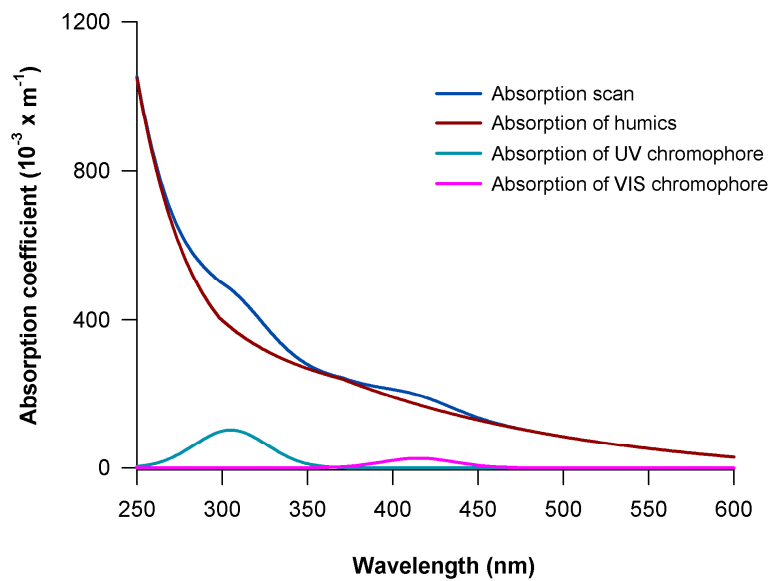


# CHROMOPHORES TOOLBOX TUTORIAL



See details in *Catalá, T. S., I. Reche, C.L. Ramón, Á. López-Sanz, E. Calvo, X. A. Álvarez-Salgado, Chromophoric signature of long-lived microbial by-products in the dark global ocean. Submitted*

## Contents

1. Introduction to the Chromophores-Toolbox
2. Chromophores-Toolbox Interface and key concepts
  - 2.1. Chromophores identification analysis
  - 2.2. Ch-UV
  - 2.3. Ch-VIS
  - 2.4. Chromophores-Toolbox components
  - 2.5. Chromophores-Toolbox interface
3. Getting started
4. Inputs to the Chromophores-Toolbox: Import data
5. Chromophores-Toolbox analysis
6. Outputs from the Chromophores-Toolbox

## 1. Introduction to the Chromophores-Toolbox

Absorption coefficients at specific wavelengths,  $a_\lambda$ , are used as proxies of the concentration of CDOM and a wide variety of spectral indices and slopes have provided key information on the origin and molecular structure of CDOM. As the characterisation of specific chromophores is still lacking in the literature we have developed a toolbox to obtain distinct chromophores within CDOM datasets.

The toolbox is adapted to obtain two distinct chromophores centred around 300 nm (named Ch-UV) and 415 nm (named Ch-VIS) by fitting the measured absorption coefficient spectra to the Röttgers and Koch's equation [Röttgers and Koch, 2012] and statistically isolating the absorption signal attributable to these chromophores from the standard decreasing exponential curve.

## 2. Chromophores-Toolbox Interface and key concepts

### 2.1. Chromophore identification analysis

In order to locate and quantify the absorption chromophores centred around 300 nm (Ch-UV) and 415 nm (Ch-VIS), we developed this Matlab toolbox that obtains the parameters that best fit (least squares sense) the following equation [Breves *et al.*, 2003; Röttgers and Koch, 2012]:

$$a_\lambda = b_1 e^{-b_2(\lambda-\lambda_0)} + b_3 \frac{1}{\sigma\sqrt{2\pi}} + e^{\frac{-(\lambda-\lambda_0)^2}{2\sigma^2}} + b_4 \quad (1)$$

The equation was fitted within different wavelength ranges for each chromophore. The **first and third terms** of the equation models the typical exponential decay of any CDOM spectrum with increasing wavelength. It consists of a pre-exponential term (**b<sub>1</sub>**), an exponential slope (**b<sub>2</sub>**), and an absorption parameter to correct for offsets in the absorption at longer

wavelength ( $b_4$ ). The **second term** is a Gaussian function to model the absorption spectra of the UV or VIS chromophore, with  $b_3$  being the height at the reference wavelength  $\lambda_0$  of the chromophore and  $\sigma$  the width of the Gaussian function, respectively. The absorption coefficient of the UV or VIS chromophore at their respective  $\lambda_0$  ( $a_{\text{Ch-UV}}$  or  $a_{\text{Ch-VIS}}$ ) is calculated as  $b_3 \frac{1}{\sigma\sqrt{2\pi}}$ .  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$ ,  $\lambda_0$  and  $\sigma$  are optimised with the Matlab toolbox.

## 2.2. Ch-UV

To obtain the parameters of the UV chromophore equation (1) is applied to the wavelength range from **250 to 400 nm**. This chromophore is related with the absorption of nitrate in the UVB region of the spectrum [Johnson and Coletti, 2002] or may be caused by the absorption of deoxygadusol, a mycosporine-like precursor synthesized via bacteria, cyanobacteria, phytoplankton, macroalgae (red, brown and green), plants and fungi with an absorption maximum at 294 nm [Shick and Dunlap, 2002].

## 2.3. Ch-VIS

To obtain the parameters of the VIS chromophore equation (1) is applied to the wavelength range from **350 to 600 nm**. A plausible source for the absorption of the VIS chromophore is the respiratory enzyme cytochrome c oxidase. This enzyme presents an absorption maxima around 410 to 415 nm and is one of the most important components of cellular respiration, which reduces molecular oxygen to water coupled to the pumping of protons across the mitochondrial or bacterial membrane [Yoshikawa et al., 2011].

## 2.4. Chromophores-Toolbox components

The Chromophores-Toolbox is a GUI (Graphical User Interface)-based Matlab toolbox. The Chromophores-Toolbox consists of three parts: (1) a main script related to the Matlab GUI interface (CHROMOPHORES.fig), which is called '**CHROMOPHORES.m**'. This script works as a connection (set of commands) between the interface components and the calculus done by Matlab; (2) A folder called '**functions**', which contains the Matlab functions that do the fitting

analysis (*Ch\_UV\_fit.m*, *Ch\_VIS\_fit.m*, *rottgers\_koch.m*) and the scripts that sends information to the interface log-window (*WriteinWindow.m*); and (3) a folder called 'Help', which contains information to help the user with the toolbox (this tutorial).

### 2.5 Chromophores-Toolbox interface

The Chromophores-Toolbox interface is divided in two main modules, (1) one for the UV chromophore and (2) another for the visible chromophore (see Fig. 1). Each module is divided, in turns, into the toolbox inputs (initial parameters), outputs (end parameters) and a "GO" command.

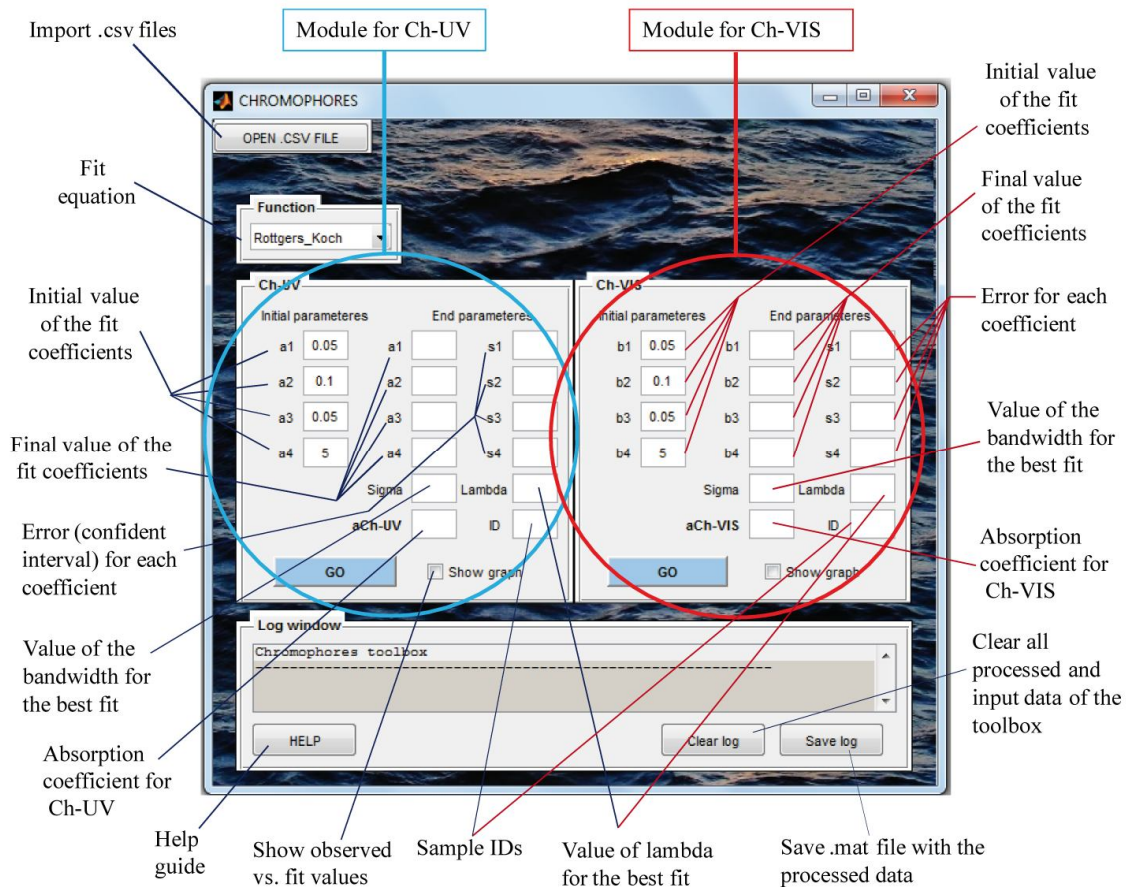


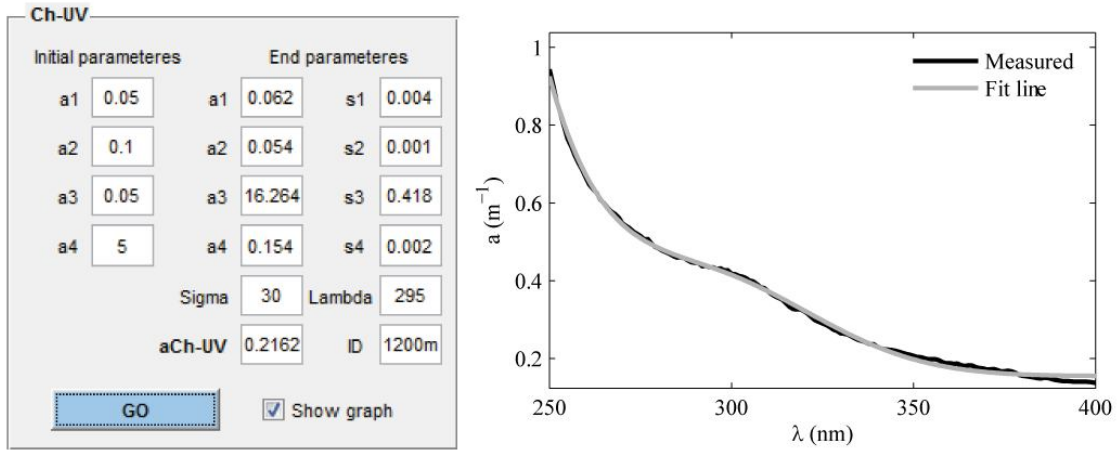
Fig. 1. Chromophores-Toolbox interface

Inputs to the toolbox are the initial value of the  $a_1$ - $a_4$  and  $b_1$ - $b_4$  fitting coefficients for the Ch-UV and Ch-VIS, respectively. These coefficients correspond to the initial value of coefficients  $b$  in Eq. 1. These initial values are set by default to be  $a_1 = 0.05$ ,  $a_2 = 0.1$ ,  $a_3 = 0.05$

and  $a_4 = 5$  (the same for the  $b$  coefficients, see Fig. 1) but can be changed by the user, before running the analysis, by clicking on the respective coefficient box and introducing a new initial value. Other inputs to the toolbox are the range of '*lambda*'  $\lambda_0$  and '*sigma*'  $\sigma$  (Eq. 1 and Fig. 1) to be tested. For Ch-UV,  $\lambda_0$  is allowed to vary within the range 280:5:320 nm and for Ch-VIS,  $\lambda_0$  is allowed to vary within the range 390:5:440 nm.  $\sigma$  is allowed to vary within the range 15:5:30 nm for both the Ch-UV and Ch-VIS. Although these ranges are fixed in the toolbox, and thus, there is not a direct access to them through the GUI interface, the user can change them by opening the script *CHROMOPHORES.m* and by modifying the values of L (wavelength) and D (bandwidth) in lines 408 and 409 for Ch-VIS and lines 1024 and 1025 for Ch-UV of that .m file.

Outputs of the toolbox (end parameters) consist of the final value (best fit) of the  $a_1$ - $a_4$  and  $b_1$ - $b_4$  fitting coefficients and their respective errors (confident intervals)  $s_1$ - $s_4$ , together with the coefficient of determination  $R^2$ , the value of  $\lambda_0$  (box *Lambda*),  $\sigma$  (box *Sigma*) and the absorption coefficient of the chromophore  $-a_{\text{Ch-UV}}$  or  $a_{\text{Ch-VIS}}$  for which the best fit was achieved. The corresponding sample ID (e.g. depth, time, name of the sample) is shown in the box *ID*.

By clicking the '*GO*' button, the fit analysis is run. If the box *Show Graphs* is ticked the toolbox will display a graph showing the measured absorption coefficients together with the best-fit line for each evaluated sample (e.g. Fig. 2).



**Fig. 2.** Example of the interface outputs and graph for the measured and calculated absorption coefficients

The interface also contains a button called '**OPEN .CSV FILE**' to import one or several .CSV files with the measured absorption coefficients (see Section 4 for details), a box called '**FUNCTIONS**' where the fit equation is defined (Eq. 1), a '**Log Window**' where the toolbox displays each of the actions that is carrying out, a '**HELP**' button which contains the basic information to use the Chromophores-Toolbox (this tutorial), a '**Clear log**' button, which erases all the information (imported data and outputs) and a '**Save log**' button, which allows the user to save the outputs in .mat format of the last analysis carried out by the toolbox.

### 3. Getting started

The Chromophores-Toolbox can be downloaded from '<http://ecologia.ugr.es/pages/herramientas/toolbox-matlab?lang=en>'. Once the file is unzipped it is important to maintain the structure (distribution) of the files within the folder *Chromophores\_Toolbox* for a correct function of the toolbox. Once the user has chosen the directory where to locate the toolbox there are two ways to run it:

- (1) Open Matlab and in *Current Folder* set the path to the toolbox: Eg: *Current Folder*:  
 C:\User\Chromophores\_Toolbox. Once Matlab is in the same directory as the toolbox,

write 'CHROMOPHORES' in the Command Window. The GUI interface should open.

- (2) Open the CHROMOPHORES.m file in the *Matlab Editor* and run the script (press the play button or go to Debug-Run CHROMOPHORES.m). The GUI interface should open.

It is important not to change the Matlab directory while the toolbox is being used. Otherwise an error would appear in the Matlab command window.

#### **4. Inputs to the Chromophores-Toolbox: Import data**

The Chromophores-Toolbox is set to load and read .csv files with information on the measured wavelengths and CDOM absorption coefficients. The first column of the .csv file should contain the wavelength values and the subsequent columns, the absorption coefficient values for each of the samples. The first row should contain identification for each sample (sample ID). This sample ID could be a measured depth, time, etc. Fig. 3 shows an example of an input .csv. It is important to keep the distribution shown in Fig. 3. The analysis will be done no matter the resolution of the sample, i.e. the separation between consecutive values of wavelengths (e.g. each 1 nm, each 5 nm, etc). An example of input .csv file can be downloaded from (<http://ecologia.ugr.es/pages/herramientas/toolbox-matlab?lang=en>).



	A	B	C	D	E	F	G	H	I
1		4000m	3000m	2500m	1700m	900m	470m	400m	200m
2	250	0.0532	0.0453	0.0475	0.0469	0.0469	0.0467	0.042	0.0482
3	251	0.0513	0.0435	0.0458	0.0451	0.0452	0.0451	0.0402	0.0465
4	252	0.0501	0.0424	0.0445	0.0438	0.0437	0.0436	0.0388	0.0453
5	253	0.0485	0.0406	0.0428	0.0422	0.0422	0.042	0.0372	0.0436
6	254	0.0475	0.0395	0.0416	0.0411	0.0411	0.0407	0.036	0.0424
7	255	0.0464	0.0386	0.0408	0.0404	0.04	0.04	0.0351	0.0417
8	256	0.0454	0.0374	0.0395	0.0391	0.0389	0.0387	0.0341	0.0406
9	257	0.0443	0.0365	0.0386	0.038	0.0376	0.0378	0.0328	0.0391
10	258	0.0438	0.0358	0.0379	0.0374	0.0374	0.0371	0.0322	0.0386
11	259	0.0431	0.035	0.0371	0.0366	0.0364	0.0365	0.0315	0.0379
12	260	0.0423	0.0344	0.0365	0.0362	0.0357	0.0358	0.031	0.0373
13	261	0.0417	0.0336	0.0356	0.035	0.0351	0.0353	0.0302	0.0365
14	262	0.0412	0.0329	0.0348	0.0345	0.0343	0.0345	0.0295	0.036
15	263	0.0408	0.0322	0.0346	0.0339	0.0338	0.0338	0.0288	0.035
16	264	0.0407	0.032	0.0341	0.0337	0.0339	0.0335	0.0289	0.0351
17	265	0.0392	0.0314	0.033	0.0327	0.0326	0.0326	0.0276	0.0341
18	266	0.0383	0.0301	0.0322	0.032	0.0323	0.032	0.027	0.0333
19	267	0.0383	0.0305	0.0323	0.0319	0.0318	0.0319	0.0267	0.0329
20	268	0.0378	0.0295	0.0314	0.0312	0.0312	0.0313	0.0262	0.0323
21	269	0.0374	0.0293	0.0312	0.0306	0.0308	0.031	0.0259	0.0323
22	270	0.0366	0.0284	0.0302	0.0298	0.0306	0.0302	0.0246	0.031
23	271	0.0354	0.0277	0.0292	0.0288	0.0296	0.0288	0.0238	0.03
24	272	0.0354	0.0267	0.0288	0.0287	0.0291	0.0287	0.0242	0.0298
25	273	0.0349	0.0263	0.0288	0.0281	0.0285	0.0281	0.0234	0.0293
26	274	0.0341	0.026	0.0277	0.0275	0.0283	0.0281	0.023	0.0287
27	275	0.0335	0.0254	0.0271	0.0267	0.0275	0.0269	0.0224	0.0275

Fig. 3. Example of .csv file to be loaded by the toolbox

The user can import one or several .csv files. In case the user loads several .csv files, the analysis for Ch-UV and Ch-VIS will be carried out for all samples within a given .csv file and all the imported .csv files.

## 5. Chromophores-Toolbox analysis

The Chromophores-Toolbox fits the data in  $x$  (wavelengths  $\lambda$ ) and  $y$  (absorption coefficients) to the function in Eq. 1., with the fitting method of non linear squares, where  $b_1$ - $b_4$  are the coefficients to estimate, and  $\lambda_0$  and  $\sigma$  are the problem parameters. In order to start the fitting iterations, an initial value of the fit coefficients should be given. These initial values are set by default to be  $b_1 = 0.05$ ,  $b_2 = 0.1$ ,  $b_3 = 0.05$  and  $b_4 = 5$ , but they can be directly changed by the user by changing the initial value of these parameters in the interface (see section 2.5). In order to obtain results, which are realistic in terms of a physical interpretation,

the coefficients  $b_1$  and  $b_3$  (the height at  $\lambda_0$ ) are forced to be  $\geq 0$  in the model. By default all coefficients have an upper-limit value of 20 and a lower-limit value of -10 (except for  $b_1$  and  $b_3$  which is 0). These limits, however, can be modified by the user in lines 393 to 399 and lines 1010 to 1016 (example below) in the '*CHROMOPHORES.m*' file for Ch-VIS and Ch-UV, respectively. The toolbox will show a warning message if the calculated value of  $a_3$  or  $b_3$  is too close ( $<0.001$ ) to their upper limit, suggesting the user to increase the upper limit for these coefficients.

#### **%2. Fit Conditions:**

```
optionsb = fitoptions('Method','NonlinearLeastSquares',... ;  
    'StartPoint',b0,...  
    'Lower',[0 -10 0 -10],...  
    'Upper',[20 20 20 20],...  
    'MaxIter',1000,...  
    'MaxFunEvals',1000);
```

## **6. Outputs from the Chromophores-Toolbox**

The Chromophores-Toolbox generates two folders, one for each chromophore under study, called '*Ch\_UV*' and '*Ch\_VIS*', in the same directory as the imported .csv files, and where the outputs related to the UV chromophore and the visible chromophore will be located, respectively. By default, the Chromophores-Toolbox generates one .csv file after each analysis called '*Parameters\_ChUV\_[FileName]<sub>1</sub>\_[FileName]<sub>end</sub>.csv*' or '*Parameters\_ChVIS\_[FileName]<sub>1</sub>\_[FileName]<sub>end</sub>.csv*', depending on the chromophore, with the information on the fitting coefficients, the goodness of the fitting, the final values of  $\lambda_0$  and  $\sigma$ , and the calculated value for  $a_{\text{Ch-UV}}$  or  $a_{\text{Ch-VIS}}$ .  $[FileName]_1$  and  $[FileName]_{\text{end}}$  are the names of the first and the last imported .csv files. For example, for three imported files called '*Example1.csv*', '*Example2.csv*' and '*Example3.csv*', the Chromophores-Toolbox would generate a .csv file called '*Parameters\_ChUV\_Example1\_Example3.csv*' after the analysis of the UV chromophore. If only one .csv file is imported, the Chromophores-Toolbox will generate a .csv

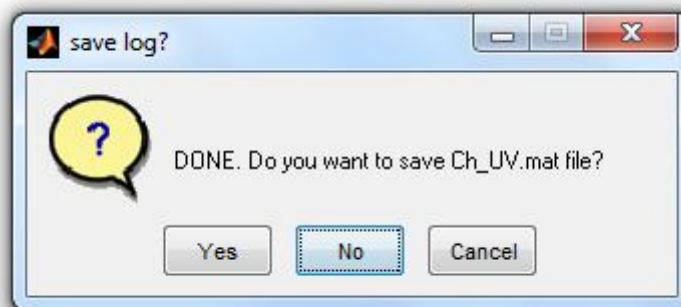
file called 'Parameters\_ChVIS\_[FileName].csv'. Fig. 4 shows an example of the 'Parameters\_ChUV\_[FileName]\_1\_[FileName]\_end.csv' generated by the Chromophores- Toolbox. Note that the 'Parameters\_ChUV\_[FileName]\_1\_[FileName]\_end.csv' is generated for each 'GO' command. Thus, it will contain the information for several imported files if more than one .csv file has been loaded into the toolbox.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Example1.csv												
2	Sample ID	a1	a2	a3	a4	sa1	sa2	sa3	sa4	R2	Lambda(nm)	Sigma(nm)	aCh-UV(1/m)
3	4000m	0.4163774	0.0249467	0.4254075	0.0540878	0.0075177	0.0004444	0.2175398	0.0039217	0.9993193	285	15	0.0113142
4													
5	Example2.csv												
6	Sample ID	a1	a2	a3	a4	sa1	sa2	sa3	sa4	R2	Lambda(nm)	Sigma(nm)	aCh-UV(1/m)
7	4000m	0.0539624	0.0565319	12.853368	0.0910721	0.002831	0.0012817	0.3435305	0.0018264	0.9984967	295	30	0.1709251
8	3000mR	0.0520134	0.0564041	13.19717	0.1143315	0.002885	0.001355	0.349053	0.0018516	0.998365	295	30	0.175497
9	2000m	0.056018	0.0629067	15.892677	0.1225537	0.0036107	0.0017577	0.4159339	0.002102	0.9981089	290	30	0.211342
10	1200m	0.061799	0.0539483	16.263709	0.1543773	0.0036596	0.0014437	0.4184411	0.0021198	0.9982788	295	30	0.216276
11	600m	0.0789688	0.0509089	16.04699	0.18835	0.0040352	0.001241	0.4307794	0.0020442	0.9986868	295	30	0.2133941
12	370m	0.1188646	0.0383593	11.538134	0.220477	0.0051962	0.000943	0.4644906	0.0019384	0.999007	300	30	0.153435
13	320m	0.119889	0.0407392	11.602415	0.1354628	0.0058734	0.0011548	0.5007417	0.0017863	0.9989714	295	30	0.1542898
14	200m	0.0954071	0.046578	12.841082	0.2021197	0.0055953	0.0014118	0.5426603	0.0023044	0.9982674	295	30	0.1707617
15	150m	0.1905735	0.0350771	10.735813	0.1430252	0.0141942	0.0018141	0.9944459	0.002784	0.9986966	290	30	0.1427657
16	120m	0.438854	0.0221344	1.7511749	0.1425604	0.0088459	0.0004351	0.2268303	0.0051522	0.9993324	290	15	0.0465745
17	100m	0.6370804	0.0225964	1.6038449	0.1538181	0.0049764	0.0001909	0.1400158	0.0028435	0.999811	280	15	0.0426561
18	67m	0.3897257	0.0378461	1.8475391	0.0809438	0.0160487	0.0013242	0.5906189	0.0050372	0.9978379	280	15	0.0491374
19	40m	0.4448933	0.0349249	1.3397668	0.1098756	0.0144798	0.0010113	0.5165056	0.0050853	0.9983023	280	15	0.0356326
20	3m	0.3610886	0.0394763	1.394694	0.0570373	0.0173741	0.001573	0.6494296	0.0051314	0.9973361	280	15	0.0370935
21													

**Fig. 4.** Example of 'Parameters\_ChUV\_[FileName]\_1\_[FileName]\_end.csv' generated by the Chromophores Toolbox. In this example two .csv files called 'Example1.csv' (with one sample) and 'Example2.csv' (with 14 samples) were loaded into the toolbox.

It is important to keep the output .csv files closed while the toolbox is doing the analysis. Otherwise an error would appear in the Matlab Command Window.

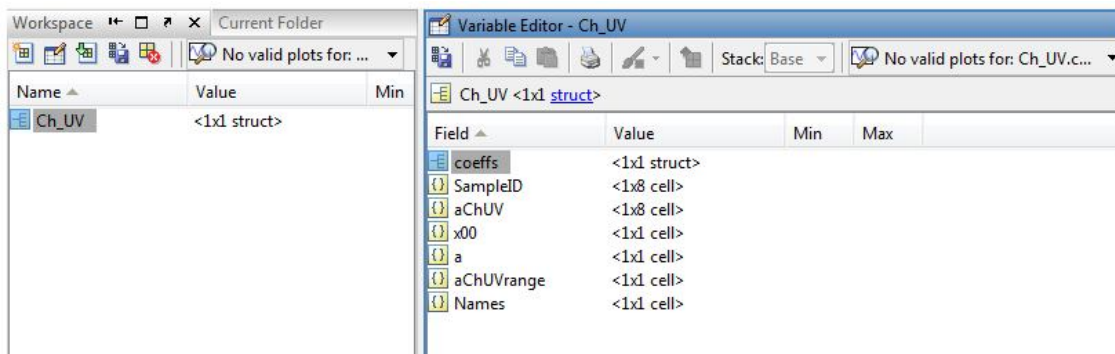
Alternatively, after each analysis ('GO' command), the Chromophores-Toolbox gives the option to save output data in .mat format. After each analysis a question dialog box will open, asking whether the user would like to save the output data in .mat format (Fig. 5).



**Fig. 5.** Question dialog box after the analysis for the Ch-UV

If the user presses 'Yes', a '*Ch\_UV.mat*' file or a '*Ch\_VIS.mat*' will be generated and the user will select the directory to save it. This .mat file will contain the following information (Fig. 6):

- Variable *Coeffs*: This contains all the variables in the '*Constants\_Ch\_UV.csv*' (or '*Constants\_Ch\_VIS.csv*') for each sample and each imported .csv file.
- Variable *SampleID*: all the sample IDs
- Variable *x00*: Range of wavelengths for which absorption coefficients were measured (x-variable) for each sample and .csv file ( $\lambda$  in Eq 1). Units = nm.
- Variable *a*: Best fit of the absorption coefficients (Eq. 1) for each sample and .csv file (y-variable). Units =  $m^{-1}$ .
- Variable *aChUVrange* or *aChVISrange*: Absorption coefficients for the Ch-UV range or Ch-VIS range, respectively (term 2 in Eq. 1). Units =  $m^{-1}$ .
- Variable *aChUV* or *aChVIS*: absorption coefficients for the UV or visible chromophores. Units =  $m^{-1}$ .
- Variable *Names*: Name of each imported .csv file.



**Fig. 6.** Example of data in a *Ch\_UV.mat* file.

If the user presses 'No' or 'Cancel' in the question dialog box, there is still the possibility to save the *.mat* file by clicking on the interface 'Save log' button. The question dialog box will reopen.

## References

1. Breves, W., R. Heuermann, and R. Reuter (2003), Enhanced red fluorescence emission in the oxygen minimum zone of the Arabian Sea, *Ocean Dynam.*, 53, 86–97, doi:10.1007/s10236-003-0026-y.
  2. Röttgers, R., and B.P. Koch (2012), Spectroscopic detection of a ubiquitous dissolved pigment degradation product in subsurface waters of the global ocean, *Biogeosciences*, 9, 2585–2596.
  3. Johnson, K.S., and L.J. Coletti (2002), In situ ultraviolet spectrophotometry for high resolution and long-term monitoring of nitrate, bromide and bisulfide in the ocean, *Deep-Sea Res. I*, 49, 1291–1305.
  4. Shick, J.M., and W.C. Dunlap (2002), Mycosporine like amino acids and related gadusols: Biosynthesis, accumulation, and UV-protective functions in aquatic organisms, *Ann. Rev. of Physiology*, 64, 223–262.
- Yoshikawa, S., K. Muramoto, K., and Shinzawa-Itoh (2011), Proton-Pumping Mechanism of Cytochrome c Oxidase, *Ann.u Rev. Biophys.*, 40, 205–223.