Thomas Rutten Projects

Manual EasyClus© v1.30

EasyClus© software for flow cytometric data analysis
working in a Matlab® environment
to perform overall data analysis
and manual or automatic clustering,
automatic selection sets processing and
classification of clusters on basis of a database.

by
Thomas Rutten
December 2019

TRP 2019.029

Content

1. Introduction

This manual describes the software application EasyClus v1.30, which is version 1.30 (Thomas Rutten Projects 2019) for the automated processing of data derived from flow cytometry and especially CytoSense. Version 1.30 has only been released to one customer with a custom specific option. The software is written within the Matlab® environment. Several upgrades have been made since 2009 leading to this latest version 1.30 EasyClus©.
The basic version of MATLAB® (2018a and newer) is required to run the application. This will be checked in the next months. EasyClus© runs in principle in Windows, OS2 (Apple) and Linux, because Matlab and thus EasyClus can run under these systems. However, the CytoBuoy software only runs under Windows, which is necessary to open their .cyz files.
EasyClus provides a lot of data processing possibilities of particular phytoplankton data ranges from manual to automated clustering, storing results tables and Cytogrammes(scatter plots) and the use of a database for the identification of phytoplankton species.
The software application consists of smaller software programs (algorithms) that are invoked through a menu selection window (EasyClus). Specific knowledge of Matlab® is not required. This user manual describes how to install and to use the EasyClus software.

Some specifications:

- ✓ There are three unsupervised and three supervised clustering methods and one hybrid clustering method.
  Unsupervised: uses 'auto clustering GO' and 'auto clustering DESIGN 1 or 2', which puts events together into clusters on basis of the offered data and used clustering settings, without any prior knowledge.
  Supervised: uses a-priori knowledge by LASSO or RULES or FIX grids. Data file formats that can be read are listmode .cyz, csv-files (CytoSense output), or the FCS 2.0 and 3.0 formats, *. mat format.
  Hybrid model to create selections sets on basis of unsupervised clusters.
  Manual clustering: uses manual selections, and selections can be compared to database stored species.
- ✓ A database can easily be created for recognition of species
- ✓ Unique clustering principles are used to recognize which events belong to each other and which events do not.
- ✓ Clustering and recognition of species by the database is performed fast within seconds.
- ✓ More files can be processed automatically

There are also other generic information modules provided:

- EasyClus LIVE for the full automatic real time online flowcytometric analysis, data analysis and reportage of interesting facts and figures to a website (only available with a LIVE license).
- New modules (images to clusters) will be released in a future version soon.

1.1 What's NEW in EasyClus v1.30

EasyClus is improved every day and new features are added:

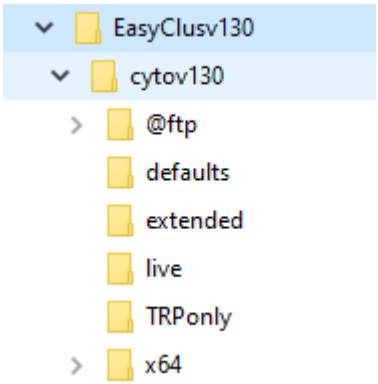
- ✓ NEW: Several new options to build a database easily including images
- ✓ NEW: Fst method to build a file temporary database including images
- ✓ NEW: (Supervised) clustering directly from saved phytodatabase.mat
- ✓ NEW: Fast option to use CytoSense image cropping
- ✓ NEW: Building a file from several files including profiles and images
- ✓ NEW: Several options to build a database including images
- ✓ NEW: Import data option to ignore specific non informative data headers
- ✓ NEW: Several options to build a database including images
- ✓ NEW: Tool to count single cells from particles by using average single cell values
- ✓ NEW: Improved EasyClus X-tool for fast analysis of data
- ✓ NEW: Google maps – plotting needs personal Api Key (new policy Google) to get maps. ApiKey input menu added.
- ✓ NEW: EasyClus can be used in IOS (Apple) and Linux (after converting files to .mat
- ✓ EasyClus has been made suitable for latest CytoSense software
- ✓ EasyClus has been made suitable for latest Matlab version

2. Installation/ upgrading of EasyClus

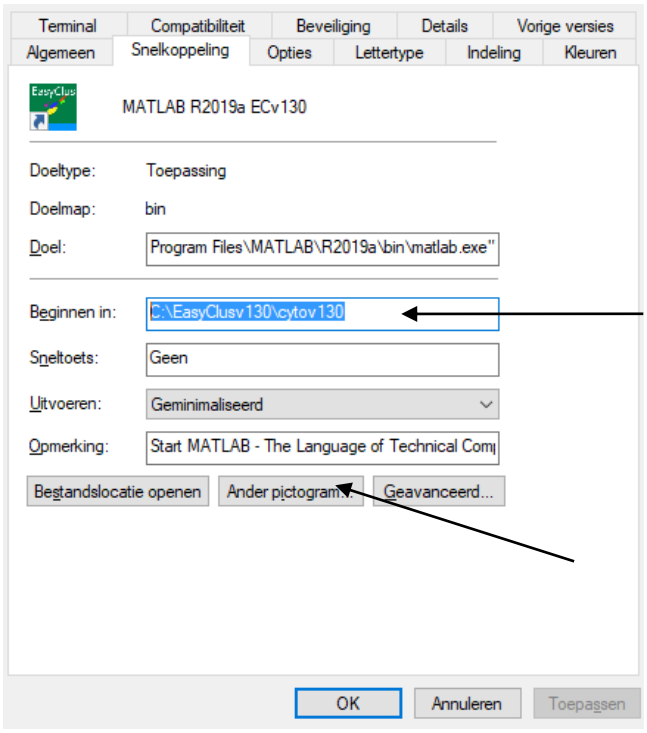
2.1 Installation

Note 1: Make sure that you have Matlab (2018a or higher) working on your computer.
Note 2: MATLAB uses a dot as decimal separator stabbing. This should be changed by 'Country settings' of the Configuration menu.

Step 1:
Copy all (unzipped) new version EasyClus files including directories to a new directory on your computer.



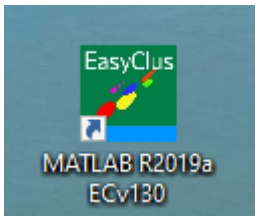
Step 2:
Make a copy of existing Matlab shortcut to your desk screen.
Change the properties of the newly created (Matlab) shortcut (see figure).
Click on the Desktop "shortcut matlab.exe 'icon with the right mouse button. Click the right mouse button on the menu 'Properties'.
Change the text under "Start in" in the directory where the Matlab application EasyClus algorithms are available(in this example in C:\EasyClusv130\cytov130.
In this case, it is assumed that the flow cytometry EasyClus algorithms are placed in the in C:\EasyClusv130\cytov130 folder.



Change directory in short cut (Start in:) to directory where new version EasyClus is located

not necessary, but nicer: Change icon in EasyClus icon (in this example in EasyClus128v1\cytov128 directory)

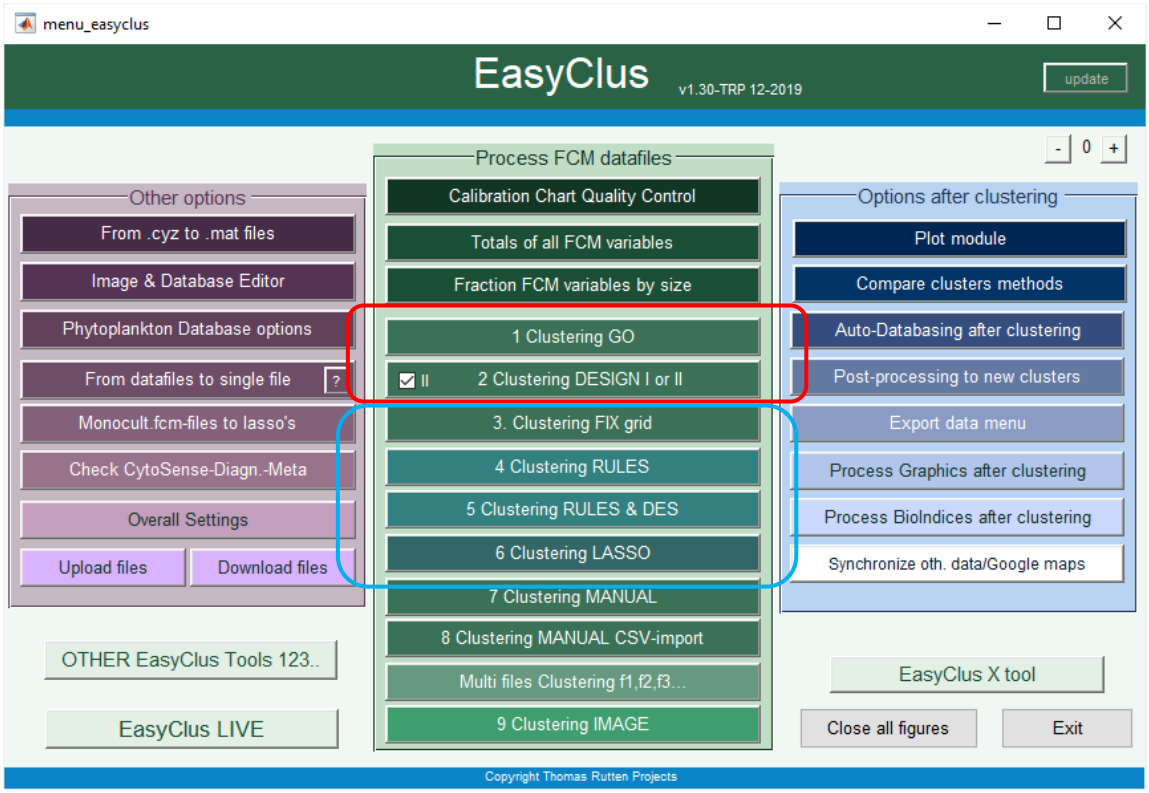
Step 3:
That's it. Double click your freshly made icon



* (If no menu appears EasyClus can always be started by typing 'easyclus (enter) in the main Matlab window.

3. The application EasyClus

After starting up Matlab®, the EasyClus menu appears on the screen. Processing commands are executed using the buttons.



= Unsupervised clustering (no trained data set)

= Supervised clustering (no trained data set)

<u>3. Unsupervised clustering methods</u>

1. GO-method : PCA based
2. DESIGN1 : uses scatterplots & density & all FCM vars
3. DESIGN2 : uses scatterplots & density & some FCM vars

<u>4. Supervised clustering methods</u>

1. LASSO : using automated produced selection sets
2. RULES : using rules that characterizes each unique cluster
3. FIX : using own virtual FCM species by preset barriers

13	14	15	16
9	10	11	12
5	6	7	8
1	2	3	4

4. Hybrid (DES2+RULES) : unsupervised as well as supervised

There are a lot of 'tricks' behind the buttons. At the right, in the blue window, there are post processing options: e.g. the plotting of figures, clustering options, which means that you can

further process your data after clustering of FCM datafiles. E.g. the matching of clusters with another database, or the making of a new database with all unique clusters found in a dataset or graphical representations on basis of clustering results.

3.1 In general

The approach of FCM data-analysis is based on three stages (actually even more, but that's for later..)

3.1.1. Exploration of the instrumental performance during analyses
3.1.2. Exploration of field data – first impression
3.1.3 Detailed data analysis of field data

3.1.1. Exploration of instrumental performance during analyses

Don't blindly trust your instrument. For good laboratory practice, even a thermometer has to be checked on a regular basis. So, check the instrumental performance of your flowcytometer too by using stable (= keep their optical properties similarly during at least half a year) calibration beads regularly e.g. at the beginning, in the middle and at the end of an analyses day. Check the performance of your instrument by using the CHECK-DIAGNOSIS button (purple field) and the CALIBRATION CHART (first button in green window) module. Beads that are found easily by specific fluorescence properties are visualized and should lay in between ascertained levels in so called calibration charts.

3.1.2. Exploration of field data – first impression

You have a lot of data e.g. sampled at one location during several months or collected during a field cruise.
Start with getting a fast impression of your data by having an overall analysis of all data. Use TOTALS OF ALL FCM PARAMETERS (button 2) and/or FRACTIONATE FCM PARAMETERS OVER SIZE (button 3) to get a fast answer to some questions like:

- How diverse are my samples?
- Is there relation of my flowcytometric data towards other overall techniques such as chlorophyll-a fluorescence, total phytoplankton counts per ml or other parameters?
- Which samples show obvious behaviour with respect to sudden increase or decrease of specific parameters or within a specific fraction towards other samples?

3.1.3 Detailed data analysis of field data

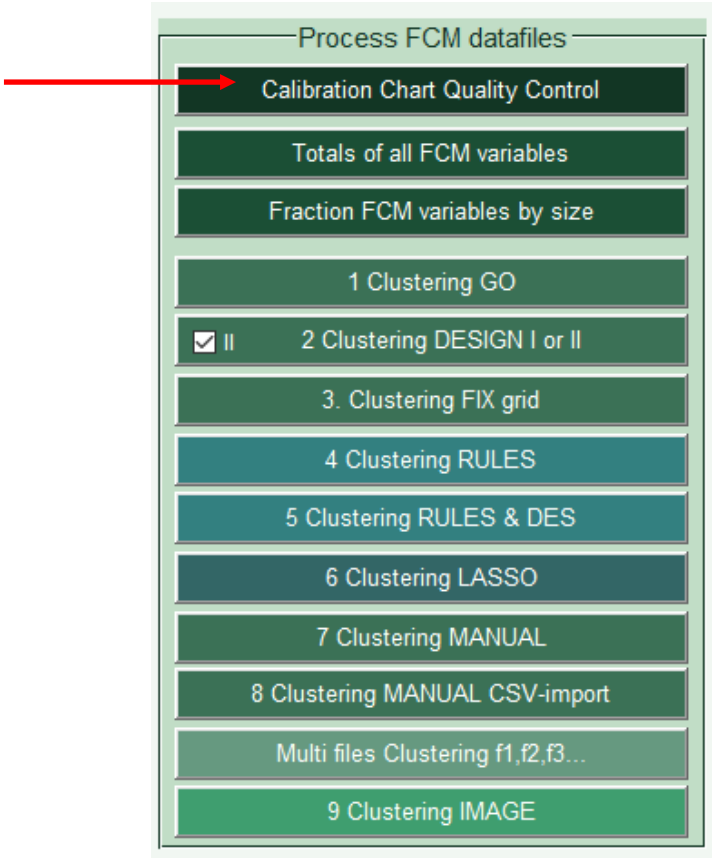
Ever since you have more feeling with your data, you can start with the automatic clustering techniques to find out if there are specific species or species groups that can be discriminated in the samples, but be aware that a flowcytometer cannot distinguish ALL types of species on basis of the available optical variables (taxonomists sometimes need an electronic microscope to distinguish between species).

Select some specific datafiles to start the clustering and to start to fill your database with found clusters or species. When clusters are processed (whether after manual or

after automatic clustering), fingerprints, i.e. kind of mean values of all particles in each cluster of each of the offered flowcytometric variables, are calculated and stored. These values are compared to a database. An adjustable number of FCM variables per cluster are matched with stored fingerprints of species data in the database. The maximum is the number of FCM variables from a FCM data file. The more FCM variables should match, the more critical requirements for matching are taken in consideration.
The degree of 'matching' is expressed in a similarity factor. The minimum value of the similarity factor is adjustable. The lower this value (<1000), the less stringent requirements are used for similarity.
The number of found clusters is visualized in scatter plots. The number of recognized species from the database is also visualized in scatter plots.

Start clustering with all datafiles and visualize the clustering results by the PROCESS SAMPLE INFORMATION AFTER CLUSTERING button.
All results are saved as txt files and / or jpg files in the folders or \ cluster \ cluster \ figure.
After this you can use the autodatabase-option for even more precise cluster definition.

3.2 EasyClus BEADS – Calibration charts

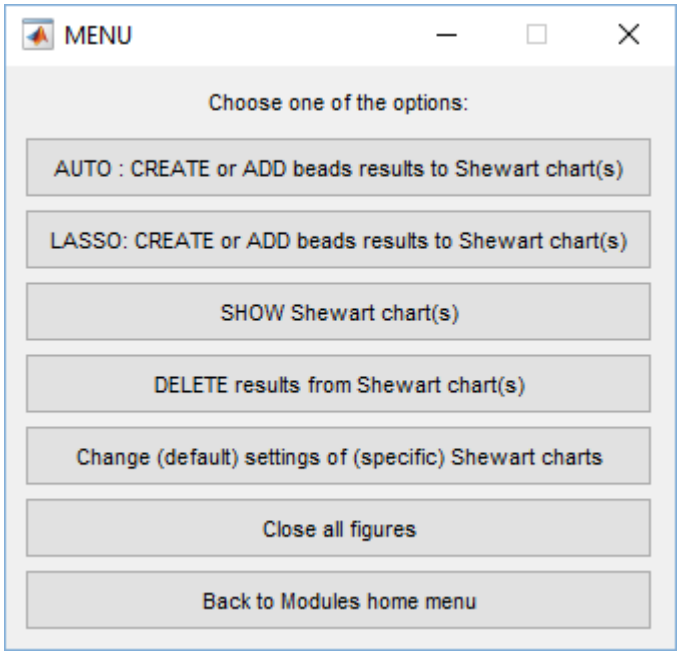


The calibration charts module enables the user to process calibration charts of (robust) references such as beads to validate the short as well as the long term stability of your instrument, which is usually required in laboratories working in a ISO 17025 environment.
In general when an operator establishes a new analysis method, the performance characteristics of the analysis method are studied and registered such as the reproducibility and repeatability. The range of distribution of a method provides us a so called warning limit and action limits, which are used to check if the method is valid and ready for analysis. Beads that are stable over a long period, that do not stick, that have a small size distribution and that fluoresce in all optical channels are usually used as a reference standard. Their average values should fit within the predefined warning and/or action limits before further analysis can take place.

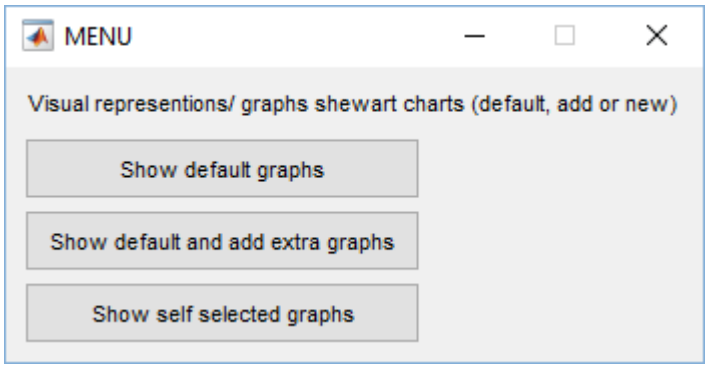
It is strongly recommended to validate your instrument every analysis day at least before and at the end of the day and to record the results in a calibration chart during a long period.

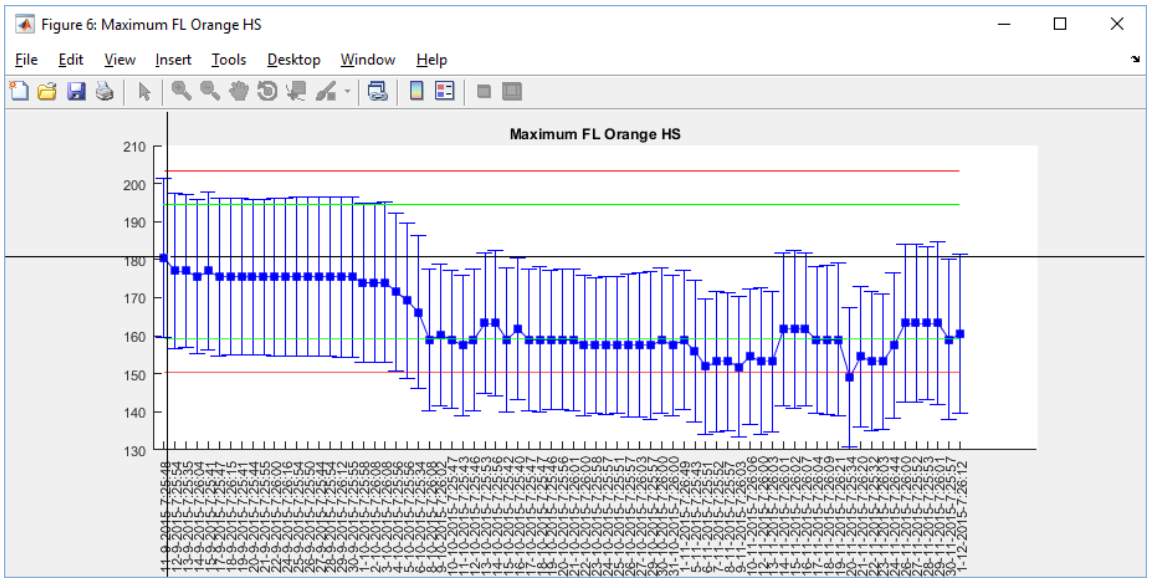
This module helps you to:
- isolate the reference beads,
- determine their average fingerprint in each FCM parameter,
- draw fingerprint results in a calibration chart and give the graphical representation of your chart including their allowed limits

There are two methods used for the recognition of beads. The AUTOMATIC mode (unsupervised) and the PREDEFINED mode (lasso) method. The charts show you which of the FCM parameters are within or outside the predefined limits at once.

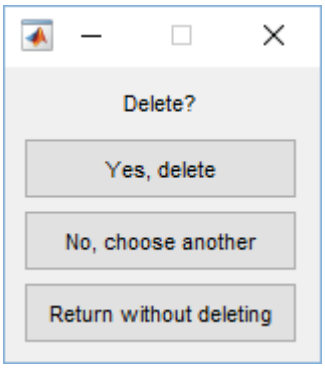


1) The AUTOMATIC mode uses unsupervised clustering and each cluster is compared to common bead characteristics to recognize possible beads in your sample. The smallest sized cluster with beads chosen in case of more than one cluster of beads is found.
2) The PREDEFINED mode uses selections sets or lasso's for clustering and finding beads in a sample. These lasso's should be processed before e.g. by the autoprocessing of lasso's on basis of other clustering methods .
3) SHOW Shewart charts is what is says. You can visualize the beads results by selecting the variables from stored calibration charts.
4) DELETE results from Shewart charts gives you the possibility to remove data – files from a Shewart chart. After selecting a Shewart chart calibration file, you choose graphs to be used to visualize the data AND to remove certain data by pressing any key (to activate the figure) and to pointer a data-event, which should be removed.

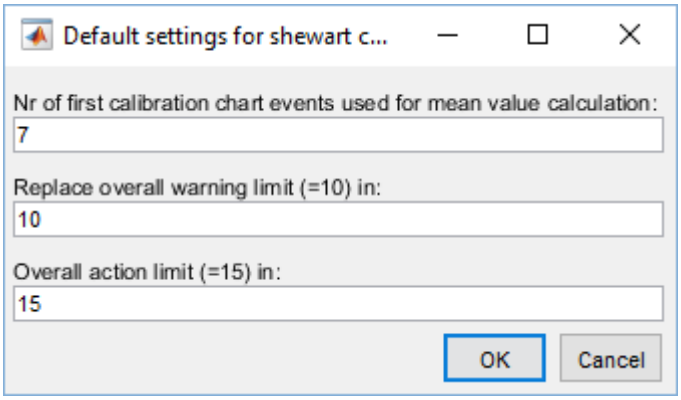




After choosing a specific event, following menu appears in order to confirm the removal process .



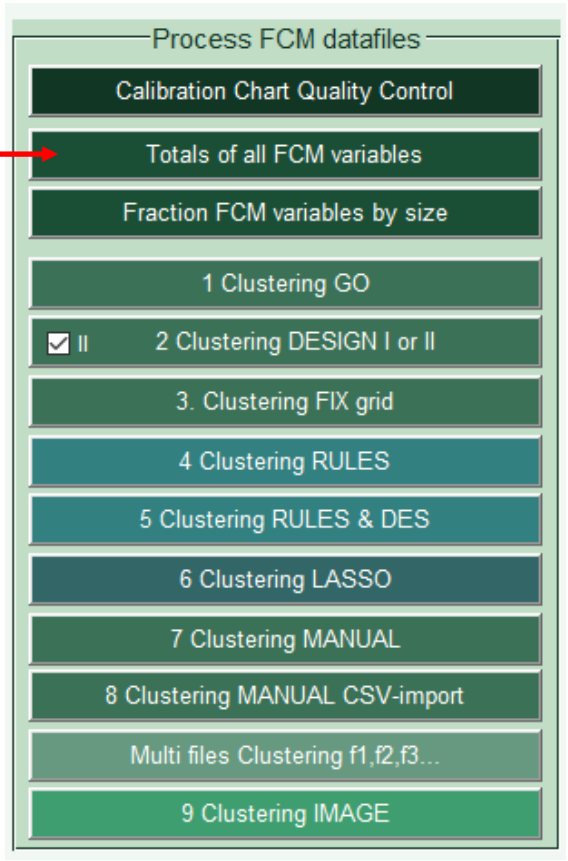
5) Change settings of Shewart charts. Settings of the lasso-method can be changed here. The green warning limit is set to 10%, the red action limit to 15% of the mean of the first 7 files. Recalculation with other numbers can be done here.

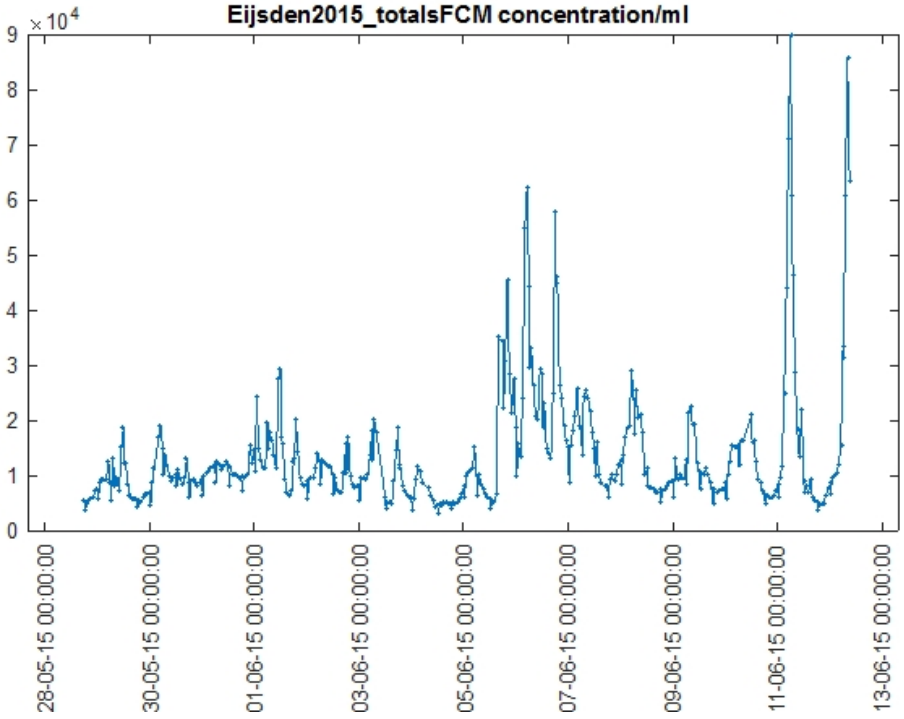


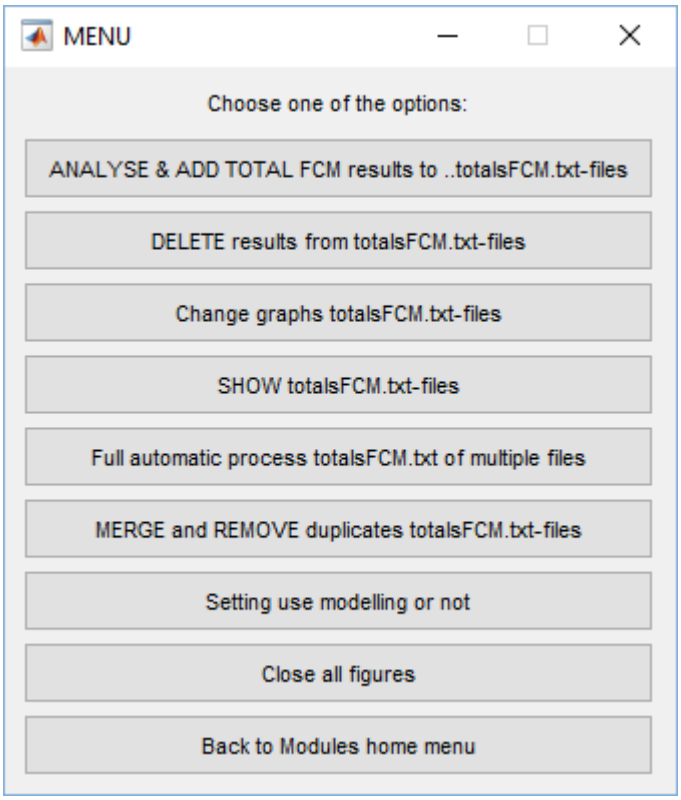
There is also the settings to visualize more or other graphs as default.

It is obvious that the system is stable if the longer term beads averages fluctuate inside the green and red limit lines. If it is outside during more measurements, there might be something wrong (air, fouling, other PMT levels, old-wrong beads, alignment).

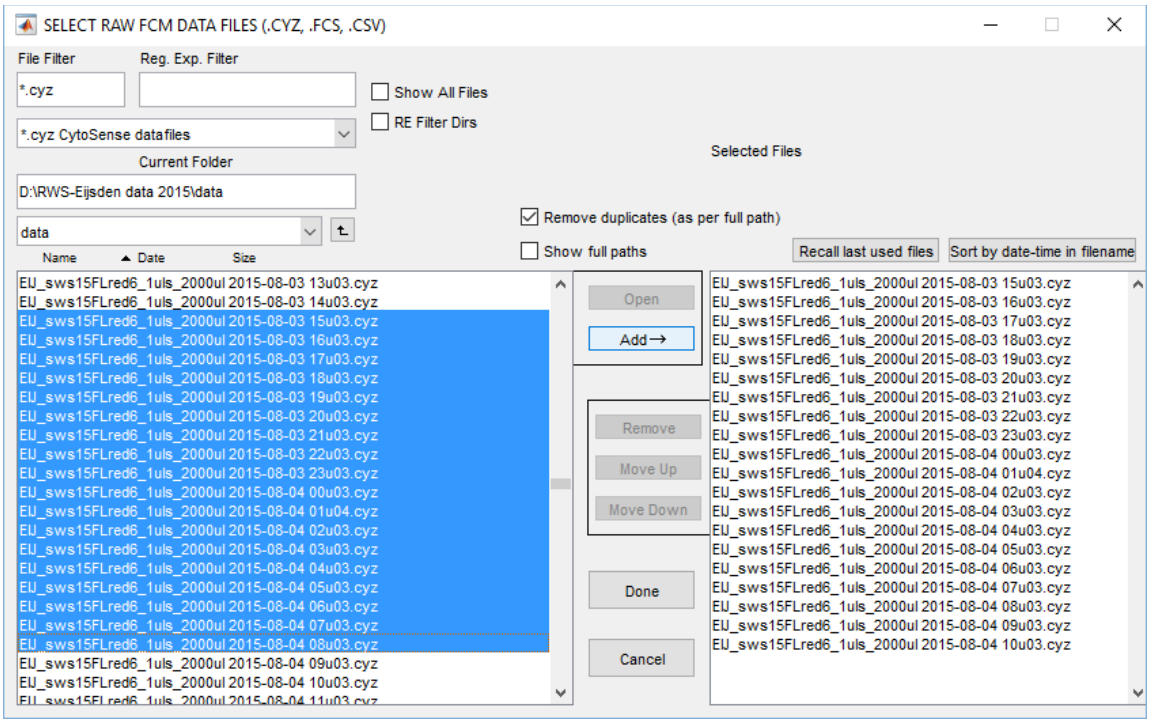
3.3 EasyClus Sample(s) Totals – concentrations, chlorofyl/ml



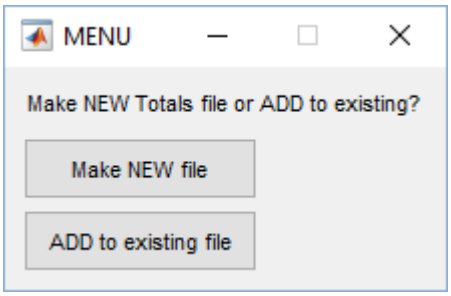




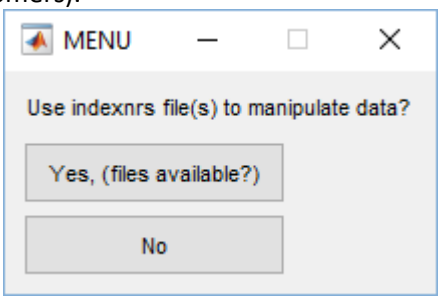
Ignore the 'ANALYSE & ADD' button, because it has been replaced by the 'MULTI ANALYSE & ADD' button. This MULTI etc. is usually used to create a new totalsdata file or to add files to an existing totalsdata file. The advantage is that you can use more than one file.
In FULL automatic process - mode you can start a new file or add it to an existing one. In case of a new file you start with selecting a (cyz)-FCM files.



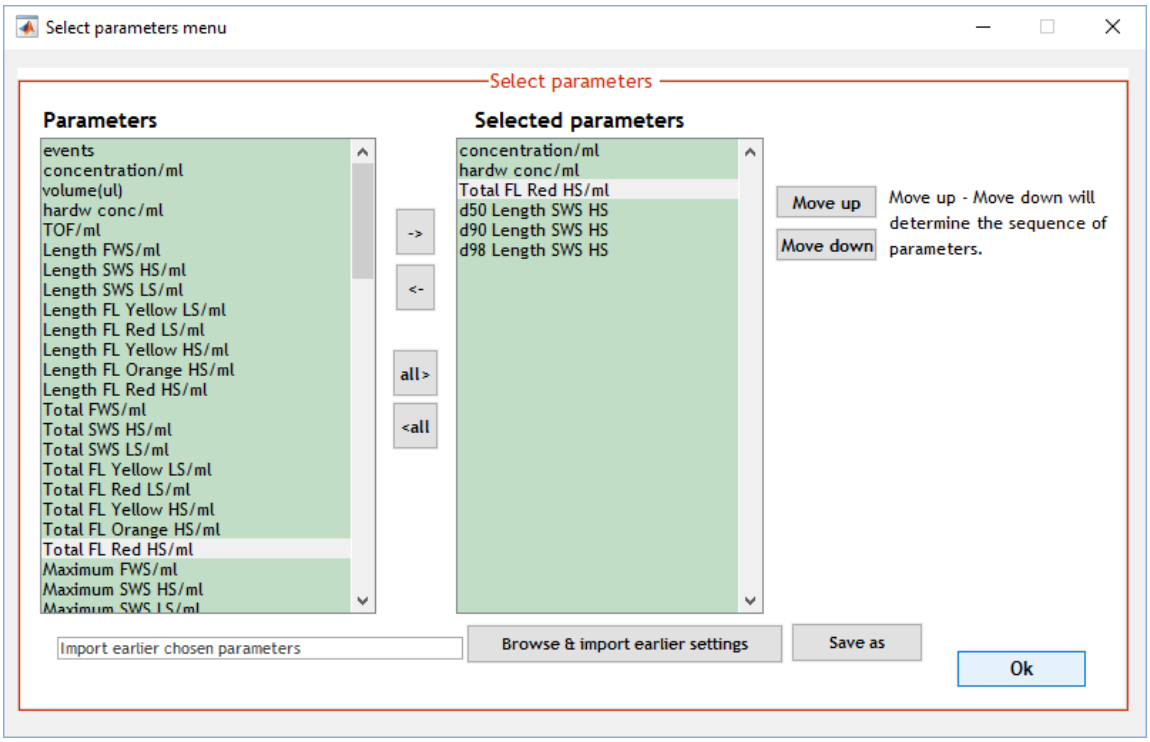
Then choose to start a new file OR add these files to an existing totalsFCMdata-file. We choose the NEW button here.



Choose here 'No' in the menu below.
(In some cases it might be useful to exclude certain events such as beads. For this, you need to cluster files before. In future releases, this part will be worked out in more detail together with customers).

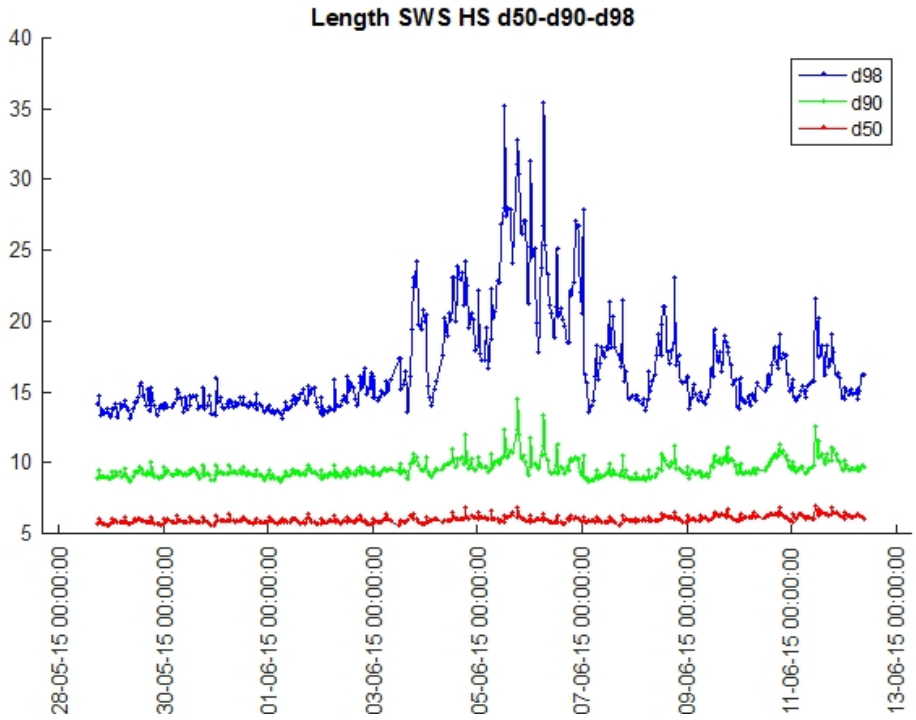


The last question is which parameters should be drawn in figures as a default:

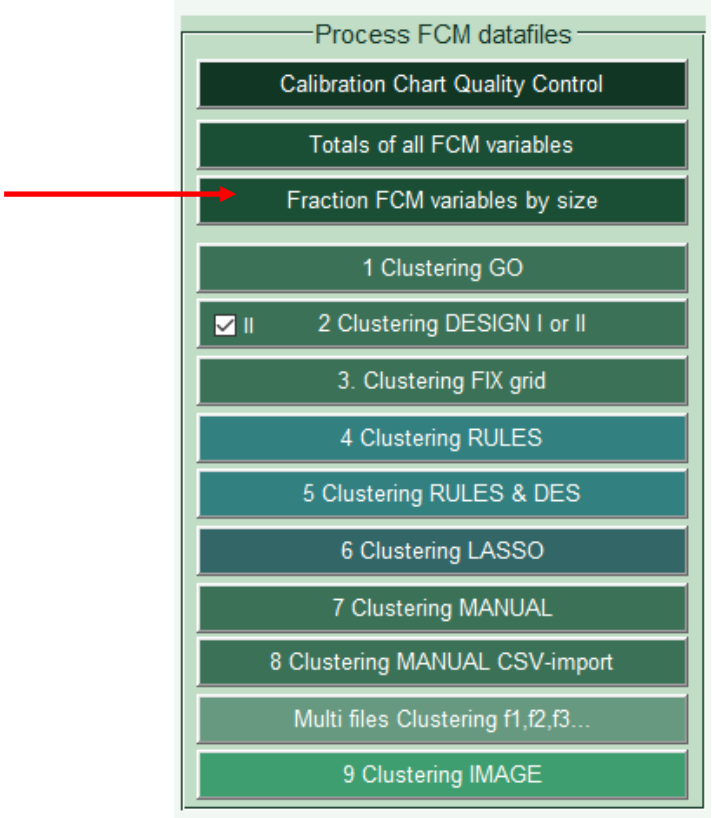


After this, running of all files is started. If you have more than 1000 files, it can take a while (more than minutes, dependent of the used computer).

An example of an output is given in the figure below. The d50 (=median) is the median size of the measured particles using all set triggers. The d90 is the size, where 90% of all particles is smaller than the given size at d90 and 10% is bigger than this value, d98 is the size, where 98% of all particles is smaller than the d98 size and only 2% is bigger than this value. These sizes might be characteristic of the particles at this station.



3.3 EasyClus Sample(s) Fractions – concentrations, chlorophyll/ml over size fractions



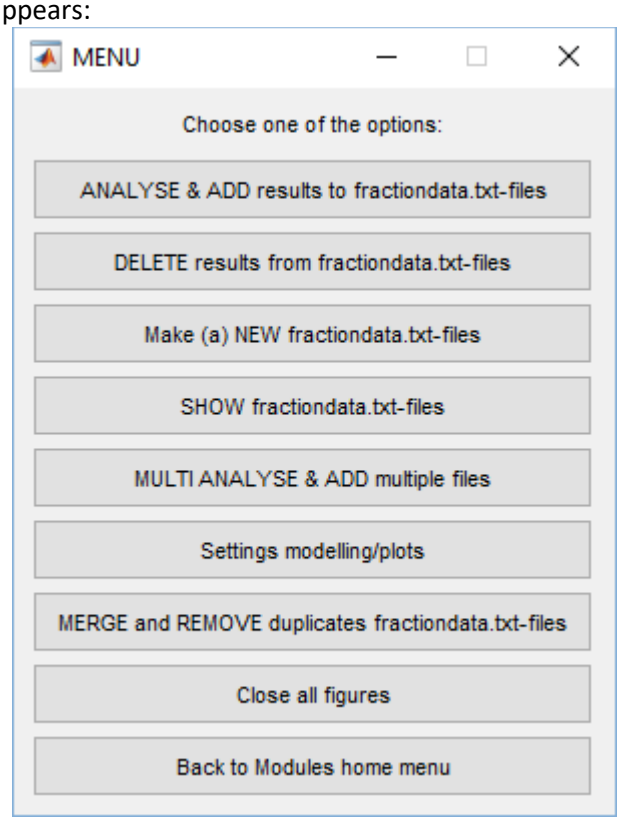
This module enables the user of flowcytometric data to produce general overviews of typical ecological characteristics. The contribution of all species distributed in several size fractions as a function of the chlorophyll-a fluorescence (biomass) or abundance is given. The results are represented in cumulated stacked histograms. Each size fraction represents its own color and results are stored as a .txt or .mat file. The gradient of e.g. the chlorophyll-a concentration (per size fraction) changing in time or location gives a nice overview how samples mutually differ.
In principle all FCM parameters that are available can be processed as a function of size fractions. The number or sizes of the size fractions can also be chosen by the operator.

This module gives you some dynamic general information on the microorganism dynamics, how size fractions contribute to the total process or how they change. The species specific information is not given.

The size of particles is based on the input variables given by the flowcytometer. The parameter that has a (linear) relation with size is given to the software. A calibration line or information about the relation between the calculated and the real size is usually needed and given to the software too. In the case of CytoSense usually the 'Length FWS or 'Length SWS' parameter is used.
The operator should be aware that the FCM-size used for the size fractions in this module, is the size of a particle in elongated direction. This direction is achieved for each particle by the hydrodynamic focusing principle of flowcytometry. The size of chain forming cells or colonies is the size of the whole chain or the whole colony that passes the laser beam as one single (larger) particle.

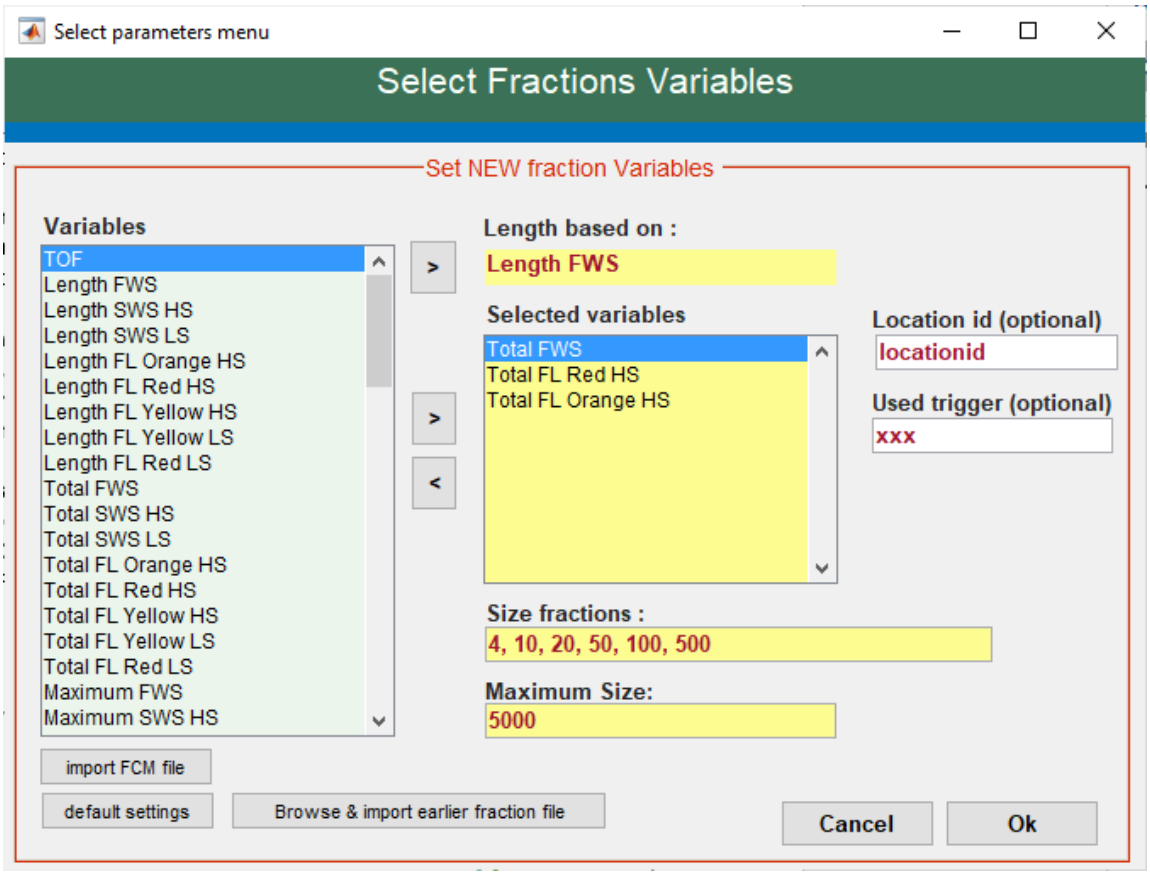
If a calibration line is used for recalculating the size, a formula should be implemented in the software by the user or with help of Thomas Rutten Projects.

Following menu appears:

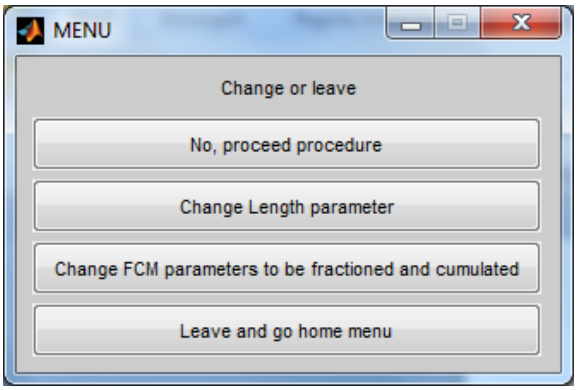


Forget 'ANALYSE & ADD' and 'Make (a) NEW fractiondata file', because they have been replaced by the 'MULTI ANALYSE & ADD' button. This MULTI etc. is usually used to create a new fractiondata file or to add files to an existing fractiondata file. The advantage is that you can use more than one file.
In MULTIANALYSE mode you can start a new file or add it to existing ones. In case of a new file you start with selecting a (cyz)-FCM file, select the FCM length parameter, which is used as an estimation of length of particles.

Select the FCM variables which should be calculated and summed as a function of their size distribution.



Settings can be changed manually by pressing in the box that should be changed and typing. Be aware that if you change the absolute values of the size fractions, to use a dash (-) in between two values.
Change of the 'Length based FCM parameter' (parameter used for size) or change of the FCM parameters that should be fractionated can be changed here, but better is to change them in the next menu to avoid typing errors. The next menu can be found after pressing the 'OK' button.



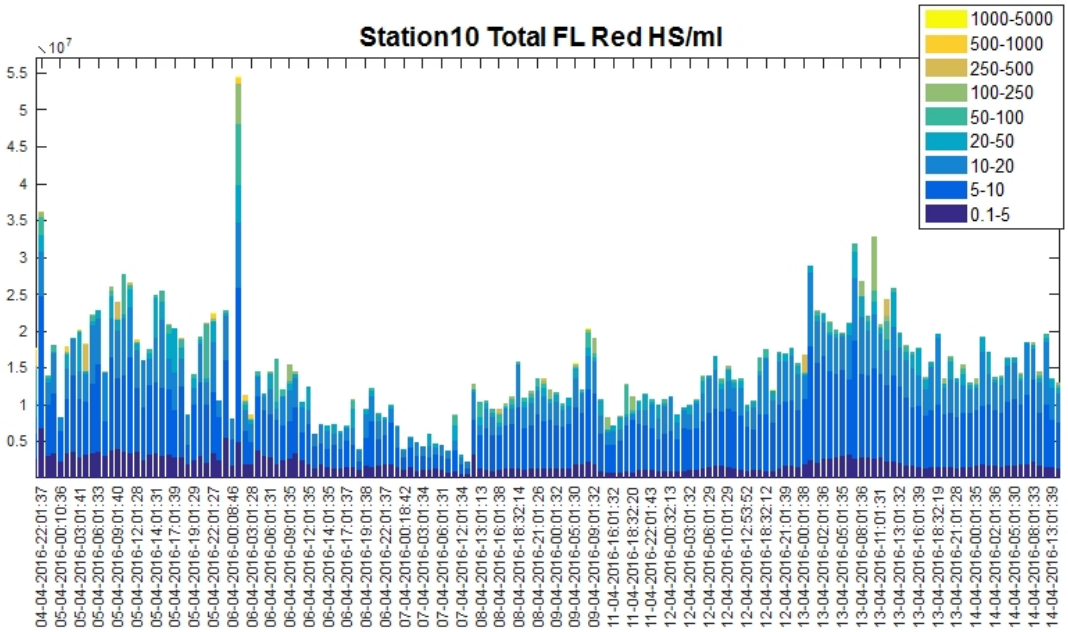
If the settings are well and you don't want to change anything, always choose 'No', proceed procedure for saving your new fraction data*.txt file.

Also after using button 2 (new Length parameter) or 3 (change or add FCM parameters that should be fractionated), the saving procedure should be proceeded by button 1.

After button 1: The settings and the new fractiondata*.txt file is stored. A proposal for the ‘save as’ name is being done: ‘Location name’_fractiondata.txt.
Press op ‘Save’ in order to save.

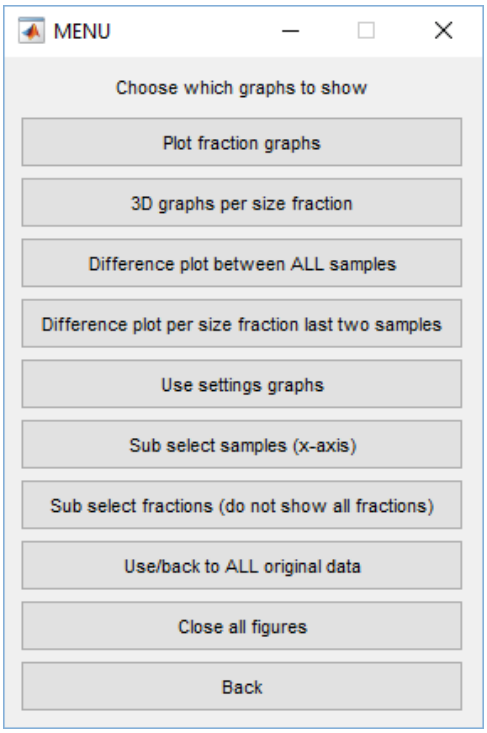
Button 2: Change Length Parameter. If the size parameter needs to be redefined, choose this button. If so, a representative FCM data file will be imported to read the used FCM parameter names for your specific FCM instruments. After that you are able to choose the Length Parameter.

Button 3: Change FCM Parameters to be fractionated/cumulated. Use this button to redefine the FCM parameters that need to be processed as a function of size fractions. If so, a representative FCM data file will be imported to read the used FCM parameter names for your specific FCM instruments. After that you are able to choose the FCM parameters.

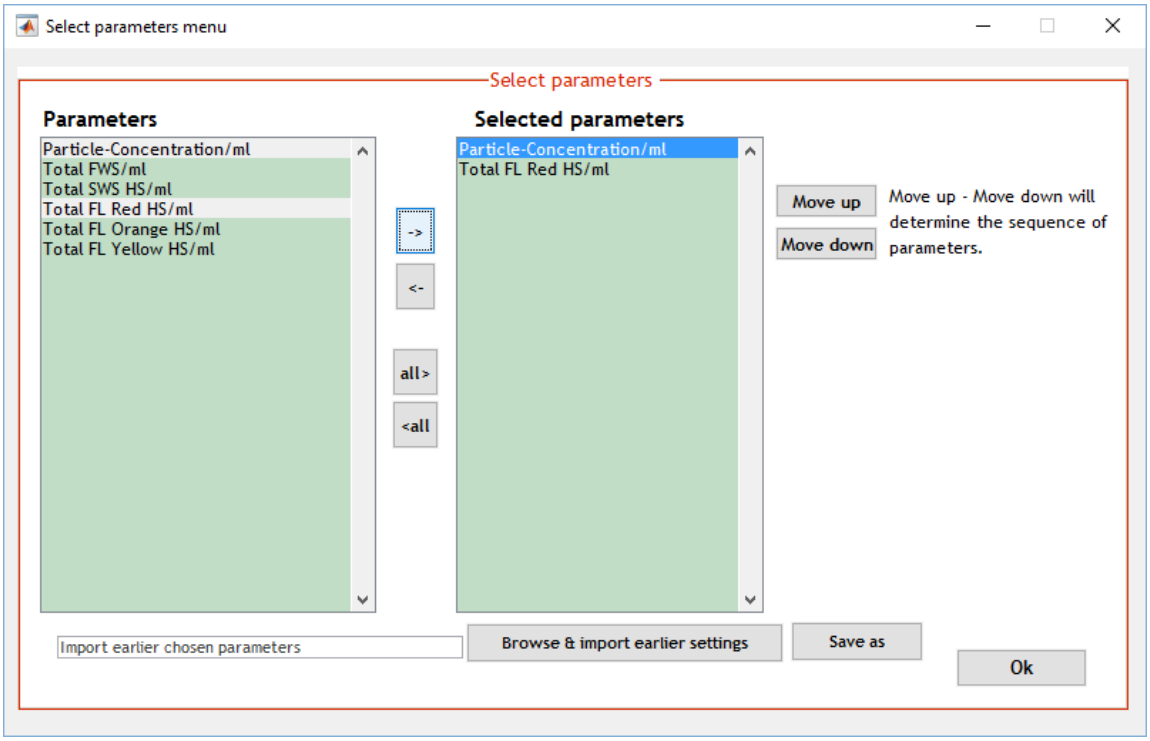


The other buttons
The DELETE button and the CLOSE figures button do what they say and do not need extra explanation.

The SHOW button starts with a menu:



Plot fraction graphs: Choose one or more of the parameters which available by the menu and press OK:



Other graphs are possible too (3D, difference, selection of specific fractions)

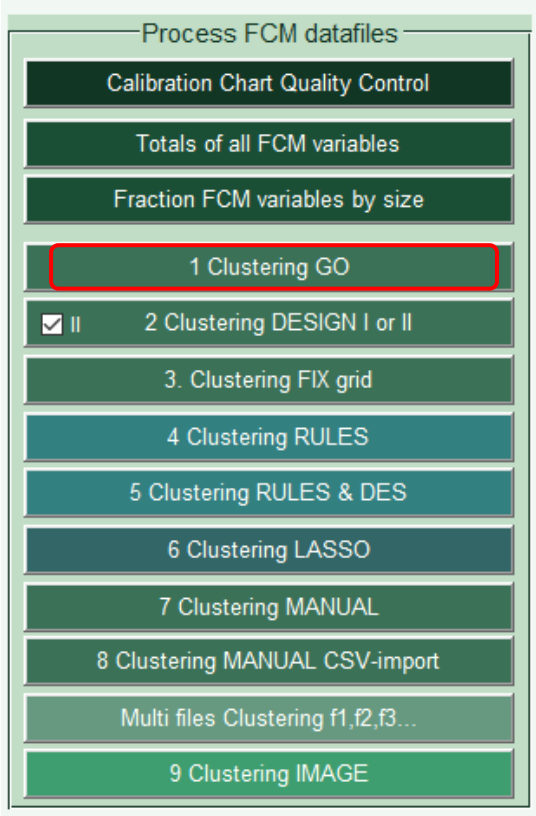
MERGE & Remove duplicates: Sometimes you have semor ethan one fractionfile and/or fraction files contining duplicates. This option puts files together and filter out duplicates.

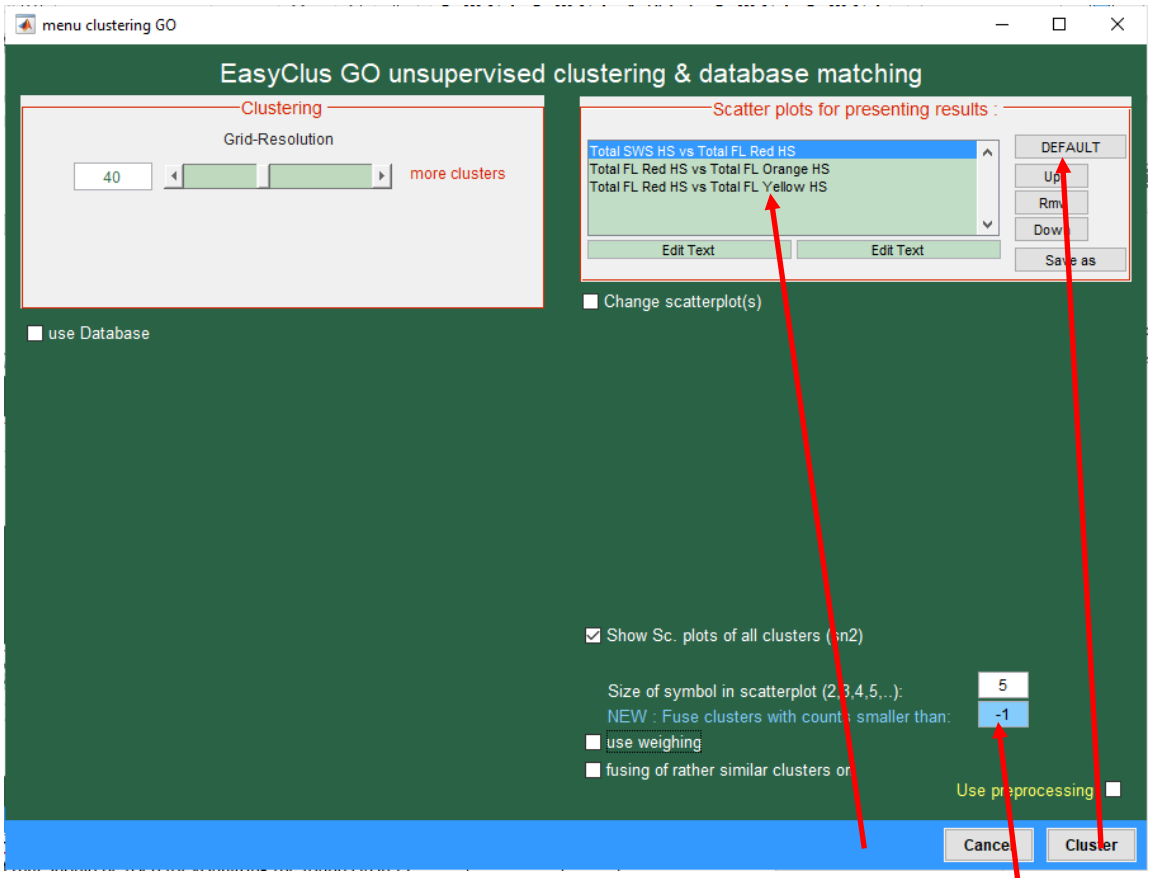
3.5 EasyClus menu Clustering

3.5.1 EASY and SIMPLE: Auto-clustering EasyClus Unsupervised **Method GO** and with similarity database matching (single file) - button 'Auto Cluster Unsupervised & Database recogn.'

To cluster automatically

Cluster are being identified on basis of principal component analysis, density distributions and neighbouring distances.



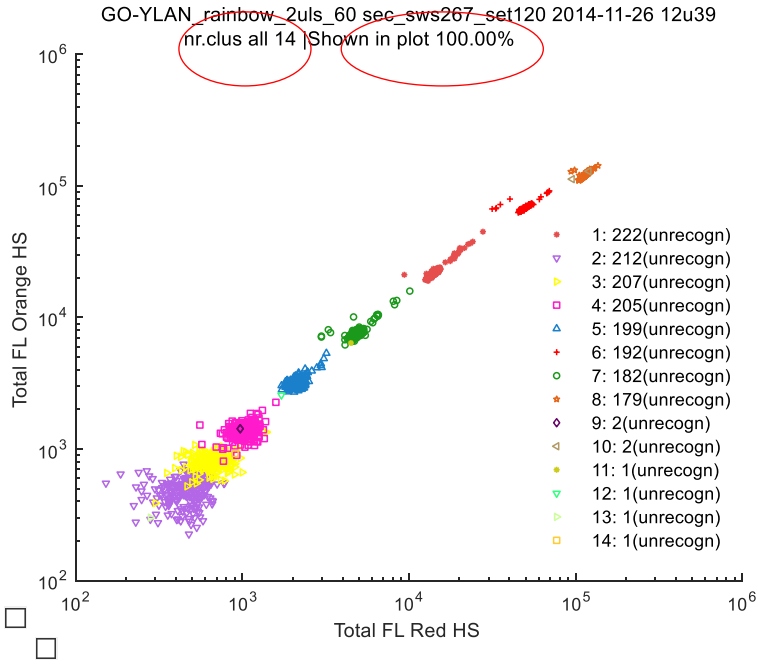


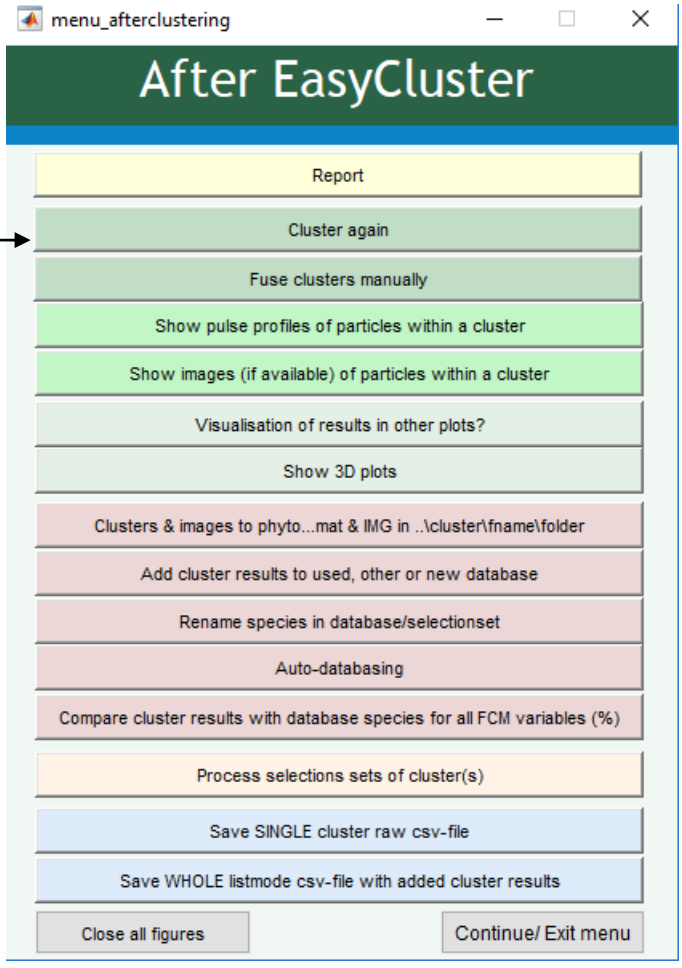
No 'pre'information is required. The only information needed are the scatterplots (or just press DEFAULT) that should be used for visualizing the found clusters.
After clustering you can increase or decrease the number of clusters by shifting the left upper button.

This is nr counts of clusters to be merged in a cluster 'merge'. Set to -1 to ignore merging clusters

- You can use a database by activating the 'use Database checkbox'.
- You can use 'weighing' in for Database matching by activating the 'use weighin checkbox'.
- You can merge clusters by activating the 'fusing of rather similar clusters on checkbox'.
- The preprocessing tick box enables the preprocessing of your data. (see 3.3.4.1)

After clicking the 'Cluster' button the clustering start and the result is visualized in scatterplots. Here clustering of 8 same sized multifluor rainbow beads are clustered.



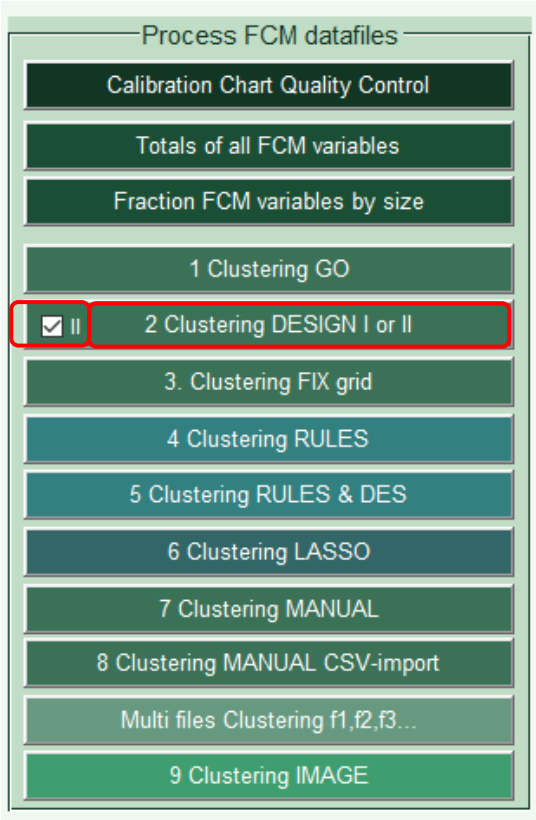


Press to cluster again with other settings

All other buttons are explained one paragraph later (Autocluster DESIGN).

3.5.2. QUITE EASY: Auto-clustering EasyClus Unsupervised Method DESIGN I or DESIGN II and with similarity database matching (single file) - button 'Auto Cluster Unsupervised & Database recogn.'

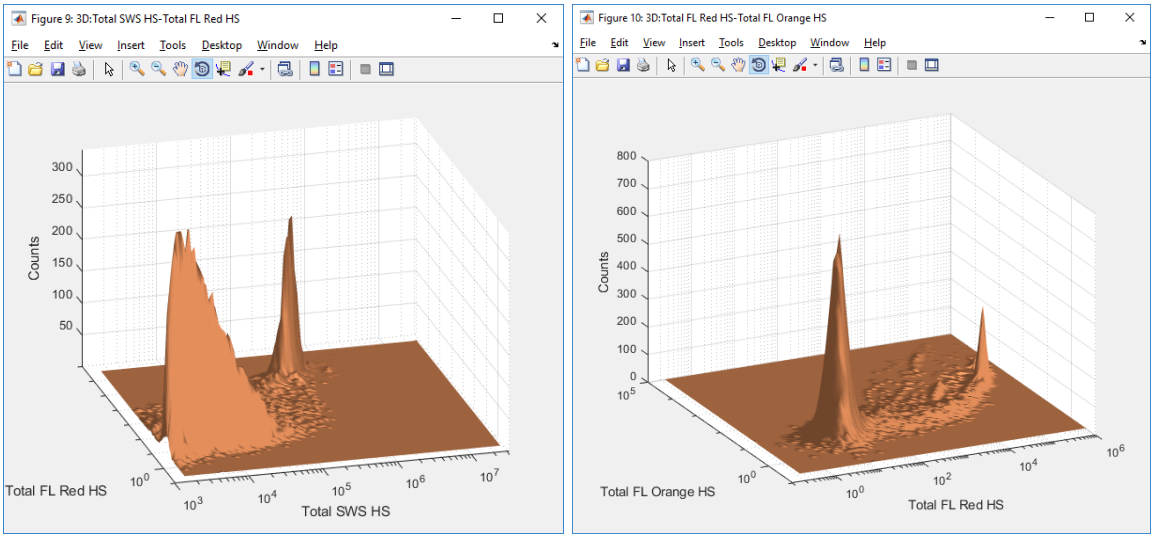
To cluster automatically according to certain specified institutions



Checkbox off = DESIGN I

DESIGN I gives more clusters than DESIGN II

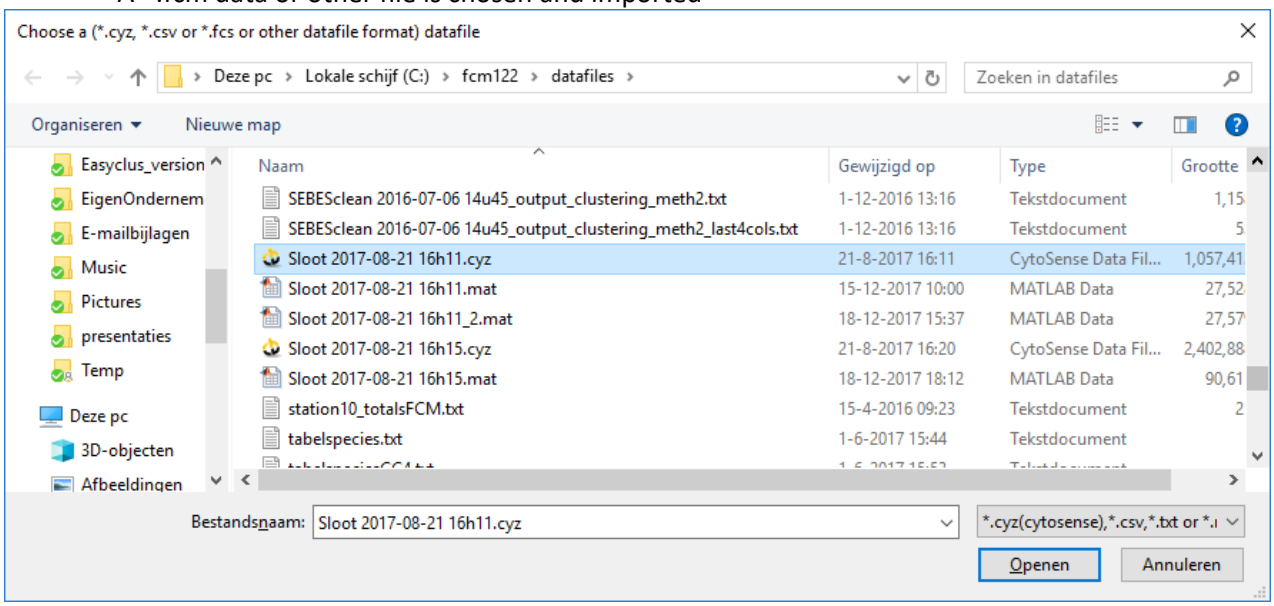
Cluster are being identified on basis of density distributions in two dimensional dotplots combined with the allocation of encircling neighbouring events according to the best fit similarity indexing principle.
For this, only comparison of single grids with each other instead of the whole cluster grids, are taken into account. Assignment to earlier recognized clusters is done on basis of the highest similarity index.



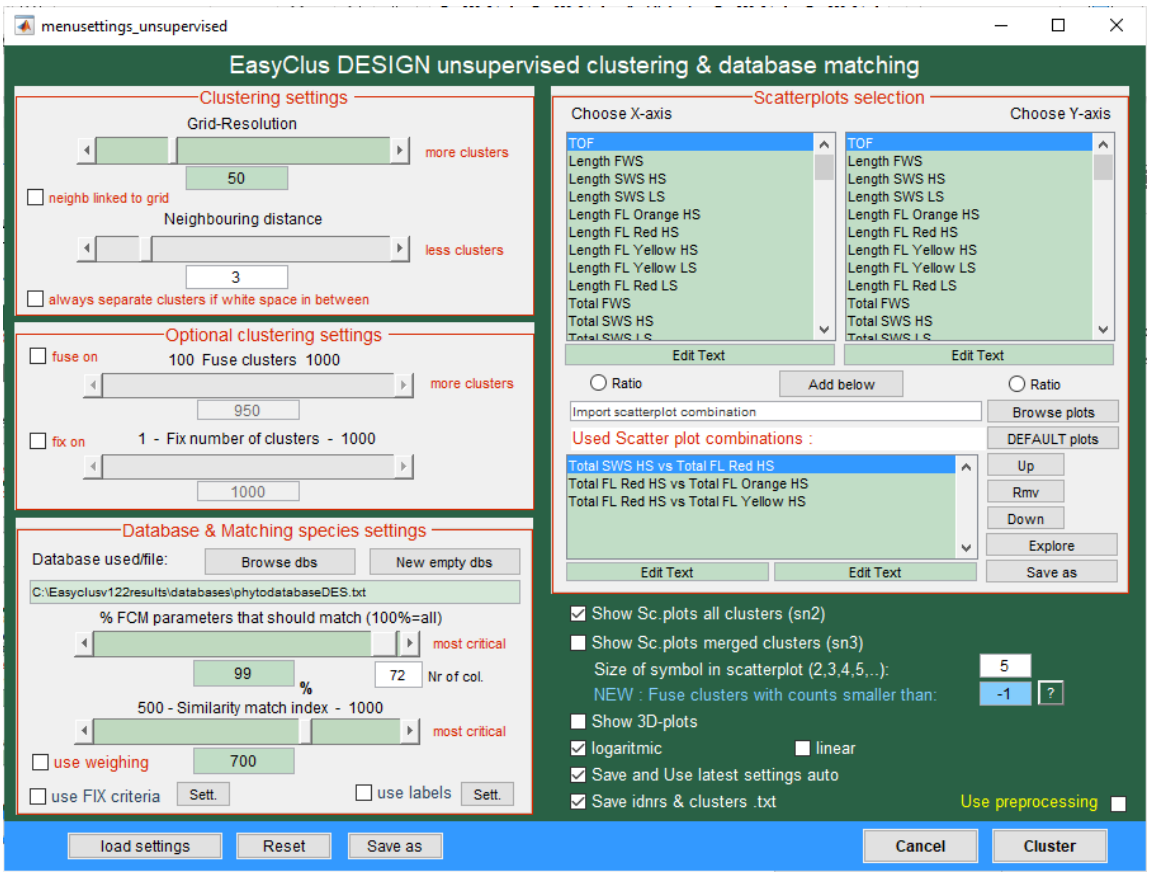
3D visualisation of a scatterplot and showing the number of events in the z-direction.

Start

A *.fcm data or other file is chosen and imported



Subsequently, the standard default settings menu appears. Clustering according to these settings will take place.



Clustering:

- Grid-Resolution: The grid or the number of grids in x-and y-direction that notionally is laid in a scatter plot. The higher the number, the finer the grid.
- Neighbour recognition layer: The number of surrounding 'lattice layers' which seeks to events that already form a cluster. If there are already defined clusters found within this neighbouring layers, then events will be assigned to these "neighbouring-outcome group". The higher the value, the more particles or events will be assigned to a neighbouring cluster. Value is automatically coupled to grid but can be setted manually by de-tick the checkbox.
- checkbox: always separate cluster ...: says what it says, but usually not used.
- Fuse factor fuses clusters that are quite similar up to the lowest allowed similarity value given here. This option is sometimes used.
- Maximum clusters that should be defined is hardly used but can decrease the number of clusters to a setted value here.

Database and fingerprinting:

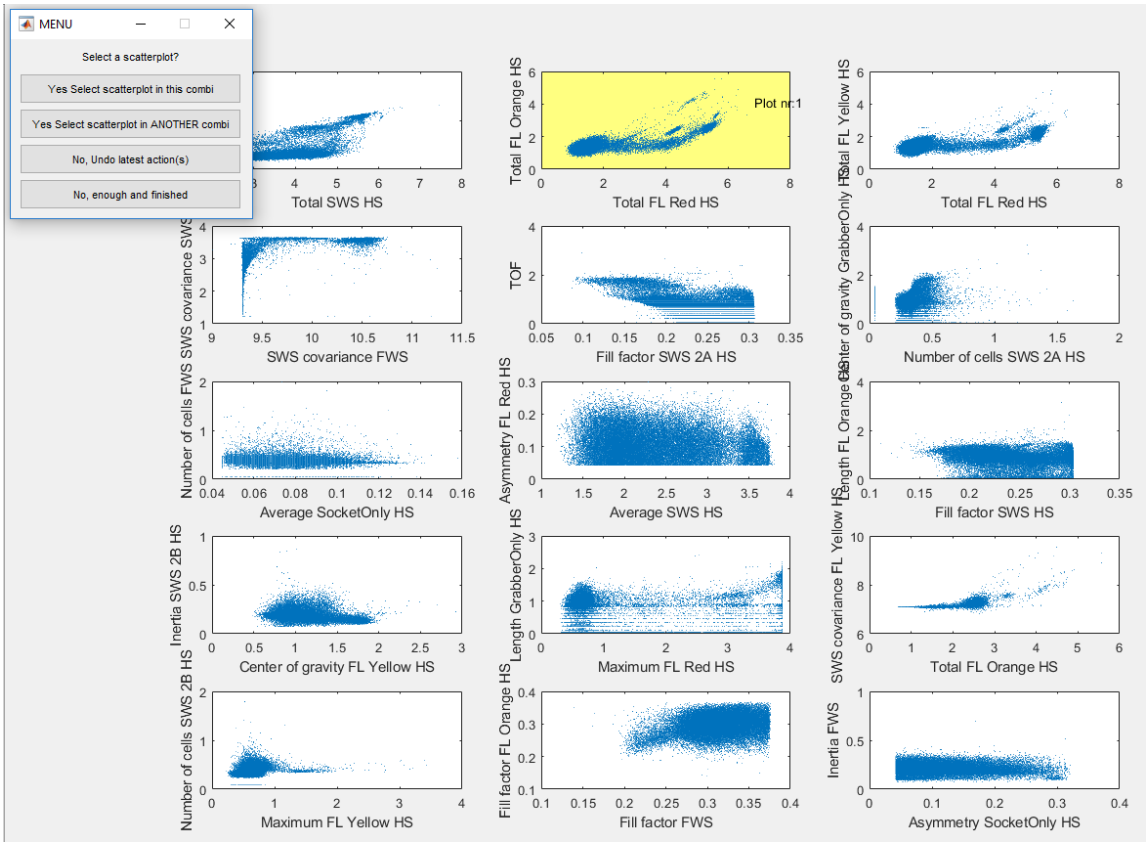
- Database.txt or .mat: Database is a .txt containing only attributes data (Length, Total, Maximum etc.) OR phytoxxx.mat made by EasyClus containing attributes data, profiles data and/or images.
- Nr columns used for similarity indexing: Number of FCM variables of a cluster that should match (at least) with species in the database. Increase makes the matching process more critical.
- Similarity Minimum: Minimum value used for classification of a cluster with species in the database. The assigment to database species will be more critical by increasing the similarity minimum value

Scatterplot combinations that are used as a start for the clustering process

Usually press ' DEFAULT scatterplots' to define good discriminators OR set scatterplots on basis of your own 'EXPERT judgement' how you visualize your data for manual clustering. 'EXPLORE' can be used too, but this is not recommended.

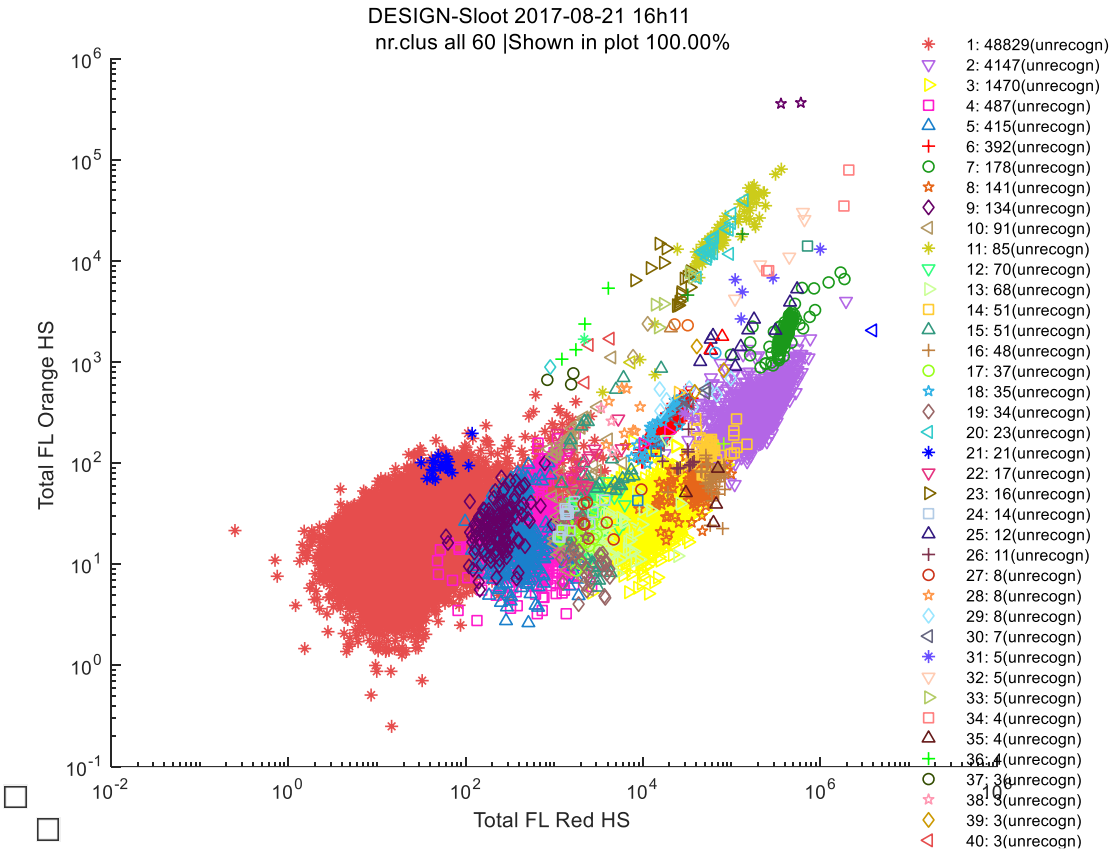
- Used Scatter combinations: Adjust left (x) and right (y). Storing and/or browse to earlier stored combinations is possible. Ratio is possible too, use '/' in between parameters.
- Explore scatterplot combinations: If you want to change the used axes of the scatterplot combinations or the number of scatterplots, you can change it here by choosing one of the relative discriminating combinations shown here.
- Lowest limit of cluster size: The minimum number of events that will be visualized in a scatter plot i.e. only clusters > value will be shown.

When the Explore scatterplot combinations option is activated, the software calculates relative powerful discriminating scatterplot combinations, which will be shown. You can select the combinations you would like to use or calculate new combinations (second button 'Yes select scatterplot in ANOTHER combi').

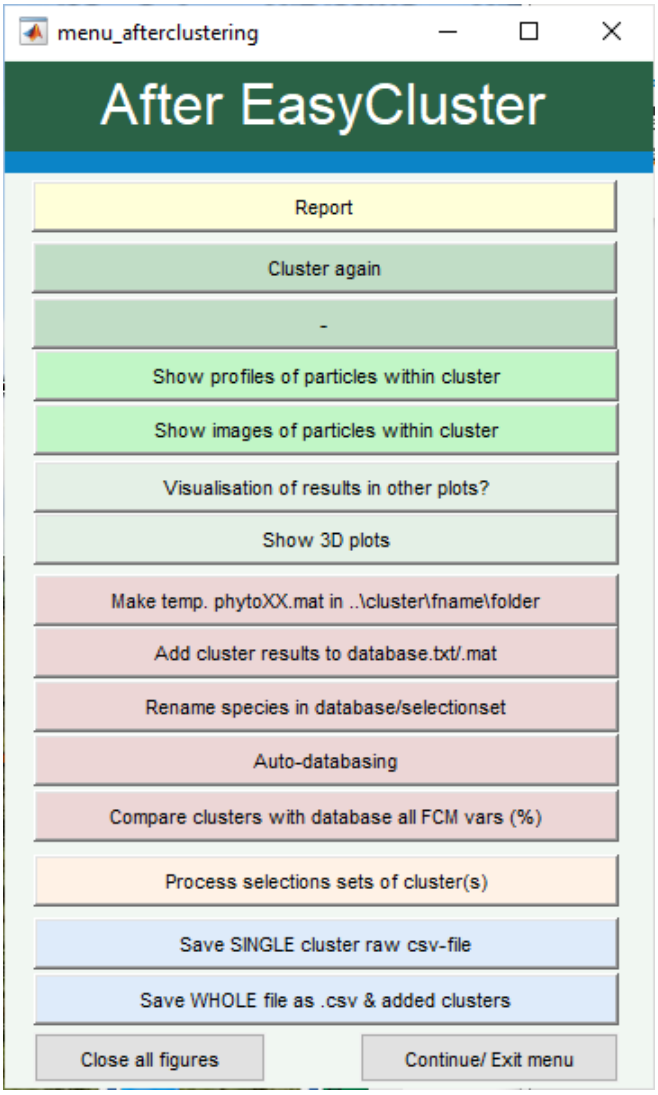


Press latest button 'No, enough and finished' to go back to the previous menu

After all settings are set, the clustering can start (Continu and Start Processing).
A clustering result example is given below.
The first series of figures represent earlier given scatterplot combinations in order to visualise the clusters.



Cluster results are compared with the species in the database for matching according to the given settings. In this example there was an empty database, so all clusters are assigned to ‘unrecogn’.
If there is matching with species in the database, the second series of figures represent given scatterplot combinations in order to visualize the clusters that are recognized after comparison with species in the database.
All results are stored as a txt file (name file with text ‘result’ added) in the (..\cluster) directory as well as the scatterplots (jpg-files) in the ..\cluster\figure directory.



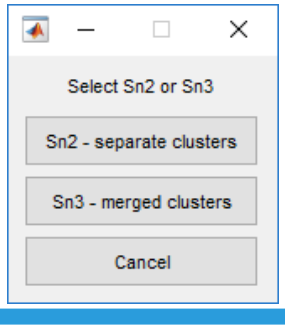
After processing clusters the following menu pops up:

Report: In concept: Report of clusters combined with images.

Cluster again: Cluster again with other settings

- or Fuse clusters manually: To merge clusters manually afterwards, which should be one cluster to your opinion

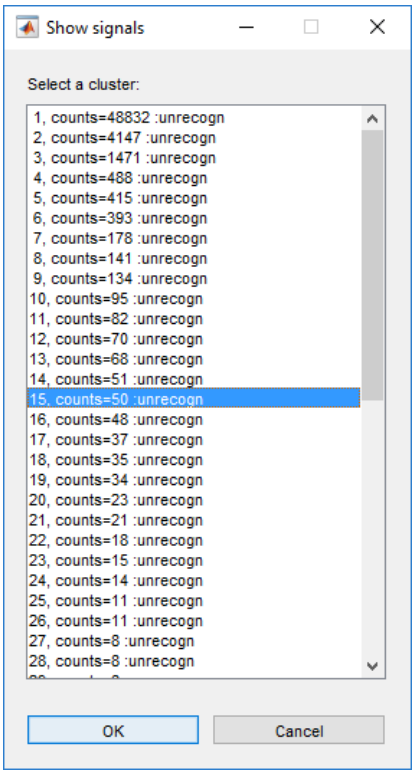
Show pulse profiles of particles within a cluster (Only cyz files CytoSense):
To select clusters and visualize the detector signals of events within a cluster.

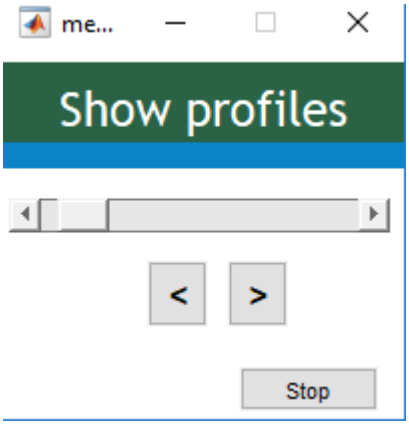


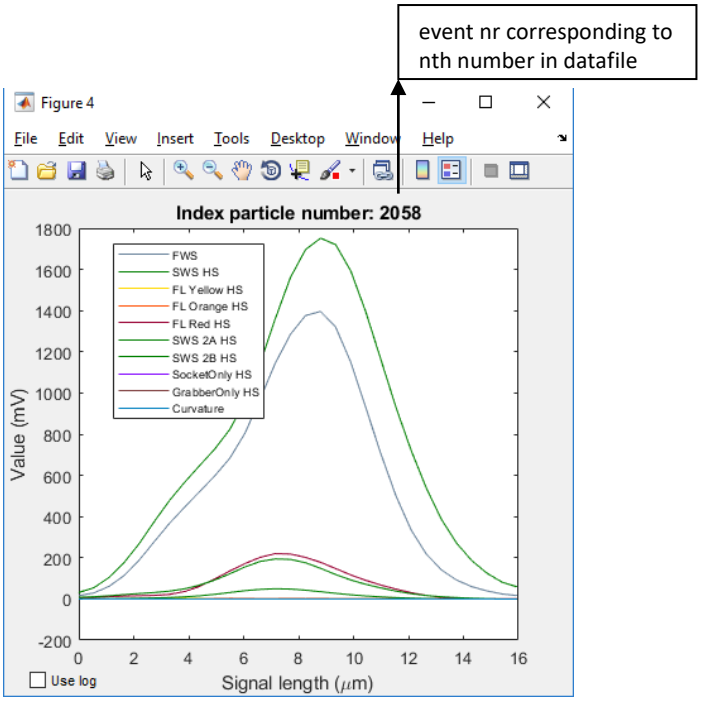
Sn2 : means clusters found before database matching

Sn3 : means clusters found after database matching

We choose Sn2

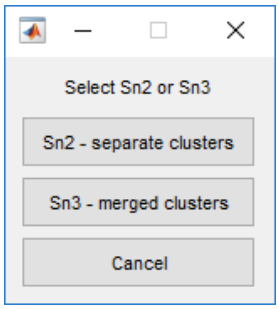




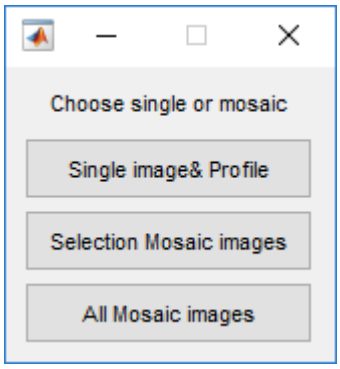


next pulse profile or stop drawing pulse profiles

Show images of particles in a cluster: (Only CytoSense files that are imported as data.cyz files and when images are available). Enables you to visualize all available images one after each other for each selected cluster OR all images of a cluster projected in a mosaic window.



We choose Sn2

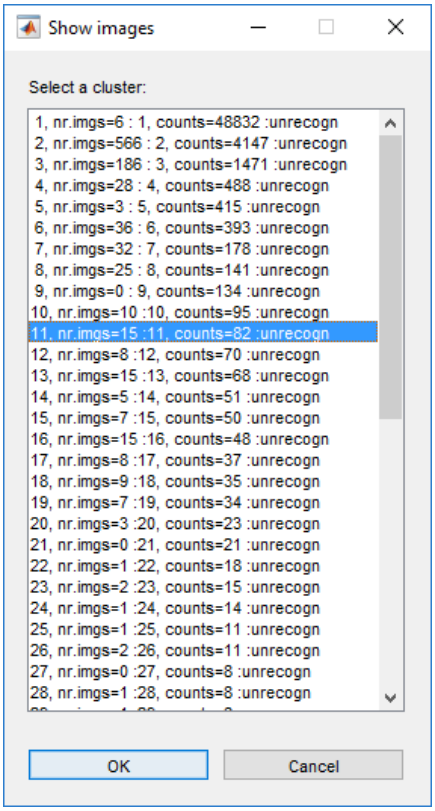


1 One by one combination of image & profile

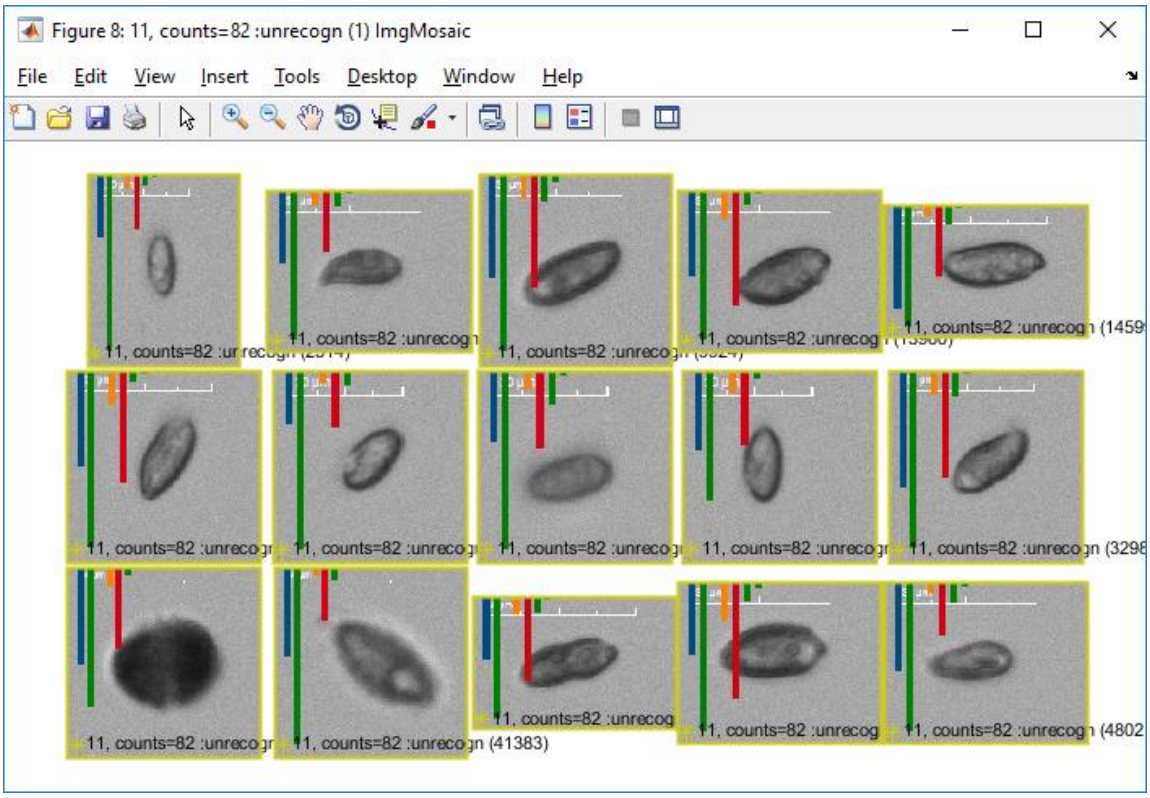
2 Cluster by Cluster of images within chosen cluster

3 All Clusters with images within cluster

We choose option 2



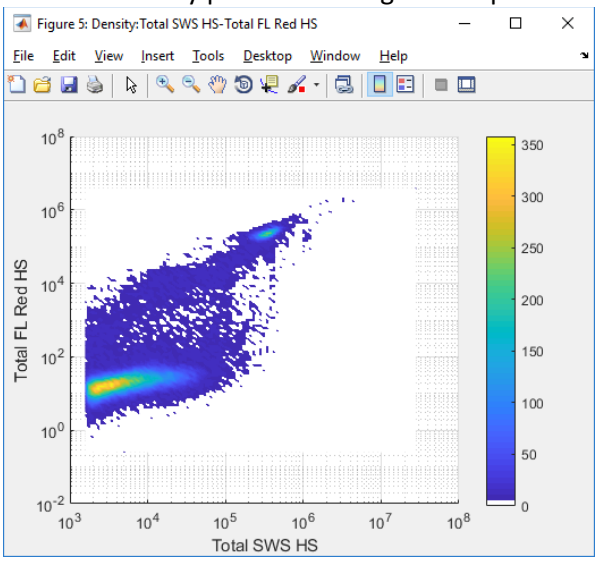
… and have a look in images within Cluster 11



Images found within cluster 11 after unsupervised clustering.

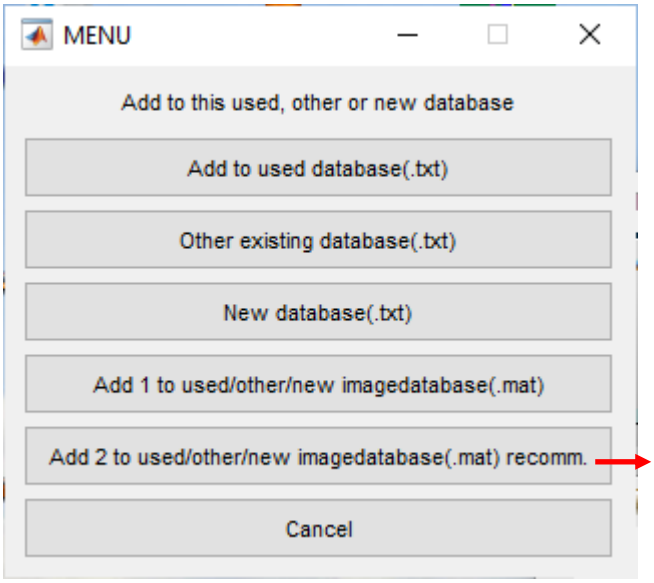
Visualisation of results in other plots'
Draw more scatterplots without changing the clustering results.

Show 3D plots
Draw 3D density plots of existing scatter plots



Clusters & images to phyto…mat & IMG in \cluster\fname\folder
This new option processes full automatically a temporary database with images attributes & profiles (IAP) and saves this to a cluster\filename\phyto_datetimestring_XXX.mat database with images in the folder cluster\filename\databasename\ folder. Only events with an image are stored!

Add cluster results to database .txt/.mat: The individual cluster fingerprint can be added to an existing or new database e.g. if the operator is sure that this is a certain species, so this species fingerprint can be used to recognize this species in other samples too. It is recommended NOT to use the old .txt database option anymore.



1 For attributes database.txt only : not recommended anymore

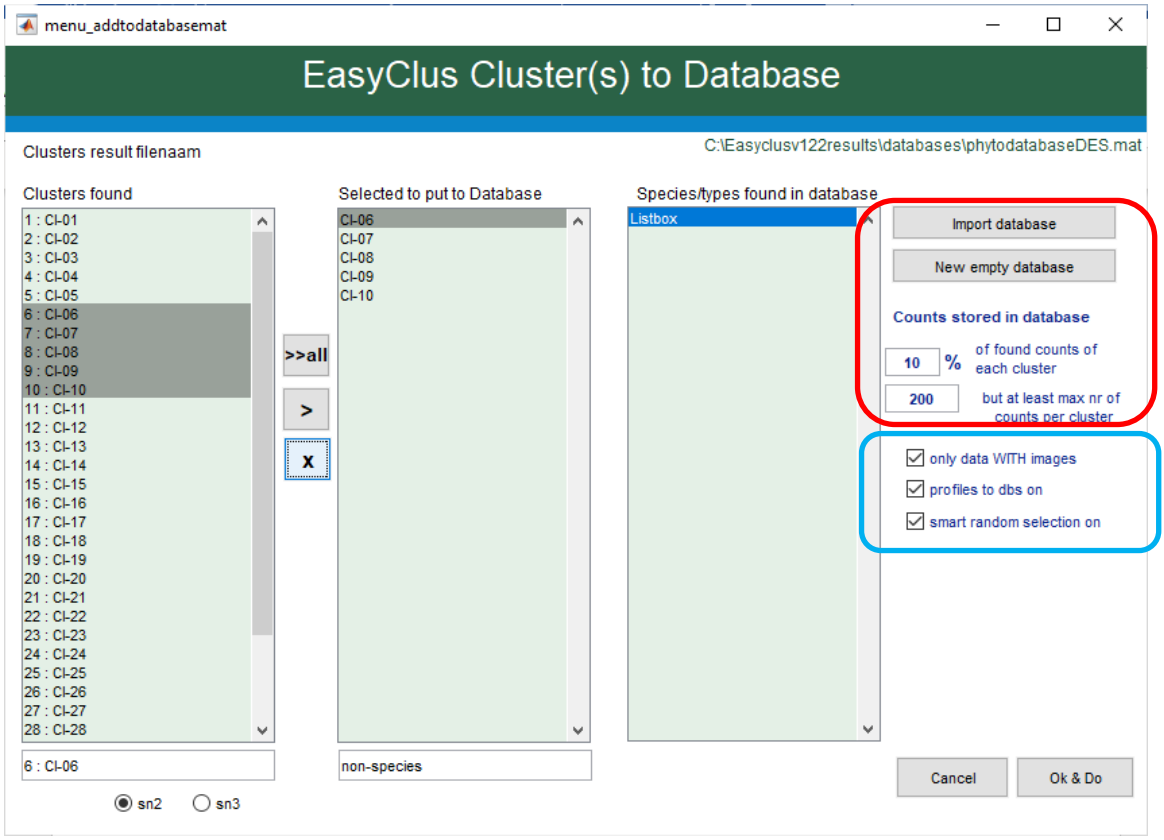
2 For attributes database.txt only : not recommended anymore

3 For attributes database.txt only : not recommended anymore

4 For IAP database.mat : storing by selecting images

5 For IAP database.mat : <u>recommended</u> method

We choose option 5 and we select the clusters 1 to 15 to be stored in the database



= import existing .mat database or make a new empty .mat database
= 10% means that 10% of each cluster (counts) are stored with a minimum of 200.
= 200 particles (at random) of each cluster are stored at minimum. Values can be changed

= if checkbox 'only particles WITH images'='on', only particles with an image are stored are stored in the database. Usually, not all clusters contain particles with an image meaning that particles are not stored in the database (in case checkbox is on). This option 'off' means that all clusters are put in the database, but be aware that particles in the database without an image cannot be validated by the image.

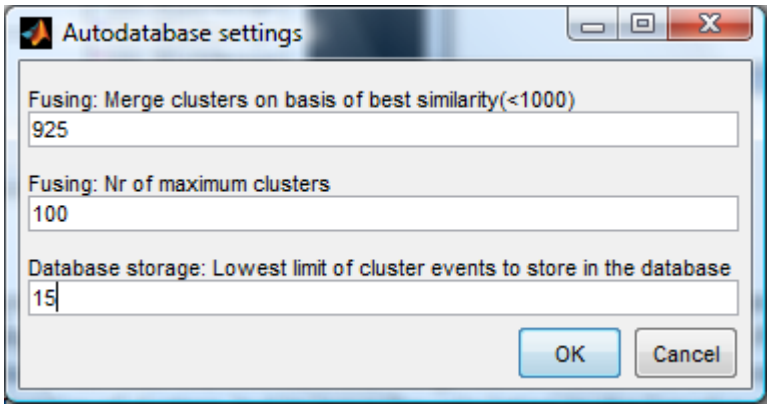
= it is possible to switch off the storage of profiles to the database.mat, however this is not recommended.

= smart selection on : if there are more particles (with or without an image) for storage in the database, this option finds out the which particles to store, also more outlier particles for instance.

The name of each cluster can be changed by the edit box below. Be aware that the automatic name (Cl-01, Cl-02, etc.) can already be available in the database. If recognition with the database is not used, this automatic name can belong to another species type than the one with this name in the database from a previous sample. If recognition by the database is used, and this is recognized (sn3), the name belongs to this species type.

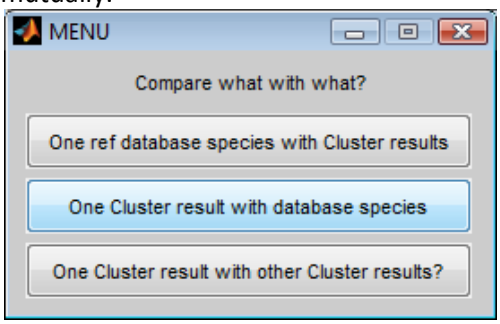
Rename species in database/selection set: To rename the earlier defined cluster or species name of a species, which has been stored in the database (or selection set if lasso method is used).

Auto-databasing: Only database.txt thus only attributes and NOT recommended anymore! Option to add unique cluster fingerprints automatically to a new database. Give minimum nr of events that a cluster should have to store it in the database (in this example=15). Clusters are matched firstly according to given matching criteria, the fusing-merge factor (in this example = 925). When the criteria are 'heavier-more critical' (closer to 1000) more separate clusters will be stored. If a match is found, only the cluster will be stored in the database. Clusters containing events less than given minimum number will not be added to the database. The software can be forced to fuse clusters to a number of given clusters. If this value (Nr of max clusters = 100 in example) is lower than the found clusters, this option will be used. If the number of cluster is not reached, decrease the merge cluster value. It is recommended not to use this fixed nr of clusters, but to leave the calculation to the EasyClus on basis of the similarity index (=default option).

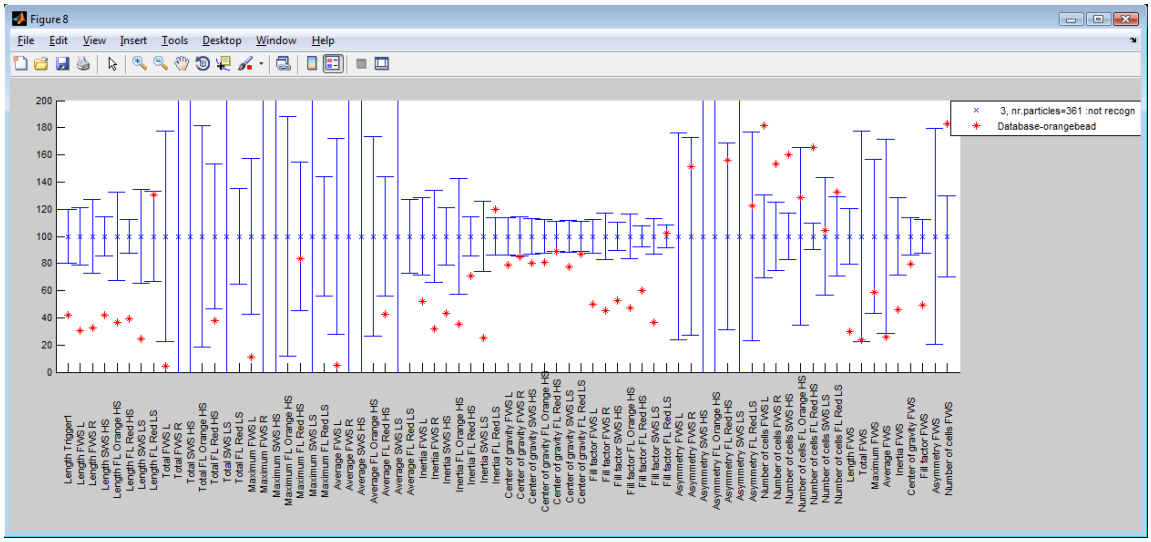


Possible settings for autodatabasing of found clusters.

Compare cluster results with database species for all FCM variables (%): Option to relatively compare all cluster fingerprints with database fingerprints or cluster results mutually.

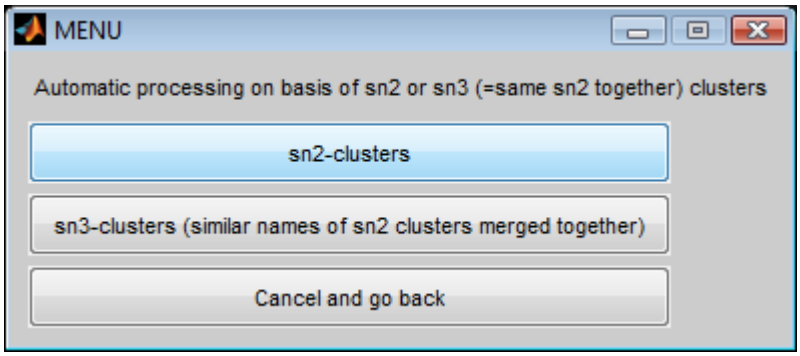


To check the relative values towards another databasespecies, or clustered events.

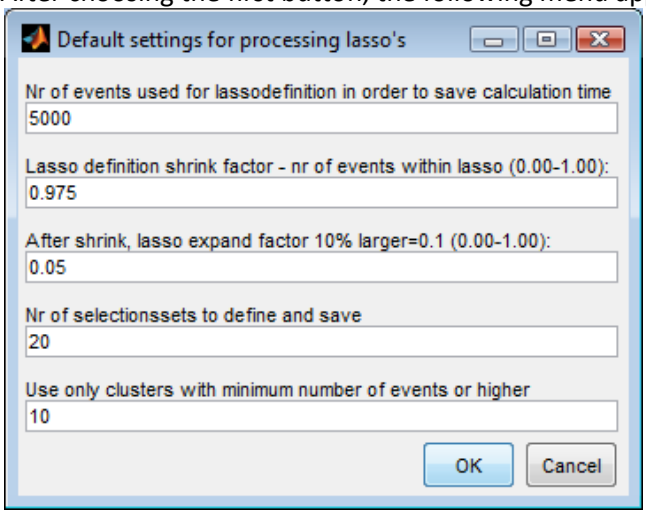


Process selections sets of a cluster: For (automatically) process lasso's or selectionsets on basis of found cluster(s), which can be used in the autolasso (semi supervised) cluster method. Another method is using the phytodatabase.mat option.
Following menu appears:

- Press first Sn2-cluster button if you decide to use all found clusters, without being recognized or fused together on basis of database-matching.
- Press second Sn3-cluster button if you decide to use all found AND fused clusters, which have been recognized and merged together on basis of database-matching.



After choosing the first button, the following menu appears:

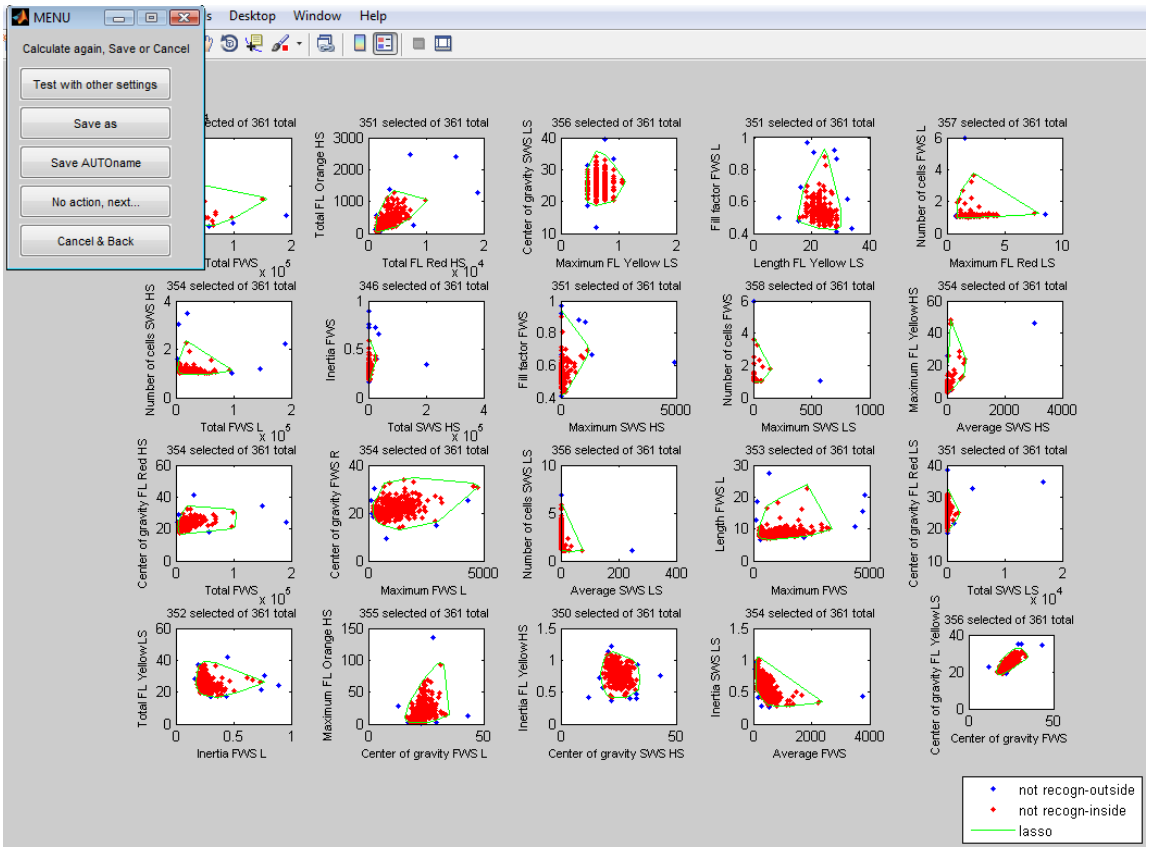


You can change settings in this window, that are used for processing lasso's or selections sets.

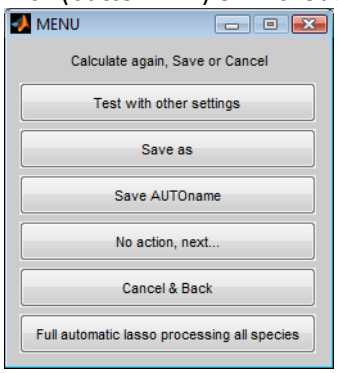
- Lasso shrinking is used to put outliers out of the selection.
- Lasso expand is used to expand the lasso, to be able to handle some instrumental variation.
- Nr of selections or lasso's is the number of lasso's that you would like to store.

Select one species or more species to calculate the selectionsets.

The calculation of lasso's will start.



The 'Test with other settings' button is used to recalculate the lasso's with other settings.
If you have selected more than one species, all lasso can be processed automatically with (button 1..4) or without (button 6) intervention of the operator.



SAVE single cluster raw csv.file: To save the events of one specific cluster as a raw data file.

Save WHOLE listmode csv-file with added cluster results: Save the whole raw FCM data file including clustering results (last columns in the data file) e.g. to check the results in Excel. 'filename_output_after_clustering.txt' file. This option is valuable in order to validate the clustering classification results.

An extra file is saved here with the extension ‘last4cols.txt’ and stored in the same \datafile\-directory. This file can be used in CytoClus to make EasyClus labelling visible in the CytoClus file.

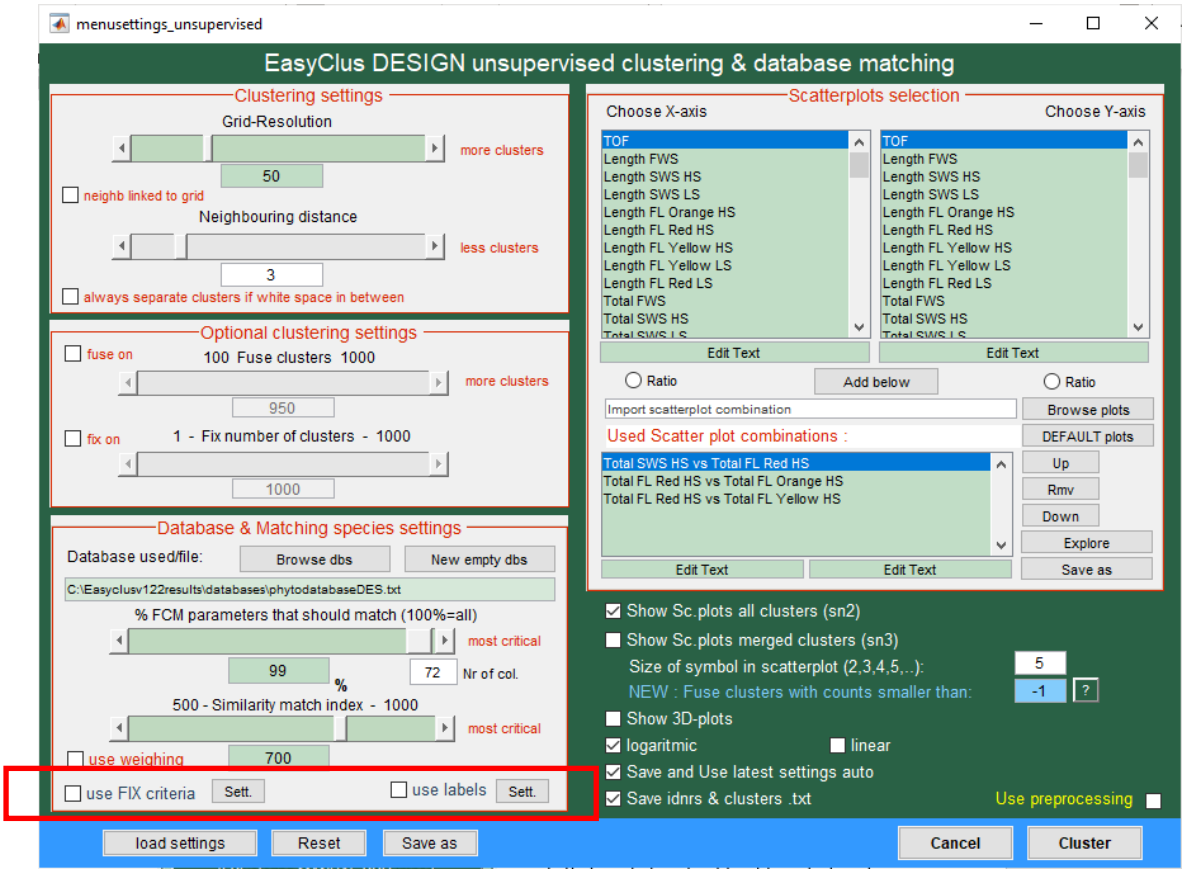
Close all figures: To close all figures e.g. scatterplots that are processed by EasyClus

Stop: To stop the latest clustering method and leave

End of this procedure and the EasyClus appears again.

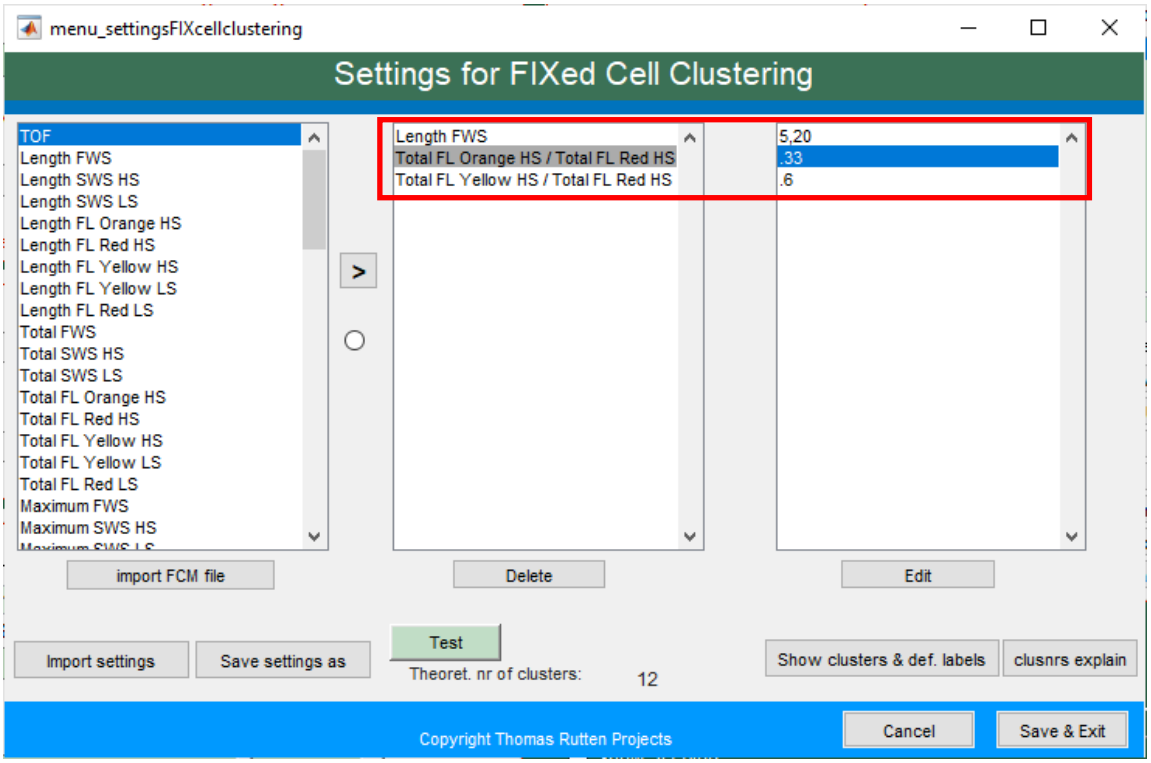
NEW : Cluster clusters by other own given criteria and labels

Clusters can be clustered and/or labeled by use FIX criteria. Use Sett to add criteria e.g. particles <=5 um, 5-20 um, > 20um in (pico, nano, micro) in combination with use label to change the criterium name to your ‘given’ name.

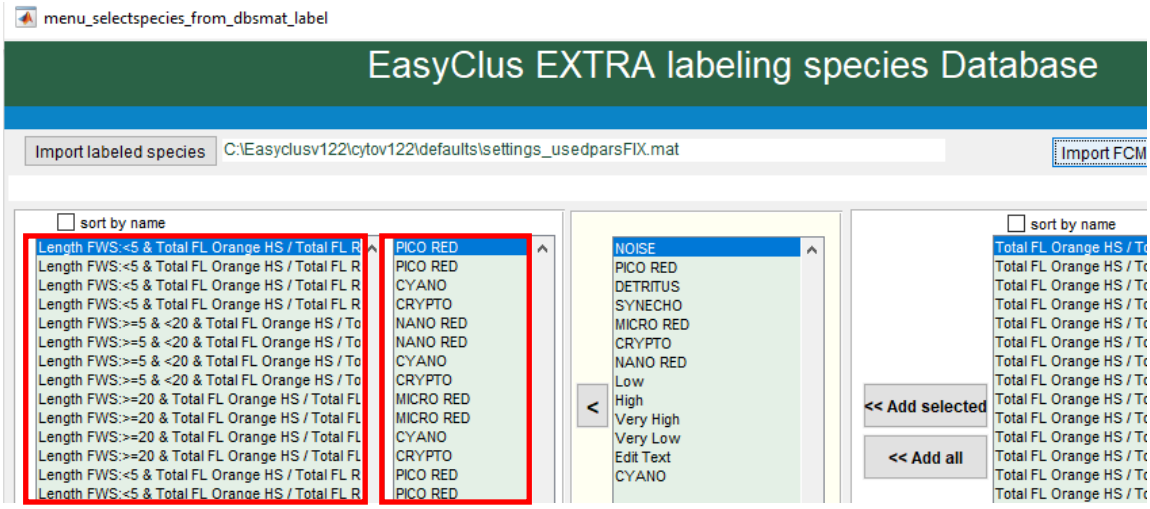


The average criteria value(s) of all data of all particles within a cluster are used to assign the whole cluster to the FIXed criteria. This might be very useful to avoid sharp cut off borders through unsupervised clusters.

Criteria :

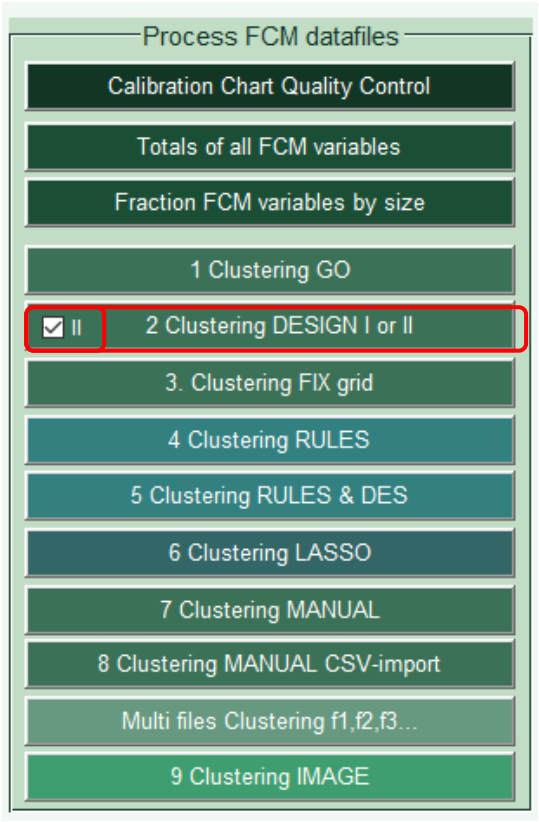


Add label : all clusters with this criteria name are labelled the given other name:



This method might be very useful to put unsupervised clusters into predefined

DESIGN II

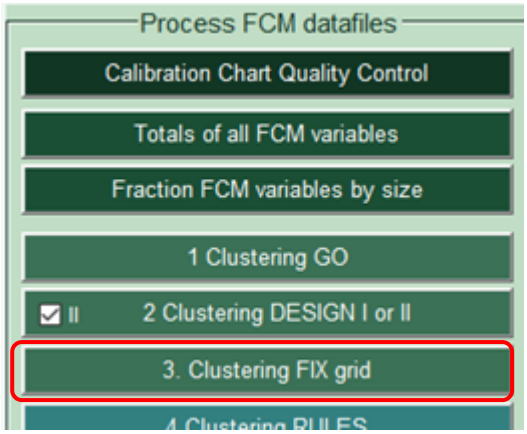


Checkbox on = DESIGN II
DESIGN II gives less clusters than DESIGN I

DESIGN II is almost similar to DESIGN I, but gives less clusters than DESIGN I. In DESIGN II we consider the so called doubt events, which are events that can belong to at least two clusters. In DESIGN II, these doubt events are assigned to the most fitting cluster at the end the scatter plots clustering processes, in stead of during each scatter plot clustering process. DESIGN II is therefore more focused on the main clusters and doubt events are assigned to main clusters. In DESIGN I main clusters can split up by doubt events more easily causing more clusters than DESIGN II. DESIGN II is the recommended method of unsupervised clustering.

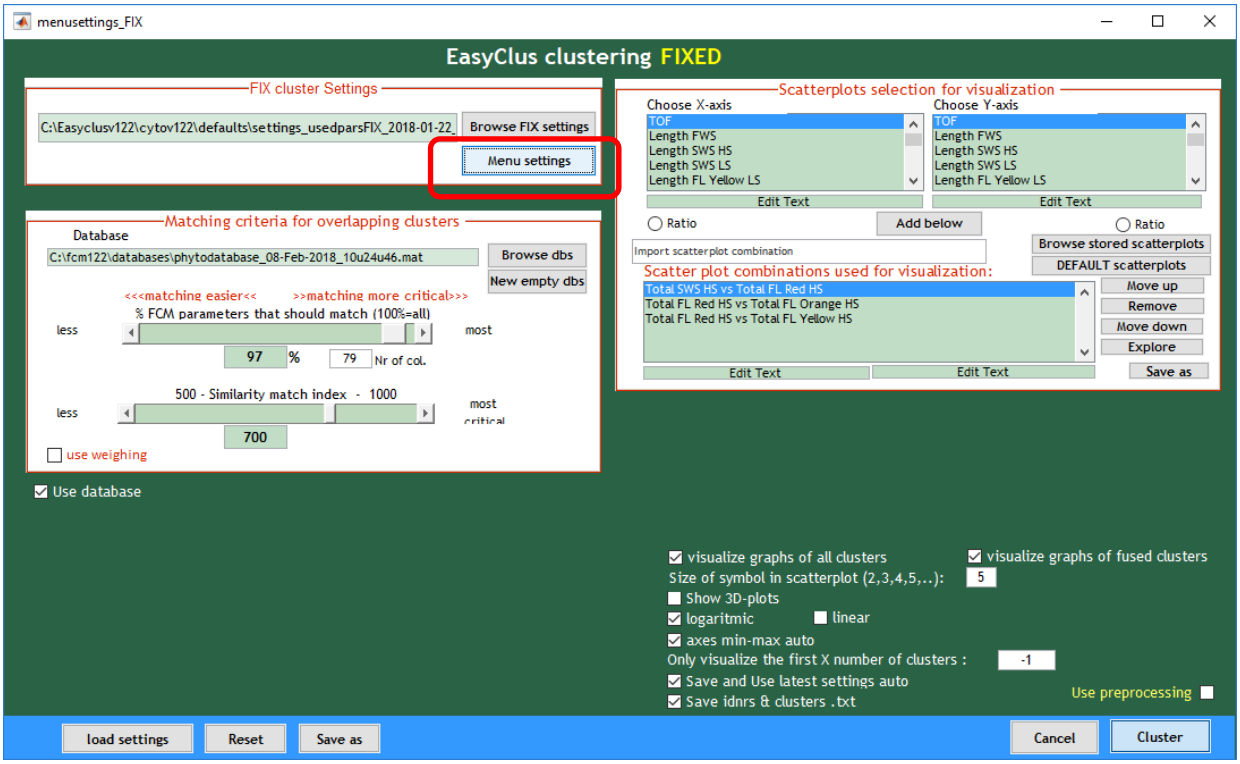
3.5.3.Auto-clustering EasyClus semi-supervised Method FIX grid cells

(Nothing to do with chemical fixative phytoplankton methods)



Very simple and straightforward clustering method. The idea is quite simple. In the multispaced FCM dataset, predefined selection criteria are defined for all kind of FCM variables in such a way that each particle will fit in a specific set of criteria. So all FCM data will fit in one of the fixed cells in the multidimensional space. Each cell or space element has its specific unique number and will only change after redefining the fixed cell criteria. Each cell is a kind of virtual species, based on the definition of the used criteria. Advantage of this method is that unique clusters will be similar independent of the data or used instrument or time

series. Precondition is that the instrument is stable and beads are measured within predefined limits.



Set FCM attributes 'boundaries': Give boundaries to specific FCM variables and all data will be clustered by your self -chosen boundaries.

Example : FCM Length = 4 , 20, 100 means that particles will be divided in :

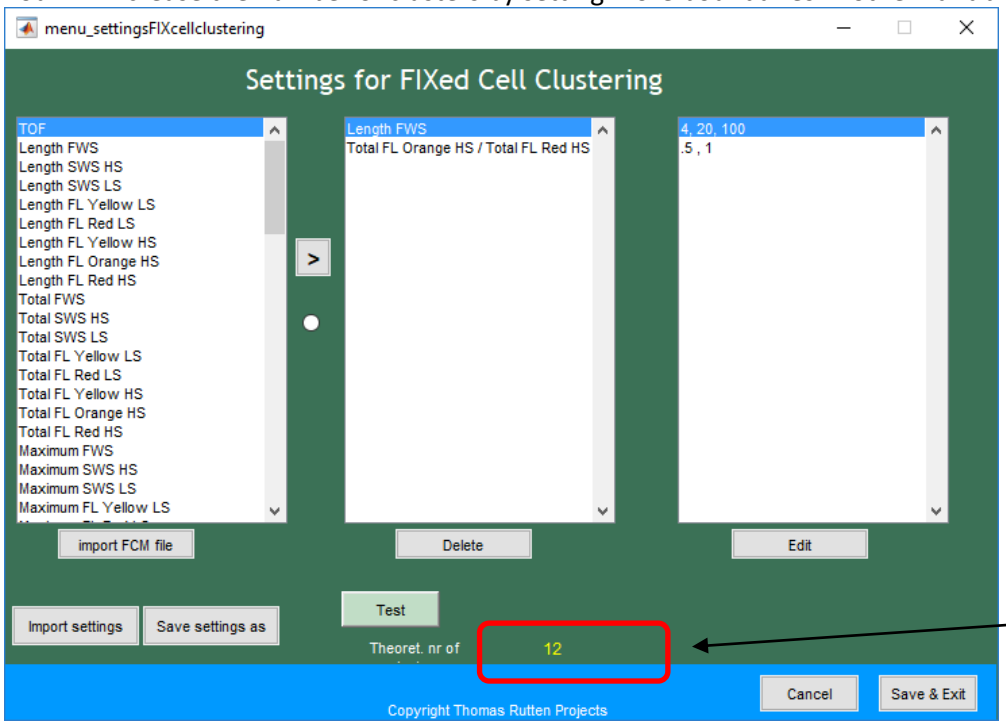
< 4 um

>= 4 & < 20 um

>= 20 & < 100 um

>= 100 um

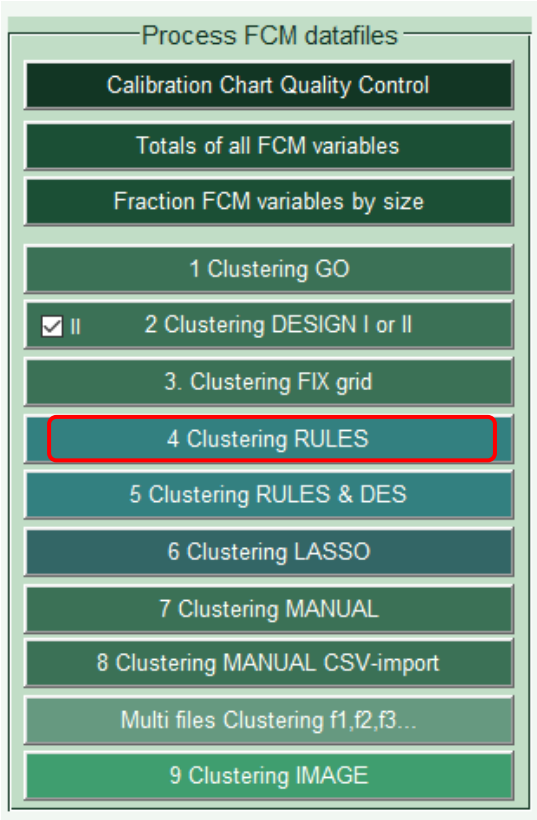
You will increase the number of clusters by setting more boundaries in other variables.



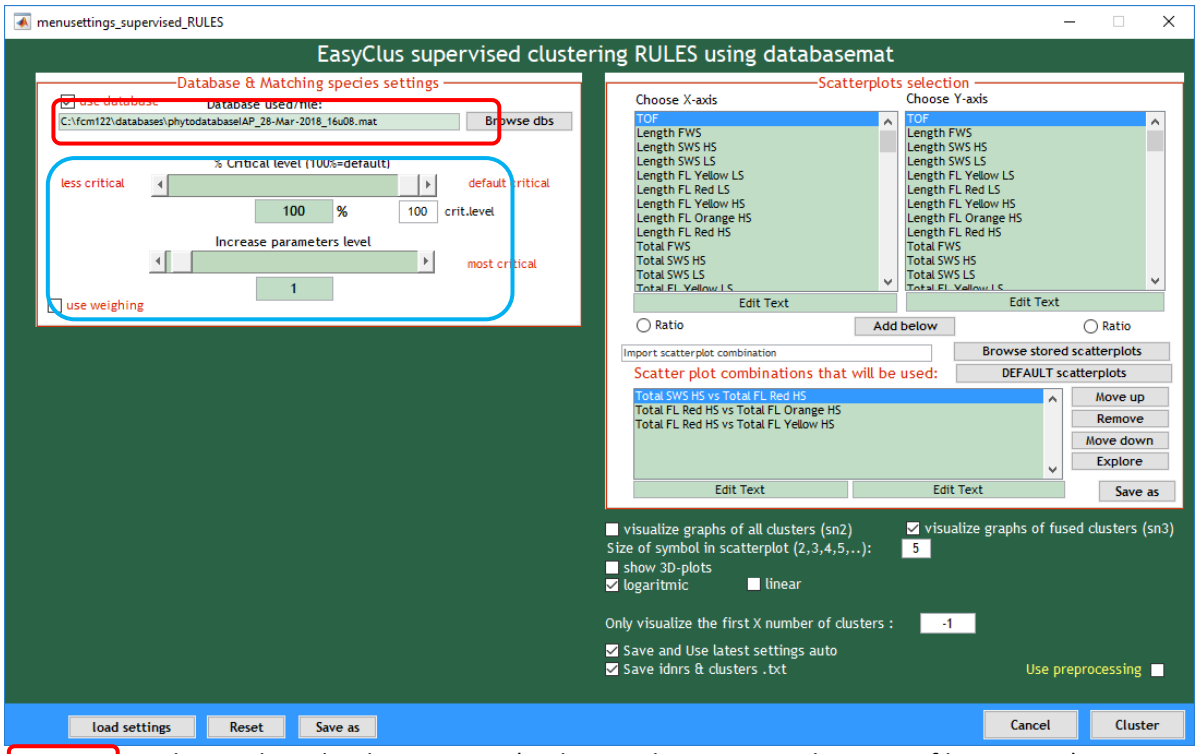
Theoretical Maximum number of clusters = 12

3.5.4 Auto-clustering (single file) – button – RULES Clustering

To cluster automatically on basis of the (new) database phytodatabaseXX.mat format



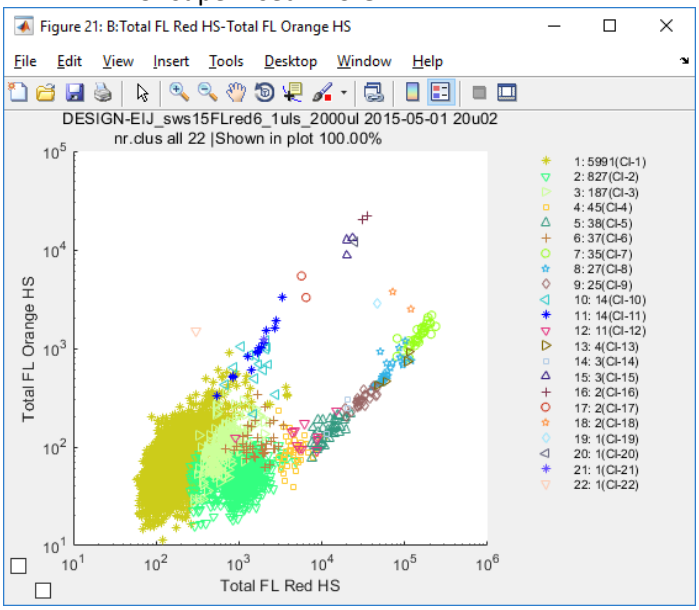
This supervised clustering method needs training data, which is derived from the **phytodatabaseXX.mat** containing species (clusternrs) combined with attributes data (Length, Total, Maximum ..) , signal profiles data and/or images. Images are recommended in order to be able validate the image with the stored species name.
The RULES method uses, as it says itself, specific rules which characterizes each species in the database towards all other. These rules are based on the flowcytometric attribute data. A smart EasyClus algorithm will find the most discriminating combination of variables for each species in the database. It might be possible that not all species can be discriminated due to a lack of discriminators between species. This method has shown to be very effective supervised clustering.



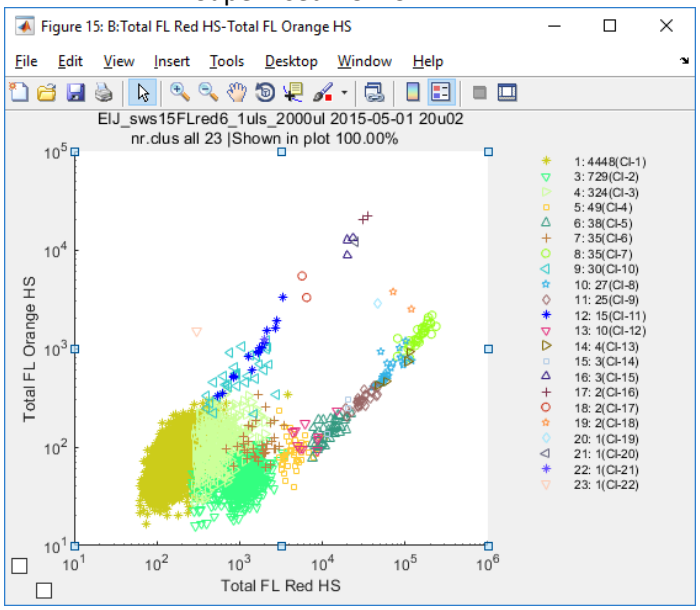
= select a phytodatabaseXX.mat (with stored species-attributes-profiles-images)

= settings for rules method 100% means that boundaries are used at 100% = identical. The boundary values are set less critical by decreasing values. Increase parameterslevel is default 0, means that the default number for each species calculated by the software is used. Increasing this number means that an extra criterium is added to discriminate a species (more precise, but also chance of missing it due to extra criteria).

Unsupervised DESIGN II



Supervised RULES



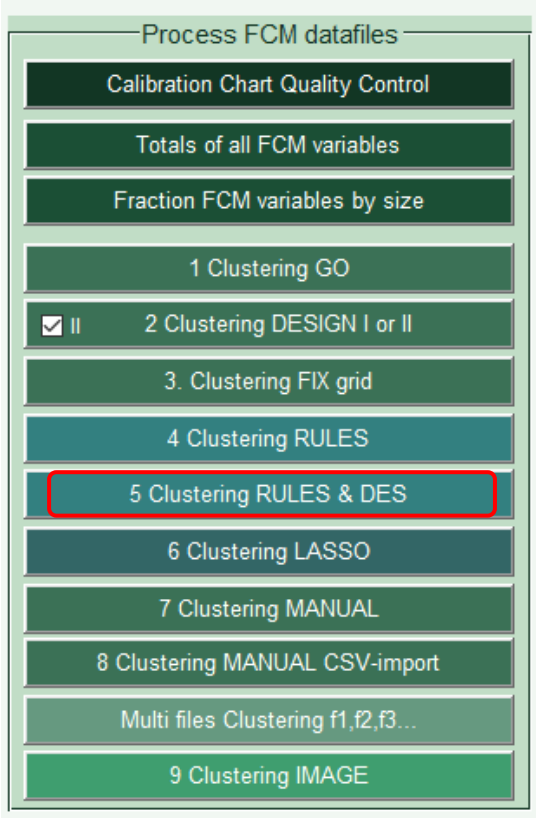
RULES clustering based on (200 events from each unsupervised cluster) phytodatabaseXX.mat.

Clusters meth 1	Nr ev.	Clusters meth 2	Nr ev.	%match
1	5991	1	4421	0.73794025
2	827	3	719	0.86940747
3	187	4	170	0.90909094
4	45	5	45	1
5	38	6	38	1
6	37	7	32	0.86486489
7	35	8	35	1
8	27	10	27	1
9	25	11	25	1
10	14	9	14	1
11	14	12	14	1
12	11	13	10	0.90909094
13	4	14	4	1
14	3	15	3	1
15	3	16	3	1
16	2	17	2	1
17	2	18	2	1
18	2	19	2	1
19	1	20	1	1
20	1	21	1	1
21	1	22	1	1
22	1	23	1	1

Particles found in unsupervised and found by RULES method. %match in last column show good results.

3.5.5 Auto-clustering (single file) – button – RULES & DESIGN II Clustering (HYBRID)

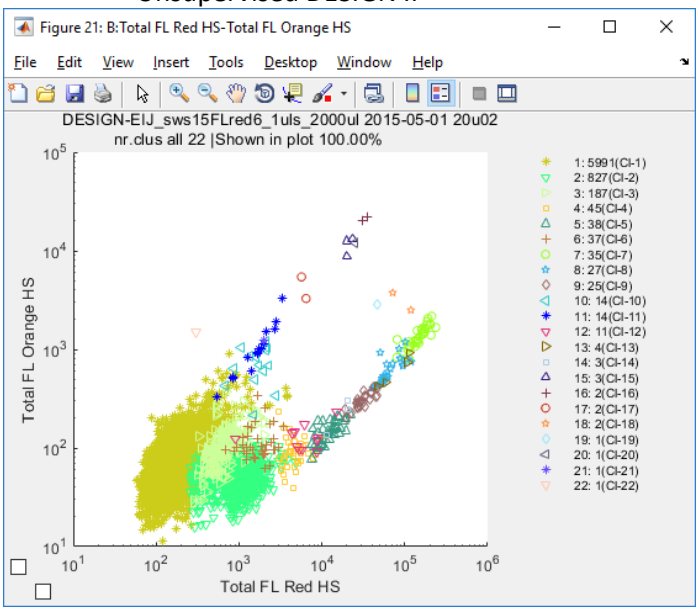
To cluster automatically on basis of the (new) database phytodatabaseXX.mat format



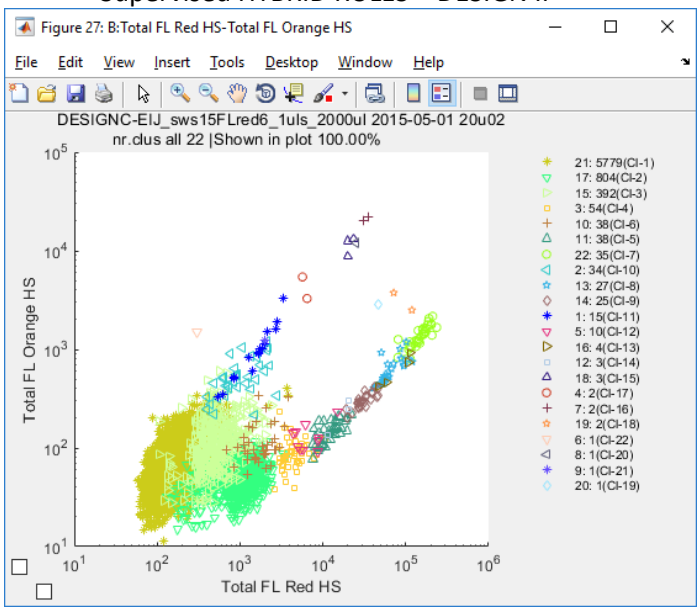
This concept HYBRID method uses DESIGN II clustering, producing unsupervised clusters AND uses the RULES clustering, producing clusters as well. This method is introduced (as a test case) to find out what to do with particles that are not in a database and particles that are close to supervised clusters, but not recognized as a species in the database.

Example 1: If a cluster of 400 particles is found by unsupervised clustering, and 300 particles in this cluster match with one species of the supervised clustering, it is considered that all 400 particles might be assigned to this cluster, because no other species had a match.
Example 2: If a cluster of 200 particles is found by unsupervised clustering, and 0 particles in this cluster match with one species of the supervised clustering, this 200 particles cluster remains a unique cluster, while it was missed in the supervised method.

Unsupervised DESIGN II



Supervised HYBRID RULES + DESIGN II



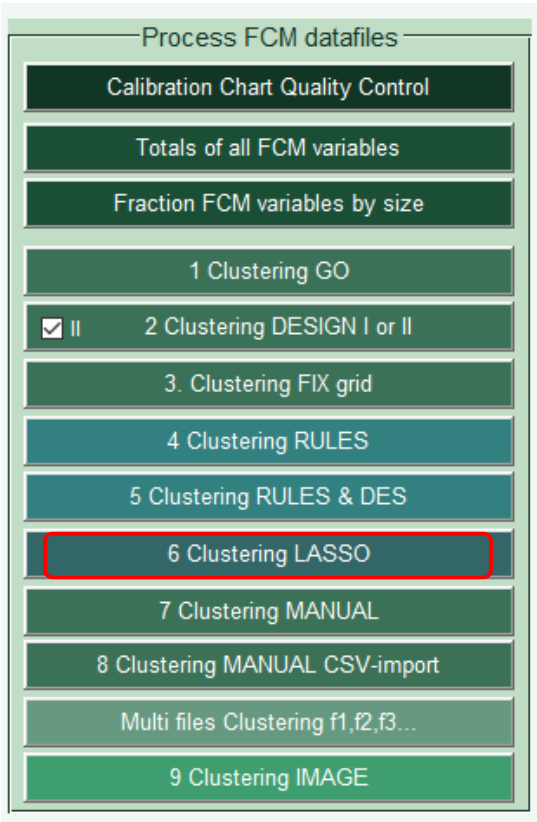
RULES clustering based on (200 events from each unsupervised cluster) phytodatabaseXX.mat.

Clusters meth 1	Nr ev.	Clusters meth 2	Nr ev.	%match
1	5991	21	5748	0.95943916
2	827	17	769	0.92986697
3	187	15	170	0.90909094
4	45	3	45	1
5	38	11	38	1
6	37	10	32	0.86486489
7	35	22	35	1
8	27	13	27	1
9	25	14	25	1
10	14	2	14	1
11	14	1	14	1
12	11	5	10	0.90909094
13	4	16	4	1
14	3	12	3	1
15	3	18	3	1
16	2	7	2	1
17	2	4	2	1
18	2	19	2	1
19	1	20	1	1
20	1	8	1	1
21	1	9	1	1
22	1	6	1	1

Particles found in unsupervised and found by HYBRID RULES & DESIGN II method. %match in last column show very good results

3.5.6 Auto-clustering (single file) – button – LASSO Clustering

To cluster automatically on basis of a new database phytodatabaseXX.mat OR stored lasso's or selections sets of species.



The lasso clustering uses predefined data (training set) to cluster particles. Particles which fit in the number of chosen predefined selections sets (lasso's) are assigned to these lasso's.

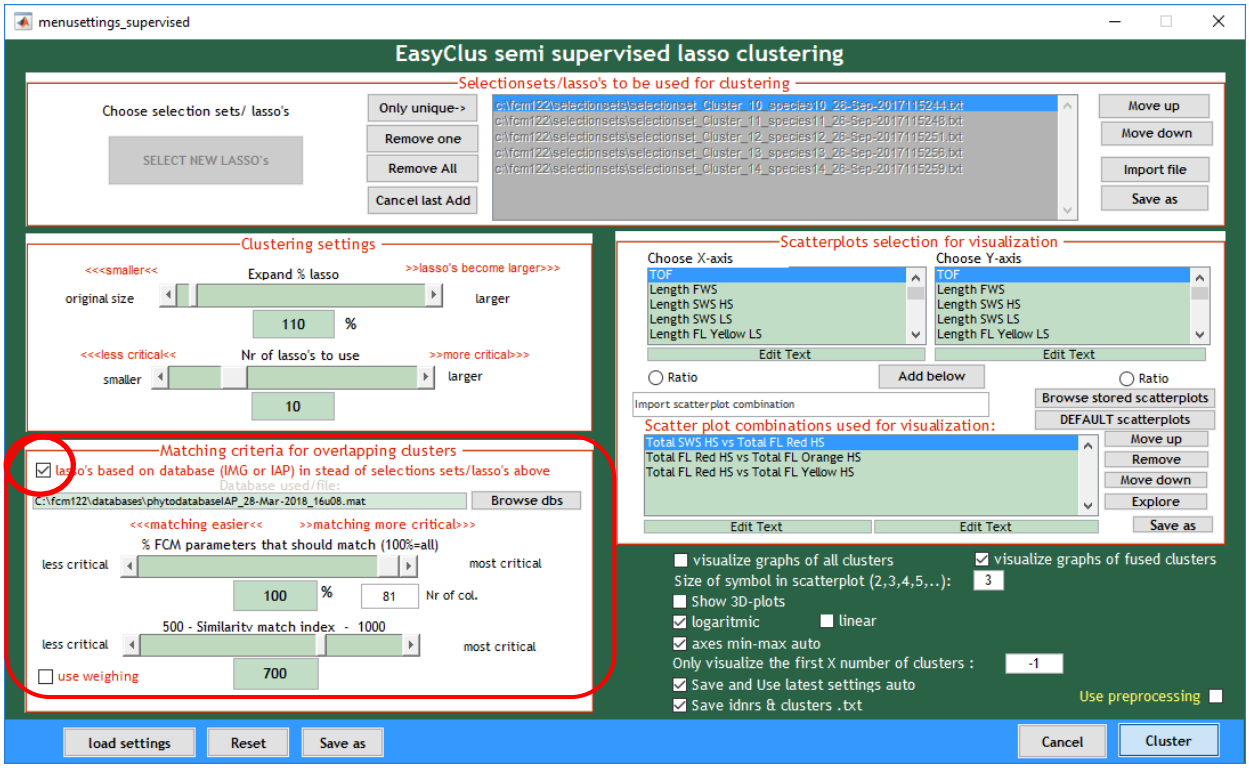
How to get lasso's ?

- Recommended: Selection sets or lasso's can are calculated (by EasyClus) out of a previous filled phytodatabase.mat (phytoxx.mat) containing attributes data as well as profiles as well as images.
- Selections sets can also be processed by storing selectoinsets during the the unsupervised clustering method or the manual method. These lasso's are stored selections sets of clusters of data in bivariate plots and stored under the directory fcm\selectionsets\.

Start of lasso-clustering:

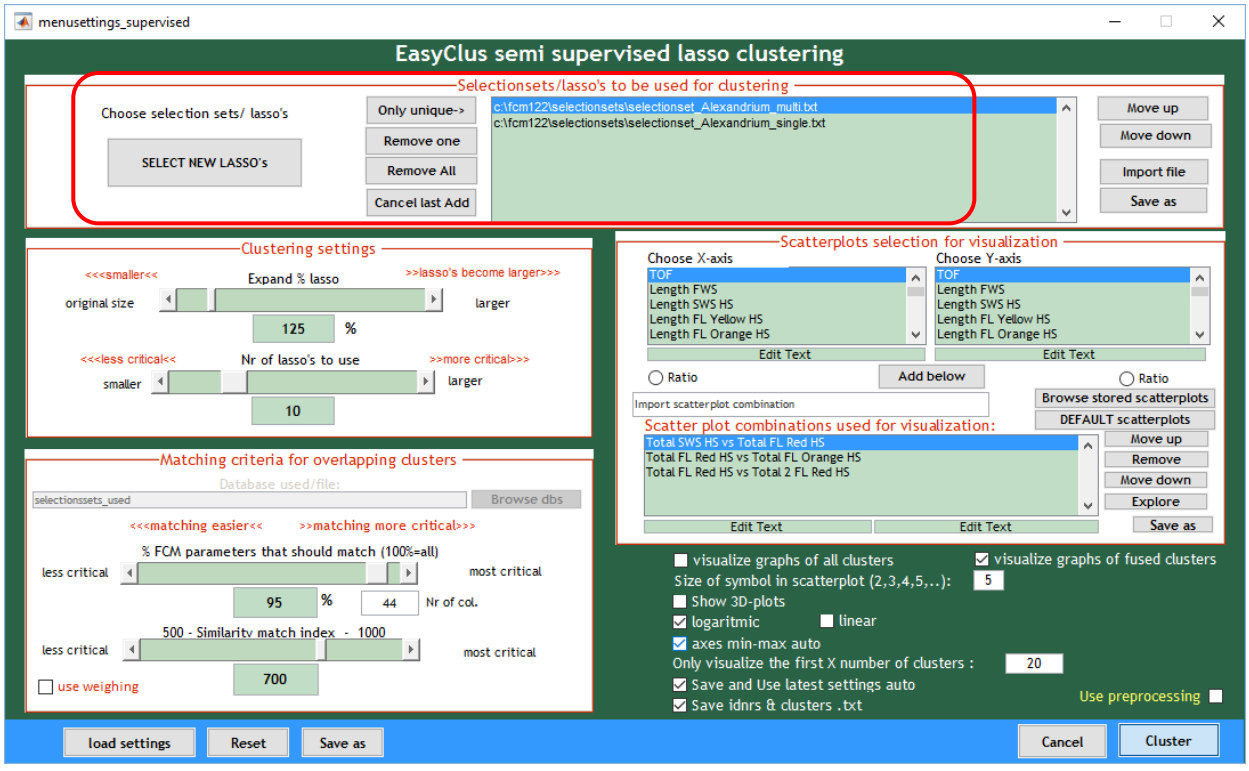
After importing a flowcytometric datafile, for instance a cyz- data flowcytometer file, the autolasso settings menu appears on your screen. These settings can be changed to improve the results.

Either start Lasso clustering by selecting a phytodatabaseXX.mat



(The phytodatabase.mat selections sets are calculated from the database and this process is repeated for each file, which is in fact unnecessary, because the database has not changed. Although this takes only a few seconds, it is better that this is fixed in a next EasyClus release to save time.)

Either start Lasso clustering by selecting selectionsets



Select the selection sets or lasso's that you want to use by pressing the button and selecting the selection sets. The requested selections sets are default stored in the c:\fcm\selectionset directory. Be aware that only the selections sets names starting with the name mentioned in the defaults setting menu will be shown. Press show all files to 'show all files'. Press 'done' to validate the selections.

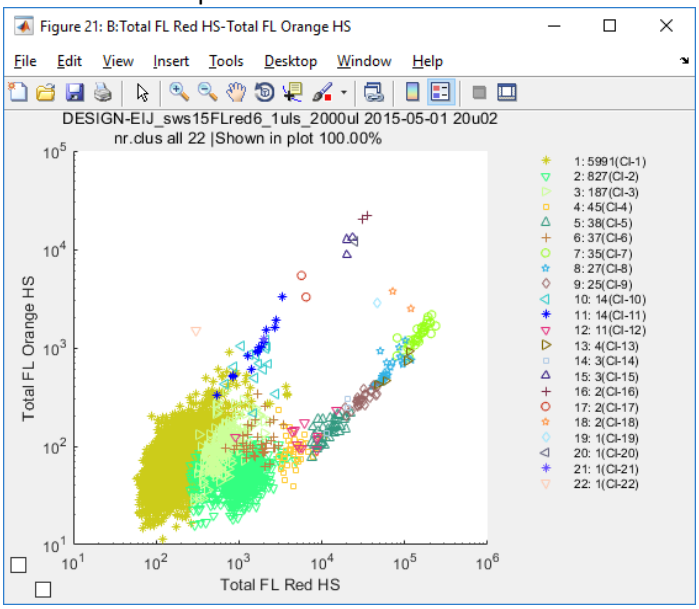
- Expand factor: is the value to expand the predefined lasso's or selections sets in any direction. Expanding 125% means expanding the selections sets to 1250% of the stored original selection set. Expanding to a certain extent is used to avoid exclusion of events that are outside selections sets due to variation of the instrument or due to some biological variation.
- Nr of selections sets used for clustering: is the number of stored selections sets that are used to select events (that should fit in each of these selections sets) out of raw data. Increasing the number means that you use more selections sets and thus meaning that events should fit in all selection sets before assigning them to this selection set. Increasing this number means 'being more critical'.
- Nr of columns: The assignment of events for overlapping clusters (events fits in selection sets A as well in B) is performed using the fingerprint matching criteria. The overlapping events are assigned to the selection set, which has the highest similarity matching index. The number of columns used here is the minimum number of parameters needed for positive matching with species data stored in the database. Increase means being more critical.
- Similarity minimum: the same as before. Minimum value needed for the matching of the overlapping events. Increase means more critical. If too critical, the assignment cannot be done and the overlapping events are assigned to 'not recognized'.
- Scatterplot combinations: Only used for visual representation of clustering results before and after recognition.

Other boxes are the same as under the unsupervised method.

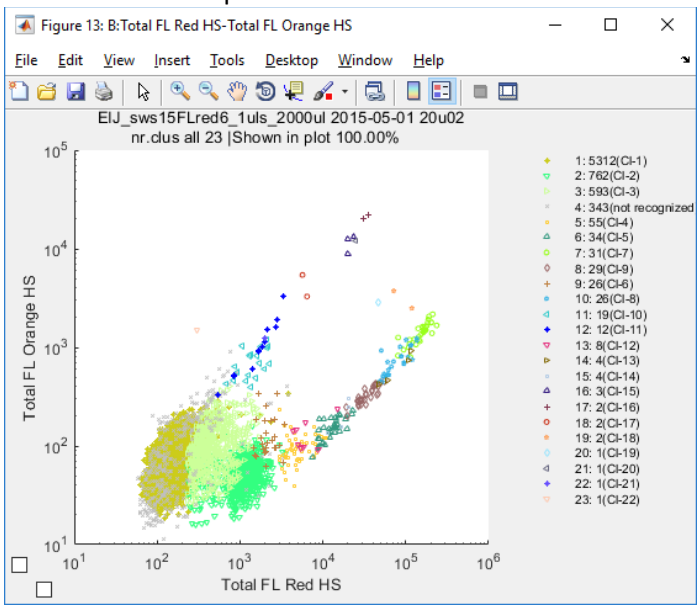
Clusters that are processed by using this 'autocluster semi supervised' lasso-clustering have several output configurations. Al l results are saved as a *.txt file in the c:\fcm\cluster directory.
Output in MATLAB screen:
Sn2: Result after matching each individual cluster.
Sn3: Result after matching each individual AND after cumulating similar species with the highest similarity index (only third column species) .
Sn: Recognition of species on basis of the species dependent used selections sets, WITHOUT using the database matching method . In case of overlapping selections sets of species, more than species will be represented in the output results.

After autolasso clustering, it is possible to repeat the processing using different settings.

Unsupervised DESIGN II



Supervised LASSO



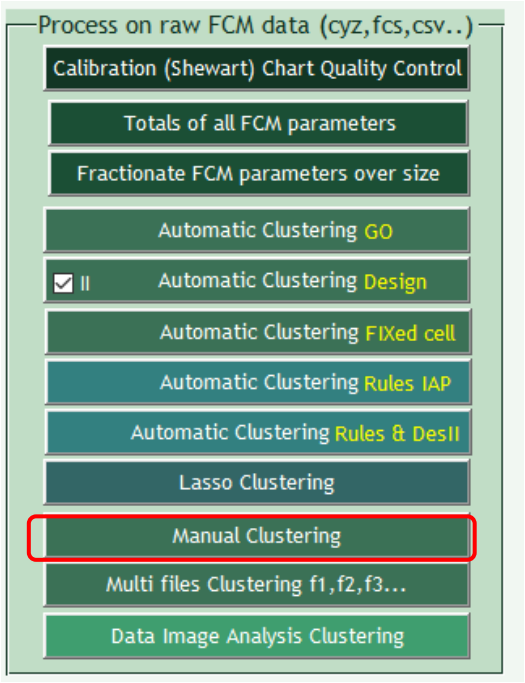
LASSO clustering based on (200 events from each unsupervised cluster) phytodatabaseXX.mat.

Clusters meth 1	Nr ev.	Clusters meth 2	Nr ev.	%match
1	5991	1	5301	0.88482726
2	827	2	717	0.86698914
3	187	3	161	0.86096257
4	45	5	43	0.95555556
5	38	6	31	0.81578946
6	37	9	20	0.54054052
7	35	7	31	0.88571429
8	27	10	22	0.81481481
9	25	8	24	0.95999998
10	14	11	12	0.85714287
11	14	12	12	0.85714287
12	11	13	8	0.72727275
13	4	14	4	1
14	3	15	3	1
15	3	16	3	1
16	2	17	2	1
17	2	18	2	1
18	2	19	2	1
19	1	20	1	1
20	1	21	1	1
21	1	22	1	1
22	1	23	1	1

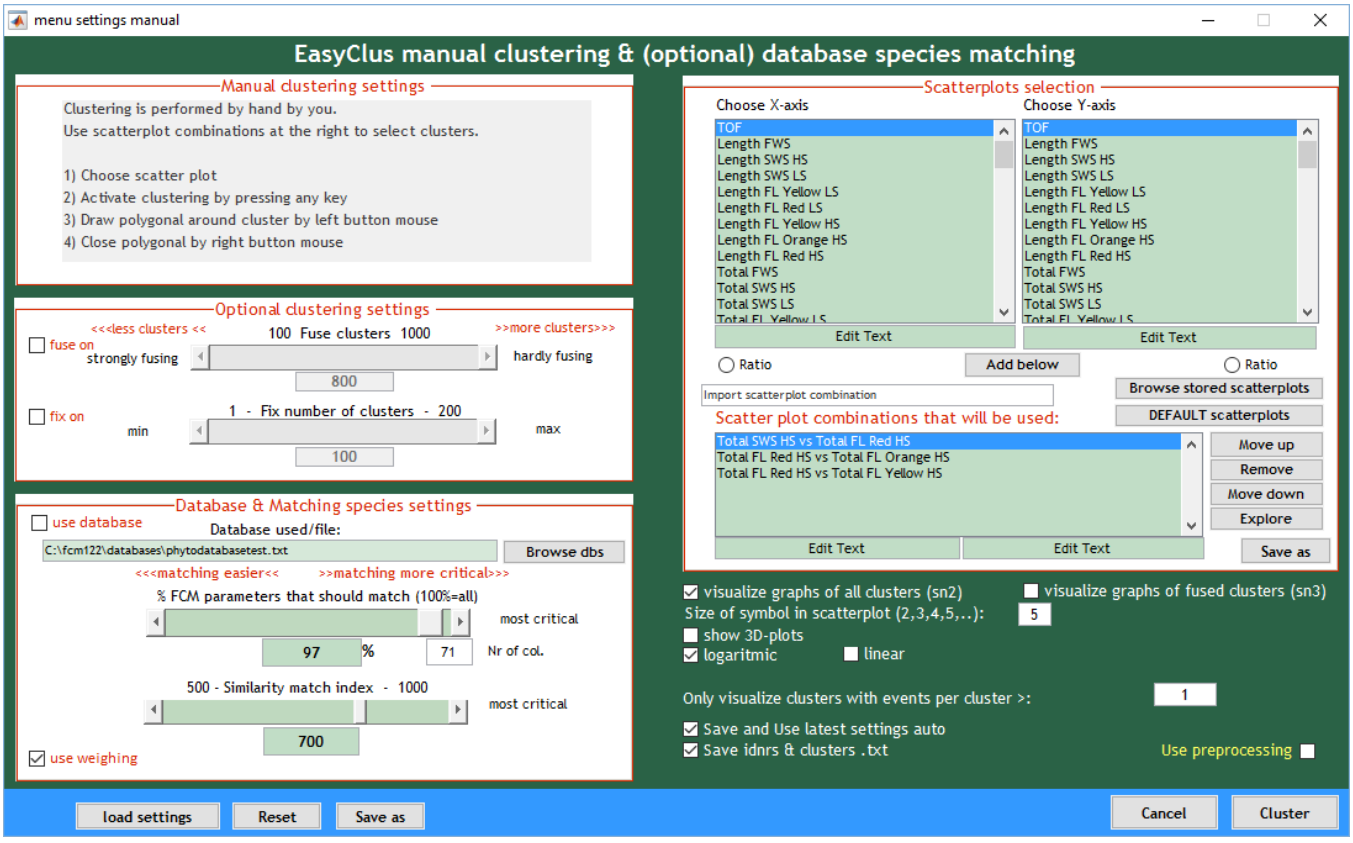
Particles found in unsupervised and found by LASSO method. %match in last column show good results.

3.5.7 Clustering by hand - button MANUAL CLUSTERING

To manually select clusters (to cluster, to compare cluster results with a database, to store events of a cluster separately, to use cluster results and add them to the database)

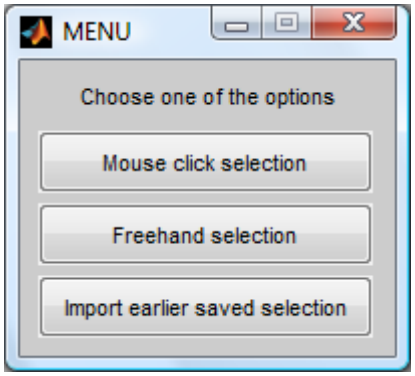


After activation, a fcm data file should be loaded. After loading the file, the scatter plots combinations should be selected for manual cluster selection. This may be previously used scatterplot combinations or new chosen combinations.

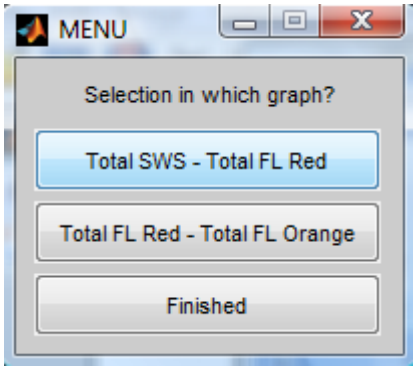


Press Cluster to start clustering.

The Scatter plots will appear and a selection method menu.

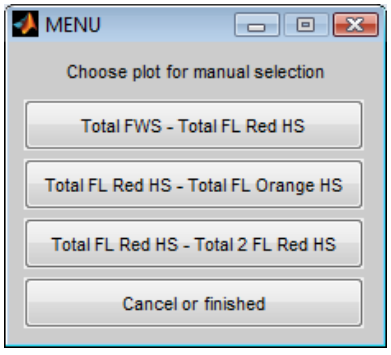


Confirm which selection method you use (by Mouse click, by freehand OR by earlier stored MANUAL cluster selections):



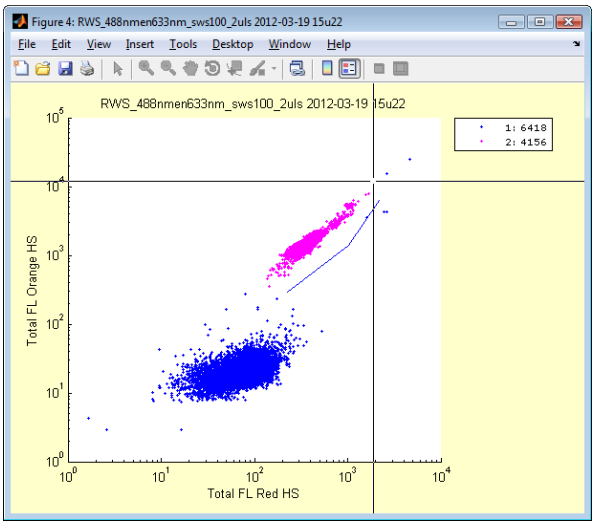
Select which Scatterplot (in this example 2 scatterplot combinations are chosen) to use to start clustering…

Continue and Start Processing:

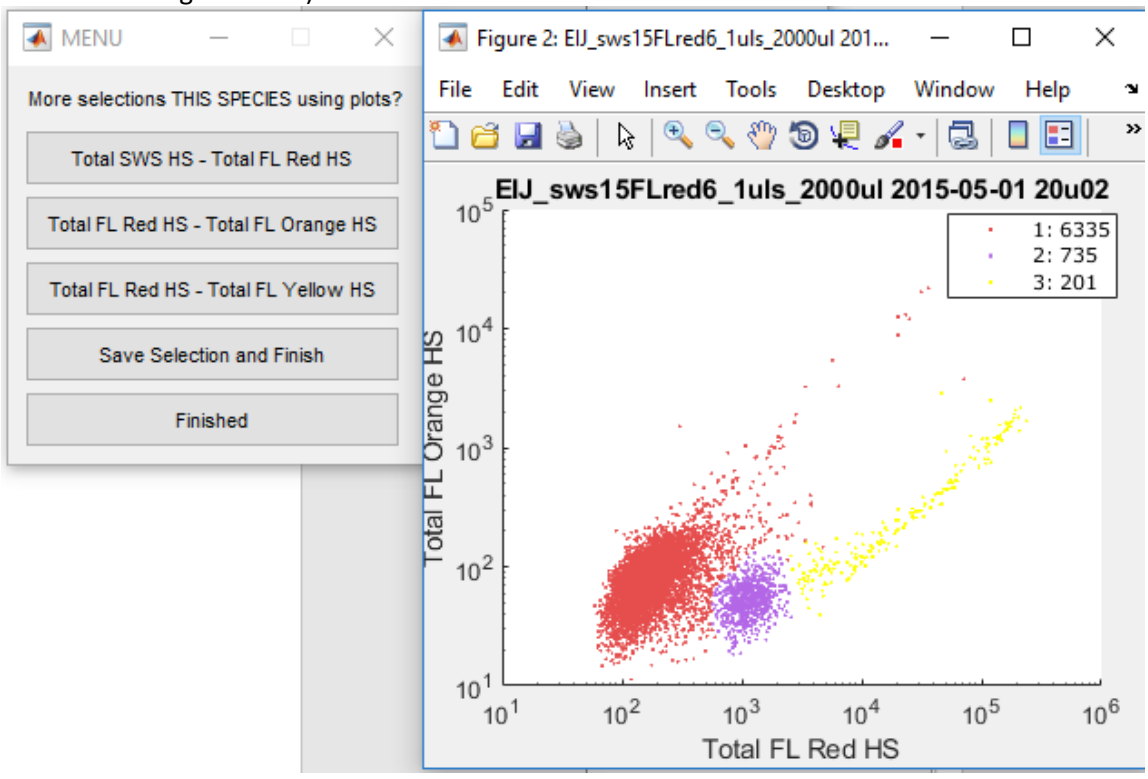


Selection of a cluster is done by choosing one scatter plot: To do so, click the requested scatterplot in the menu:

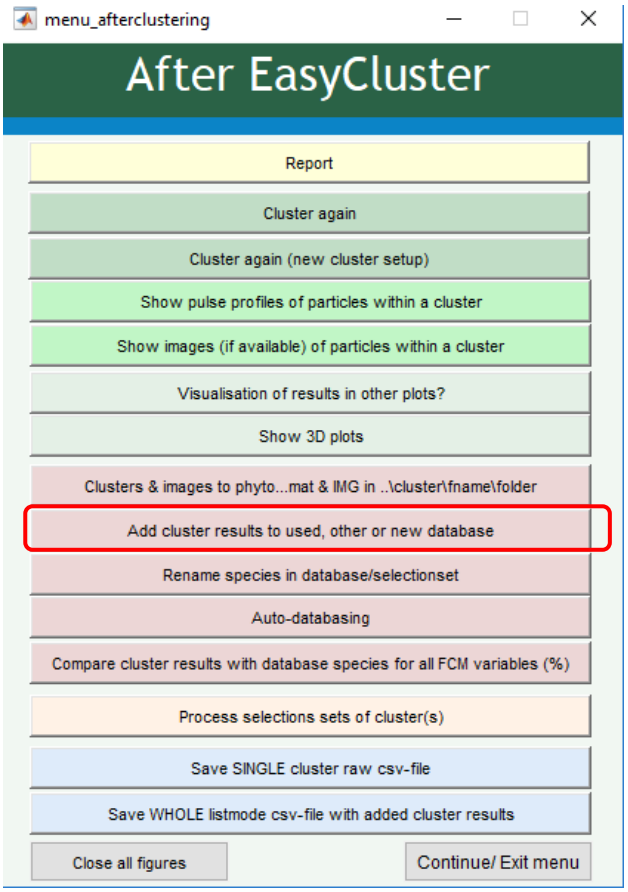
Use the left mouse button to draw your selection. Use the right mouse button to close the lasso.



Repeat the procedure for each manual selection. If you have made one or more selections in one or more scatterplots and you're finished press 'Cancel or finished'.
If the selections are completed, all selected clusters are shown. Overlapping selections appear as a new cluster with a new color.
The events and median values of each of the clusters of each FCM variable are calculated and stored (.. \ cluster), the scatter plots are stored as jpg files (.. \ cluster \ figure).
After drawing clusters, the following menu appears (this menu is also used in other clustering methods):



Save Selection and Finish will store your selections sets as drawn by hand for each cluster, which can be used (later) in the LASSO cluster method.
Better is to store the particles within the clusters in a phytodatabaseXX.mat, so it can be edited by the database editor and used in all other clustering methods.



Button '**Cluster again'**
Use the same selected clusters but change settings for instance of the database.

Button: **Cluster again new cluster setup'**
To start all over again with clustering. Previous selections will be erased.

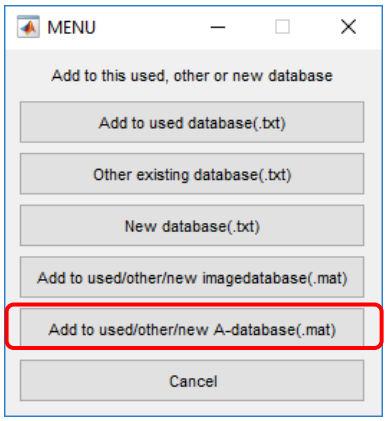
Button: Only .cyz or .mat (from cyz) files CytoSense '**Show pulse profiles of particles within a cluster'**
To select clusters and visualize the detector signals of events within a cluster.

Button: Only .cyz files or .mat (from cyz) CytoSense '**Show images (if available) of particles within a cluster'**
To select clusters and visualize the images of events within a cluster.

Button: '**Visualisation of results in other plots'**
Draw more scatterplots without changing the clustering results.

Button: '**Show 3D plot'**
Draw density and 3d plots.

Button: '**Add cluster result to used, other or new database'**
To add selected cluster(s) to a stored or new phytodatabaseXX.mat by shown option below.

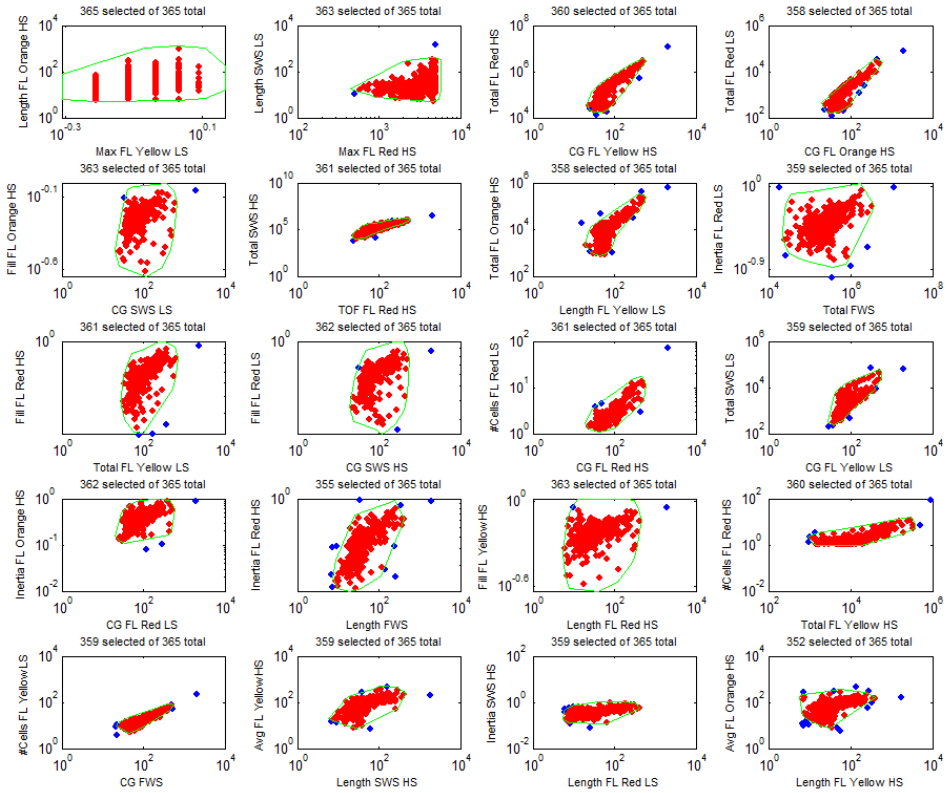


Button '**Rename species databases/ selectionsets**
To rename species in database (better to do this in database editor) or selectionsets.

Button '**Auto-databasing'**
To store automatically all clusters in a new database.**txt**. After choosing this option, some settings need to be confirmed or changed, that are used to match if clusters have mutual comparison to a certain degree.

Button: '**Compare cluster results with database …**'
To compare cluster results mutually or with database species.

Button **'Process selectionsets of a cluster'**
A lot of automatically generated selectionsets in several scatterplots are produced and can be stored on basis of cluster result.



Button '**Save SINGLE cluster events in raw csv-file'**
The events of a cluster can separately be saved as a txt/csv file (in .. \ data files). This can be useful for checking.

Button '**Save whole listmode csv-file with clustering results'**
The whole raw file is saved including the clustering results as well as the indexnumber.

Button: **'Close all figures'**
To close all figures e.g. scatterplots that are processed by EasyClus

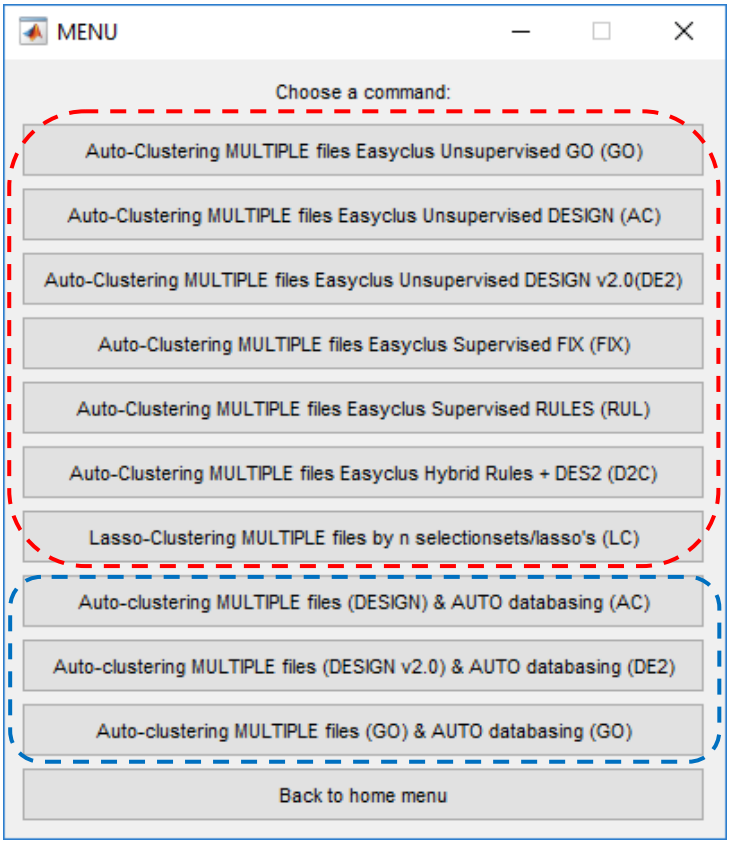
Button: '**Continue/ Exit method'**
To stop the latest clustering method and leave to the EasyClus home menu.

The results are stored in the same txt file (.. \ cluster) and scatter plots are stored as jpg files (.. \ cluster \ figure).

3.5.8. Auto-clustering (multiple files) – button – Cluster multiple files

To cluster automatically more than one file but several files in a directory one after each using similar settings

If you have several files to process and you want to use the same clustering settings for each file, this option allows you to process several files automatically. This option imports and 'auto' clusters files automatically just after each other and can save a lot of time especially if you have quite a lot of files. For the unsupervised methods GO, DESIGN I and DESIGN II, there is the combination of autoclustering and adding of (only) unique) clusters to a database.txt (only attributes) automatically.

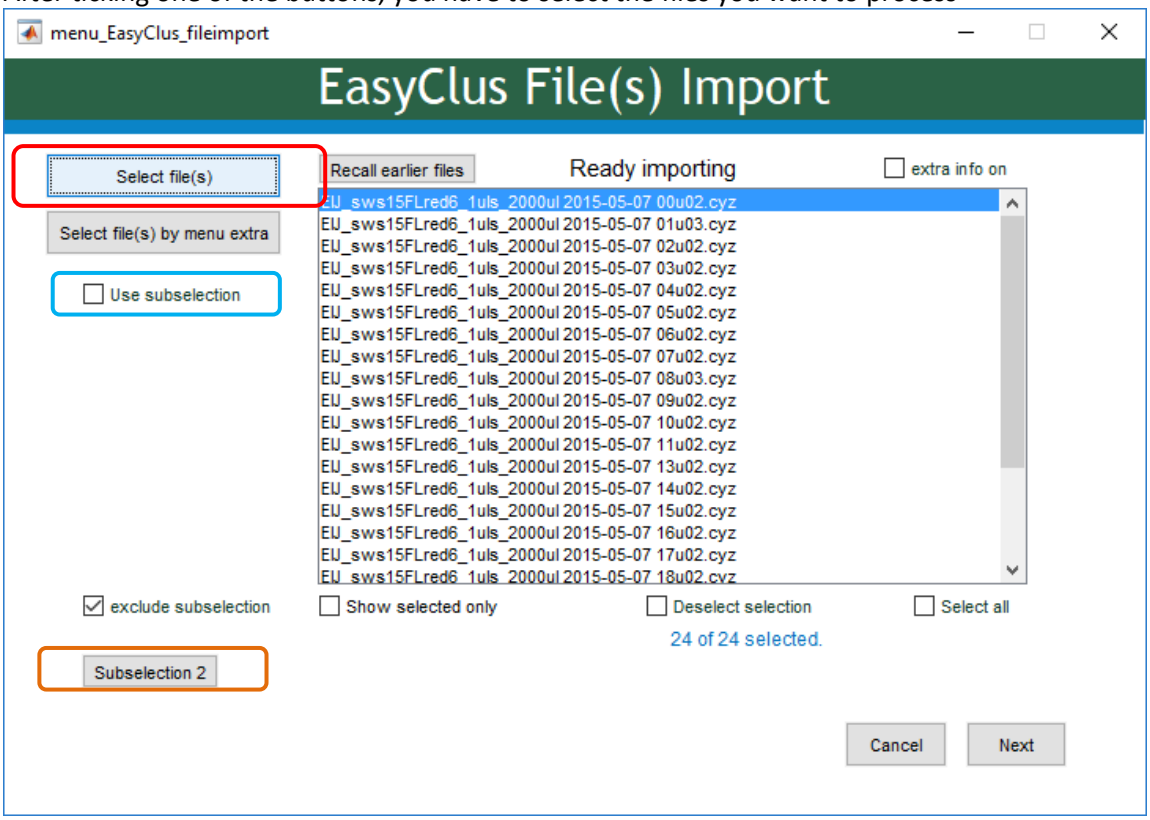


After choosing the clustering method, a menu appears which kind of files you want to process. (Call Thomas Rutten Projects if other file formats are needed here)

Choose the files you would like to process using previous window. You can select filter, sort or specify all the data files using the next window.

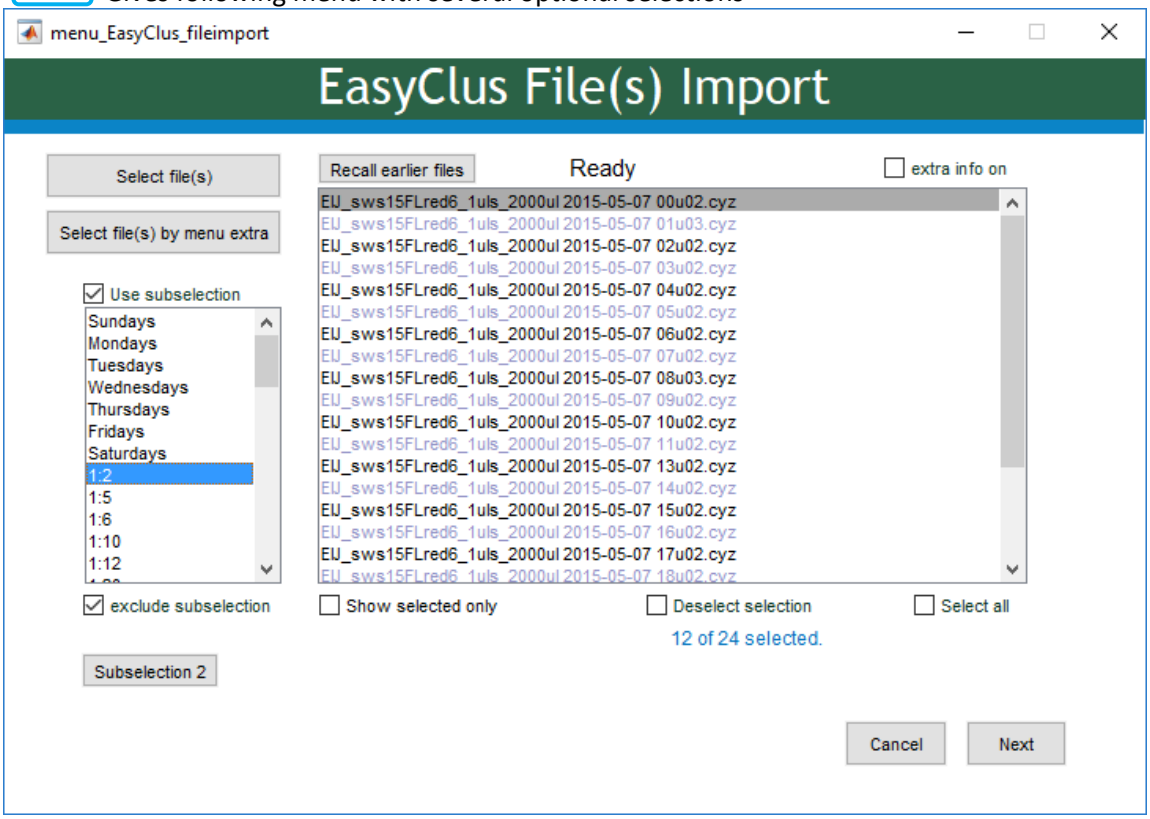
3.5.8.1 Multiple clustering

After ticking one of the buttons, you have to select the files you want to process

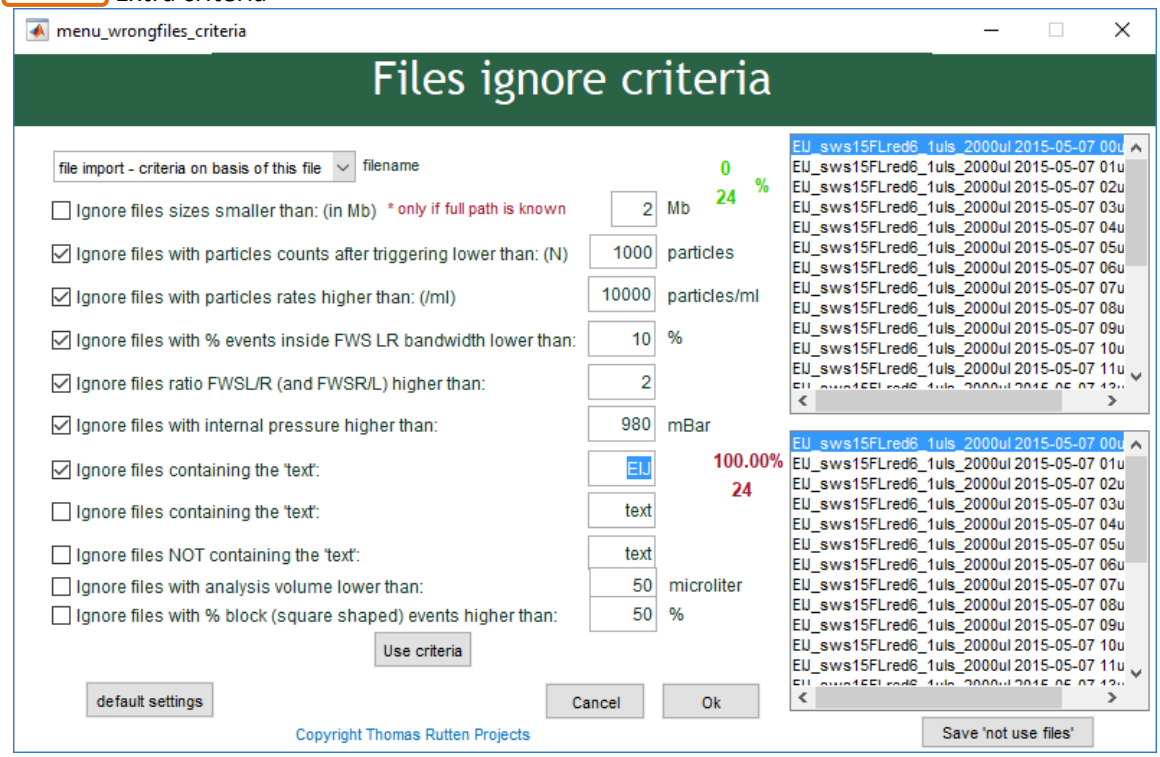


From this menu you have several subselection possibilities by the use subselection checkbox and/or Subselection 2 button

Gives following menu with several optional selections



Extra criteria

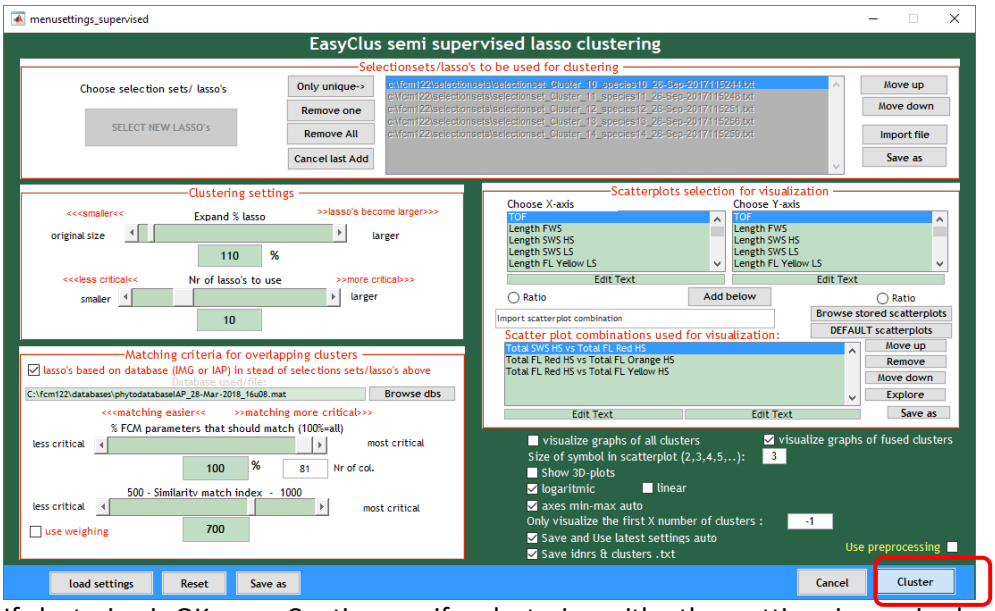


Multiple files clustering is can be done by :

- unsupervised autoclustering GO
- unsupervised autoclustering DESIGN I
- unsupervised autoclustering DESIGN II
- supervised autoclustering FIX
- supervised autoclustering RULES
- supervised autoclustering RULES & HYBRID (not yet implemented)
- supervised autoclustering LASSO

The processing from this point is similar as described in 3.4.1 to 3.4.6

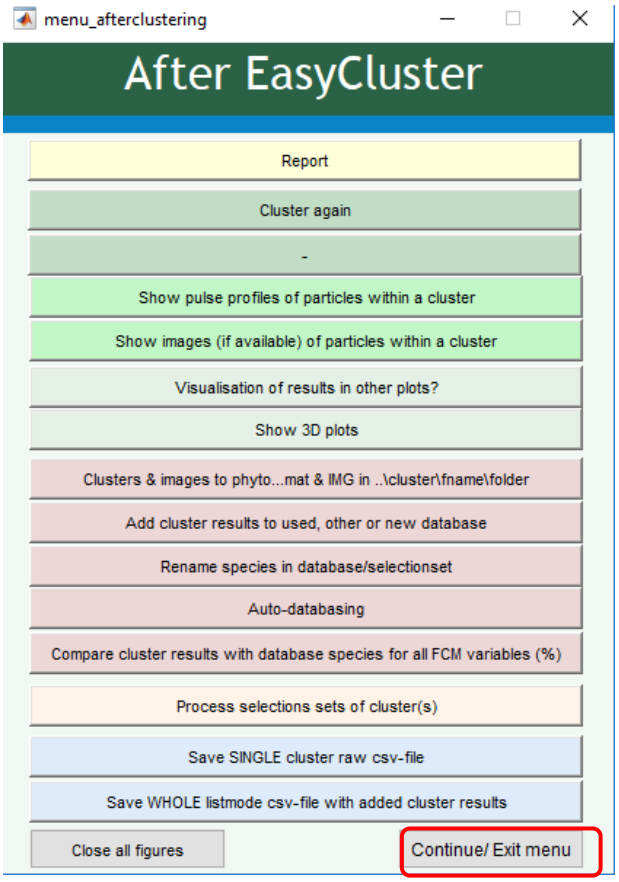
After that, the clustering default settings menu appears once, just before processing the first file.
Example by supervised autoclustering LASSO:

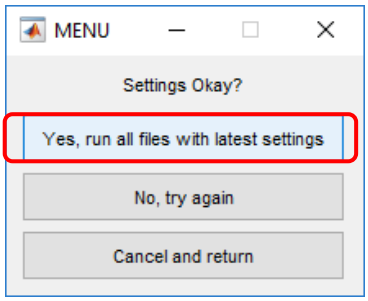


If clustering is OK press Continue or if reclustering with other settings is required press Cluster again

Cluster results before and after matching with the given database, are represented in scatterplots and stored in the ..\cluster directory as well as the scatterplots, which are saved as jpg-files in the ..\cluster\figure\ directory.

Results of the autolasso option, are stored for each individual species in one file in order of processed filename.





If the procedure is finished, the 'EasyClus' home menu appears in the Matlab screen.

3.5.8.2 Multiple clustering with database

The same as the first option but now an auto database.txt (only attributes) is produced. Each sample is (unsupervised) clustered and all clusters are automatically put in a database (named cluster_1, cluster_2 etc.) according to the given settings. E.g. field sample 1: 20 clusters are found and put to the database.
Field sample 2: 22 clusters are found, let's say that 19 clusters are recognized by the database, thus 3 unique clusters left and they are automatically added to the database too and so on until all field samples e.g. 3, 4, 5, ... 20 are analyzed. At the end the database is filled with all unique species. Repeating the unsupervised clustering for all field samples, followed by classification of the freshly made database gives an overview of the changes of clusters between the samples.

This feature is a three step procedure to unsupervised clustering (method 3b) all selected data files, followed by auto-databasing of all unique clusters and finally repeating the (unsupervising) clustering combined with database recognition. Recognition is done on basis of the new database that is built in step 2. Cluster names are like 'cluster_1', 'cluster_2' etc. and each cluster has its unique symbol and color for easy comparison between scatter plots. Also a scatter plot movie is made (..\cluster\figure\ *.avi) starting with file1, which shows the changing of clusters in the files.

The output results are stored in ..\cluster*.txt.

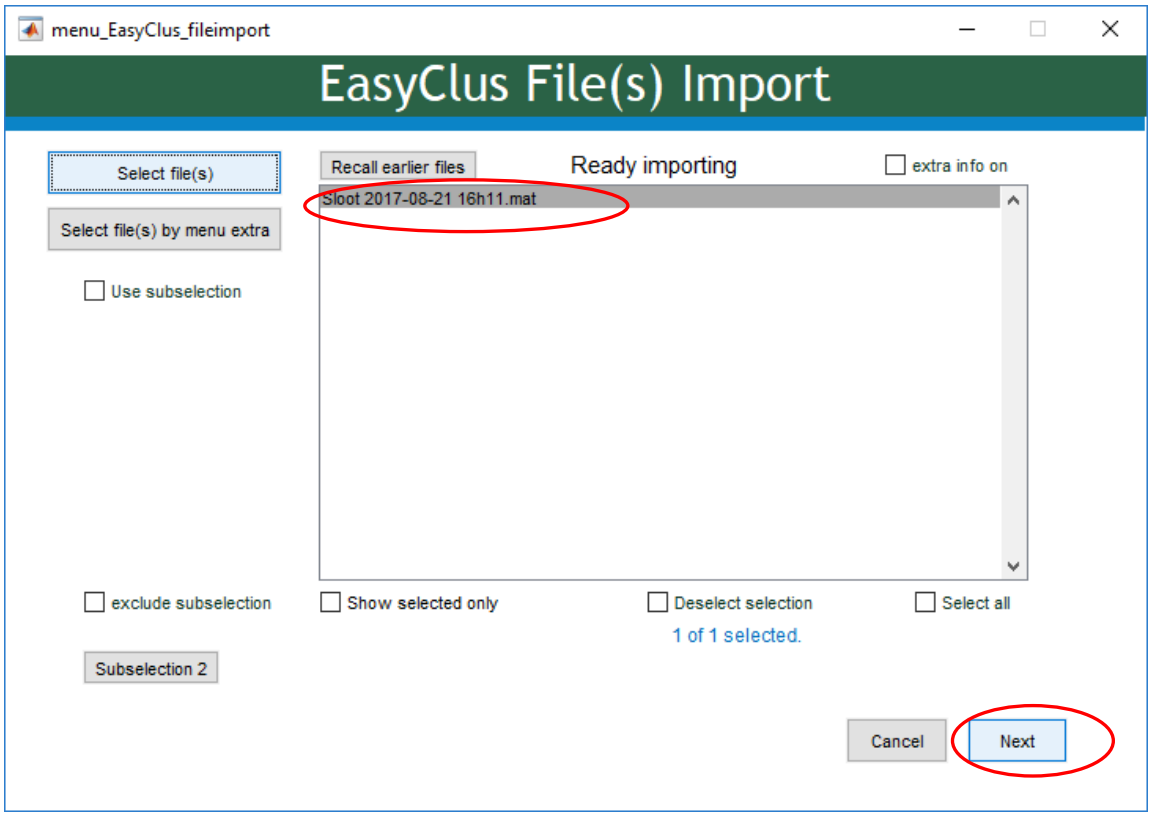
Note: After a new database is automatically created, it might be interesting to visualize this database in scatter plots with the option 'show database in scatter plot' (4th button under 5).

REMARK : In a future release this autodatabase option will automatically create a phytodatabase.mat (attributes, profiles and images) database, possibly by creating a big mixed datafile by taking a subsample out of each file first and using this as a blue print for clusters. Point of discussion here is the best strategy to follow.

3.5.9 Data Image Analysis Clustering – button

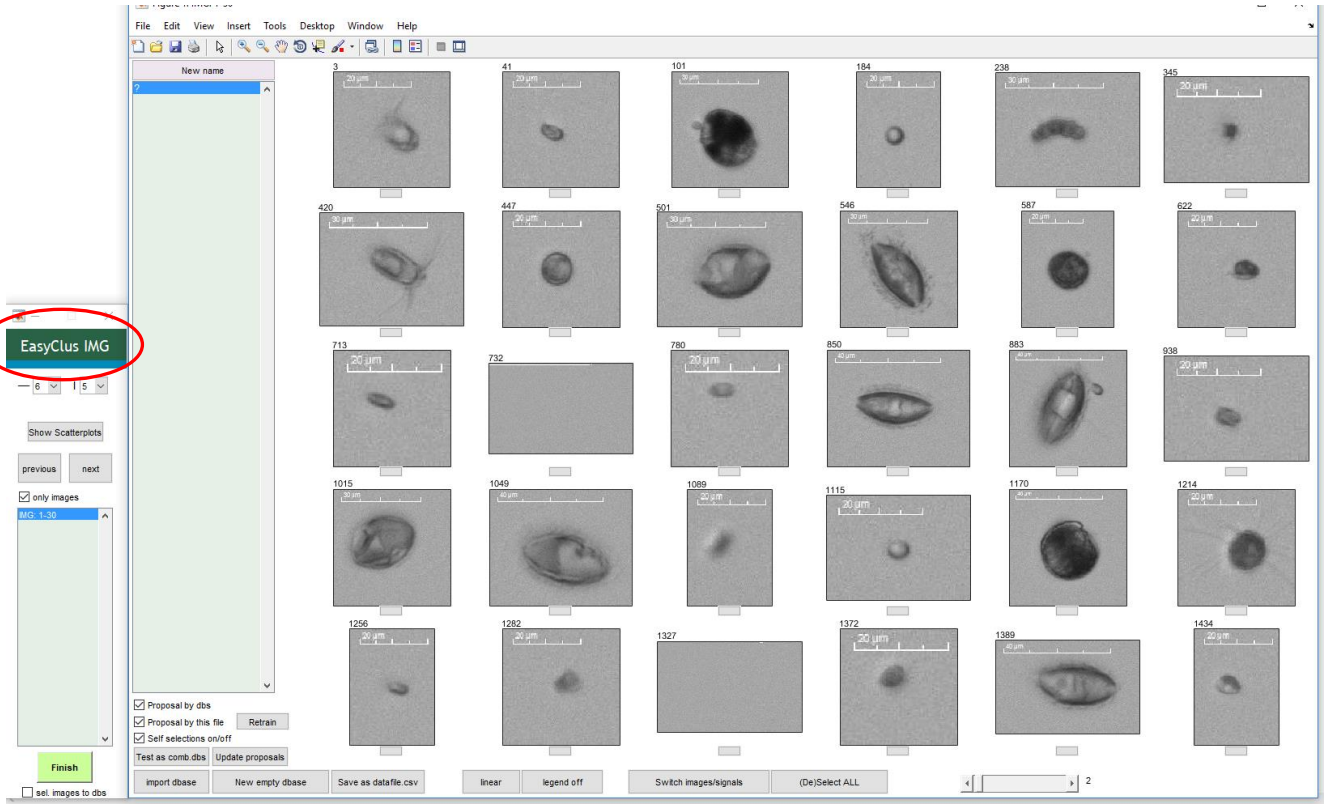
To cluster FCM data on basis of the recognition of images by the operator and successive training of the EasyClus software.

This module starts with showing you the images while you 're giving a name to them and so you are learning the software which images is which species. On the background other FCM-fingerprint data is used to build up automatically a database. During the image classification by the operator, the software tries to give a species proposal, which will take after about assigning five images to the same species. After building up this image based database, all other data (without images) can be classified.

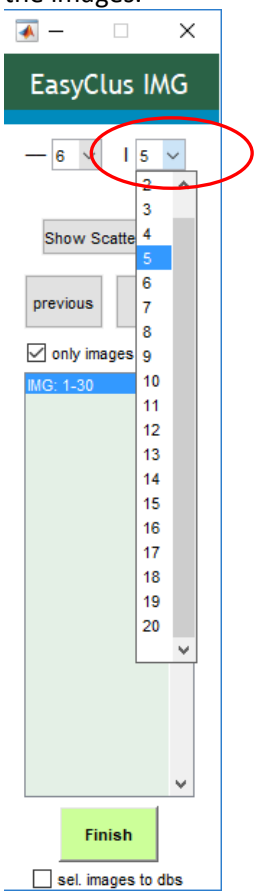


The following menu's appear:
Left the EasyClus IMG file handling menu to switch between pages of images within the file.
The image menu, a page with images (here 30 images) (or signals) of particles available within this file.

Dependent of the size of your screen, you can change the number of images in horizontal and/or vertical direction

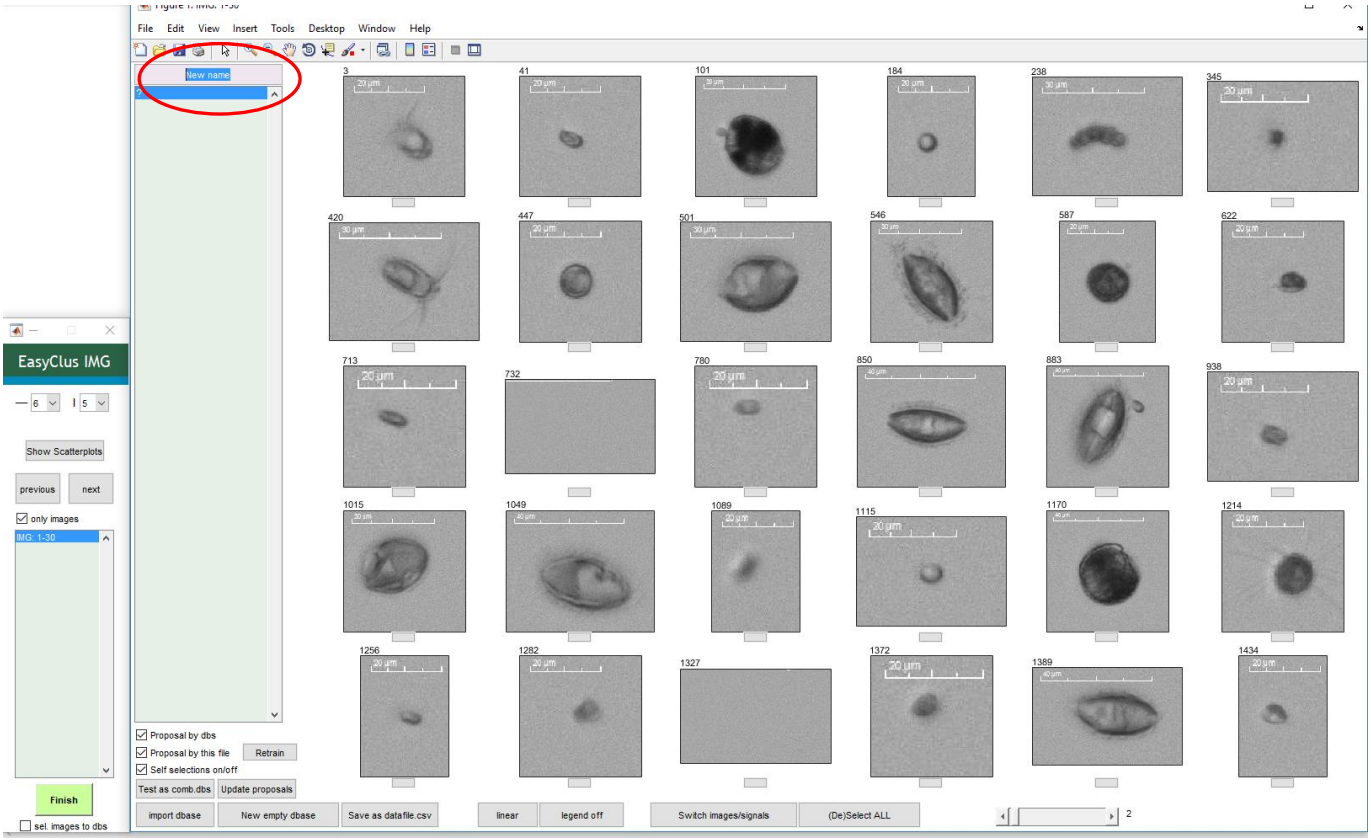


Changing the number of images in horizontal or vertical direction is done in the EasyClus IMG menu. This is just for your own convenience and about the image size you would like to have. Increasing the number of images in horizontal and/or vertical direction decreases the size of the images.

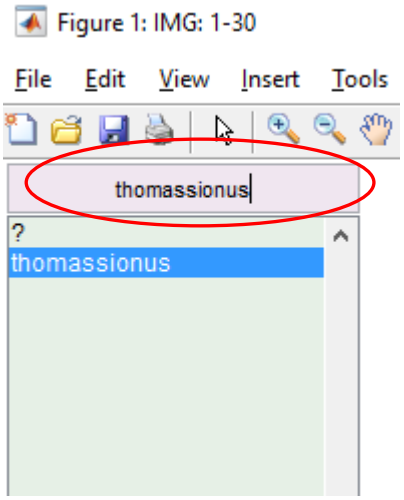


Define a species name in the left upper part (if it is not in the list yet).

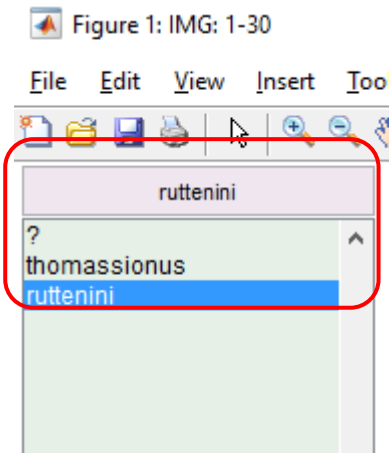
The database does not show any species name, so we start with adding a (recognized) species name to it.



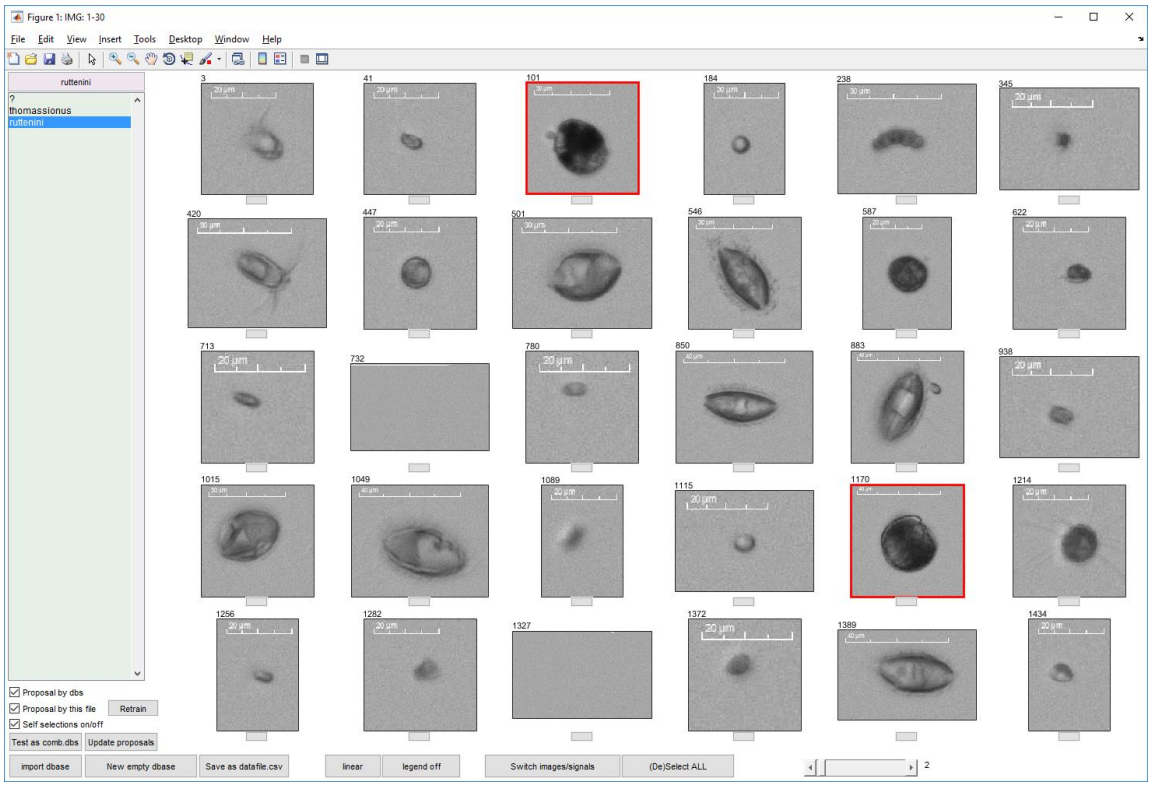
Type a name (here ‘thomassionus’) and press enter



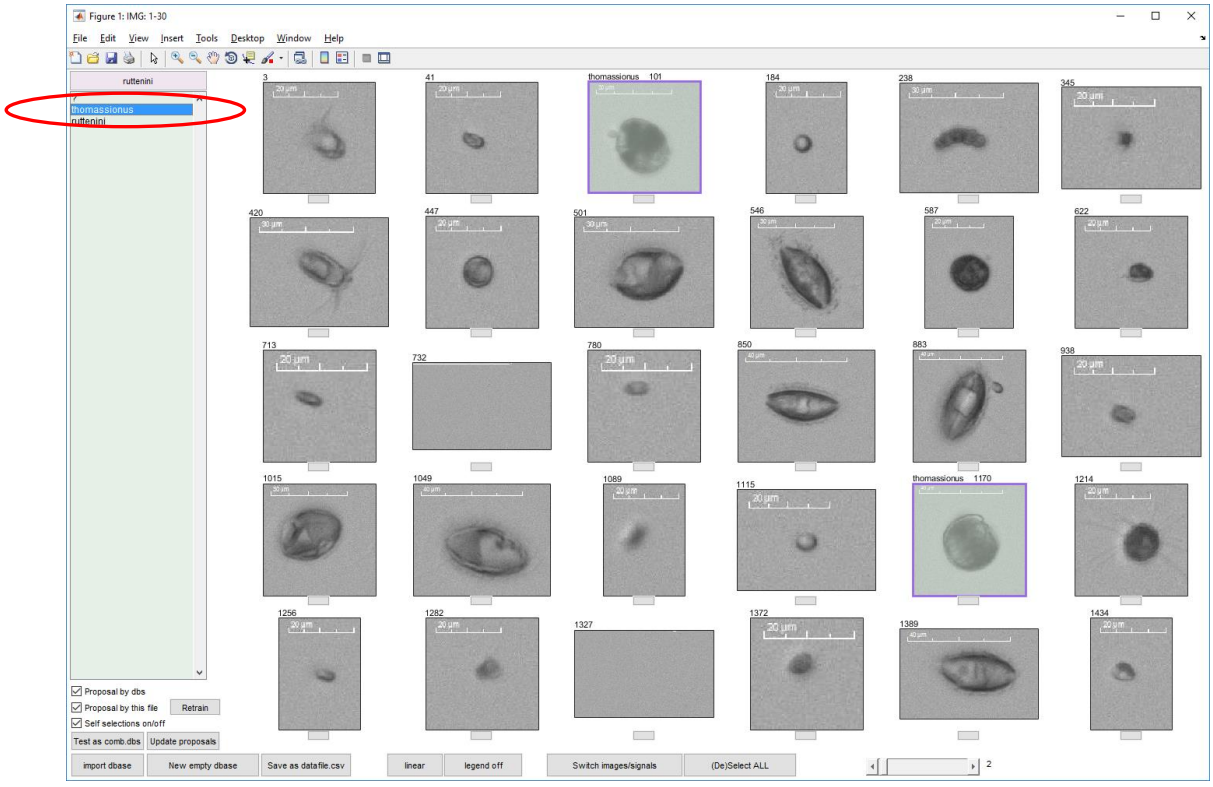
We add second name, ‘ruttenini’



Now we start to with the classifying process of the images.
Selected images are classified by clicking the right name in the database list.
So select the images by clicking the images by the left mouse button so that a red square appears (see below) ...

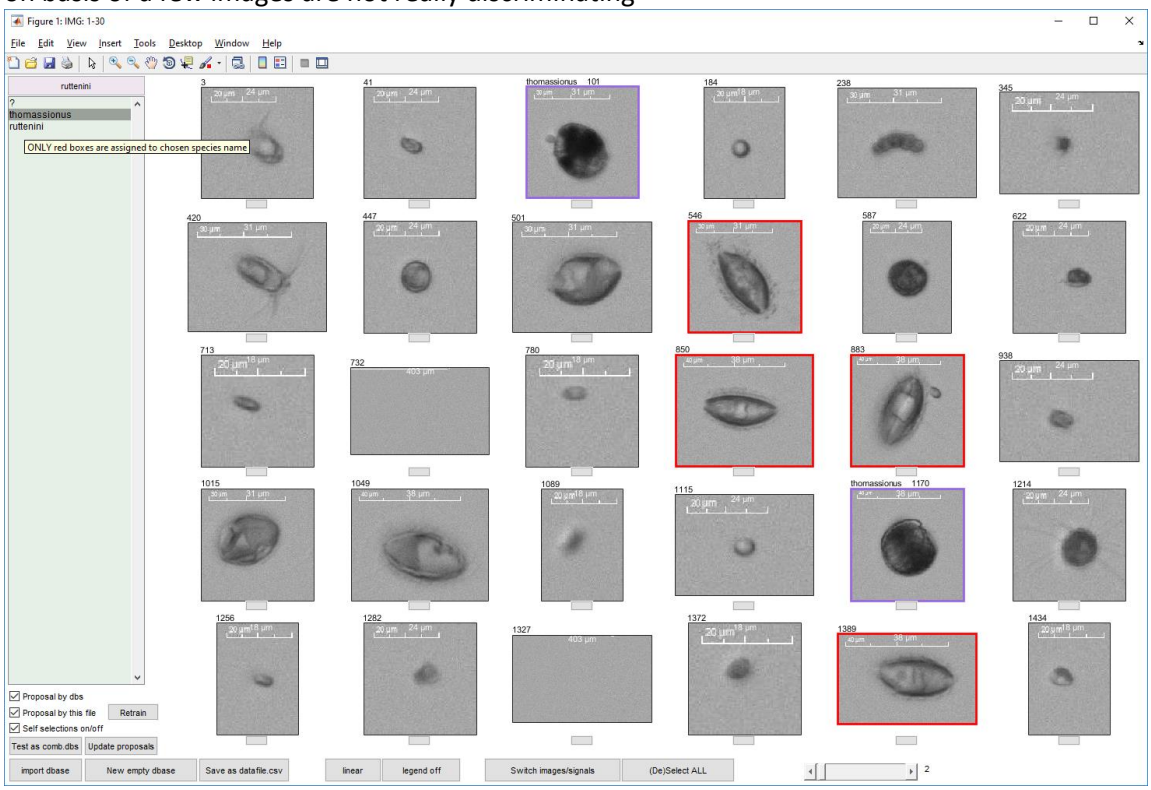


.. and click in the list for the right species. The red square changes to a (specific) other color. These images are assigned to ‘thomassionus’ and change from red boxes to another color (here purple) with the name in the title followed by the id or index number used in the file.

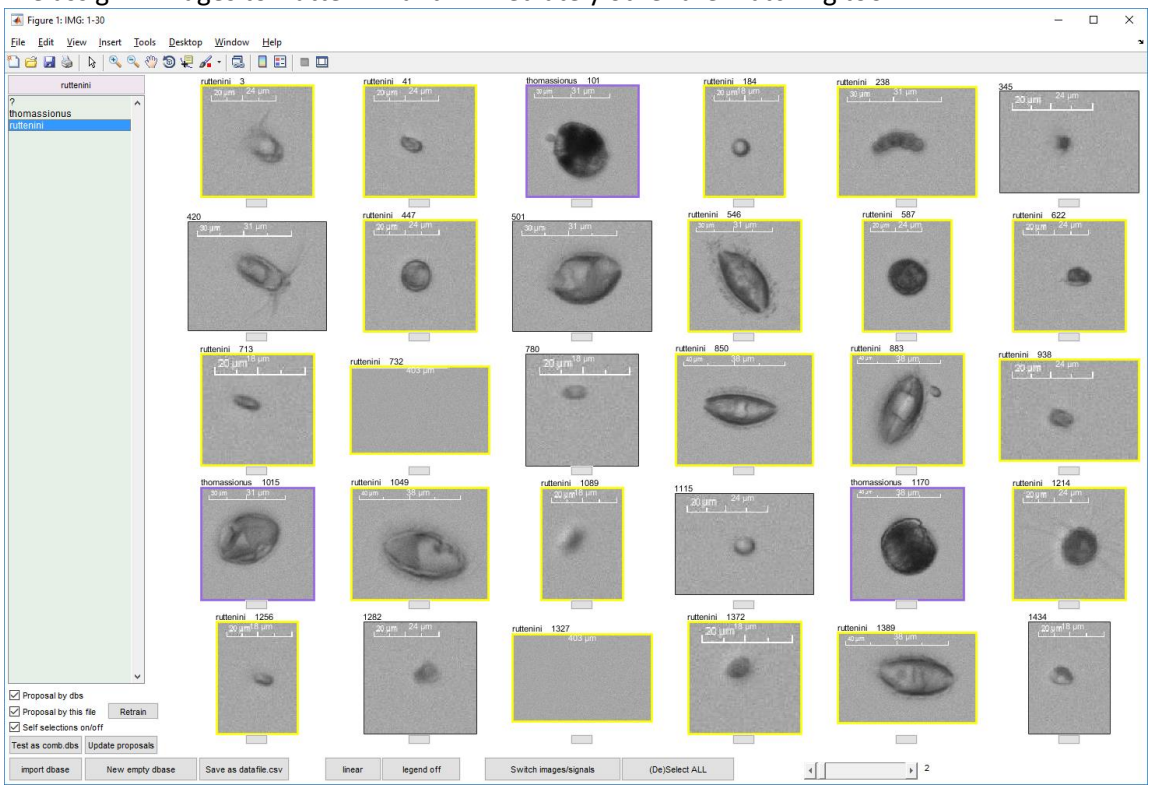


We go on with another species. Select the images by clicking them (red box) followed by clicking in the database list. The (yellow) images are assigned to the second species ‘ruttenini’.

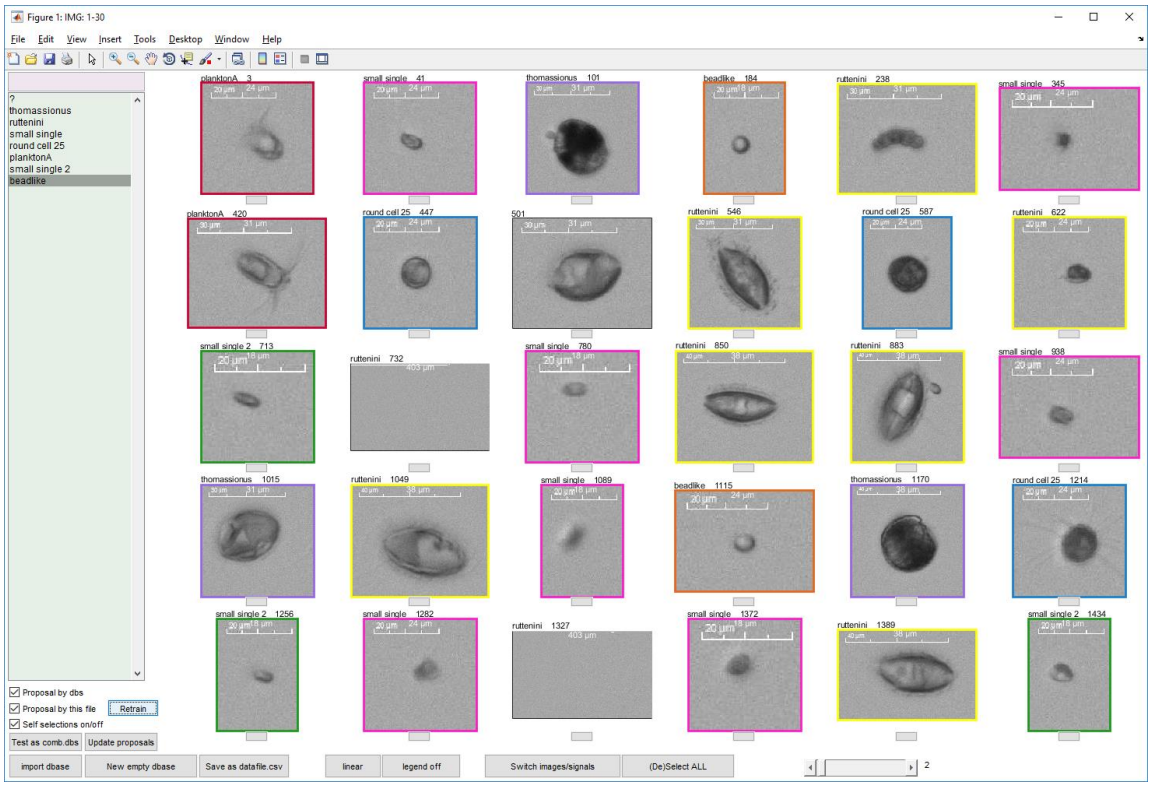
Immediately many other images (which have not been chosen by the author) are assigned to the same species according to the 'on the fly' learning algorithm working in the background. We already that some are not okay but that's only because we have just started and the criteria on basis of a few images are not really discriminating



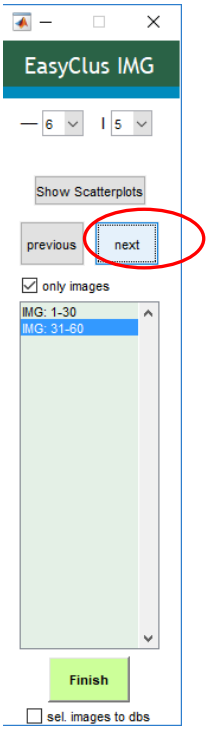
We assign 4 images to 'ruttenini' and immediately other are matching too.



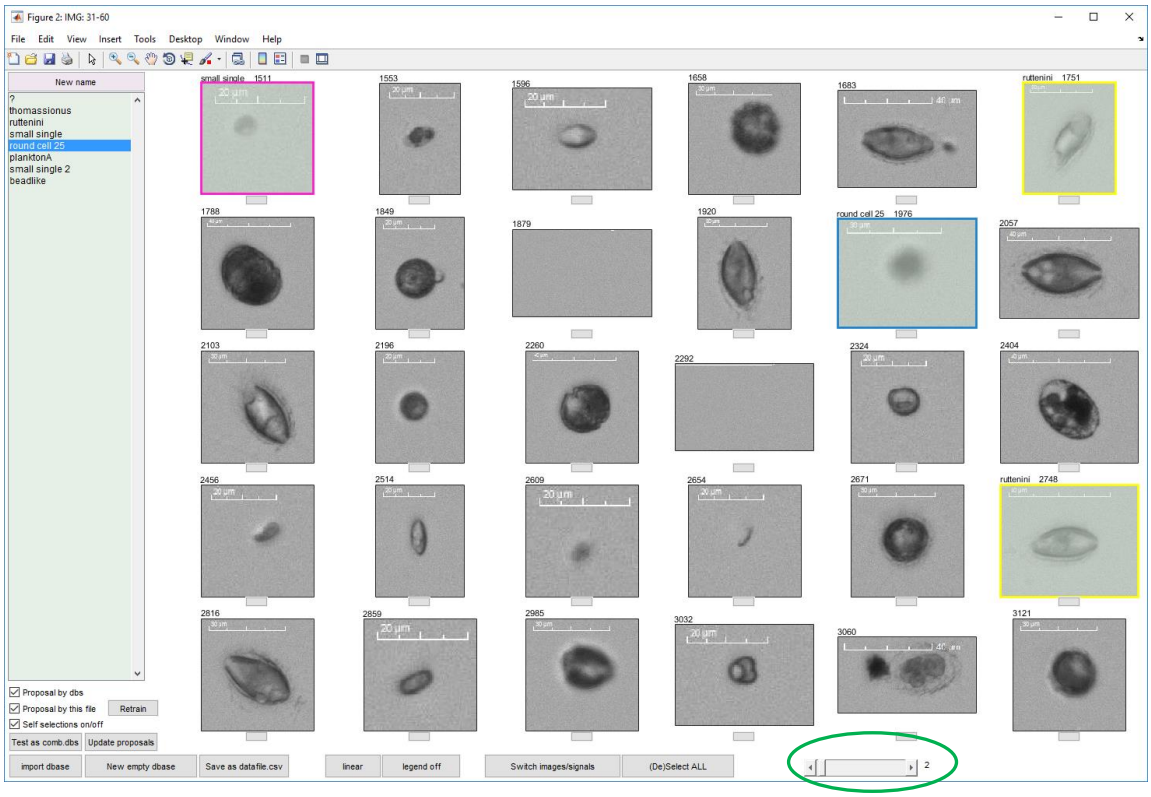
We assign more names to images by repeating the process



Let's see what more images we have in the (cyz) file. Press 'NEXT' in the EasyClus IMG menu positioned at the left on your screen.

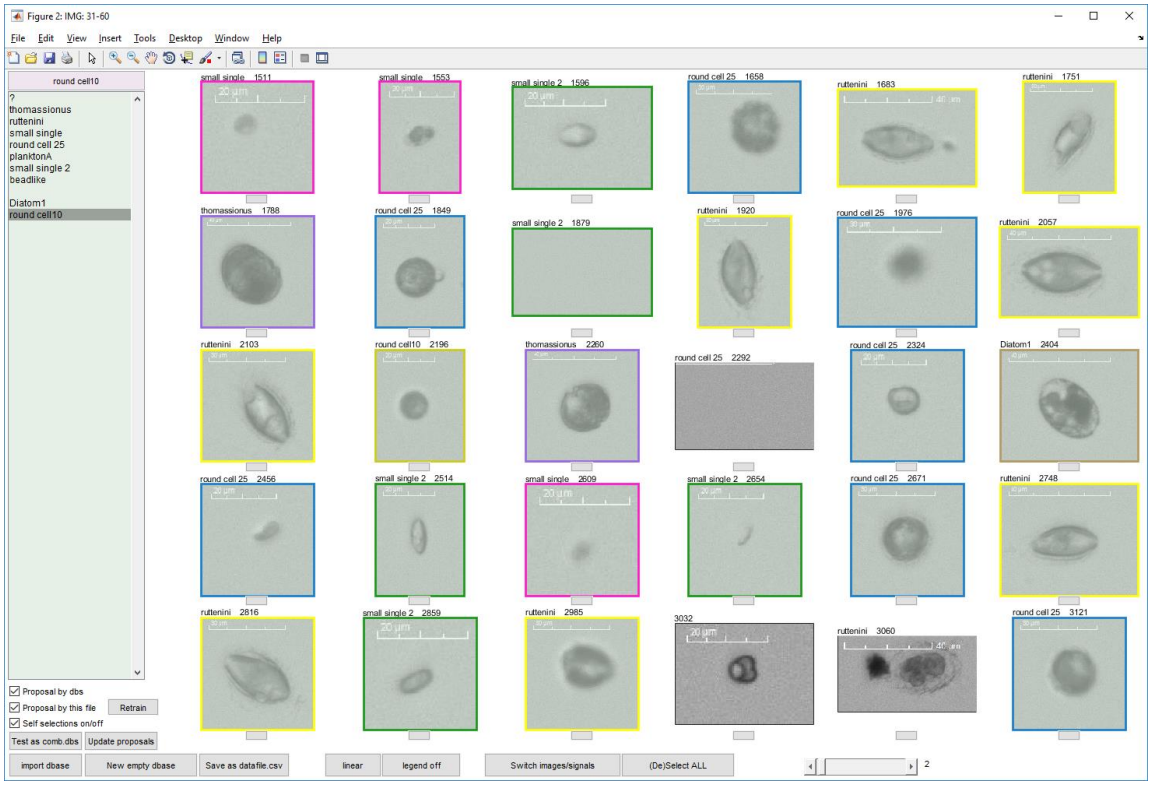


Images from the file are loaded and represented in a second images window. Already ‘on the fly’ matching is done and given in this window. Only four images are being recognized on basis of previous images assignment

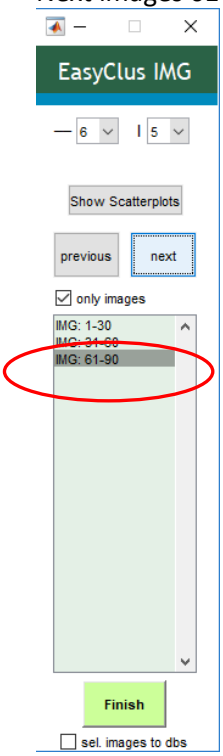


These assignments are only proposals! Be aware that the auto image classifying results might be wrong, they are just based on a few assigned species of your input! The more species you assign, the better the proposal result will be. The assigning procedure becomes more critical by increasing the slider value (shift to the right) in the green encircled area

We assign more images by manual selection in order to train our 'on the fly' learning database better.

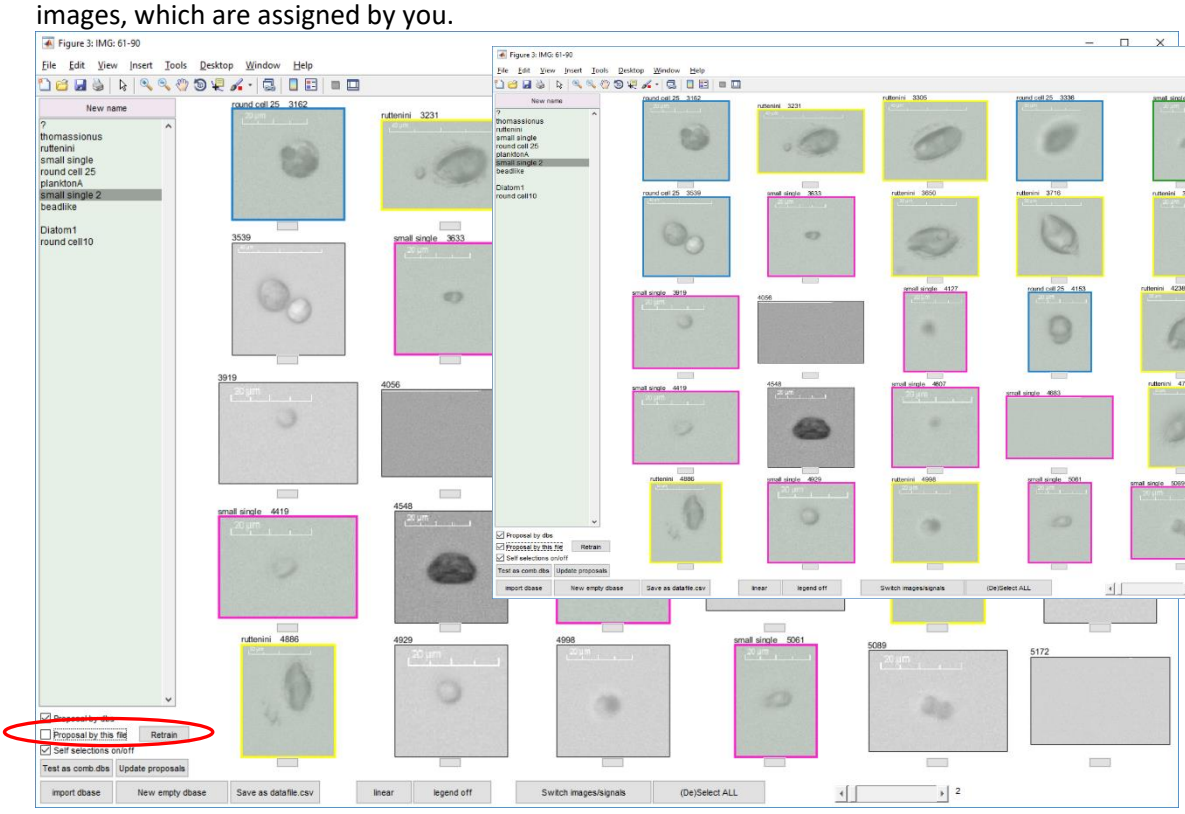


Next images 61-90 are loaded…



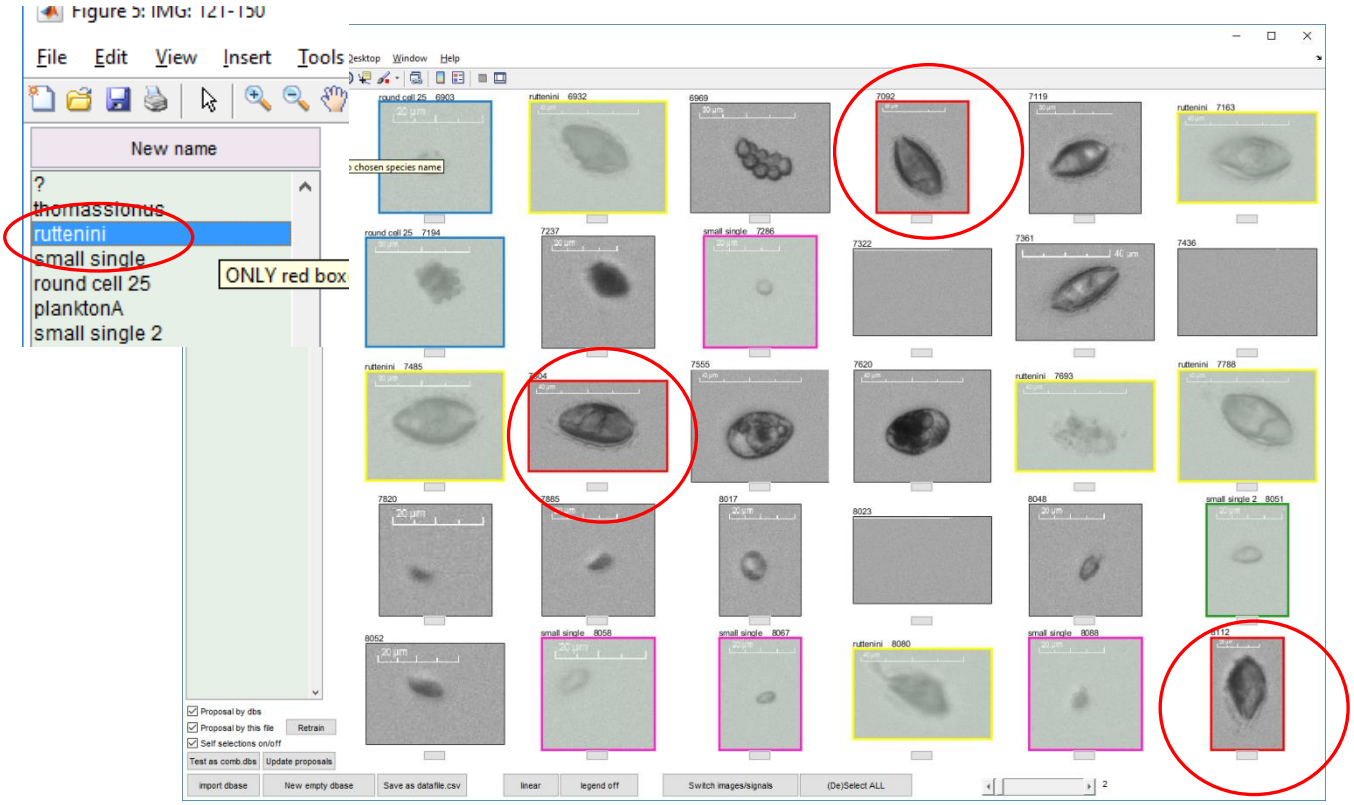
We assign or confirm more images to species names …

We can set the software proposal 'off' to stop automatic assigning and/or to have a look at the images, which are assigned by you.

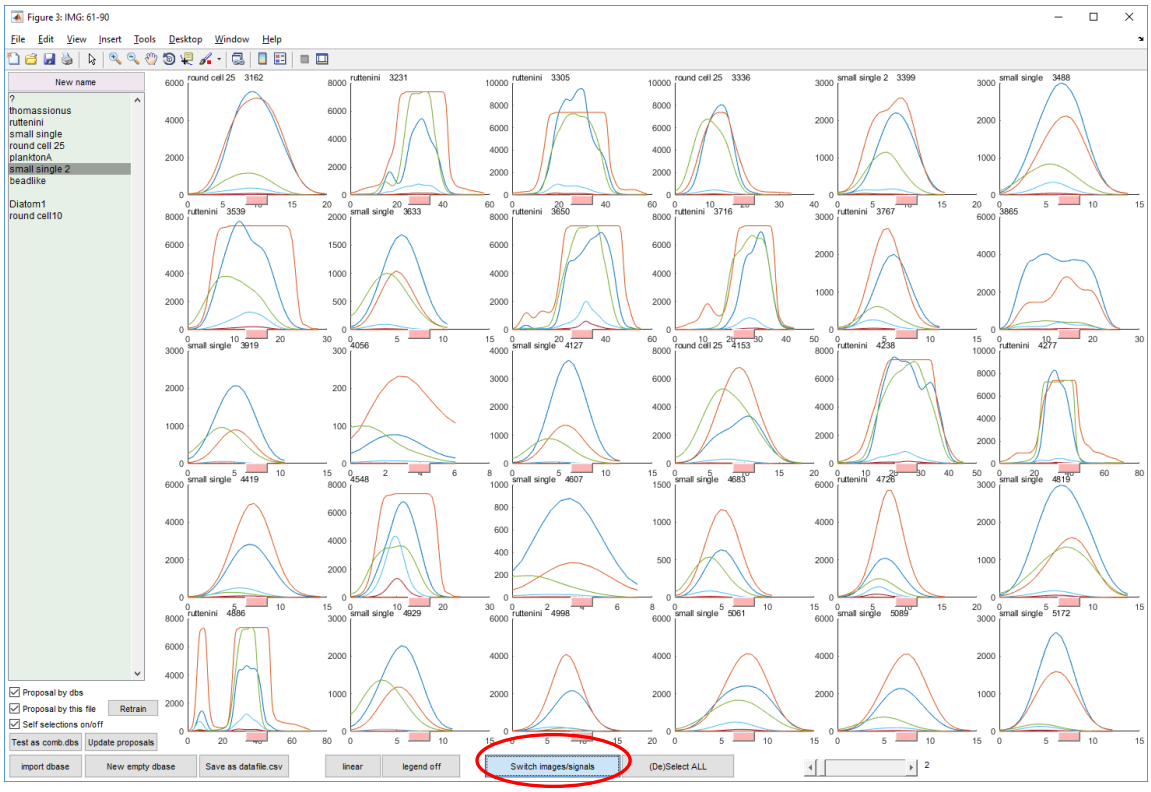


Put the software proposal checkbox to ‘on’ again, to look what the software proposes and rename those images which are wrong. Otherwise they will be stored in the database (afterwards) under the wrong name.

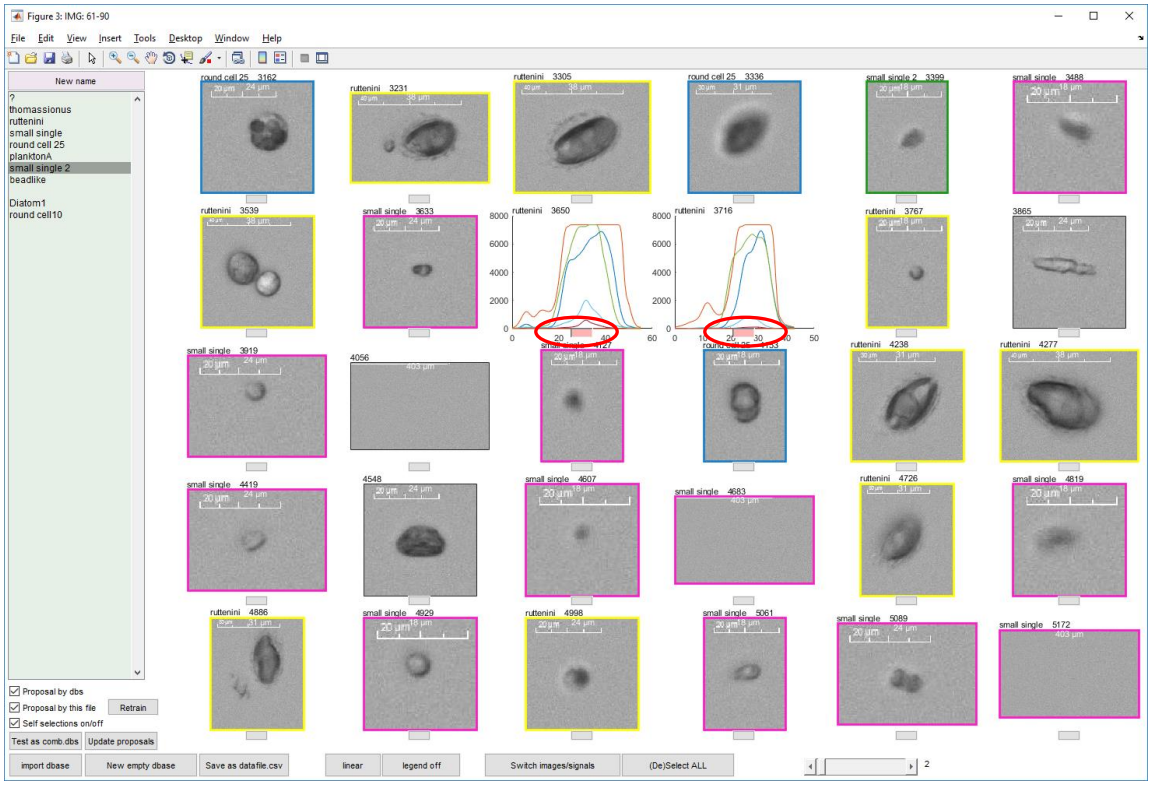
Rename wrongly assigned images by clicking on the image until there is the red box, and then confirm the new name by clicking on the right database name in the list.



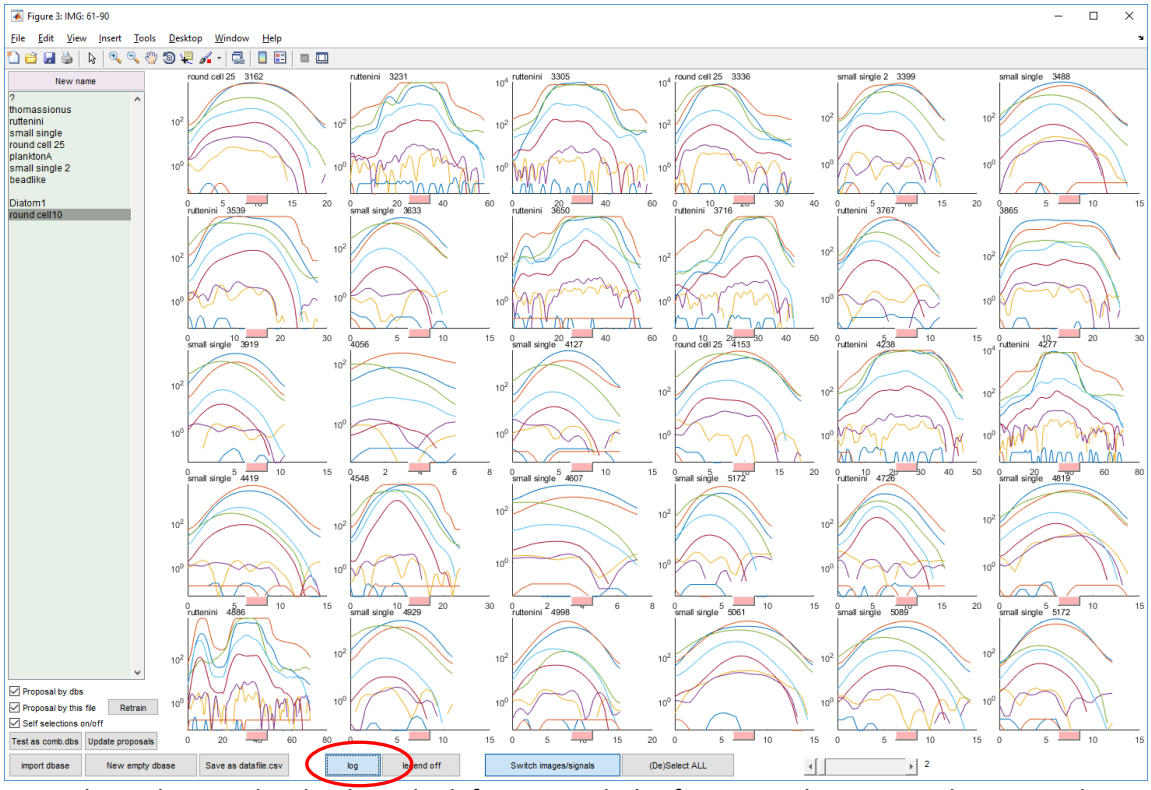
Flow cytometric signals or profiles of the images can be visualized all at once by the button 'Switch images/profiles' in the middle below ...



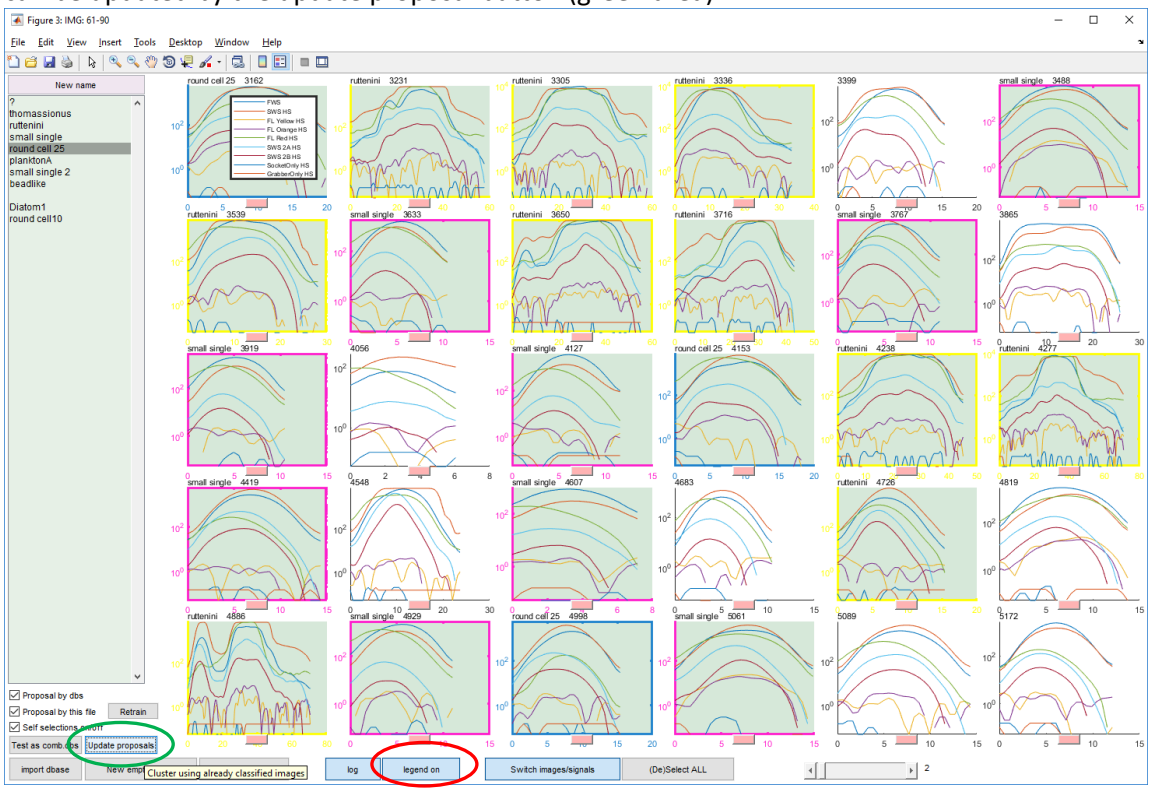
or by indivudual switching with the button just in the middle below the image.



Switch from a linear to logaritmic axis in the signal height by the middle below 'Linear' or 'Log' button.

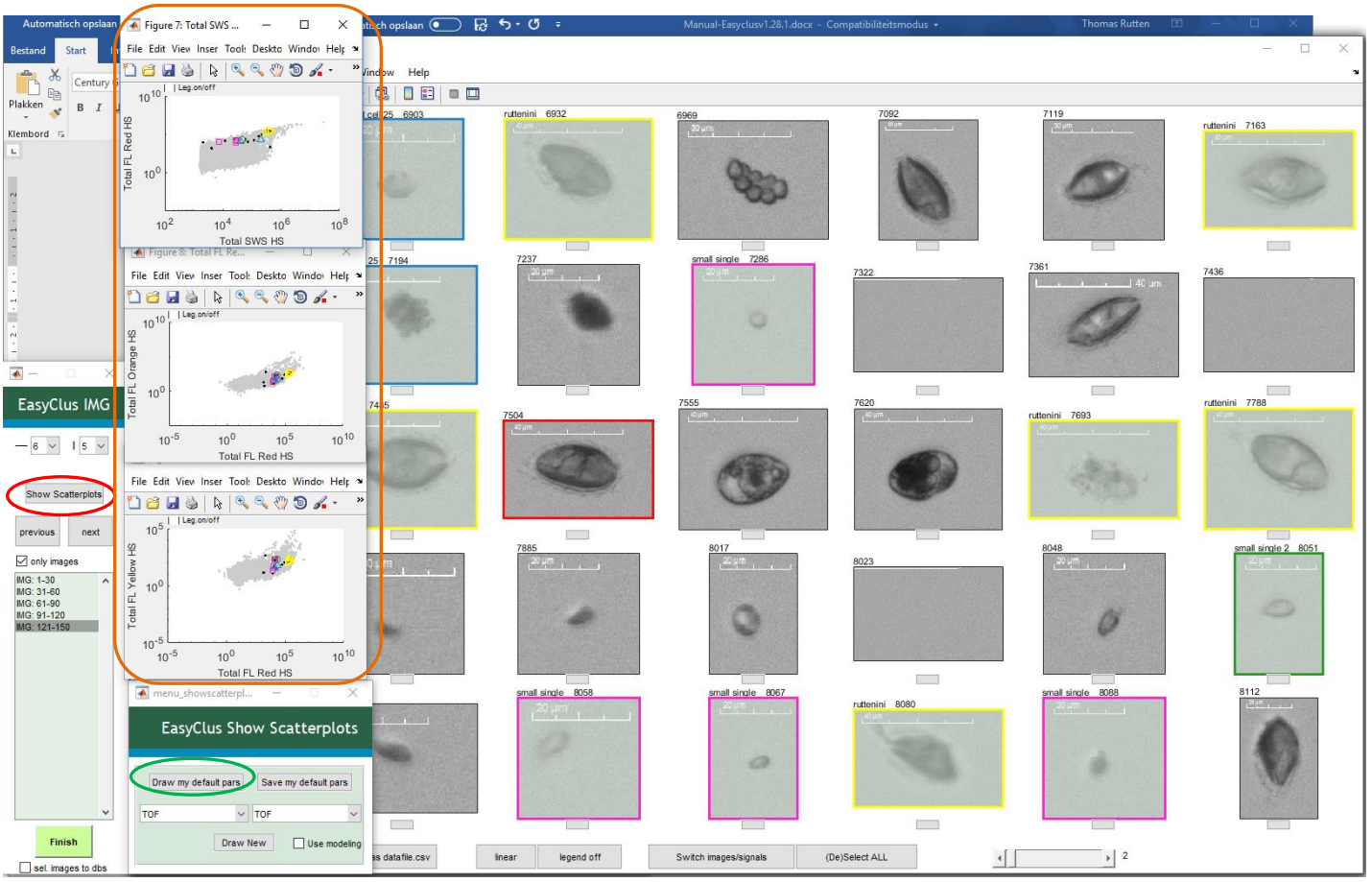


Legends can be visualized only in the left-upper subplot figure. It takes too much time to plot them for each subplot figure and it is similar for each figure. The assignment to species names can be updated by the update proposal button (green area)

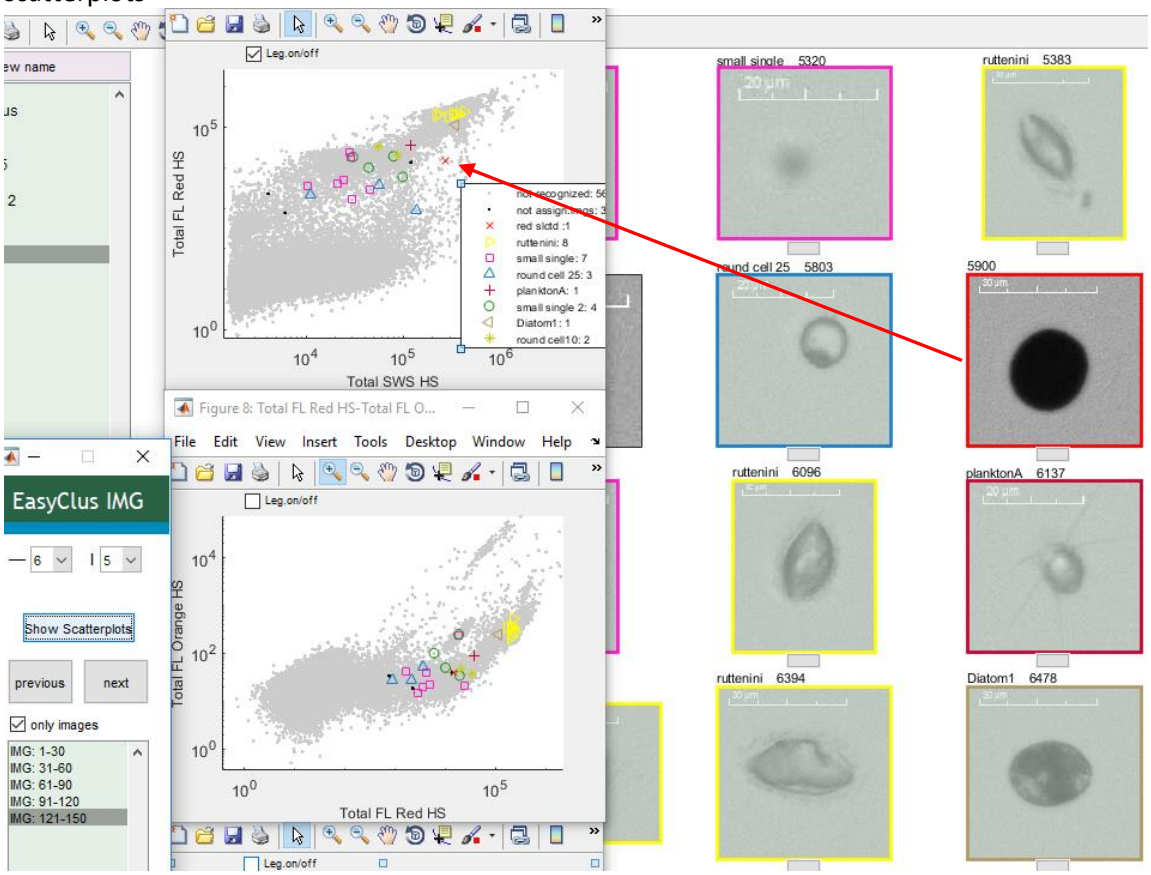


Show scatterplot(s) of images and assigned species. It can be helpful to know where the FCM variables of the images are in the scatterplots. The 'Scatterplot menu' appears after clicking the 'Show Scatterplots' button (red area).

In the menu, just press 'Draw my default pars' (green area) and some scatterplot combinations are drawn (orange area). Press 'legend on' in the scatterplots to show the legends. The fifth series (images 121-150) is represented in the scatterplots



Second series (images 21-39): The assigned species and not assigned events are shown in the scatterplots



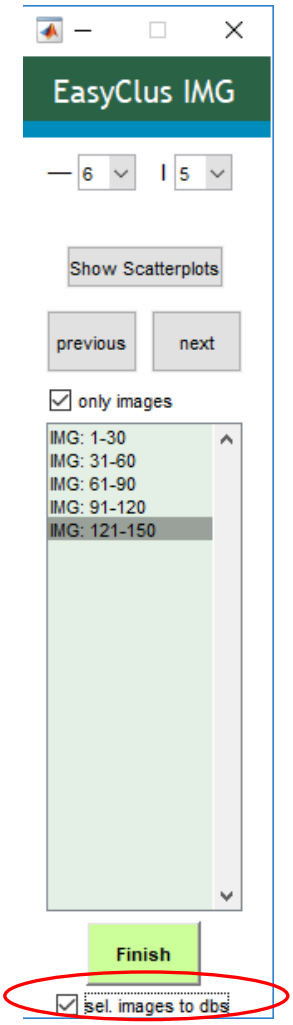
How to go on?

1. We open new images windows
2. Check what proposal is done on basis of earlier assignments
3. and manually change them if necessary

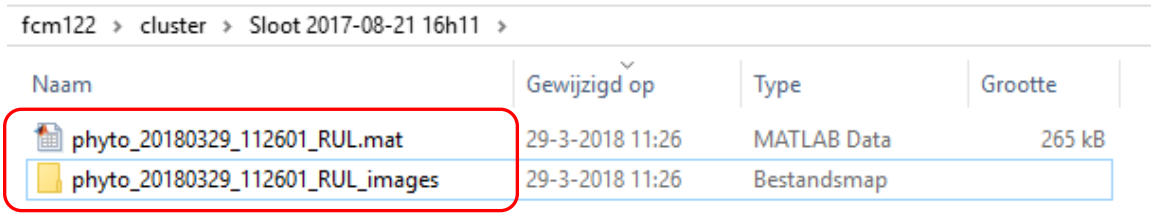
If we are happy with the propsal results we can store the results into a new phytodatabaseXX.mat database.

In fact nothing is stored in the database yet. This is only the pre-phase, the selection and assign process. But for now, you are finished and would like to store the results in the database.

You are finished and want to save the results to the database. Just put the checkbox ‘save to database’ to ‘on’ and press the ‘Finish’ button. All results, selections of images with a specified species name, are stored in the database. Images are stored as a separate jpg.file in the databasename folder.

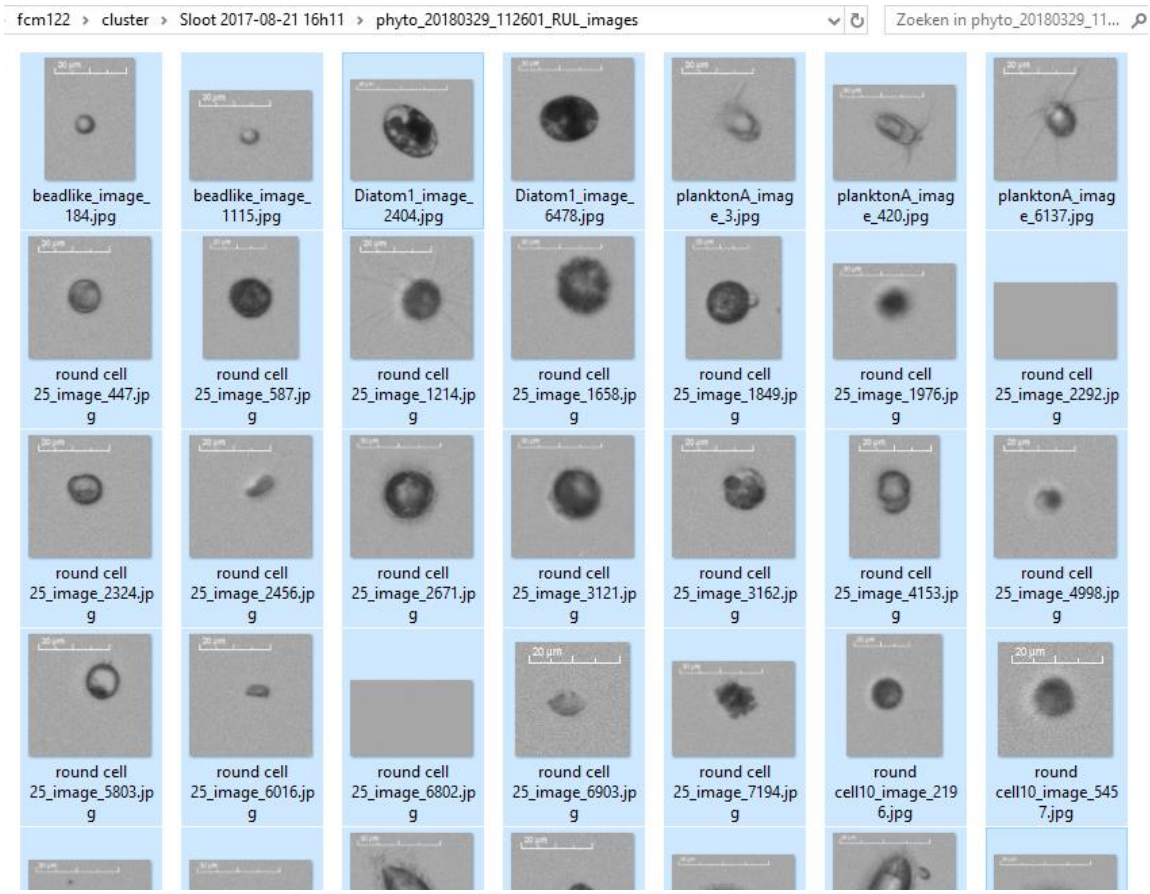


The database storage option is a little bit confusing here (will be solved later), because all the assigned images are stored in fcmXX\cluster\FCMfilename\phyto…RUL.mat and images in fcmXX\cluster\FCMfilename\ phyto…RUL _images\

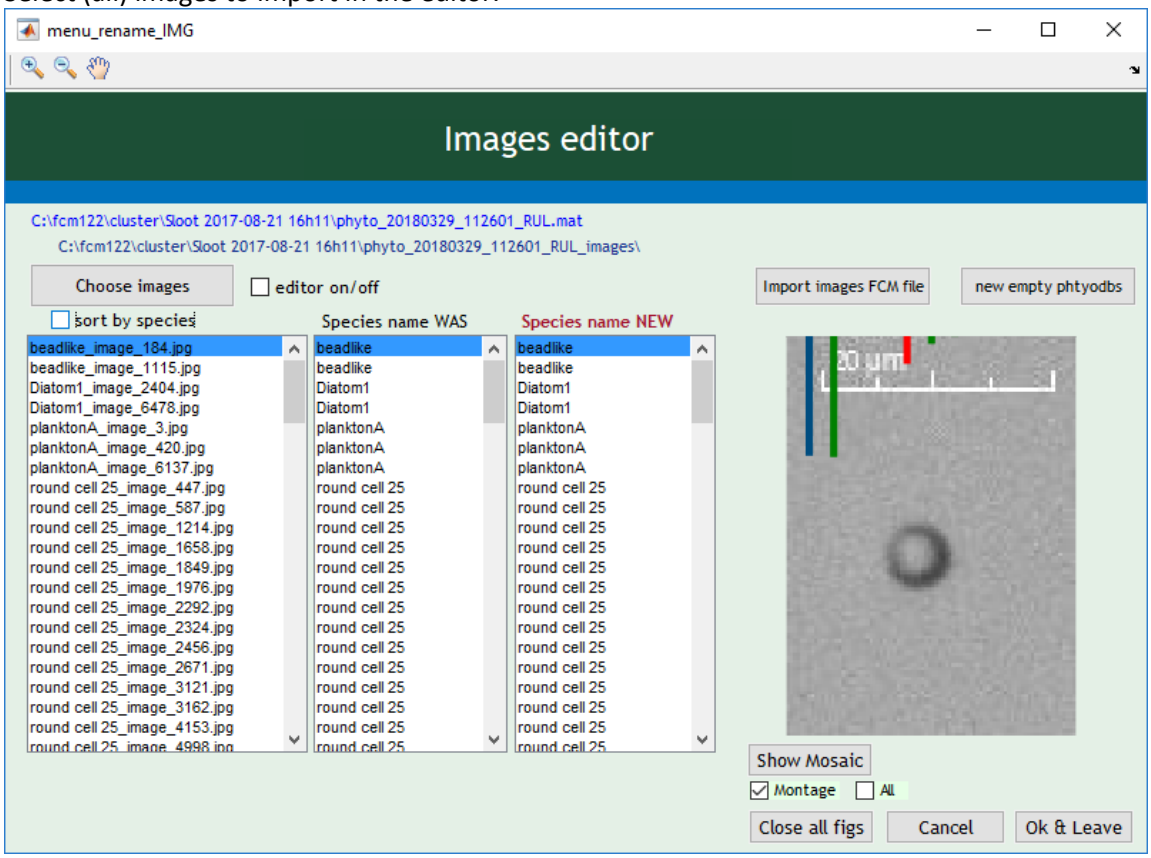


Database-images are saved as jpg-file in this folder using the species name and id in the database.

If wanted, an image-editor can be opened immediately by selecting the assigned images from the phyto…RUL_images\ folder. The images are loaded in the image editor for renaming or other editing procedures.

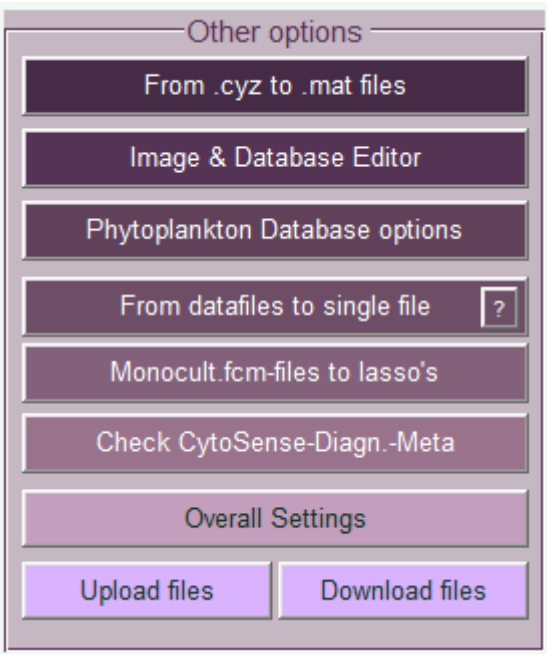


Select (all) images to import in the editor.



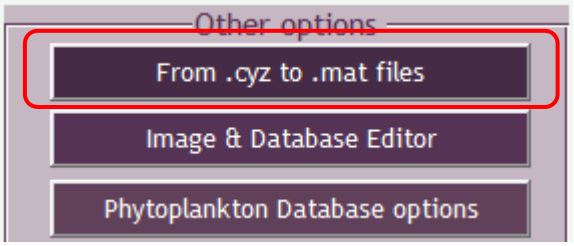
This last option gives you the opportunity to check the results and store them as a definite validated database.

4. Purple window – Other Options



4.1 From .cyz to .mat files

Option to convert .cyz files to (EasyClus) .mat files

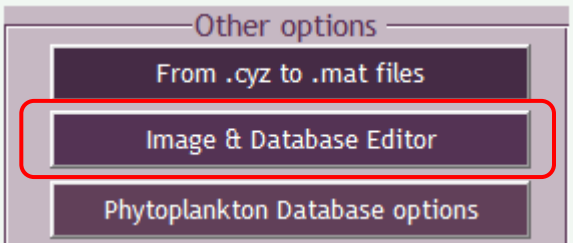


The .cyz files can be slow when importing in EasyClus especially when they are big. The import of attribute data is usually fast enough, but the import of images and signal profiles is very slow. Also the size of the .cyz file especially if a lot of images are included, can take a lot of computer memory. The profiles signals are used for modelling of saturated peaks and (later) new tools will be developed working with profile signals (e.g. new attributes data). The reload time of .mat files in EasyClus is much faster than .cyz files. It is recommended to convert .cyz files to .mat files for big files.

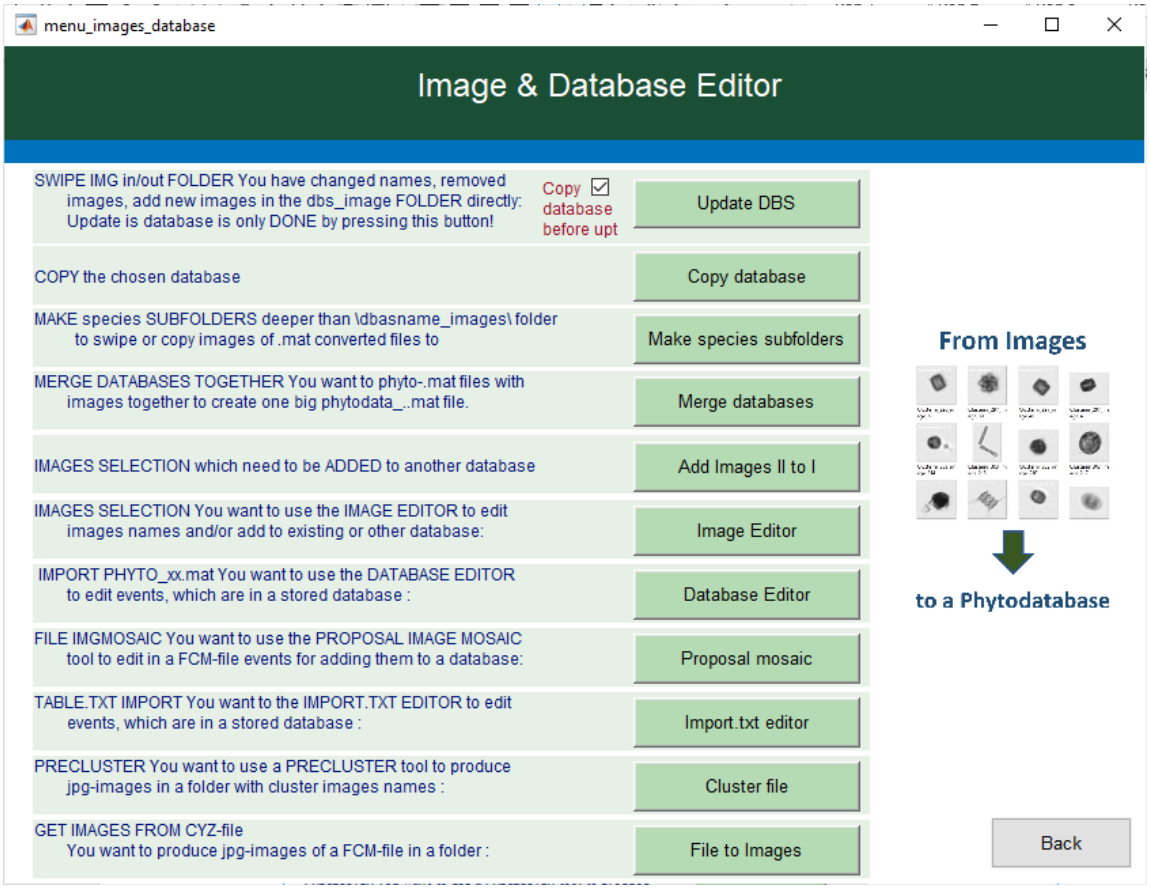
Procedure is simple. Select the .cyz files and press ok.

4.2 Image & Database Editor

Multipurpose Editor for (new) phytodatabaseXX.mat format (attributes, profiles, images)



The menu below appears when clicking the Image & Database editor, but it is more than only an editor.



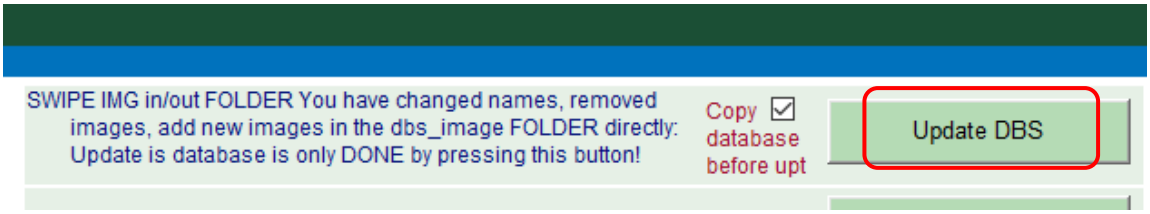
I have chosen to create or make a database by several different methods. Which method is used depends on what you the user fits the best:

- by clustering in EasyClus producing a temporary database
- by swiping species images to species folders in the database images folder
- by selecting and changing the image names (in the image editor) of the .jpg image files produced by EC
- by editing in an existing database in the database editor produced by EC
- by importing a text table with filename – speciesname - indexnrs in FCM file.
- By merging databases with each other

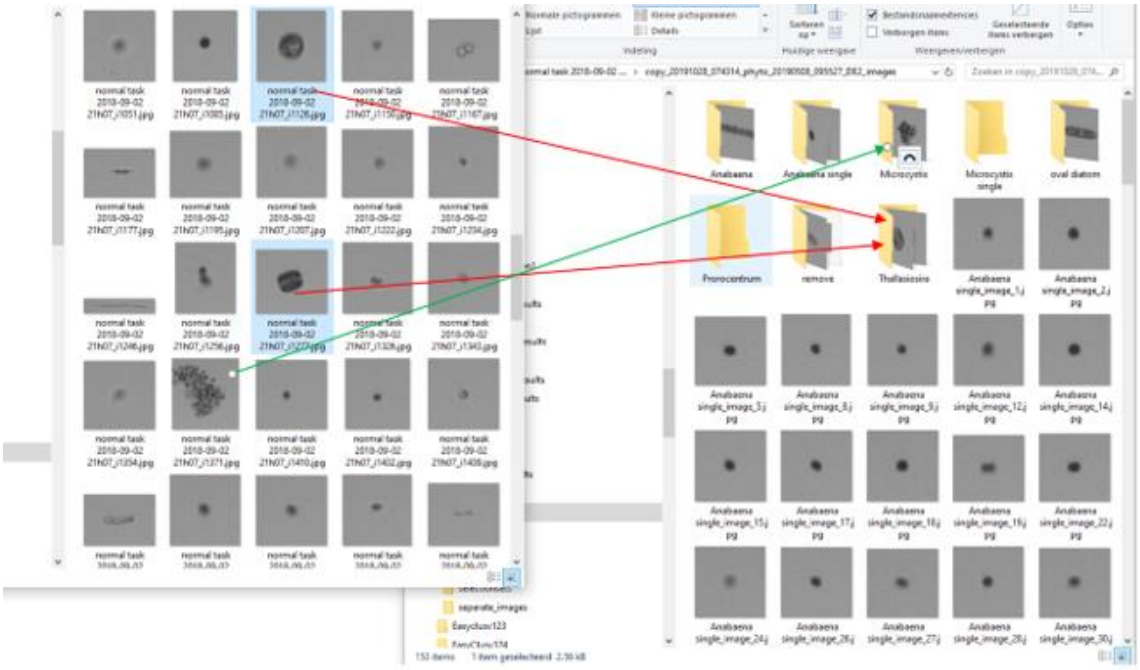
All options are possible in this menu and it works as plug & play.

4.2.1. Update DBS

Check if all information is in the phytodatabaseXX.mat file

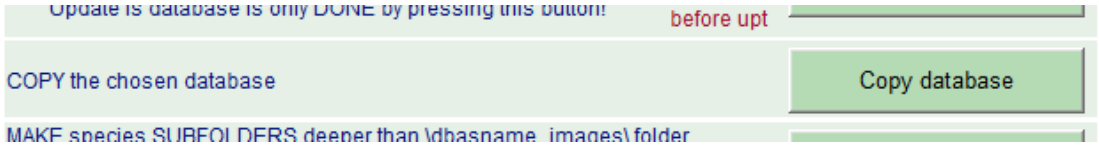


The update option checks if all information of all events in the database are complete and available. That is the combination of attributes data as well as signal profile as well as image for each event. If not, the Update DBS operation tries to find the original file to complete the information. This .cyz or .mat file is supposed to be in the \Easyclusvxxresults\datafiles\ folder or a folder lower than the database. Copies of species are deleted, removed marked species are deleted too, images from the database image folder that have been deleted will be deleted too.



4.2.2 Copy database

existing phytodatabase.mat as a backup Easily copy



4.2.3 Make subfolder of unique species in the database

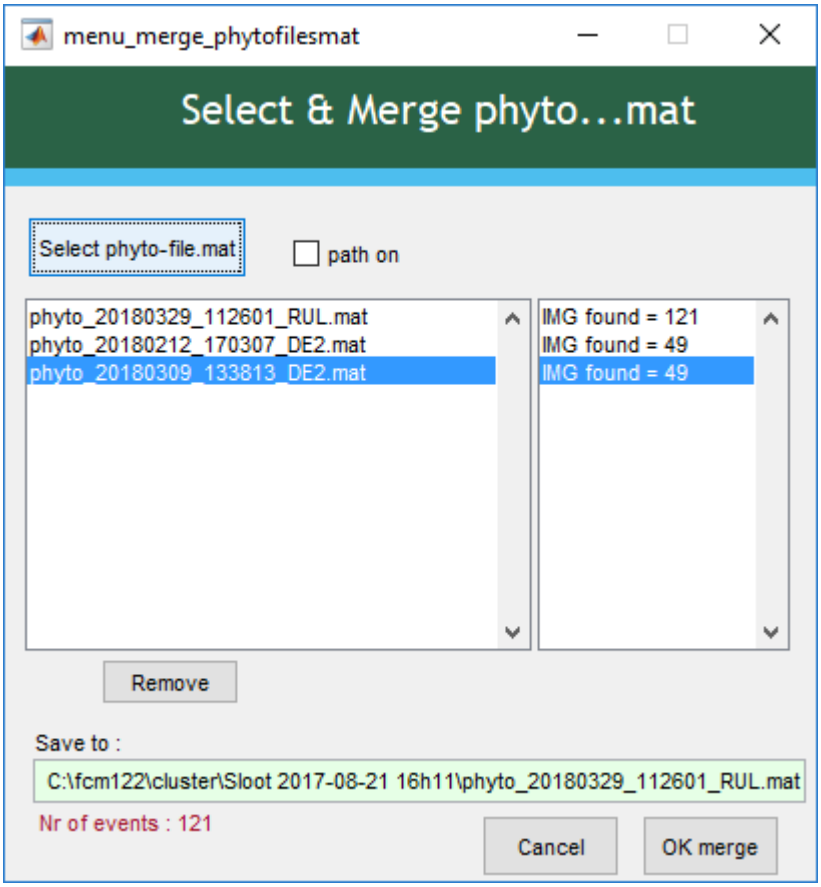
Make automatically empty subfolders of unique species in the phytodatabase.mat main _images folder

Images of species are NOT stored in subfolders, but in the phytodatabase.mat main _images folder.
Swiping (or copy-paste) of new images of (new) FCM files to the subfolders in the database_images folder with a specific unique species name, can be used to easily add new images (and data) to this database. Also new species types can be added this way by adding a new species type folder name filled with images of this new species type.
Images can also be removed from the database by swiping them to the a subfolder with the name 'remove'.
To activate the images from subfolders to database process, use the 'Update DBS' button.
IMPORTANT : New images which are not yet in the database, and which are put in a subfolder to be added to the database need the original FCM file (.cyz or .mat) for attributes data and profiles information. These FCM files are supposed to be available in the 'Easyclusvxxxresults\datafiles\' folder.

4.2.4 Merge databases

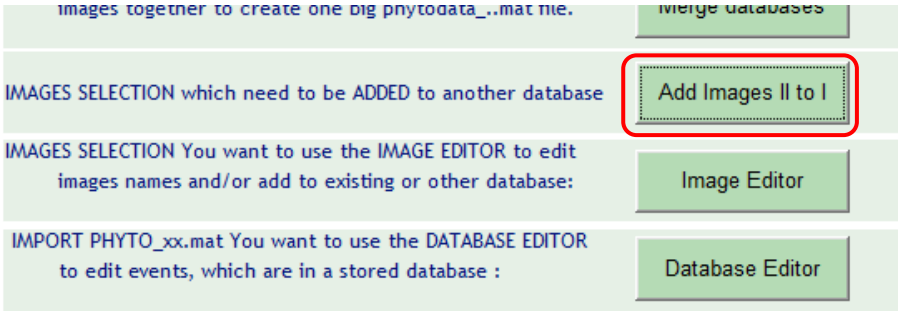
Fast tool to easily merge two or more phytodatabaseXX.mat file with each other

Simply select phytodatabaseXX.mat files and press OK merge. In principal the first database name chosen, will be the main database and all other databases will be added to this one.

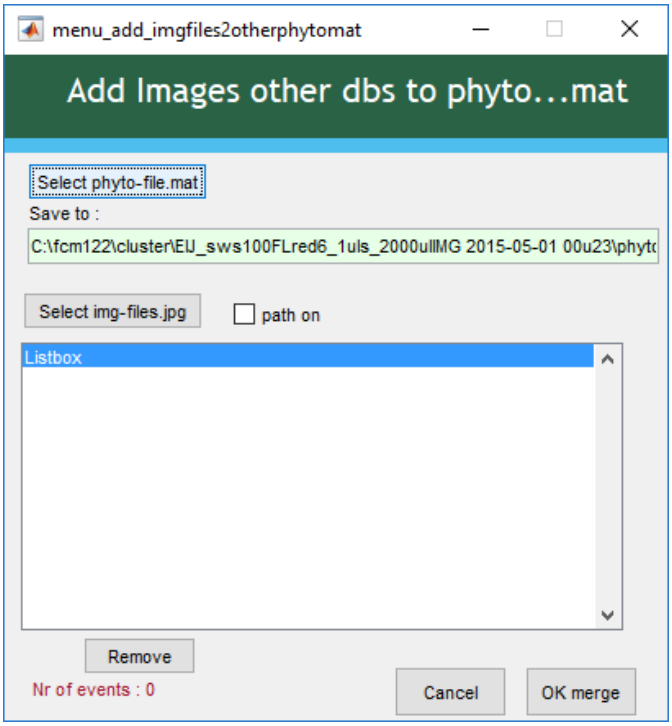


4.2.5. Add Images II to I

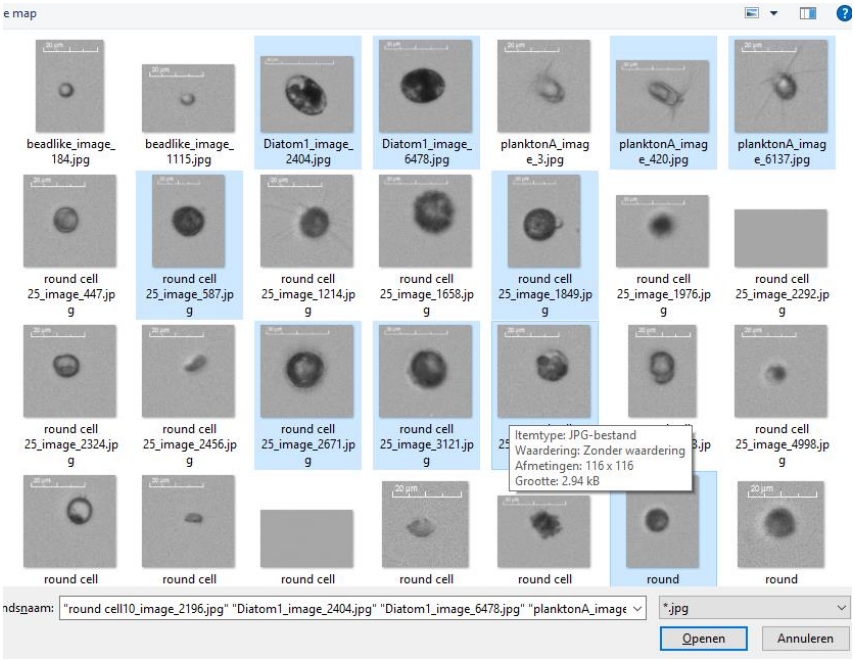
Tool to add images directly to a database



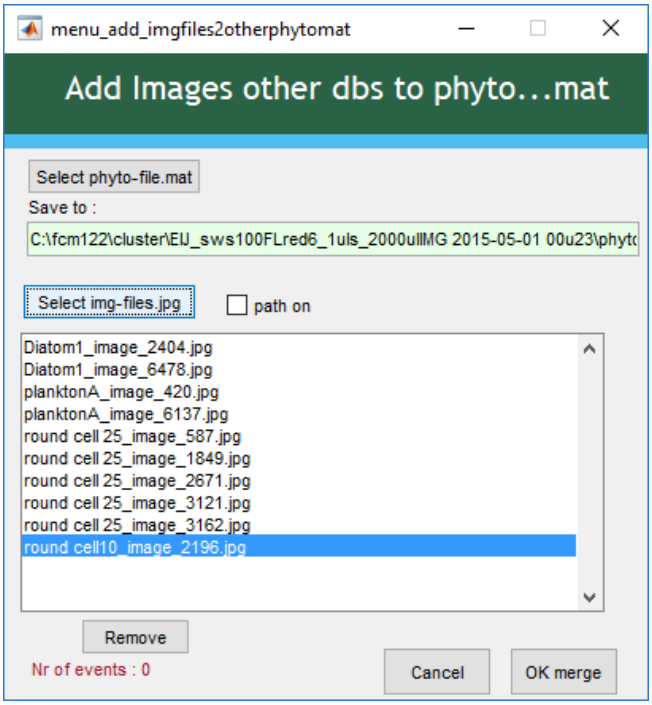
Select phytodatabaseXX.mat and select images, for example from another database, that should be added to this database.



Selection of images (only blue ones) ….

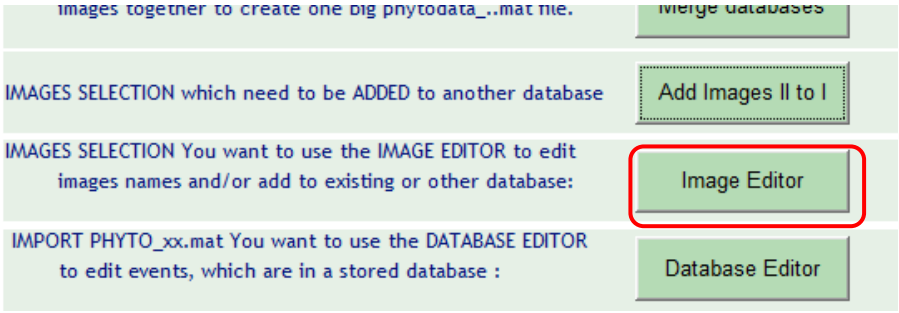


And press OK merge to add and save the images (and in the background the attributes and profiles) into the chosen database.

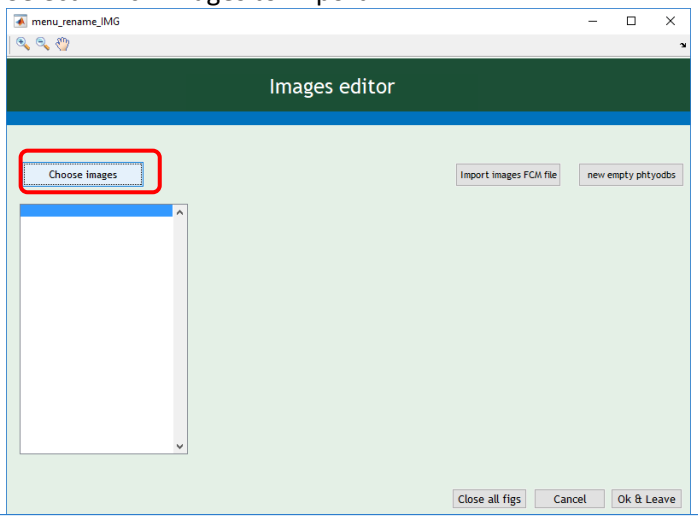


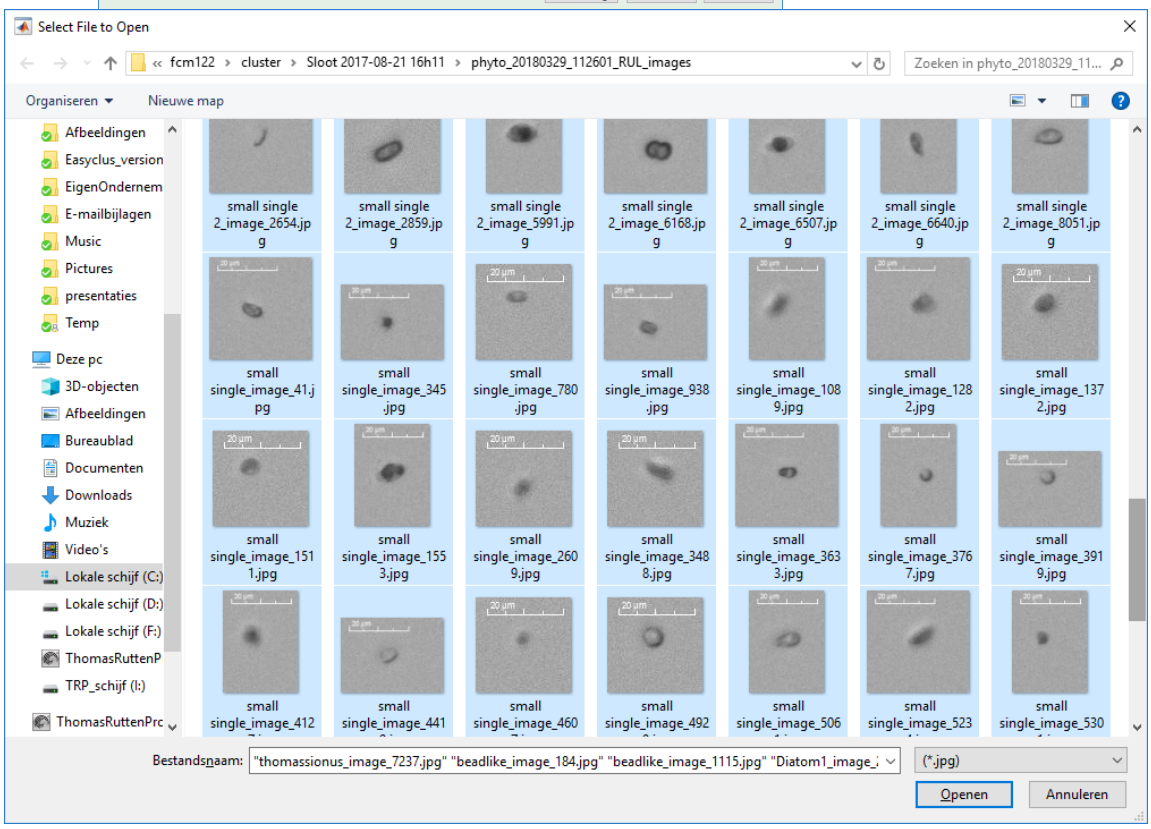
4.2.6. Image Editor

Tool to choose, rename, add images directly to a (new) database

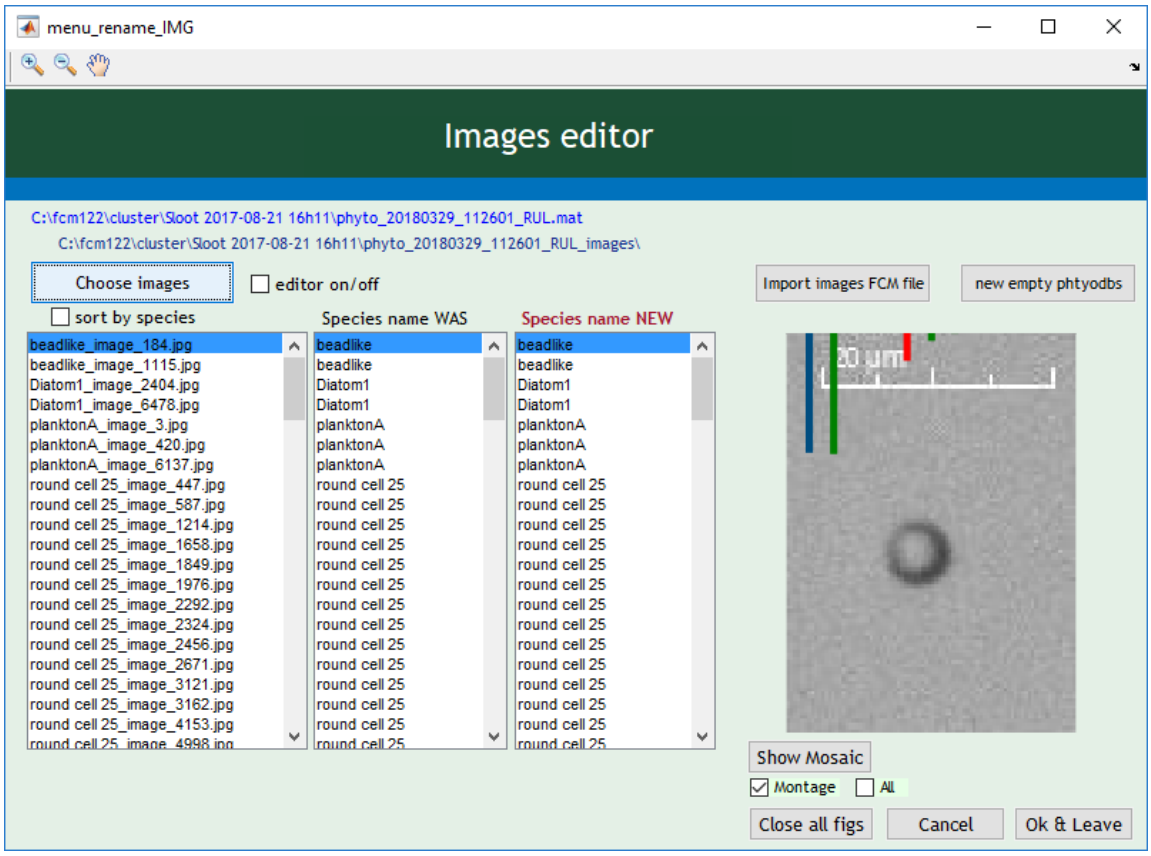


Select which images to import:

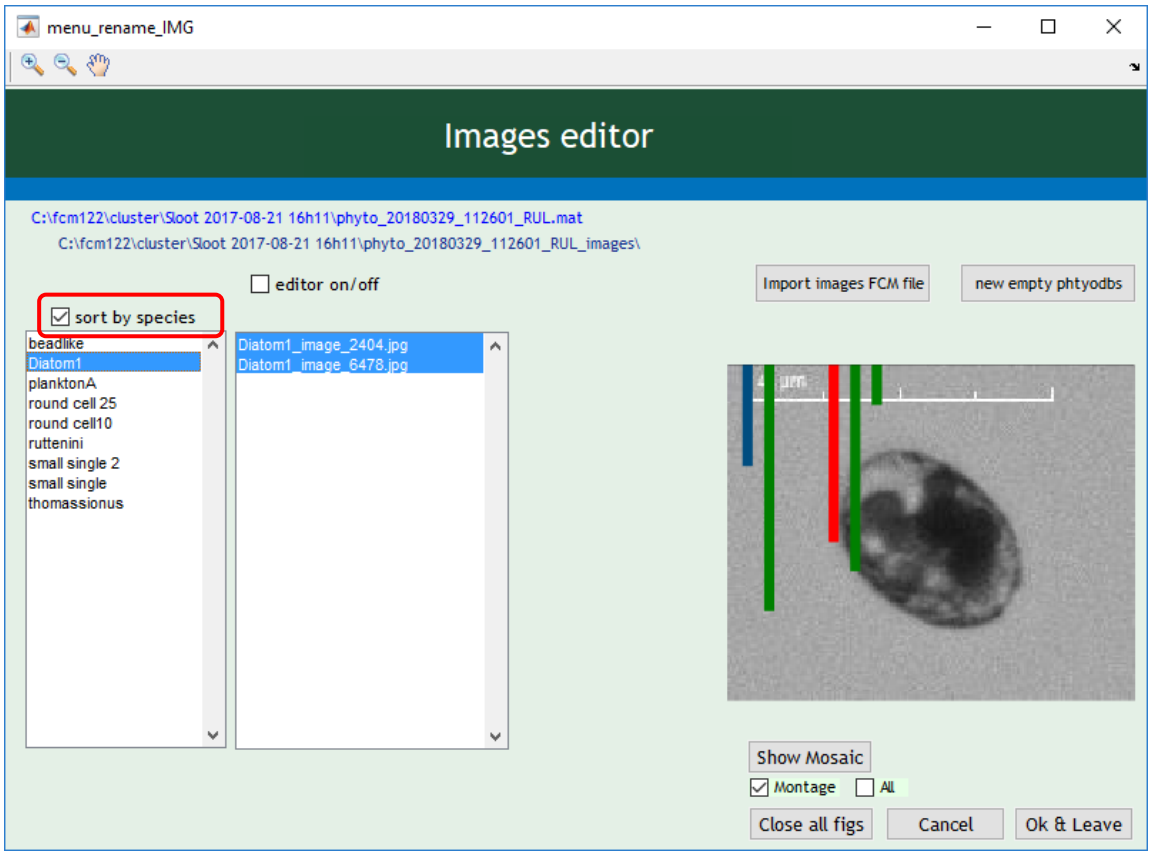




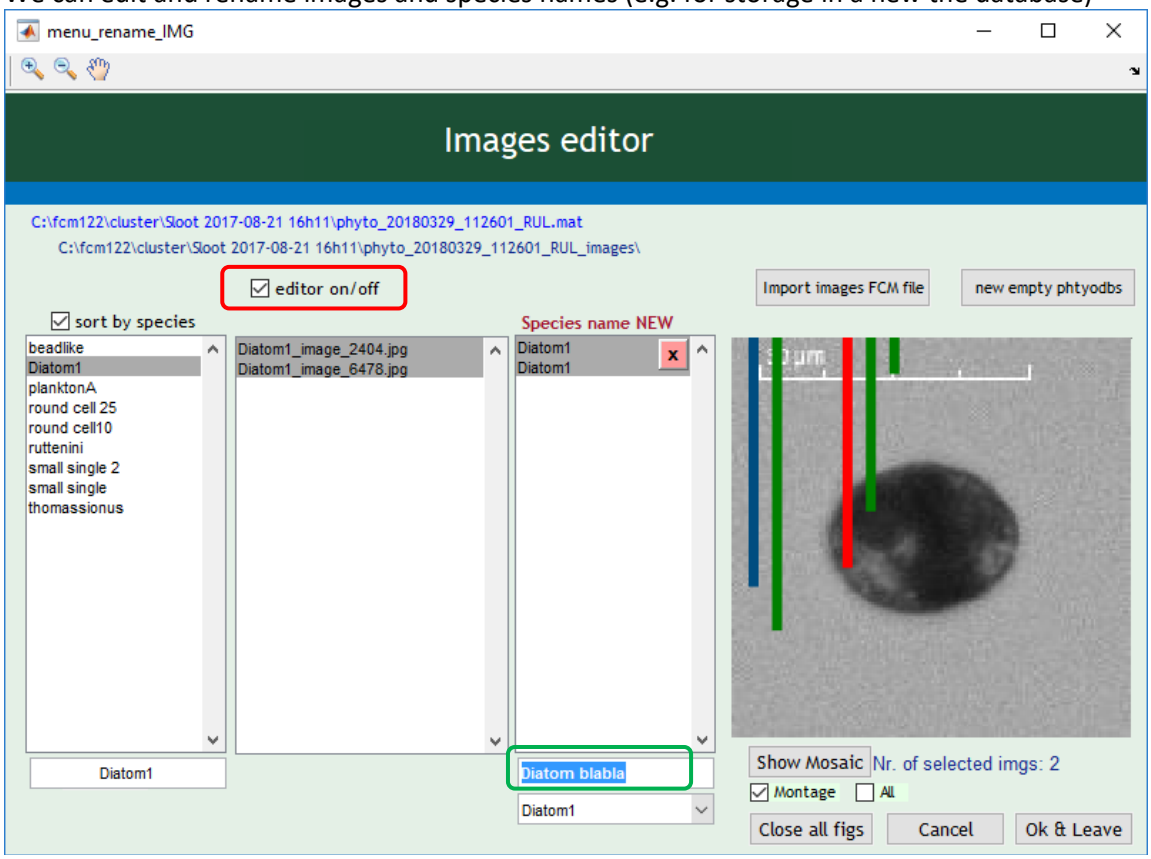
After importing they appear in the Images editor menu

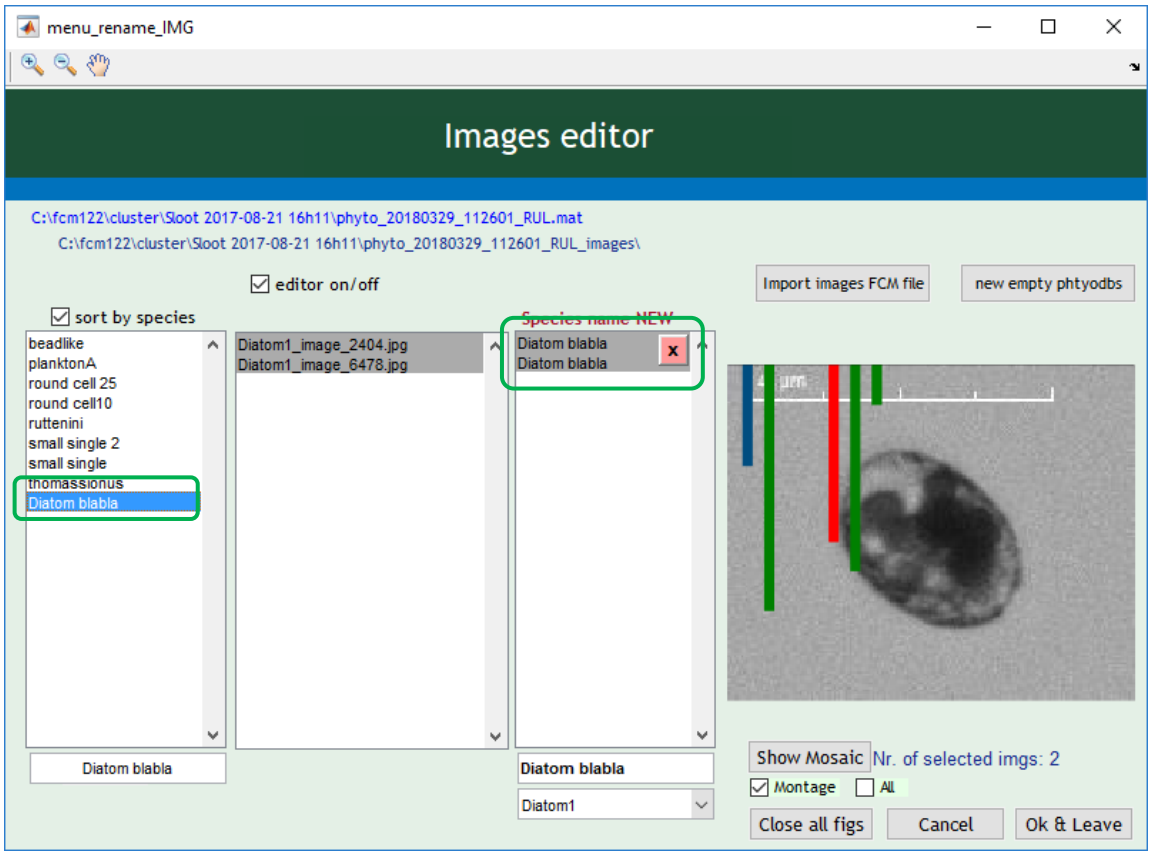


We can easily sort them by name

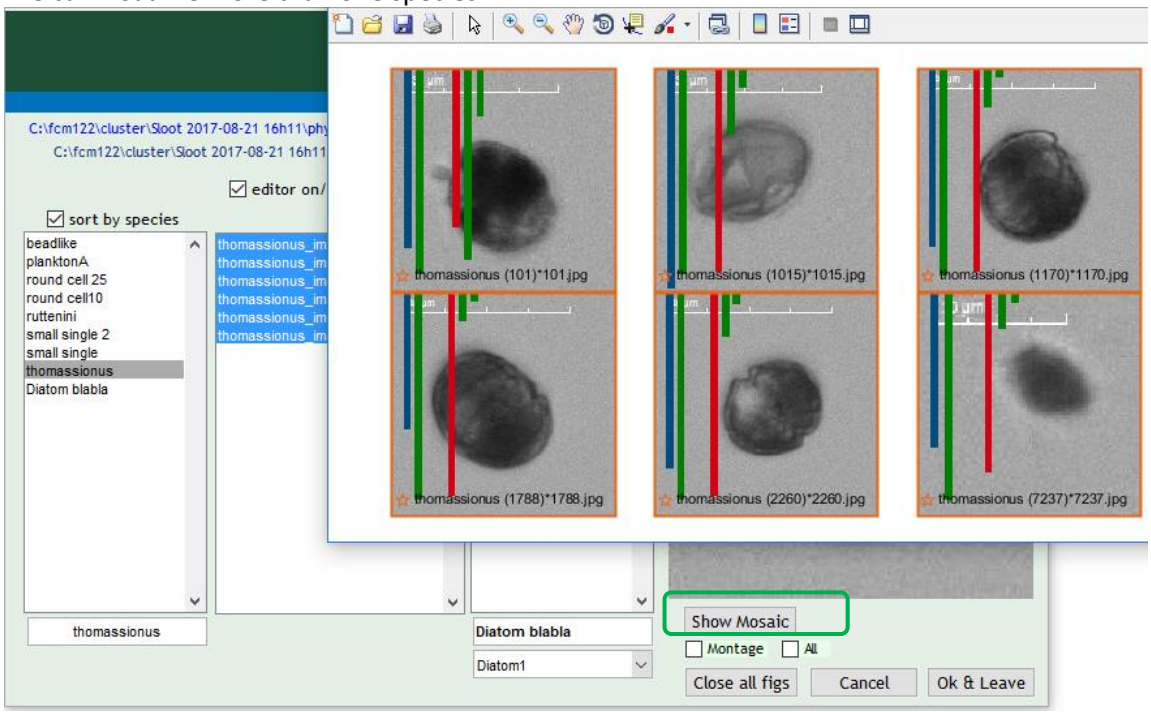


We can edit and rename images and species names (e.g. for storage in a new the database)

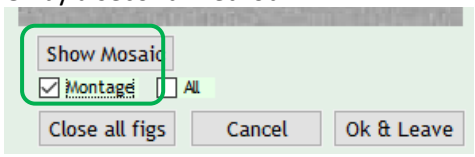


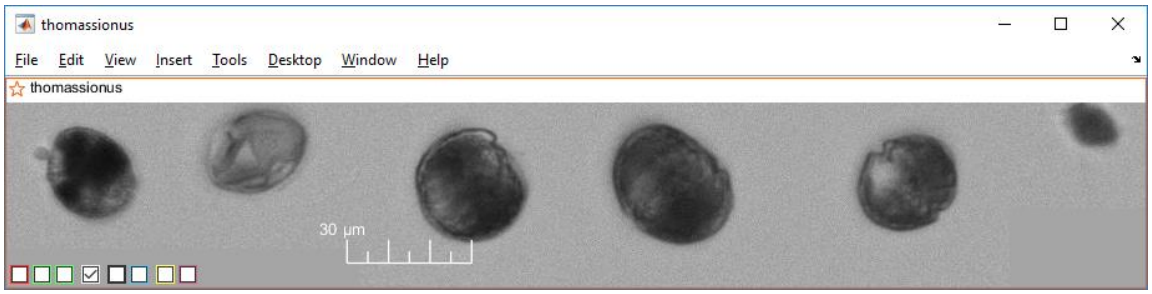


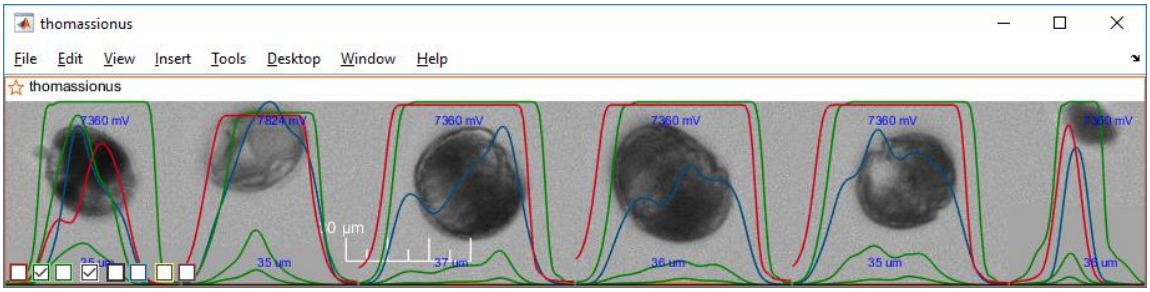
We can visualize more than one species:



Or by a second method



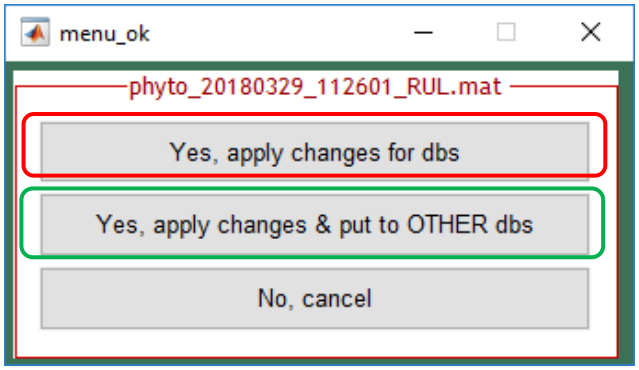




Results are saved to a new or existing database after changes has been applied and images editing has been finished.

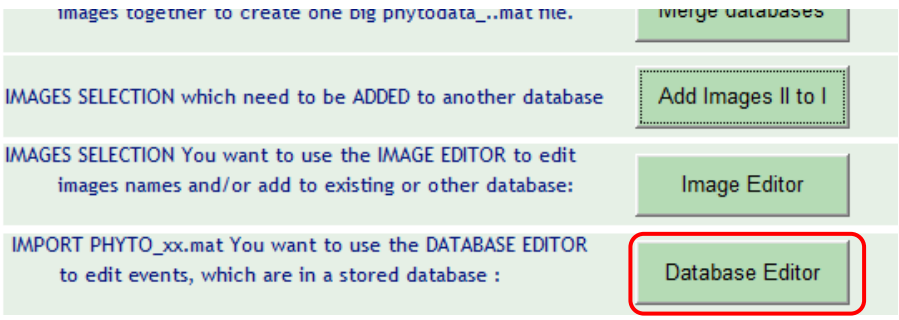
= save to existing database, where selected images belong to.

= save to other database, where selected should be copied to.

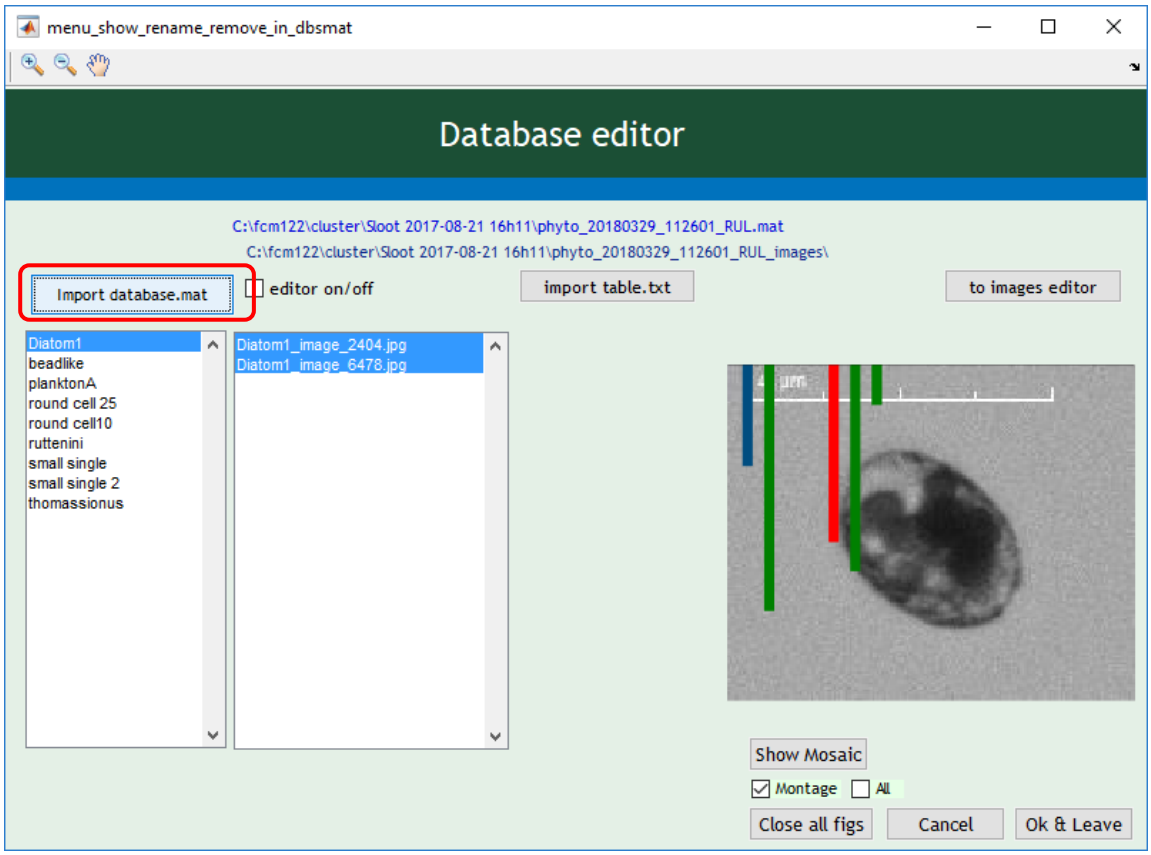


4.2.7. Database Editor

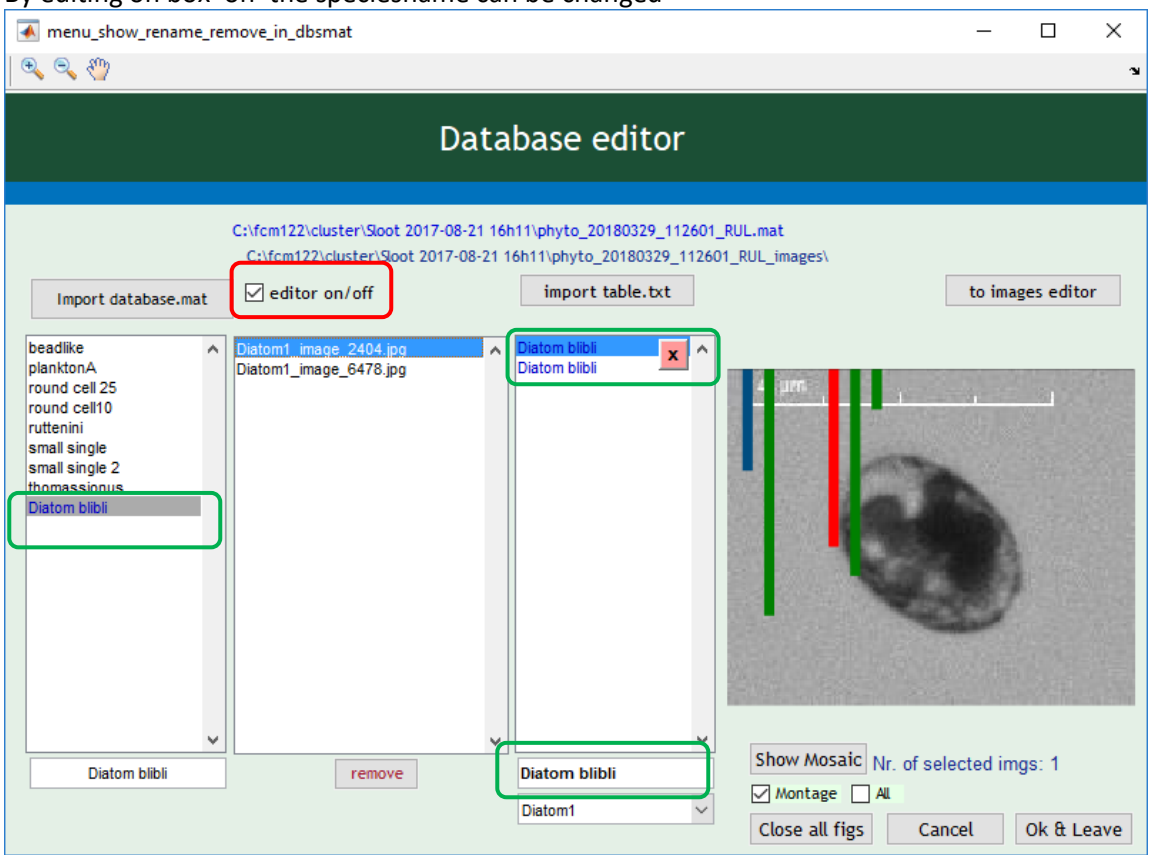
Tool to edit, choose, rename, add species and events directly in a database



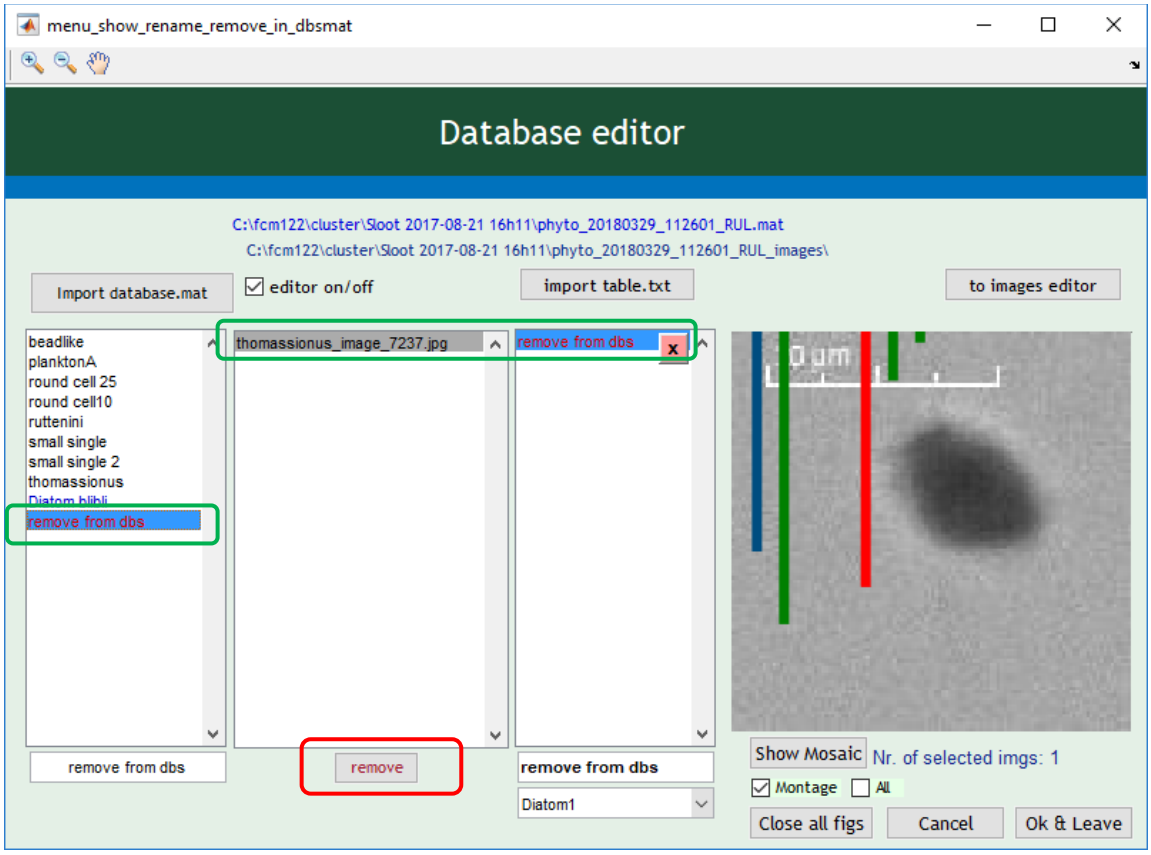
Select which database to import:



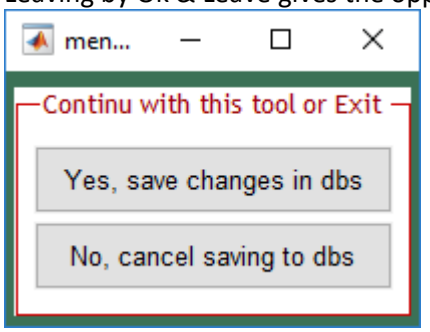
By editing on box ‘on’ the speciesname can be changed



Or events can be removed

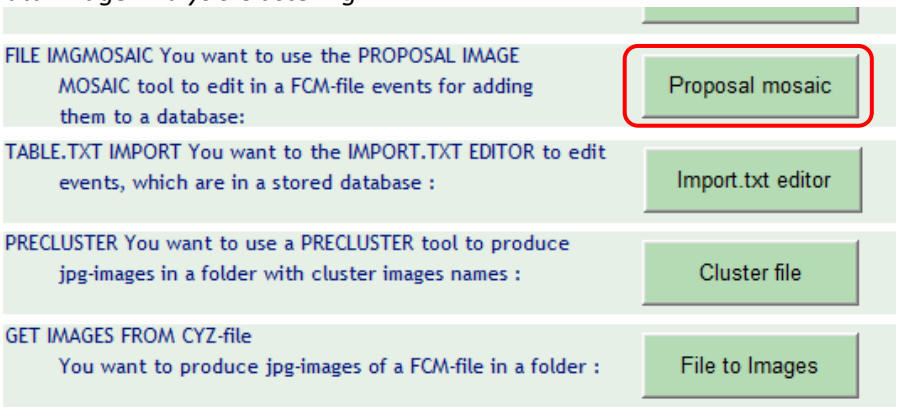


Leaving by Ok & Leave gives the opportunity to save the changes to existing database or not.



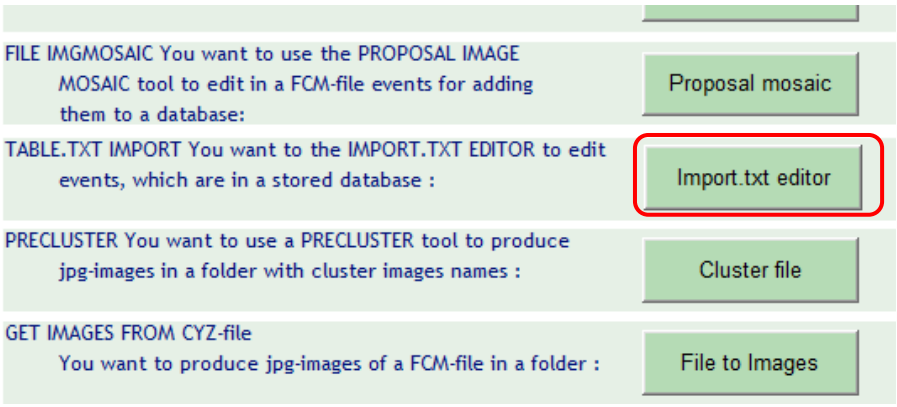
4.2.8 Proposal mosaic

Tool to make a database 'on the fly' = the method described is already described in 3.4.9 Data Image Analysis Clustering



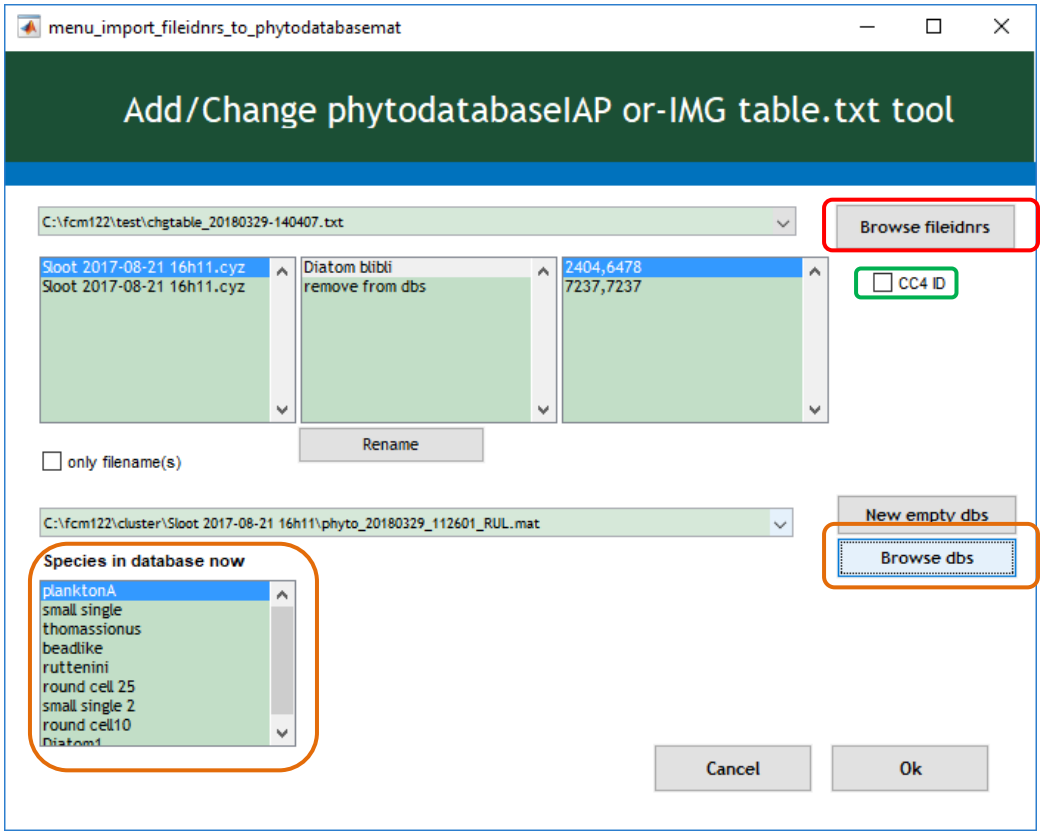
4.2.9 Import txt editor

Tool to make a database by a stored .txt file in combination with the used FCM-files described in the table



An user is allowed to produce a table with FCM-filename (preferably with pathname) and species name that belongs to a specific IDnr (see table below). In CytoClus images can also be easily viewed as well as the scatterplots and pulse profile. This table should be save as .txt table and can be imported here. More than one filename is allowed.

	A	B	C	D	E	F
1	LL_Zirfaea_sws15flr 2uls_600 sec 2017-04-12 07h31.cyz	Asterionella	34	53	79	110
2	LL_Zirfaea_sws15flr 2uls_600 sec 2017-04-12 07h31.cyz	Prorocentrum	124	488		
3	LL_Zirfaea_sws15flr 2uls_600 sec 2017-04-12 07h31.cyz	remove from dbs	440			
4						



= Import .txt file with table containing filename – speciesname – ID nrs

= CytoClus ID start with 0,1,2.. EasyClus ID start with 1,2,3 … If ID from CC4 put checkbox on

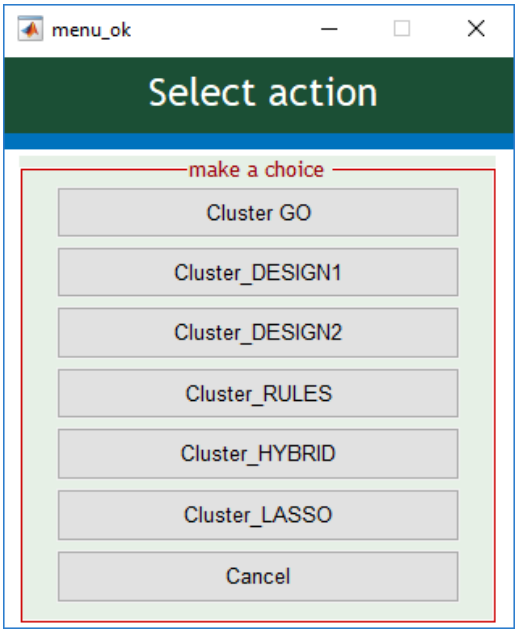
= Import phytodatabaseXX.mat database, names within database are shown

Press OK to import events in table to the database. FCM-files are supposed to be in path-filename (if path is in table), otherwise the filename is expected to be in fcmxx\datafiles\

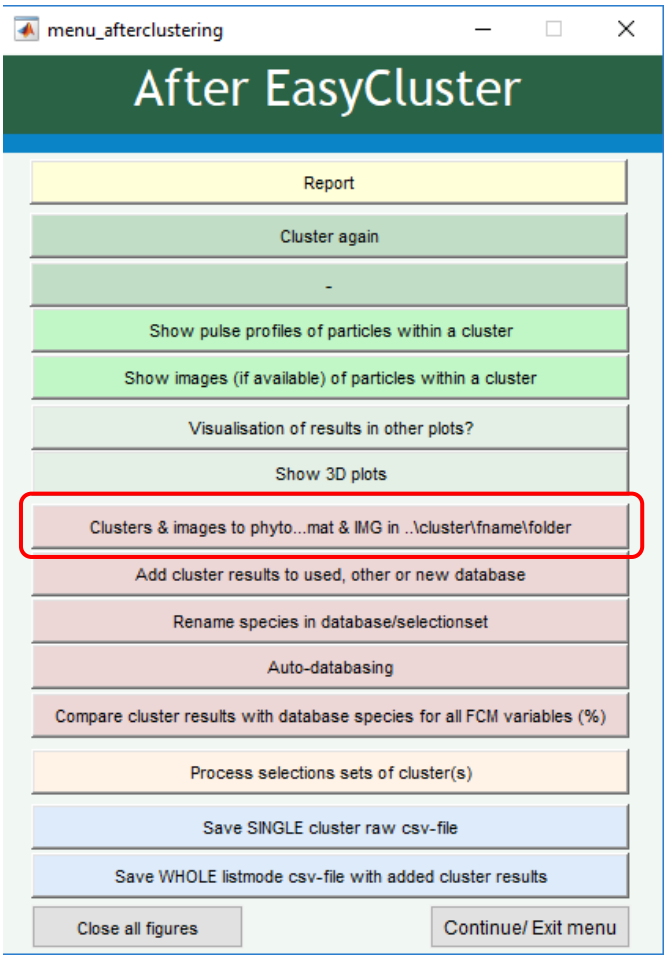
4.2.10 Cluster file

Tool to cluster FCM-files and to produce a temporary phytodatabaseXX.mat database

Preclustering on basis of FCM data can help the user a lot by classifying the data on basis of images. Therefore this option is put here to precluster on basis of the already known EasyClus clustering methods.



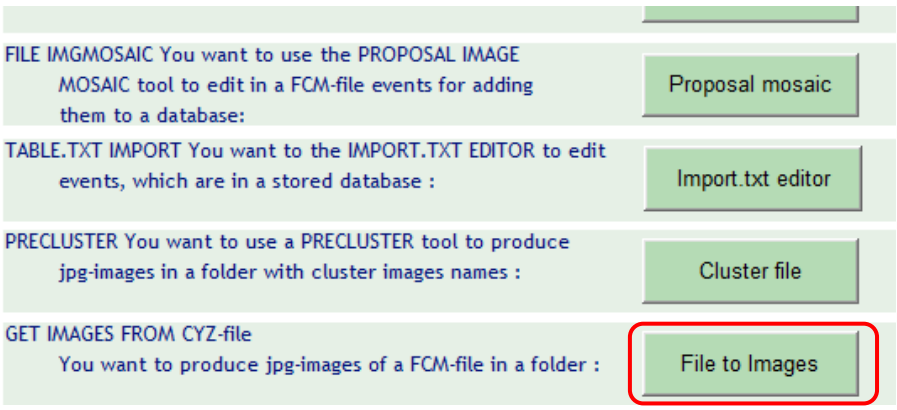
After clustering with one of these cluster options, the temporary database is made in fcmxx\cluster\filename by clicking the button 'Cluster & images to phyto…mat & IMG. Be aware, ONLY particles WITH an image are stored here!
Recommended procedure is to use and store only data with an image, because this can be validated while assigning names to it by the image recognition.



In \fcmxx\cluster\filename\phyto …XX.mat and in \fcmxx\cluster\filename\phyto…XX_images\

4.2.11 File to Images

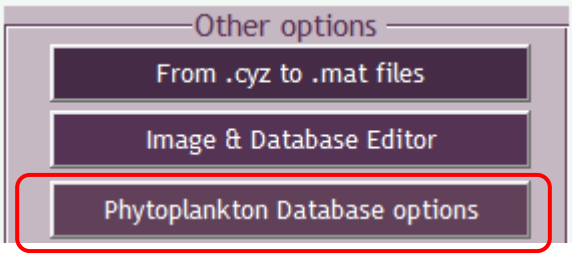
Tool to extract jpg image files from .cyz files



Select a .cyz file and OK

4.3 Database related options

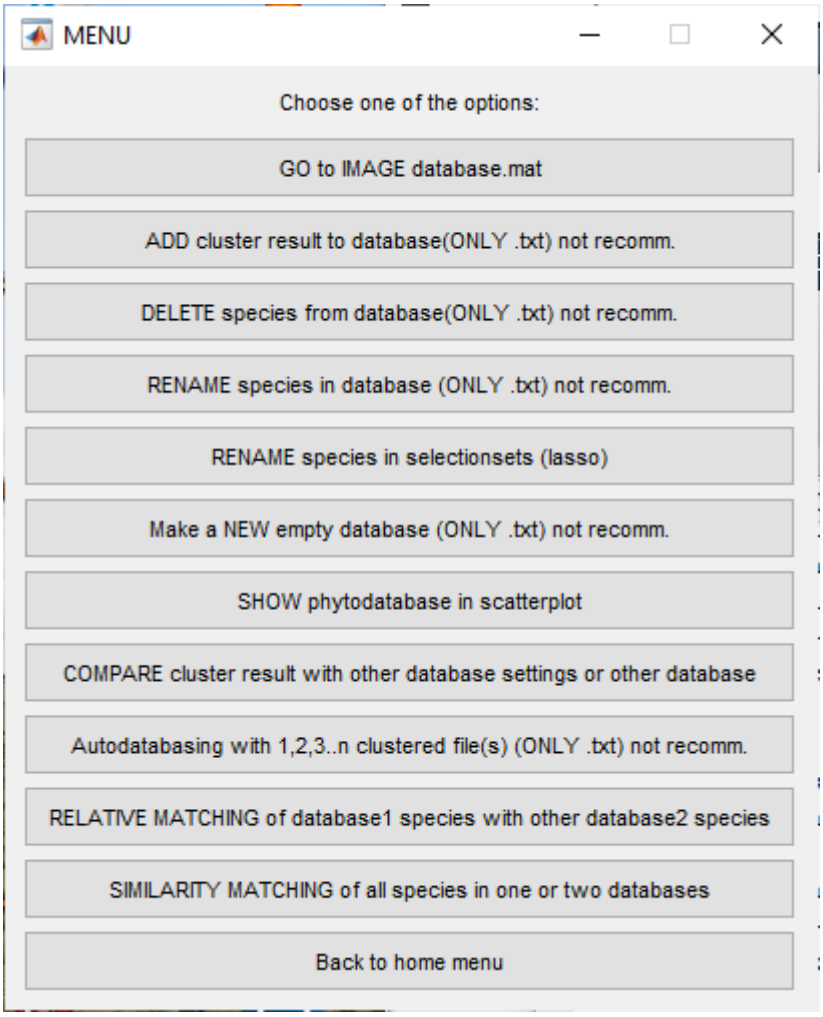
About handling (ADD species,DELETE species, NEW, SHOW, COMPARE cluster results with other settings)



ADD cluster result to database, DELETE species from database/
make a NEW database/ SHOW database in scatterplot(s)/ COMPARE cluster results with other settings

Database handling menu: to exchange information with one or more stored databases

After choosing this option, following menu appears:



4.3.1 GO to IMAGE database.mat (you will be forwarded to the database editor menu)

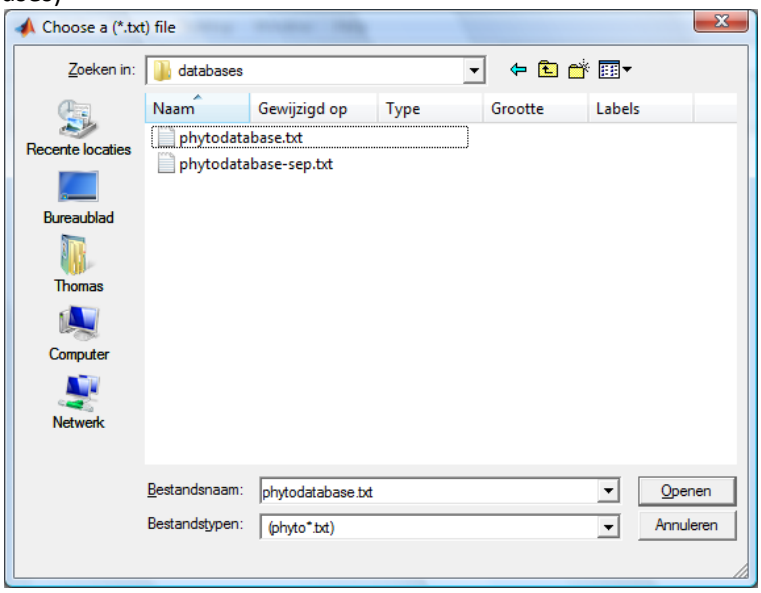
For those who have an image in flow module and EasyClus LIVE

Option ‘GO ..’ starts with choosing the image database (name) where species cluster profiles, cluster images and attributes data (fingerprints) are saved. Several option appears in the next menu to handle your image-profiles database.

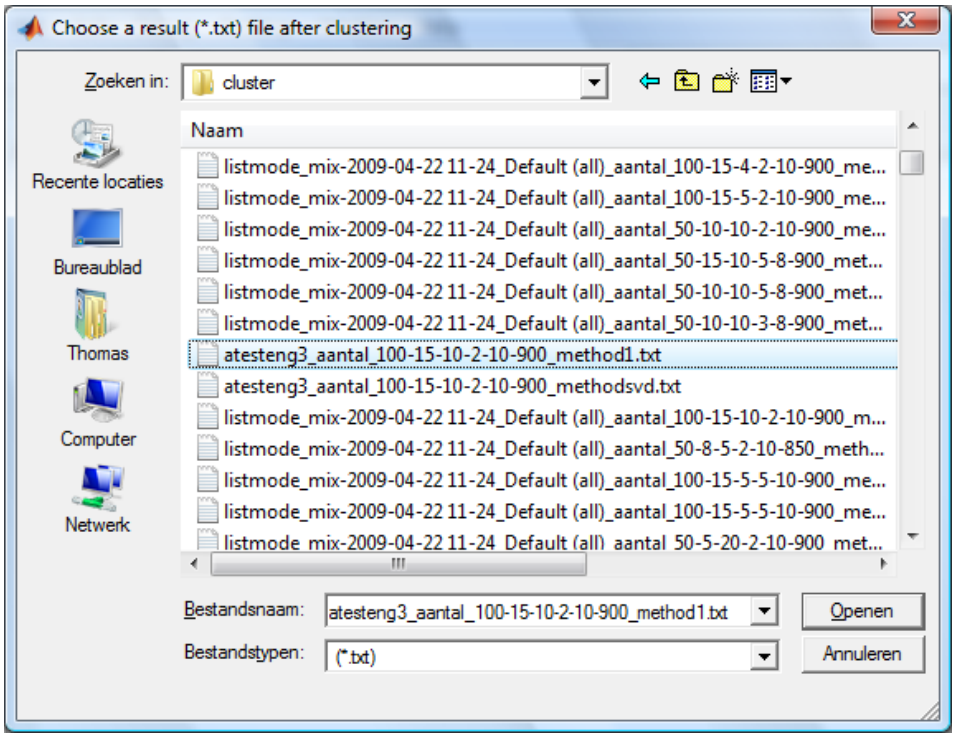
4.3.2 ADD cluster result to database.txt (menu database) (.txt database is not recommended anymore)

To add cluster results (txt-files after clustering) to a database.

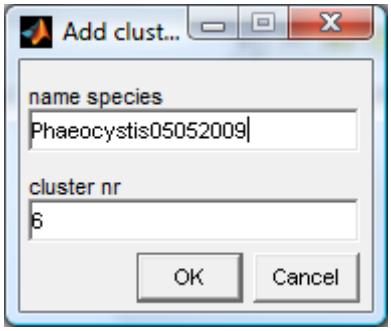
Option ‘ADD ..’ starts with choosing the database (name) where species cluster fingerprints should be saved to. Default database name is phytodatabase.txt (in ..\databases)



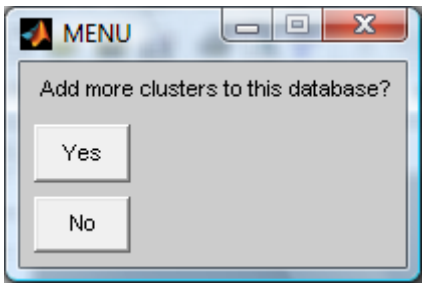
After this, a cluster fingerprint file-*txt file (e.g. from ..\cluster), processed by manual, autocluster or autolasso, should be chosen. In this file, the clusters and their median-fcm-values are stored. Be aware that you know which cluster number corresponds to a species (e.g. validated by imaging in flow or microscopy), which need to be stored in the database.



The species name and corresponding cluster number in the file you imported just before, need to be given here:



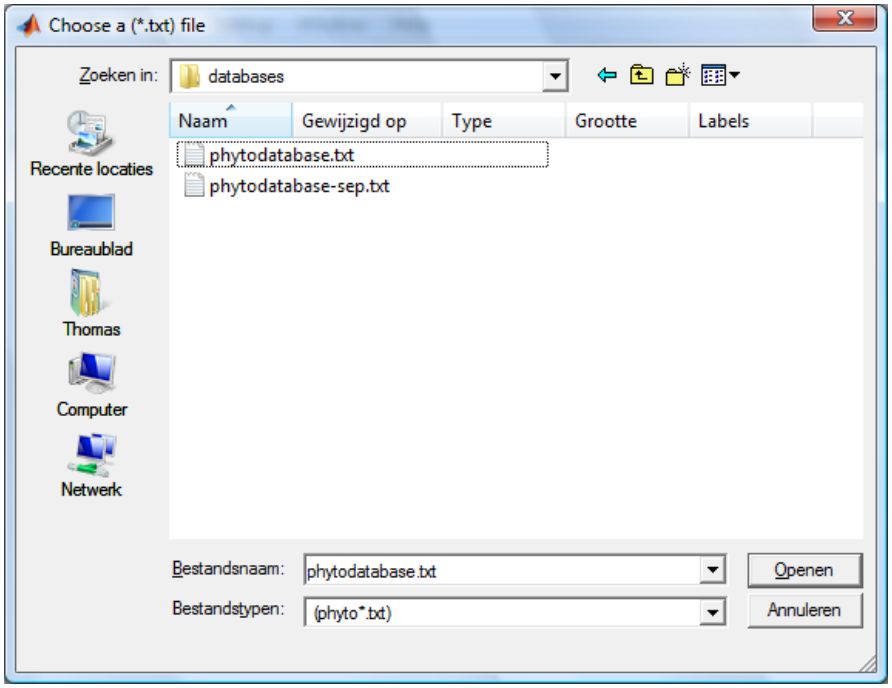
After 'OK', the values are added and saved in the database. If needed, you can add more species names and corresponding cluster numbers of the already imported cluster result .txt file, to the database.



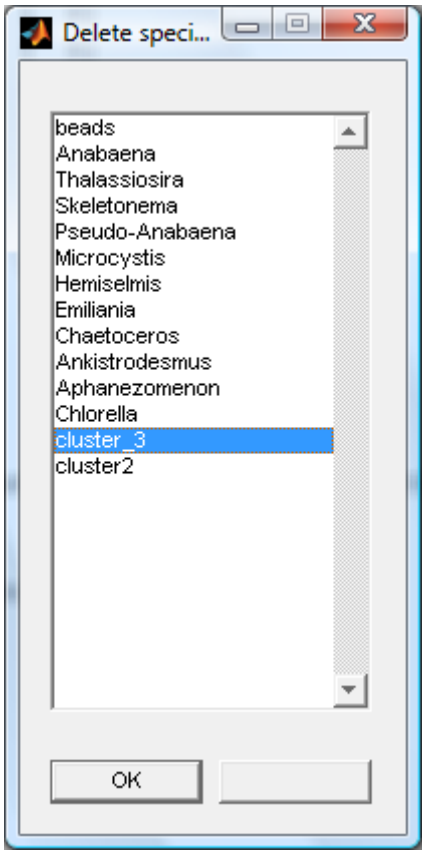
4.3.3 DELETE species from database.txt (menu database) (.txt database is not recommended anymore)

To delete species or other events from a database.

After choosing the 'DELETE' option, the database (name) is asked where species should be deleted from. Default database name is phytodatabase.txt (in ..\databases)
It is recommended to make a copy of the database, which will be changed.



After this, a menu is opened containing all species in the chosen database. Choose a species, which should be deleted from the database and press 'OK'.

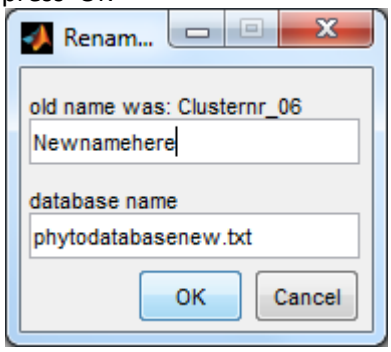


After this, the selected species is removed from the database.

4.3.4 RENAME species from database.txt (menu database) (.txt database is not recommended anymore)

To rename species or other events from a database.

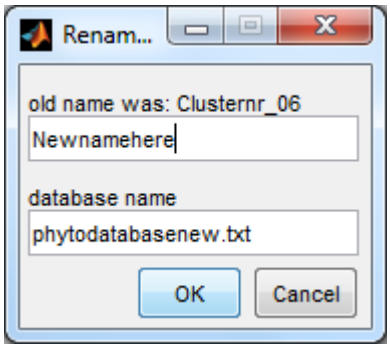
Select database and select species and following menu appears. Change name and press 'OK'



4.3.5 RENAME species in selectionsets

To rename species or other events in selections sets (lasso method)

Select database and select species and following menu appears. Change name and press 'OK'

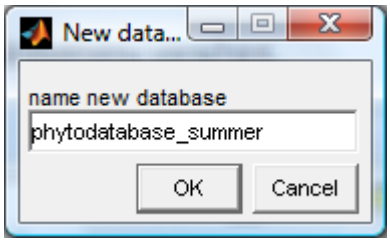


4.3.6 Make a NEW empty database.txt (menu database) (.txt database is not recommended anymore)

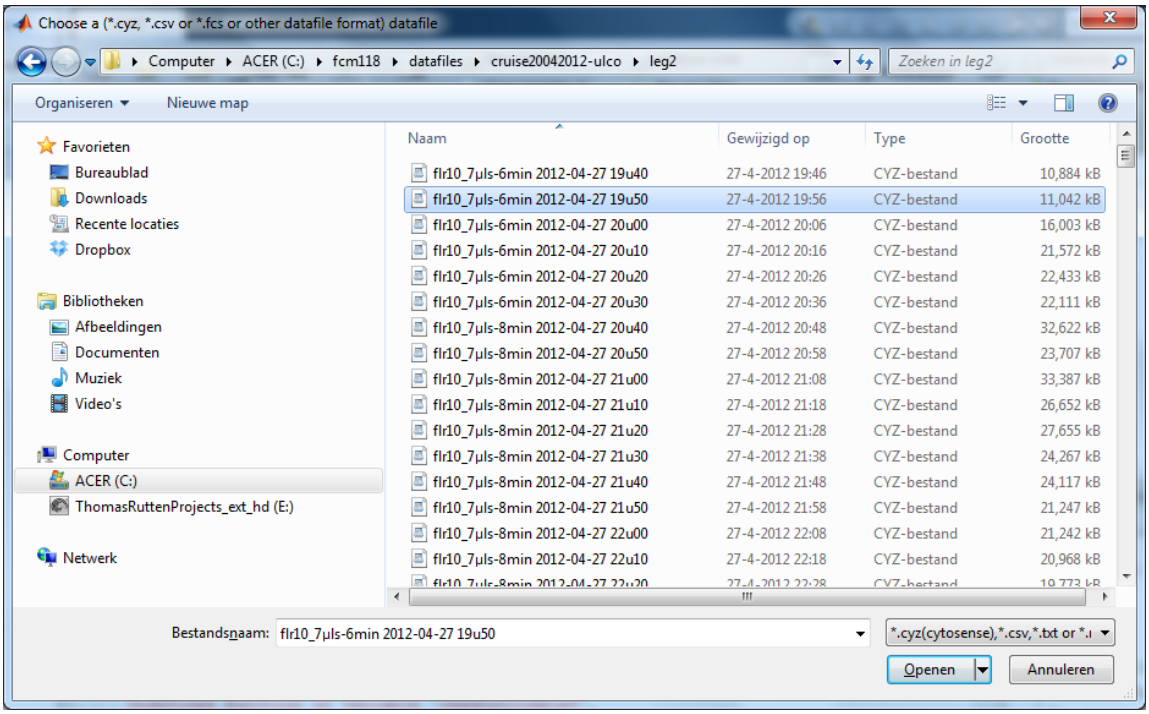
To create a new empty database

In order to create a new (empty) database, a flowcytometric file is needed containing as much as variables that are exported by the flowcytometric instrument. The names of these variables (flowcytometric parameters) are used in the new database. These variables are the column names of the imported flowcytometric file and are used as column names in your new database.

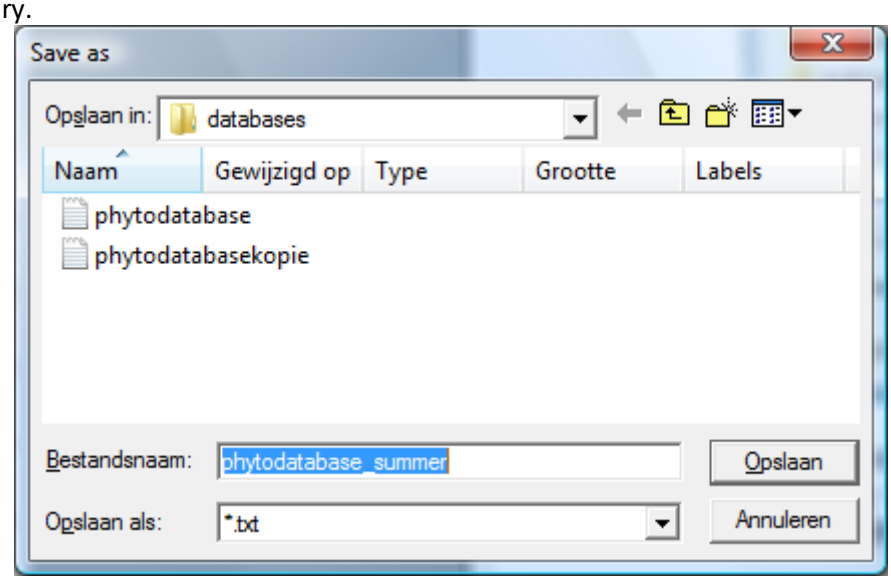
After importing the flowcytometric, the new database name menu appears. Define a new database name in the menu and press 'OK'.



After choosing the database, a flowcytometric file need to be selected.



A confirmation window appears. You have the possibility to change the name or directory.

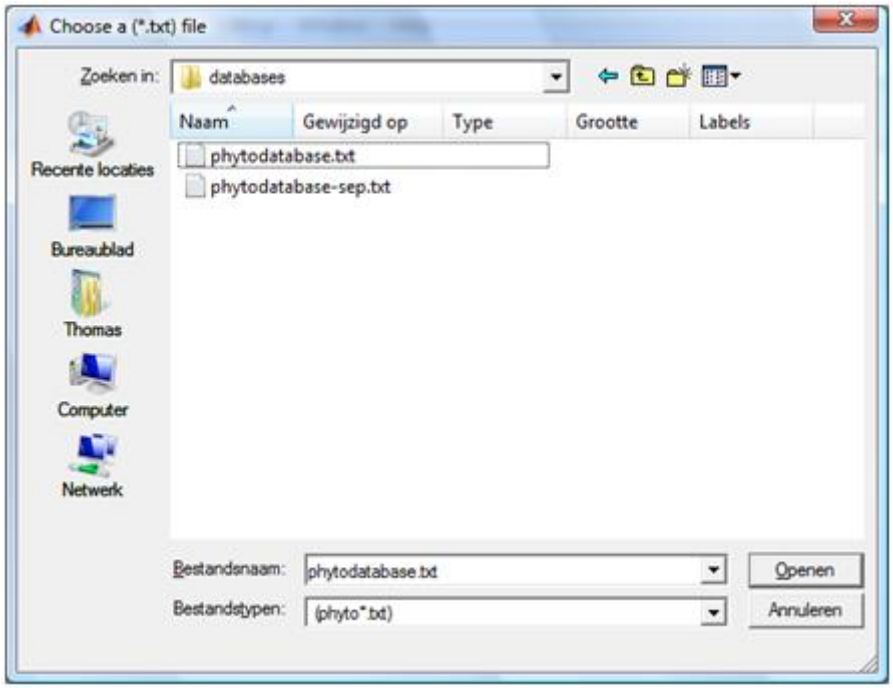


Press ‘Save’ and a new database is created.

4.3.7 Show phytodatabase.txt in scatterplot (hoofdmenu)

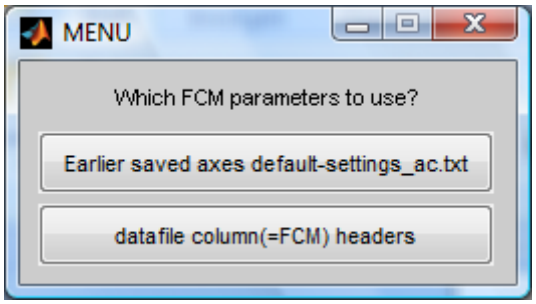
To visualize species from a database in scatterplots.

The option ‘SHOW ..’ asks for a databasename to be visualized. Default database name is phytodatabase.txt (in ..\databases)

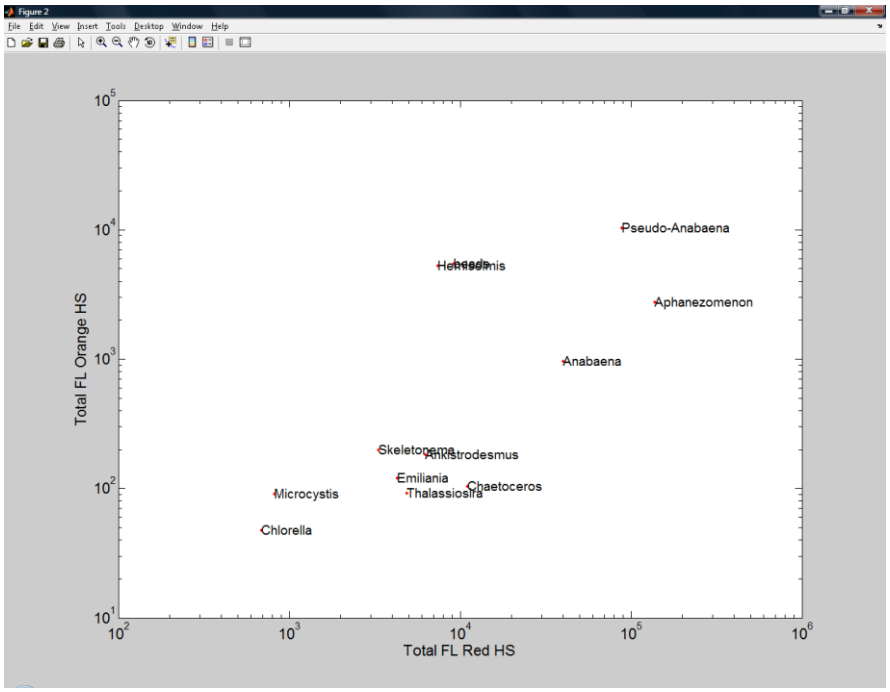


Each species stored in the database is visualized in scatterplots (from the defaults-parameters_ac.txt) file by projecting their names next to a red dot representing the species database values and therefore 'position' in the scatterplots.

The scatterplot combinations are selected using the menu below.



Example of a representation of database species in a database.

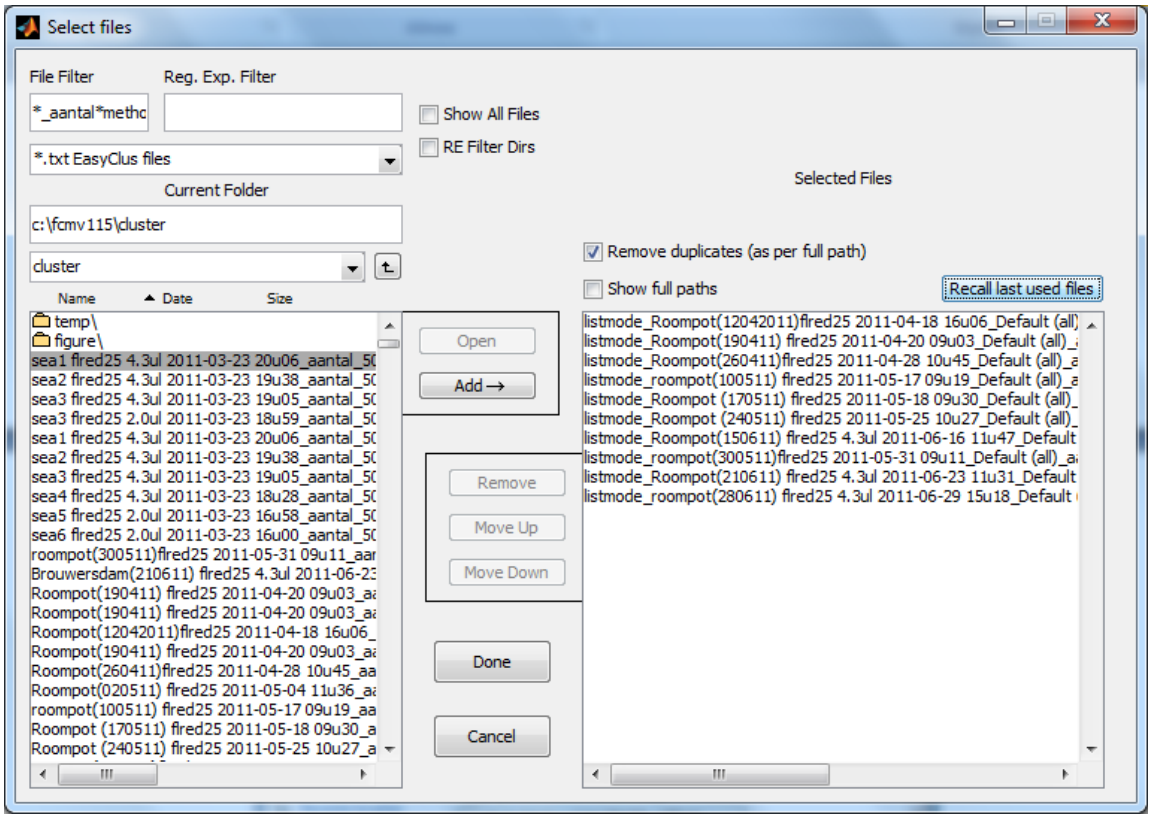


Finished.

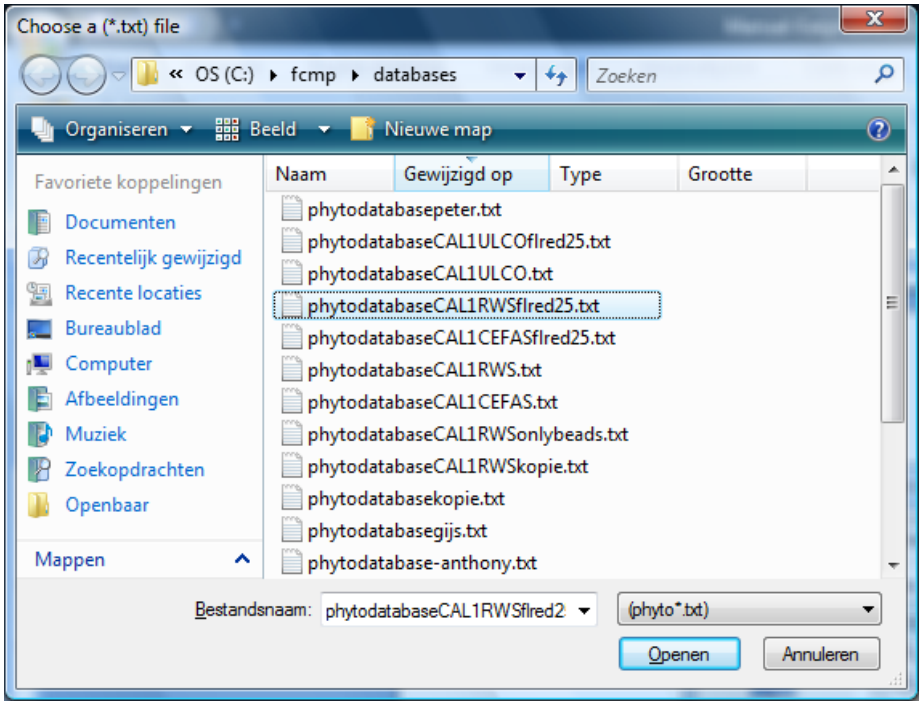
4.3.8 COMPARE one or more cluster results with other data settings or database.txt

To compare cluster results (after method 3a, 3b or 3c) with other database settings or with another database, with already processed clusters.This could be useful to process new data very fast without unsupervised clusterin, because you decided the clustering was okay, butyou would like to compare the calculated clusters with for example a new database or other recognition settings .

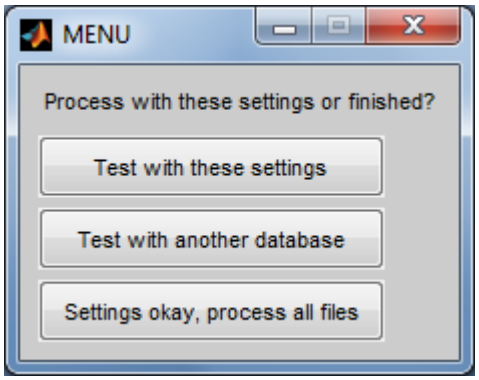
The option 'COMPARE ..' asks for an already cluster results file (usually stored in the ..\cluster\ *.txt directory, which will be imported for further processing.



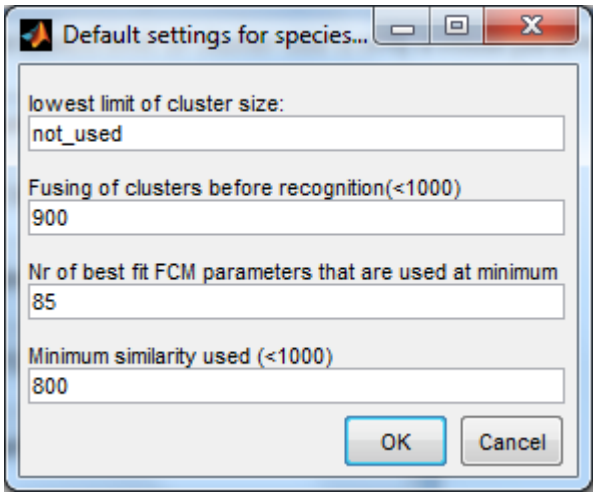
After this, a database should be selected, which will be used for matching cluster results with species stored in this database.



Following menu appears. You can choose whether you want to use default settings database matching settings, another database or return to the database menu



Pressing other settings gives you the opportunity to change database matching settings such as the number of FCM variables that should match at minimum and the minimum similarity value used for positive matching. The lowest limit of cluster size is not used.

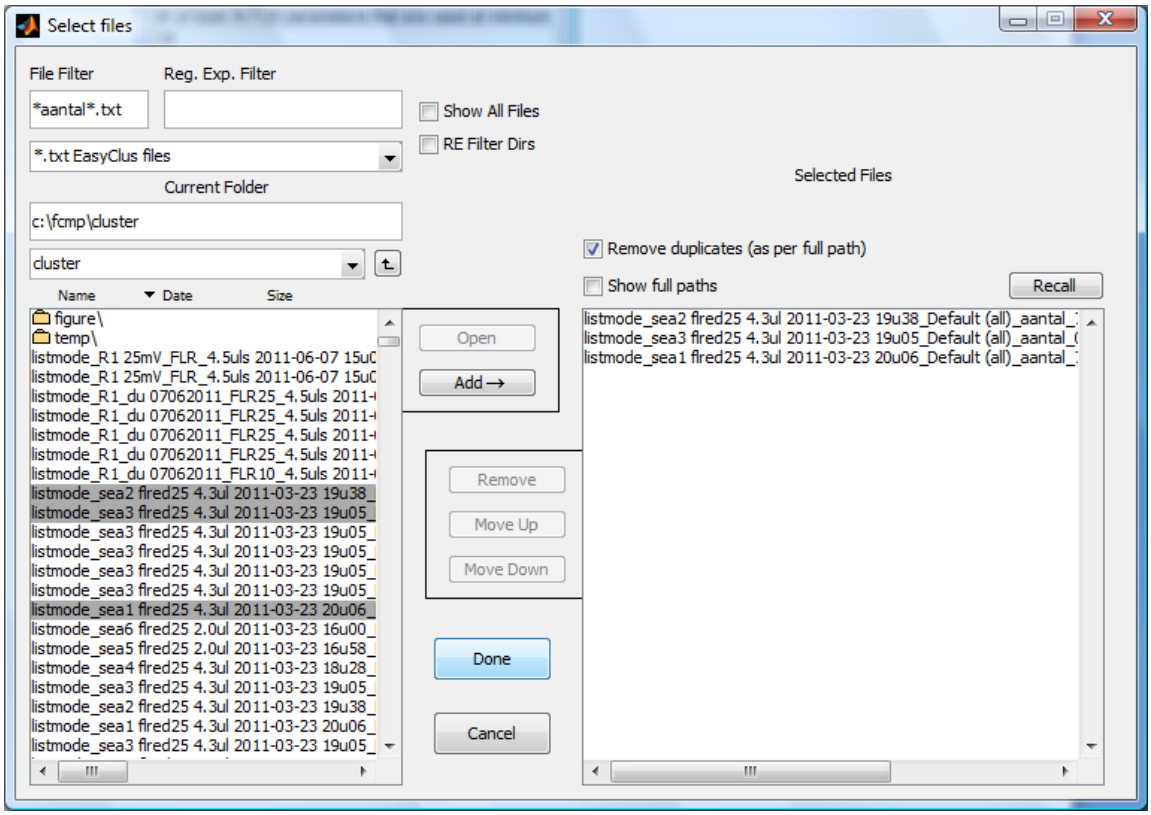


Rematched cluster results are saved in \cluster\ with the extension 'postproc'

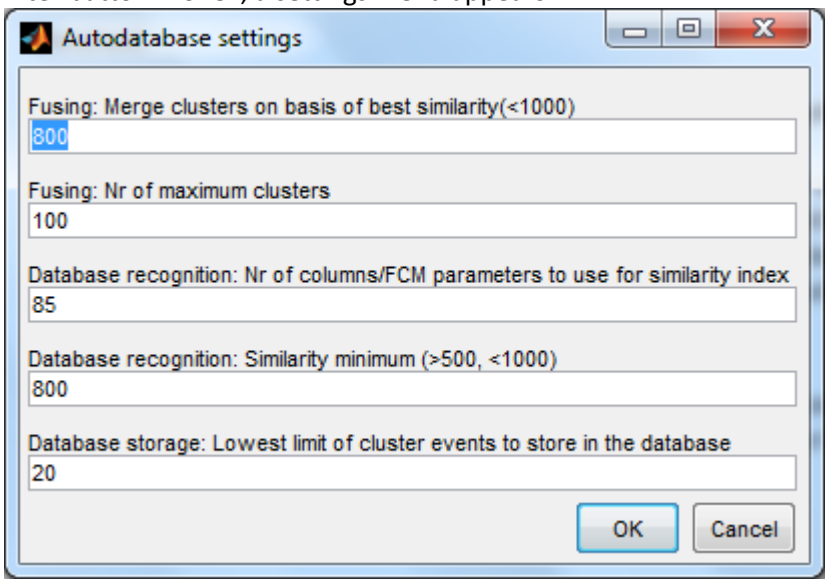
4.3.9 Autodatabasing with 1,2,3 ..n clustered files

After autoclustering (method 3a, 3b or 3c) of several FCM datafiles, it is possible to create and add unique clusters to a new database automatically.

The results after autoclustering are saved in ..\cluster*.txt and the procedure starts with the selection of these cluster result files:



After button 'Done' , a settings menu appears:

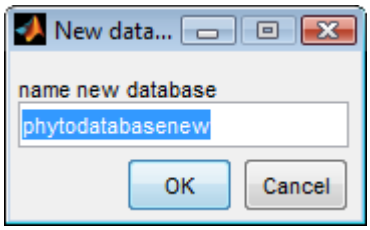


Only clusters higher than the given minimum value of events that each cluster (of events) should contain, will be stored in the database(in example =20).

Each cluster fingerprint is fused according to the 'fusing value given (in example=800)' and matched with the database fingerprints according to setting 'Nr of best fitting .., Minimum similarity ..'. There is a possibility to force the EasyClus software to a maximum number of clusters (example=100), but this method is not recommended, because it is better to fuse on basis of similarity, rather than the maximum nr of clusters.
If there is a match with a fingerprint that is already present in the database, the cluster is not unique and will therefore not be saved in the database. If the cluster is unique, there is no match with any of the fingerprints in the database and the cluster fingerprints will

be saved in the database. Increasing the 'Nr of best fit …' and/or 'Minimum similarity ..' will make the matching process more critical and therefore more unique clusters will be found.

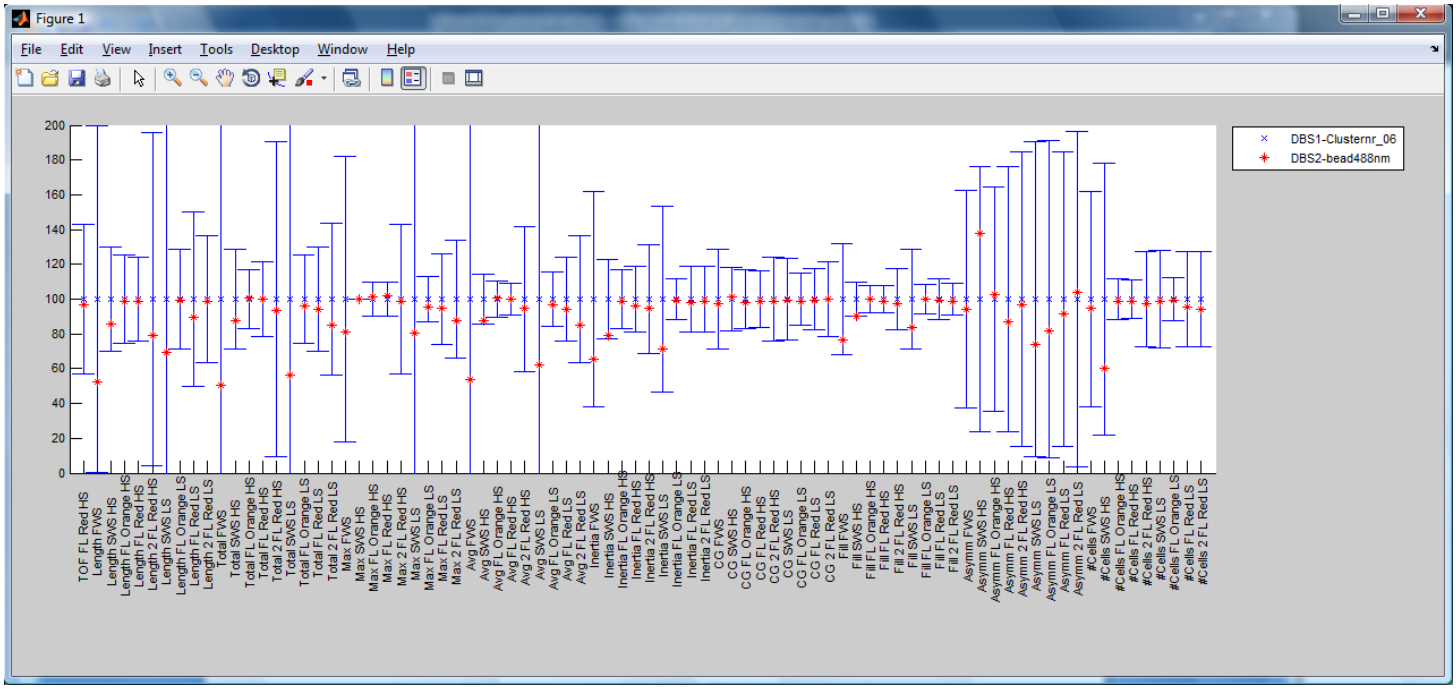
The new name of the 'auto'database , e.g. 'phytodatabasenew' should been given here.



After 'OK' the procedure will be started and in the Matlab command window, the result will be shown such as the number f unique found clusters.

4.3.10 RELATIVE MATCHING of database.txt 1 species with database.txt 2 species

This option is not default available: To validate the database fingerprints (within one or between two different databases) with each other in all FCM parameters together. The blue cross is the average reference fingerprint and set to 100%, the blue vertical line is the relative standard deviation. The red asterix (*) is the relative value of the second fingerprint compared to the reference fingerprints.
You are able to match more than one fingerprint with the reference.



4.3.11 Similarity MATCHING of database.txt 1 with database.txt 2 (all mutual species)

to compare whole databases with each other (which species match with species from another database)

Select database 1, select database 2. Mutual similarity results appears in the Matlab command window. If 'Canceled' in menu Choose database 2, database 1 species are mutually compared with each other.

4.3.12 From ‘database.txt menu’ to ‘Home Menu’

Return to ‘EasyClus Home Menu’ via button ‘Back to home menu’.

4.4 From several files to one (bigger) file – button – ‘from datafiles to single file (randomly)’

To make a merged data file out of several sample files by subsampling

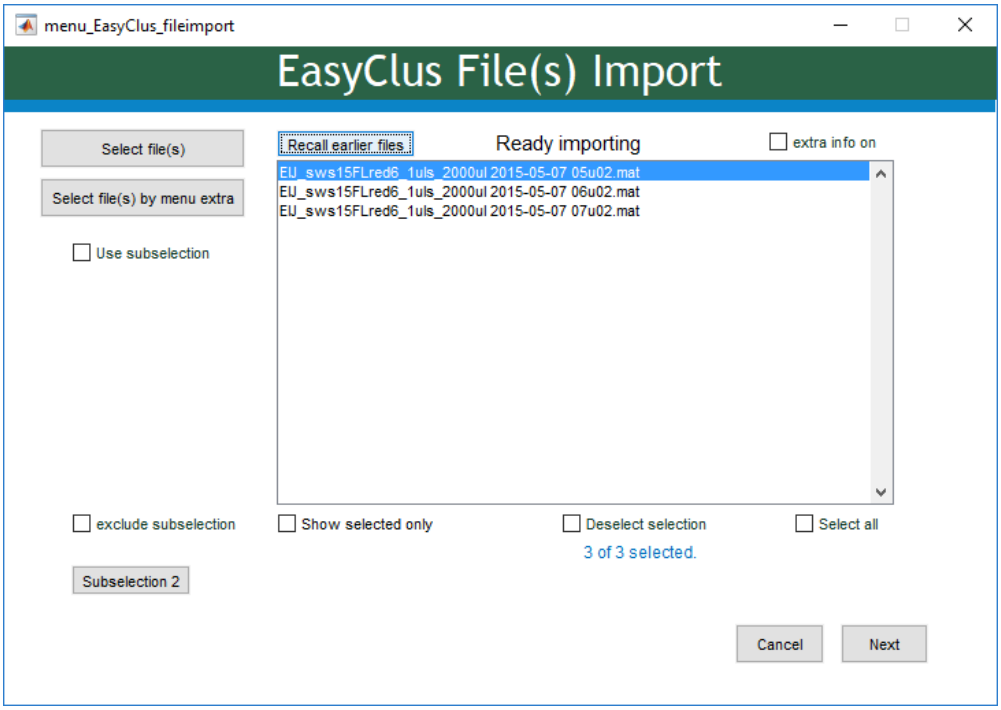
NEW Use ‘?’ for an explanation how this option works

Goal : Merge several (.cyz or .mat) files together to one .mat, .csv, .txt or .xlsx file. In case of merging several files to one .mat file, there is the option to merge ONLY data with IMAGES, so one file of ONLY IMAGES (and pulses) out of several files can be made. This can be useful for making a database.
Important : This option is NOT meant for quantitative studies (about counts and concentrations) because it is just put together. *Merge pico-nano-micro (other command elsewhere in EasyClus) with different trigger is meant for merging different triggered files.*

Sometimes it might be useful to take a sub selection of randomly chosen events in each file and to put all these sub selections derived from several files in one big file. The new file created is a collection of events representing all species found throughout a whole set of files. This can be useful to have an idea of the variation in species in several samples or files, but also to collect the variation of the same species caused by environmental circumstances or instrumental variation.

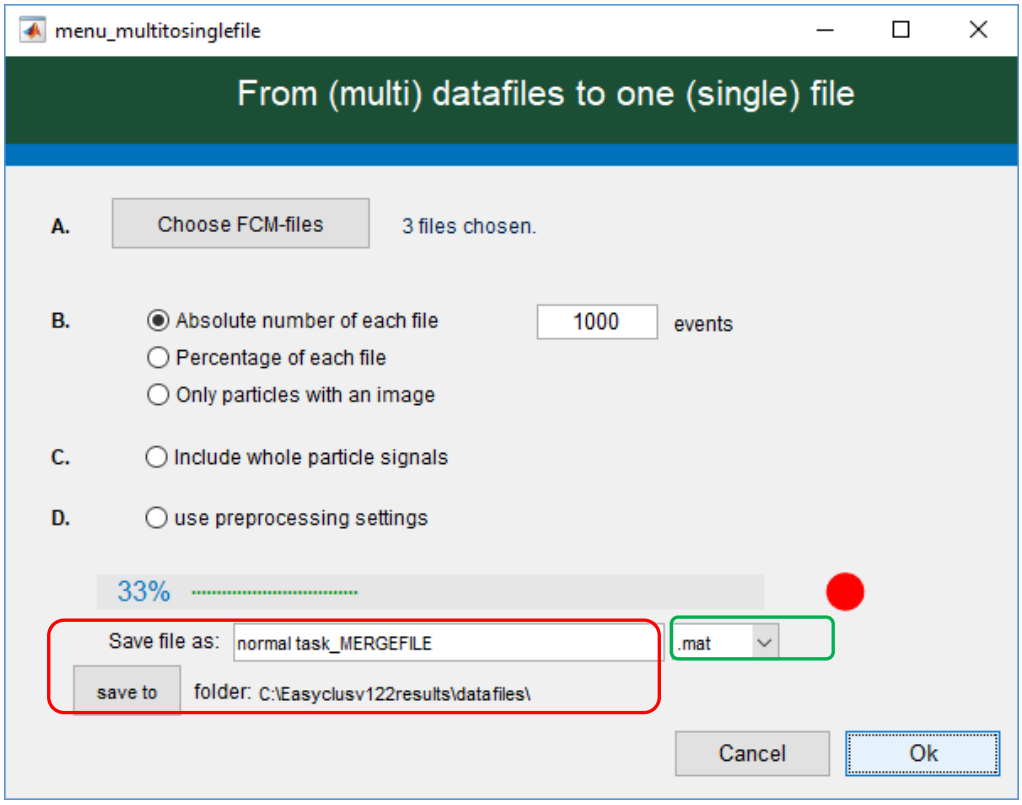
New, since version v1.30 is the option to merge particles which have an image only out of several files. For example to create a big file only containing particles with an image. This can be useful for clustering and building a database quickly.

Example 1 : merge files without image storage, using a fixed number (ABSOLUTE) of counts in each file:



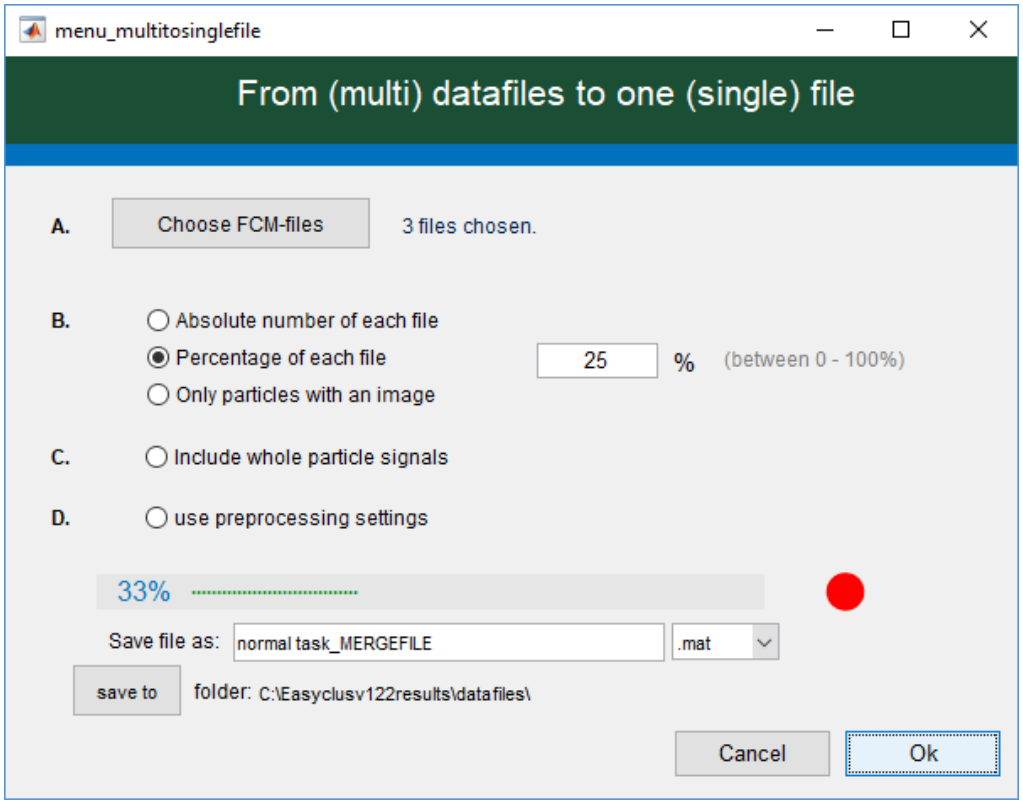
If started, the operator is ask to take a percentage of each file (10%) or an absolute number of events per file (e.g. 1000 using 20 files, e.g. 500 using 50 files)

Here 1000 events of each file are selected.



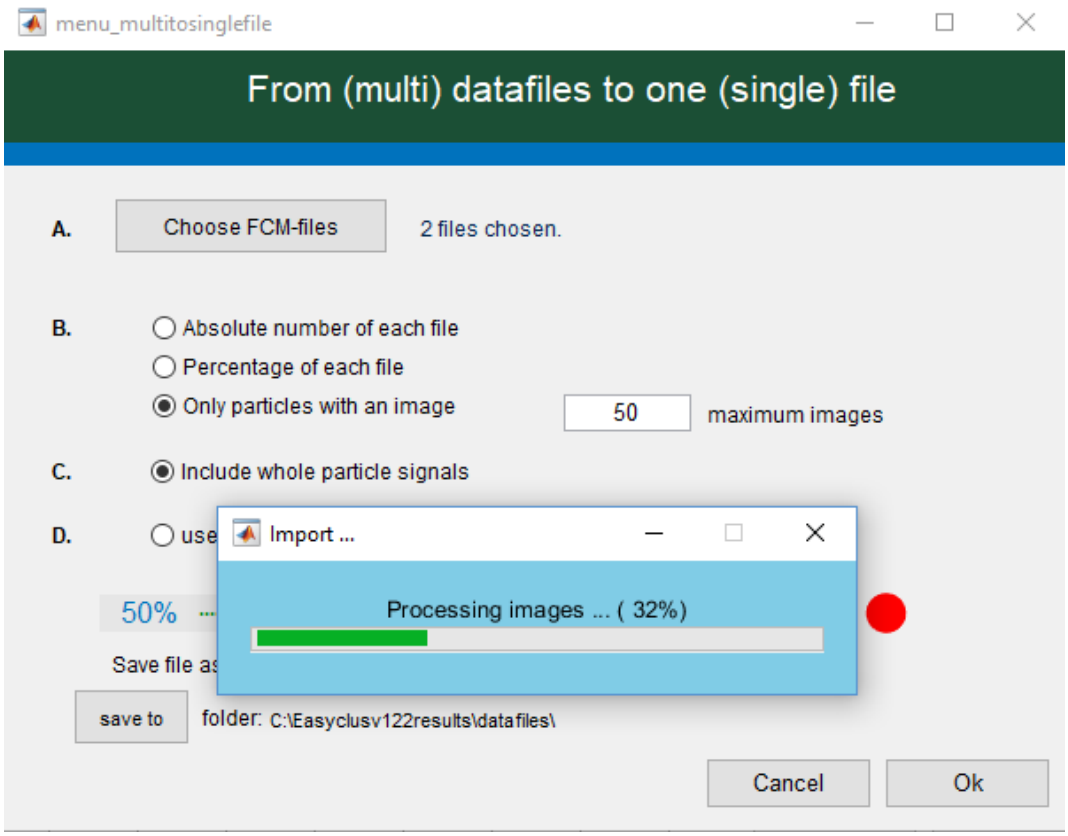
Merged events file is saved as given in
As .mat or csv-file or .txt, or.xlsx file.

Example 2 : Here 25% of all events in each file are selected



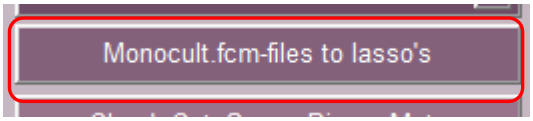
Important to mention here is that the selected events of each file are marked as a clustered file in a xxcoll_AC2_idnrs.mat file, which can be used for comparing cluster results with the original merged file here. This can be useful when merging monoculture files with each other.

Example 3 : Here ONLY IMAGES of all files are selected with a maximum of 100 images of each file



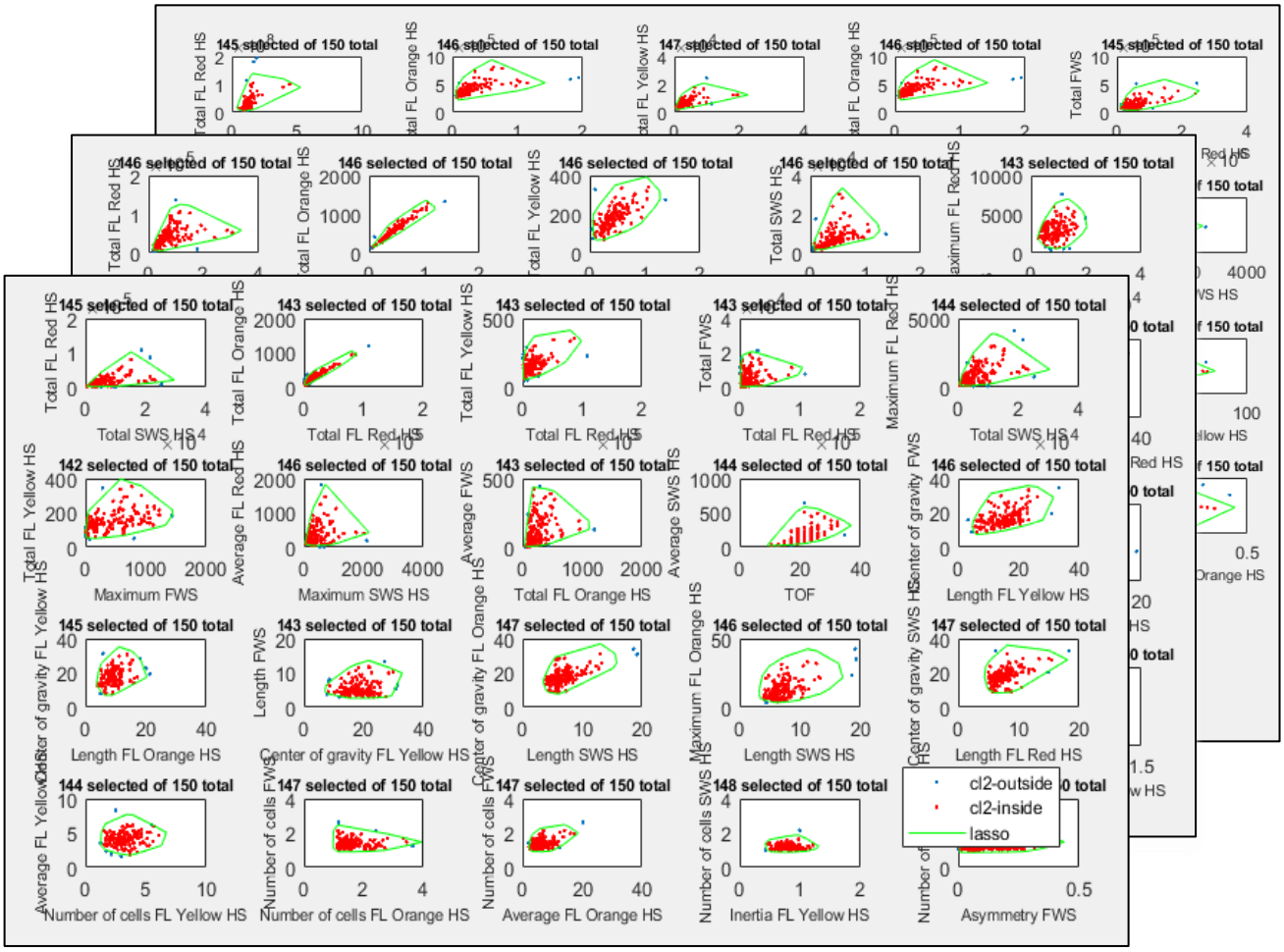
4.5 From pure datafiles to selection sets

Tool to produce LASSO's from monoculture datafiles



When using the LASSO clustering, you need selection sets or lasso's of species. These selection sets can be processed after the other cluster methods by the 'process selection set' button.
If you have cleaned monoculture FCM data and you 're 100% sure that it is not contaminated, this method is the fastest way to produce selections sets. Just select the (cleaned) datafiles and the selection sets or lasso's are produced automatically.

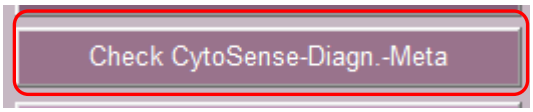
Select files and press OK



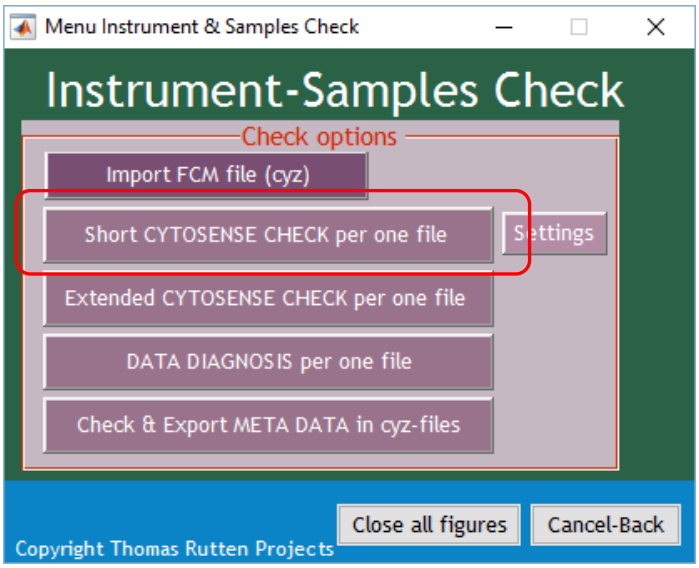
4.6 Check CytoSense

Tool to check instrument performance and meta data settings

Button 'Check CytoSense' can be used to check a lot of features features of the CytoSense such as temperature, curvature optics and block shaped signals.



The next window appears:
Import FCM-file first



Button 1: Short CYTOSENSE check

This is a traffic light check for each cyz-file. Green is ok. Red is not ok.

Button 2: Extended CYTOSENSE check

This is a strongly recommended check if you're not an expert. On basis of your results, EasyClus gives you recommendations, critical remarks and consult to your instrument settings and analyses settings and results.

Button 3: DATA DIAGNOSIS

Gives you an idea how your data in one file is distributed or if there are sudden unexpected changes in your data.

Button 4: Check & Export META DATA

A very useful tool to select all kinds of meta data out of your files and export them as .txt files for plotting, or exporting (by a menu selection) for storage in big data storage systems.

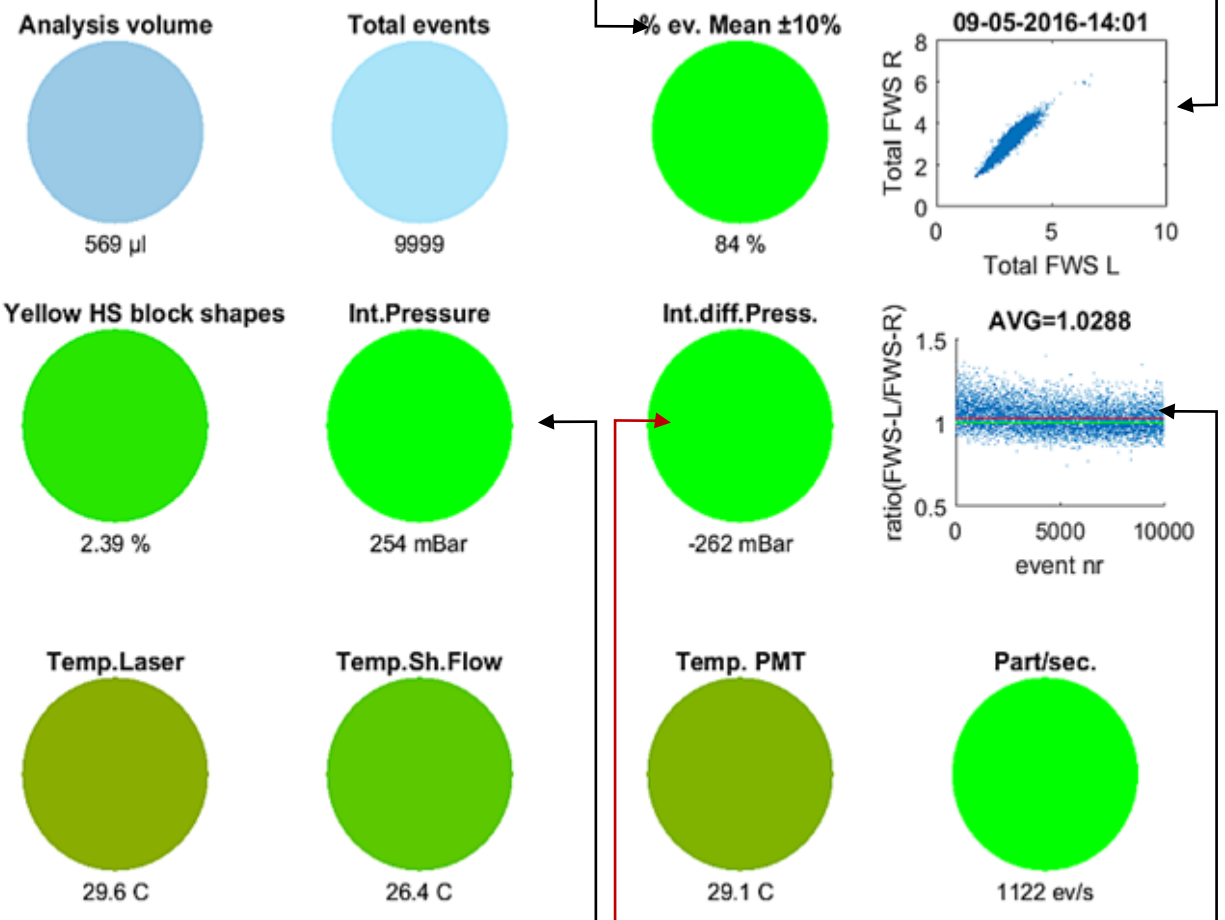
4.6.1 Traffic light check

Analysis volume : All data in the FCM file is taken from the given volume of sample in microliters (10^{-6} l) and is lower than the FCM pumped volume (not given here).

% ev.Mean 10%: The percentage of particles between .9 and 1.1 of the ratio Total FWS R / Total FWS L. If it is 84%, it means that 84% of all particles is between .9*ratio and 1.1*ratio. If it is less than 30%, 70% of all particles is outside this bandwidth, which might be an indication of fluctuations of the core for example caused by air in the fluid path.

Total events: The number of particles (= events) that has been stored according to the trigger settings.

Plot 1: Scatter plot Total FWS L vs. Total FWS R. Each particle gives Forward scatter on the left and right curvature detector. The value of the total signal area left and right is prejected in this scatter plot. It should be a linear relation. Title is the date and time of the measured sample.

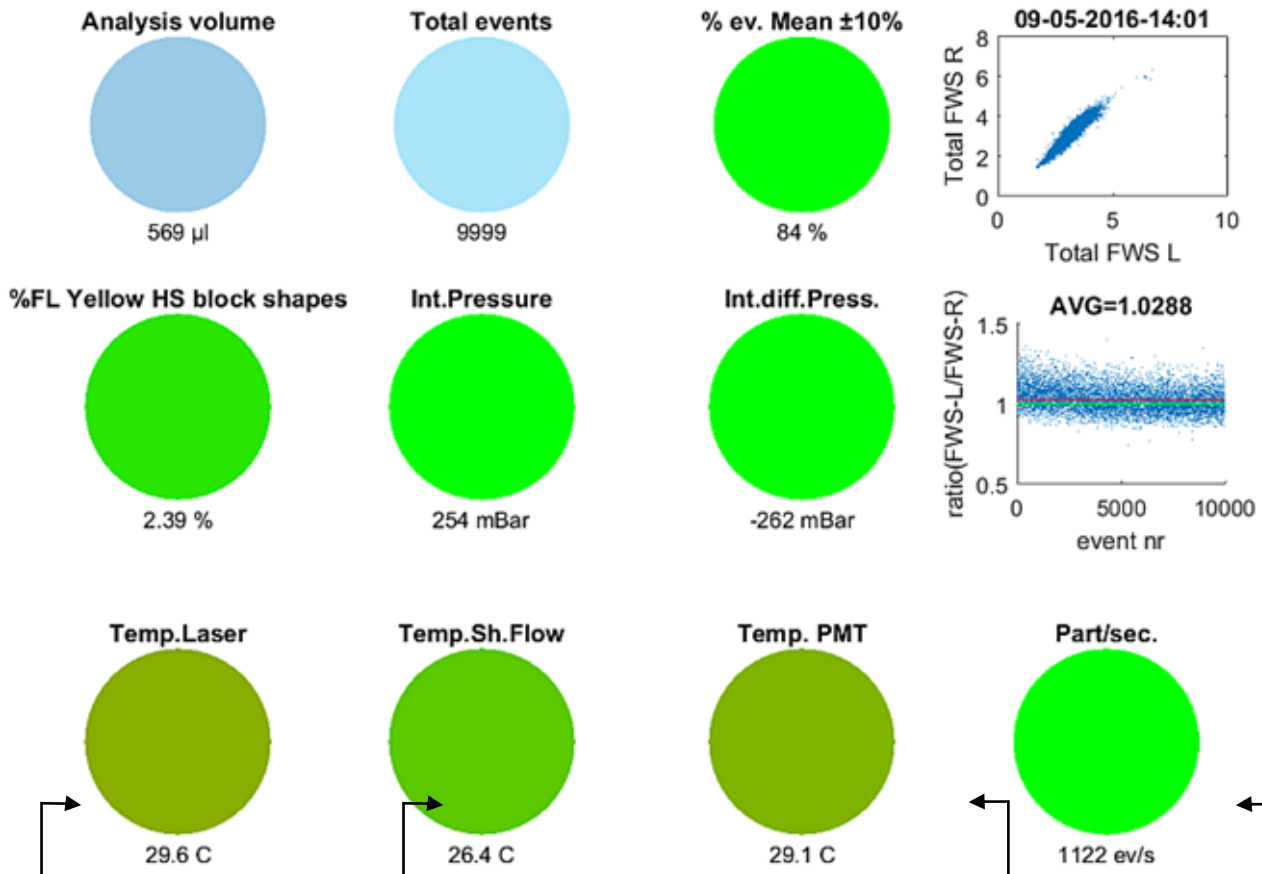


% Par block shapes: The highest found percentage of block shaped signals in the given parameter. Sometimes caused by very small signals just above threshold, sometimes by other not well known reasons.

Plot 2: Scatter plot event nr vs. ratio (FWS-L/FWS-R). The ration value of FWS-L/FWS-R. 1=perfect, but some difference is not a problem becasue it is very sensitive and some particles are irregular shaped. The AVG=1.0288 in the title is the average ration of all particles.

Int. Pressure: Internal Pressure of the fluidics in mBar. If it becomes higher than 900 mBar, it indicates that the internal pressure is increasing towards the outside pressure, e.g. by blockage of the fluidics path (filter, kuvette channel).

Int. diff. Pressure: Internal Differential Pressure of the fluidics inlet in mBar. If it becomes higher than -900 mBar, it indicates that the outside pressure is increasing towards the inside pressure, e.g. by blockage of the fluidics path (inlet tube).



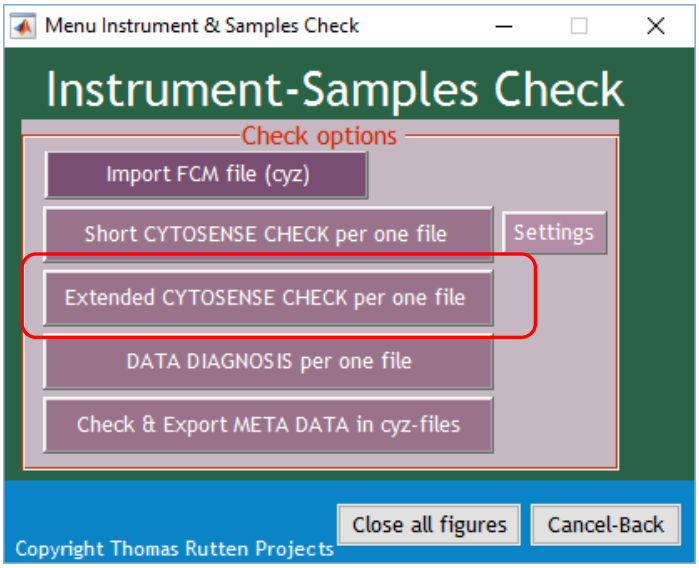
Temp. Laser: Temperature Laser is the temperature in Celsius in the laser to check if it is not too high. Temperatures higher than around 36-39 °C are high.

Temp. Sh. Flow: Temperature of the Sheath Flow (internal cleaned sample stream which takes the sample to the laser) in Celsius. Temperatures higher than around 35 °C are high.

Temp. PMT: Temperature inside the PMT detection house in Celsius. PMT's are temperature sensitive (more noise with increasing temperature). Temperatures should not fluctuate too much. Higher than around 35 °C is not recommended.

Part/sec.: Particles per second that crosses the laser beam. Each particle from approximately .2 µm and larger, which crosses the laser beam, will give signal in the detectors. If the particles are too close after each other, they will be interpreted as one signal. More than one particle in the laser beam at the same time is called coincidence and should be avoided. The chance of coincidence increases from particles rates of 5000/sec. and higher. Decrease the sample rate to decrease the particle rate.

4.6.2 Extended CYTOSENSE Check

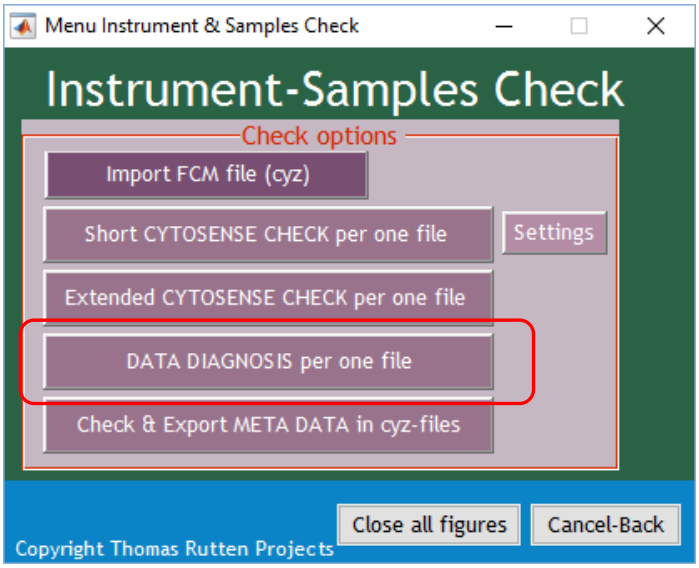


The extended check gives the FCM-file settings overview as text in the Matlab EasyClus command window, an overview of some analyses by EasyClus about the data and meta data and warnings and recommendations to improve the analysis and instrumental settings.

Example of extended file check (just a part if it..)

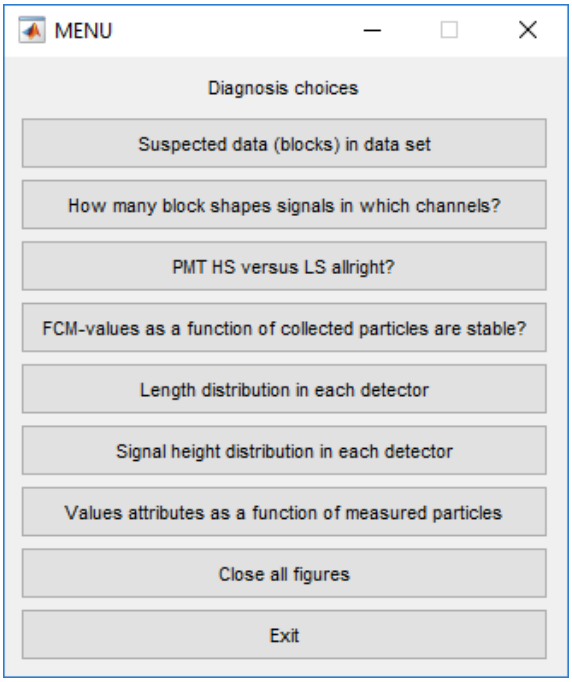
```
Nr of particles or events    : 2226
Nr of FCM variables/ columns : 51
Nr of found images is        : 39
.............................

------------------------->  <- info --------------------------

1. The cytometric analysis is performed on machine MarseilleIII
2. The CytoUSB software used is CytoUSB Version: 5.7.6.2
3. The filename is
4. The analysis is performed on 06-07-2016 time 14:45:49
5. The number of measured particles is 2226
6. The total concentration found (e.g. only SWS trigg.) is 488 particles/microliter
7. The concentration found (all used triggers) is 7 particles/microliter
8. The preconcentration (before analysis) found (all used triggers) is 7 particles/microliter
9. The flush time found (between samples) is 60 seconds
10. The effective analysis volume is 304.9754 microliter
11. The pumped volume is 591 microliter
12. The particle-rate is 960 particles passing the laser beam/second
13. The used sample flow rate is 1.9682 microliter/second
14. The used block size memory is 4 kB
15. Trigger 1: SWS at threshold level 30 mV
16. Trigger 2: Maximum FL Red at threshold level 30 mV
17. Temperature sensor PMT is : 29 C
18. Temperature sensor Sheath is : 27 C
19. Temperature sensor Sheath is : 30 C
20. Pressure sensor difference is : -373 mBar
21. Pressure sensor absolute is : 369 mBar
22. Sample pump 1: Last calibration of the sample pump was : 04-Jul-2016
23. Alignment check 1 by curvature detector is 1.0346 (1=perfect)
```

4.6.3 Data diagnosis



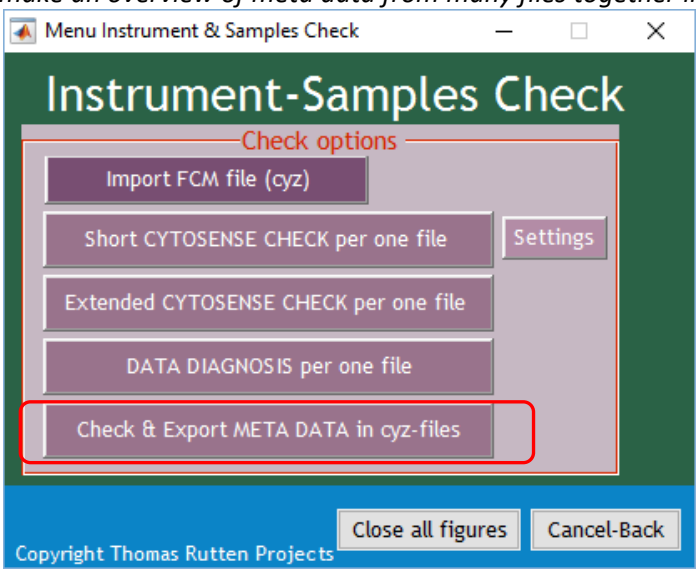
The data diagnosis tool enables you to look at your cytometric data in a different way. For instance the first button (suspected blocks) is a tool that can be used to figure out if randomly distributed and expected data contains blocks of not randomly distributed data.

There are more options but not explained in detail here. Just try and figure out.



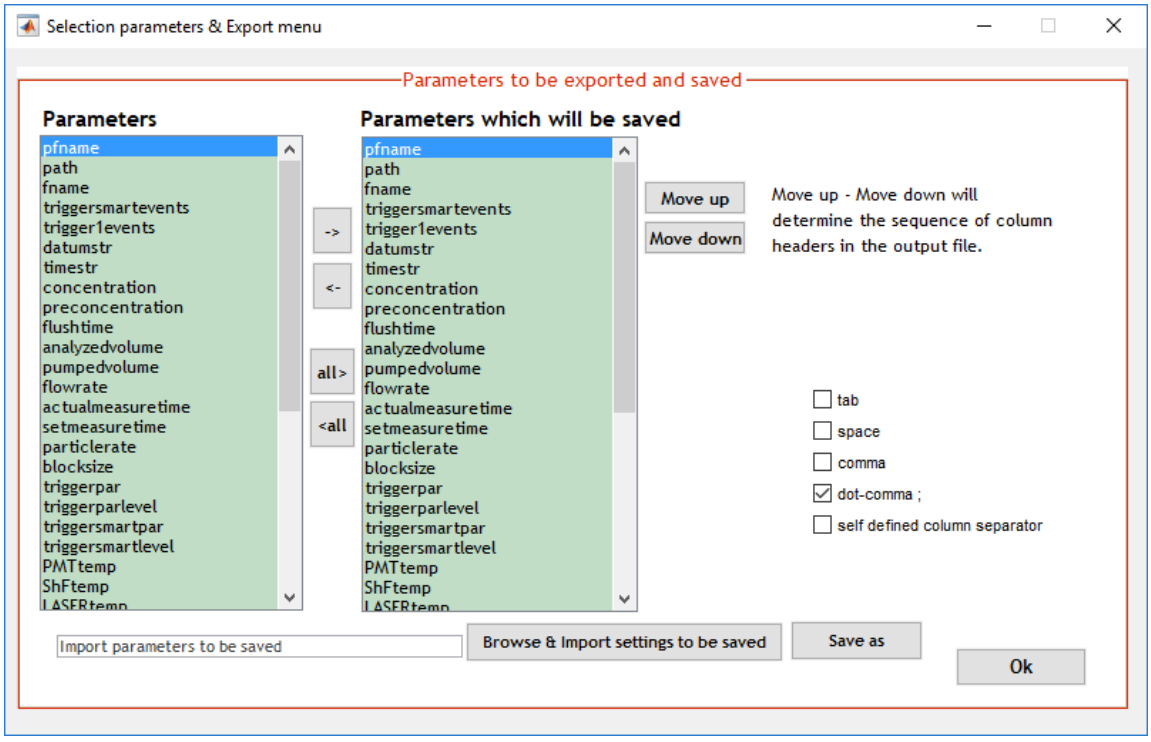
4.6.4 Check & export meta data

Tool to make an overview of meta data from many files together in one file.



Meta data derived from many files can be put in a separate file :

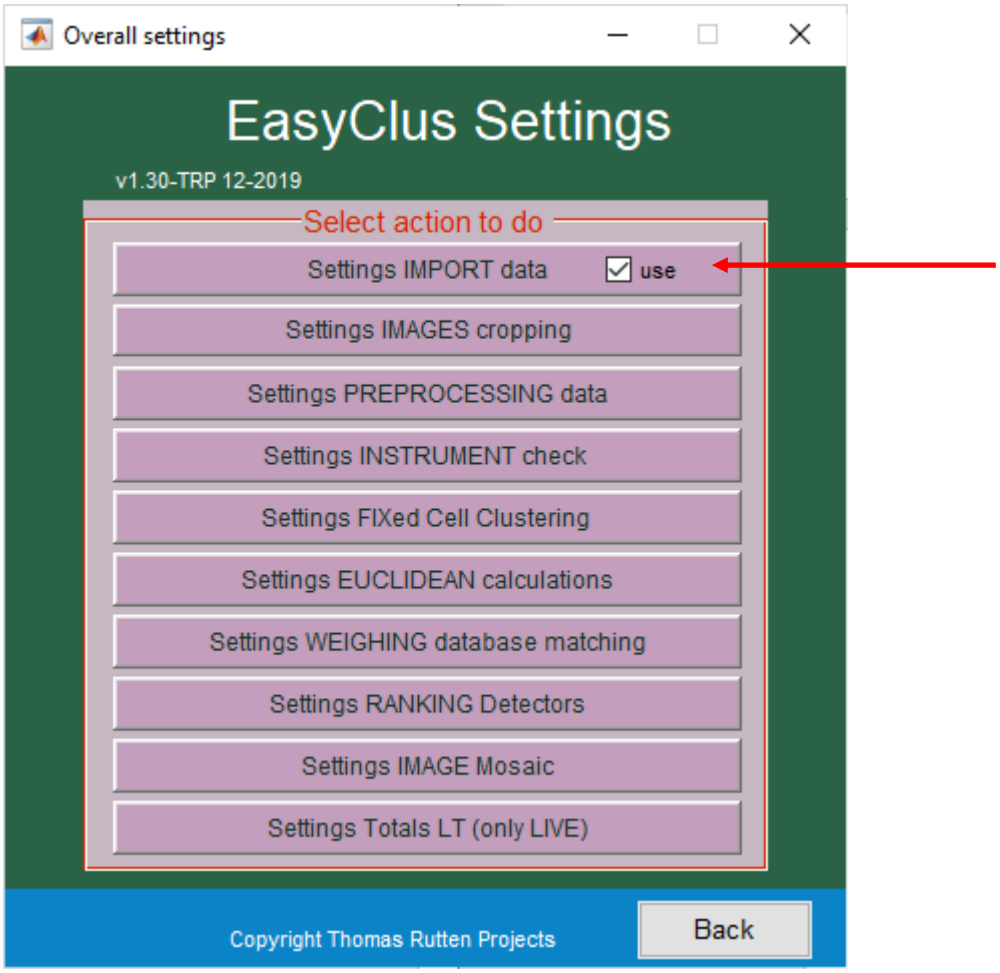
Select cyz-files and after that select what meta data parameters to store in one big file.



Very useful tool to analyze your analysis and instrumental (meta) data quickly, also in combination with plot module tool.

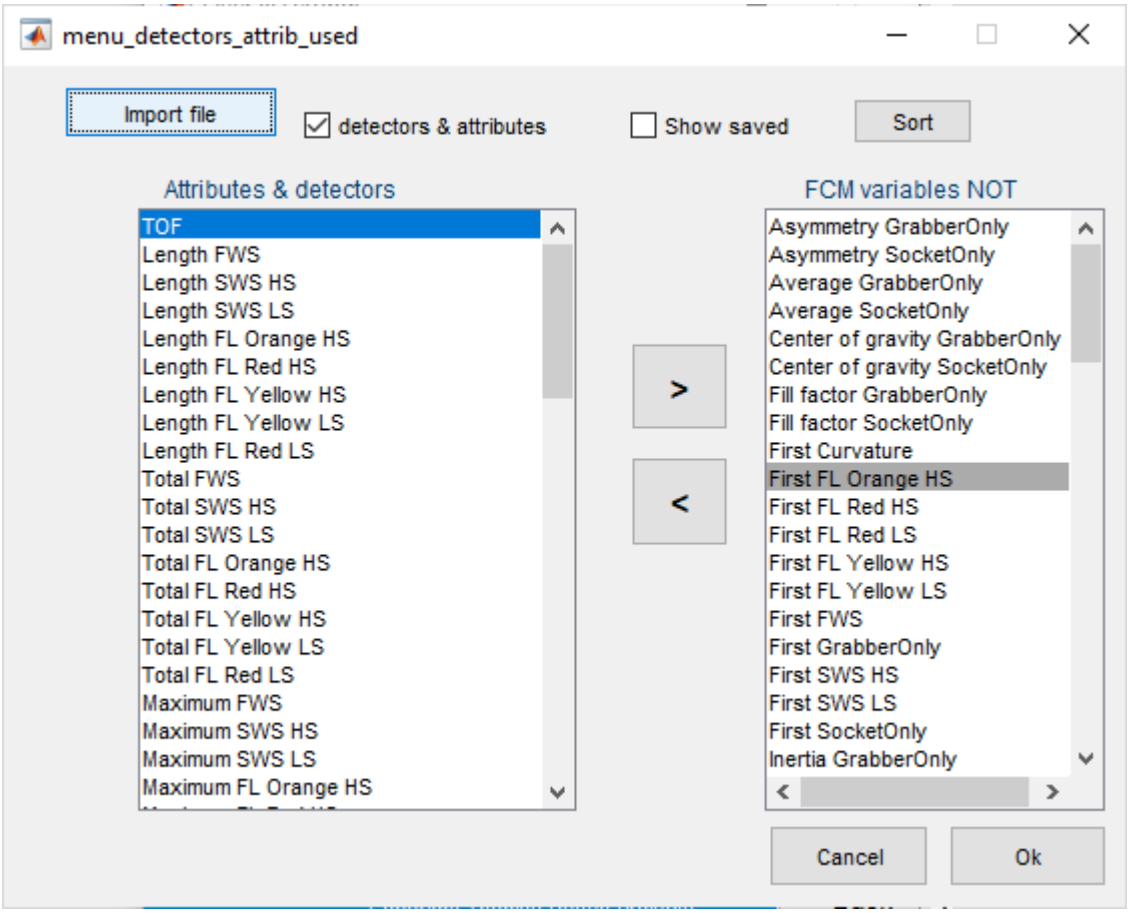
4.7 Overall Settings (very important tool to understand)

Overall settings is introduced to set all kinds of settings:

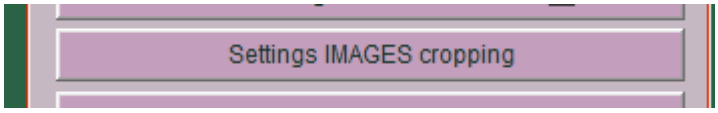


4.7.1 Settings IMPORT data

The cyz-files contain particles profiles. Attributes like Length, Total, Maximum and more are calculated from these profiles for each detector and are used for the clustering and characterization of species. Due to unknown reasons, new surplus attributes data have been added since 2018 to this list such as ‘First’, ‘Last’, ‘SWS Covariance’, and ‘Time of Arrival’, which might negatively affect the clustering and/or classification process. Why? They add redundant and non-specific information to the existing data, giving more weight to non-informative data. Therefore it is decided to add the opportunity to EasyClus to sett off the use of specific data. In the default mode of EasyClus, the headers starting with ‘First’, ‘Last’, ‘SWS Covariance’, and ‘Time of Arrival’ are ignored.
Nevertheless, the user is able to set its own data on or off by this settings menu.

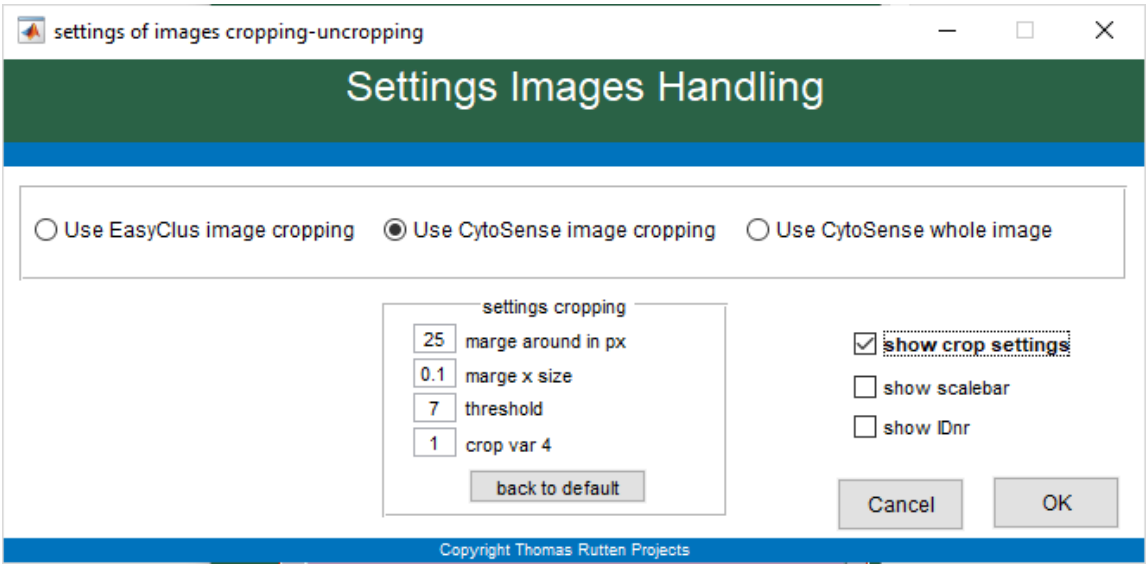


4.7.2 Settings IMAGES cropping



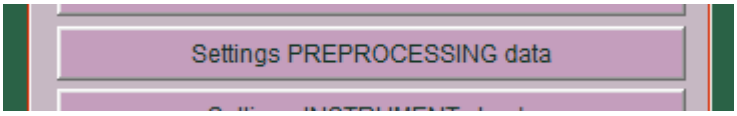
The images made by the CytoSense-Sub, are meant for small as well large particles. The size of the image itself is therefore usually big compared to the particle in it, which makes the visualization of the particles time consuming (a lot of zoom needed) and the storage of the images unnecessary big. Therefore the cropping - the automatic cut off the particle out of the image - of the images is needed and strongly recommended.
This can be done by the CytoSense software or by the EasyClus software. In the past the latter method was recommended, but from 2019 and later the newer CytoSense method is recommended because of speed and effectiveness.

Choose therefore Use CytoSense image cropping (=default). Default settings are used here. If they do not meet the cropping results you like, you can change them here.

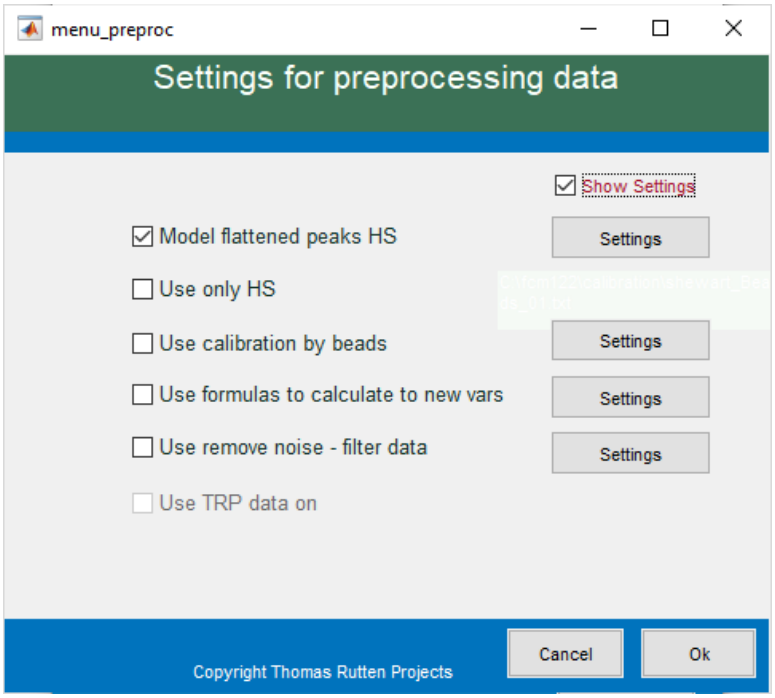


For both methods some settings can be adjusted.

4.7.3 Settings PREPROCESSING of raw data



The preprocessing of raw FCM-data is possible to improve the quality of data. There are several options (cyz-file related):
- Profiling flattened HS signals on basis of LS signals
- Use of only HS signals and ignore LS values
- Correct FCM data towards assigned values of reference beads and the closests reference beads results
- Calculate new variables (eg. chlorophyll) by using own formulas for each particle



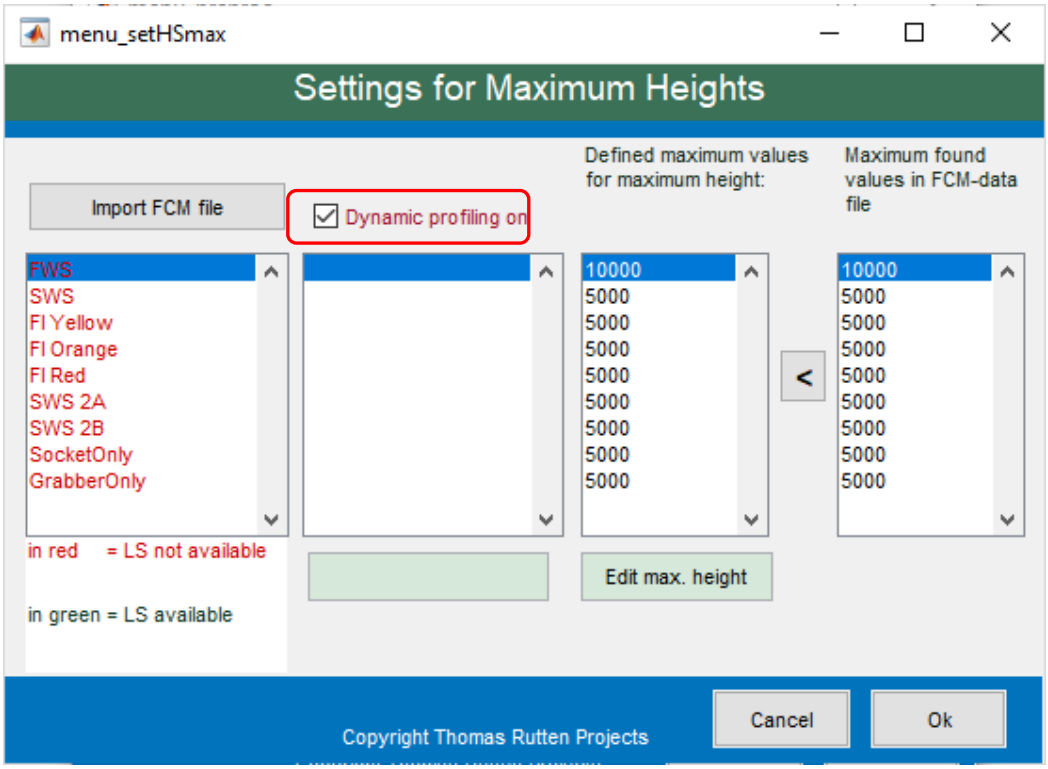
Preprocessing is performed on raw data according to the 'tick on' checkboxes in the settings menu below and their underlying settings.

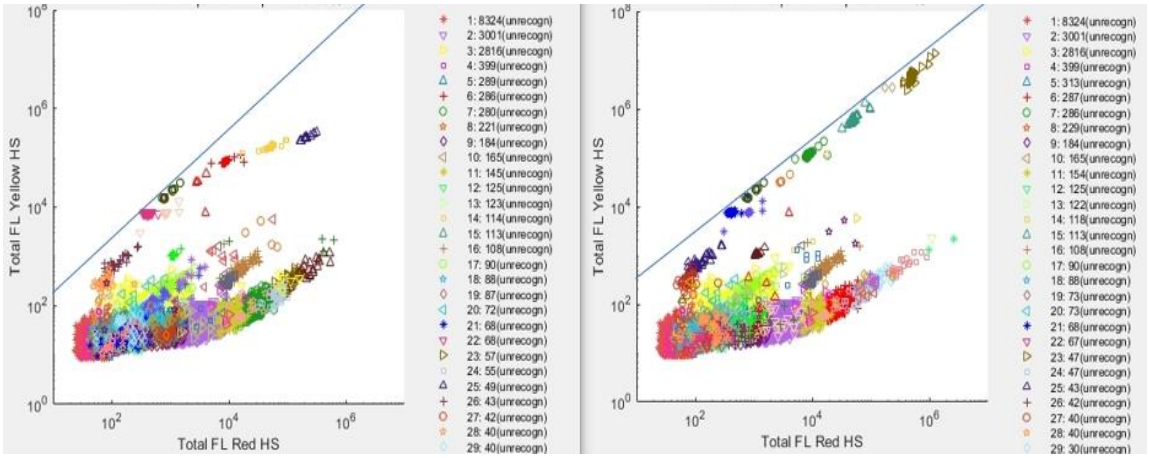
Settings Profiling flattened peaks HS signals on basis of LS signals

After importing a file, the settings menu shows which FCM HS detectors have a LS detector (in green) and which have not (in red). Those who have HS & LS detectors will be 'profiled' for flattened peaks. The maximum detector value possible should be defined in the column in the centre and is default taken as 5000 mV. This value is for some CytoSense too high, for some too low.

For signal without twin LS-detector, the flattened peaks will be redefined by estimation as well.

= new: tries to shape saturated peak on basis of other detectors if not LS is available





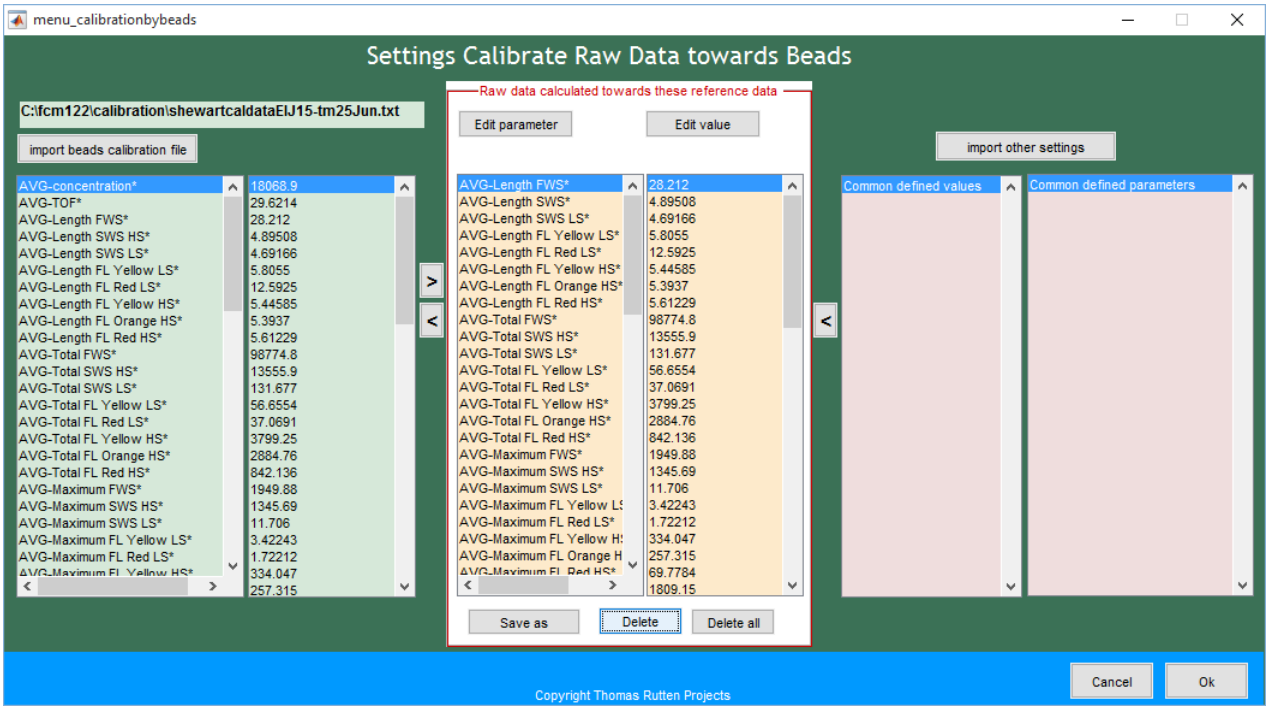
Result without and with pulse saturation correction (system without LS detector)

Settings use calibration by beads

Raw FCM data attributes (Length, Total, Maximum, ..) is recalculated on basis of the closest beads results of the given Shewart-calibration chart (file) and the given average reference values of the same bead in each attribute. These reference values are given in the centered columns in given menu below. At the left the average results of the shewart calibration chart beads, used for processing. These average values can be used as reference values by the arrow (>) 'button', but other values can also be used as a reference values (e.g. values from another machine). Previously stored reference calibration settings can be imported at the right.

So, if a sample has a maximum FL Red height for the calibration bead of 400 mV, and the reference has a value of 500 mV (centered yellow columns), all peak heights of the sample will be increased with 1.30 (=500/400).

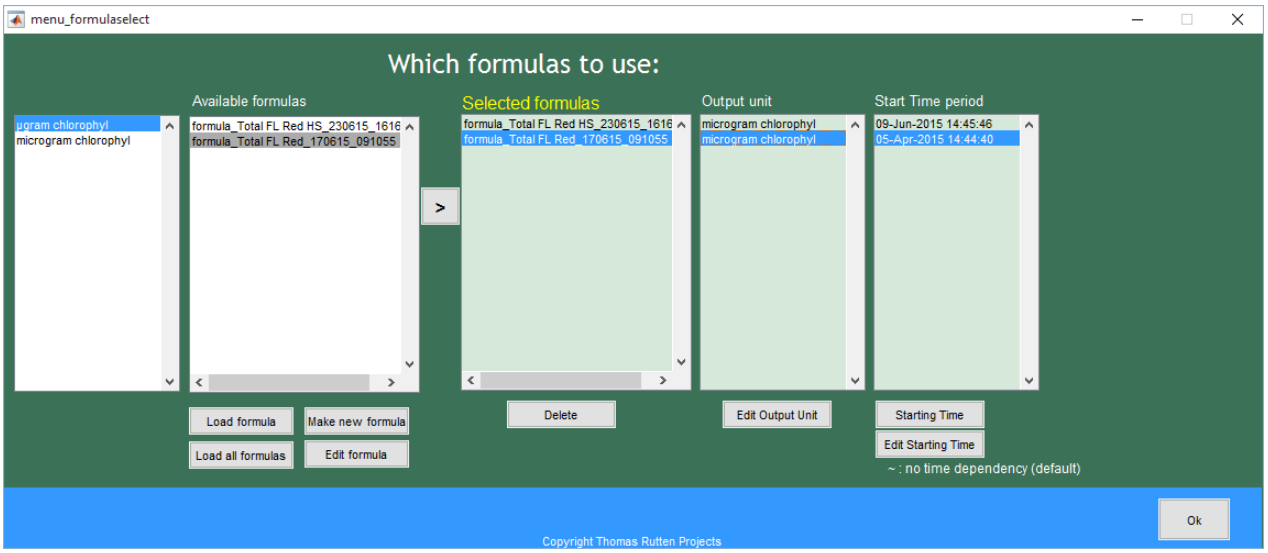
It is obvious that this 'correction procedure' will work for small changes of the instrument. Be aware that the used beads should be measured at non-saturated level, i.e. the absence of flattened peaks.



Settings use formulas to calculate new variables

This option is used to define new variables such as chlorophyll or biovolume or Length as a new column in the raw FCM data. The calculation of a new variable(s) is based on (a) given formula(s).

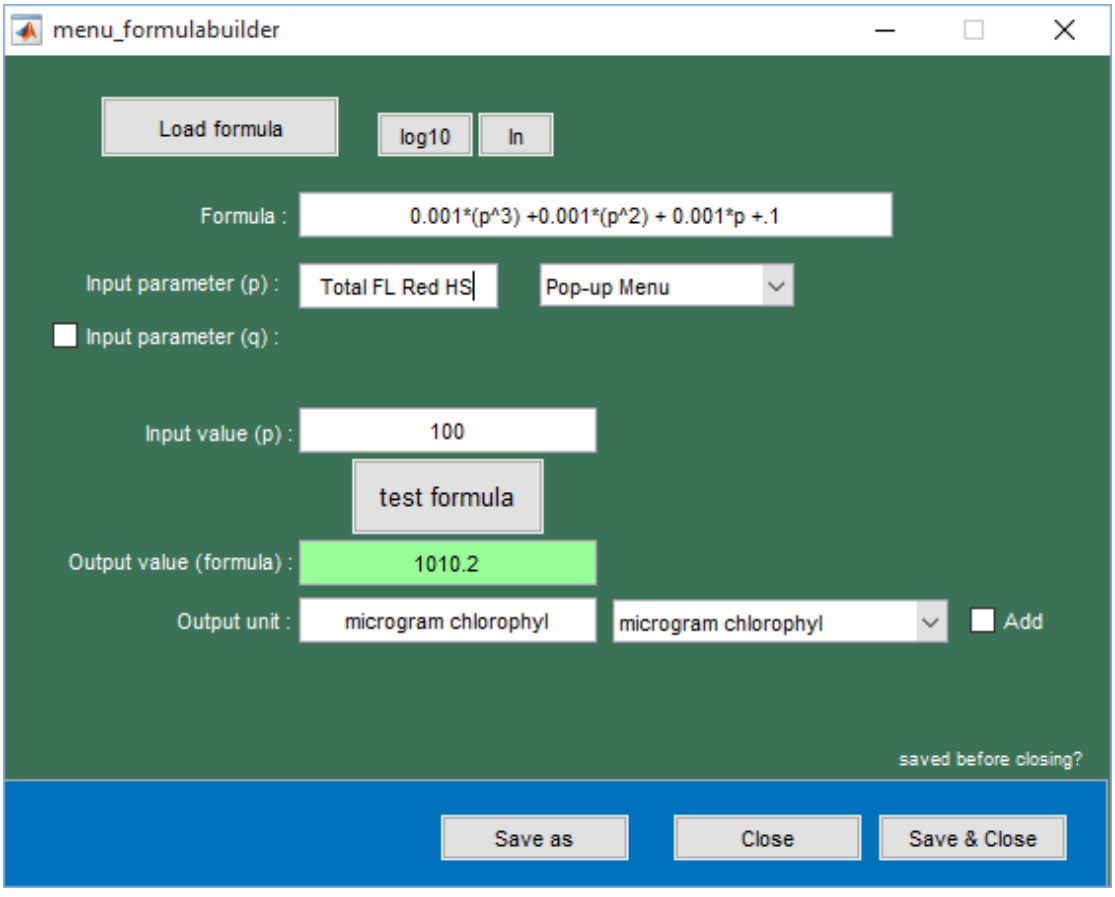
Load one or more formulas (made in the formula builder) and select the formula(s) you would like to use. The selected formula(s) appear ion the centered and right (green background) columns. The 'Start Time period' column gives the opportunity to use different formulas for the same output parameter. Use ~ sign if you want to use the same formula independent of time.



Setting a formula by the formula builder:

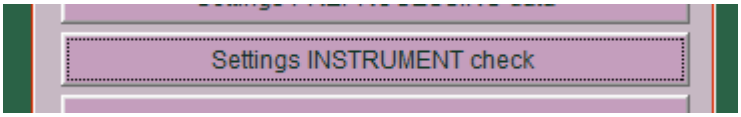
The input variable 'p' in the formula in this example is the 'Total FL Red HS' parameter in your data. It is strongly recommended to test the formula by the 'test formula' button. For this, give an input variable (here: 100) and tick the 'test formula' button. The 'Output value (formula) window gives a numerical value and has a green backgroundcolor if the formula is okay, a red background if it is not okay.

Don't forget to fill in the right 'output unit' (here: microgram chlorophyl), because this will used as a column header. The' /ml' should not be used here, because an individual particle is an absolute unit, not relative. The formula is used to recalculate each particle!

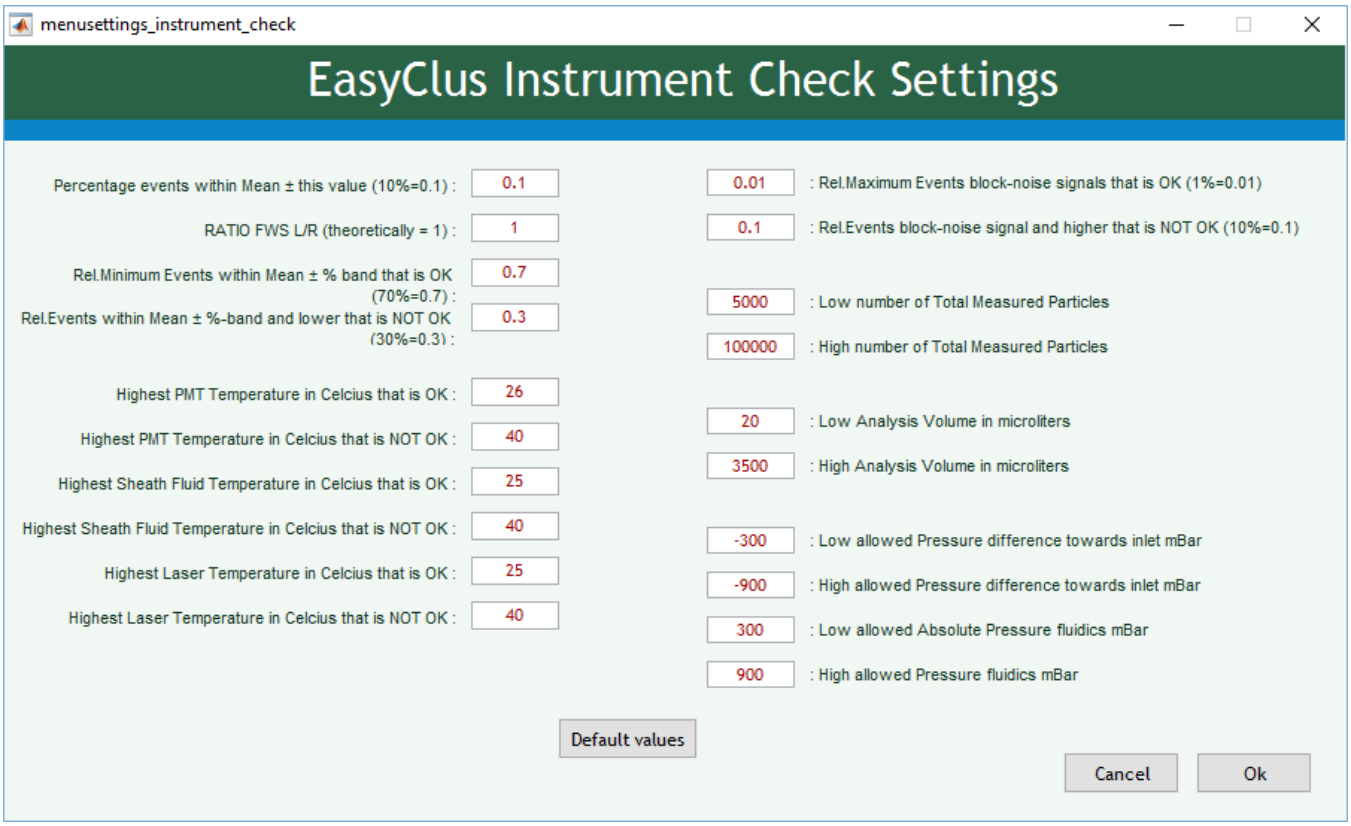


4.7.4 Settings INSTRUMENT Check

The default settings of the instrument check can be changed by the settings Instrument check button.

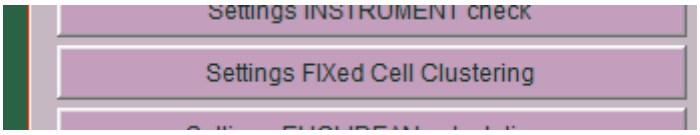


After ticking, the menu below appears and the criteroia values can be changed and saved by clicking ok.



4.7.5 Settings FIX Grid Clustering

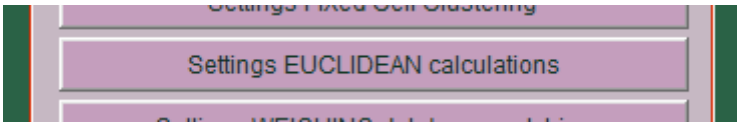
Menu to adjust the FIX Grid clusters (previously called fixed cells) cluster criteria, which will be used to define virtual FCM species in unique clusters (cells)**.**



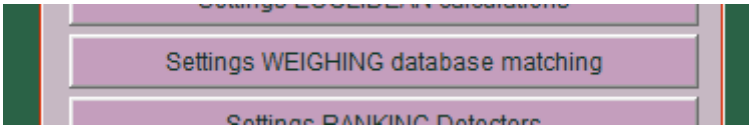
4.7.6 Settings EUCLIDEAN calculations

Euclidean distances are the calculated distances between the clusters, between particles of clusters and several characteristics are calculated here as a measure of the cluster variety of samples. An important thing to do here is the definition of the maximum expected distance (value) in FCM variables. This is done by these settings

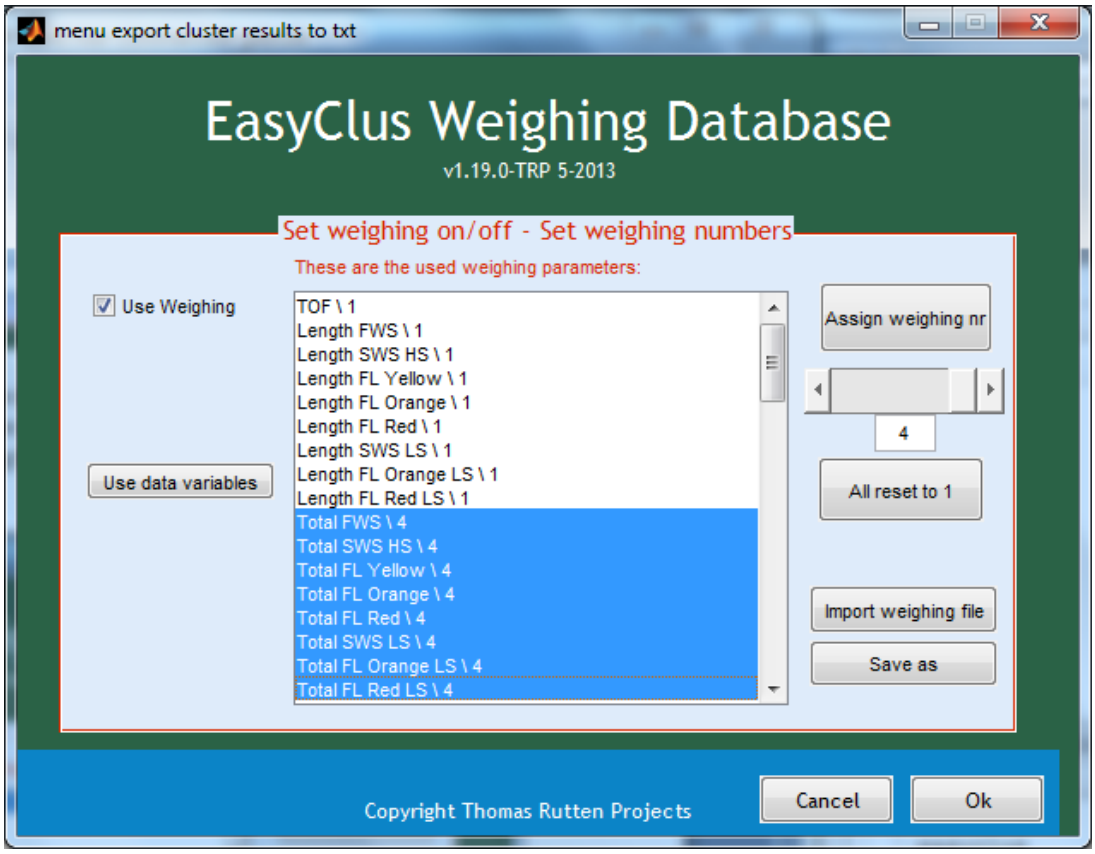
Menu to adjust the Euclidean distance calculations of found clusters after clustering.



4.7.7 Settings WEIGHING database matching



Weighing for database matching can be put on or off (0) and the degree of weighing can be adjusted here. Weighing means that variables with higher values counts heavier. In this example, the Total variables are set to a weighing value of 4.
Select Totals, go to slider for adjusting the value and confirm by pressing 'Assign weighing nr'. Press OK to save and leave.



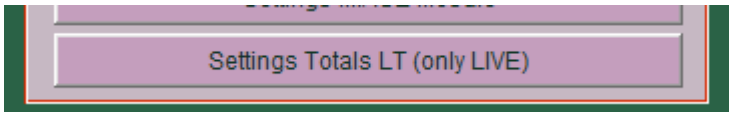
4.7.8 Settings RANKING Detectors

This menu is meant for a release in the future to give a ranking to the detectors, which are more important to your opinion than others. It is not used now.

4.7.9 Settings IMAGE Mosaic

This option is the number of images that you would like to project horizontally in a mosaic images plot.

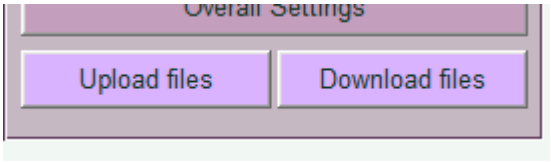
4.7.10 Settings Totals LT (only LIVE)

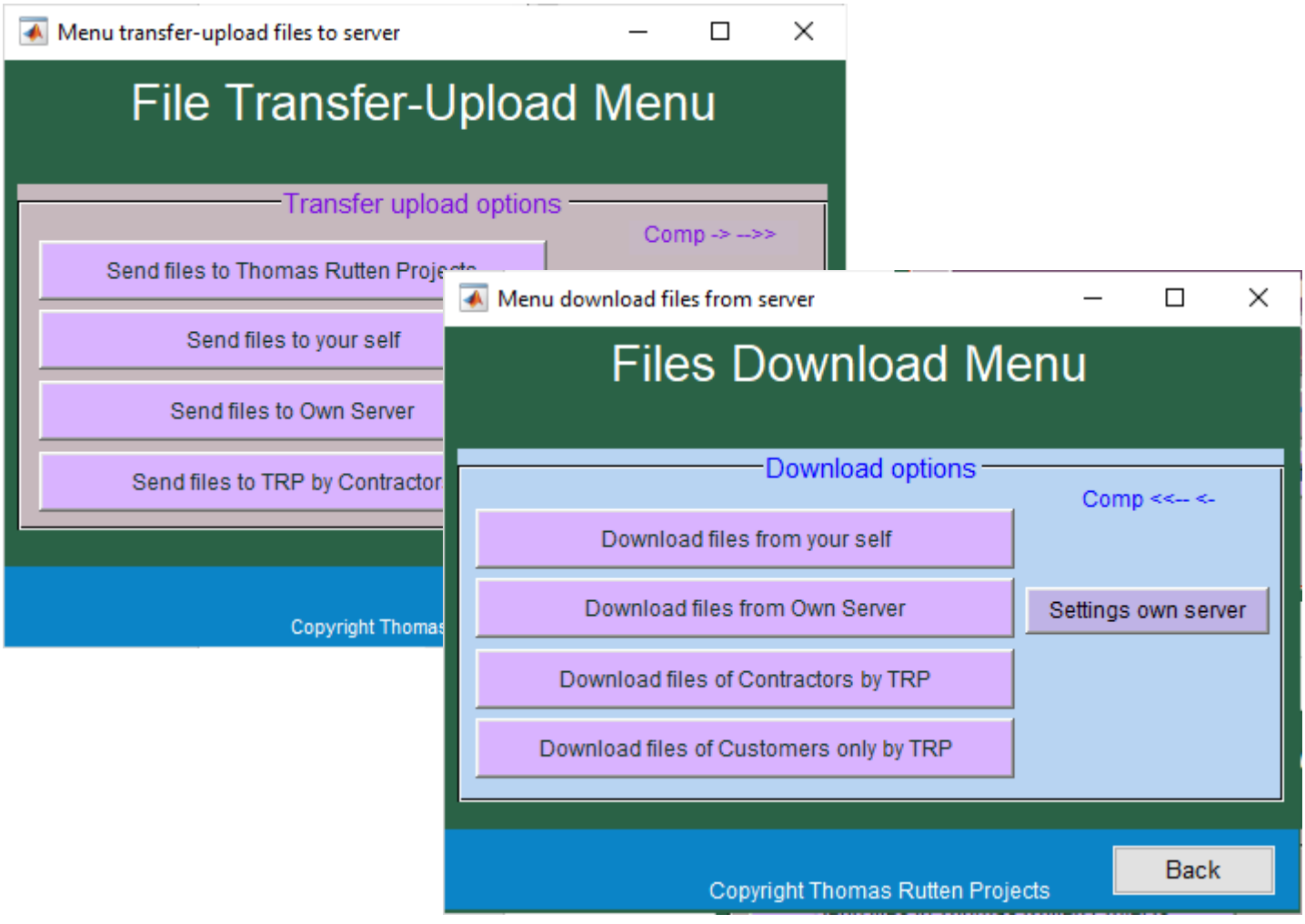


Settings to define the parameters you would like to store of the 'Totals' data. The usual 'Totals' data contains all sums of available FCM variables, which can be quite a lot and redundant. Therefore there is an option in LIVE to store only own selected sum variables, to keep the data storage smaller.

4.8 EasyClus Upload and download service

Options to upload and download cyz-files to Thomas Rutten Projects, your own server or transfer files from one computer to another (with EasyClus and the same license)

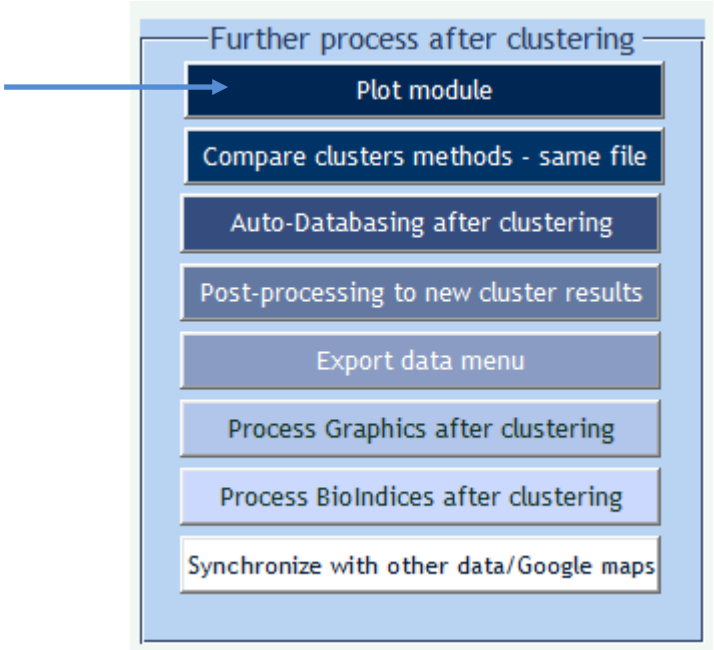




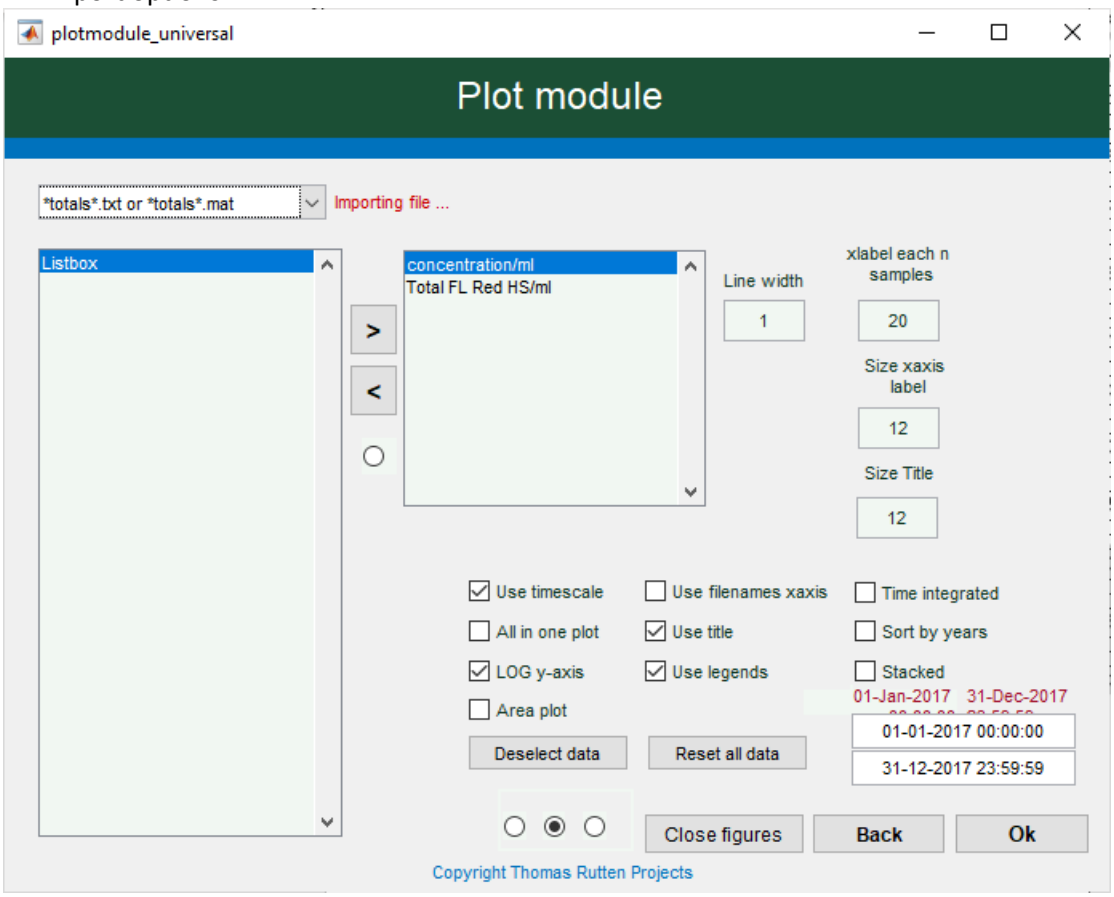
5. Blue window - post processing & graphics

5.1 Plot module

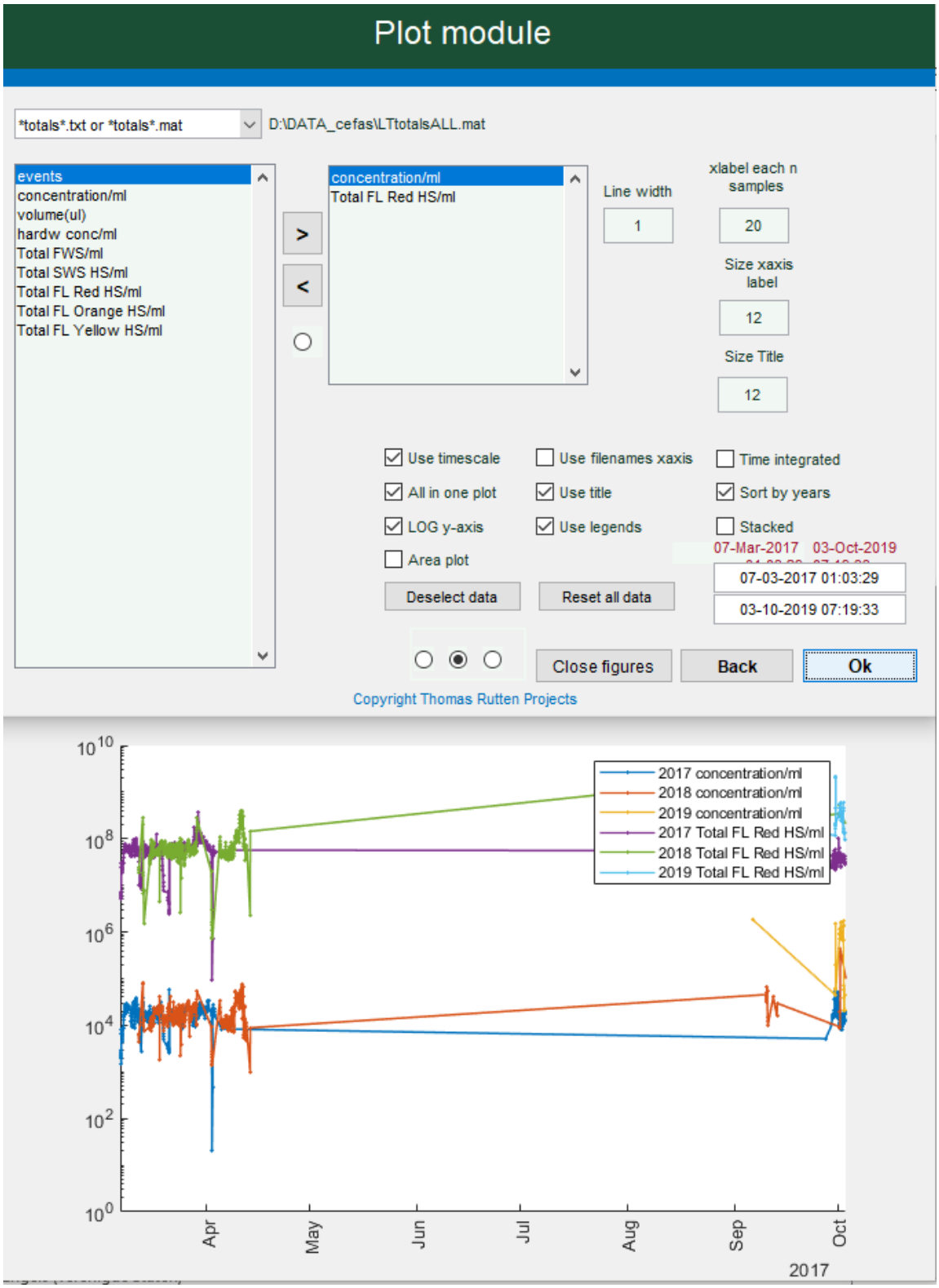
The (universal) plot module enables you to make graphs of all kinds of data processed by EasyClus.



Import options



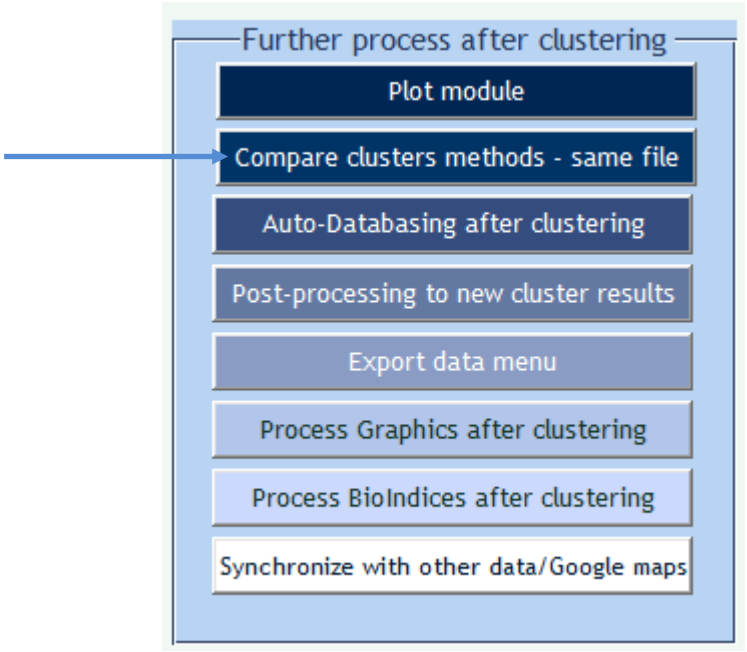
Drawing options: Here: together in one plot, time scaling used, log yaxis used



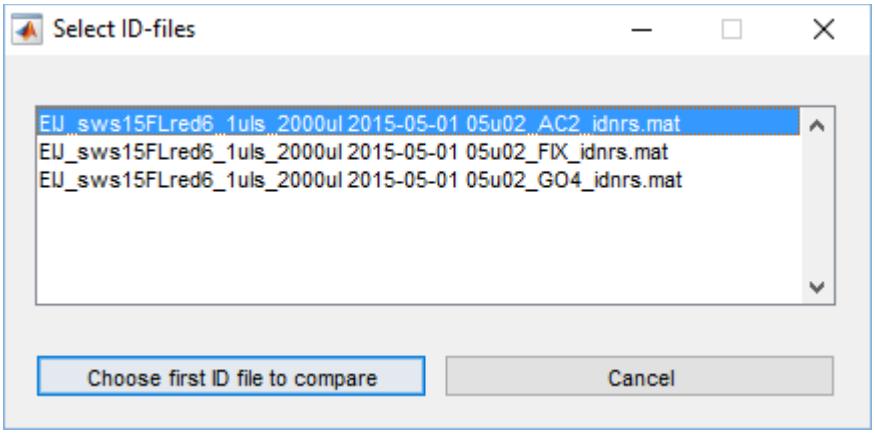
5.2 Compare cluster methods - same file

Option to mutually compare cluster results of the same file with each other.

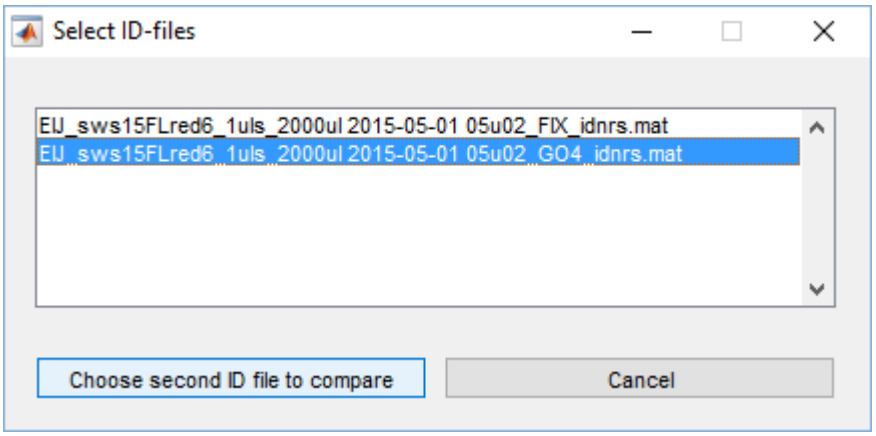
The comparison of the different clustering methods is made easy by the 'compare clusters methods' option. After clicking this button, you are asked to select a file (which is clustered by you previously). All previous used clustering results of this file are searched in the 'usual' cluster results directory. If there are two or more cluster results for this file, you are asked to select two results, which should be compared.



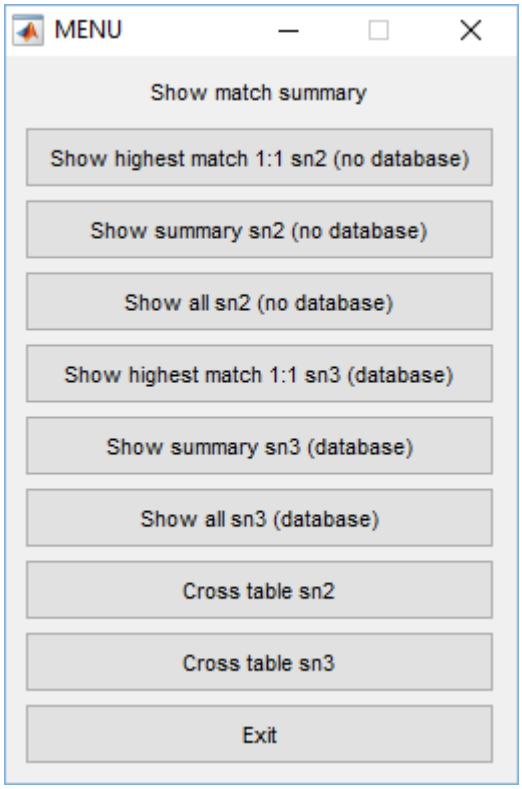
After selecting a file, the clustering results that are available are shown. Select the first result to which the second will be compared.



Select the second result, which will be compared to the first



After that, select how much information to show. 1:1 means, the highest mutual match for each cluster of the first chosen cluster result. 'Summary' gives more information and 'all' gives every event match to the other. The cross table shows the 'cross table' of the matching results which is also stored as a txt file.
Sn2 are the clustered results by each applied cluster method.
Sn3 are the clusters after database matching.



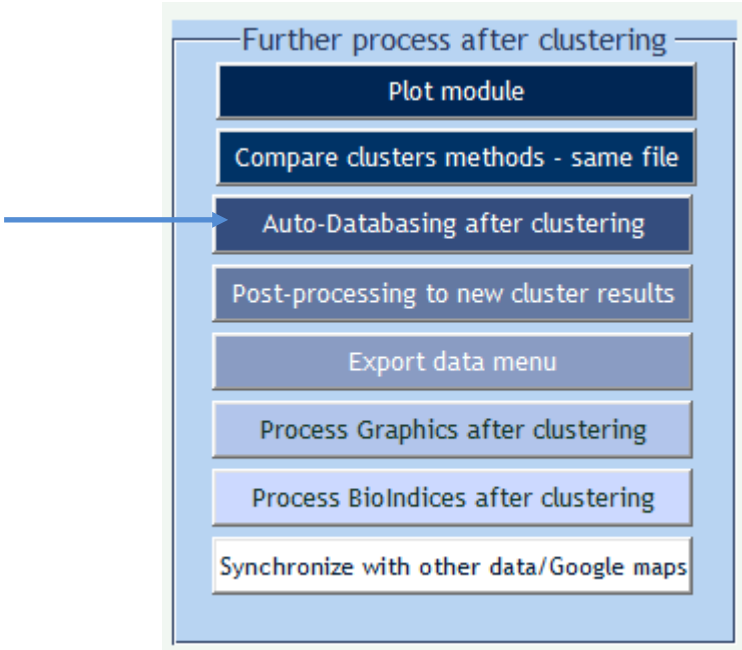
An example is given below:

	A	B	C	D	E	F	G	H
	Clusters meth GO4	Nr ev.	Clustername		Clusters n	Nr ev.	Clustername	%match
	1	384	Clusternr_01		1	329	Clusternr_01	0.857
	2	348	Clusternr_02		2	340	Clusternr_02	0.977
	3	308	Clusternr_03		3	290	Clusternr_03	0.942
	4	296	Clusternr_04		4	296	Clusternr_04	1.000
	5	257	Clusternr_05		5	197	Clusternr_05	0.767
	6	218	Clusternr_06		6	214	Clusternr_06	0.982
	7	209	Clusternr_07		7	207	Clusternr_07	0.990
	8	196	Clusternr_08		8	190	Clusternr_08	0.969
	9	194	Clusternr_09		9	152	Clusternr_09	0.784
	10	179	Clusternr_10		10	160	Clusternr_10	0.894
	13	48	Clusternr_13		11	48	Clusternr_13	1.000
	14	29	Clusternr_14		12	24	Clusternr_14	0.828
	15	21	Clusternr_15		13	20	Clusternr_15	0.952
	16	21	Clusternr_16		14	7	Clusternr_16	0.333
	19	63	not recognized		15	37	not recognized	0.587
	17	21	Clusternr_17		13	16	unrecogn	0.762
	19	11	unrecogn		17	9	unrecogn	0.818
	20	10	unrecogn		35	2	unrecogn	0.200
	23	4	unrecogn		30	1	unrecogn	0.250
	24	3	unrecogn		32	1	unrecogn	0.333
	25	3	unrecogn		45	1	unrecogn	0.333
	27	2	unrecogn		28	2	unrecogn	1.000
	28	2	unrecogn		36	1	unrecogn	0.500
	29	2	unrecogn		46	1	unrecogn	0.500
	31	2	unrecogn		29	2	unrecogn	1.000
	33	1	unrecogn		20	1	unrecogn	1.000
	35	1	unrecogn		41	1	unrecogn	1.000
	% particles that mutually match	0.90617						
	Total nr of particles	2971						
	Not matched particles	460						

All results are automatically stored in the fcm\cluster\ directory

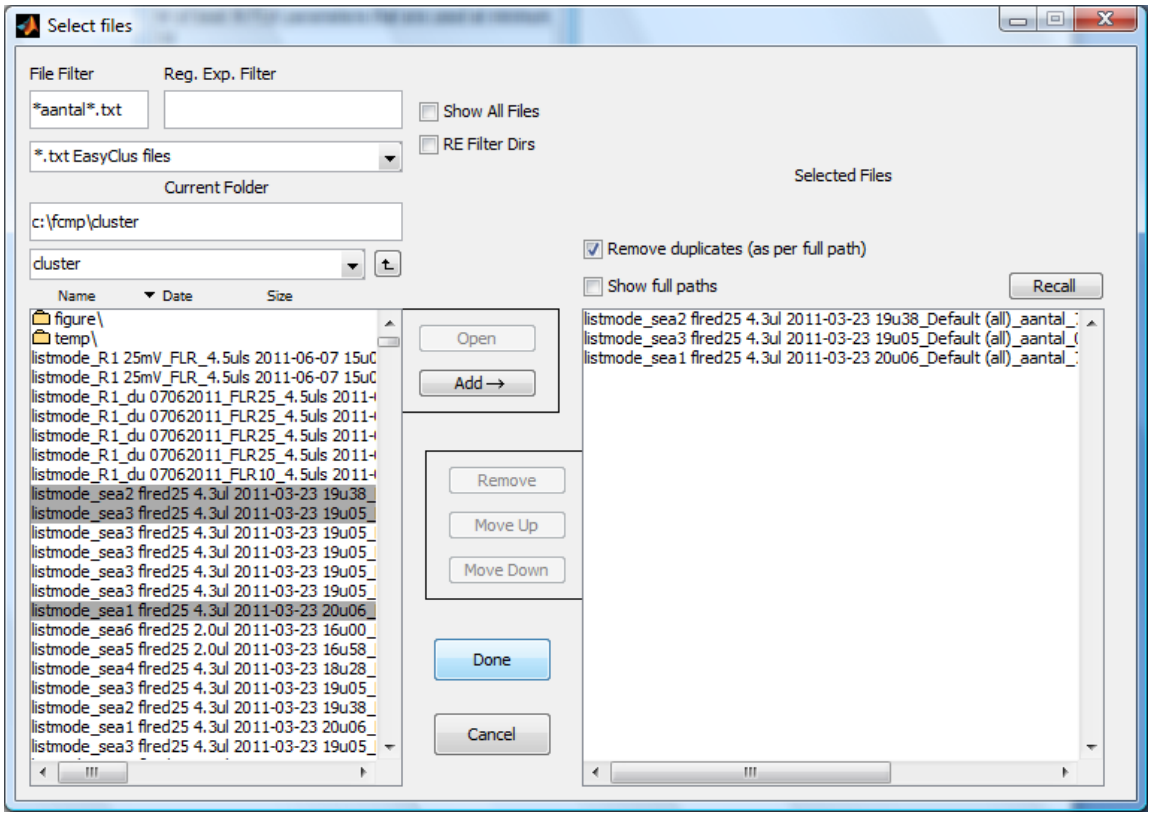
5.3 Autodatabasing with 1,2,3 ..n clustered files (.txt database is not recommended anymore)

Option to build a database automatically on basis of unique clusters found with one of the clustering methods

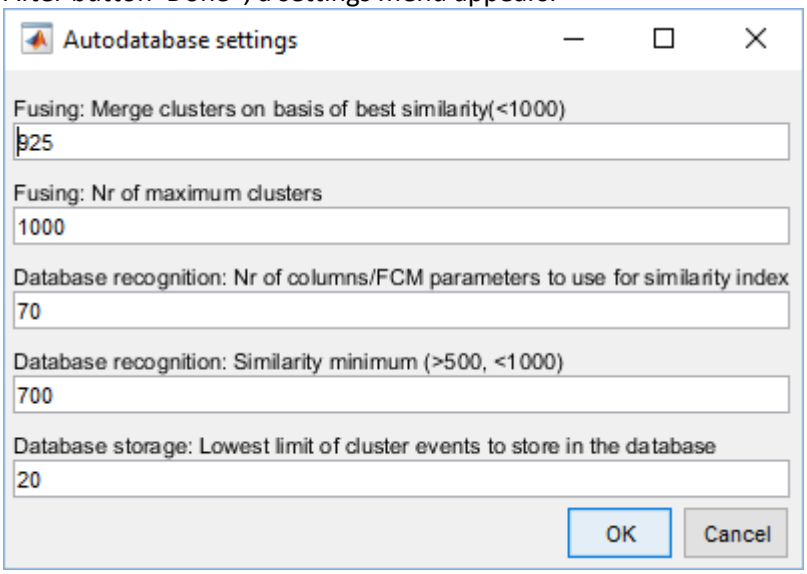


After autoclustering (method GO, DESIGN) of several FCM datafiles, it is possible to create and add unique clusters to a new database automatically.

The results after autoclustering are saved in ..\cluster*.txt and the procedure starts with the selection of these cluster result files:



After button 'Done' , a settings menu appears:

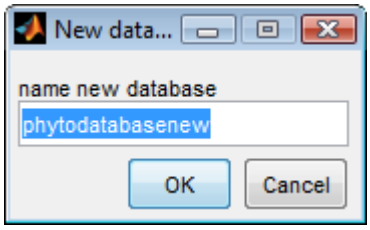


Only clusters higher than the given minimum value of events that each cluster (of events) should contain, will be stored in the database(in example =20).

Each cluster fingerprint is fused according to the 'fusing value given (in example=925)' and matched with the database fingerprints according to setting 'Nr of best fitting .., Minimum similarity ..'. There is a possibility to force the EasyClus software to a maximum number of clusters (example=1000), but this method is not recommended, because it is better to fuse on basis of similarity, rather than the maximum nr of clusters.

If there is a match with a fingerprint that is already present in the database, the cluster is not unique and will therefore not be saved in the database. If the cluster is unique, there is no match with any of the fingerprints in the database and the cluster fingerprints will be saved in the database. Increasing the 'Nr of best fit …' and/or 'Minimum similarity ..' will make the matching process more critical and therefore more unique clusters will be found.

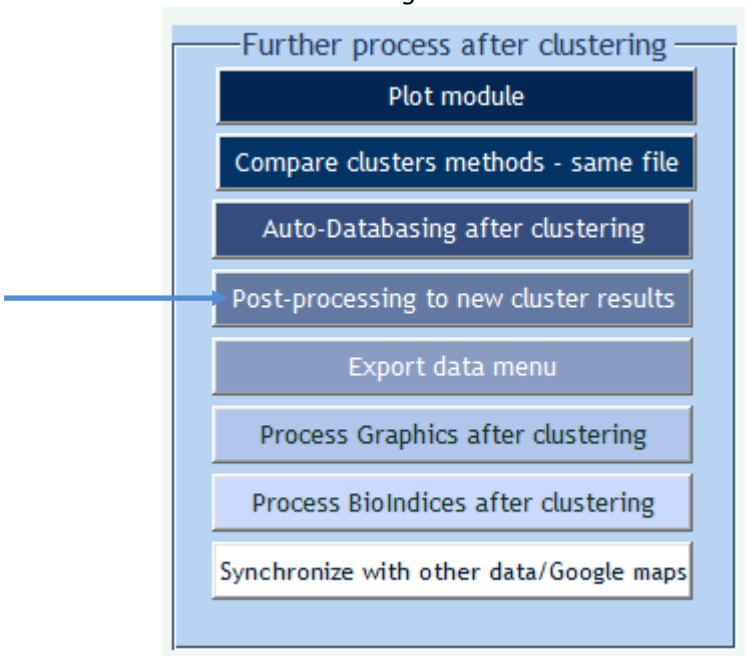
The new name of the 'auto'database , e.g. 'phytodatabasenew' should been given here.



After 'OK' the procedure will be started and in the Matlab command window, the result will be shown such as the number of unique found clusters.

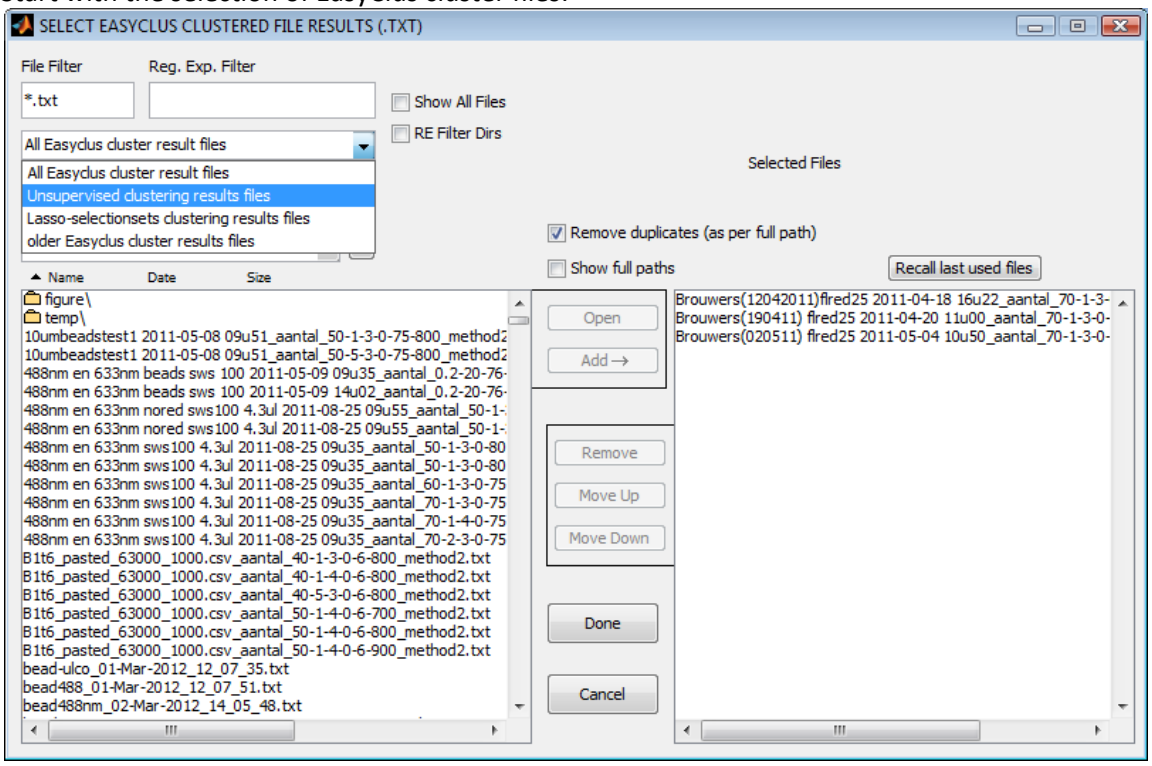
5.4 Post processing to new cluster results

Option to process new cluster results based on existing data clustering results e.g. with another database or database settings



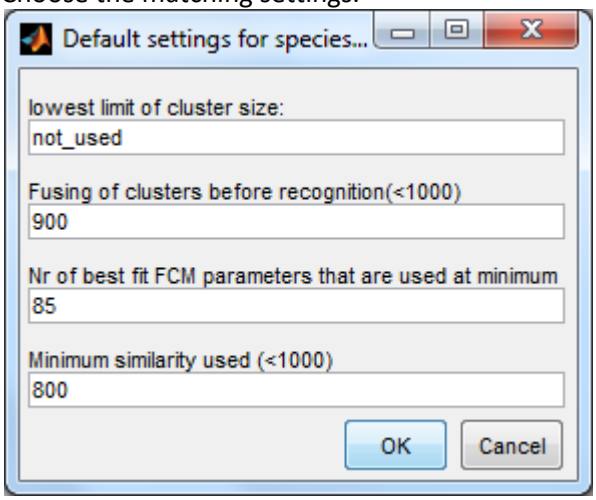
The post processing to new cluster results can only be performed on already collected clustering results (original data clusters) derived from one of the clustering methods. Post clustering is performed on basis of the original data clusters, e.g. using another database or using other database settings or using another fusing value. This method saves time if the original data-clustering is okay, but you like to play with the database matching criteria.

Start with the selection of EasyClus cluster files:

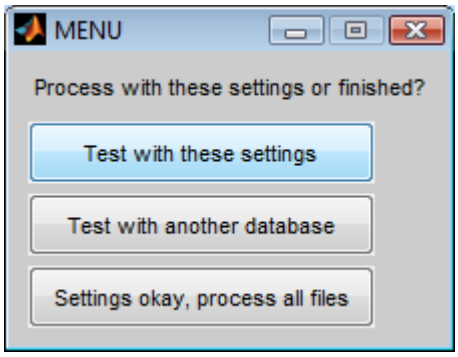


Choose database to match found clusters with.

Choose the matching settings.



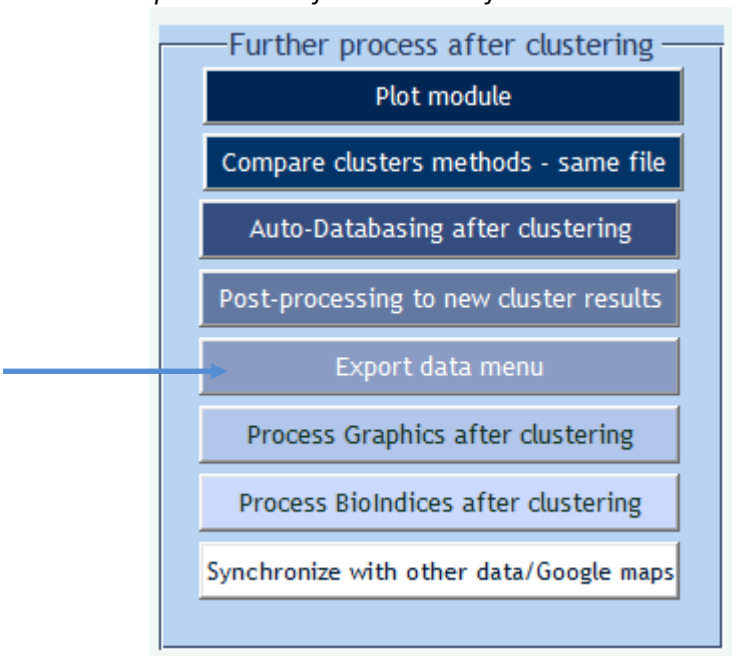
Test or change settings or process on all selected files.



These new clustering output results are saved in \cluster\oldfname_postprocessing_nr

5.5 Export data menu

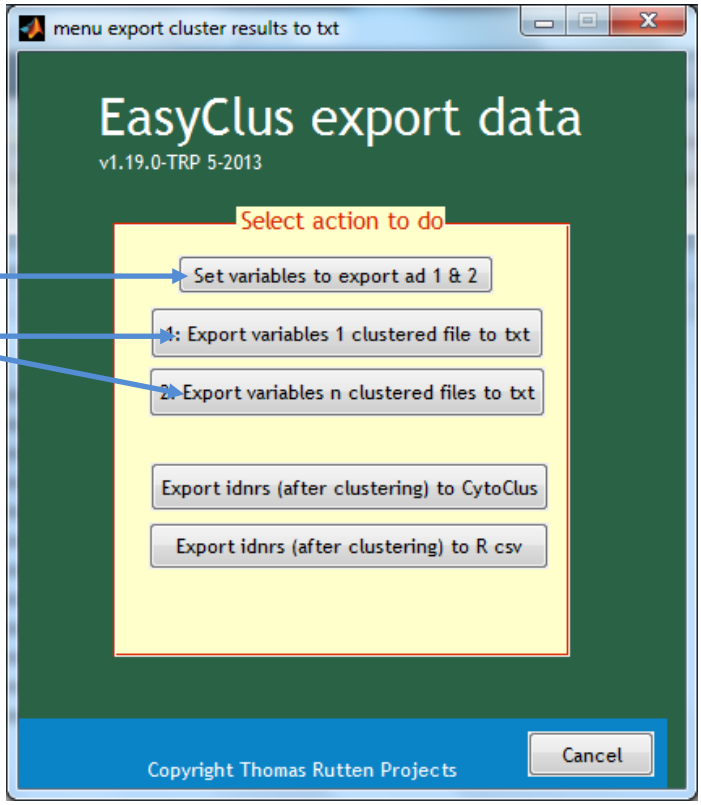
Option to make export txt-data files on basis of clustered FCM data files.



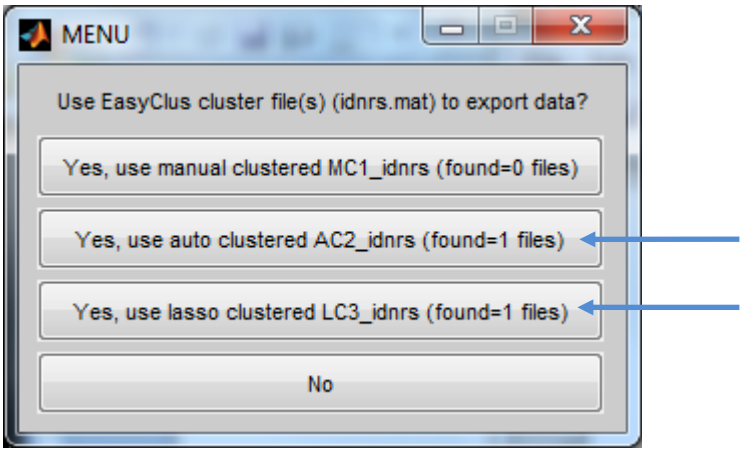
Following menu appears:

1
Set FCM variables which need to be exported

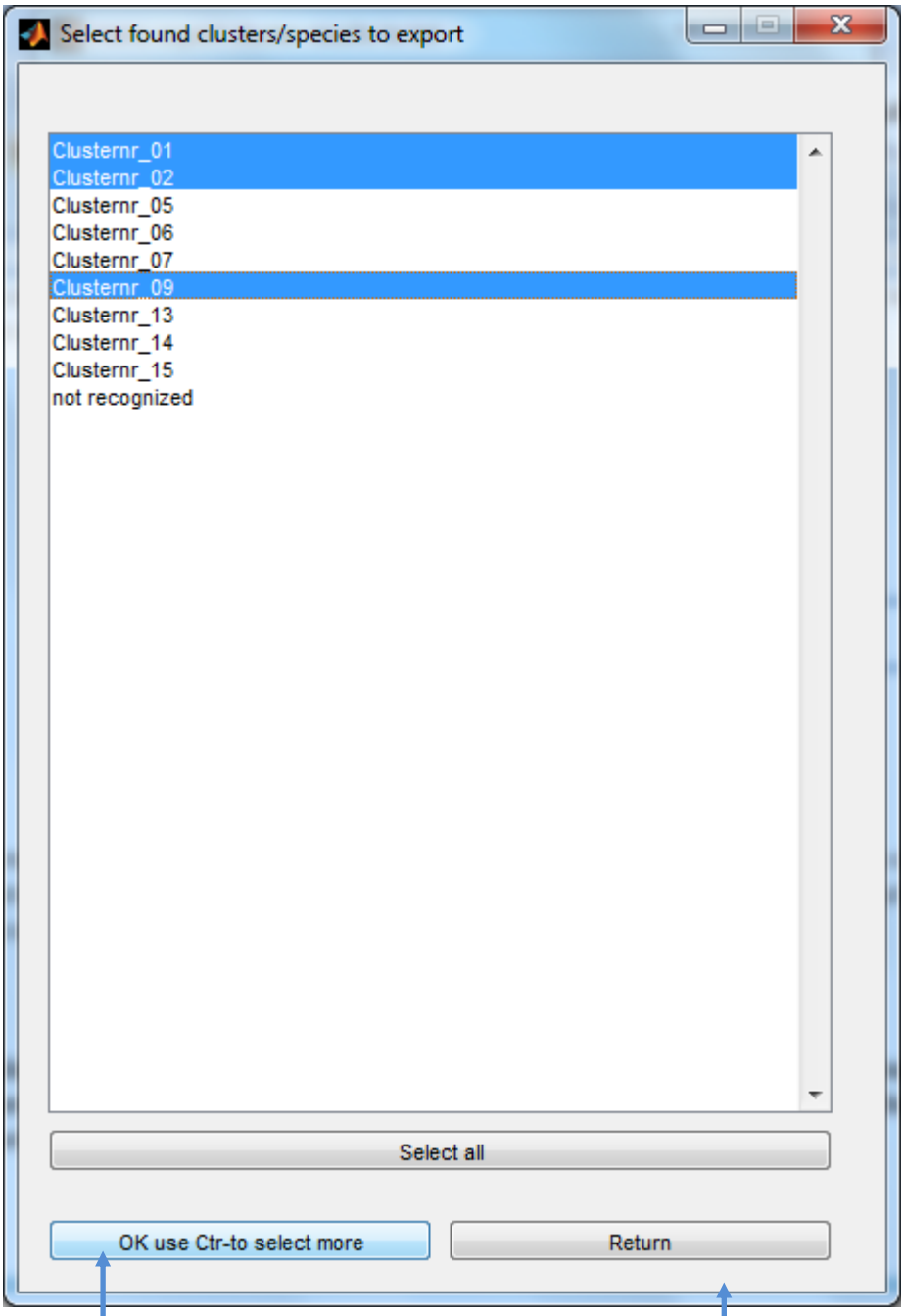
2
Select FCM-data files that are clustered and should be exported.



Ad 2: If you select one or more original FCM –datafiles, EasyClus will look for the most recent (!) available clustered files

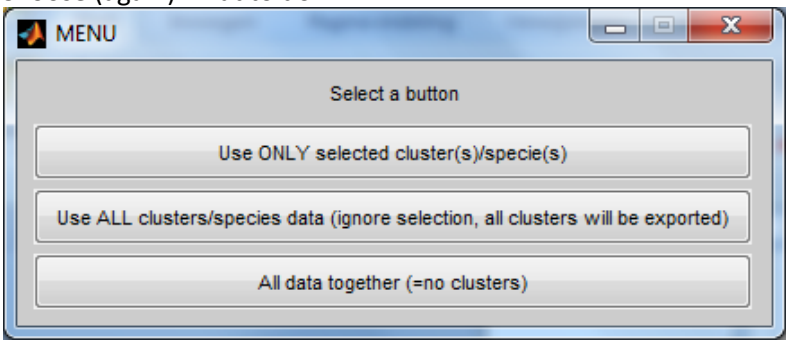


Two clustered (by EasyClus) files are available for the original FCM-data file. Select one of the files and extract the parameters (chosen under 'set variables') out of it for the clustered groups you want.



1. press OK to confirm selected clusters/species and 2. press Return to go on

Choose (again) what to do:

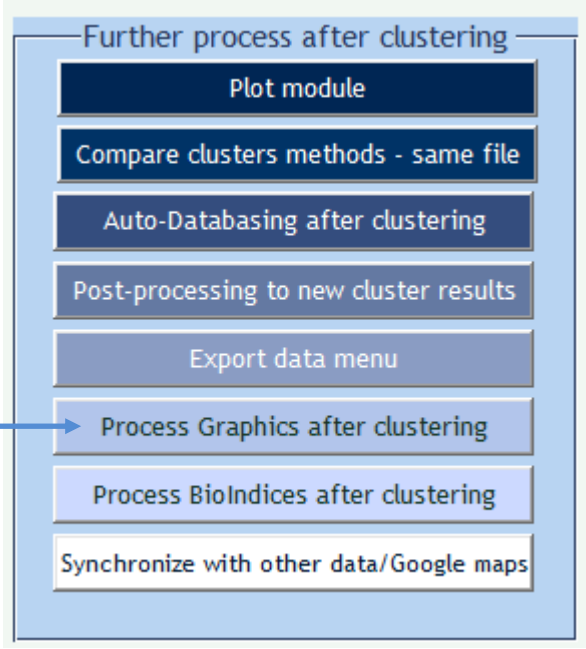


Results are saved in fcm\cluster\exportdata\ filename_EasyClus_method
and looks for instance like this:

Filenames	Volume [ul]	Used cluster m	Date	Time	Instrument	Data export by	Trigger param.	Trigger l
Maas_sws20f	3.98E+02	EasyClus_aut	17-4-2013	17:07:19	Riza CS-2013	Thomas	SWS HS	
Cluster names	Particles	Mean Length I	Mean Length	Mean Total FV	Mean Total FL	Mean Total FL	Mean Total FL Red	
Clusternr_01	2810	9.32E+00	8.12E+00	3.84E+04	2.78E+02	2.79E+04	2.47E+04	
Clusternr_02	1612	6.09E+00	6.38E+00	1.11E+04	2.43E+02	7.59E+02	3.35E+02	
Clusternr_09	2	2.44E+01	2.04E+01	1.16E+05	8.45E+02	1.58E+05	1.55E+05	
Cluster names	Particles	StD. Length F	StD. Length S	StD. Total FW	StD. Total FL Y	StD. Total FL (	StD. Total FL Red	
Clusternr_01	2810	5.74E+00	6.10E+00	2.03E+04	1.91E+02	2.20E+04	1.97E+04	
Clusternr_02	1612	4.39E+00	5.06E+00	1.68E+04	3.22E+02	1.14E+03	2.78E+02	
Clusternr_09	2	1.05E+01	6.50E+00	4.36E+02	1.99E+02	5.76E+04	3.09E+04	
Cluster names	Particles	Min. Length F	Min. Length S	Min. Total FW	Min. Total FL Y	Min. Total FL (	Min. Total FL Red	
Clusternr_01	2810	4.09E+00	4.04E+00	1.42E+03	3.86E+01	9.94E+02	1.10E+03	
Clusternr_02	1612	2.00E-01	2.24E+00	2.21E+01	1.60E+01	7.05E+01	6.04E+01	
Clusternr_09	2	1.69E+01	1.58E+01	1.15E+05	7.05E+02	1.17E+05	1.33E+05	
Cluster names	Particles	Max. Length F	Max. Length S	Max. Total FW	Max. Total FL Y	Max. Total FL (	Max. Total FL Red	
Clusternr_01	2810	9.58E+01	9.45E+01	2.34E+05	3.32E+03	1.77E+05	1.62E+05	
Clusternr_02	1612	8.49E+01	8.43E+01	2.09E+05	5.28E+03	2.13E+04	4.10E+03	
Clusternr_09	2	3.18E+01	2.50E+01	1.16E+05	9.86E+02	1.98E+05	1.77E+05	
Cluster names	Particles	Median Length	Median Length	Median Total F	Median Total F	Median Total F	Median Total FL Red	
Clusternr_01	2810	8.07E+00	6.79E+00	3.37E+04	2.38E+02	2.06E+04	1.77E+04	
Clusternr_02	1612	4.63E+00	5.00E+00	6.20E+03	1.69E+02	4.96E+02	2.38E+02	
Clusternr_09	2	2.44E+01	2.04E+01	1.16E+05	8.45E+02	1.58E+05	1.55E+05	
Cluster names	Particles	Mode Length I	Mode Length	Mode Total FV	Mode Total FL	Mode Total FL	Mode Total FL Red	
Clusternr_01	2810	4.09E+00	5.91E+00	17678	1.56E+02	1.77E+04	1.10E+04	
Clusternr_02	1612	2.00E-01	2.24E+00	2.79E+03	8.93E+01	2.03E+02	1.98E+02	
Clusternr_09	2	1.69E+01	1.58E+01	1.15E+05	7.05E+02	1.17E+05	1.33E+05	
Cluster names	Particles	Sum Length F	Sum Length S	Sum Total FW	Sum Total FL Y	Sum Total FL (	Sum Total FL Red	
Clusternr_01	2810	2.62E+04	2.28E+04	108003256	7.81E+05	78444872	69461656	
Clusternr_02	1612	9.82E+03	1.03E+04	17827296	3.92E+05	1.22E+06	5.40E+05	

5.6 Calculate cluster-species concentrations/ absolute cluster-species numbers/ visualize results of several files and put these in one file – button –Process sample(s) information after clustering

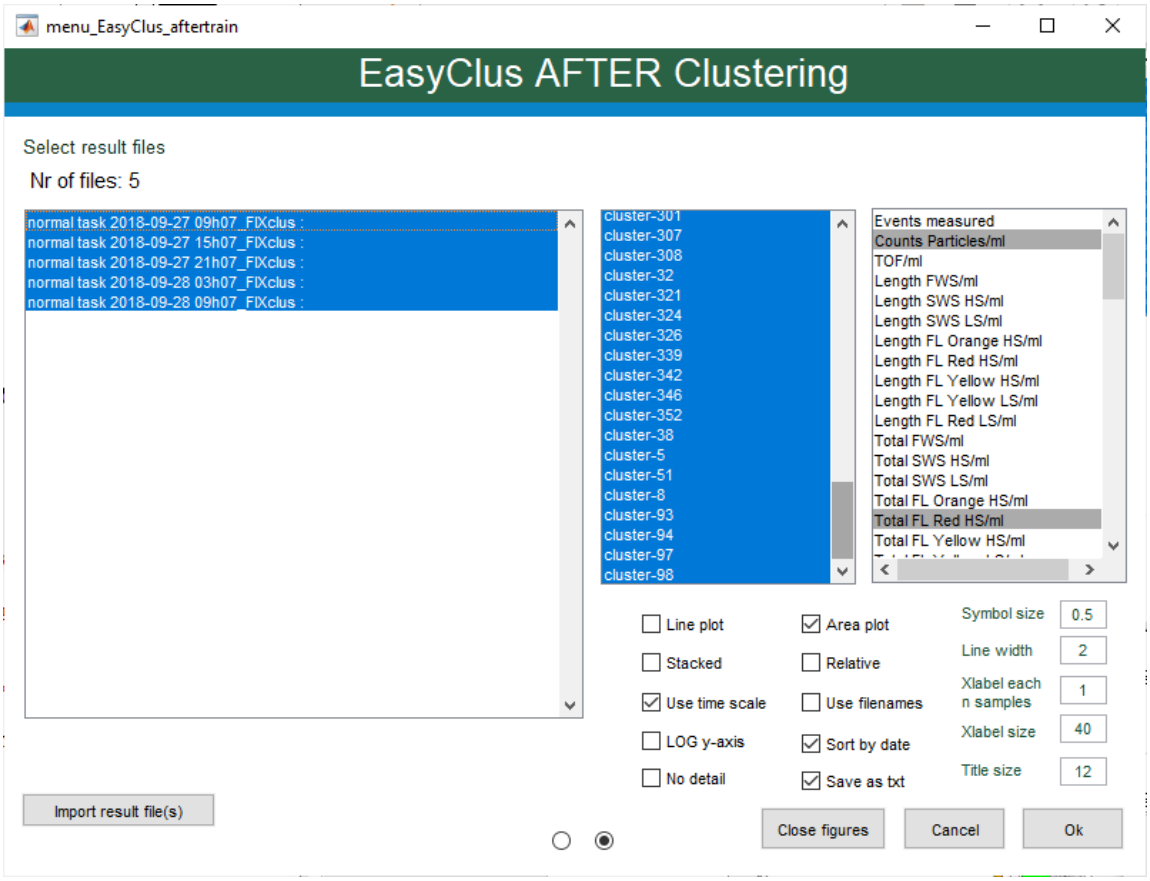
To determine concentrations or absolute numbers (if concentrations are not provided) of species from a database after clustering or produce nice figures and graphics of the results.



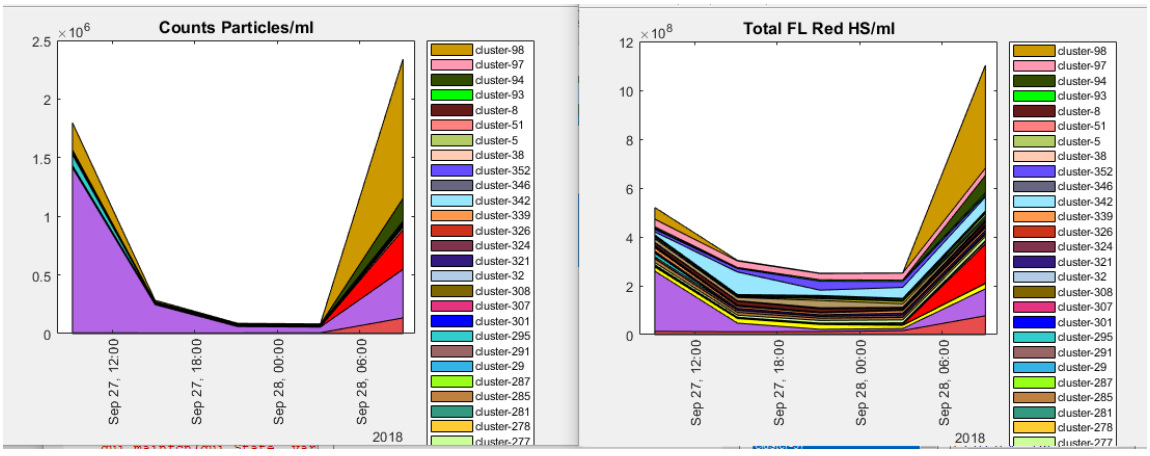
The 'Calculate species concentration' button gives the opportunity to calculate (species) concentrations out of cluster results.txt files, that are produced after EasyClus clustering. Species concentrations are put in one file together including its filename. If total sample concentrations are not provided, one has the possibility to extract the absolute numbers of species out of the cluster result files.
Needed for this button:

- clustered result(s) *.txt files saved after EasyClus clustering methods (manual, unsupervised with database, semi supervised lasso.

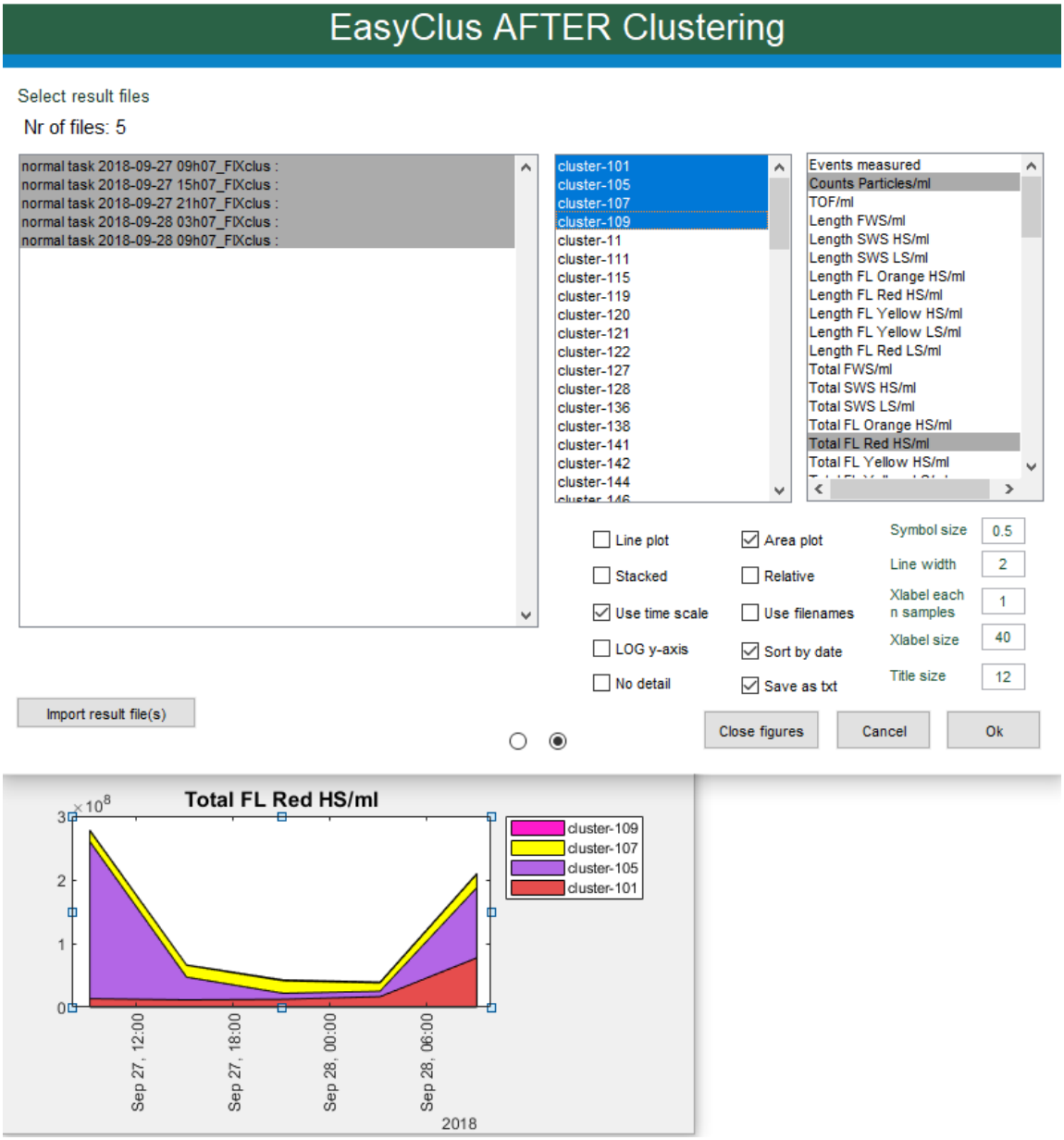
Start with the selection of EasyClus cluster files:



All unique species which have been found in the samples are collected.
Choose the FCM variables that you would like to visualize in the right listbox.
After that choose clusters or species (middle box) to be visualized.
Choose settings such as sorted by time, area plot etc.
Press OK to plot.



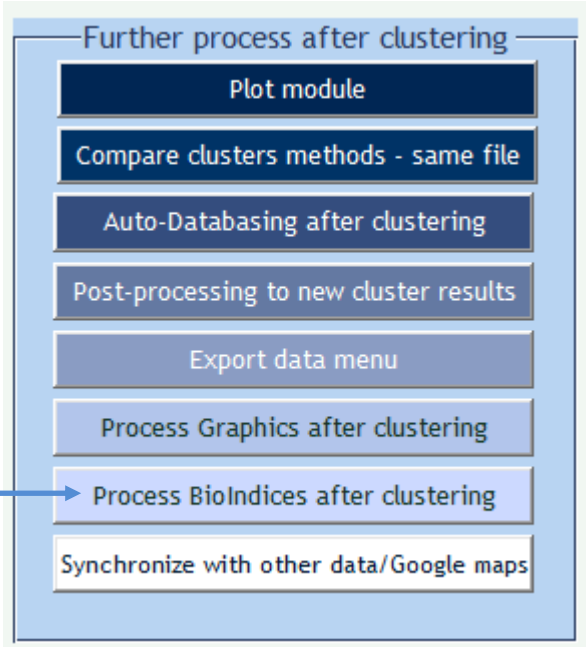
A subselection of clusters-species is done by choosing only those of interest for you



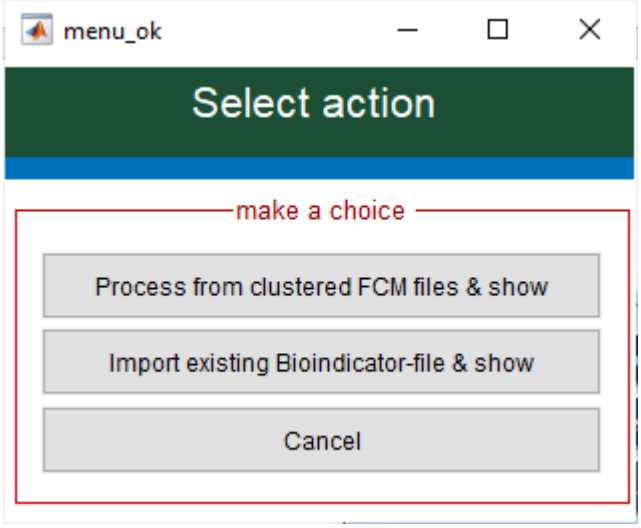
IMPORTANT to save as txt in a file!
Press checkbox ‘save as txt’ to ‘on’ to store chosen results as a file i.e. sorted by clusters and chosen variables. Results are saved in/as :\Easyclusvxxxresults\cluster\results_clusters_yyyymmdd_HHMMSS.txt

5.7 Process sample(s) BioIndices after clustering

To make figures of 'FCM' bioversity in samples on basis of EasyClus clustering results. If there are a few dominant clusters/species the biodiversity will be low, if there is a rather equally distribution of a lot of clusters, the biodiversity will be high.

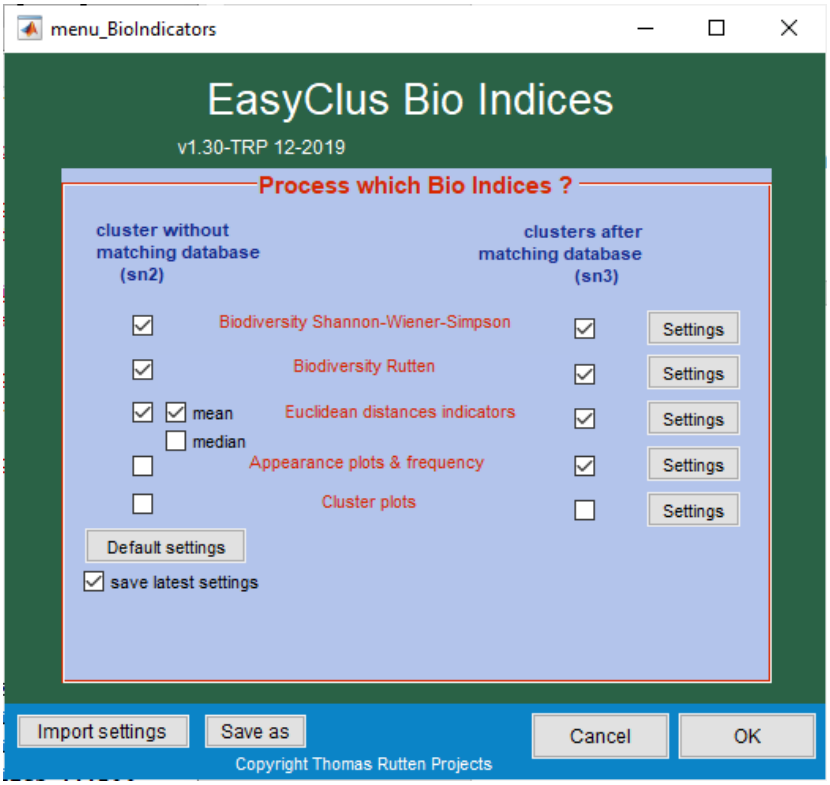


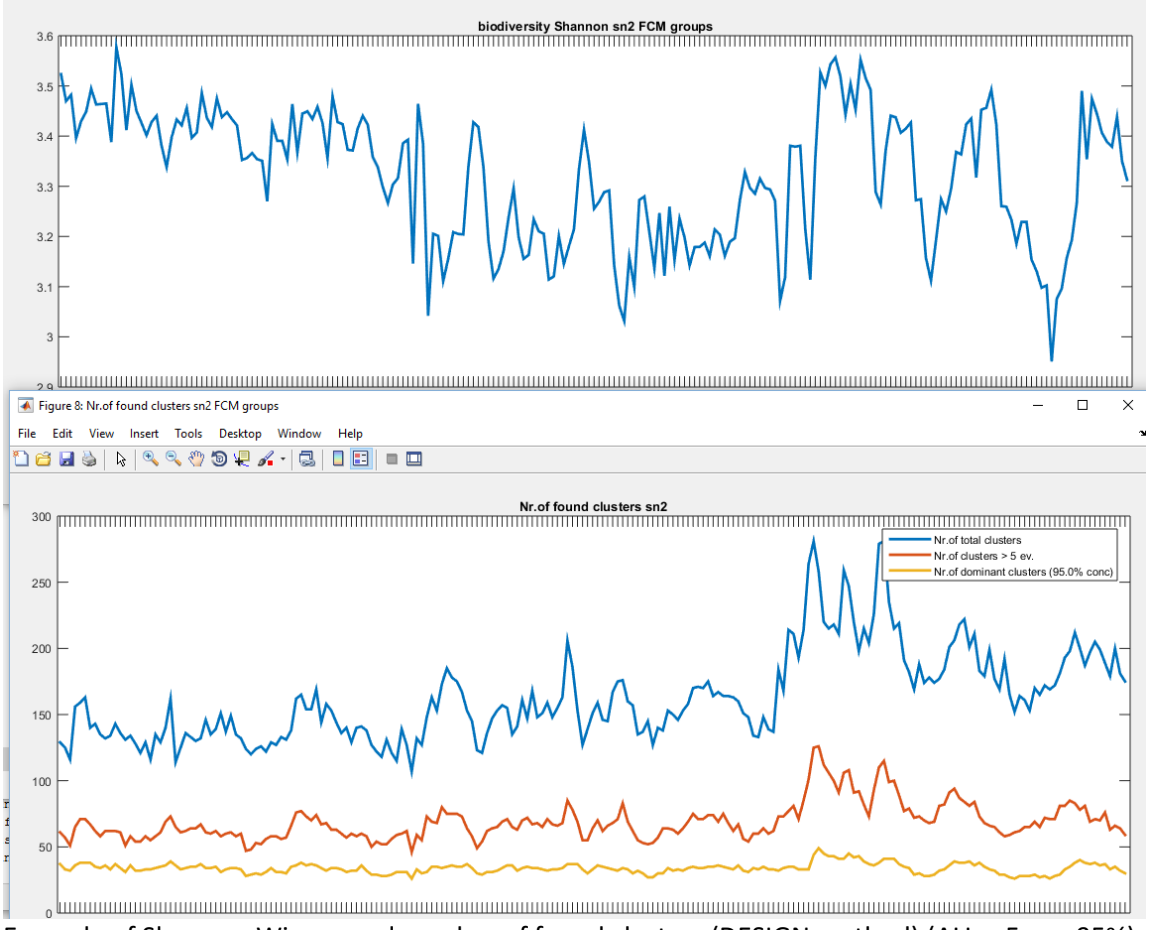
After selecting the Process Bioindicator button, you have to choose whether you want to process bioindicators on basis of clustered files results OR import an already previously generated bioindicator –file.



The first option ('Process from ..') is used for calculating bioindicator indices (such as Shannon-Wiener) out of previously clustered results. After a selection of cluster results files, a menu appears which ask you to calculate specific bioinidicators or not. When the processing is confirmed, you are asked to add (by ADD) these results to another (saved) bioindicatorfile or only store these processed bioindicators (NEW).

The plotmenu of bioindicators appears after processing, Here you can choose what to plot.

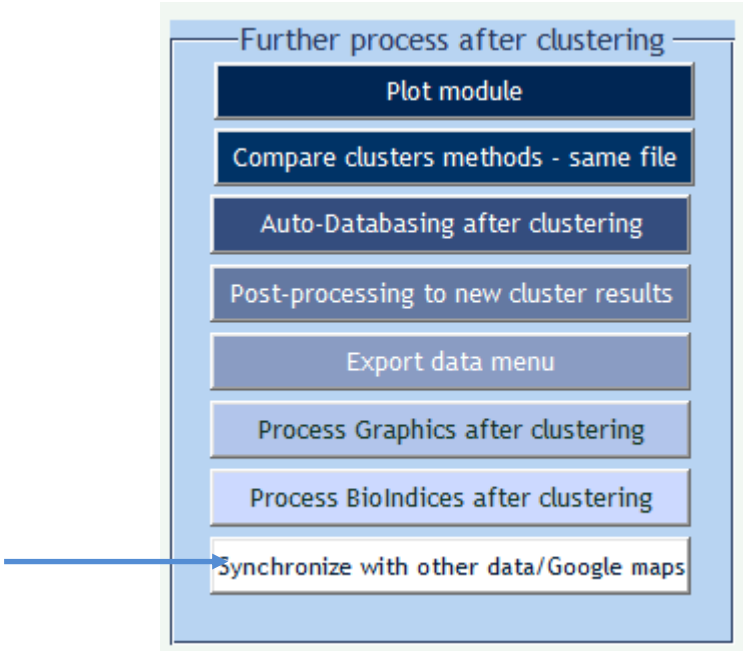




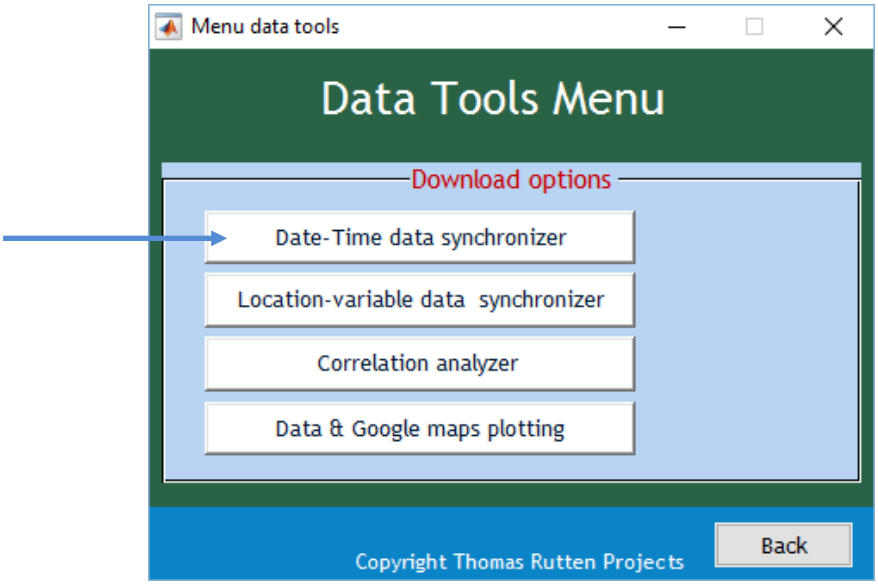
Example of Shannon-Wiener and number of found clusters (DESIGN method) (ALL, >5 ev., 95%)

5.8 Synchronize with other data/Google maps

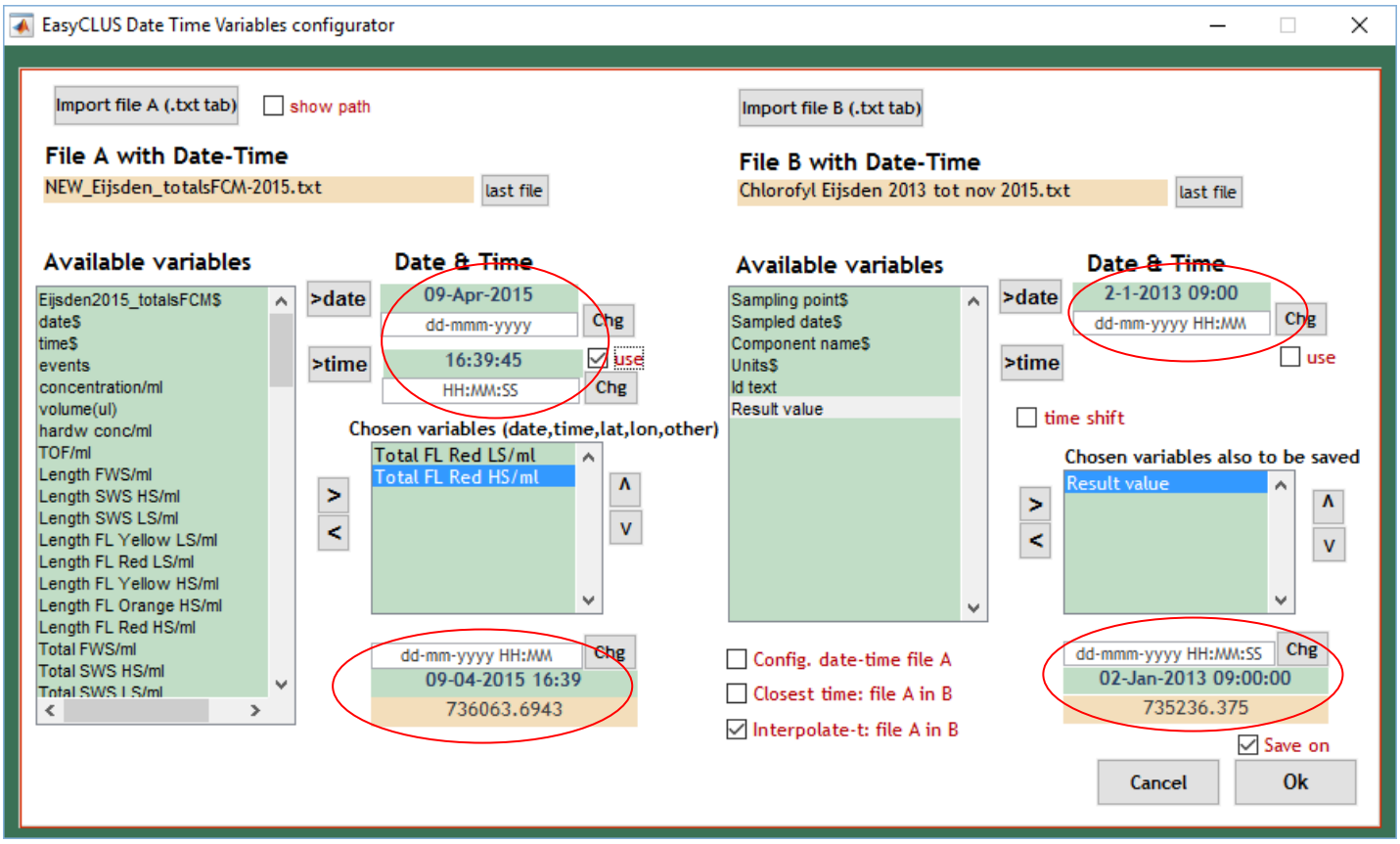
Data tool box to synchronize data with other (external) data on basis of date and time and possibility to project transect data in Google maps.



First button starts a module to synchronize two different data with date and/or time files.

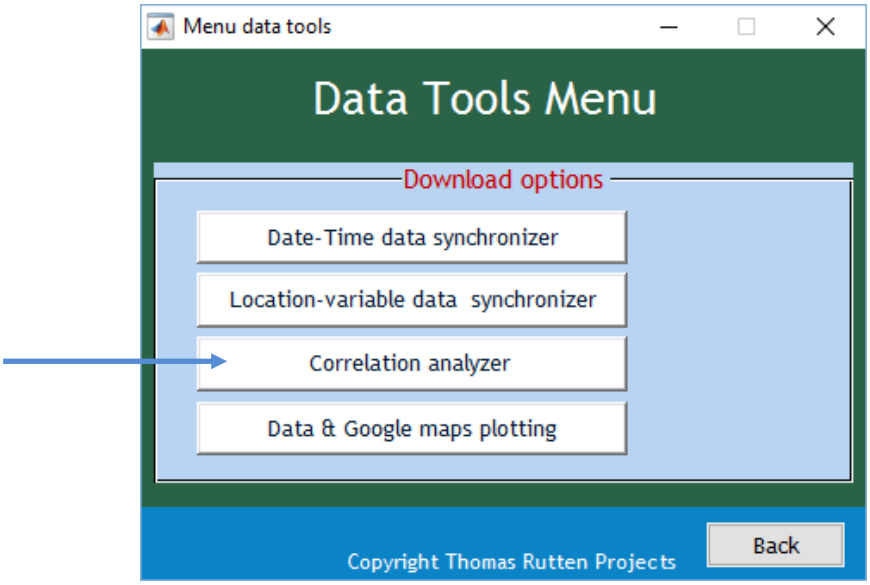


The date and time definition of the file at the left side should be defined in the left part. If it is correct, the right date time and numerical value is represented below.
At the right it is the same. Define the right date and time configuration how it is used in the second file and it is confirmed in the part below.
The variables which should be synchronized and merged can be selected by the arrows.
If one of the files contains GPS latitudes-longitudes; they are also recalculated by using the date time synchronization.

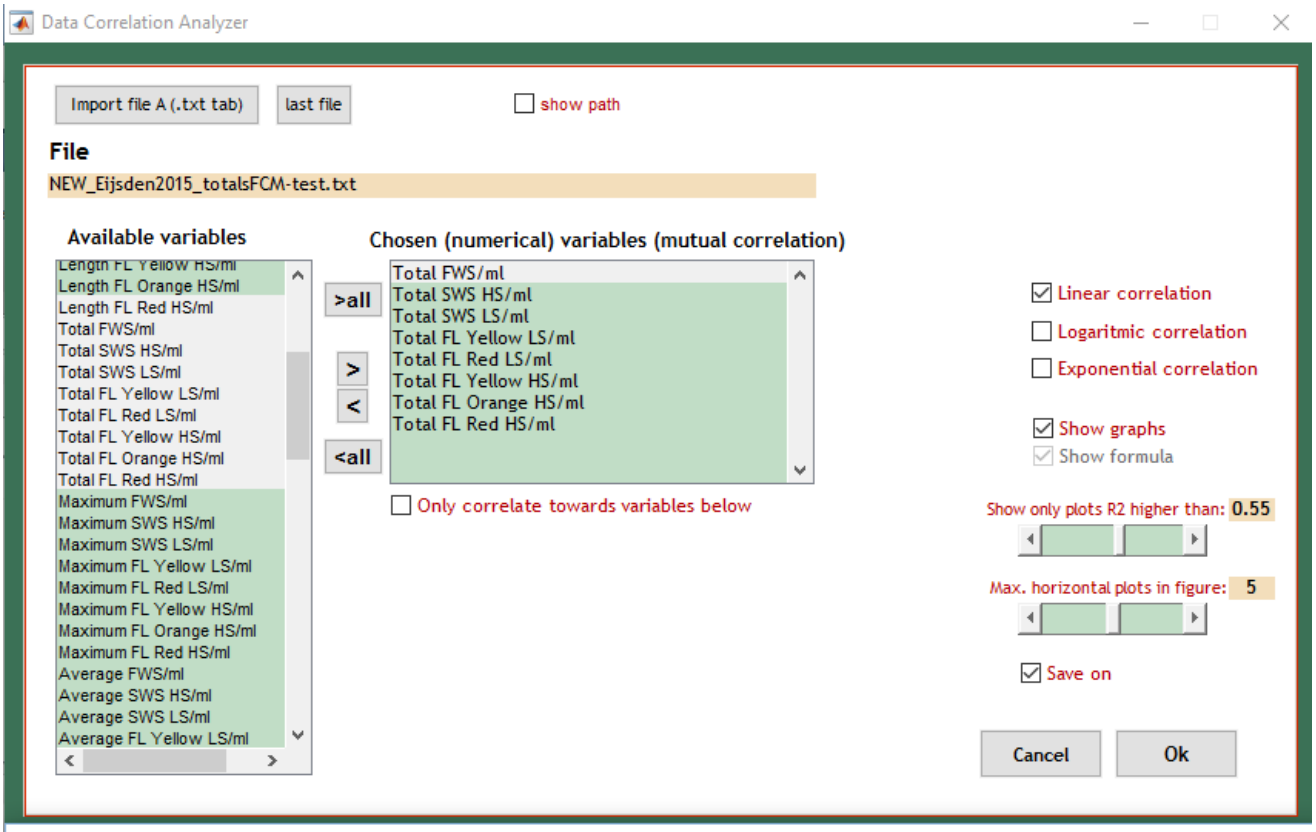


The second button is not in use yet.

The correlation analyzer does what is says. Find out if you have any correlation in your data.



Select a file, choose the variables you would like to evaluate and press ok. According to the given criteria, graphs with high to low correlation are visualized.


Data Correlation Analyzer
Import file A (.txt tab)
last file
show path
File
NEW_Eijsden2015_totalsFCM-test.txt
Available variables
Length FL Orange HS/ml
Length FL Red HS/ml
Total FWS/ml
Total SWS HS/ml
Total SWS LS/ml
Total FL Yellow LS/ml
Total FL Red LS/ml
Total FL Yellow HS/ml
Total FL Orange HS/ml
Total FL Red HS/ml
Maximum FWS/ml
Maximum SWS HS/ml
Maximum SWS LS/ml
Maximum FL Yellow LS/ml
Maximum FL Red LS/ml
Maximum FL Yellow HS/ml
Maximum FL Orange HS/ml
Maximum FL Red HS/ml
Average FWS/ml
Average SWS HS/ml
Average SWS LS/ml
Average FL Yellow LS/ml
>all
>
<
<all
Chosen (numerical) variables (mutual correlation)
Total FWS/ml
Total SWS HS/ml
Total SWS LS/ml
Total FL Yellow LS/ml
Total FL Red LS/ml
Total FL Yellow HS/ml
Total FL Orange HS/ml
Total FL Red HS/ml
Only correlate towards variables below
Linear correlation
Logaritmic correlation
Exponential correlation
Show graphs
Show formula
Show only plots R2 higher than: 0.55
Max. horizontal plots in figure: 5
Save on
Cancel
Ok

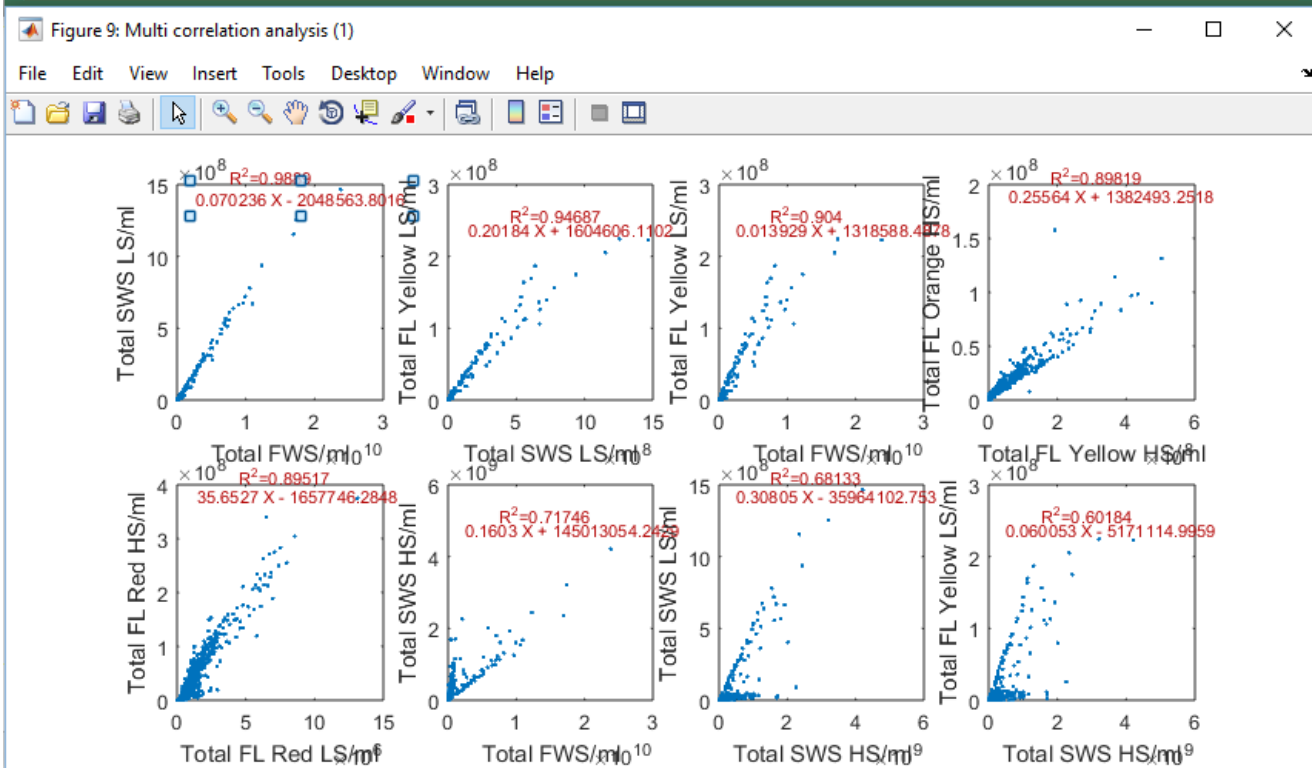
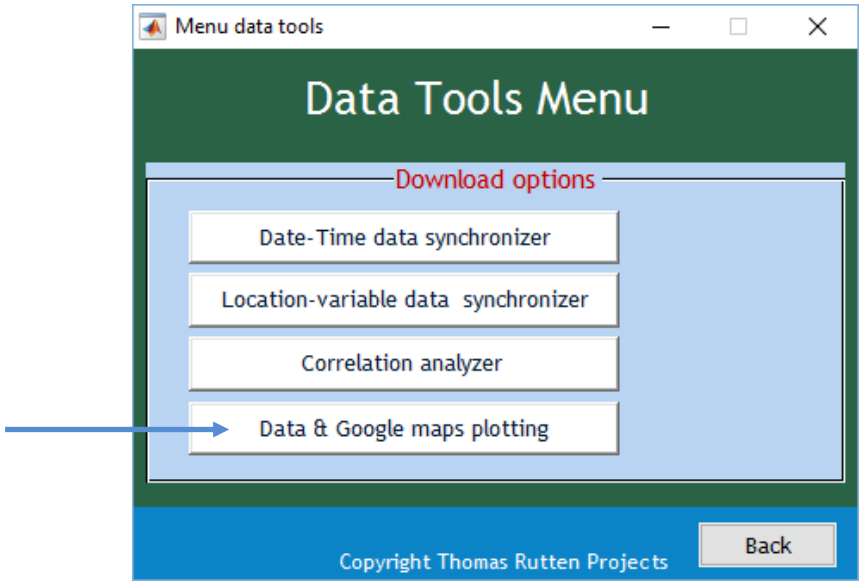


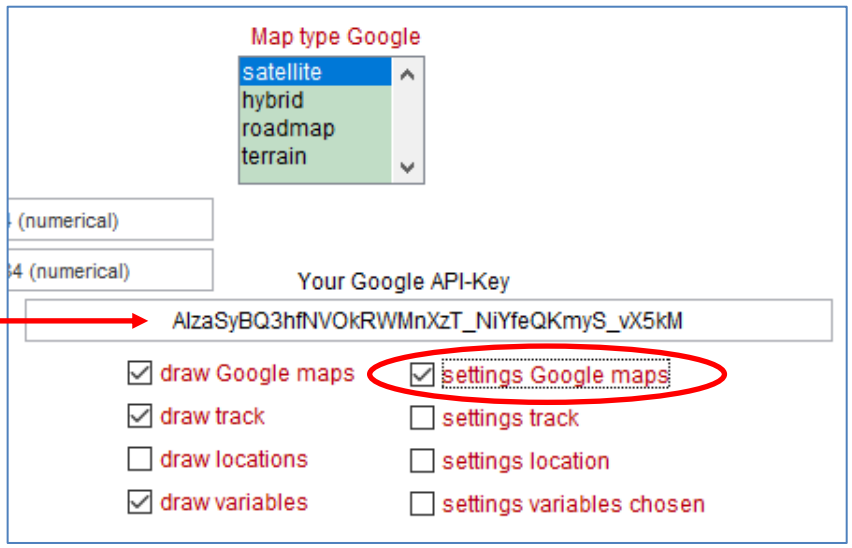
Figure 9: Multi correlation analysis (1)
File Edit View Insert Tools Desktop Window Help
Total SWS LS/ml
Total FWS/ml
R²=0.9809
0.070236 X - 2048563.8016
Total FL Yellow LS/ml
Total SWS LS/ml
R²=0.94687
0.20184 X + 1604606.1102
Total FL Yellow LS/ml
Total FWS/ml
R²=0.904
0.013929 X + 1318588.4478
Total FL Orange HS/ml
Total FL Yellow HS/ml
R²=0.89819
0.25564 X + 1382493.2518
Total FL Red HS/ml
Total FL Red LS/ml
R²=0.89517
35.6527 X - 1657746.2848
Total SWS HS/ml
Total FWS/ml
R²=0.71746
0.1603 X + 145013054.2459
Total SWS LS/ml
Total SWS HS/ml
R²=0.68133
0.30805 X - 35964102.753
Total FL Yellow LS/ml
Total SWS HS/ml
R²=0.60184
0.060053 X - 5171114.9959

Plotting data as an overlay over Google maps.
Policy change by Google from mid 2018: Google changed the policy of downloading maps for free. You need to make a personal API-key, which have to be built in the EasyClus software and which gives you (still) the opportunity to download many maps for free. However, the API-key activation needs an Google invoice billing system activation usually by a using credit card, which might give problems if you don't want that. The maps are still for free for the first thousands (>20000?) of maps. It works for the author now. To get an API Key by Google see here.

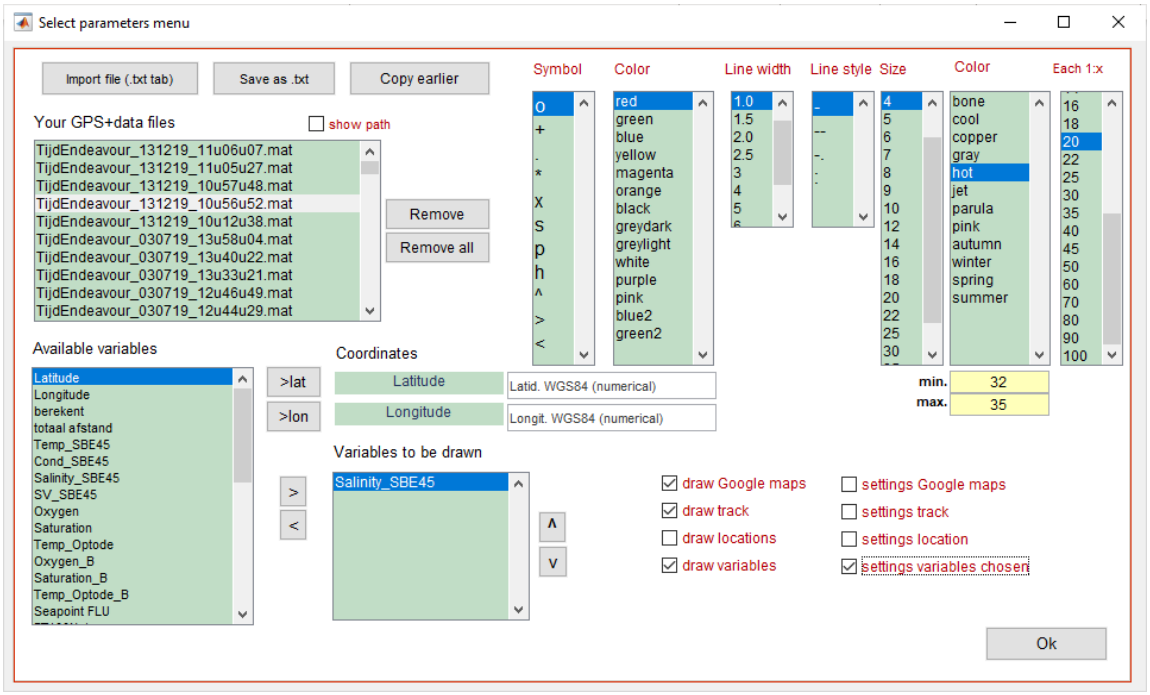


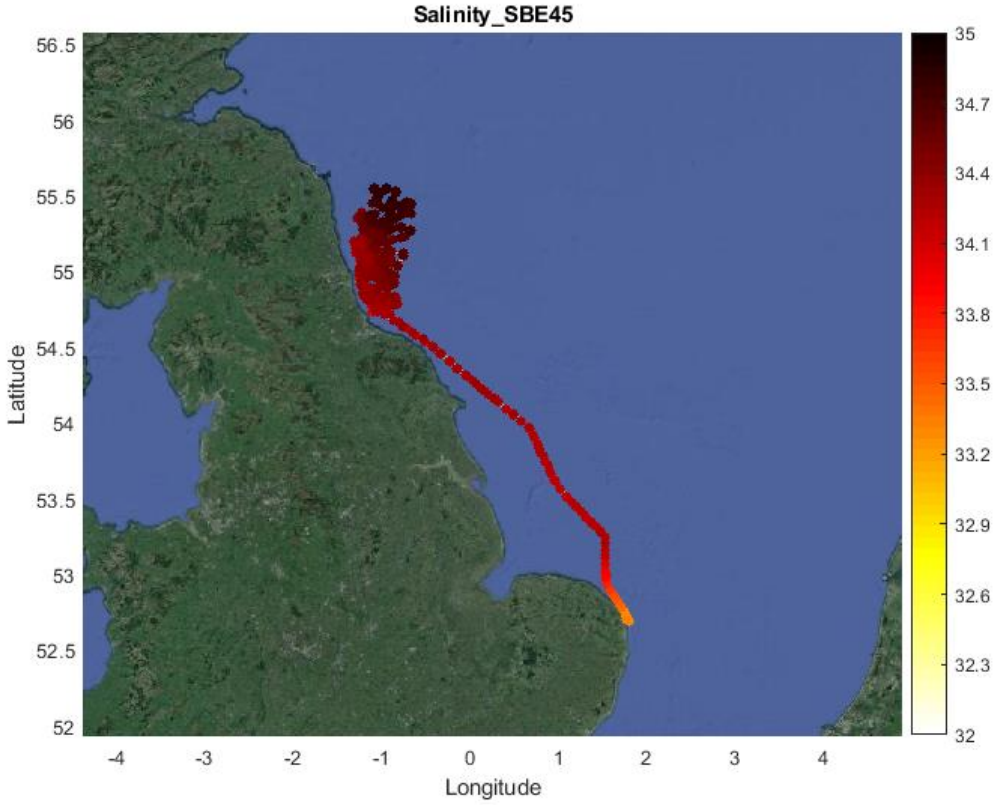
Import a file with data and with latitude and longitude coordinates. Select the variable(s) you would like to project over a map by the arrow. Tick the desired layout and press ok.

Your Google API Key (from your Google account) should be addressed (copy – paste) here



Plot Google menu

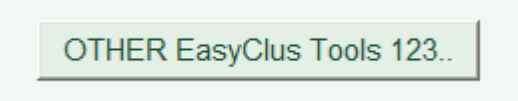




And this is what might be possible if your Google API key works.

6. Other EasyClus Tools 123..

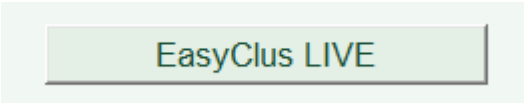
Extra module with some extra features.



- Merge pico-nano-micro different triggered files. Use '?' button to open local manual.
- From particle counts to single cell counts. Use '?' button to open local manual.
- Download images from website. Use '?' button to open local manual.

7. EasyClus LIVE

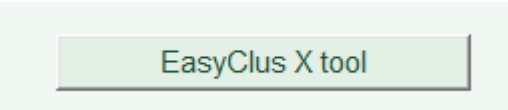
Extra module that should be purchased separately and which enables the ONLINE importing of new samples, ONLINE data processing on FCM-data, ONLINE control of the FCM functioning, ONLINE clustering and databasing function and the ONLINE reportage of chosen results to a website.

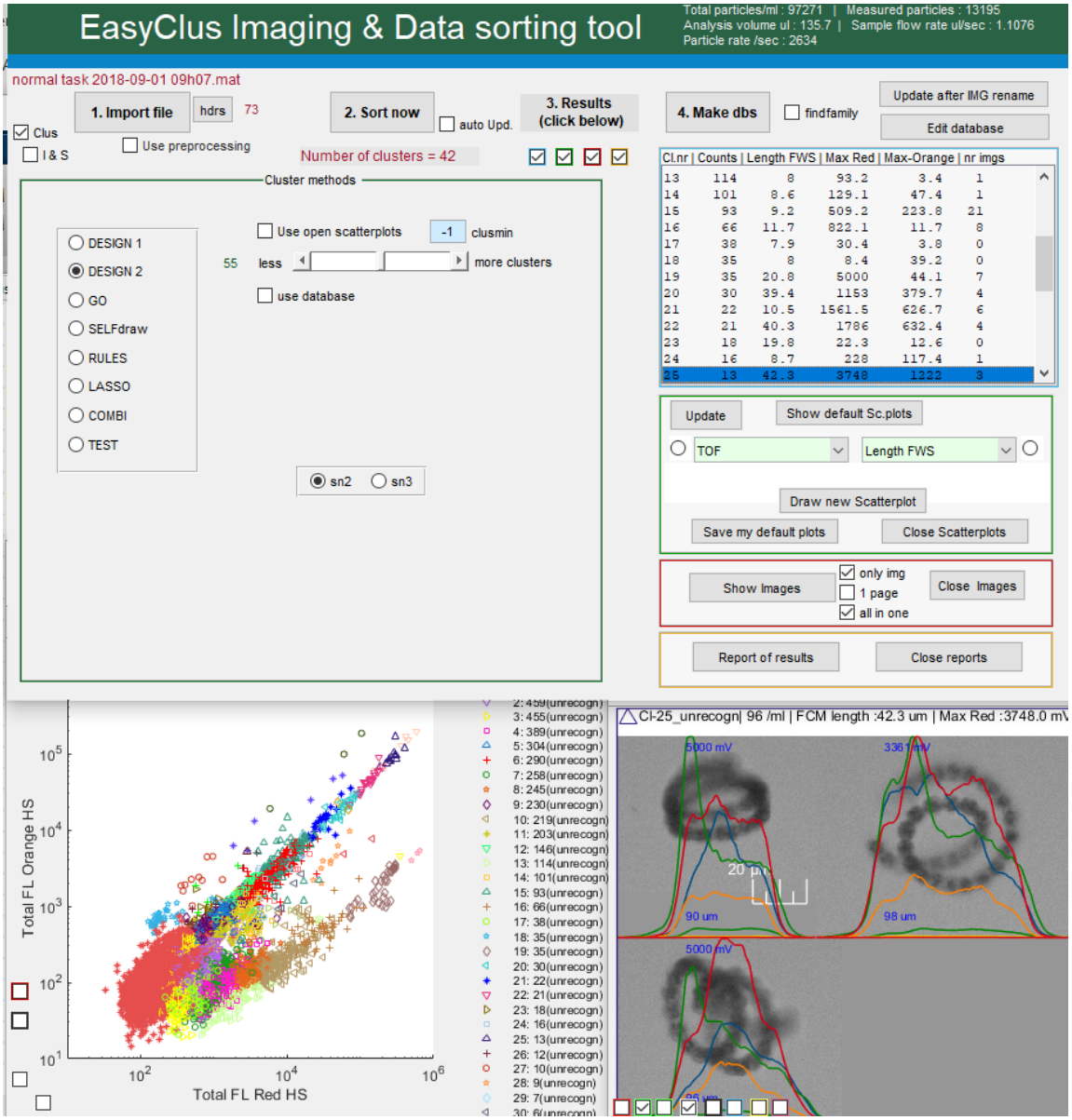


EasyClus LIVE is explained in another manual, but data processing works similar as the usual EasyClus. People who work with the usual EasyClus will understand the LIVE version easily. The LIVE-module is a plug & play module, because it uses a lot of default settings and generates files by itself.

8. The EasyClus X-tool

Very useful and easy tool to cluster by several methods and visualize results.





9. Final buttons

9.1 Close all figures

To close all figures that are open in Matlab

Press this button to close all open figures. Figures can also be closed individually one after each other by pressing the red cross in the upper right.

9.2 Exit

To leave the EasyClus menu and to return to the Matlab® command window

Press 'Exit' to leave the EasyClus menu and to return to the Matlab® command window. After this you can leave Matlab®.

To restart EasyClus from the Matlab command window, just type EasyClus and pres enter

10. Extra

Changing txt-files in Excel

If txt-files are changed in Excel, use the option ‘Save as’ text (TAB is delimiter).

Change decimal delimiter en column delimiter of datafile

In Matlab Command window (via ‘Exit’ button EasyClus), type ‘changesymbols’ and press ‘enter’.

The decimal delimiter comma ‘,’ is changed in dot ‘.’ and the column delimiter is chosen by the user. One file or several files can be processed. The new listmode_*.csv file is ‘saved’ in the given directory ..\new\same naam.csv .

Starting EasyClus v1.30 menu by typing ‘EasyClus ‘ without quotes in Matlab, followed by pressing ‘Enter’.

11. Possible problems and what to do

Possible problems	Description in Matlab	What to do?
EasyClus stops and give error messages ->	??? Error using ==> fprintf Invalid file identifier -1.	This usually happens if a file is already opened (e.g. in Excel), while EasyClus tries to save results to the same filename. Close opened file and type EasyClus in Matlab and try again. .
One of the menus of EasyClus is on my screen and stops continuation. I can't close this menu. Help!	No message	Try to close Matlab the rude way by using Ctr-Alt-Del, Open 'Task manager', look in the list for 'Matlab.exe' and close process. Matlab closes. If you want to proceed with EasyClus, start Matlab again.
I manipulated my database using Excel and my database only has 2 characters after delimiter.	No message	Database is saved using Excel saved, but maybe Excel was set standardly to only two characters behind delimiter. Put more characters by using 'Cell properties'.
I manipulated a csv-file and now I can't import this file anymore and I 've a EasyClus crash?	Different (red) error messages possible.	1)Be aware that your ',' and '.' are set right. '.' For decimal delimiter is crucial because Matlab uses this. 2)'Save as'-> use csv-file format. 3) If you have left text in columns except for the variable names in row 1 OR if the column header is called 'REF ..', EasyClus expects numbers.
Can't import cyz-file and EasyClus crash.	mscorlib error cyz-file error	CytoSense dll-problem. Do you have copied the dirname\x64 directory or dirname\x86 directories both filled with dll's to the EasyClus directory (path like dirname\extended\ directory?)
Can't import cyz	mscorlib error cyz-file error	Is .NET software installed? Needed for running CytioSenSe dll's. Solution: Install CytoClus (Cytobuoy software) and .NET is installed by itself.
Can't import cyz	mscorlib error cyz-file error	Look to properties of x64 and x86 directory or dll-files to check if they are blocked by your computer. Sometimes this happens for downloaded files automatically.
Can't import cyz	mscorlib error cyz-file error	If you're using LINUX this is a problem for importing cyz-files, because .NET is not supported by LINUX. In this case it is

		recommended to use other datafile formats or to use a Windows operating computer (necessary for using the CytoSense).
Can't import cyz	mscorlib error cyz-file error	Sometimes the x64 & x84 directory including files received after downloading or after unzip are corrupted due to unknown reasons (one time happened before). Solution is to use your latest working x64 & x86 directories incl files. Replace corrupted by your own ones.
Can't import cyz	Import problem in readcytodata2	Do you have the latest Cytosense.dll? Sometimes dll's are changed by CytoBuoy.
EasyClus version has changed and an error occurred	Probably there is a second 'dirsettings' file available for Matlab. There should be one dirsettings-file in the main directory of EasyClus.	Search in all Matlab-paths for dirsettings by 'edit'- 'find files' – 'type dirsettings in box and look in 'entire Matlab path'. If found, delete it.
Fprintf or print error to \figure\ or to \cluster MATLAB:print:CannotCreateOutputFile	This happened once using an old external harddisk and saving results to this harddisk AND containing a lot of files.	Get rid of bunch of files in these directories. E.g. rename figure-?figure2 and make new empty figure-directy
Slow EasyClus-writing-saving	This happened once using an old external harddisk and saving results to this harddisk AND containing a lto of files.	Get rid of bunch of files in these directories. E.g. rename figure-?figure2 and make new empty figure-directy
Out of memory	This can happen if a file is very big and your computer has not enough RAM-memory. Files up to 140.000 events have been processed <100 Mb)	
Open GL software / hardware	Matlab graphics uses your hardware graphics driver. If this driver is not found it will give a message openGL software used. If this happens again and again, update your graphical hardware driver.	Type : opengl info and find out what Matlab uses software : 'true' = used software : 'false' = not used
Open GL software / hardware	Force openGL hardware to be used by Matlab always	Type : `opengl('save','hardware')` (next startup it is activated)
Open GL software / hardware	Force openGL software to be used by Matlab always	Type : `opengl('save','software')` (next startup it is activated)