SSL4Eco: A Global Seasonal Dataset for Geospatial Foundation Models in Ecology

Supplementary Material

A-1. SSL4Eco Dataset Construction

In this section, we provide more details on our dataset construction protocol.

Spatial Sampling. We use the same approach as Major-TOM [29] for sampling locations uniformly across the landmass. Our locations correspond to the center of the grid cells.

Seasonal Sampling. As explained in Section 3.1 and Figure 2, we define 4 seasons as intervals between Greenup, Maturity, Senescence, Dormancy, and next Greenup variables. The definition of these EVI variables can be found in Section A-3. For each variable, we calculate the median day in the available years. The EVI product from the MCD12Q2 v6.1 [30] product has missing values in non-vegetated and some evergreen areas (*e.g.* tropics), for which we expect low seasonal variation. We populate these with a nearest-neighbor approach by searching across geographical space.

For each location and season, we preselect all Sentinel-2 tiles across the 6 years of data available 2017-2024. The broad range of years was chosen to account for high cloud coverage in some areas (*e.g.* tropics in wet seasons). Following previous work [89], we remove the tiles with less than less than 20% cloud coverage. Finally, we choose the date and tile with the lowest cloud coverage for the location-season at hand. If fewer than four seasonal images are available for a location due to cloud filtering, we use the 2 or 3 images that are available with less than 20% cloud coverage. Locations with only one image are excluded, accounting for 3% of initially sampled locations, mostly in the tropics and Antarctica. Hence, final patches may be clouded, but the construction process ensures that the overall dataset has less than 20% cloud coverage.

We stress that the scope of this work is to study impact of spatiotemporal sampling compared to existing widely-used 4-date seasonal datasets such as SeCo [58] and SSL4EO [89]. As such, we follow the standard preprocessing procedure of these datasets regarding cloud filtering and the number of seasonal dates per year fair comparison across the computer vision literature. However, realistic Earth Observation applications would require methods capable of handling arbitrarily sampled, potentially clouded, time series of satellite observations. We leave this exploration of the required dataset and models for further work.

Data Source. Several open-access satellite products support vegetation monitoring.

- Landsat missions [91] offer a long-term multispectral record at 30 m resolution, with a 16-day revisit cycle (reduced to 8 days since 2013).
- MODIS [45, 76] provides more spectral bands and a 1–2 day revisit rate, though at a coarser 250–1000 m resolution.
- Since 2015, Sentinel-2 [70] has been delivering 10 m global imagery with a 5-day maximum revisit period, balancing high spatial and temporal resolution. The Sentinel-2 instrument captures spectral bands indicative of ecological patterns, such as red-edge wavelengths sensitive to vegetation stress and chlorophyll content [18].
- Radar sensors may provide diverse ecological insights depending on their frequency: C-band such as Sentinel-1 [83] detects foliage, topography, and moisture, while L-band such as ALOS PALSAR [75] can characterize wood structure.

In this paper, we chose Sentinel-2 due to its widespread use for large-scale vegetation monitoring [46, 52, 80], but we believe our conclusions remain applicable and may be extended to other satellite products in future works. We leave the exploration of our proposed spatiotemporal sampling for multimodal representation learning for future work.

Downloading. The SSL4Eco dataset is downloaded from Google Earth Engine using code from SeCo [58] and SSL4EO-S12 [89] with altered data source, seasonality, and data distribution. We use the Sentinel-2A MSI collection which, compared to Sentinel-2C, has atmospheric correction and depicts more accurately features on the ground [70]. We use harmonized version of the product instead of the original one, as it corrects for normalization issues in 2022. We use Sentinel tiles with less than 20% cloud coverage.

A-2. Implementation Details

In this section, we provide more details on the implementation and training of our models.

Input Bands. Our models SeCo-Eco and MoCo-Eco are trained to take as input the 8 Sentinel-2 bands for ecological applications. Specifically, we use the B2, B3, B4, B5, B6, B7, B8, and B8A bands. While B2-B4 provide information on foliage color, which helps to assess seasonality and plant health, B5-B7 capture red-edge wavelengths sensitive to vegetation stress and chlorophyll content, and B8 and B8A in near-infrared range are useful to distinguish non-vegetated areas. In addition, we also include the NDVI index as a remote sensing-based proxy of vegetation productivity and biomass [69]. As a result, our models expect 9 channels as input.

We leave the exploration of pretraining on our SSL4Eco sampling with more bands or modalities for future work.

Weighted Sampling. Despite the uniform global sampling of SSL4Eco, some locations may have more interesting geographical and seasonal dynamics than others. In order to drive the pretraining towards regions with richer ecological patterns, we use a weighted sampling in our pretraining dataloader. Specifically, we assign a $\div 4$ weight to non-vegetated areas, identified as mean NDVI < 0.1 in all seasons (17% of SSL4Eco), focusing less on deserts and ice packs. We oversample mountain regions with a $\times 2$ weight, identified with the GMBA Mountain Inventory [57] (16% of SSL4Eco), focusing more on ecologically diverse areas, as mountain regions harbor the highest diversity and heterogeneity of ecoregions.

Pretraining. We pretrain SeCo-Eco using the hyperparameters and code provided by Mañas *et al*. [58], using MoCo v2 [11], with minor changes: we replace the RGB input with multispectral images and set the length of the negative examples queue to 65 536, following the implementation of Wang *et al*. [89].

We pretrain MoCo-Eco using the hyperparameters and code provided by Wang *et al*. [89], adapted for a single A100 GPU with batch size of 256.

Finally, we modify the random seasonal sampling found in the implementations of SeCo [58] and SSL4EO [89]. When randomly selecting seasons at batch construction time, both use:
`np.random.choice(..., replace=True)`,
although we believe:
`np.random.choice(..., replace=False)`
is the correct implementation of their respective methods, as this avoids contrasting an image against itself.

A-3. EVI-based Seasonality

We use the Enhanced Vegetation Index (EVI) from the MCD12Q2 v6.1 [30] product of the MODIS [45] satellite mission to define our local, phenology-informed seasons. Similar to NDVI, the EVI index is commonly used to quantify the greenness of an area, but is more sensitive in areas with dense vegetation cover. Figure A-1 illustrate a typical EVI curve over the year, and Table A-1 details how the Greenup, Maturity, Senescence, and Dormancy seasonality variables are defined. For each location in our dataset, we choose 4 images, one for each season, close to the middle between the four EVI-derived variables. See the MCD12Q2 user guide [44] for more details on EVI variables.

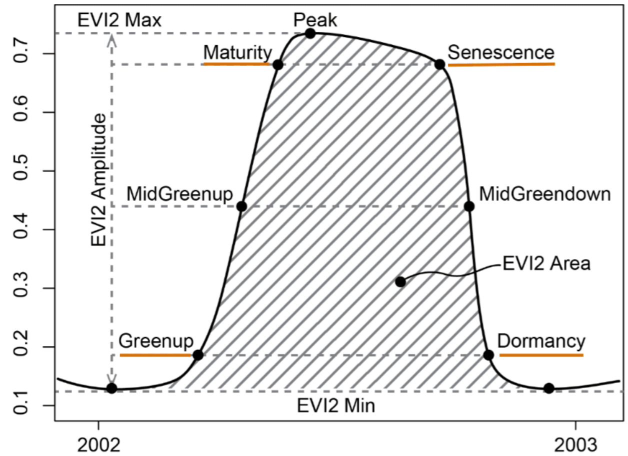


Figure A-1. Enhanced Vegetation Index (EVI) curve of the vegetation cycle at a given location. Based on this curve, the Greenup, Maturity, Senescence, and Dormancy seasonality variables are defined as detailed in Tab. A-1. Image taken from [44].

Name	Definition - Date when...
Greenup	EVI first crossed 15% of segment EVI amplitude
Maturity	EVI first crossed 90% of segment EVI amplitude
Senescence	EVI last crossed 90% of segment EVI amplitude
Dormancy	EVI last crossed 15% of segment EVI amplitude

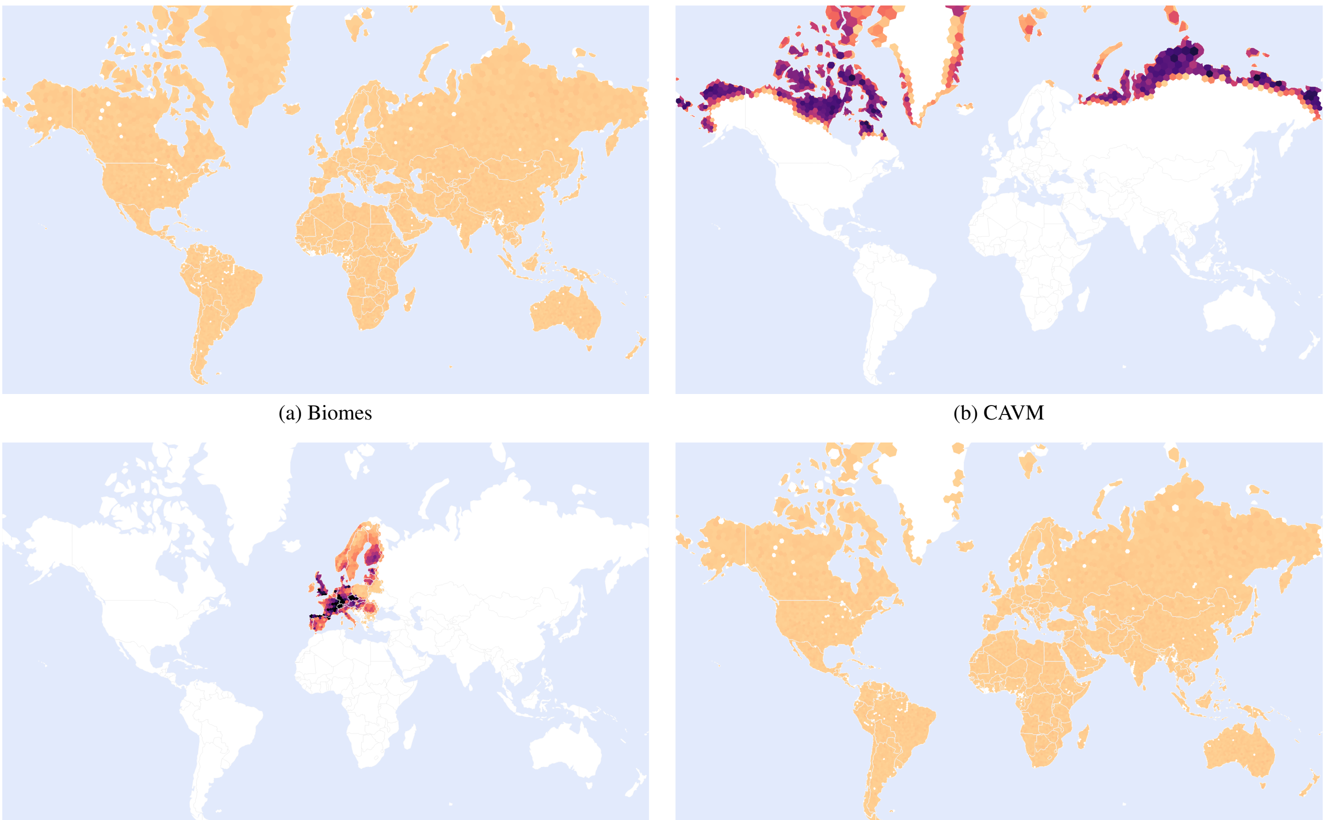
Table A-1. Definition of the Greenup, Maturity, Senescence, and Dormancy seasonality variables based on the EVI curve (Fig. A-1).

A-4. Calendar Ablation

Our temporal sampling of SSL4Eco described in Section 3.1 makes the assumption that pretraining on EVI-based seasonal samplings rather than calendar seasons yields better features for ecological downstream tasks. To verify this claim, we assemble the SSL4Eco-Calendar dataset, which follows the same spatial sampling as SSL4Eco, but with a temporal sampling based on calendar dates following SSL4EO-S12 [89]. We derive SeCo-Calendar from this dataset, by using the same pretraining recipe and backbone as for our SeCo-Eco, and compare in Table A-2 their respective performance across downstream tasks. We observe that our proposed EVI-based seasonal sampling yields representations which overall perform better than calendar-based sampling on most downstream tasks. In particular, EU-Forest (+1.5 micro F1), TSAI (+1.9 macro F1), and Biomes (+0.9 macro F1) prove

Model	BE10% (micro mAP) ↑	CLEF (micro F1) ↑	EU-Forest (micro F1) ↑	TSAI (micro F1) ↑	Biomes (macro F1) ↑	CAVM (macro F1) ↑	BioMassters (mean R^2) ↑	Chelsa (mean R^2) ↑
SeCo-Calendar	**85.3** $\pm$ 0.0	22.4	34.2 $\pm$ 0.1	40.8 $\pm$ 0.0	55.2 $\pm$ 1.0	58.7 $\pm$ 0.8	**75.7** $\pm$ 0.0	80.6 $\pm$ 0.5
SeCo-Eco (ours)	**85.3** $\pm$ 0.0	**22.7**	**35.7** $\pm$ 0.4	**42.7** $\pm$ 0.0	**56.1** $\pm$ 0.7	**59.4** $\pm$ 1.0	75.1 $\pm$ 0.0	**81.1** $\pm$ 0.4

Table A-2. Linear probing comparison of SeCo-Eco and SeCo-Calendar pretrained on EVI-based and calendar-based seasonal samplings, respectively. EVI-based samplings overally yields better features for downstream macroecological tasks, with the exception of the BioMassters dataset. **Best**.



(a) Biomes

(b) CAVM

(c) EU-Forest

(d) CHELSA

Figure A-2. Spatial distribution of the four new downstream tasks created for this work. We sample Biomes and CHELSA locations uniformly across the landmass. Meanwhile, the CAVM dataset is located in arctic regions and EU-Forest is limited to Europe.

to benefit from the finer phenology-informed features of SeCo-Eco. These results validate the importance of temporal sampling and the definition of local seasonality to capture local ecological patterns.

A-5. Downstream Tasks

We illustrate in Figure A-2 the spatial distribution of the samplings used for the new downstream tasks proposed in this paper: Biomes, CAVM, EU-Forest, and CHELSA

A-6. Detailed Results

Beyond evaluating performance with the most established metric per dataset, we provide further experimental results on an expanded set of metrics.

Model	BE10% [82]							
	Macro F1 ↑		Micro F1 ↑		Macro mAP ↑		Micro mAP ↑	
	LP	30-NN	LP	30-NN	LP	30-NN	LP	30-NN
SeCo [58]	56.3 ± 0.3	36.0 ± 0.1	68.9 ± 0.2	44.7 ± 0.1	64.5 ± 0.2	62.4 ± 0.2	79.2 ± 0.0	77.8 ± 0.1
SatMAE [16]	58.9 ± 0.7	39.0 ± 0.1	69.3 ± 0.3	47.5 ± 0.1	66.2 ± 0.3	65.1 ± 0.2	79.7 ± 0.2	79.6 ± 0.0
Satlas [5]	55.7 ± 1.2	37.3 ± 0.1	67.3 ± 0.7	45.9 ± 0.1	64.8 ± 0.2	62.2 ± 0.2	77.9 ± 0.2	77.9 ± 0.0
Croma [31]	59.9 ± 0.5	37.2 ± 0.1	70.7 ± 0.2	46.1 ± 0.1	67.1 ± 0.1	63.6 ± 0.3	80.7 ± 0.2	79.1 ± 0.0
SSL4EO [89]	$\underline{63.1} \pm 0.2$	$\underline{39.6} \pm 0.1$	$\underline{72.5} \pm 0.2$	$\underline{47.9} \pm 0.1$	$\underline{71.1} \pm 0.3$	$\underline{67.8} \pm 0.2$	$\underline{83.2} \pm 0.1$	$\underline{81.1} \pm 0.0$
DOFA [93]	59.9 ± 0.6	37.8 ± 0.2	70.1 ± 0.2	46.1 ± 0.1	66.9 ± 0.2	62.7 ± 0.2	80.1 ± 0.0	77.3 ± 0.1
SeCo-Eco (ours)	$\mathbf{66.8} \pm 0.3$	$\mathbf{41.4} \pm 0.1$	$\mathbf{75.0} \pm 0.1$	$\mathbf{49.9} \pm 0.1$	$\mathbf{74.1} \pm 0.2$	$\mathbf{71.7} \pm 0.2$	$\mathbf{85.3} \pm 0.0$	$\mathbf{84.0} \pm 0.0$

Table A-3. Linear probing and K-Nearest Neighbor performance across multiple metrics for the BigEarthNet-10% task. **Best**, second best.

Model	EU-Forest [60]							
	Macro AUROC ↑		Macro F1 ↑		Micro AUROC ↑		Micro F1 ↑	
	LP	5-NN	LP	5-NN	LP	5-NN	LP	5-NN
SeCo [58]	82.6 ± 0.0	63.9 ± 0.3	12.3 ± 0.7	18.2 ± 0.3	90.6 ± 0.1	77.6 ± 0.2	31.3 ± 0.9	30.6 ± 0.2
SatMAE [16]	$\underline{84.6} \pm 0.2$	$\mathbf{66.7} \pm 0.4$	$\mathbf{15.0} \pm 0.7$	$\mathbf{21.0} \pm 0.3$	$\underline{91.6} \pm 0.1$	$\mathbf{79.8} \pm 0.2$	$\mathbf{35.7} \pm 0.9$	$\mathbf{33.3} \pm 0.1$
Satlas [5]	81.1 ± 0.3	62.7 ± 0.3	10.1 ± 0.4	17.5 ± 0.3	89.6 ± 0.1	76.7 ± 0.2	29.8 ± 1.5	30.0 ± 0.2
Croma [31]	82.9 ± 0.3	63.6 ± 0.3	12.2 ± 0.7	18.1 ± 0.3	90.5 ± 0.2	77.8 ± 0.2	32.3 ± 0.9	30.9 ± 0.2
SSL4EO [89]	83.9 ± 0.0	65.0 ± 0.3	11.6 ± 0.4	19.3 ± 0.3	91.2 ± 0.2	78.5 ± 0.2	32.6 ± 0.1	31.5 ± 0.2
DOFA [93]	83.1 ± 0.1	63.1 ± 0.5	13.5 ± 0.5	17.6 ± 0.5	90.7 ± 0.1	77.3 ± 0.3	$\underline{34.8} \pm 0.9$	29.9 ± 0.3
SeCo-Eco (ours)	$\mathbf{84.8} \pm 0.2$	$\underline{65.6} \pm 0.2$	$\underline{14.8} \pm 0.6$	$\underline{19.9} \pm 0.2$	$\mathbf{91.7} \pm 0.1$	$\underline{79.0} \pm 0.1$	$\mathbf{35.7} \pm 0.4$	$\underline{32.4} \pm 0.2$

Table A-4. Linear probing and K-Nearest Neighbor performance across multiple metrics for the EUForest task. **Best**, second best.

Model	TreeSatAI [1]							
	Macro F1 ↑		Macro MAP ↑		Micro F1 ↑		Micro MAP ↑	
	LP	5-NN	LP	5-NN	LP	5-NN	LP	5-NN
SeCo [58]	10.1 ± 0.0	24.3	24.3 ± 0.0	20.5	23.4 ± 0.0	35.2	44.6 ± 0.0	34.6
SatMAE [16]	$\mathbf{21.0} \pm 0.1$	**33.7**	$\mathbf{36.8} \pm 0.1$	**35.8**	$\mathbf{46.8} \pm 0.3$	**43.7**	$\mathbf{58.0} \pm 0.1$	**52.3**
Satlas [5]	17.8 ± 0.0	30.1	32.4 ± 0.0	27.9	42.9 ± 0.0	40.8	54.2 ± 0.0	45.4
Croma [31]	$\underline{20.3} \pm 0.0$	30.1	$\underline{34.9} \pm 0.0$	27.8	$\underline{43.8} \pm 0.0$	40.7	$\underline{56.6} \pm 0.0$	45.6
SSL4EO [89]	18.2 ± 0.0	<u>30.2</u>	33.1 ± 0.0	28.4	42.3 ± 0.0	<u>40.9</u>	54.5 ± 0.0	<u>46.0</u>
DOFA [93]	14.7 ± 0.0	26.2	28.7 ± 0.0	21.9	35.1 ± 0.0	37.3	50.8 ± 0.0	37.5
SeCo-Eco (ours)	19.2 ± 0.0	29.7	34.3 ± 0.0	<u>29.0</u>	42.7 ± 0.0	40.6	54.8 ± 0.0	45.7

Table A-5. Linear probing and K-Nearest Neighbor performance across multiple metrics for the TreeSatAI task. Due to the fixed splits, no standard deviation can be reported for K-Nearest Neighbor probing. **Best**, second best.

Model	Biomes [65]									
	Macro Acc ↑		Macro AUROC ↑		Macro F1 ↑		Micro Acc ↑		Micro F1 ↑	
	LP	10-NN	LP	10-NN	LP	10-NN	LP	10-NN	LP	10-NN
SeCo [58]	40.0 ± 0.4	35.4 ± 0.7	91.2 ± 0.6	79.8 ± 1.0	41.6 ± 0.5	36.9 ± 1.0	62.7 ± 0.5	59.2 ± 0.5	62.7 ± 0.5	59.2 ± 0.5
SatMAE [16]	49.9 ± 1.0	46.1 ± 0.5	93.7 ± 0.4	88.8 ± 0.4	51.4 ± 1.1	47.8 ± 0.7	69.0 ± 0.5	66.7 ± 0.6	69.0 ± 0.5	66.7 ± 0.6
Satlas [5]	47.1 ± 1.4	45.9 ± 0.7	92.8 ± 0.5	88.4 ± 0.4	48.3 ± 1.6	47.6 ± 0.9	65.6 ± 0.8	65.1 ± 0.5	65.6 ± 0.8	65.1 ± 0.5
Croma [31]	46.2 ± 1.8	41.2 ± 0.5	92.2 ± 0.4	85.7 ± 0.6	47.2 ± 1.4	42.2 ± 0.6	65.7 ± 0.7	61.7 ± 0.3	65.7 ± 0.7	61.7 ± 0.3
SSL4EO [89]	$\underline{51.3} \pm 0.9$	$\underline{48.2} \pm 0.5$	$\underline{94.3} \pm 0.6$	$\underline{89.6} \pm 0.8$	$\underline{53.4} \pm 1.0$	$\underline{49.7} \pm 0.5$	$\underline{70.4} \pm 0.5$	$\underline{67.6} \pm 0.6$	$\underline{70.4} \pm 0.5$	$\underline{67.6} \pm 0.6$
DOFA [93]	48.1 ± 1.4	41.8 ± 0.4	92.9 ± 0.3	85.7 ± 0.6	49.7 ± 1.3	43.0 ± 0.5	66.4 ± 0.6	61.8 ± 0.5	66.4 ± 0.6	61.8 ± 0.5
SeCo-Eco (ours)	**53.9** ± 0.7	**49.3** ± 0.7	**95.5** ± 0.4	**90.0** ± 0.7	**56.1** ± 0.7	**51.2** ± 0.9	**72.9** ± 0.5	**69.4** ± 0.4	**72.9** ± 0.5	**69.4** ± 0.4

Table A-6. Linear probing and K-Nearest Neighbor performance across multiple metrics for the biomes classification task. **Best**, <u>second best</u>.

Model	CAVM [73]									
	Macro Acc ↑		Macro AUROC ↑		Macro F1 ↑		Micro Acc ↑		Micro F1 ↑	
	LP	20-NN	LP	20-NN	LP	20-NN	LP	20-NN	LP	20-NN
SeCo [58]	53.2 ± 0.6	50.3 ± 0.6	87.3 ± 0.3	85.6 ± 0.3	54.5 ± 0.7	52.1 ± 0.7	61.4 ± 0.6	60.6 ± 0.5	61.4 ± 0.6	60.6 ± 0.5
SatMAE [16]	55.2 ± 1.6	54.0 ± 0.6	88.3 ± 0.3	87.9 ± 0.3	56.4 ± 1.5	55.8 ± 0.7	63.0 ± 0.5	63.5 ± 0.5	63.0 ± 0.5	63.5 ± 0.5
Satlas [5]	52.7 ± 2.1	51.5 ± 0.4	87.6 ± 0.3	86.6 ± 0.3	53.8 ± 2.0	53.2 ± 0.5	61.2 ± 0.5	61.2 ± 0.5	61.2 ± 0.5	61.2 ± 0.5
Croma [31]	52.7 ± 1.3	50.1 ± 0.7	87.4 ± 0.3	85.6 ± 0.4	53.7 ± 1.2	51.6 ± 0.8	61.0 ± 0.7	60.3 ± 0.6	61.0 ± 0.7	60.3 ± 0.6
SSL4EO [89]	$\underline{56.0} \pm 0.5$	$\underline{55.0} \pm 0.6$	$\underline{88.9} \pm 0.3$	$\underline{88.2} \pm 0.3$	$\underline{57.5} \pm 0.6$	$\underline{56.9} \pm 0.7$	$\underline{63.7} \pm 0.6$	$\underline{63.7} \pm 0.5$	$\underline{63.7} \pm 0.6$	$\underline{63.7} \pm 0.5$
DOFA [93]	55.3 ± 1.8	51.7 ± 0.5	88.2 ± 0.4	87.0 ± 0.3	56.5 ± 1.6	53.6 ± 0.6	62.4 ± 0.8	62.2 ± 0.4	62.4 ± 0.8	62.2 ± 0.4
SeCo-Eco (ours)	**58.1** ± 1.2	**58.0** ± 0.7	**89.9** ± 0.3	**89.2** ± 0.4	**59.4** ± 1.0	**59.5** ± 0.8	**65.3** ± 0.5	**65.6** ± 0.6	**65.3** ± 0.5	**65.6** ± 0.6

Table A-7. Linear probing and K-Nearest Neighbor performance across multiple metrics for the CAVM classification task. **Best**, <u>second best</u>.

Model	BioMassters [61]					
	Mean R^2 ↑		Mean MAE ↓		Mean RMSE ↓	
	LP	1-NN	LP	1-NN	LP	1-NN
SeCo [58]	51.3 ± 0.0	-19.2	3.9 ± 0.0	$\underline{7.0}$	5.8 ± 0.0	$\underline{11.0}$
SatMAE [16]	59.5 ± 0.6	-18.0	3.6 ± 0.0	$\underline{7.0}$	5.3 ± 0.0	$\underline{11.0}$
Satlas [5]	62.5 ± 0.9	-17.8	3.3 ± 0.1	$\underline{7.0}$	4.9 ± 0.1	$\underline{11.0}$
Croma [31]	58.5 ± 0.2	-18.1	3.5 ± 0.0	$\underline{7.0}$	5.3 ± 0.0	$\underline{11.0}$
SSL4EO [89]	$\underline{71.4} \pm 0.0$	$\underline{-16.8}$	$\underline{2.8} \pm 0.0$	**6.9**	$\underline{4.2} \pm 0.0$	**10.9**
DOFA [93]	63.1 ± 0.4	-18.3	3.2 ± 0.0	$\underline{7.0}$	4.8 ± 0.0	$\underline{11.0}$
SeCo-Eco (ours)	**75.2** ± 0.1	**-16.3**	**2.5** ± 0.0	**6.9**	**3.8** ± 0.0	**10.9**

Table A-8. Linear probing and K-Nearest Neighbor performance across multiple metrics for the BioMassters task. Due to the fixed splits, no standard deviation can be reported for K-Nearest Neighbor probing. **Best**, <u>second best</u>.

Model	CHELSA Climate [47] - Temperature & Precipitation							
	Temp MAE ↓		Temp R^2 ↑		Prec MAE ↓		Prec R^2 ↑	
	LP	10-NN	LP	10-NN	LP	10-NN	LP	10-NN
SeCo [58]	572.3 ± 1.1	547.8 ± 1.7	63.1 ± 0.3	61.3 ± 0.3	33380.8 ± 291.5	30725.5 ± 171.8	60.3 ± 0.7	60.7 ± 0.8
SatMAE [16]	<u>482.0</u> ± 2.3	411.4 ± 1.2	<u>74.4</u> ± 0.2	<u>76.1</u> ± 0.2	30999.5 ± 314.9	<u>27087.1</u> ± 135.4	65.2 ± 0.4	<u>67.1</u> ± 0.5
Satlas [5]	595.1 ± 3.4	474.7 ± 3.6	62.1 ± 0.4	69.4 ± 0.7	36698.8 ± 685.3	29535.8 ± 95.1	55.9 ± 1.0	62.4 ± 0.7
Croma [31]	511.5 ± 2.5	505.5 ± 1.6	71.1 ± 0.2	66.4 ± 0.2	32887.8 ± 350.8	30974.2 ± 96.6	61.4 ± 0.6	60.3 ± 0.4
SSL4EO [89]	496.1 ± 1.1	<u>410.7</u> ± 0.8	72.4 ± 0.2	75.8 ± 0.3	<u>30960.7</u> ± 154.7	27989.7 ± 148.3	<u>65.5</u> ± 0.4	65.4 ± 0.4
DOFA [93]	576.0 ± 0.7	505.9 ± 0.9	63.9 ± 0.3	66.9 ± 0.3	34860.1 ± 297.0	30311.1 ± 182.9	59.7 ± 0.5	59.9 ± 0.7
SeCo-Eco (ours)	**411.4** ± 0.9	**364.8** ± 0.7	**80.7** ± 0.2	**80.5** ± 0.2	**27695.5** ± 74.8	**25946.7** ± 72.6	**70.2** ± 0.3	**69.5** ± 0.4

Table A-9. Linear probing and K-Nearest Neighbor performance overview for the CHELSA Climate task. We break down the predictions for temperature and precipitation. **Best**, <u>second best</u>.

Model	CHELSA Climate [47] - Evapotranspiration & Site Water Balance							
	Evap MAE ↓		Evap R^2 ↑		Swb MAE ↓		Swb R^2 ↑	
	LP	10-NN	LP	10-NN	LP	10-NN	LP	10-NN
SeCo [58]	2131.6 ± 8.7	2068.0 ± 9.8	68.9 ± 0.2	67.1 ± 0.1	24903.7 ± 137.7	23878.3 ± 143.0	80.9 ± 0.2	80.5 ± 0.2
SatMAE [16]	<u>1760.4</u> ± 6.7	1564.7 ± 5.4	<u>79.2</u> ± 0.2	80.0 ± 0.2	20999.2 ± 87.5	19055.6 ± 66.2	86.9 ± 0.1	87.7 ± 0.1
Satlas [5]	2093.3 ± 6.4	1761.5 ± 10.3	70.9 ± 0.2	75.3 ± 0.5	24115.2 ± 131.3	20772.5 ± 116.8	83.4 ± 0.2	85.5 ± 0.1
Croma [31]	1872.4 ± 27.2	1882.2 ± 5.6	76.2 ± 0.5	73.2 ± 0.2	23003.2 ± 270.2	21593.9 ± 79.0	84.6 ± 0.4	84.7 ± 0.2
SSL4EO [89]	1786.0 ± 3.4	<u>1522.8</u> ± 3.0	78.6 ± 0.2	<u>81.1</u> ± 0.2	<u>20444.3</u> ± 58.1	<u>18155.6</u> ± 42.4	<u>87.7</u> ± 0.1	<u>88.9</u> ± 0.1
DOFA [93]	2086.1 ± 3.5	1911.6 ± 6.2	71.2 ± 0.3	72.0 ± 0.3	23943.5 ± 38.3	22370.0 ± 60.8	83.4 ± 0.1	83.5 ± 0.2
SeCo-Eco (ours)	**1537.6** ± 4.2	**1391.2** ± 3.3	**83.7** ± 0.1	**83.9** ± 0.2	**18567.4** ± 90.2	**17257.4** ± 50.6	**89.6** ± 0.1	**89.9** ± 0.1

Table A-10. Linear probing and K-Nearest Neighbor performance overview for the CHELSA Climate task. We break down the predictions for evapotranspiration and site water balance. **Best**, <u>second best</u>.