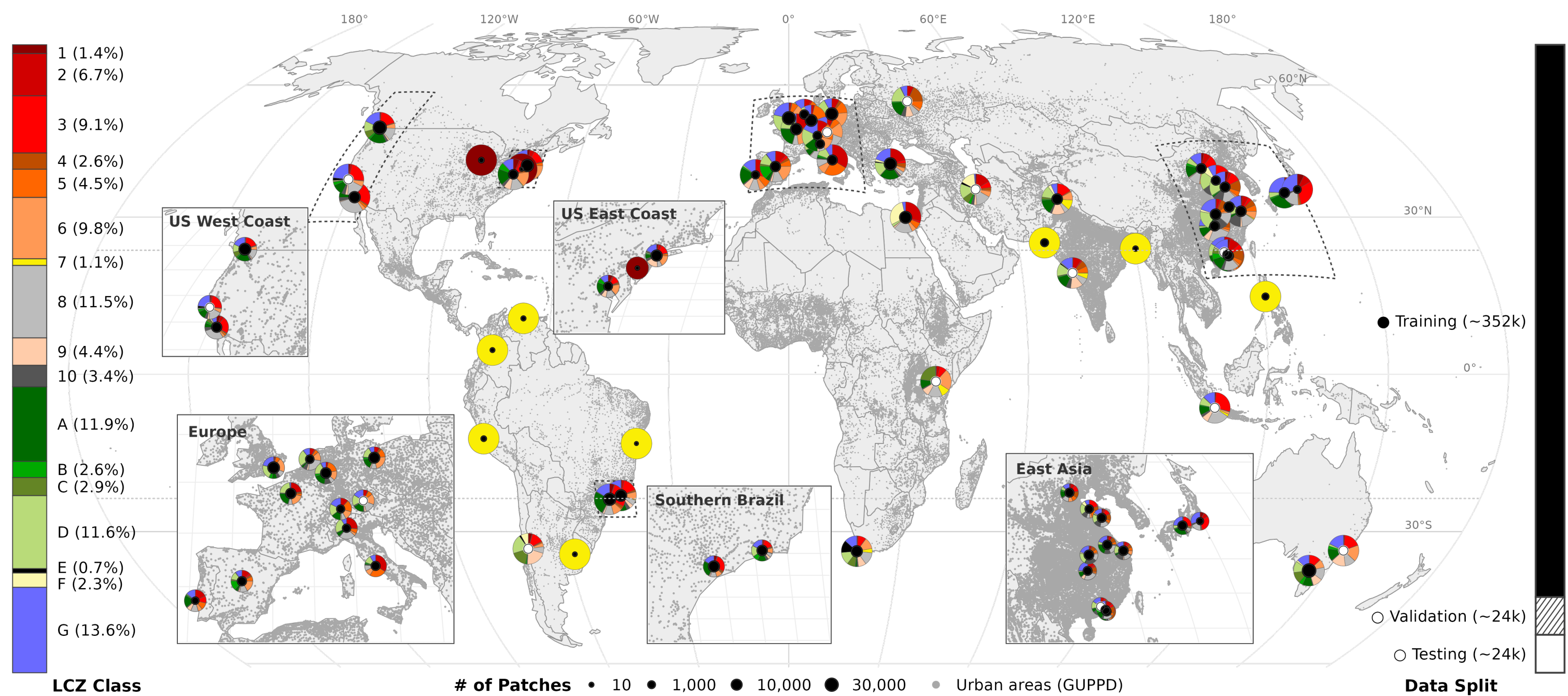


Mapping Local Climate Zones Globally with Geospatial Foundation Models

Andrés C. Zúñiga-González^{1,2}, Anil Madhavapeddy², and Ronita Bardhan¹¹ Department of Architecture² Department of Computer Science and Technology

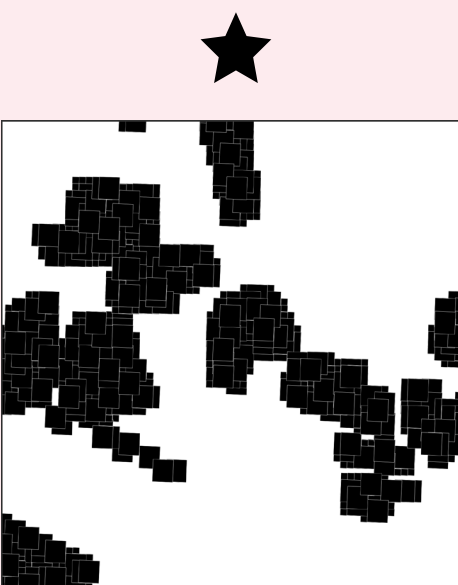
Local Climate Zones (LCZs) are a 17-class scheme describing neighbourhood-scale urban form and cover: ten built types (LCZ 1-10) defined by building height, density and impervious fraction, and seven natural types (LCZ A-G) defined by vegetation and surface cover (Stewart & Oke, 2012). Because the classes are physical rather than administrative, they are comparable across cities and cultures, which makes them a standard input for urban heat island studies and urban climate modelling; misclassification propagates directly into temperature estimates. Global coverage exists at 100 m (Demuzere et al., 2022) but relies on volunteered training data of uneven quality and is costly to update. **Geospatial Foundation Models (GFM)** offer an alternative: they compress a full year of satellite observations into a compact 10 m per-pixel **embedding**, which can be used frozen so that downstream models are small and cheap to train. Here we benchmark two GFMs: **Tessera v2** (Feng et al., 2026) and **AlphaEarth** (Brown et al., 2025), across two task framings (**semantic segmentation** and **patch classification**), two spatial splits, and several neural network architectures and configurations, using the **So2Sat LCZ42** dataset (Zhu et al., 2020), which contains ~400k LCZ labels across the world with the year 2017 as reference.



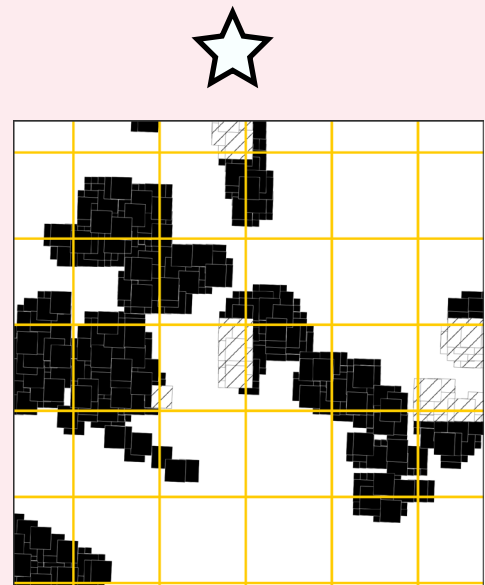
Distribution of the So2Sat LCZ42 dataset across LCZ classes, cities, and the culture-10 data split

Data Split

Cultural-10



Gridded



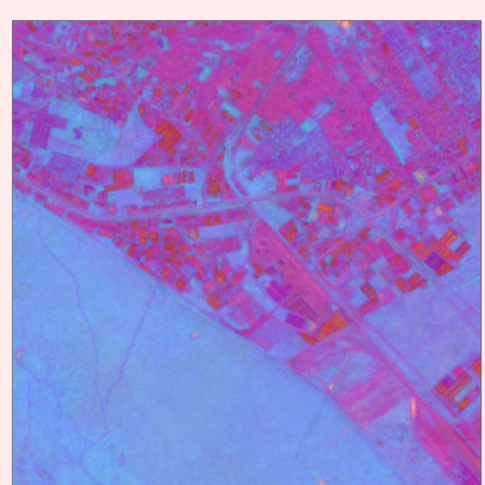
Cultural-10 (So2Sat LCZ42 original) splits by city: 42 for training, 10 for held out and cut west/east into validation and test.

Gridded lays a 128px checkerboard over every city and assigns tiles in 70/15/15, patches inherit their tile. Test stays fixed; training cities now also yield validation.

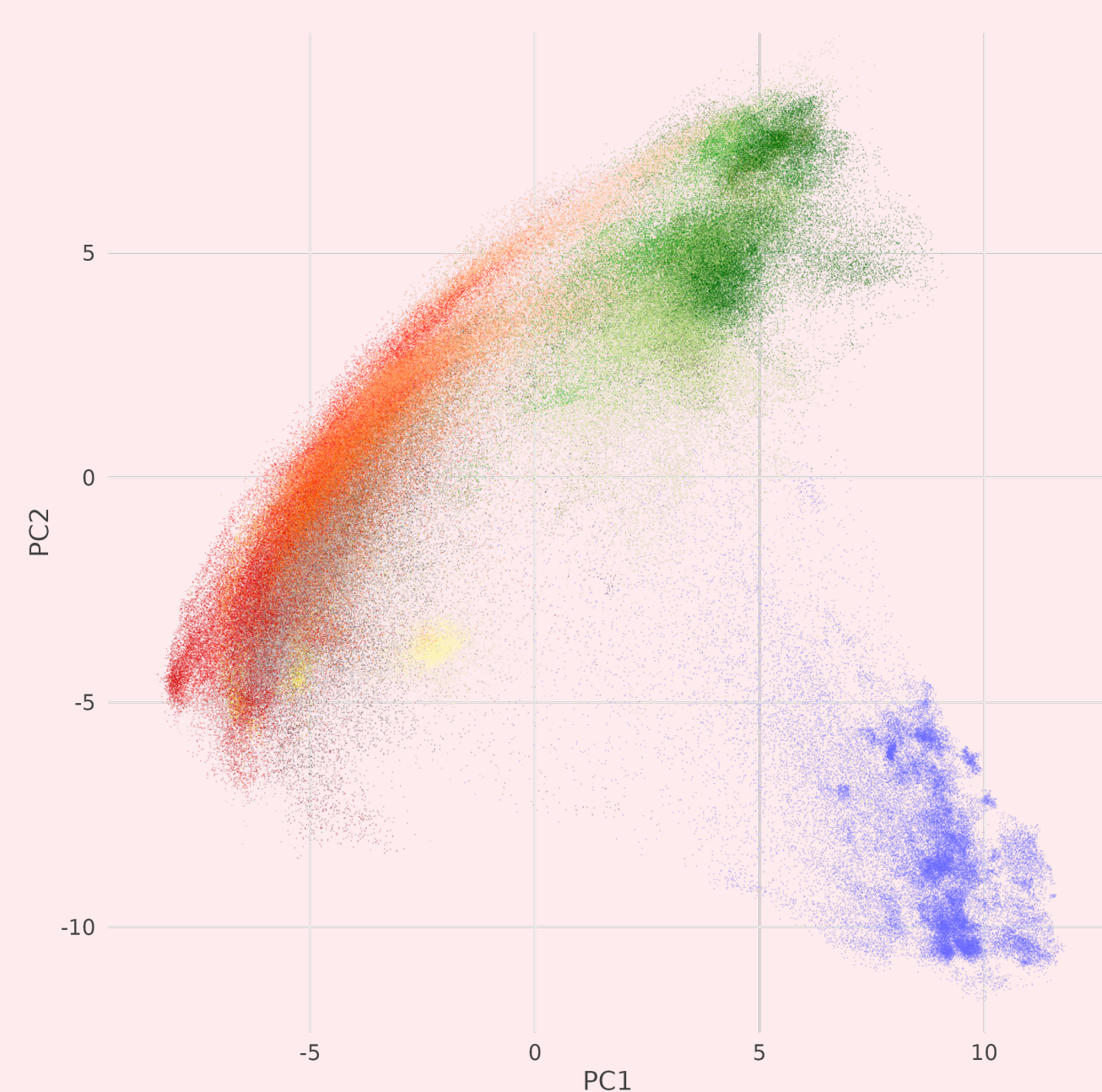
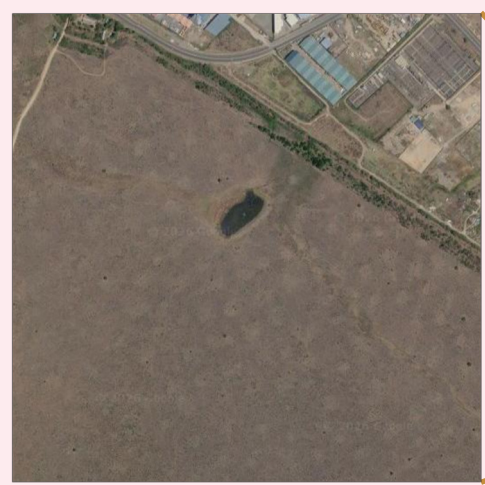
Embeddings

Tessera v2

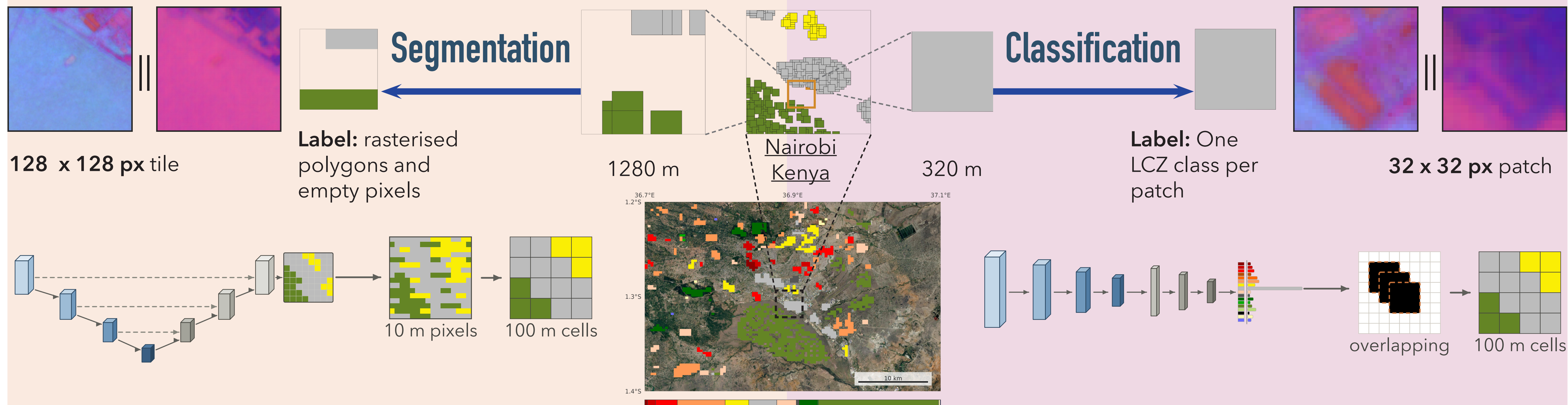
AlphaEarth



Reference



PCA representations using Global Average Pooling (GAP) of Tessera v2 embeddings of the So2Sat LCZ42 patches

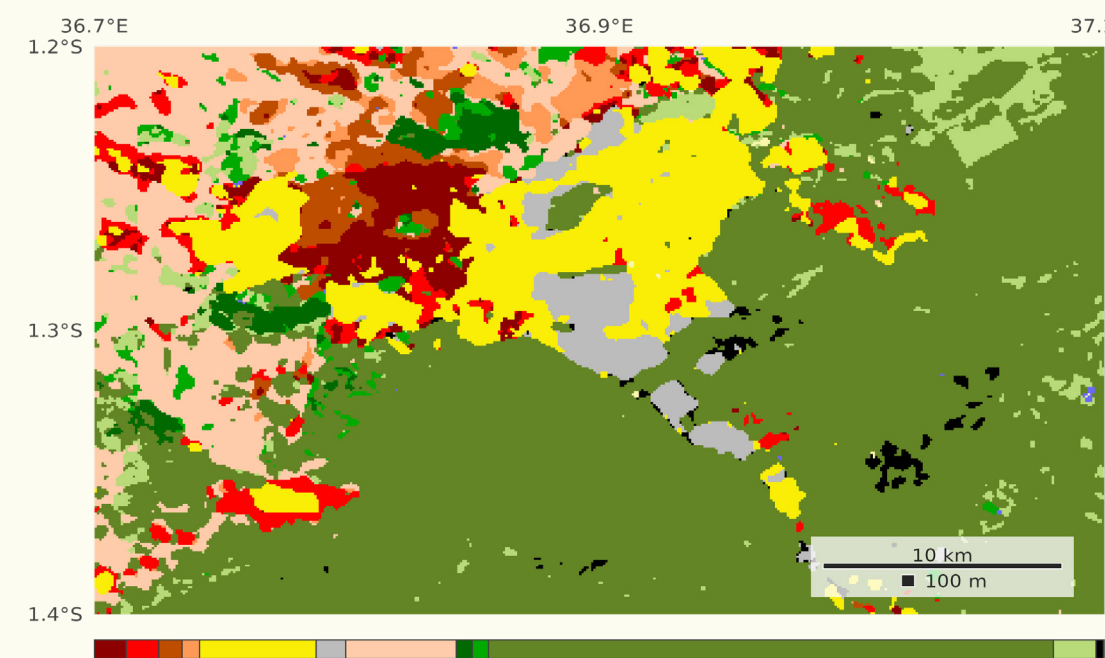


Predicted LCZ class																
	1	2	3	4	5	6	7	8	9	10	A	B	C	D	E	F
1	0.61	0.01	0.09	0.11	-	-	-	-	-	-	-	-	-	-	-	0.06
2	0.29	0.33	0.11	0.02	0.01	0.01	0.06	0.15	-	0.01	-	-	-	-	-	0.00
3	0.03	0.01	0.51	0.00	0.01	0.13	0.16	0.06	0.00	-	0.00	0.01	0.00	0.01	-	0.03
4	0.15	0.10	0.23	0.16	0.10	0.05	0.00	0.04	0.01	0.01	-	0.03	0.07	-	-	0.04
5	0.05	0.07	0.26	0.02	0.22	0.17	0.01	0.07	0.03	-	0.01	-	0.05	0.01	-	0.00
6	0.01	-	0.15	0.03	0.01	0.52	0.03	0.02	0.06	0.00	0.02	0.01	0.10	0.02	-	0.02
7	0.02	-	0.04	-	-	0.00	0.43	0.41	-	-	-	-	0.07	0.01	0.00	-
8	0.01	0.02	0.06	0.01	0.01	0.03	0.06	0.65	0.01	0.01	0.01	0.00	0.05	0.04	0.00	0.03
9	0.00	0.01	0.04	0.01	0.03	0.30	0.04	0.00	0.37	-	0.01	0.01	0.09	0.04	-	0.03
10	-	0.06	0.03	0.00	0.01	0.02	0.01	0.45	0.02	0.04	0.01	0.01	0.03	0.11	0.04	0.04
A	0.00	-	0.00	0.00	-	0.01	-	0.00	0.11	-	0.66	0.03	0.16	0.02	-	0.00
B	0.00	-	0.03	-	0.01	0.13	-	0.05	0.16	-	0.04	0.10	0.17	0.19	-	0.05
C	-	-	-	-	0.00	-	0.00	0.01	-	0.01	-	0.63	0.28	-	0.06	-
D	-	-	0.00	0.00	-	0.01	0.02	0.01	0.11	-	0.00	-	0.84	-	0.01	0.00
E	-	-	-	-	-	0.08	0.39	-	0.15	0.01	-	0.02	0.02	0.25	0.04	0.04
F	-	-	-	-	-	0.02	0.03	0.03	0.00	0.00	0.00	0.01	0.15	0.05	0.69	0.01
G	0.03	0.00	0.00	0.01	-	0.00	-	0.02	0.01	0.02	0.01	0.00	0.00	0.02	-	0.03

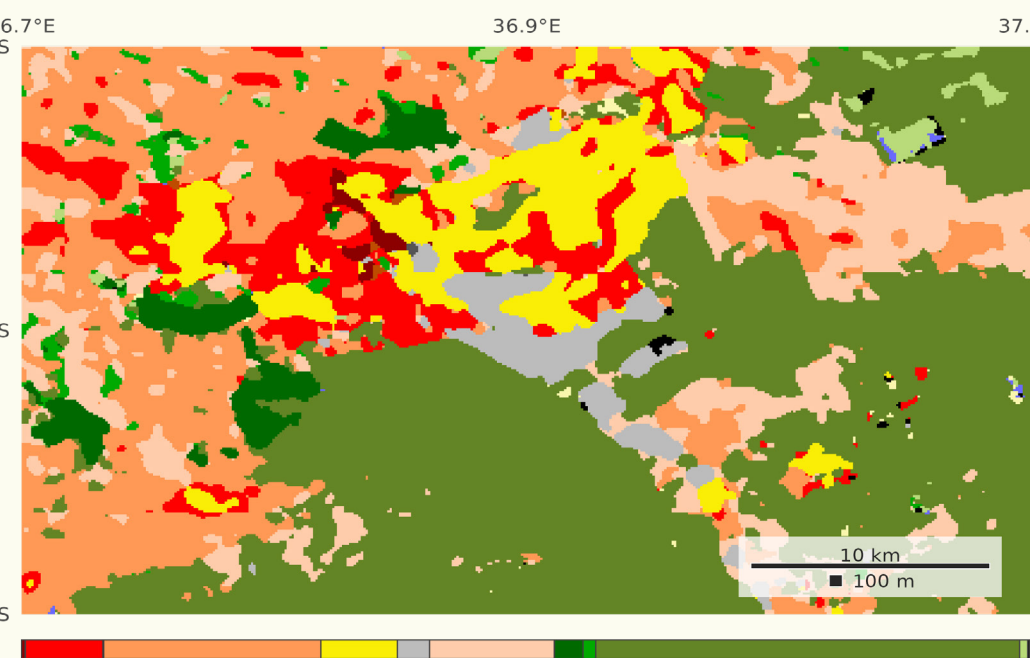
Model	# Params	OA	OAU	Macro F1	K
GAP	2.2K	0.581	0.558	0.449	0.542
ShallowCNN-16/32	23.7K	0.644	0.618	0.559	0.612
ShallowCNN-32/64	56.6K	0.646	0.625	0.572	0.616
MobileNetV3-S 0.5x	603.6K	0.670	0.643	0.560	0.640
MobileNetV3-S 1.0x	5.98M	0.644	0.633	0.552	0.644
ResNet18	11.2M	0.633	0.612	0.544	0.644
ResNet34	21.4M	0.624	0.588	0.549	0.591
GAP	1.1K	0.536	0.480	0.460	0.498
ShallowCNN-16/32	14.5K	0.558	0.485	0.475	0.523
ShallowCNN-32/64	38.2K	0.546	0.458	0.491	0.513
MobileNetV3-S 0.5x	594.4K	0.633	0.588	0.533	0.577
MobileNetV3-S 1.0x	1.5M	0.572	0.505	0.498	0.538
ResNet18	11.2M	0.577	0.497	0.390	0.529
ResNet34	21.3M	0.533	0.417	0.445	0.498
GAP	2.2K	0.698	0.652	0.594	0.671
ShallowCNN-16/32	23.7K	0.794	0.759	0.734	0.776
ShallowCNN-32/64	56.6K	0.816	0.783	0.758	0.799
MobileNetV3-S 0.5x	603.6K	0.844	0.808	0.788	0.830
MobileNetV3-S 1.0x	1.6M	0.850	0.809	0.800	0.838
ResNet18	11.2M	0.840	0.802	0.802	0.834
ResNet34	21.4M	0.804	0.804	0.797	0.811
GAP	1.1K	0.679	0.614	0.574	0.651
ShallowCNN-16/32	14.5K	0.736	0.678	0.657	0.713
ShallowCNN-32/64	38.2K	0.802	0.763	0.739	0.824
MobileNetV3-S 0.5x	594.4K	0.841	0.798	0.778	0.827
MobileNetV3-S 1.0x	1.5M	0.838	0.794	0.773	0.821
ResNet18	11.2M	0.832	0.785	0.782	0.824
ResNet34	21.3M	0.833	0.788	0.782	0.817

Predicted LCZ class																
	1	2	3	4	5	6	7	8	9	10	A	B	C	D	E	F
1	0.61	0.07	0.02	0.20	0.00	-	0.01	0.06	-	0.02	-	-	-	-	-	-
2	0.07	0.49	0.17	0.02	0.10	0.00	0.04	0.09	-	0.02	-	-	-	-	-	-
3	0.03	0.06	0.61	0.01	0.02	0.12	0.11	0.02	-	0.00	0.00	-	-	-	-	-
4	0.02	0.02	0.03	0.76	0.03	0.00	0.00	0.01	0.00	0.12	-	-	-	-	-	-
5	0.00	0.06	0.17	0.20	0.39	0.14	-	0.01	0.03	0.01	-	0.00	-	-	-	-
6	0.01	0.00	0.23	0.02	0.09	0.51	0.01	0.01	0.06	0.00	-	-	-	0.08	-	-
7	0.00	0.07	0.33	0.01	0.00	0.00	0.46	0.12	-	0.01	-	-	-	-	0.00	-
8	0.01	0.00	0.02	0.01	0.01	0.00	0.01	0.90	0.00	0.03	-	-	-	-	0.00	0.00
9	-	-	0.06	0.01	0.01	0.41	-	0.00	0.40	-	0.01	0.01	0.00	0.08	-	0.01
10	0.01	-	0.01	0.07	0.04	0.00	-	0.26	0.00	0.58	0.00	-	0.02	0.01	0.01	-
A	-	-	-	-	-	-	-	-	-	0.85	0.07	0.02	0.06	0.00	-	-
B	-	-	0.02	-	0.10	-	0.03	-	0.22	0.41	0.04	0.16	-	0.01	0.00	-
C	-	-	0.01	-	0.00	0.01	-	0.00	0.02	0.00	0.34	0.47	0.01	0.14	0.01	-
D	-	-	-	0.02	-	0.00	-	0.00	0.01	0.02	0.00	0.02	0.01	0.88	-	0.04
E	-	-	-	-	-	-	0.03	-	0.17	-	-	-	-	-	0.80	0.00
F	-	-	-	-	-	-	-	-	0.00	0.00	-	-	-	-	0.15	0.02
G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00	0.99

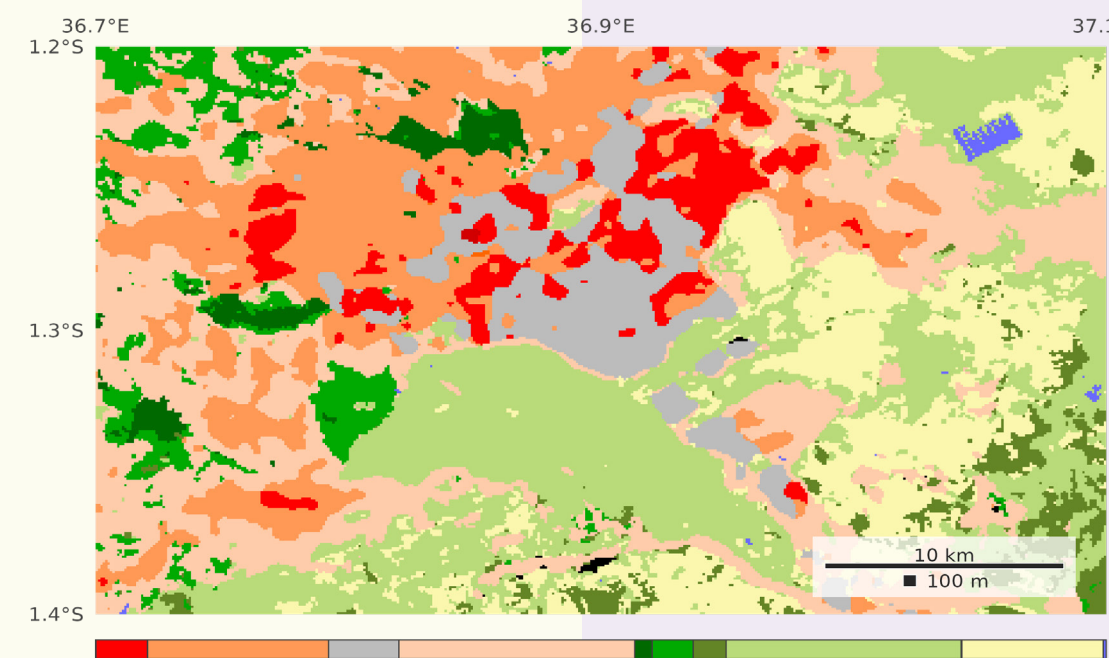
U-Net-3/32



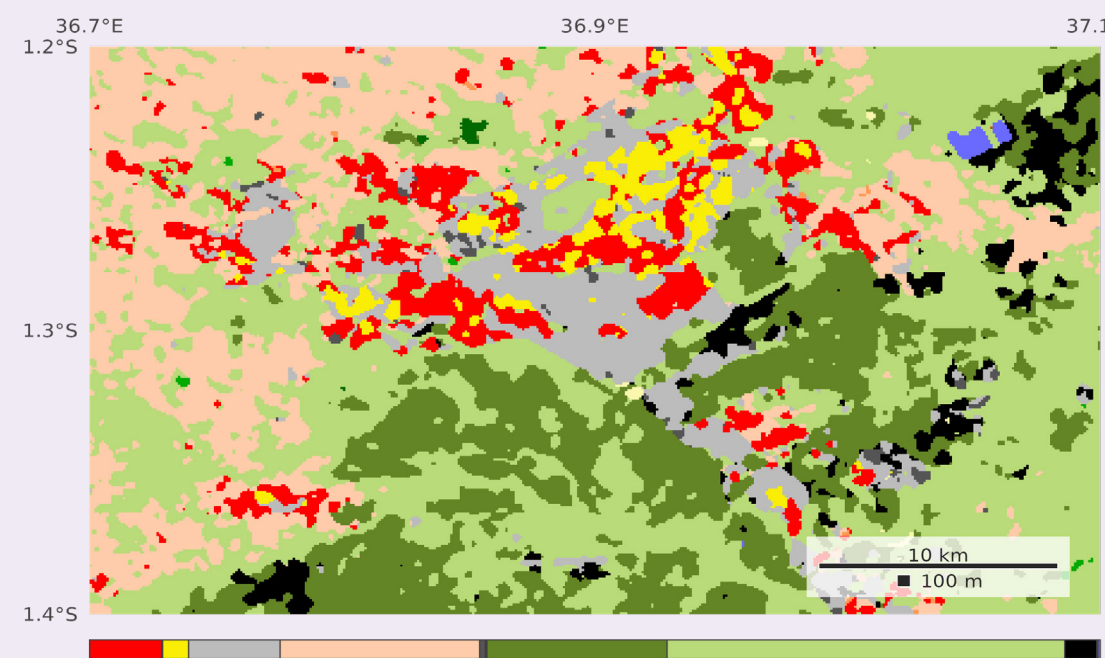
U-Net-3/32



Demuzere et al., 2022



MobileNetV3-S 1.0x



MobileNetV3-S 1.0x

