NL 2021 Selected Samples prediction using NL 2021 training set,
Learning with selected samples classes with no low regional training instances, no extra training categories,
No Calanoida (civ–vi), Cyclopoida, Zooplankton classes in learning set

Confusion Matrix - In percent of Actual Value

Actual Values	Acartia spp.	Evadne spp.	Temora spp.	Pseudocalanus spp.	Oithona spp.	Bryozoa (larvae)	Podon/Pleopsis spp.	Calanoida (ci–ciii)	Oikopleura spp.	Centropages spp.	Echinodermata (larvae)	Harpacticoida– epibenthic	Gastropoda (larvae/Limacina)	Bivalvia (larvae)	Fritillaria spp.	Microcalanus spp.	Chaetognatha	Polychaeta (larvae)	Calanus spp.	Eurytemora spp.
Acartia spp.	73%	<1%	7%	8%	3%		8%		<1%	<1%	<1%		<1%	<1%		<1%				
Evadne spp.		94%	<1%	<1%	<1%	<1%	2%	1%			<1%	<1%	<1%			<1%				
Temora spp.	14%	2%	57%	9%	2%		<1%	16%			<1%				<1%					
Pseudocalanus spp.	20%		13%	58%	<1%			8%			<1%				<1%			<1%	<1%	
Oithona spp.	11%	1%	<1%		80%	<1%		3%			<1%		3%							
Bryozoa (larvae)		38%				50%	1%	4%		2%	1%	2%		<1%	<1%		<1%			
Podon/Pleopsis spp.	1%	16%	21%	2%		<1%	19%	17%					<1%	2%		21%				
Calanoida (ci–ciii)	12%	9%	11%	8%	3%	<1%	7%	42%						5%		2%	<1%			
Oikopleura spp.			<1%	4%	2%			2%	64%		<1%			24%				2%		
Centropages spp.	62%		7%	1%				2%	26%								1%			
Echinodermata (larvae)	6%			18%		2%	6%	2%		22%				42%			2%			
Harpacticoida– epibenthic	21%		6%	9%	9%		15%				40%									
Gastropoda (larvae/Limacina)		9%				9%					60%	11%				13%				
Bivalvia (larvae)		2%			5%	5%					14%	75%								
Fritillaria spp.				9%			6%	3%	3%				79%							
Microcalanus spp.		10%	3%	10%			52%			3%			16%		6%					
Chaetognatha						32%					11%		57%							
Polychaeta (larvae)	5%	15%	10%	5%	5%		15%						45%							
Calanus spp.			21%				7%						71%							
Eurytemora spp.	60%		10%	30%																

Predicted Values

Classification Report Matrix
max 20000 learning objects per class

	precision	recall	f1-score
Acartia spp. (n=2490–train=5448)	0.78	0.73	0.75
Evadne spp. (n=1931–train=2845)	0.90	0.94	0.92
Temora spp. (n=1416–train=5148)	0.67	0.57	0.62
Pseudocalanus spp. (n=1044–train=4552)	0.64	0.58	0.61
Oithona spp. (n=345–train=1409)	0.67	0.80	0.73
Bryozoa (larvae) (n=248–train=119)	0.93	0.50	0.65
Podon/Pleopsis spp. (n=230–train=201)	0.40	0.19	0.26
Calanoida (ci–ciii) (n=130–train=3713)	0.08	0.42	0.13
Oikopleura spp. (n=115–train=761)	0.85	0.64	0.73
Centropages spp. (n=88–train=40)	0.92	0.26	0.41
Echinodermata (larvae) (n=50–train=276)	0.61	0.22	0.32
Harpacticoida– epibenthic (n=47–train=136)	0.56	0.40	0.47
Gastropoda (larvae/Limacina) (n=47–train=110)	0.53	0.60	0.56
Bivalvia (larvae) (n=44–train=71)	0.79	0.75	0.77
Fritillaria spp. (n=34–train=3447)	0.25	0.79	0.39
Microcalanus spp. (n=31–train=80)	0.24	0.16	0.19
Chaetognatha (n=28–train=67)	1.00	0.57	0.73
Polychaeta (larvae) (n=20–train=452)	0.11	0.45	0.18
Calanus spp. (n=14–train=213)	0.62	0.71	0.67
Eurytemora spp. (n=10–train=88)	0.00	0.00	0.00
macro avg	0.58	0.52	0.50
weighted avg	0.74	0.70	0.71

Predictions of discarded taxa from training

Predicted Taxa \ Actual discarded Taxa	Aglantha spp. (medusa) n=9	Calanoida (civ-vi) n=1370	Cyclopoida (unid) n=2	Decapoda-non brachyura (larvae) n=1	Foraminifera n=1	Hydrozoa (medusa) n=47	Monstrillidae n=1	Oncaeidae spp. n=6	Ostracoda n=2	Sarsia spp. (medusa) n=4	Tortanus spp. n=2	Zooplankton (unid) n=318
Acartia spp.	0.0%	24.7% (n=339)	0.0%	0.0%	0.0%	0.0%	100.0% (n=1)	16.7% (n=1)	0.0%	0.0%	50.0% (n=1)	0.6% (n=2)
Bivalvia (larvae)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	9.1% (n=29)
Bryozoa (larvae)	0.0%	0.4% (n=5)	0.0%	0.0%	0.0%	2.1% (n=1)	0.0%	0.0%	50.0% (n=1)	0.0%	0.0%	4.4% (n=14)
Calanoida (ci-ciii)	0.0%	32.0% (n=438)	0.0%	0.0%	0.0%	4.3% (n=2)	0.0%	16.7% (n=1)	50.0% (n=1)	0.0%	0.0%	6.9% (n=22)
Calanus spp.	33.3% (n=3)	0.1% (n=1)	0.0%	100.0% (n=1)	0.0%	10.6% (n=5)	0.0%	0.0%	0.0%	25.0% (n=1)	50.0% (n=1)	0.6% (n=2)
Echinodermata (larvae)	0.0%	0.0%	0.0%	0.0%	0.0%	4.3% (n=2)	0.0%	0.0%	0.0%	0.0%	0.0%	0.9% (n=3)
Evadne spp.	0.0%	5.0% (n=68)	50.0% (n=1)	0.0%	0.0%	25.5% (n=12)	0.0%	0.0%	0.0%	0.0%	0.0%	34.3% (n=109)
Fritillaria spp.	0.0%	1.3% (n=18)	0.0%	0.0%	0.0%	25.5% (n=12)	0.0%	0.0%	0.0%	25.0% (n=1)	0.0%	3.8% (n=12)
Gastropoda (larvae/Limacina)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	8.8% (n=28)
Harpacticoida- epibenthic	0.0%	0.4% (n=6)	0.0%	0.0%	0.0%	0.0%	0.0%	33.3% (n=2)	0.0%	0.0%	0.0%	0.3% (n=1)
Microcalanus spp.	0.0%	0.3% (n=4)	50.0% (n=1)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.9% (n=3)
Oikopleura spp.	11.1% (n=1)	0.1% (n=2)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Oithona spp.	0.0%	10.6% (n=145)	0.0%	0.0%	0.0%	0.0%	0.0%	16.7% (n=1)	0.0%	0.0%	0.0%	2.5% (n=8)
Podon/Pleopsis spp.	0.0%	0.8% (n=11)	0.0%	0.0%	100.0% (n=1)	2.1% (n=1)	0.0%	0.0%	0.0%	0.0%	0.0%	11.9% (n=38)
Polychaeta (larvae)	55.6% (n=5)	0.7% (n=10)	0.0%	0.0%	0.0%	25.5% (n=12)	0.0%	0.0%	0.0%	50.0% (n=2)	0.0%	9.7% (n=31)
Pseudocalanus spp.	0.0%	12.6% (n=172)	0.0%	0.0%	0.0%	0.0%	0.0%	16.7% (n=1)	0.0%	0.0%	0.0%	3.8% (n=12)
Temora spp.	0.0%	11.0% (n=151)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.3% (n=4)

Relative Abundance of Top Taxonomic Instances per Sample