The final publication is available at

dx.doi.org/10.1016/j.talanta.2008.12.073

Additional Information

DETERMINATION OF PHYTOPLANKTON COMPOSITION USING ABSORPTION SPECTRA

Martínez-Guijarro R.[a*], **Romero I.**[a], **Pachés M.**[a], **del Río J.G.**[a], **Martí C.M.**[a], **Gil G.**[a], **Ferrer-Riquelme A.**[b], **Ferrer J.**[a]

[a] Institute for Water Engineering and the Environment [Polytechnic University of Valencia], Camino de Vera s/n 46022 Valencia, Spain

[b] Department of Statistics and Applied Operative Research and Quality [Polytechnic University of Valencia], Camino de Vera s/n 46022 Valencia, Spain

[*] Corresponding author. Tel.:34-96-3877616; fax: 34-96-3877618. E-mail address: mmarting@hma.upv.es

Abstract

Characterisation of phytoplankton communities in aquatic ecosystems is a costly task in terms of time, material and human resources. The general objective of this paper is not to replace microscopic counts but to complement them, by fine-tuning a technique using absorption spectra measurements that reduces the above-mentioned costs. Therefore, the objective proposed in this paper is to assess the possibility of achieving a qualitative determination of phytoplankton communities by classes, and also a quantitative estimation of the number of phytoplankton cells within each of these classes, using spectrophotometric determination.

Samples were taken in three areas of the Spanish Mediterranean coast. These areas correspond to estuary systems that are influenced by both continental waters and Mediterranean Sea waters. 139 Samples were taken in 7-8 stations per area, at different depths in each station. In each sample, the absorption spectrum and the phytoplankton classes (Bacyllariophyceae (diatoms), Cryptophyceae, Clorophyceae, Chrysophyceae, Prasynophyceae, Prymnesophyceae, Euglenophyceae, Cyanophyceae, Dynophyceae and the S*ynechococcus* sp) were determined.

Data were analysed by means of the Partial Least Squares (PLS) multivariate statistical technique. The absorbances obtained between 400 and 750 nm were used as the independent variable and the cell/l of each phytoplankton class were used as the dependent variable, thereby obtaining models which relate the absorbance of the sample extract to the phytoplankton present in it. Good

results were obtained for diatoms (Bacillarophyceae), Chlorophyceae and Cryptophyceae.

1.- Introduction

Phytoplankton is one of the organic components of natural waters and, therefore, phytoplankton diagnosis is important to assess the ecological condition of coastal waters (Fadeev et al. [2]).

As explained in Millán-Núñez et al. [8], the light absorption by particulate matter, including phytoplankton in the ocean, is of great significance, since the absorption and dispersion of light causes colouring of the sea. The particle variability allows to determine the attenuation of light, the primary productivity and the biomass of phytoplankton pigments. Moreover, some authors (Yentsch and Phinney, [12]; Nelson and Prezelin, [10]; Cleveland, [1]) showed that changes in the optical characteristics of masses of water were related to cellular biochemical processes inherent in the consumption of energy by photosynthesis. Many of them concluded that there is a non-linear relationship between the light absorption coefficient by phytoplankton and the chlorophyll *a* concentration. This relationship is complicated due to phytoplankton properties, such as, size, particles shape and accessory pigments.

Phytoplankton contains pigments (chlorophylls a, b and c, carotenoids and phycobiliproteins) in different proportions; thus, phytoplankton identification on the basis of the absorption spectrum depends on pigment composition in cells. Pigment composition in the chloroplasts provides a way to classify the algae group. For example, coloured chloroplasts in diatoms (Bacillariophyceae) are usually yellow-brown due to xantophylls, whereas blue-green colours (cyanophytes) are variable, within a range that encompasses from blue-green to red due to phycobiliproteins, phycoerythrins (red) and phycocyanins (blue).

Pigments that form the basis for chemical taxonomical discrimination absorb light in wavelengths in the visible spectrum (Moberg et al. [9]). Therefore, an alternative way to obtain qualitative information about phytoplankton composition is to analyse samples absorption through the visible spectrum. Qualitative information about phytoplankton classes and photoadaptation is included in such spectrum data, but the key is the interpretation thereof (Moberg et al. [9]). In general, it is better to interpret or evaluate the spectrum data by means of chemometrics, a branch of chemistry devoted to extracting information from large sets of data.

Furthermore, if the presence of blooms or toxic species of phytoplankton (such as some dinoflagellates) is detected in advance in the phytoplankton population, the harmful effects produced could be mitigated and even prevented (Kirkpatrick et al., [6]). Microscopic examination of water samples is the main method used to detect such toxic groups. Unfortunately, this method is slow, laborious and intermittent. For this reason, optical detection methods and automated methods have been developed to determine the presence of these species in the phytoplankton population in a quick and continuous manner.

Some laboratory works suggest that it is possible to perform group discrimination on the basis of cellular absorption. For example, Johnsen et al. [5] used discriminant analysis to classify the absorption spectrum amongst 31 bloom-formers (which represent the four main groups of phytoplankton with respect to accessory chlorophylls, for example Cl b, Cl c_1 and/or Cl c_2, Cl c_3 and non-accessory chlorophylls), thereby distinguishing dinoflagellates and toxic prymnesiophytes which contain Cl c_3 from taxons that do not have this pigment.

It could be argued that, if the absorption spectrum of each individual pigment is known, the absorption spectrum of the phytoplankton may be easily reconstructed from the concentration of pigments. However, the relation between pigments concentration and phytoplankton absorption coefficients is not linear, due to the "package effect". This effect is caused by the fact that pigments are not in solution, but rather, packed inside the cells (and in cells inside chloroplasts).

The package effect varies with cell size, with intracellular concentration of several pigments and with wavelength. Due to these sources of variation (which depend on environmental factors), it is difficult, if not impossible, to precisely model and predict this effect for natural populations. Given the complexity of this effect, it is expected that neuronal network techniques will make it possible to approach the relation model between pigment concentrations and absorption spectra using current measurements that implicitly take the package effect into consideration.

As explained in Perry and Darling [11], phytoplankton, other particles and chromophoric dissolved organic matter are susceptible to radiometric optical sensors because they absorb, disperse, attenuate and fluoresce light with optical pattern characteristics (models). Phytoplankton, as a photosynthetic organism, absorbs electromagnetic radiation primarily within the blue, blue-green and red bands of the visible spectrum and absorption coefficient is determined by pigment composition. Because they are particles, phytoplankton disperse light. The manner in which dispersion of the spectrum takes place is dependent on size (of the phytoplankton), composition and absorption spectrum. Other non-algal organic particles, such as bacteria and detritus, are relatively weak absorbers, with the maximum absorption in the UV region. As is the case of phytoplankton, the way in which dispersion of the spectrum takes place is dependent on the size distribution. Suspended sediments generally disperse more than they absorb, although a strong absorption has been observed in mineral sediments, particularly in iron-rich minerals.

2.- Experimental

2.1.- Sampling

The study was performed in three areas of the Eastern coast of Spain. These areas correspond to estuarine systems which are formed when the fresh/brackish waters of the continental systems flow into the Mediterranean Sea. Samplings campaigns were performed on 29 March 2006 at the area called Almenara plume, on 30 March 2006 at the Albufera plume in Valencia and on 17 May 2006 at the Estany plume in Cullera.

Water samples were collected in 2-litre polyethylene bottles and 250-ml glass bottles. They were kept refrigerated until arrival at the laboratory, which never took longer than 12 hours. A total of 139 samples were collected, corresponding to 7-8 stations per area, and in each station samples were taken at different depths.

2.2.- Analytical techniques

Determination of the absorption spectrum was performed on a 90% acetone extract obtained filtering water samples through a cellulose acetate membrane (Millipore 0.45 µm HAWP04700). Subsequently, they were frozen to break the cells and to facilitate the pigments release. Filters with the retained particulate material were introduced in 6 ml of 90% acetone. On this extract absorbance was determined at 1-nm intervals, at wavelengths between 400 and 750 nm. A 1-cm quartz cuvette and a Perkin Elmer Lambda 35 spectrophotometer were used. Absorbance values obtained at 750 nm were subtracted from the values between 400 and 749 nm, in order to eliminate the absorbance which is not caused by the pigments.

In order to analyse phytoplankton communities, epifluorescence microscopic count method was used. Samples contained in a 250-ml glass bottle were fixated with glutaraldehyde until a final concentration of 2%. They were filtered with 0.2-µm membranes (Millipore GTTP), filters were washed with distilled water to eliminate the retained salt and, subsequently, they were dehydrated with successive washes with 50%, 80%, 90% and 99% ethanol. Each dried filter was placed onto a drop of immersion oil in the centre of a slide and 2 more drops were added on the top side of the filter. Finally, a coverglass was placed on the top of the filter (Fournier [3]). Phytoplankton counts were performed by epifluorescence microscopy with a Leica DM2500, using the 100x-oil immersion objective. A minimum of 300 cells were counted and at least 100 cells of the species or genera more abundant were counted with an error lower than 20% (Lund [7]).

2.3.- Statistical techniques

Statistical technique called PLS (Partial Least Squares) was used for the multivariate analysis of the experimental data obtained (Geladi and Kowalski, [4]).

To develop the model that relates absorbances to phytoplankton classes, 2/3 of the total number of samples (139) were used to fit the model, and the remaining one-third of samples (46) were used to validate it. Even though the statistical programme used, SIMCAP 9.0, makes an internal estimation of the model prediction ability by means of "cross-validation", the only way to be absolutely sure regarding a model prediction ability is to make external predictions, that is, to make predictions for an independent system of observations.

After developing this first model, a second model was developed which made it possible to verify that the results improved when a filter called OSC (Orthogonal Signal Correction) was applied to the X variables (absorbances). The purpose of this method was to correct the matrix of X data by removing the information that was orthogonal to the Y response matrix, that was, the information which was not related to the response of interest. This preliminary pre-processing method was jointly applied to all the spectra in the calibration set. Subsequently, the correction made on the X matrix was applied to the external set of data in order to verify the true prediction ability of the model built from the corrected data.

3.- Results and discussion

At the three sampled areas, it was found that the closest stations to the continental water outflow exhibited a vertical salinity and chlorophyll *a* gradient, due to freshwater surface layer present in those stations. This layer was not very thick; in most cases it was less than 40 cm thickness, for that reason practically all the samples analysed were saline (36.84 ±1.68 g/kg) (salinity and chlorophyll a date are not showed).

Samples had a phytoplankton composition with a high content of diatoms, Prasinophyceae and Chlorophyceae as compared to the rest of eukaryotes cells, as it is shown in figure 1 for the Albufera plume. Although we have not included prokaryotes (colonial Cyanophyceae and S*ynechococcus* genus) in figure 1, it is worth mentioning that at the Albufera plume area there are Cyanophyceae which are not present in the other two areas.

Absorption spectra of the pigment-containing extracts in each sample had a similar shape, with peaks at 440 and 664 nm, which correspond to chlorophyll *a*, the main pigment present in phytoplankton cells. An example is shown in figure 2 for the Albufera plume. Differences should be studied in the accessory pigments (chlorophylls b and c, carotenoids, etc.); for that reason the area between 400 and 500 nm would have to be enlarged, since these pigments maximum absorption peaks are in this region.

A first model that related the absorption spectra of the sample extract to phytoplankton composition (model 1) was developed. Subsequently, in order to improve the results obtained with model 1, a second model was developed, but in this case the OSC filter was applied to the X variables (absorbances).

To decide the adequate number of components for model 1, it must be taken into account the general model fit, which is defined by parameter R^2, and the prediction ability thereof, which is defined by parameter Q^2. Table 1 shows that the eigenvalue of the first component (81.7) is greater than the one from the other two components; furthermore, $Q^2_{(cum)}$ values decreases for the third component and, consequently, the model ability prediction. Also, when components 2 and 3 are used, the increase in R^2 is very small. For all these reasons, it is sufficient to use only one component in this model. Results obtained with model 1 are shown in figure 3a, where it can be observed that very good results for the individual model responses for phytoplankton classes are achieved for diatoms (Bacillariophyceae), Chlorophyceae and Cyanophyceae, both in terms of the model fit (R^2) and its ability to predict these phytoplankton classes content (Q^2). However, results for Cyanophyceae were discarded because the presence of this class was only significant in samples from the Albufera plume area, whereas in the remaining samples most of the Cyanophyceae values were equal to the technique detection limit.

As previously discussed, model 2 was obtained by applying the OSC filter to the X variables previous to performing the PLS analysis. Results obtained with this model improve the former. Table 1 shows that eigenvalue of the first component (92.9), as was the case with model 1, is very large by comparison to

the following ones. In this case, it was decided to use 3 components because, in addition to taking into consideration the R^2 and Q^2 values (table 1), this was the number of components that led to the best individual response results. Figure 3b shows that this model produced adequate results also for Cryptophyceae.

Figure 4 shows diatoms, Cryptophyceae and Chlorophyceae values predicted with model 2 against the real contents of the samples. In this figure, it can be observed that good predictions were obtained, since the R^2 of the linear fit between predicted and real values was 0.9197 for diatoms, 0.7808 for Cryptophyceae, and 0.9758 for Chlorophyceae.

Once obtained model 2, it was validated by means of the external set of samples, which were used to make predictions of the phytoplankton content. As in the case of the calibration set, the OSC filter was first applied to the absorbances. This set corresponded to one-third of the total samples taken, which were not used for the model fit. Prediction results for these samples can be seen in figure 5. In this case, good predictions were still obtained for diatoms and Chlorophyceae, but not for Cryptophyceae (figures 5a, 5c and 5b, respectively), since the predicted values obtained were not acceptable.

4.- Conclusions

In this paper, the absorption spectra of the sample extracts led to good results in determining of diatoms and Chlorophyceae content. An acceptable model was also obtained for Cryptophyceae, although in this case, in order to obtain good results, the absorbance values must be processed prior to applying the multivariate statistical technique. Phytoplankton classes for which good results were not obtained were due to either they had a limited presence in most of the samples studied or they had low pigments cellular quota.

In order to improve this work, a study should be performed on samples with different phytoplankton abundance and composition to the samples already studied, since it is possible that the majority presence of certain classes makes it

difficult to determine other classes with type and pigments content that have overlapping or masking spectra.

Phytoplankton determination through extract samples absorption spectra is a simple and cheap method that, albeit having limitations, may be used to complement microscopic counts.

References

[1] Cleveland, J.S., 1995. Regional models for phytoplankton absorption as function of chlorophyll a concentration. J. Geophys. Res. 100(7): 13333-13344.

[2] Fadeev V.V., D.V., Maslov, D.N. Matorin, R. Reuter, and T.I. Zavyalova, 2000. Some peculiarities of fluorescence diagnostics of phytoplanctonkton in coastal waters of the Black sea. [checked 13-04-2007]: Available in: http://las.physik.uni-oldenburg.de/eProceedings/vol01_1/01_1_fadeev1.pdf.

[3] Fournier, R., 1978. Membrane filtering. In: Sournia, A. (ed), Phytoplankton Manual. Monographs on oceanographic methodology. Unesco: 197-201

[4] Geladi, P. and B.R. Kowalski, 1986. Partial Least-Squares Regression: A tutorial. Anal. Chim. Acta, 185: 1-17.

[5] Johnsen, G., O. Samset, L. Granskog, and E. Sakshaug, 1994. In vivo absorption characteristics in 10 classes of bloom-forming phytoplankton: Taxonomic characteristics and responses to photoadaptation by means of discriminant and HPLC analysis. Mar. Ecol. Prog. Ser., 105: 149-157.

[6] Kirkpatrick, G.J., D.F. Millie, M.A. Moline and O. Schofield, 2000. Limnol. Oceanogr., 45(2): 467-471.

[7] Lund, J.W.G., Kipling, G. and Le Cren, E.D., 1958. The inverted microscope method of estimating algal numbers and the statistical basis of estimations by counting. Hydrobiologia, 11 (2), 143-170.

[8] Millán-Núñez, E., M.E. Sieracki, R. Millán-Núñez, J.R Lara-Lara, G. Gaxiola-Castro and C.C Trees, 2004. Specific absorption coefficient and phytoplankton biomass in the southern region of the California Current. Deep-Sea Res. II 51: 817-826.

[9] Moberg, L., B. Karlberg, K. Sϕrensen and T. Källqvist, 2002. Assessment of phytoplankton class abundance using absorption spectra and chemometrics. Talanta 56: 153-160.

[10] Nelson, N.B. and B.B. Prezelin, 1993. Phytoplankton light absorption and package effect in California coastal waters. Mar. Ecol. Prog. Ser., 94: 217-227.

[11] Perry M.J. and I.C. Darling, 2003. Optical sensors. [checked 22-03-2007]: Available in: http://www.geo-prose.com/ALPS/white_papers/perry.doc.

[12] Yentsch, C.S. and D.A. Phinney, 1989. A bridge between ocean optics and microbial ecology. Limnol. and Oceanogr., 34: 1694-1705.

A	R^2X	R^2X(cum)	Eigenvalues	R^2Y	R^2Y(cum)	Q^2	limit	Q^2(cum)	significance	Iterations
Model 1										
0	Cent			Cent						
1	0,878	0,878	81,70	0,369	0,369	0,303	0,05	0,303	R1	3
2	0,0788	0,957	7,33	0,026	0,395	0,00295	0,05	0,305	R2	9
3	0,0123	0,969	1,14	0,036	0,431	-0,0144	0,05	0,295	R2	25
Model 2										
0	Cent			Cent						
1	0,999	0,999	92,9	0,373	0,373	0,303	0,05	0,303	R1	2
2	0,000452	1	0,0421	0,0608	0,434	0,0545	0,05	0,341	R1	9
3	9,05e-005	1	0,00842	0,0806	0,515	0,0832	0,05	0,396	R1	11

**Table 1**: PLS results of model 1 and model 2 with X filtered variables (OSC)
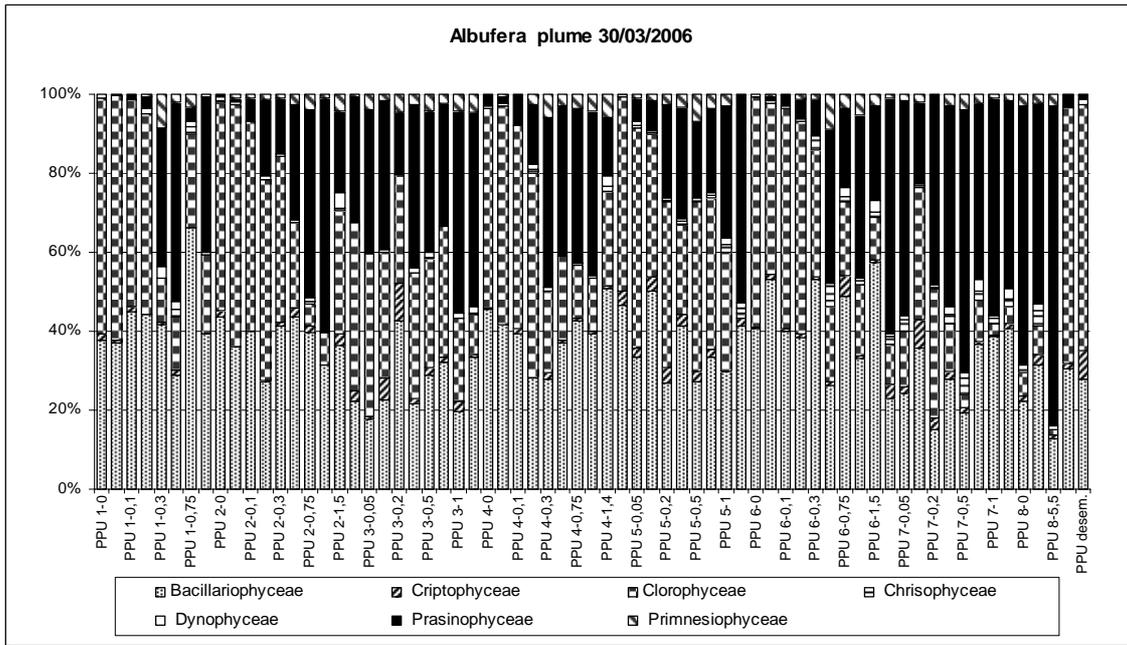
Figure 1: Phytoplanktonic composition (eukaryota) of the Albufera plume samples
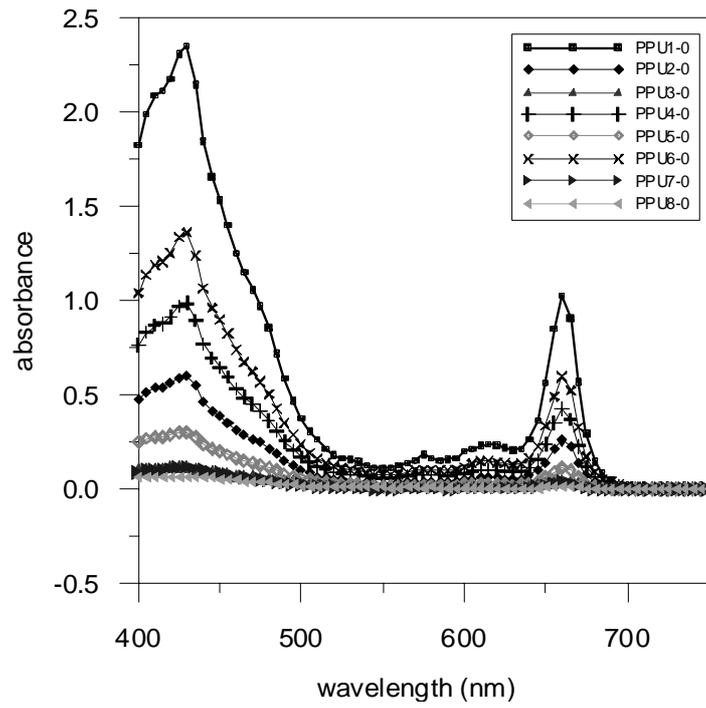
**Figure 2**: Absorption spectrum of superficial samples from the Albufera plume.
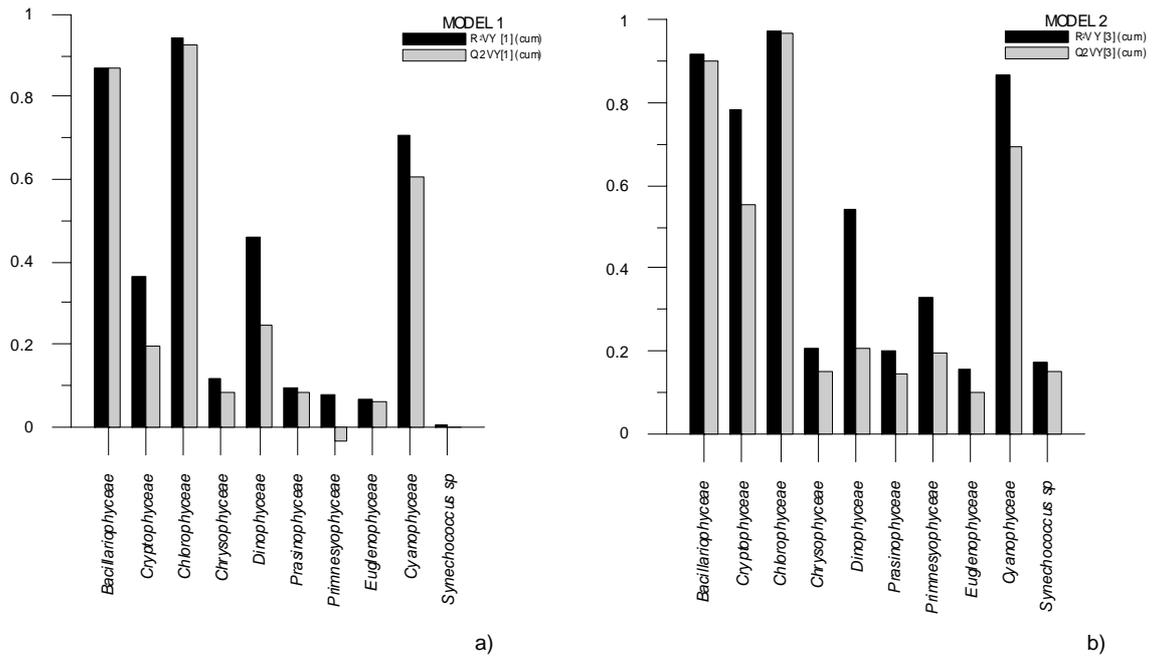
Figure 3: Statistical adjustment (R^2) and prediction capacity (Q^2) for each individual responses.
a) model with 1 component; b) model with 3 components (OSC filter)
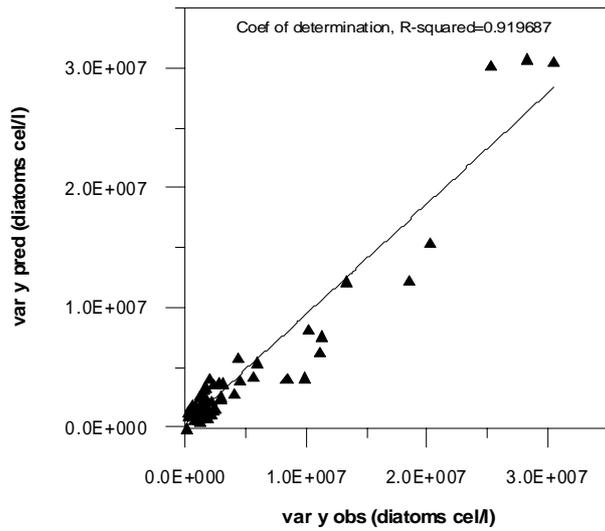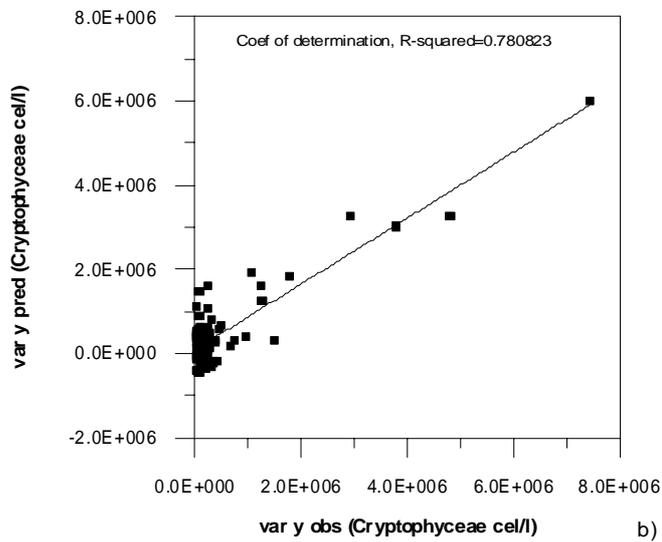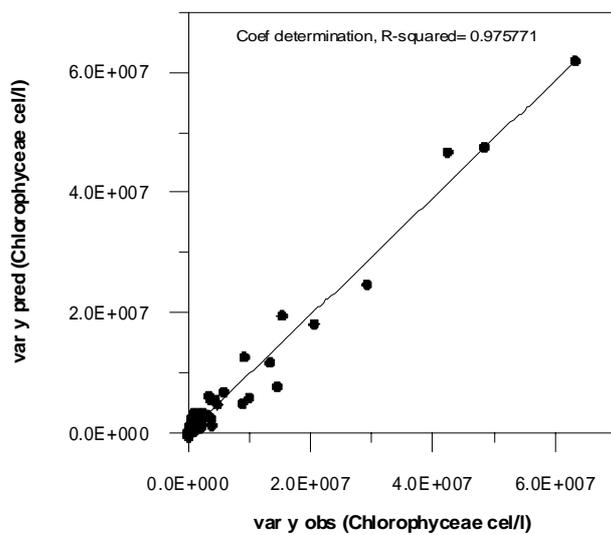
Figure 4: Graphical representation of the observed values versus the predicted values from model 2 with samples used for the adjustment of the model. a)diatoms (*Bacillariophyceae*); b)*Cryptophyceae*; c)*Chlorophyceae*
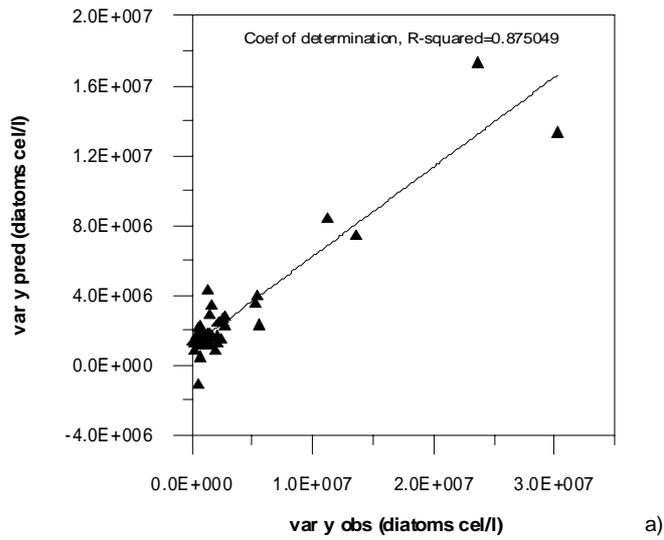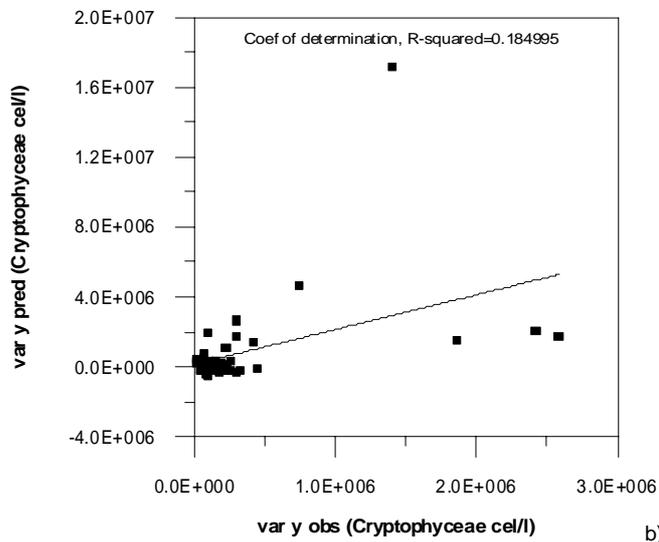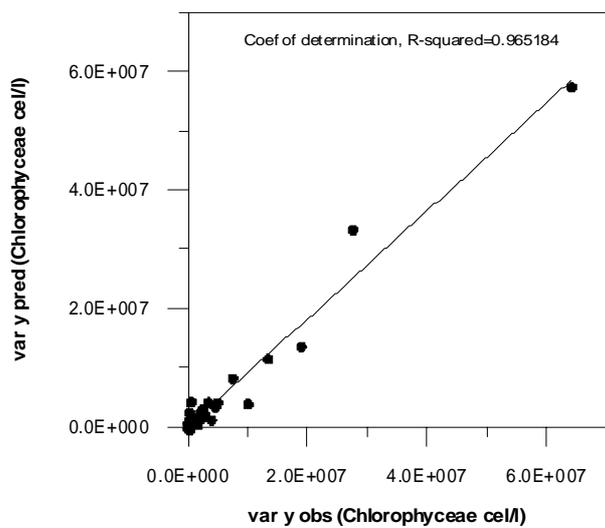
Figure 5: Graphical representation of the observed values versus the predicted values from model 2 for an external group of samples. a)diatoms (*Bacillariophyceae*); b)*Cryptophyceae*; c)*Chlorophyceae*