Supplementary Material
When Less is More: Evaluating Structural Pruning in Geospatial Foundation Models

Amina Said [*] Julia Dietlmeier [†] Margaret McCaul [†] Noel E. O'Connor [*][†]

A. Dataset Setup

Below are dataset-specific details on experimental set up. TerraTorch [3] is the main library used in setting up the datasets, data modules and training of the models. While Torch-Pruning [1, 2] library is used in implementing the global structural pruning on the three Geospatial Foundation Models (GeoFMs). The experimented GeoFMs are namely: Clay-v1, Prithvi-L and TerraMind-B. Below are information about the four datasets used, ranging from regression, segmentation and classification task.

The Above Ground Biomass [9] is a regression dataset, contains 6 bands and a patch size of 224×224. The bands used are blue, green, red, near-infrared (NIR) narrow, short-wave infrared 1 and 2 (SWIR1, SWIR2). A batch size of 4 is used in all models. AGB dataset has one class and root mean squared error (RMSE) is the loss function that is monitored in training. Tab. 1 shows the performance of all the models and techniques used in the regression task.

Burn Scars [10] is a binary segmentation dataset and part of PANGAEA datasets [8]. It has a patch size 512×512 and contains 6 bands, same bands as of AGB. Cross entropy (CE) provided in TerraTorch for binary segmentation as well, is used as the loss function. A batch size of 4 is used. Fig. 1 shows the prediction masks from each of the best performing pruning strategy. Tab. 2 shows the performance of all binary segmentation models, across different techniques.

The **m-Cashew-Plantation** [6] is one of the Geo-Bench [7] datasets and 6 out of the 13 bands are used for training the models. The bands used are the same as those used in AGB. A batch size of 8 is used on the 256×256 patches, with 7 different classes and CE loss function. Fig. 2 shows the prediction masks of the semantic segmentation from each of the best performing pruning strategy. Tab. 3 shows

performance of the models across different experimented techniques.

The **m-EuroSAT** dataset [4, 5] is also part of Geo-Bench [7]. It contains 13 bands but 6 bands are used for training, same bands as used in AGB. A batch size of 32 is used on the dataset. The dataset contains 10 different classes. The performance of classification is based on the accuracy of the correctly predicted classes.

B. Experimental Setup

For consistency, Linear Decoder with upsampling size of 16 is used across all GeoFMs for full fine-tuning (FT) and Parameter Efficient Fine-Tuning (PEFT). Full fine-tuning models are trained for 150 epochs and 100 epochs for PEFT and after-prune fine-tuning. Early stopping halts training if there is no improvement on the validation loss for 30 epochs. Learning rate (LR) scheduler is used, with a patience of 4 epochs, after which if there there no improvement, it reduces the LR by a factor of 0.5.

Global structural pruning is applied across all the three experimented pruning techniques. Pruning is compared across the different techniques and pruning ratios. The full FT model is loaded and pre-processing is carried out to ensure model handles the structural change after pruning. This entails patching the attention layers and wrapping the model to return a valid tensor. The ignored layers are then defined to ensure sensitive layers such as the patch embeddings that encode input data and final layers used in task-specific final output, are not pruned to preserve sanity of the model. Thereafter, pruning is defined based on importance and pruning ratio. Isomorphic [1, 2] is also used to ensure consistent pruning rate across layers and the model is pruned. After pruning, the model is updated to match changes after pruning and update the model's weight. The pruned model is then fine-tuned to improve on its performance. The performance is evaluated by testing on the test datasets, while monitoring energy consumed and carbon emitted.

[*]Research Ireland Centre for Research Training in Machine Learning (ML-Labs), Dublin City University.
Correspondence to: amina.said3@mail.dcu.ie.
[†]Insight Research Ireland Centre for Data Analytics, Dublin City University.
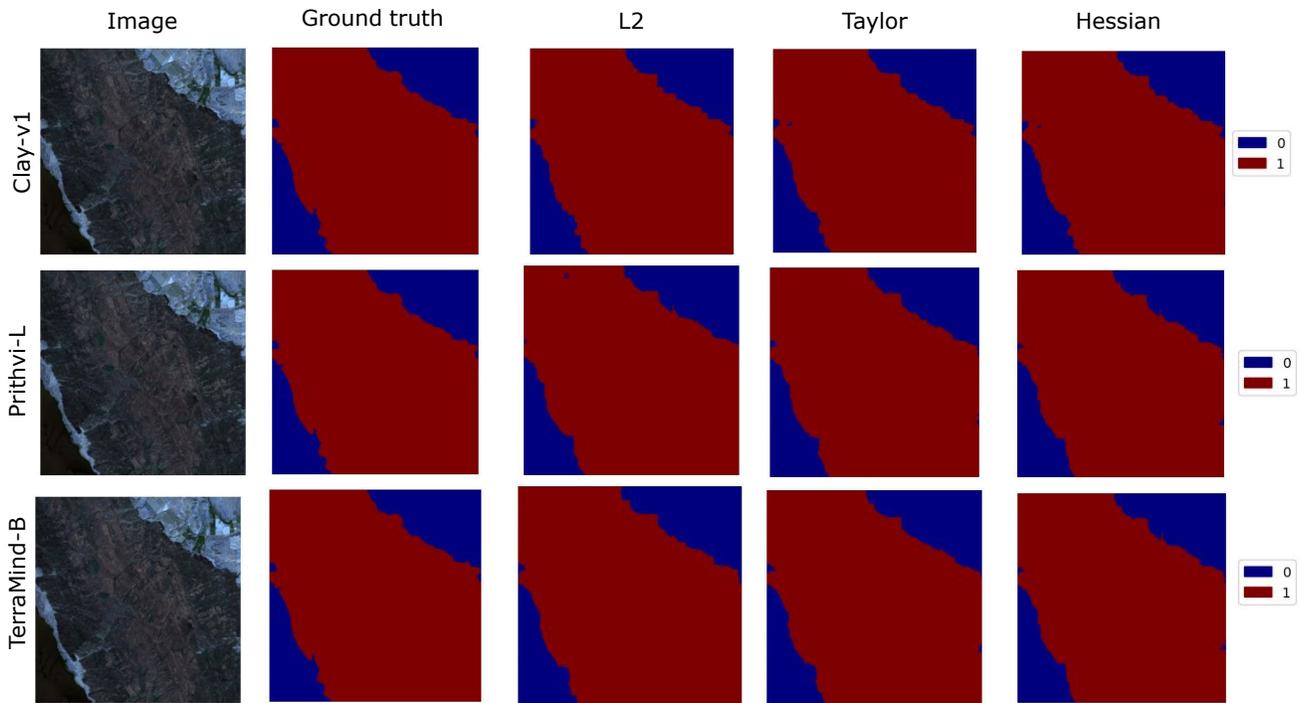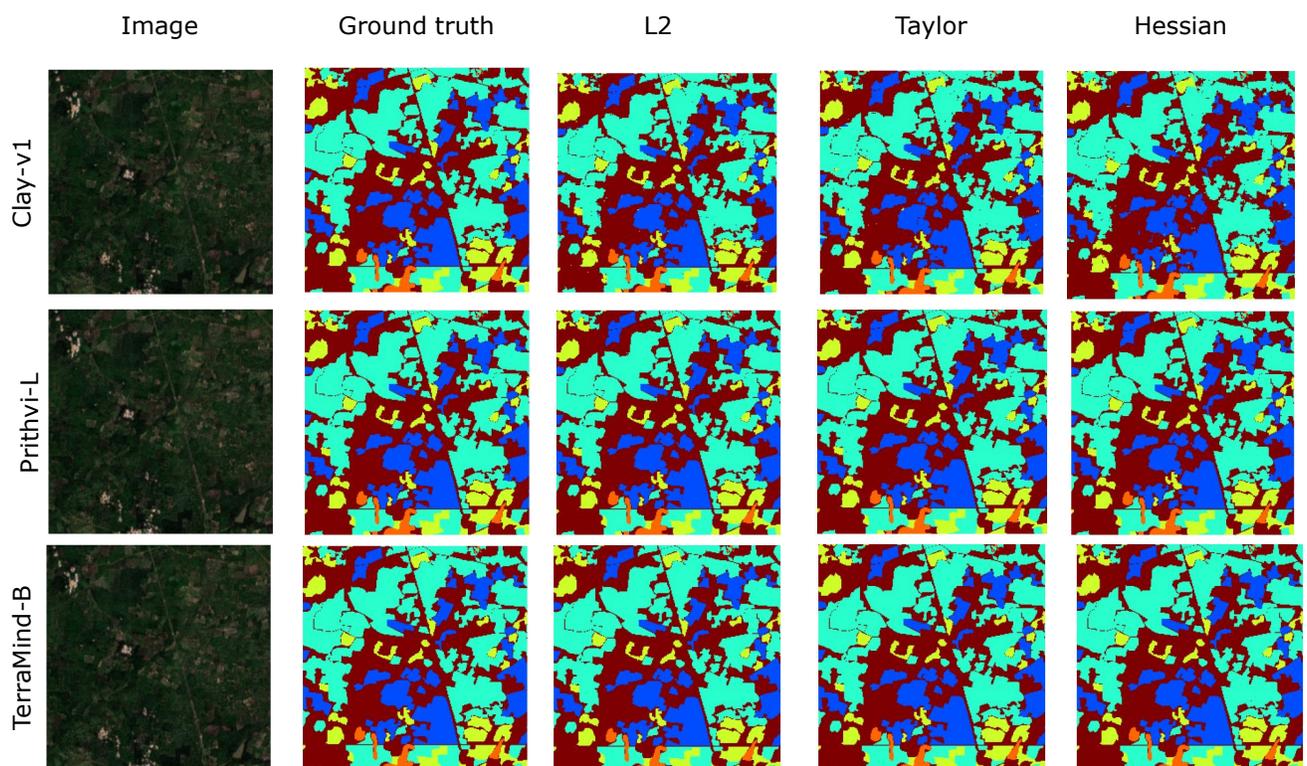
Figure 1. Binary segmentation

Figure 2. Semantic segmentation

Table 1. Regression task performance values showing RMSE (Mg/Ha), train time (h), inference (s), number of paramaters (M), inference energy consumption (kWh), inference carbon emission (kgCO$_2$eq), model size (MB) and compression rate.

Model	Technique	RMSE	Train	Inference	Params.	Energy	Carbon	Model size	Compres.
U-NET	Full FT	31.69	1.6	6.7	39	3.1E-04	8.9E-05	157	-
Terramind-S	Full FT	32.60	2.0	4.4	23	2.5E-04	7.2E-05	90	-
Clay-v1	Full FT	31.86	3.9	7.2	95	4.6E-04	1.3E-04	379	-
	LoRA	31.99	4.6	7.4	4	5.1E-04	1.5E-04	385	-
	L2 0.3	31.87	1.4	7.1	77.8	4.2E-04	1.2E-04	311	1.2
	L2 0.5	31.96	1.4	9.4	66.4	6.5E-04	1.9E-04	265	1.4
	L2 0.7	32.01	1.5	8.6	55.1	6.1E-04	1.8E-04	220	1.7
	L2 0.9	32.58	4.0	8.9	43.7	6.0E-04	1.8E-04	174	2.2
	Taylor 0.3	31.96	1.4	7.2	77.8	4.1E-04	1.2E-04	311	1.2
	Taylor 0.5	31.79	1.5	9.0	66.4	6.1E-04	1.8E-04	265	1.4
	Taylor 0.7	32.05	2.0	9.3	55.1	6.1E-04	1.8E-04	220	1.7
	Taylor 0.9	32.15	2.0	5.9	43.8	3.3E-04	9.6E-05	175	2.2
	Hessian 0.3	32.00	2.3	9.3	77.7	6.1E-04	1.8E-04	310	1.2
	Hessian 0.5	31.90	2.0	9.1	66.4	6.1E-04	1.8E-04	265	1.4
	Hessian 0.7	32.04	1.6	8.8	55.1	6.1E-04	1.8E-04	220	1.7
	Hessian 0.9	32.70	2.3	5.9	43.7	3.4E-04	9.8E-05	174	2.2
Prithvi-L	Full FT	32.95	4.5	5.4	308	2.7E-04	7.8E-05	1200	-
	LoRA	32.44	7.8	4.3	10	2.8E-04	8.2E-05	1300	-
	L2 0.3	33.05	2.6	4.7	211	2.5E-04	7.2E-05	844	1.4
	L2 0.5	32.88	2.6	4.3	152	2.2E-04	6.4E-05	609	2.0
	L2 0.7	32.75	1.9	3.9	92	1.8E-04	5.4E-05	368	3.3
	L2 0.9	32.57	2.7	4.8	106	3.2E-04	9.3E-05	425	2.8
	Taylor 0.3	32.74	2.1	4.7	212	2.5E-04	7.1E-05	848	1.4
	Taylor 0.5	32.29	2.1	4.8	153	2.9E-05	8.4E-06	613	2.0
	Taylor 0.7	32.97	1.9	3.9	92	2.5E-04	7.4E-05	368	3.3
	Taylor 0.9	32.64	2.0	4.0	98	1.9E-04	5.5E-05	392	3.1
	Hessian 0.3	32.58	3.4	4.6	213	2.4E-04	7.1E-05	852	1.4
	Hessian 0.5	32.83	3.2	5.2	151	3.3E-04	9.6E-05	607	2.0
	Hessian 0.7	32.54	2.1	4.6	105	2.9E-04	8.4E-05	419	2.9
	Hessian 0.9	32.86	3.5	4.9	109	3.1E-04	9.0E-05	438	2.7
Terramind-B	Full FT	32.73	1.8	3.8	89	2.5E-04	5.4E-05	355	-
	LoRA	33.20	3.8	3.7	4	1.7E-04	5.0E-05	362	-
	L2 0.3	32.76	2.9	3.7	72	1.6E-04	4.5E-05	287	1.2
	L2 0.5	32.68	2.9	3.6	61	2.1E-04	6.1E-05	242	1.5
	L2 0.7	33.01	3.0	3.5	49	1.5E-04	4.3E-05	196	1.8
	L2 0.9	33.01	2.8	3.6	38	1.6E-04	4.5E-05	151	2.4
	Taylor 0.3	32.81	2.9	3.9	62	2.3E-04	6.8E-05	248	1.4
	Taylor 0.5	32.85	2.9	3.3	45	1.4E-04	4.1E-05	181	2.0
	Taylor 0.7	32.70	3.0	3.7	28	2.1E-04	6.0E-05	109	3.3
	Taylor 0.9	32.69	2.5	3.2	30	1.4E-04	4.0E-05	118	3.0
	Hessian 0.3	32.84	2.7	3.5	62	1.5E-04	4.4E-05	247	1.4
	Hessian 0.5	32.67	2.7	3.9	45	2.1E-04	6.2E-05	181	2.0
	Hessian 0.7	32.74	2.6	3.2	32	1.4E-04	4.0E-05	129	2.8
	Hessian 0.9	32.90	2.5	3.2	40	1.4E-04	4.1E-05	160	2.2

Table 2. Binary segmentation performance values showing mIoU (%), train time (m), inference (s), number of paramaters (M), inference energy consumption (kWh), inference carbon emission (kgCO$_2$eq), model size (MB) and compression rate.

Model	Technique	mIoU	Train	Inference	Params.	Energy	Carbon	Model size	Compres.
UNET	Full FT	88.59	34.3	3.1	33	1.7E-04	5.0E-05	130	-
TerraMind-S	Full FT	88.54	24.8	2.8	23	1.5E-04	4.4E-05	90	-
Clay-v1	Full FT	89.93	136.3	6.8	96	9.4E-04	2.7E-04	380	-
	LoRA	89.53	137.1	5.2	5	3.7E-04	1.1E-04	386	-
	L2 0.3	92.08	21.6	4.8	78	3.1E-04	9.1E-05	312	1.2
	L2 0.5	91.09	13.5	4.7	67	3.1E-04	9.1E-05	267	1.4
	L2 0.7	91.35	18.6	4.5	55	2.9E-04	8.4E-05	221	1.7
	L2 0.9	90.55	22.4	4.4	44	2.8E-04	8.2E-05	176	2.2
	Taylor 0.3	91.68	18.2	4.9	78	3.3E-04	9.5E-05	312	1.2
	Taylor 0.5	91.73	16.1	4.5	67	3.1E-04	9.0E-05	267	1.4
	Taylor 0.7	91.71	19.7	5.0	55	3.1E-04	9.0E-05	221	1.7
	Taylor 0.9	91.26	19.6	5.6	44	3.7E-04	1.1E-04	176	2.2
	Hessian 0.3	91.60	18.7	4.7	78	3.1E-04	9.1E-05	312	1.2
	Hessian 0.5	92.31	22.5	4.7	67	3.1E-04	8.9E-05	266	1.4
	Hessian 0.7	91.36	21.3	4.7	55	3.0E-04	8.6E+00	221	1.7
	Hessian 0.9	91.90	30.1	4.7	44	2.9E-04	8.3E-05	176	2.2
Prithvi-L	Full FT	91.35	102.9	5.5	309	3.0E-04	8.8E-05	1200	-
	LoRA	91.45	161.1	4.8	11	2.8E-04	8.3E-05	1300	-
	L2 0.3	91.61	14.9	5.2	211	2.5E-04	7.1E-05	847	1.4
	L2 0.5	91.42	13.3	4.4	152	2.0E-04	5.9E-05	608	2.0
	L2 0.7	90.43	15.5	3.8	92	1.8E-04	5.1E-05	369	3.3
	L2 0.9	90.28	15.7	4.0	106	1.9E-04	5.5E-05	426	2.8
	Taylor 0.3	91.92	14.1	4.9	211	2.3E-04	6.7E-05	847	1.4
	Taylor 0.5	91.18	13.2	4.4	152	2.0E-04	6.0E-05	608	2.0
	Taylor 0.7	91.46	13.4	3.8	92	1.8E-04	5.2E-05	369	3.3
	Taylor 0.9	91.47	12.5	4.0	109	1.9E-04	5.5E-05	439	2.7
	Hessian 0.3	91.45	13.5	5.6	213	2.6E-04	7.5E-05	853	1.4
	Hessian 0.5	90.51	13.5	4.5	153	2.0E-04	5.9E-05	612	2.0
	Hessian 0.7	91.11	16.8	4.0	96	1.9E-04	5.5E-05	384	3.1
	Hessian 0.9	89.76	15.0	4.0	110	1.9E-04	5.5E-05	439	2.7
Terramind-B	Full FT	90.51	29.9	4.9	89	4.2E-04	1.2E-04	356	-
	LoRA	90.38	42.3	3.6	5	2.2E-04	6.4E-05	363	-
	L2 0.3	91.70	10.0	3.5	72	1.9E-04	5.6E-05	288	1.2
	L2 0.5	90.73	10.1	3.3	61	1.8E-04	5.2E-05	242	1.5
	L2 0.7	90.46	12.9	3.3	50	1.8E-04	5.2E-05	197	1.8
	L2 0.9	89.41	13.4	3.2	38	1.8E-04	5.2E-05	152	2.3
	Taylor 0.3	91.20	10.8	3.4	62	1.8E-04	5.3E-05	248	1.4
	Taylor 0.5	89.85	10.0	3.3	45	1.8E-04	5.2E-05	178	2.0
	Taylor 0.7	89.64	13.1	3.2	28	1.8E-04	5.1E-05	111	3.2
	Taylor 0.9	91.07	10.7	3.3	34	1.8E-04	5.1E-05	137	2.6
	Hessian 0.3	89.77	11.0	3.5	62	1.2E-05	5.6E-05	247	1.4
	Hessian 0.5	90.06	10.9	3.3	45	1.2E-05	5.0E-05	180	2.0
	Hessian 0.7	89.81	14.2	3.2	35	1.2E-05	5.0E-05	141	2.5
	Hessian 0.9	90.08	11.9	3.3	38	1.2E-05	5.2E-05	151	2.4

Table 3. Semantic segmentation performance values showing mIoU (%), train time (m), inference (s), number of paramaters (M), inference energy consumption (kWh), inference carbon emission (kgCO$_2$eq), model size (MB) and compression rate.

Model	Technique	mIoU	Train	Inference	Params.	Energy	Carbon	Model size	Compres.
UNET	Full FT	83.54	34.3	0.9	33	4.8E-05	1.4E-05	130	-
TerraMind-S	Full FT	79.58	25.2	0.8	23	4.0E-05	1.2E-05	93	-
Clay-v1	Full FT	77.09	48.6	1.1	92	1.6E-04	4.6E-05	370	-
	LoRA	62.00	25.3	0.9	2	6.4E-05	1.9E-05	376	-
	L2 0.3	76.04	40.6	1.9	75	1.6E-04	4.6E-05	301	1.2
	L2 0.5	75.21	38.9	1.9	64	1.2E-04	3.6E-05	256	1.4
	L2 0.7	72.47	46.3	1.8	53	1.4E-04	4.0E-05	210	1.8
	L2 0.9	62.08	45.3	1.2	41	5.3E-05	1.5E-05	165	2.2
	Taylor 0.3	75.66	36.5	2.0	75	1.6E-04	4.5E-05	301	1.2
	Taylor 0.5	74.44	33.8	1.6	64	1.4E-04	4.0E-05	256	1.4
	Taylor 0.7	73.69	40.7	1.8	53	1.5E-04	4.3E-05	210	1.8
	Taylor 0.9	67.08	48.4	1.9	41	1.4E-04	4.2E-05	165	2.2
	Hessian 0.3	73.35	37.4	2.3	75	1.9E-04	5.4E-05	301	1.2
	Hessian 0.5	72.91	41.4	1.9	64	1.5E-04	4.4E-05	256	1.4
	Hessian 0.7	70.16	41.8	1.2	53	5.8E-05	1.7E-05	210	1.8
	Hessian 0.9	57.81	39.7	0.8	41	4.8E-05	1.4E-05	165	2.2
Prithvi-L	Full FT	80.69	47.6	3.0	307	8.9E-05	2.6E-05	1200	-
	LoRA	79.29	78.9	1.3	7	8.1E-05	2.3E-05	1300	-
	L2 0.3	81.13	35.9	2.4	210	9.3E-05	2.7E-05	840	1.4
	L2 0.5	80.53	37.5	2.0	149	7.6E-05	3.0E-07	599	2.0
	L2 0.7	79.98	31.2	1.6	91	7.0E-05	2.8E-07	363	3.3
	L2 0.9	76.25	35.1	1.8	105	7.6E-05	3.0E-07	420	2.9
	Taylor 0.3	80.64	33.9	2.7	210	1.1E-04	3.1E-05	841	1.4
	Taylor 0.5	80.95	33.2	2.0	151	7.8E-05	2.3E-05	604	2.0
	Taylor 0.7	80.00	36.3	1.8	90	8.2E-05	2.4E-05	359	3.3
	Taylor 0.9	76.45	42.5	2.4	108	1.8E-04	5.3E-05	433	2.8
	Hessian 0.3	80.97	44.4	2.7	211	1.0E-04	3.0E-05	945	1.3
	Hessian 0.5	80.24	35.3	2.5	151	1.0E-04	2.9E-05	606	2.0
	Hessian 0.7	80.08	37.1	1.8	94	1.6E-04	4.8E-05	365	3.3
	Hessian 0.9	76.45	33.8	1.7	102	1.3E-04	3.6E-05	411	2.9
Terramind-B	Full FT	81.01	27.8	1.4	89	6.0E-05	1.7E-05	355	-
	LoRA	76.48	44.7	1.0	3	5.8E-05	1.7E-05	362	-
	L2 0.3	80.86	25.2	1.5	72	1.0E-04	3.0E-05	287	1.2
	L2 0.5	80.95	27.3	1.2	61	8.7E-05	2.5E-05	241	1.5
	L2 0.7	77.66	20.3	1.3	49	7.7E-05	2.2E-05	196	1.8
	L2 0.9	77.17	35.8	1.0	38	4.8E-05	1.4E-05	151	2.4
	Taylor 0.3	80.93	24.2	1.1	61	5.0E-05	1.5E-05	244	1.5
	Taylor 0.5	80.81	24.6	1.3	45	1.0E-04	2.9E-05	179	2.0
	Taylor 0.7	80.53	29.9	1.2	28	9.0E-05	2.6E-05	111	3.2
	Taylor 0.9	78.22	37.6	0.9	32	4.3E-05	1.2E-05	126	2.8
	Hessian 0.3	80.84	16.3	1.1	62	4.8E-05	1.4E-05	245	1.4
	Hessian 0.5	80.64	17.7	0.9	45	4.6E-05	1.3E-05	178	2.0
	Hessian 0.7	80.58	20.8	0.9	30	4.3E-05	1.3E-05	119	3.0
	Hessian 0.9	77.71	23.2	0.9	30	4.3E-05	1.3E-05	120	3.0

Table 4. Classification task performance values showing accuracy (%), train time (m), inference (s), number of paramaters (M), inference energy consumption (kWh), inference carbon emission (kgCO$_2$eq), model size (MB) and compression rate.

Model	Technique	Accuracy	Train	Inference	Params.	Energy	Carbon	Model size	Compres.
ResNET50	Full FT	90.90	14.1	2.0	24	4.6E-04	1.3E-04	94	-
TerraMind-S	Full FT	93.60	11.1	1.7	23	8.8E-05	2.6E-05	91	-
Clay-v1	Full FT	95.50	31.9	6.3	92	4.4E-04	1.3E-04	368	-
	LoRA	93.00	41.4	6.0	2	4.9E-04	1.4E-04	374	-
	L2 0.3	96.20	30.0	7.0	75	8.3E-04	2.4E-04	300	1.2
	L2 0.5	95.80	24.4	7.1	64	7.7E-04	2.2E-04	254	1.4
	L2 0.7	95.10	33.4	8.3	52	7.9E-04	2.3E-04	209	1.8
	L2 0.9	95.20	27.3	3.3	41	2.6E-04	7.7E-05	164	2.2
	Taylor 0.3	96.20	31.0	6.9	75	8.0E-04	2.3E-04	300	1.2
	Taylor 0.5	96.10	31.8	6.4	64	7.2E-04	2.1E-04	254	1.4
	Taylor 0.7	96.00	26.8	6.3	52	6.9E-04	2.0E-04	209	1.8
	Taylor 0.9	95.30	23.2	3.1	41	2.4E-04	7.1E-05	164	2.2
	Hessian 0.3	96.30	24.3	4.0	75	3.6E-04	1.0E-04	300	1.2
	Hessian 0.5	95.50	20.6	5.5	64	6.9E-04	2.0E-04	254	1.4
	Hessian 0.7	94.40	20.7	6.4	52	6.9E-04	2.0E-04	209	1.8
	Hessian 0.9	94.2	23.2	3.4	41	2.8E-04	8.1E-05	164	2.2
Prithvi-L	Full FT	91.30	14.9	6.4	306	3.8E-04	1.1E-04	1200	-
	LoRA	77.60	138.1	6.3	6	2.7E-04	7.8E-05	1200	-
	L2 0.3	91.60	19.1	6.9	208	4.1E-04	1.2E-04	833	1.4
	L2 0.5	92.60	19.1	5.9	148	3.1E-04	8.9E-05	592	2.0
	L2 0.7	91.90	26.6	5.2	89	2.7E-04	7.7E-05	355	3.4
	L2 0.9	92.20	24.8	6.5	103	4.7E-04	1.4E-04	413	2.9
	Taylor 0.3	92.90	27.3	7.5	210	5.6E-04	1.6E-04	842	1.4
	Taylor 0.5	91.50	29.0	6.9	149	5.5E-04	1.6E-04	598	2.0
	Taylor 0.7	91.50	28.5	6.3	89	4.2E-04	1.2E-04	356	3.4
	Taylor 0.9	92.00	29.5	6.1	88	4.2E-04	1.2E-04	350	3.4
	Hessian 0.3	92.30	28.3	7.6	209	5.6E-04	1.6E-04	837	1.4
	Hessian 0.5	91.40	26.4	8.3	148	5.7E-04	1.6E-04	592	2.0
	Hessian 0.7	91.50	24.3	6.6	93	4.5E-04	1.3E-04	372	3.2
	Hessian 0.9	92.50	22.7	5.3	101	2.7E-04	7.9E-05	406	3.0
Terramind-B	Full FT	93.60	10.8	2.3	88	1.5E-04	4.4E-05	351	-
	LoRA	77.10	15.8	1.4	2	6.8E-05	2.0E-05	358	-
	L2 0.3	94.80	13.6	2.7	71	2.5E-04	7.4E-05	282	1.2
	L2 0.5	92.60	13.1	2.2	59	2.1E-04	6.2E-05	237	1.5
	L2 0.7	93.10	15.4	2.1	48	1.3E-04	3.8E-05	191	1.8
	L2 0.9	92.70	15.2	1.9	37	1.1E-04	3.3E-05	145	2.4
	Taylor 0.3	94.30	9.3	2.1	61	1.3E-04	3.7E-05	243	1.4
	Taylor 0.5	93.50	8.9	2.0	44	1.2E-04	3.6E-05	175	2.0
	Taylor 0.7	93.70	8.9	1.7	27	1.1E-04	3.2E-05	106	3.3
	Taylor 0.9	82.50	11.2	1.7	30	1.1E-04	3.1E-05	121	2.9
	Hessian 0.3	92.30	12.7	2.7	60	2.7E-04	7.7E-05	239	1.5
	Hessian 0.5	94.10	12.9	2.5	44	2.2E-04	6.4E-05	174	2.0
	Hessian 0.7	94.00	16.5	1.9	31	1.2E-04	3.4E-05	123	2.9
	Hessian 0.9	82.00	16.2	1.9	34	1.1E-04	3.3E-05	134	2.6

References

[1] Gongfan Fang, Xinyin Ma, Mingli Song, Michael Bi Mi, and Xinchao Wang. DepGraph: Towards Any Structural Pruning. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16091–16101, Vancouver, BC, Canada, 2023. IEEE. 1

[2] Gongfan Fang, Xinyin Ma, Michael Bi Mi, and Xinchao Wang. Isomorphic Pruning for Vision Models, 2024. arXiv:2407.04616 [cs]. 1

[3] Carlos Gomes, Benedikt Blumenstiel, Joao Lucas de Sousa Almeida, Pedro Henrique de Oliveira, Paolo Fraccaro, Francesc Marti Escofet, Daniela Szwarcman, Naomi Simumba, Romeo Kienzler, and Bianca Zadrozny. TerraTorch: The Geospatial Foundation Models Toolkit, 2025. arXiv:2503.20563 [cs]. 1

[4] Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth. Introducing EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 204–207. IEEE, 2018. 1

[5] Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019. Publisher: IEEE. 1

[6] Z. Jin, C. Lin, C. Weigl, J. Obarowski, and D. Hale. Smallholder Cashew Plantations in Benin, 2021. Version Number: 1.0. 1

[7] Alexandre Lacoste, Nils Lehmann, Pau Rodriguez, Evan David Sherwin, Hannah Kerner, Björn Lütjens, Jeremy Andrew Irvin, David Dao, Hamed Alemohammad, Alexandre Drouin, Mehmet Gunturkun, Gabriel Huang, David Vazquez, Dava Newman, Yoshua Bengio, Stefano Ermon, and Xiao Xiang Zhu. GEO-Bench: Toward Foundation Models for Earth Monitoring, 2023. arXiv:2306.03831 [cs]. 1

[8] Valerio Marsocci, Yuru Jia, Georges Le Bellier, David Kerekes, Liang Zeng, Sebastian Hafner, Sebastian Gerard, Eric Brune, Ritu Yadav, Ali Shibli, Heng Fang, Yifang Ban, Maarten Vergauwen, Nicolas Audebert, and Andrea Nascetti. PANGAEA: A Global and Inclusive Benchmark for Geospatial Foundation Models, 2025. arXiv:2412.04204 [cs]. 1

[9] Michal Muszynski, Levente Klein, Ademir Ferreira da Silva, Anjani Prasad Atluri, Carlos Gomes, Daniela Szwarcman, Gurkanwar Singh, Kewen Gu, Maciel Zortea, Naomi Simumba, Paolo Fraccaro, Shraddha Singh, Steve Meliksetian, Campbell Watson, Daiki Kimura, and Harini Srinivasan. Fine-tuning of Geospatial Foundation Models for Aboveground Biomass Estimation, 2024. arXiv:2406.19888 [cs]. 1

[10] Christopher Phillips, Sujit Roy, Kumar Ankur, and Rahul Ramachandran. HLS Foundation Burnscars Dataset, 2023. 1