Global seasonal pre-training dataset (SSL4Eco) and self-supervised model (SeCo-Eco) for ecological applications

Elena Plekhanova[1], Damien Robert[2], Johannes Dollinger[2], Emilia Arens[2], Philipp Brun[1], Niklaus Zimmermann[1]

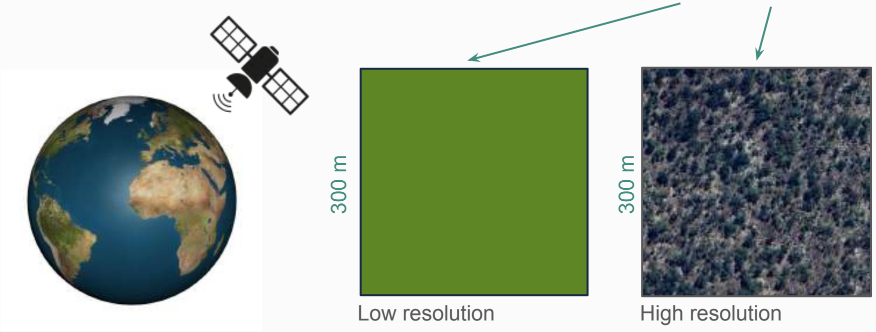
[1] Swiss Federal Research Institute, WSL, Switzerland
[2] University of Zurich, Switzerland

Remote sensing for ecological applications

Tasks

- land cover classification
- species distribution modelling

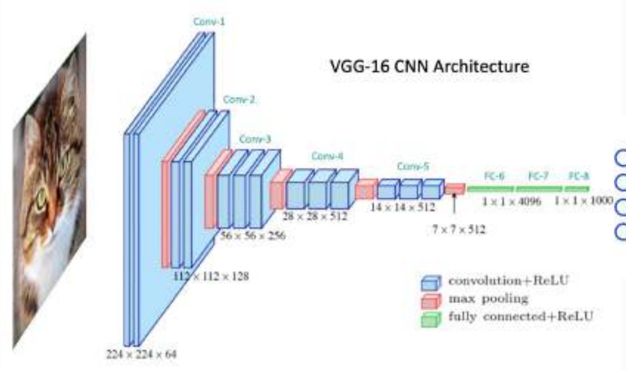
Remote sensing for ecological applications

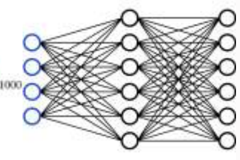


Remote sensing for ecological applications

Low resolution

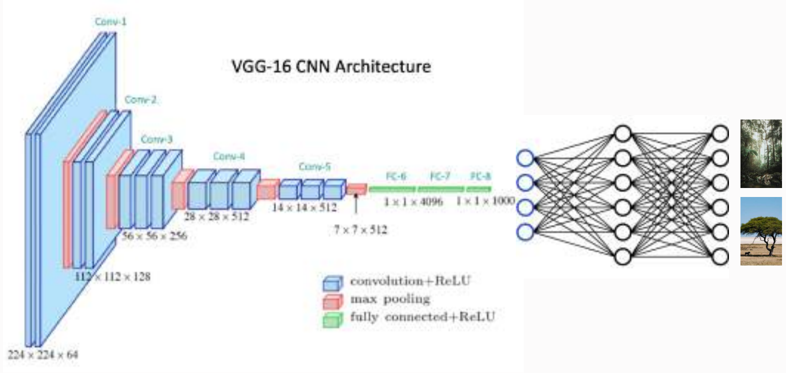
High resolution





Remote sensing for ecological applications

Low resolution
High resolution



Objective

01

Design Sentinel-2 pre-training dataset global and seasonal

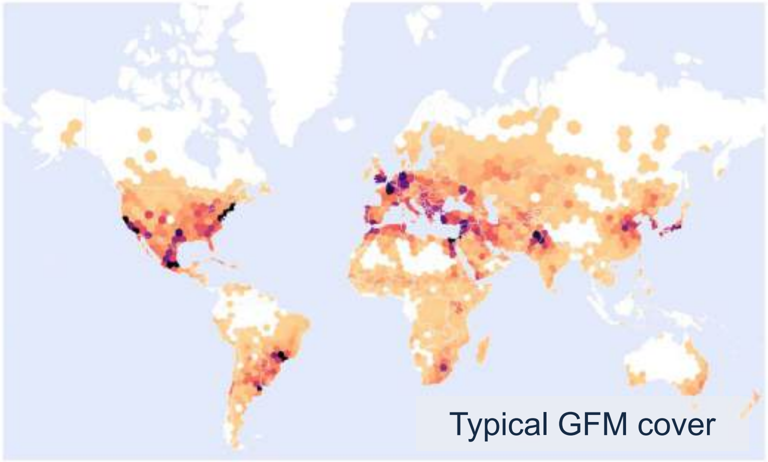
02

Train Geospatial Foundation Model (GFM)

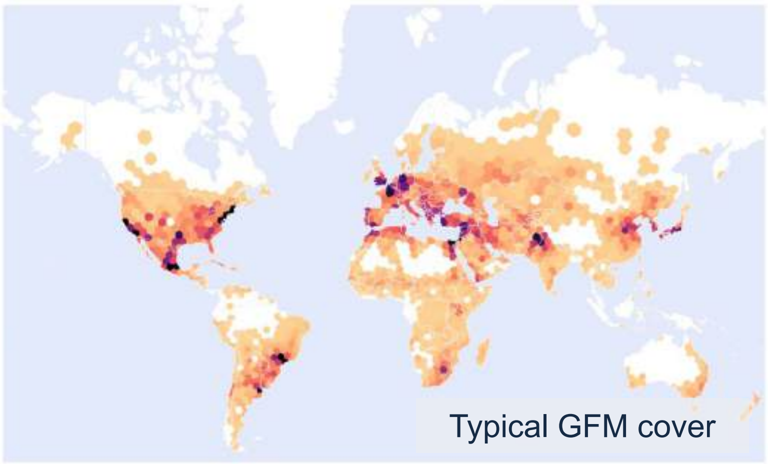
03

Test on ecologically relevant benchmarks

Geographical distribution



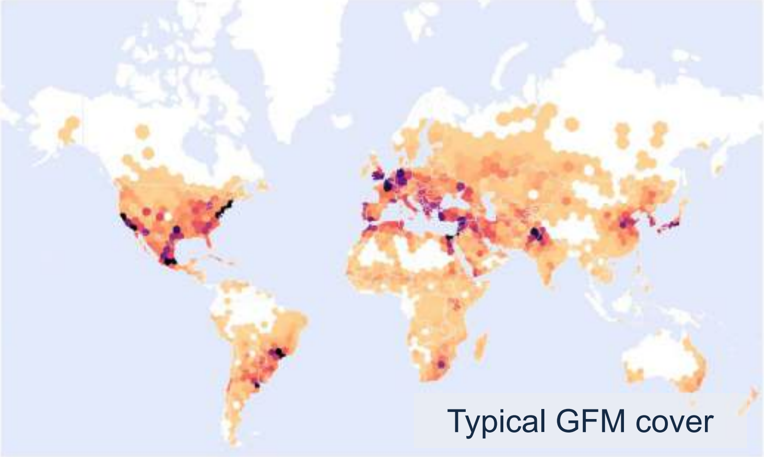
Geographical distribution

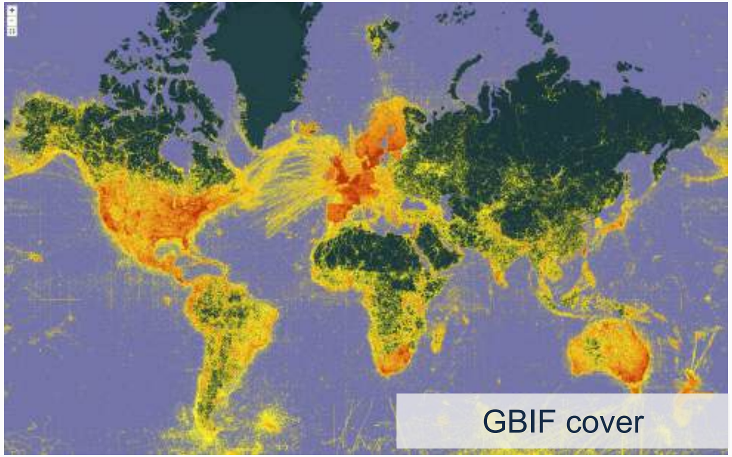


Sampling bias

- centred on cities
- missing entire ecosystems

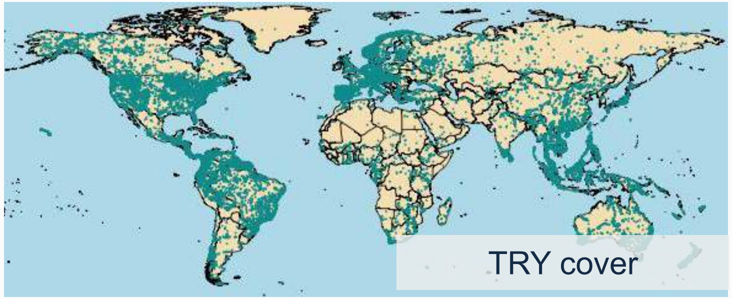
Geographical distribution



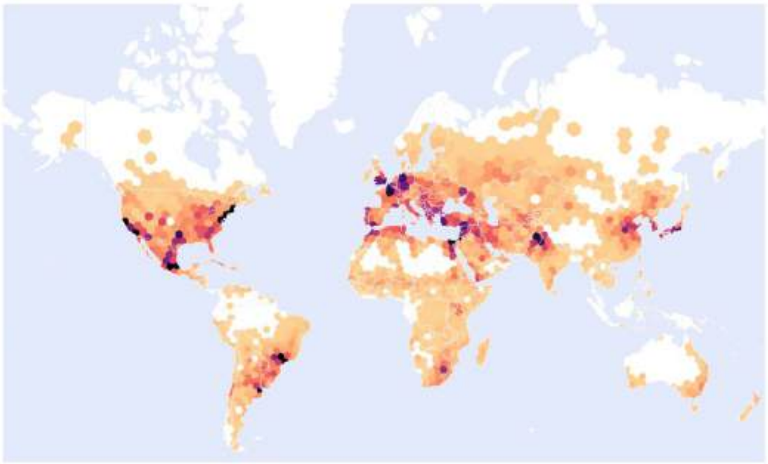


Sampling bias

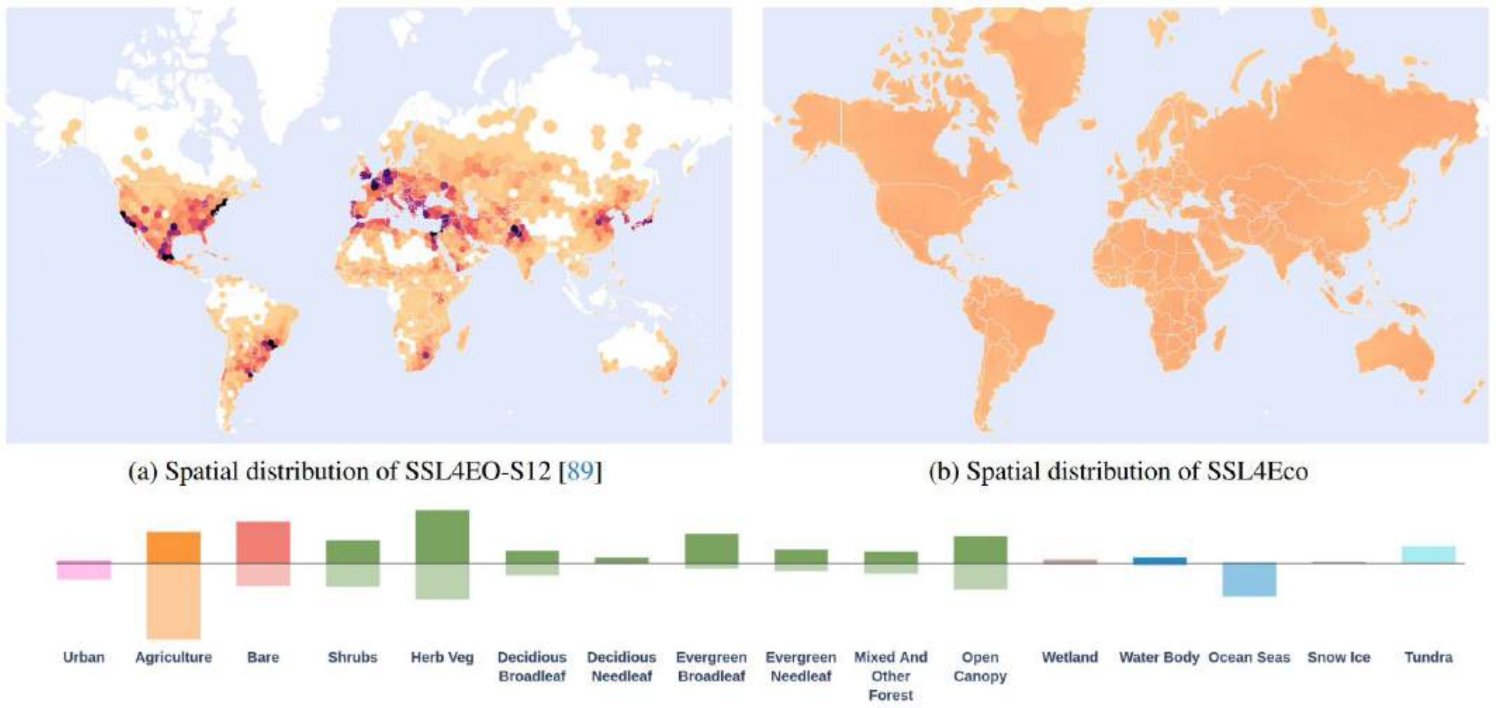
- centred on cities
- missing entire ecosystems



Geographical distribution



Geographical distribution of pretraining dataset



(a) Spatial distribution of SSL4EO-S12 [89]

(b) Spatial distribution of SSL4Eco

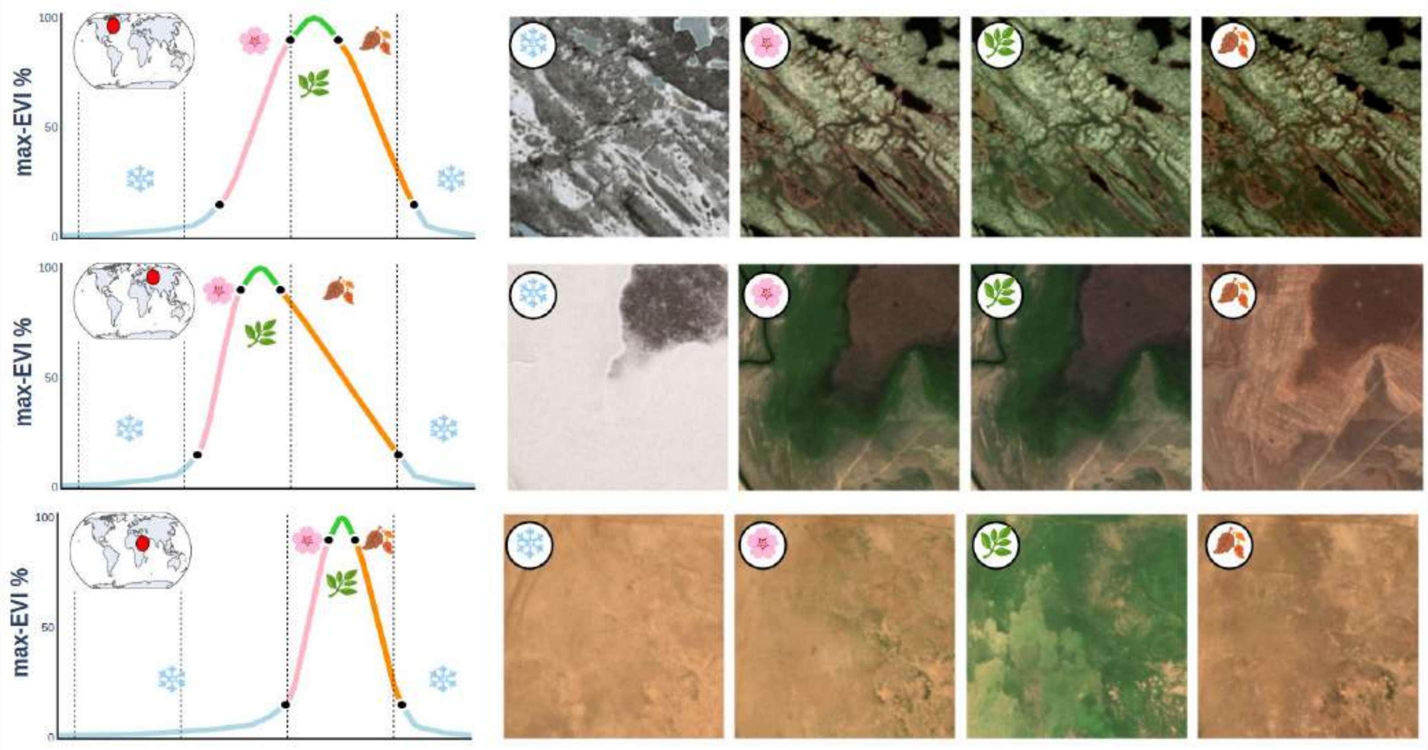
(c) Copernicus land cover [55] distribution for SSL4Eco (upwards) and SSL4EO-S12 [89] (downwards)

Seasonality

How to pick seasons?

- at random
- calendar date
- phenology curve

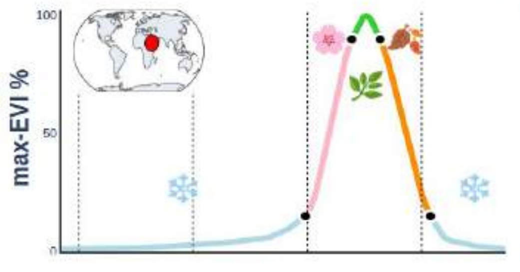
Seasonality



(a) EVI-based seasons (b) Seasonal images

Pre-training dataset

SSL4Eco



(a) EVI-based seasons

(b) Spatial distribution of SSL4Eco

GFM model

Momentum Contrast (MoCo) He et al., *CVPR* 2020

+

WHAT SHOULD NOT BE CONTRASTIVE IN CONTRASTIVE LEARNING

↓ Xiao et al., *ICLR* 2021

Seasonal Contrast (SeCo) Mañas et al., *ICCV* 2021

learns to capture seasons
instead of being invariant to seasons

SeCo-Eco - ResNet50 trained on SSL4Eco with Seasonal Contrast technique

GFM model

Momentum Contrast (MoCo)

He et al., *CVPR* 2020

+

WHAT SHOULD NOT BE CONTRASTIVE IN CONTRASTIVE LEARNING

Xiao et al., *ICLR* 2021

↓

Seasonal Contrast (SeCo)
learns to capture seasons
instead of being invariant to seasons

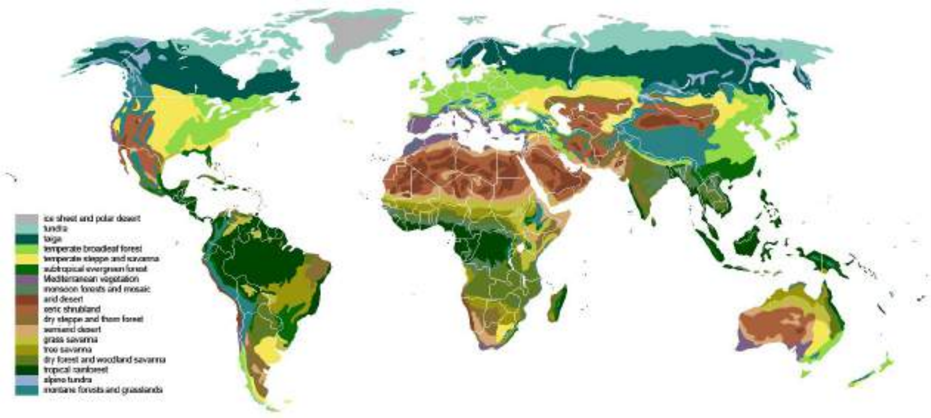
Mañas et al., *ICCV* 2021

SeCo-Eco - ResNet50 trained on SSL4Eco with Seasonal Contrast technique

New benchmarks

Biomes

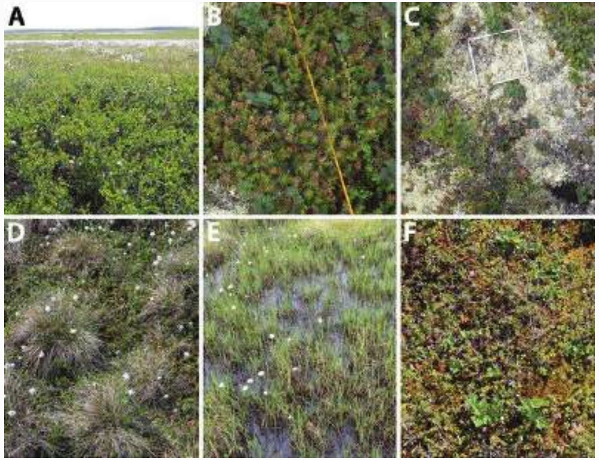
Image classification into one of 15 global biomes
based on Olson et al.

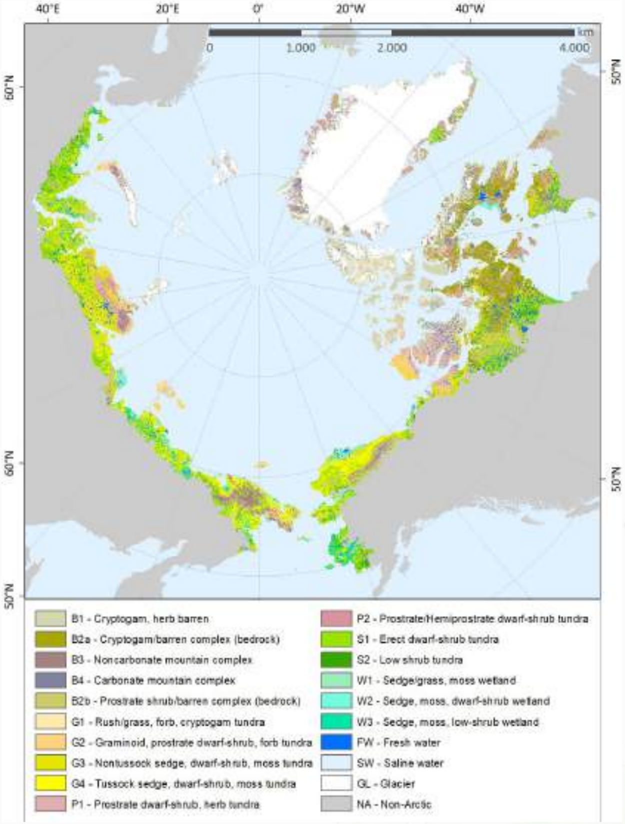


New benchmarks

CAVM

Classification of an image into one of Arctic vegetation types based on CAVM.



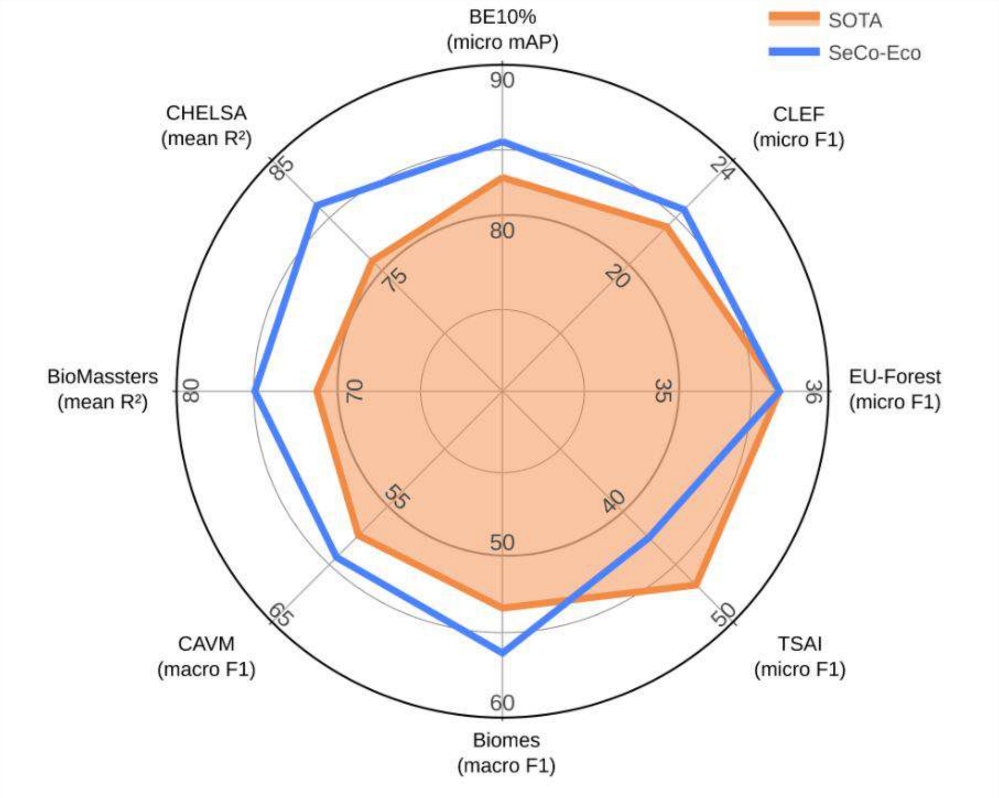


Results

Model	Biomes (macro F1)↑		CAVM (macro F1)↑	
	LP	10-NN	LP	20-NN
SeCo [58]	41.5 ± 0.5	36.9 ± 1.0	54.4 ± 0.7	52.1 ± 0.7
SatMAE [16]	51.3 ± 1.1	47.7 ± 0.7	56.3 ± 1.4	55.8 ± 0.7
Satlas [5]	48.3 ± 1.6	47.6 ± 0.9	53.8 ± 2.0	53.2 ± 0.5
Croma [31]	47.1 ± 1.4	42.2 ± 0.6	53.6 ± 1.2	51.6 ± 0.8
SSL4EO [89]	$\underline{53.3} \pm 1.0$	$\underline{49.7} \pm 0.5$	$\underline{57.5} \pm 9.6$	$\underline{56.9} \pm 0.6$
DOFA [93]	49.7 ± 1.3	42.9 ± 0.5	56.4 ± 1.6	53.5 ± 0.6
SeCo-Eco (ours)	$\mathbf{56.1} \pm 0.7$	$\mathbf{51.1} \pm 0.9$	$\mathbf{59.4} \pm 1.0$	$\mathbf{59.5} \pm 0.8$

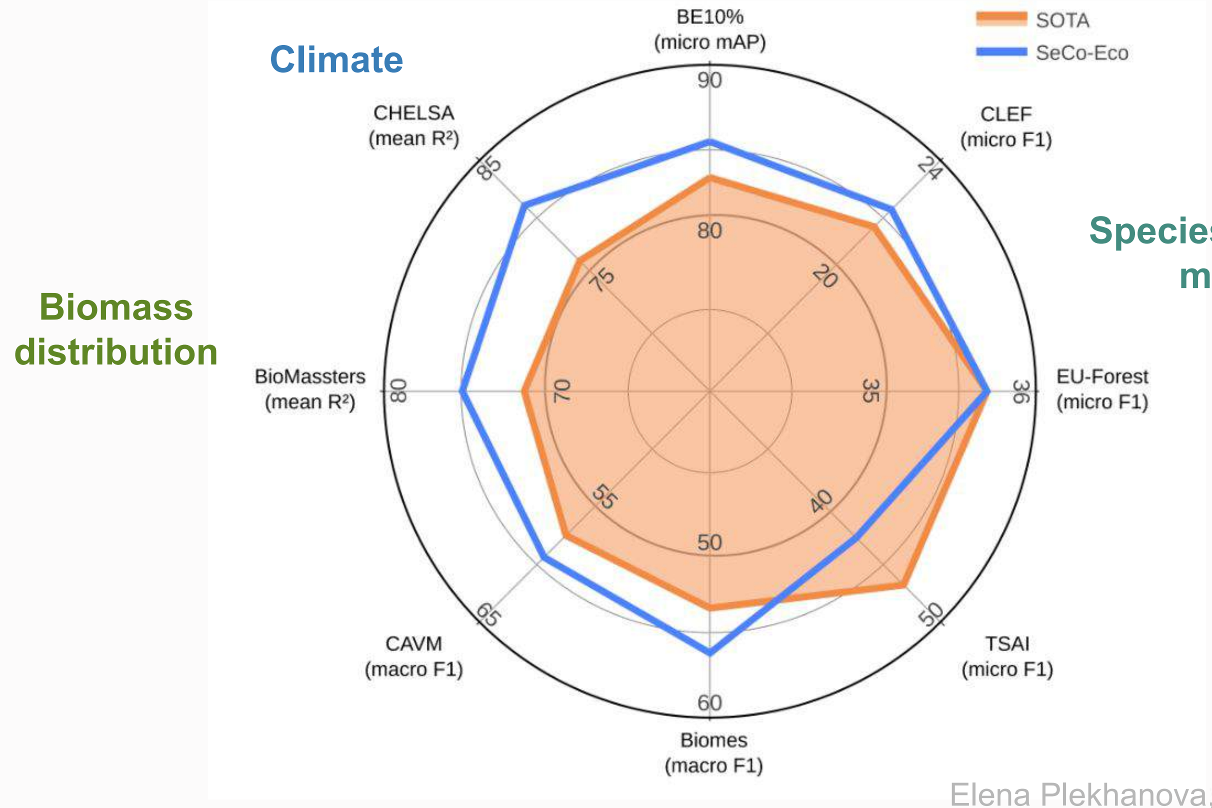
Table 4. Linear probing and K-Nearest Neighbor comparison of state of the art models with our SeCo-Eco pretrained on our SSL4Eco on classification of two land cover datasets: global biomes and Arctic vegetation types [73]. **Best**, <u>second best</u>.

Results



SOTA
SeCo-Eco
BE10%
(micro mAP)
90
80
CLEF
(micro F1)
24
20
EU-Forest
(micro F1)
36
35
TSAI
(micro F1)
50
40
Biomes
(macro F1)
60
50
CAVM
(macro F1)
65
55
BioMassters
(mean R²)
80
70
CHELSA
(mean R²)
85
75

Results



Climate
Biomass distribution
Species distribution modelling
SOTA
SeCo-Eco
BE10%
(micro mAP)
90
80
CLEF
(micro F1)
24
20
EU-Forest
(micro F1)
36
35
TSAI
(micro F1)
50
40
Biomes
(macro F1)
60
50
CAVM
(macro F1)
65
55
BioMassters
(mean R²)
80
70
CHELSA
(mean R²)
85
75

Takeaways

Recommendation for future GFM design

- geographical sampling
- EVI-based seasonality

Practical outcomes

- SSL4Eco pretraining dataset
- SeCo-Eco model for ecological tasks
- ecological benchmarks
- easy to combine with other data modalities

Takeaways

Recommendation for future GFM design

- geographical sampling
- EVI-based seasonality

Practical outcomes

- SSL4Eco pretraining dataset
- SeCo-Eco model for ecological tasks
- ecological benchmarks
- easy to combine with other data modalities

CVPR

EARTHVISION 2025

June 11/12th, Nashville, TN, USA

github.com/PlekhanovaElena/ssl4eco