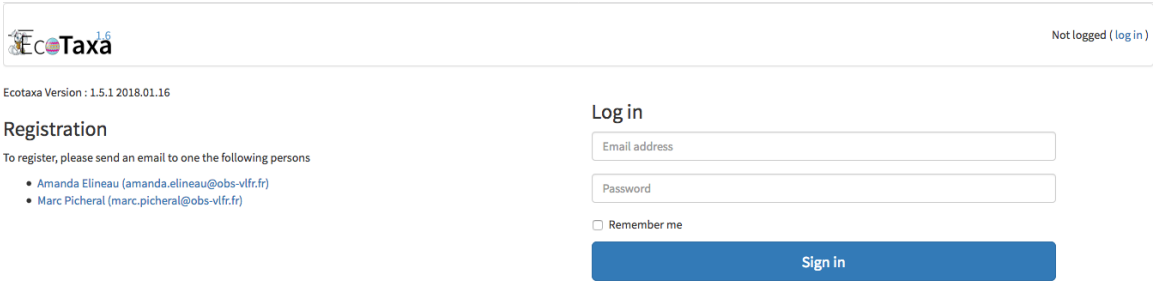
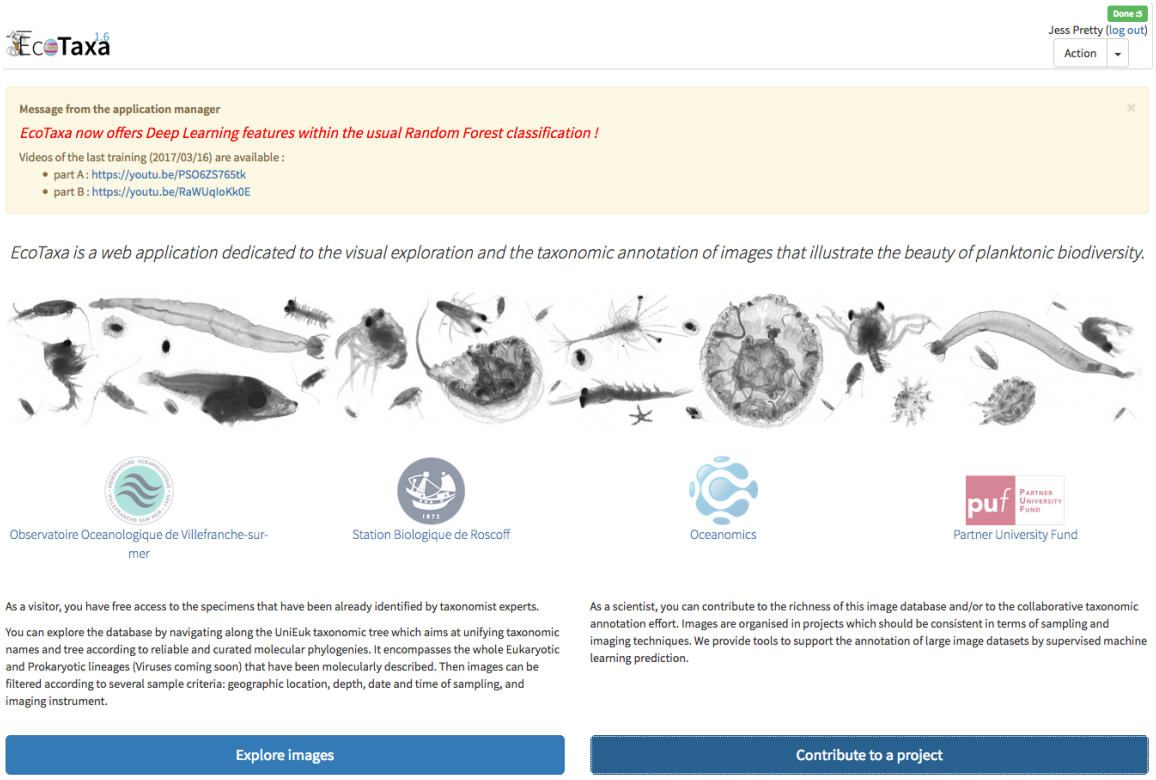
Quick Guide to Image Sorting in Ecotaxa

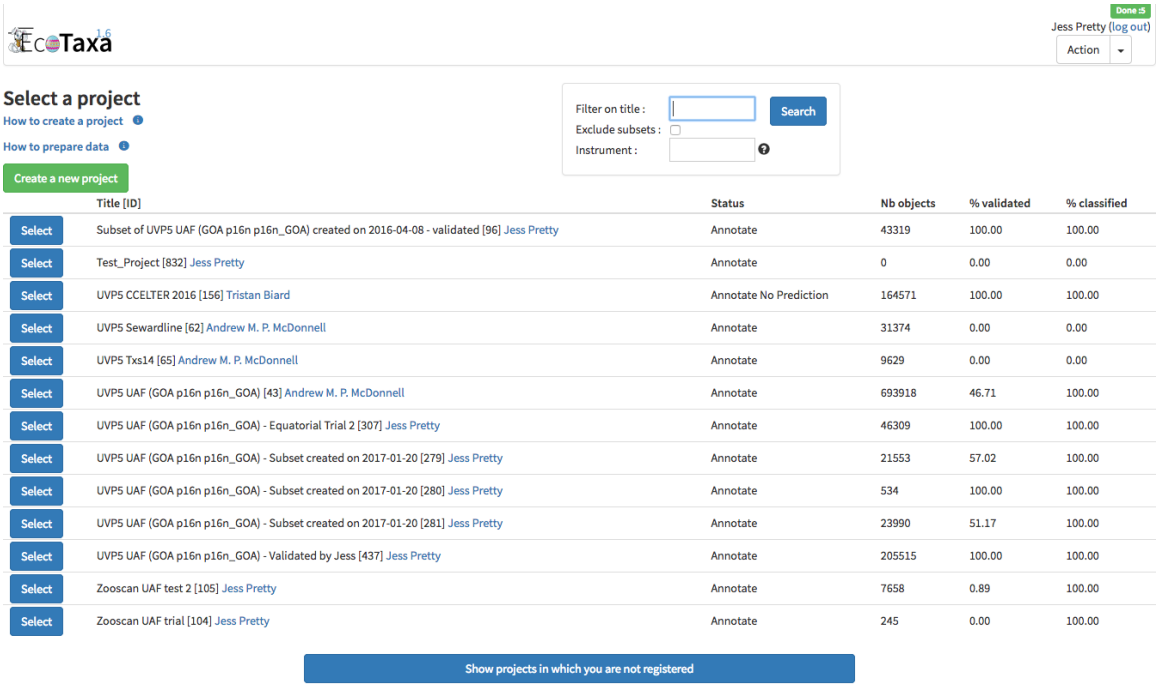
Contact Marc Picheral for permission to access Ecotaxa & create project



Log-in using credentials

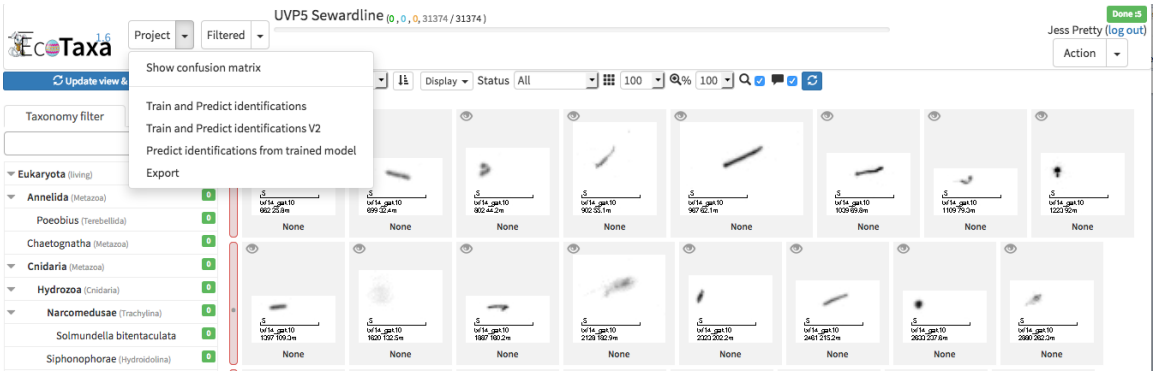


Select Contribute to a project

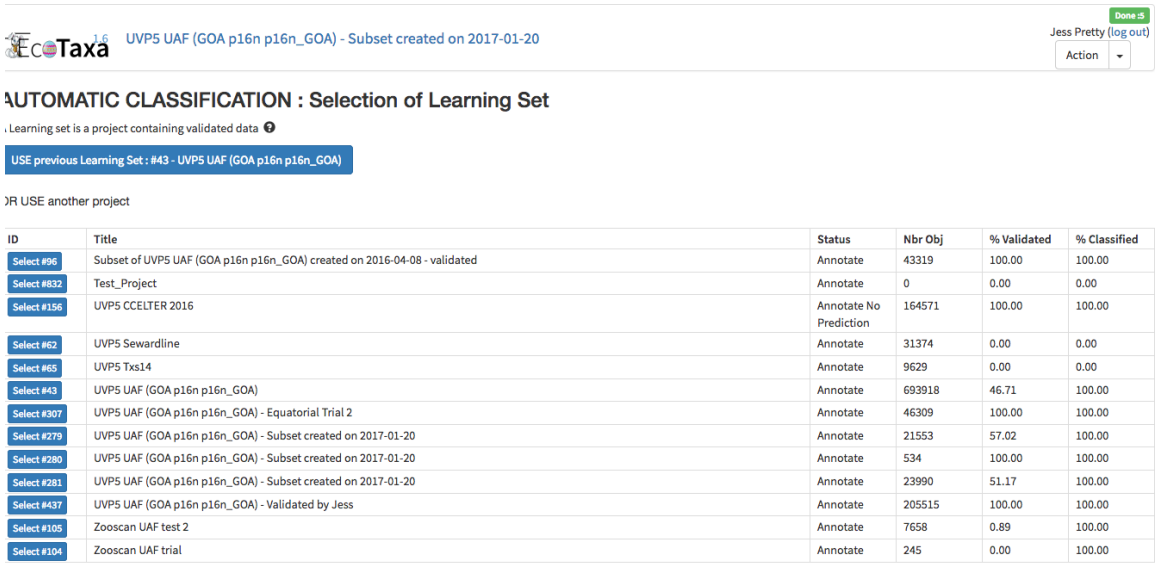


1. Once you are logged into Ecotaxa and have selected ‘Contribute to Project’, select which project you would like to sort images in.

For a new Project, you must first choose to predict the identity of the images using a previously uploaded Learning set. Follow the following steps:

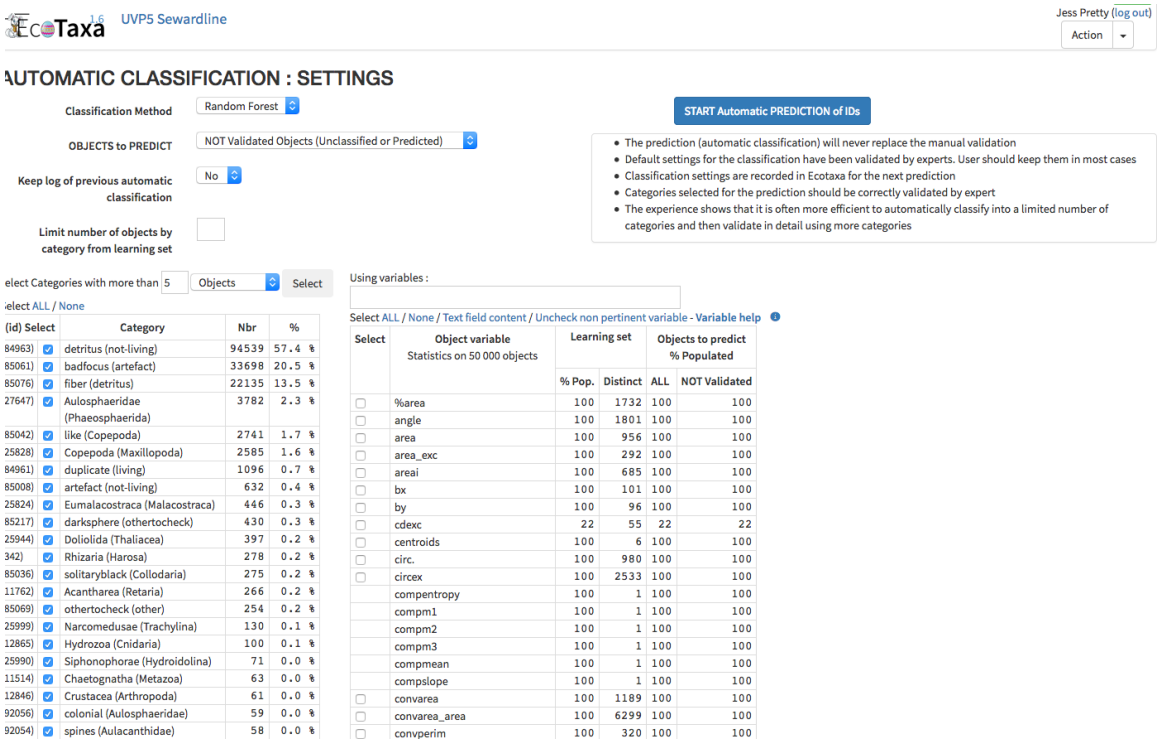


a. Select ‘Train and Predict Identifications’



ID	Title	Status	Nbr Obj	% Validated	% Classified
Select #96	Subset of UVP5 UAF (GOA p16n p16n_GOA) created on 2016-04-08 - validated	Annotate	43319	100.00	100.00
Select #832	Test_Project	Annotate	0	0.00	0.00
Select #156	UVP5 CCELTER 2016	Annotate No Prediction	164571	100.00	100.00
Select #62	UVP5 Sewardline	Annotate	31374	0.00	0.00
Select #65	UVP5 Txs14	Annotate	9629	0.00	0.00
Select #43	UVP5 UAF (GOA p16n p16n_GOA)	Annotate	693918	46.71	100.00
Select #307	UVP5 UAF (GOA p16n p16n_GOA) - Equatorial Trial 2	Annotate	46309	100.00	100.00
Select #279	UVP5 UAF (GOA p16n p16n_GOA) - Subset created on 2017-01-20	Annotate	21553	57.02	100.00
Select #280	UVP5 UAF (GOA p16n p16n_GOA) - Subset created on 2017-01-20	Annotate	534	100.00	100.00
Select #281	UVP5 UAF (GOA p16n p16n_GOA) - Subset created on 2017-01-20	Annotate	23990	51.17	100.00
Select #437	UVP5 UAF (GOA p16n p16n_GOA) - Validated by Jess	Annotate	205515	100.00	100.00
Select #105	Zooscan UAF test 2	Annotate	7658	0.89	100.00
Select #104	Zooscan UAF trial	Annotate	245	0.00	100.00

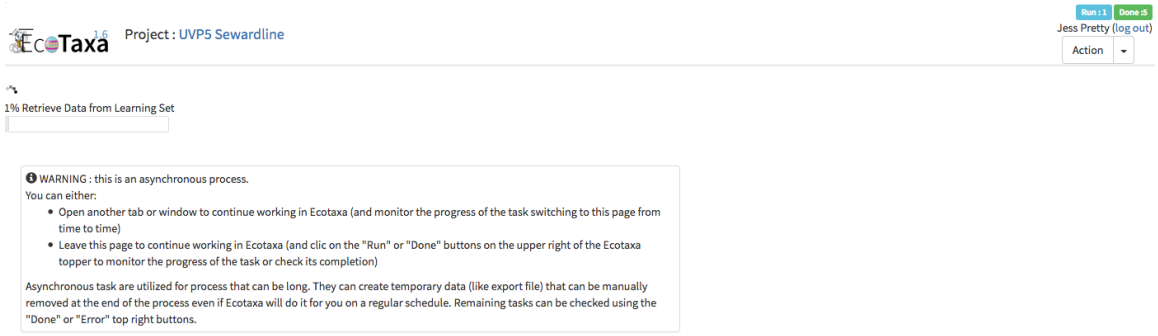
b. Select which project you would like to use to base your predictions on. The best choices are those with the majority of objects sorted, and with similar sorting levels to what you are hoping to achieve.



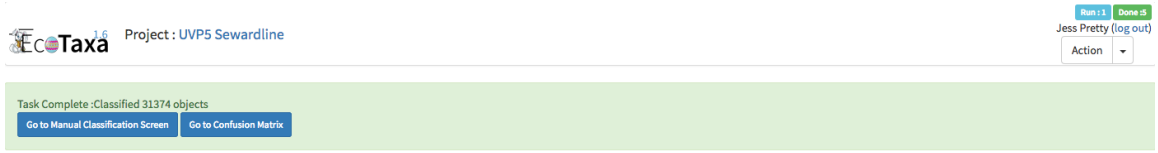
(id) Select	Category	Nbr	%
84963)	detritus (not-living)	94539	57.4 %
85061)	badfocus (artefact)	33698	20.5 %
85076)	fiber (detritus)	22135	13.5 %
27647)	Aulosphaeridae (Phaeosphaerida)	3782	2.3 %
85042)	like (Copepoda)	2741	1.7 %
25828)	Copepoda (Maxillopoda)	2585	1.6 %
84961)	duplicate (living)	1096	0.7 %
85008)	artefact (not-living)	632	0.4 %
25824)	Eumalacostraca (Malacostraca)	446	0.3 %
85217)	darksphere (othertocheck)	430	0.3 %
25944)	Doliolida (Thaliacea)	397	0.2 %
342)	Rhizaria (Harosa)	278	0.2 %
85036)	solitaryblack (Collodaria)	275	0.2 %
11762)	Acantharea (Retaria)	266	0.2 %
85069)	othertocheck (other)	254	0.2 %
25999)	Narcomedusae (Trachylina)	130	0.1 %
12865)	Hydrozoa (Cnidaria)	100	0.1 %
25990)	Siphonophorae (Hydroidolina)	71	0.0 %
11514)	Chaetognatha (Metazoa)	63	0.0 %
12846)	Crustacea (Arthropoda)	61	0.0 %
92056)	colonial (Aulosphaeridae)	59	0.0 %
92054)	spines (Aulacanthidae)	58	0.0 %

Select	Object variable Statistics on 50 000 objects	Learning set		Objects to predict % Populated	
		% Pop.	Distinct	ALL	NOT Validated
	%area	100	1732	100	100
	angle	100	1801	100	100
	area	100	956	100	100
	area_exc	100	292	100	100
	areai	100	685	100	100
	bx	100	101	100	100
	by	100	96	100	100
	cdexc	22	55	22	22
	centroids	100	6	100	100
	circ.	100	980	100	100
	circex	100	2533	100	100
	compentropy	100	1	100	100
	compm1	100	1	100	100
	compm2	100	1	100	100
	compm3	100	1	100	100
	compmean	100	1	100	100
	compslope	100	1	100	100
	convarea	100	1189	100	100
	convarea_area	100	6299	100	100
	convperim	100	320	100	100

c. This part is very trial and error, so choose variables you would like to sort images in (the lower right portion of the window) and press ‘START Automatic PREDICTION of IDs’.

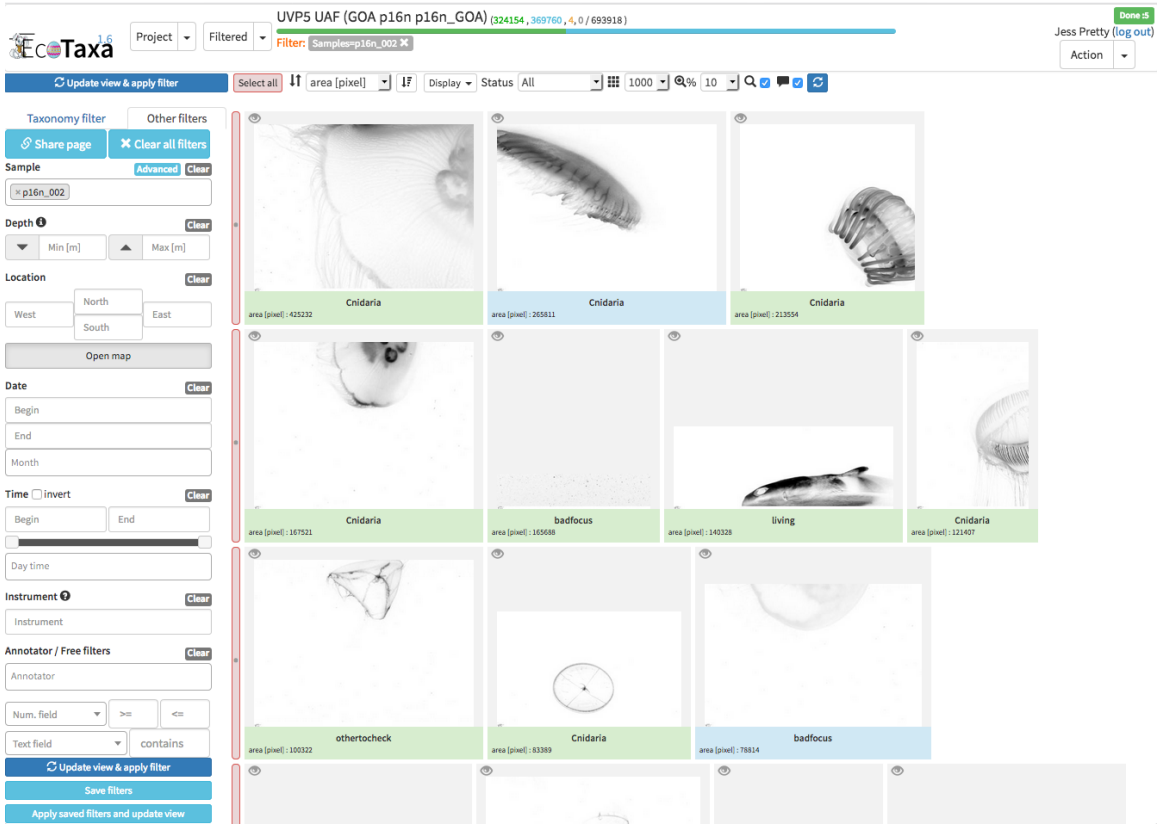


d. This screen pops up while Ecotaxa is predicting your images and shows progress. You can open another window and continue work on another project, but don’t exit out of this window.



e. Once this task is complete, you are ready to start validating and sorting images!

- If you don’t like the sorting, or want to try a different one then repeat steps above and choose different variables or categories. You can do this multiple times throughout the sorting process. Continue to steps below to sort and validate images.



2. Project opens with last filters used, or entire project opens up if this is your first time opening it – in example above only one station of data is selected (p16n_002).

3. Basics of Sorting:

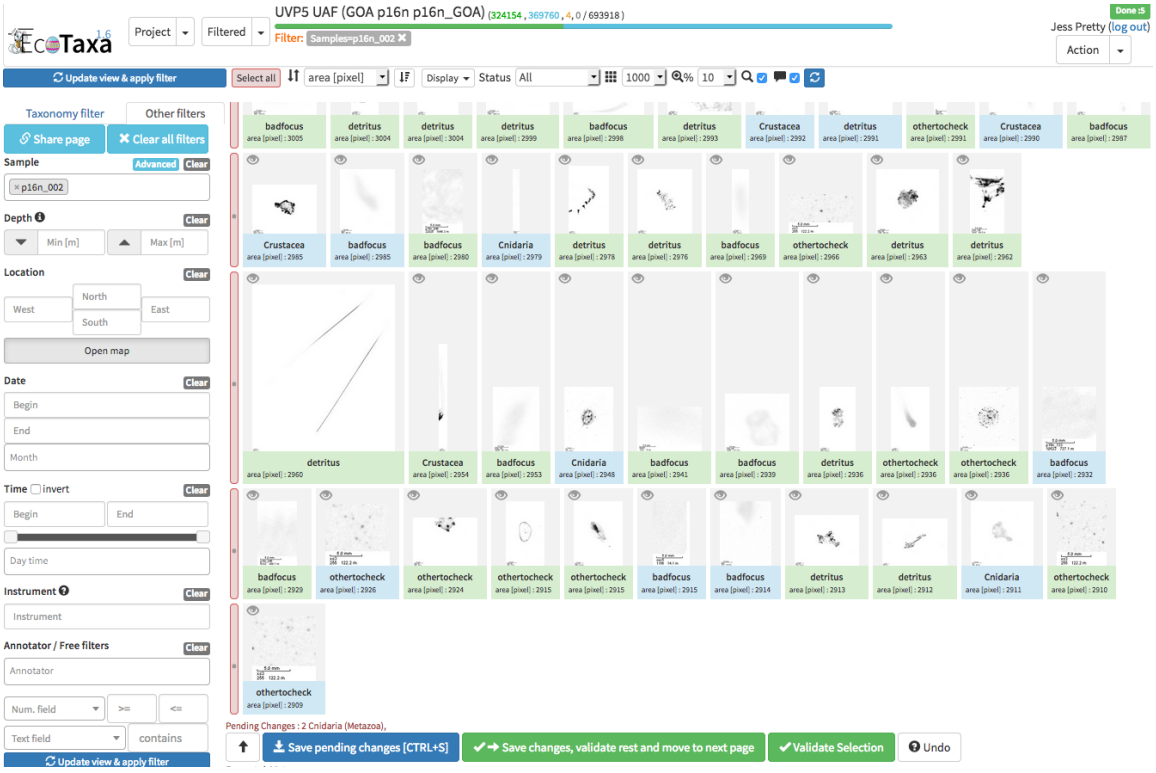
a. Click on image and begin typing name of group you would like to sort into.

To Note: You do not need to click anywhere else, the screen on the

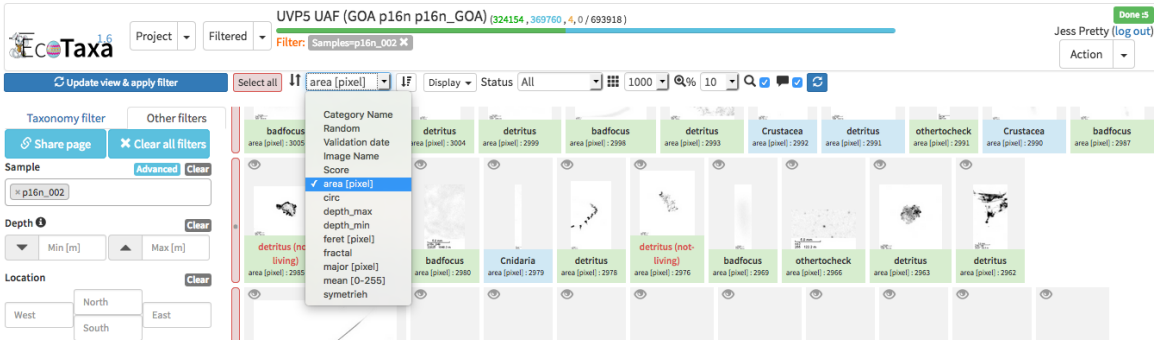
upper left-hand corner will automatically appear. Click ‘Enter’ to accept new assignment for image. The name changes to red when you have changed the assignment.

b. You can select multiple images and sort them into the same category by clicking on multiple images – the selected images will be highlighted in red. If you would like to sort them into the same category as was just used, press ‘Command + d’ (or ‘Control + d’ for PC).

c. Continue sorting as many images as desired, changing those that you would like to/need to.

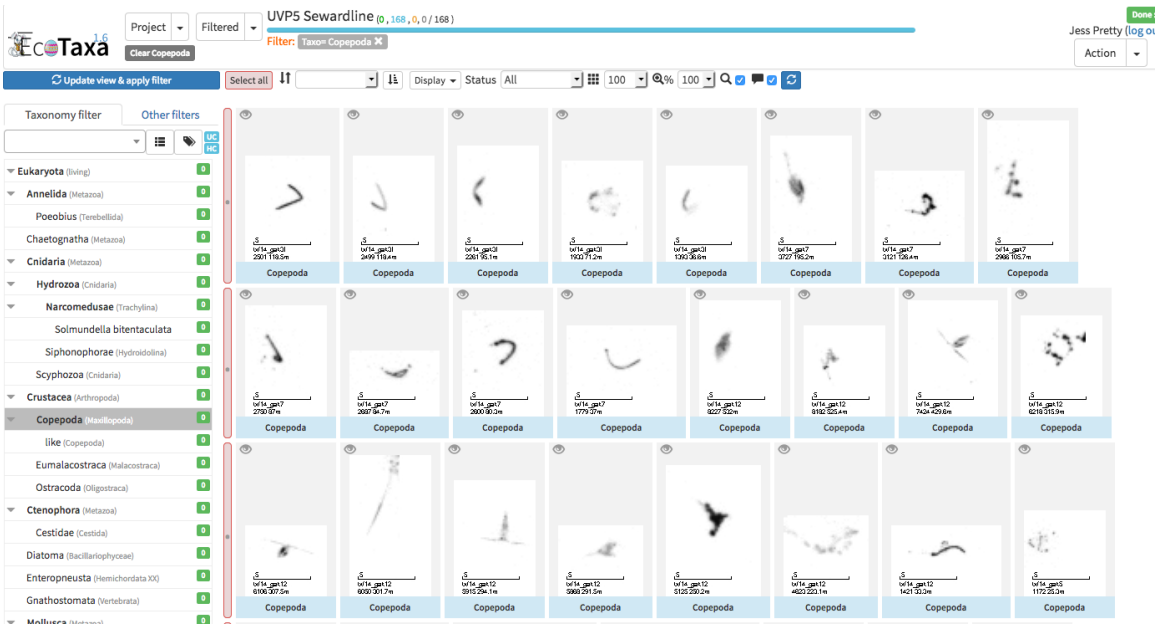


d. When finished with this page, either ‘Save pending changes’ which will save all changes made and validate those changes, or ‘Save changes, validate rest and move to next page’. Be aware of which one you choose to do!



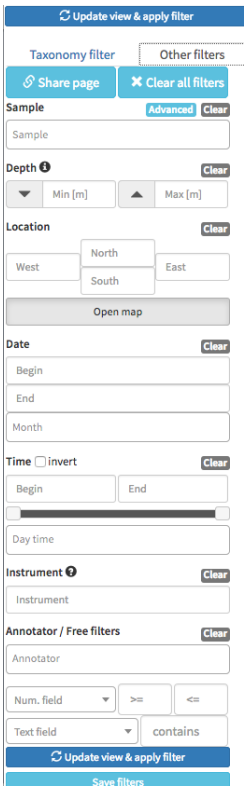
e. You can choose how to sort images in the dropdown menu showing ‘area [pixel]’. There are many options here that are useful for different things, so try them out.

f. You can select to zoom in on images (beside the magnifying glass) or zoom out as you see fit, as well as select how many images will be shown per page (shown as ‘1000’ in the above figure.



g. You may choose to only sort one type of predicted image – for example under the taxonomy tab, I chose to only look at images predicted to be copepods in the example above.

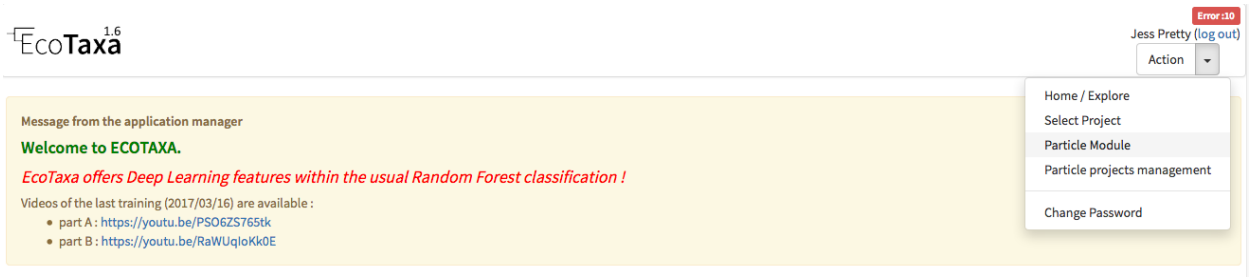
h. You may choose to only look at certain stations, regions (based on lat/lon) or depths using the 'Other Filters' tabs. Remember to select 'Update and apply filter' to apply this to your viewed images.



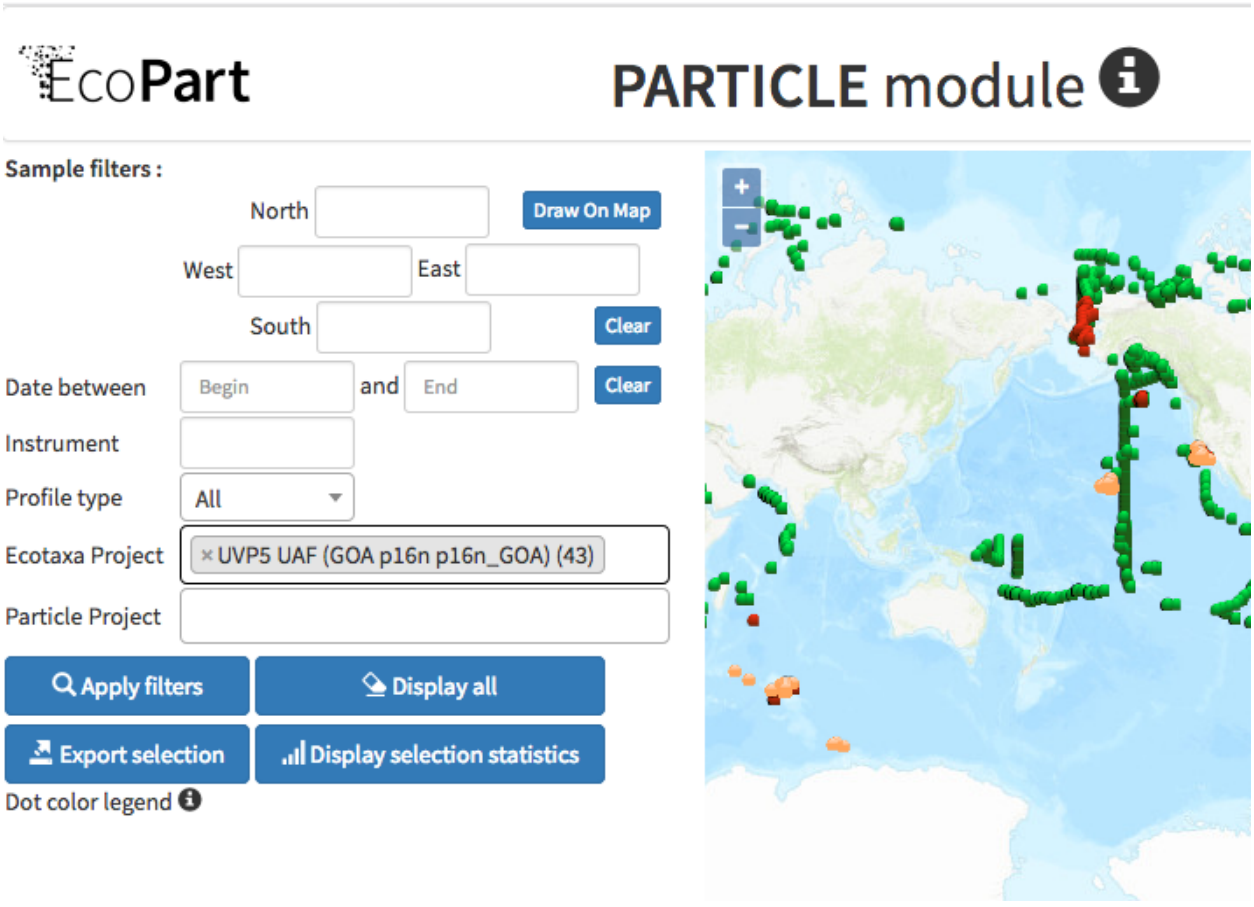
i. All filters applied will appear at the top of the screen.

After image sorting is finished, your data is ready to be exported and analyzed.

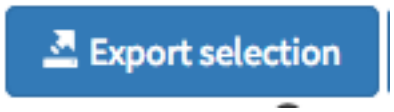
4. Exporting Data from Ecotaxa



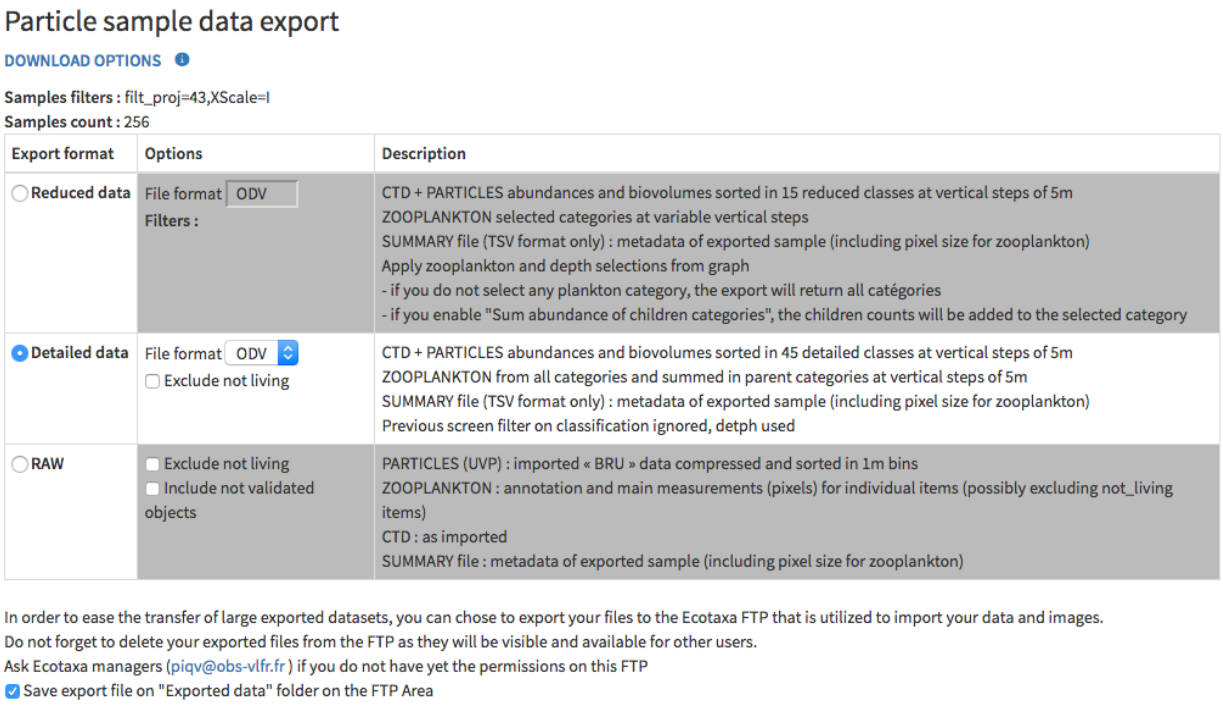
a. Choose particle module from the home screen, drop down menu



b. Select the project you would like zooplankton data from under ‘Ecotaxa Project’ & Press ‘Apply Filters’ Button



c. Select any extra filters you may want to constrain the data by, and then press ‘Export Selection’



Particle sample data export

DOWNLOAD OPTIONS

Samples filters : filt_proj=43,XScale=I
Samples count : 256

Export format	Options	Description
○ **Reduced data**	File format ODV **Filters :**	CTD + PARTICLES abundances and biovolumes sorted in 15 reduced classes at vertical steps of 5m ZOOPLANKTON selected categories at variable vertical steps SUMMARY file (TSV format only) : metadata of exported sample (including pixel size for zooplankton) Apply zooplankton and depth selections from graph - if you do not select any plankton category, the export will return all catégories - if you enable "Sum abundance of children categories", the children counts will be added to the selected category
● **Detailed data**	File format ODV ☐ Exclude not living	CTD + PARTICLES abundances and biovolumes sorted in 45 detailed classes at vertical steps of 5m ZOOPLANKTON from all categories and summed in parent categories at vertical steps of 5m SUMMARY file (TSV format only) : metadata of exported sample (including pixel size for zooplankton) Previous screen filter on classification ignored, detph used
○ **RAW**	☐ Exclude not living ☐ Include not validated objects	PARTICLES (UVP) : imported « BRU » data compressed and sorted in 1m bins ZOOPLANKTON : annotation and main measurements (pixels) for individual items (possibly excluding not_living items) CTD : as imported SUMMARY file : metadata of exported sample (including pixel size for zooplankton)

In order to ease the transfer of large exported datasets, you can chose to export your files to the Ecotaxa FTP that is utilized to import your data and images.
Do not forget to delete your exported files from the FTP as they will be visible and available for other users.
Ask Ecotaxa managers (piqv@obs-vlfr.fr) if you do not have yet the permissions on this FTP
☑ Save export file on "Exported data" folder on the FTP Area

d. Download data in ODV or TSV format by either direct download locally, or export to FTP site (suggested for large files)

e. If you’d like total zooplankton abundance, or any parameter that is not included in the download, you can edit the tsv/odv file in Excel to add parameters