JOURNAL OF PLANKTON RESEARCH | VOLUME 34 | NUMBER 6 | PAGES 454–469 | 2012

Featured Article

Improvement of plankton biovolume estimates derived from image-based automatic sampling devices: application to FlowCAM

EVA ÁLVAREZ*, ÁNGEL LÓPEZ-URRUTIA AND ENRIQUE NOGUEIRA
CENTRO OCEANOGRÁFICO DE GIJÓN, INSTITUTO ESPAÑOL DE OCEANOGRAFÍA, 33212 GIJÓN, SPAIN

*CORRESPONDING AUTHOR: eva.alvarez@gi.ieo.es

Received November 16, 2011; accepted in principle January 23, 2012; accepted for publication February 17, 2012

Corresponding editor: Roger Harris

The most commonly used biomass estimate for microalgae is obtained from cell biovolume, usually calculated from microscopically measured linear dimensions. Although reliable, this is a highly time-consuming and specialized technique. Automated sampling devices that acquire images of cells and use pattern recognition techniques to identify the images have been developed as an alternative to microscopy-based methods. There are some aspects of automatic sampling and classification methods, however, which can be improved for the analysis of field samples including living and non-living particles. In this work, we demonstrate how the accuracy of a state-of the-art technique for plankton classification (Support Vector Machine) can be improved up to 86% if a previous automated step designed to remove non-living images is included. There is a tendency with the currently applied automatic methods to misestimate cell biovolume due to the two-dimensionality of the images. Here, we present a data set of more than 500 samples to show that the greatest effect is caused by the incorrect estimation of biovolume of the chain-forming diatoms. This results in an overestimate of biomass of between 20 and 100% where chain-forming diatoms represent more than the 20% of the biomass of the sample. We show how the classification method can be adapted to provide not only taxonomic but also the morphological classification of cells in order to obtain a more reliable estimate of biovolume according to the predicted cell shape, in a way comparable with microscopy-based estimates.

KEYWORDS: biovolume; biomass; size spectra; machine learning; Support Vector Machine (SVM); FlowCAM

INTRODUCTION

The biomass size distribution in aquatic and terrestrial ecosystems has been an important focus of research in ecology during the last decades. The "particle-size" approach began to develop in the last century as a unifying theory that established a relationship between the biomass of organisms in any size category and the individual body weight of these organisms (Rodriguez and Mullin, 1986). The planktonic community soon appeared to be a very appropriate object of study to test

doi:10.1093/plankt/fbs017, available online at www.plankt.oxfordjournals.org. Advance Access publication March 16, 2012

the hypotheses of the "particle-size" approach due to its wide size range (Quiñones *et al.*, 2003). An extensive oceanic survey carried out by Sheldon *et al.* (Sheldon *et al.*, 1972) provided the first empirical evidence of a general feature of marine ecosystems: the same amount of biomass is allocated in logarithmically equal biovolume ranges (Sheldon *et al.*, 1972). The study of this relationship was carried out with a Coulter counter (Sheldon and Parsons, 1967), a device capable of counting and sizing particles in a fluid, which enabled the authors to analyse more than a hundred of un-preserved samples distributed all over the world ocean.

The introduction of automated sampling devices designed to characterize the abundance and size of planktonic organisms in a reasonable amount of time has furthered the application of this "particle-size" approach. One disadvantage, however, of these faster methods such as the Coulter counter is the lack of identification of the particles analysed. In an attempt to solve the problem of knowing what kind of particles are sampled, several automated devices that acquire images of the particles have been developed. These new technologies brought together the need to examine, identify and automatically measure large numbers of images (Benfield *et al.*, 2007). The automated classification of plankton images has specific challenges because of the highly variable proportion of taxonomic groups depending upon location, time and nature of the survey, morphological heterogeneity, variable proportion of non-living targets, different orientations, partial occlusion and single images containing more than one particle. In order to overcome these challenges, several authors have developed automatic classification techniques based on pattern recognition that have yielded very promising results. These classification techniques are general mathematical techniques that can be applied to the images obtained from different instruments, e.g. the Video Plankton Recorder (Tang *et al.*, 1998; Davis *et al.*, 2004), the Shadow Imaging Particle Profiling Evaluation Recorder (Luo *et al.*, 2004), the ZooScan (Grosjean *et al.*, 2004; Gorsky *et al.*, 2010), the FlowCytobot (Sosik and Olson, 2007) or the Flow Cytometer And Microscope (FlowCAM; Blaschko *et al.*, 2005; Zarauz *et al.*, 2009).

With the exception of holographic methods, all image-based automatic sampling devices have in common that they do not capture the three-dimensional shape of the particles. To calculate the volume of planktonic cells, the methods based on two-dimensional images usually consider all cells as spheres and use the equivalent spherical diameter (ESD) of the area projected by the cell on a flat surface. This is a good approximation in the case of spherical or elliptical cells but a poor one in the case of cylindrical cells. A cylinder with an aspect ratio of 0.5 has a volume 26% lower than the volume of a sphere with the same projected area; the underestimate reaches 54% if the aspect ratio is of 0.2. On the other hand, a discus with an aspect ratio of 0.5 has a volume 47% higher than the volume of the equivalent sphere; the overestimate reaches 136% if the aspect ratio is of 0.2. This can bias the estimation of biovolume when the particles are placed preferably in some orientations, a typical effect that occurs when particles lay on a flat surface (e.g. microscopy, scanner) or are embedded in a moving fluid (e.g. FlowCAM). The effect is dependent on the proportion of cylindrical cells in the sample. Under these conditions, the simplification of assuming that all particles are spheres can render poor biovolume or biomass estimates per group, genera or species, even with a perfect classification at these prescribed taxonomic levels.

Microscopy-based methods are considered the most accurate for plankton enumeration and biovolume estimate and have been the most extensively used methodology for the study of microplankton community size structure (Cermeño and Figueiras, 2008). The most widely used protocol is the one proposed by Utermöhl (Utermöhl, 1958), in which the preserved water sample is settled on a microscope slide. An expert taxonomist examines the slide under the microscope, identifies the taxonomic groups and estimates the abundance of each group. To estimate the biovolume of the cells, the expert needs to assign a geometrical shape to each taxonomic group, measure a certain number of linear dimensions and estimate the hidden dimensions following aspect ratios previously calculated. Hillebrand *et al.* (Hillebrand *et al.*, 1999) proposed a set of 20 geometric models to be use for the determination of microalgal biovolume. Although this set of shapes was thought to minimize the effort of microscopic measurement, it also aims to be as close to the actual shape of the organism as possible, so the shape that describes a cell can be as simple as a sphere, or complex, formed by various geometrical bodies that require the measurement of several dimensions. However, the description of a cell with a complex shape usually requires accommodating structures that do not constitute a large proportion of the total volume of the cell, such as spines or setae (Menden-Deuer *et al.*, 2000). For this reason, considering the intrinsic error of microscopic volume measurements and the sometimes inaccurate designation of composite shapes, the description of complex shapes can be approximated by three simple shapes: sphere, ellipsoid and cylinder (Menden-Deuer *et al.*, 2000). The effort needed to process samples according to this technique turns it into a highly time-consuming and specialized task, although this technique is considered the most reliable and is used as a

benchmark for the assessment of new techniques for plankton enumeration and biovolume estimation.

In order to provide biovolume estimates comparable to those obtained with microscopy-based techniques using automated classification devices, it is necessary to consider the three-dimensional shape of the cells. Our strategy in the present work has been to imitate the working procedures followed by microscopy-based techniques, that is, to predict the shape of the cells and make some assumptions about the hidden dimension. To do this, we adapted the currently used techniques to automatically classify the images not only taxonomically but also morphologically. The aim of this work is to develop a technique to classify automatically images from an automatic sampling device (FlowCAM) that permits at the same time a more accurate estimate of the biovolume of the cells. With this aim we have: (i) adapted the state-of-the-art classification techniques to classify images from FlowCAM in order to obtain morphological information of the cells; (ii) used the morphological information to improve the estimate of the biovolume of the cells digitized in two-dimensional images and (iii) used the taxonomic information to identify situations where poor size-structure characterization or biomass estimation may result.

METHOD

A total of 526 natural samples were taken with a rosette sampler system in the Bay of Biscay between August 2008 and April 2010 at 313 different sampling stations. Figure 1 shows the geographic location of all sampling stations and the season of the year when they were sampled. At all the stations, samples were collected at the surface, and additional 213 samples were obtained between 10 and 75 m. All samples were pre-filtered at sea through a 200-μm mesh-size net, kept fresh, stored in the dark and analysed with FlowCAM on board or in the laboratory within a few hours. All samples constituted our data set, subdivided into: a training set used to train the automated classifier, a test set to evaluate it and a field evaluation set to illustrate the application of the automatic classification method. The training set was built using images from 86 randomly selected samples and indicated in Fig. 1 as grey dots. The test set consisted of 17 samples, obtained over 2 years in a monthly sampling of a coastal station located in the Cantabrian Sea, so these samples cover the whole seasonal cycle. Another two samples, taken on the French shelf, were collected in triplicate and used to explore the different calculations of biovolume. Finally, the rest of the samples constituted the field evaluation set.

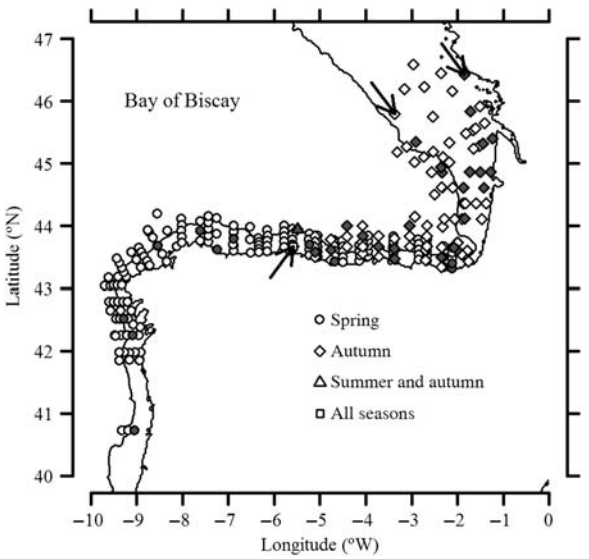

Fig. 1. Geographical location of the sampling stations. The shape of the point indicates the season when the samples were taken. Grey dots indicate samples used to build the training set. The upward arrow indicates the location of the series of samples used as the test set. Downward arrows indicate the location of samples used to explore the effect of different ways of calculating the cell biovolume.

Automatic processing of FlowCAM samples

The automatic processing of FlowCAM samples consists of several steps:

Digitalization of samples

Each sample was split in an un-concentrated subsample for analysis with the 200× magnification to digitize particles between 3 and 50 μm, and in a concentrated subsample (1 L concentrated by reverse filtration (Dodson and Thomas, 1978) down to around 20 mL) for the 100× magnification to digitize particles between 6 and 100 μm. The lower limit for both magnifications corresponds to particles covering an area smaller than 10 × 10 pixels that do not have enough resolution to be identified, even by a human expert, and thus they are not considered. Photographs can be captured following three working modes: in autoimage mode photographs are taken at a constant rate (in our case, 11 frames s^{-1}), while in fluorescence-triggered and side-scatter-triggered modes photographs are taken only when a particle that emits fluorescence or scatters light, respectively, passes in front of the camera. Samples were analysed in autoimage and fluorescence-triggered modes; side-scattered mode was not used as it has been shown to give unreliable results (Álvarez *et al.*, 2011). Thirty minutes were considered a maximum running time for each subsample. In the subsamples processed with autoimage, 20 000 photos were taken, which corresponds to

a net volume of 0.072 mL in 20× and 14.91 mL in 10×. In the subsamples processed with fluorescence triggered mode, 1 mL was processed at 0.0333 mL min^{-1} in the 20× and 10 mL at 0.3333 mL min^{-1} in the 10×. The FlowCAM software, Visual SpreadSheet, extracts each particle present in a photograph (a process called image segmentation), and stores the rectangular section on disk. The result is a plankton sample turned into a collection of images, each containing an individual particle. All rectangular sections are combined in collages that constitute the raw output of the VisualSpreadSheet.

Description of the images

We followed the philosophy of the SCOR 130 working group (Benfield *et al.*, 2007) promoting the intercomparison of plankton pictures from various origins. Zoo/PhytoImage is an open source software based on R for plankton image analysis that can be modified to meet the requirements of the user and accommodate many different imaging systems (Grosjean, 2005). We followed the instructions proposed by the Zoo/PhytoImage platform in terms of the conventional names of the samples, the inclusion of metadata and the format of the images and the data.

In order to import the raw data into the Zoo/PhytoImage convention, we used the EBImage R package from Bioconductor (Sklyar and Huber, 2006; Pau *et al.*, 2010), which provides general purpose functionality for the reading, writing, processing and analysis of images as well as tools to transform the images, segment cells and extract quantitative cellular descriptors. The collage files provided by the FlowCAM software were separated in individual images and a series of morphological features were calculated for each image. To localize the particle within each rectangular section we used the intensity-based method proposed by Tang (Tang *et al.*, 1998). Once the particle was located, new features were extracted. These included simple shape and texture descriptors (Blaschko *et al.*, 2005), invariant moments (Hu, 1962), granulometric features (Luo *et al.*, 2004; Blaschko *et al.*, 2005), edge features (Sklyar and Huber, 2006), texture features from the co-occurrence matrix (Haralick *et al.*, 1979) and Zernike features (Zernike, 1934). In total, 136 features or image attributes described each particle (see Supplementary Data).

Generation of a training and a test set

A training set is a group of example images classified by a human expert that is used to train the classification algorithm. In our case, these images come from 86 randomly selected samples of the data set. To manage the images we used the tools provided by the Zoo/PhytoImage software. The training set was established by means of the *ad hoc* method, that is, with a variable number of items within each group as chosen by the user (Davis *et al.*, 2004), so that the largest diversity of organisms could be considered. We have created three different training sets (TS): (1) for the images obtained with the 200× magnification (TS200×, composed by 3600 images); (2) for the images smaller than 20 μm in ESD obtained with the 100× magnification (TS100×$_{<20}$, 9500 images); and (3) for the images larger than 20 μm ESD obtained with the 100× magnification (TS100×$_{>20}$, 9500 images). We divided each of the training sets in a group of artefacts, a group of detritus and a variable number of living-particle groups indicated in Table I and Fig. 2. These living groups were made based on three criteria: size, taxonomy and morphology (Bakker *et al.*, 1985). Each group was homogeneous for each of the criteria, containing cells of the same size range, same taxon and same shape. The training sets TS200× and TS100×$_{<20}$ separate organisms based only on their shape and size, without taxonomic differentiation (Groups 1 to 4 in Table I). In the TS100×$_{>20}$, each living group includes individuals which are taxonomic and morphologically homogeneous (Groups 5 to 34 in Table I).

The test set is a group of images classified by a human expert and not included in the training set that is used to evaluate the efficiency of the classifier. The test set consisted of 19 samples of the data set with 70 900 images between 3 and 100 μm. The living particles were separated from detritus and artefacts by visual inspection and compared with predictions of the classifier to estimate the efficiency of the automated method.

Pretreatment of the images for the elimination of artefacts and detritus

The relative importance of non-living particles, which include artefacts (bubbles, background images and repeated images) and detritus, and the living particles differs depending on the FlowCAM working mode used to digitize the sample. The fluorescence-triggered mode does not record many non-living particles because they are not photographed due to their absence of red fluorescence, whereas autoimage mode captures a variable and usually large amount of images of non-living particles that interferes with the classification of living particles.

Instead of eliminating artefacts and detritus manually, as it is commonly done (Zarauz *et al.*, 2007), we implemented a series of cleaning filters. For the artefacts, we designed the filters using the information of the time of recording and the position in the field of view of the particles. We considered as bubbles, those images segmented with a frequency higher than 10 per second; as

Table I: Classification of FlowCAM images in groups according to size, shape and taxonomic affiliation

Taxon class	Groups	Shape class
1. Others (all <20 μm)	Spherical cells (1)	Sphere
	Elliptical cells (2)	Ellipsoid
	Discoid cells (3)	Disc
	Cylindrical cells (4)	Cylinder
2. Others >20 μm	Cells unidentified (5)	Ellipsoid
	Flagellates (6)	Ellipsoid
3. Silicoflagellates	Silicoflagellates (7)	Ellipsoid
4. Dinoflagellates	Small dinoflagellates (8)	Ellipsoid
	Prorocentrum micans (9)	Ellipsoid
	Genera *Prorocentrum* (10)	Ellipsoid
	Genera *Dinophysis* (11)	Ellipsoid
	Globose un-horned cells (e.g. *Protoperidinium*) (12)	Ellipsoid
	Globose horned cells (e.g. *Ceratium*) (13)	Ellipsoid
5. Ciliates	*Mesodinium rubrum* (14)	Ellipsoid
	Few cilia in apical position (e.g. *Laboea*, *Strombidium*) (15)	Ellipsoid
	Few cilia in lateral position (e.g. *Euplotes*, *Diophrys*) (16)	Ellipsoid
	Tintinnids with extended lorica (e.g. *Salpingella*) (17)	Ellipsoid
	Tintinnids with globular lorica (e.g. *Codonella*, *Dictyocysta*) (18)	Ellipsoid
6. Diatoms	Pennate cells (19)	Ellipsoid
	Centric cells with height < diameter (20)	Disc
	Centric cells with height > diameter (21)	Cylinder
	Cylindrical cells with round ends (e.g. *Ditylum*, *Stephanopyxis*, *Corethron*) (22)	Cylinder
	Cylindrical cells with spine-like ends (e.g. *Rhizosolenia*) (23)	Cylinder
	Wide chains with closed connected cells (e.g. *Detonula*, *Guinardia*) (24)	Cylinder
	Narrow chains with closed connected cells (e.g. *Leptocylindrus*) (25)	Cylinder
	Chains with slightly connected cells (e.g. *Eucampia*) (26)	Cylinder
	Chains with long setae (e.g. *Chaetoceros*) (27)	Cylinder
	Chains with intercellular spaces (e.g. *Skeletonema*) (28)	Cylinder
	Chains with cells connected by gelatinous threads (e.g. *Thalassiosira rotula*) (29)	Cylinder
	Chains of needle-like cells (e.g. *Nitzschia*) (30)	Cylinder
	Colonial star-shaped (e.g. *Asterionellopsis*) (31)	Sphere
	Colonial zig-zag (e.g. *Thalassionema*) (32)	Sphere
	Gelatinous spheres (e.g. *Phaeocystis*, *Thalassiosira subtilis*) (33)	Sphere
7. Crustaceans	Crustaceans (34)	Ellipsoid
8. Detritus	Faecal pellets, exuvia, dead fragments	—
9. Artefacts	Bubbles, empty images, background noise	—

The name of the groups that contains a description of the species included are based on the classification of Bakker *et al.* (Bakker *et al.*, 1985). These groups are merged, after automatic classification, in the taxon and shape classes indicated.

background images, those recorded more than 10 times in the same position of the camera field of view; and as repeated images, those recorded in consecutive photographs and with an Euclidean distance between their image attributes lower than 50. For the detritus, we have explored the 136 image attributes searching for a selected subset of them able to discriminate between detritus and living particles. In Fig. 3, two of these selected attributes are shown to illustrate the procedure we have followed: the mean value of grey intensity for each particle (Fig. 3A), that represented a lower limit for living particles, and the skewness of grey intensity (Fig. 3B), that represented an upper limit for living particles. Thus, setting a lower and an upper limit value, respectively, for these two attributes permits the identification of many non-living particles. Obviously, there is not a single image attribute that can distinguish between detritus and living particles unequivocally, but a combination of attributes can be used to establish a statistical filter to identify most of detritus particles. In total, a subset of 50 attributes were selected as they discriminate living particles in their upper limit and a subset of 10 attributes were selected as they discriminate living particles in their lower limit. The range of these attributes for the living particles in the training set were extended a 5% for the 200× images and a 2% for the 100× images in their discrimination limit, to establish a confidence interval within the living particles are expected to be included. We labelled as detritus those images in the test and field evaluation sets that show values of any of the attributes selected outside its confidence interval.

Training the algorithm

Support Vector Machine (SVM) was used to classify the rest of the particles after the pretreatment of the images. We used the library LIBSVM (Chang and Lin, 2009) with its R interface: *e1071* package (Karatzoglou *et al.*, 2006) and *ipred* package (Peters and Hothorn, 2009).

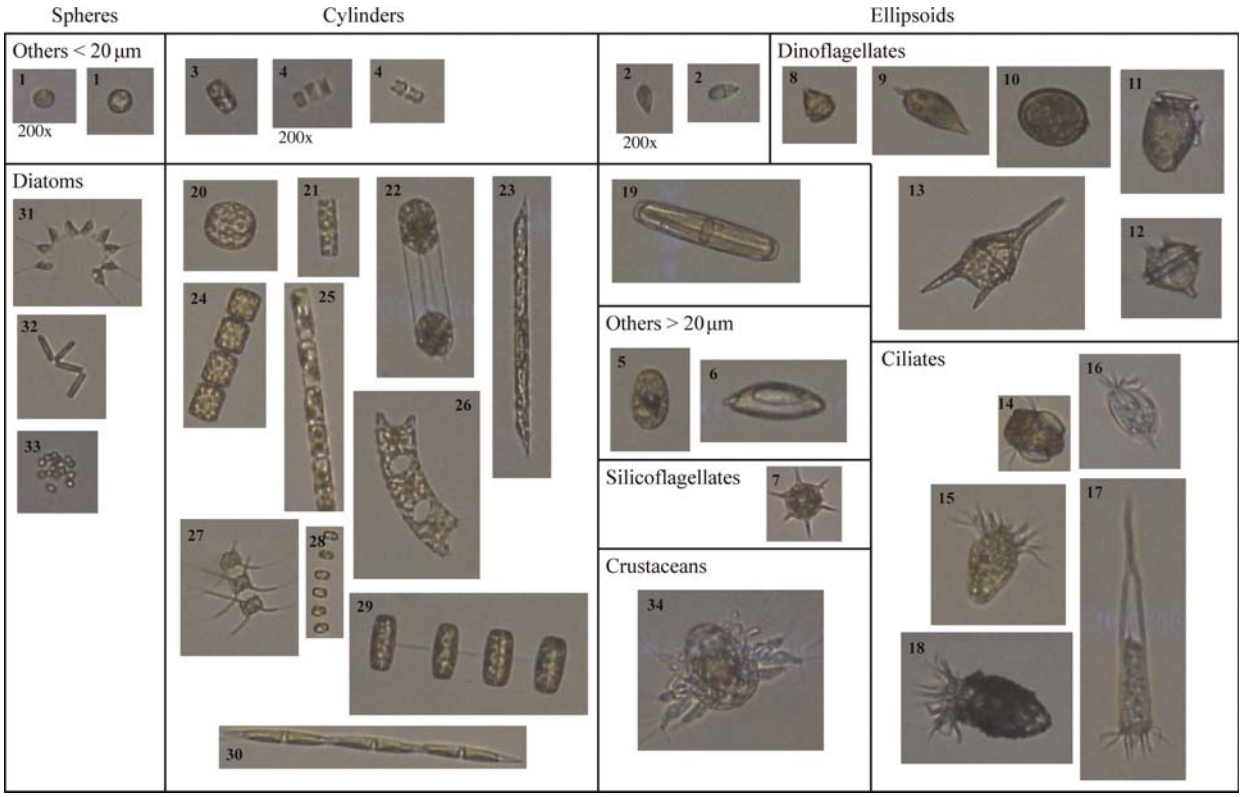

Fig. 2. Groups in the automatic classification and their correspondence in taxonomic and morphological classes. Dashed lines separate morphological classes, dotted lines taxonomic classes and solid lines both. The number in the left upper corner of the images corresponds to the description of the group given in Table I. All image magnifications are ×100 except when indicated.

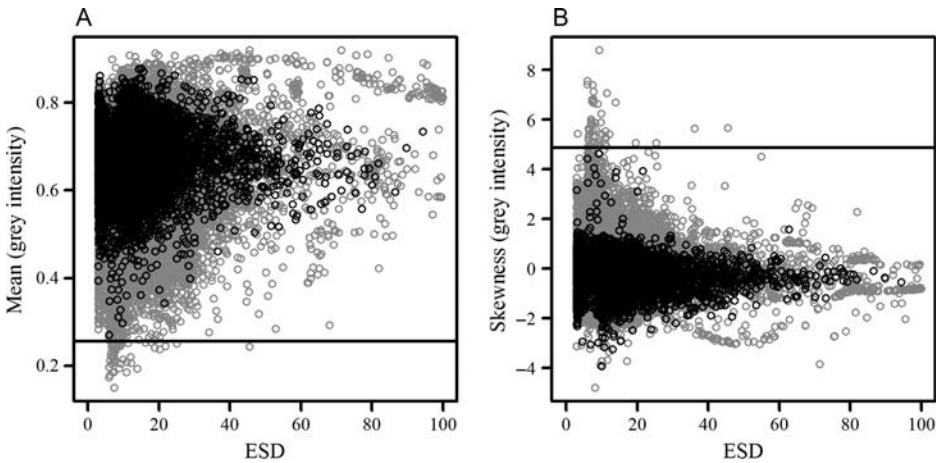

Fig. 3. Example of two attributes selected to design a statistical filter to eliminate non-living particles prior to the automatic classification. Both are texture attributes related to grey intensity that range from 0 to 1: (**A**) mean and (**B**) skewness versus the ESD of the particles. Living particles are plotted in black and detritus and artefacts in grey. Horizontal lines show the value used to discriminate living from non-living particles in the lower and the upper limit of the variable, respectively.

Each training set of images with its 136 attributes per image, scaled previously, constitutes the input for the LIBSVM. The Gaussian Radial Basis Function was used as the kernel and parameters were tuned using a grid search over parameter ranges, testing the maximization of accuracy by 5-fold cross-validation. Cross validation randomly separates the training set in several sets of images (five sets in our case); the images of each

of these sets are classified with a tool built with the remainder (four) sets. The predicted assignations are compared with the actual assignations, and the number of elements correctly classified with respect to the total number of elements gives the global accuracy of the classifier.

Predicting new samples

The three classifiers obtained (i.e. for 200x, $100x_{<20}$ and $100x_{>20}$) make up a classifier tool that can predict any FlowCAM sample described in the same way as the example images. Each image was classified in one of the 34 groups and assigned to the corresponding taxon class and shape class (Table I). The result was each particle being classified in one of the 6 taxonomic categories (diatoms, dinoflagellates, silicoflagellates, ciliates, crustaceans and other living particles) and of the 4 morphological categories (spheres, ellipsoids, cylinders and discuses) (Table I). For instance, a species of the genus *Lauderia* is included in the taxon-class diatoms but also in the shape-class cylinders, while a species of the genus *Prorocentrum* is included in the taxon-class dinoflagellates and in the shape-class ellipsoids.

Evaluation of the classifier

The most widely used technique to estimate the efficiency of the automatic classification is the cross validation, described above. The cross validation, however, has been shown to give higher efficiencies than the estimate over a test set of independent samples (Davis *et al.*, 2004; Sosik and Olson, 2007). The testing over natural samples consists in automatically classifying a set of images (i.e. the test set) that has been previously classified by an expert and are not included in the training set.

The efficiency is tested with the construction of a confusion matrix, a table where the manual classes are placed in rows and automatic classes in columns and each image counted in the cell determined by its manual and automatic assignation. The efficiency of the automatic classification was assessed by: (i) global accuracy, that is the percentage of elements correctly classified, estimated as the ratio between total elements in the diagonal of the confusion matrix (true positives) and the total number of elements, and (ii) accuracy per class, that is the ratio between the number of particles classified as belonging to a class and the actual number of particles in that class. The global accuracy can not be higher than one, whereas the accuracy per class is lower than one when the automatic classification counts fewer particles than actual ones and is higher than one when automatic classification overestimates the number of particles. It should be noted that the accuracy per class only takes into account the total number classified and does not consider whether particles are correctly classified or not. For instance, in a sample with 100 diatoms where we classify 10 diatoms correctly and 80 detrital particles as diatoms, the accuracy per class is 90% (90 of 100 particles were classified as diatoms). Two measures are used to evaluate this type of error: (i) specificity, that is the percentage of the elements that really belong to the class where they have been classified (true positives/total classified), and (ii) probability of detection, that is the percentage of elements in a class which are correctly classified (true positives/total manual). In the previous example, the specificity would have been 11.1% (10 of the 90 particles were correctly classified as diatoms), and the probability of detection 10% (10 of the 100 diatoms were classified correctly).

To estimate these parameters for the taxonomic classification, 17 samples of the test set, obtained during 2 years in a monthly sampling of a coastal station located in the Cantabrian Sea (at 46.42°N, 1.85°W, $z = 100$ m), were analysed. Only the living particles were selected and separated into six taxonomic classes (dinoflagellates, silicoflagellates, diatoms, ciliates, crustaceans and other living particles) by visual inspection, constituting a test set of 61 700 images between 3 and 100 μm. The other two natural samples of the test set, corresponding to a coastal (at 46.42°N, 1.85°W, z = 25 m) and a mid-shelf location (at 45.79°N, 3.37°W, z = 142 m) in the Bay of Biscay (October 2009), were used to estimate the parameters for the morphological classification, constituting a test set of 9200 images between 20 and 100 μm separated into four morphological classes (spheres, ellipses, cylinders and discuses).

To account for the possible variation in the accuracy of each sample (i.e. changes in taxonomic composition and variations in image quality related to lightning and focus), we have computed for each sample the abundance of living particles in the size range 3–100 μm manually and automatically classified and compared with an expected one-to-one relationship with a *t*-test.

Estimation of biovolume

Two natural samples were used to explore the effect of the different ways of calculating the biovolume of the particles. These two samples were taken in triplicate, each replicate taken using a separate Niskin bottle and analysed as described above. Only the living particles were selected, obtaining 1378, 1475 and 1470 particles in the coastal sample and 243, 170 and 205 particles in the shelf sample. Particles were separated in four shape classes (spheres, ellipses, cylinders and discuses) by visual inspection.

The biovolume of each living particle recorded by the FlowCAM was calculated in three ways: (i) the projected area of each particle was used to calculate their biovolume considering a spherical shape (i.e. area-based); (ii) the length and width were measured manually and the biovolume calculated as a revolution volume according to the shape manually assigned (i.e. shape-based); and (iii) the biovolume was calculated according to the shape predicted by the classification algorithm and the automatic measurements made by the image analysis (i.e. automatic-shape-based). Depth and width dimensions were considered equivalent since the orientation of the particles results in one or the other being measured randomly. The normalized abundance size spectra (NASS) obtained with area-based, shape-based and automatic-shape-based biovolumes were compared with a one-way ANOVA and a Tukey test. Only those bins of the spectra for which a minimum of five particles were measured were considered for the comparison.

Field evaluation

For the field evaluation set of samples ($n = 509$), the same calculations were done but only for those corresponding to automatic assignments. The total abundance of living cells was estimated and the biovolume for each living particle was calculated based on the area-based and the automatic-shape-based method. The NASS for both data sets were built and compared with a *t*-test.

The area-based and the automatic-shape-based biovolume obtained for the living particles were converted to carbon. Different conversions factors were applied for cells: smaller than 3000 μm^3, larger than 3000 μm^3 except diatoms, and diatoms larger than 3000 μm^3 (Menden-Deuer *et al.*, 2000). The abundance and biomass per taxon class were calculated and the percentage that each taxon represents in the sample was estimated. The misestimation of autotrophic biomass was calculated comparing the area-based and the shape-based biomass only for autotrophic and mixotrophic taxa.

RESULTS

Automatic classification of FlowCAM images

We used the test set to estimate the efficiency of the automated method to classify living particles. The global accuracy of the automatic processing of FlowCAM sorting living from non-living particles was 86% (Table II). The accuracy per class for both types of images is close to 1 which means that the number of particles given for each class fits the actual value. The specificity is lower because it does not consider the false positives, but it is still above 80% and the probability of detection is above the 84%.

There are no significant differences in the global accuracy of autoimage and fluorescence (87 and 85% respectively). The accuracy per class is, however, different depending on the working mode. Whereas the accuracy per class in fluorescence is close to 1 for detritus and living particles, in autoimage the misclassification of living particles is high (124%). This is caused by the large amount of detritus present in autoimage samples; even with a good classification of detritus (95%), their proportion is so high (32 000 non-living from 38 000 particles) that it causes the inclusion of many false positives in the counts of living particles.

The efficiency estimated over the confusion matrix depends on the size and composition of the test set (Solow *et al.*, 2001). Since our test set is constituted by complete samples, a very abundant sample well classified can determine the efficiency of the classification. The abundance of living particles in the size range 3–100 μm manually and automatically classified for each sample is

Table II: Confusion matrix for the automatic classification of living and non-living particles, digitized with the two FlowCAM working modes

		Automatic					
Mode	Manual	Non-living	Living	Global accuracy	Accuracy per class	Specificity	Probability of detection
Total	Non-living	38 324	5628	0.86	0.97	0.90	0.87
	Living	4408	22 525		1.05	0.80	0.84
Autoimage mode	Non-living	31 699	3701	0.87	0.95	0.94	0.90
	Living	1972	5298		1.24	0.59	0.73
Fluorescence-triggered mode	Non-living	6625	1927	0.85	1.06	0.73	0.77
	Living	2436	17 227		0.97	0.90	0.88

The parameters that account for the efficiency of the classifier are: global accuracy (true positives/total particles), accuracy per class (total classifier/total manual), specificity (true positives/total classifier) and probability of detection (true positives/total manual).

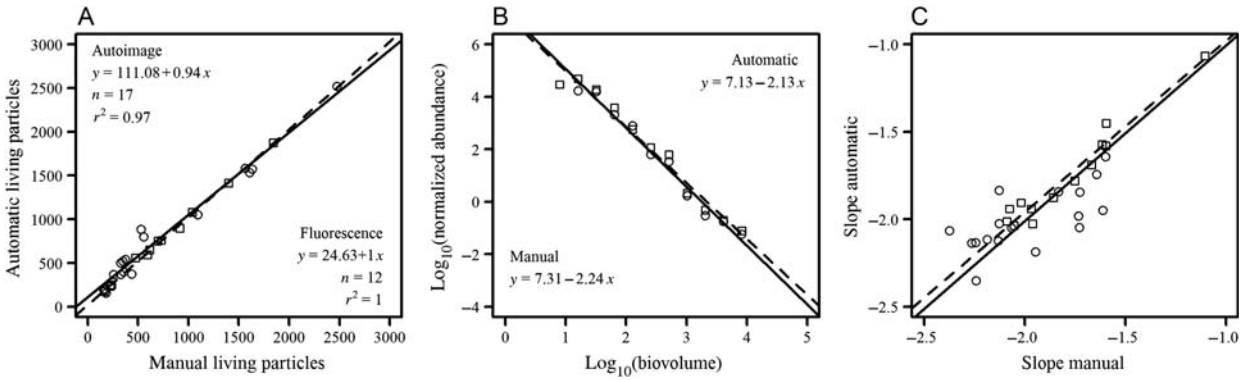

Fig. 4. Accuracy classifying living particles. (**A**) Counts of living particles manually and automatically classified in the size range 3–100 μm for each sample of the test set. Autoimage samples in dots and solid line and fluorescence-triggered samples in squares and dashed line. (**B**) NASS built with living cells manually (dots and solid line) and automatically (squares and dashed line) classified for an example sample. (**C**) Comparison of the slope of the NASS obtained with the living particles manually and automatically classified between 3 and 100 μm. Autoimage samples in dots and solid line and fluorescence samples in squares and dashed line.

shown in Fig. 4A. The relationship between the abundances is not different from 1 in autoimage ($P = 0.1868$, $n = 17$) and in fluorescence-triggered mode ($P = 0.4760$, $n = 12$). In autoimage, the r^2 is lower than in fluorescence (0.97 versus 1) given the higher error on classification explained above. A sample with 5009 living particles in manual assignment and 4668 in automatic assignment was considered an outlier and eliminated.

The NASS were built only with living particles, so the efficiency classifying living particles also determines the accuracy of the parameters of the spectra. Figure 4B shows the NASS and the parameters of the regression line fitted to the NASS, for an example sample that has been manually and automatically classified. The slope of the regression line fitted to the NASS based on manual and automatic estimates of abundance of all samples in the test set (Fig. 4C), did not differ significantly. The relationship for the whole test set of samples is not different from the expected one-to-one relationship, both for autoimage (slope automatic = 1.0056 × slope manual, $P = 0.8127$, $n = 17$) and fluorescence working modes (slope automatic = 0.9793 × slope manual, $P = 0.0564$, $n = 13$). Since the NASS include all living organisms it means that the error in the classification is smaller than the variations in abundance needed to observe a change in the parameters of the NASS.

The confusion matrix of abundance for taxon-class assignation and the parameters of efficiency are shown in Table III. The global accuracy of taxon assignation, once the detritus is eliminated, is 87%. The accuracy per class is low in classes with fewer individuals (ciliates and crustaceans) and in diatoms. The specificity is above 74% in all classes except in ciliates, a very heterogeneous class. The probability of detection is above 60% in all classes. Figure 5 shows a comparison between the counts obtained by the manual and automatic classification in each sample for the different taxonomic classes of microplankton (20–100 μm). In autoimage, the relationship was different from the expected one-to-one for ciliates, crustaceans and the "others" class, which are the taxonomic classes with fewer individuals (ciliates: $P = 0.0031$, $n = 13$; crustaceans: $P = 0.0105$, $n = 6$; others: $P < 0.0001$, $n = 15$). The relationship for diatoms, dinoflagellates and silicoflagellates was not different from the expected one-to-one (diatoms: $P = 0.2209$, $n = 14$; dinoflagellates: $P = 0.3422$, $n = 15$; silicoflagellates: $P = 0.5910$, $n = 6$). In fluorescence-triggered mode, the only relationship different from the expected one-to-one was the class "others" ($P < 0.0001$, $n = 12$).

The confusion matrix of abundance for shape-class assignation and the parameters of efficiency are shown in Table IV. The global accuracy of shape assignation, once the detritus was eliminated, is 78%, although the accuracy per class is variable. The class that is better classified is that of cylinders (104%), which is in fact the class that may cause the highest overestimate of biovolume (see below). There is high misclassification between spheres and ellipses but, since the geometrical body of a sphere is a particular case of an ellipse, the error does not greatly influence the estimate of biovolume. However, if the shape classification is used for other purposes, this must be improved. The specificity is high for cylinders and low for ellipses and spheres. The probability of detection followed the same pattern but always above the 60%.

Cell biovolume estimation

Figure 6 shows the NASS built with biovolume data calculated from area-based, shape-based and automatic-

Table III: Confusion matrix for the taxonomic classification with the parameters to account for the efficiency of the classifier

	Automatic										
Manual	Diatoms	Silicoflagellates	Dinoflagellates	Ciliates	Crustaceans	Others	Detritus	Artefacts	Accuracy per class	Specificity	Probabilityof detection
Diatoms	3590	10	35	28	1	1915	393	240	1.50	0.97	0.64
Silicoflagellates	3	244	—	2	—	74	51	4	1.18	0.89	0.76
Dinoflagellates	9	2	201	12	—	34	28	2	0.95	0.74	0.78
Ciliates	9	—	7	42	—	11	21	—	0.70	0.42	0.61
Crustaceans	1	—	—	—	3	1	7	1	1.25	0.75	0.60
Others	96	17	28	15	—	11 078	1921	419	0.86	0.84	0.99
Detritus	285	33	59	93	4	4242	27 713	889	—	—	—
Artefacts	101	—	1	4	5	28	729	6944	—	—	—

Efficiency parameters are calculated only for the living organisms groups.

shape-based methods. The NASS of the coastal station has an ordinate higher than the NASS of the shelf station given the higher concentration of particles in the coastal one. The slopes of the regression lines fitted to the NASS were different at the coastal station ($P = 0.0024$, $n = 23$), whereas they were not significantly different at the oceanic station ($P = 0.7283$, $n = 15$). The coastal station is dominated by chain-forming diatoms, and thus the proportion of cylinders relative to the other basic shapes is 75%, with an average aspect ratio (height/diameter) of 0.35. In the oceanic sample, there were a significantly lower proportion of cylinders, around 14%, with an average aspect ratio of 0.40. Given the overestimation of biovolume caused by the utilization of the area-based method, more conspicuous when the cylinders are abundant and with a large height, the regression lines were different. On the other hand, the spectra calculated with the shape-based method, with manual or automatic measures, were not different (coastal: $P = 0.6393$, $n = 21$; oceanic: $P = 0.6016$, $n = 14$), which suggests that the efficiency in shape assignment is sufficiently high to improve the biovolume estimate.

Field evaluation

The parameters of the regression line fitted to the NASS calculated with area-based and automatic-shape-based methods were different only in 13 samples (2.5%). Even when the parameters of the NASS were not different in many cases, the misestimation of biomass is higher than the 20% in 132 samples (14%). In Fig. 7A, where the concentration of carbon (μg C L^{-1}) in the sample is plotted against the concentration of carbon that corresponds to chain-forming centric or pennate diatoms, these samples are indicated with grey solid dots. It must be noted that the samples with overestimation higher than 20% are the points closer to the 1:1 relationship, that is, the samples where the biomass of chain-forming diatoms is dominant.

In Fig. 7B the percentage that the biomass of chain-forming centric or pennate diatoms represents of the total biomass in the sample against the percentage of misestimation of the autotrophic biomass is shown. The overestimate of autotrophic biomass is directly proportional to the percentage of biomass that corresponds to chain-forming diatoms. The slope of the regression line between both variables is significantly different from zero ($P < 0.0001$, $n = 882$). The percentage of chain-forming diatoms reaches 71% and their average aspect ratio (height/diameter) per sample range from 0.04 to 0.72 with an average of 0.35.

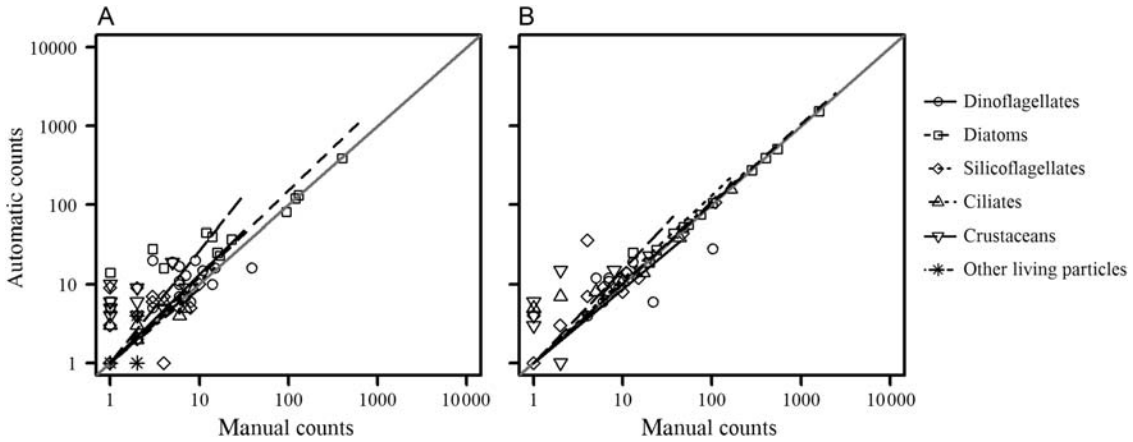

Fig. 5. Comparison of manual and automatic counts per taxonomic class (size range 20–100 μm) for the test samples analysed in (**A**) autoimage and (**B**) fluorescence-triggered modes.

Table IV: Confusion matrix for the morphological classification with the parameters to account for the efficiency of the classifier

	Automatic								
Manual	Cylinder	Discus	Ellipsoid	Sphere	Detritus	Artefacts	Accuracy per class	Specificity	Probability of detection
Cylinder	2332	12	161	25	282	593	1.04	0.96	0.92
Discus	28	208	64	1	4	64	1.27	0.88	0.69
Ellipsoid	54	8	974	571	196	89	1.20	0.72	0.61
Sphere	18	9	145	447	40	53	0.59	0.43	0.72
Detritus	241	4	256	72	1047	820	—	—	—
Artefacts	174	—	19	7	59	123	—	—	—

Efficiency parameters are calculated only for the living organisms groups.

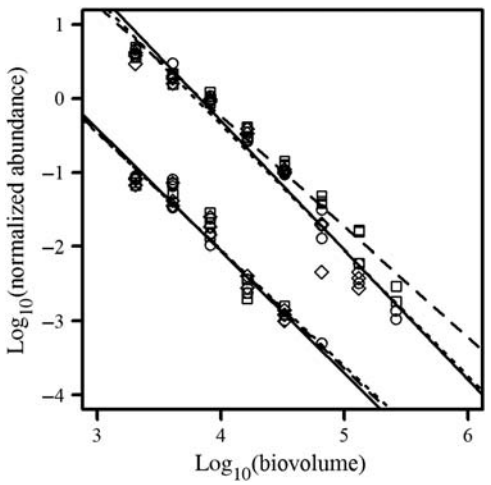

Fig. 6. NASS for FlowCAM counts of two samples, a coastal station in the upper part (i.e. higher normalized abundance) and a shelf station in the lower part of the figure. Biovolume data are calculated with area-based (squares and dashed line), shape-based (dots and solid line) and automatic shape-based methods (diamonds and dotted line).

However, the situation is not the same with the single cell centric diatoms, whose geometrical shape is expected to cause the opposite effect to chain-forming diatoms (Fig. 7C). The slope of the regression line between the underestimate of autotrophic biomass and the percentage of biomass that corresponds to single cell centric diatoms is not significantly different from zero ($P = 0.1359$, $n = 882$). The percentage of single cell centric diatoms does not exceed 35% in any sample and their aspect ratios (diameter/height) ranged from 0.29 to 0.99 with an average of 0.84.

This means that a relatively high error in biomass estimation could not be detected in the parameters of the size spectrum. However, when more that the 20% of the biomass of the sample corresponds to chain-forming centric or pennate diatoms, using the area-based method to estimate the biovolume causes an overestimate of autotrophic biomass higher than the 20%. A 20% or larger biomass corresponding to chain-forming diatoms is not unusual, very common in fact during the spring bloom at temperate and boreal locations.

DISCUSSION

The main objective of automated classification techniques is, most frequently, to estimate the abundance of

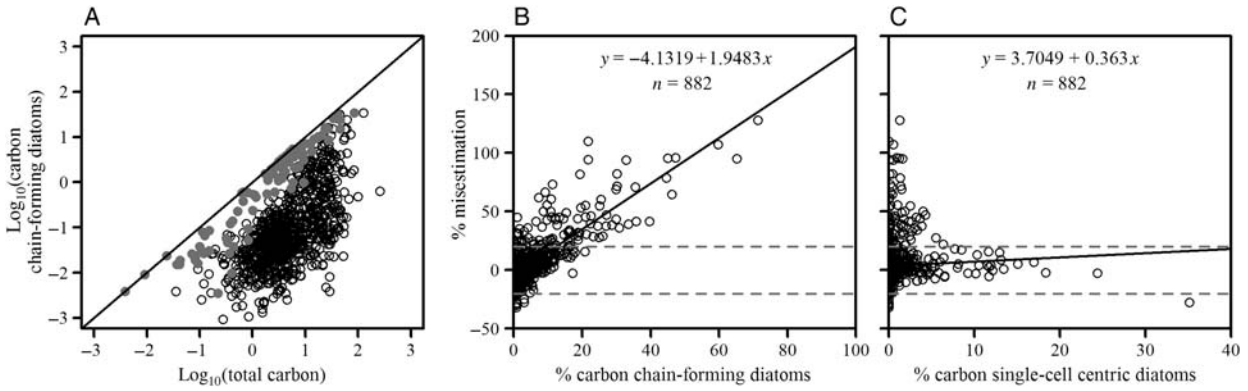

Fig. 7. (**A**) Relationship between the concentration of total carbon in the samples of the field evaluation set (μg C L^{-1}, size range 4–100 μm) and the carbon corresponding to chain-forming diatoms in the size range 20–100 μm. Samples with a misestimate of autotrophic biomass higher than 20% are marked with grey dots. Misestimation of autotrophic biomass versus the relative biomass of (**B**) chain-forming diatoms (*y* versus *x* in A) and (**C**) single-cell centric diatoms in the sample, with the lineal regression adjustment. Horizontal dotted lines show the error of 20% in the biomass estimate.

each predicted class. Sometimes, however, the research focuses on other ecosystem properties, such as biovolume, biomass or cell attributes (i.e. pigment or carbon content), for which the estimation of cell biovolume is required. The goal of the present study was to improve the estimation of cell biovolume using automated classification of FlowCAM images. With this aim we solved two challenges of the automated classification of plankton images with no additional effort using state-of-the-art automatic classification methods: (1) the need to classifying automatically and with a high specificity (i.e. correct assignation to a given class) a heterogeneous set of images, which include living and non-living particles (artefacts and detritus); and (2) the tendency to misestimate cell biovolume due to the two-dimensionality of the images. These two characteristics of the FlowCAM are common to other image-based devices for plankton identification and enumeration.

The automated classification of plankton images has improved over the last years. For instance, the methods and algorithms explored have been diverse, although in the last years SVM has been the most widely used algorithm. Other authors who have used linear discriminant analysis or random forest algorithms obtained comparable results (Irigoien *et al.*, 2009). In the first attempts at automatic classification of plankton samples the number of classified groups was below 10 while in the last years it has increased to more than 20 (Sosik and Olson, 2007). The description of images began being merely based on shape features but, later, texture features were included, increasing the accuracy of classification considerably (Tang *et al.*, 1998). The accuracy of classification has increased from below 75% to near the 90% (Gorsky *et al.*, 2010), although it depends on the digitization device.

Several aspects can be improved in automated classification of plankton samples. We have carried out a series of tests in order to improve the prediction of particle biomass by the algorithm. However, even when the results have been promising (González *et al.*, 2010), further room for improvement, may arise, according to our experience, from the refinement of the image acquisition stage. One challenge is in the digitization of high-quality images, which permits the extraction of new attributes. Some of them more class-specific, as it has been done with mesoplankton organisms, considering specific structures of each taxonomic class (Tang *et al.*, 2006). Another challenge relies on the improvement of the digitization protocols, in order to be more selective with the particles that are going to be included in the classification. When the training set used to train the classifier is large enough the accuracy of the classification is a property of the classifier and does not depend on the composition of the predicted samples (Solow *et al.*, 2001). However, even if the training set was representative and the accuracy of classification was constant, a large number of non-target images reduce the specificity and probability of detection of the classification, and hence the biovolume or biomass estimates per class are incorrect. Related to this we have shown that the fluorescence-triggered mode obtains better classification efficiencies because it is a much more selective digitization method, recording fewer detritus images. On the other hand, autoimage, a non-selective digitization method, has lower efficiencies.

Non-living particles (i.e. device-specific artefacts and detritus) form a group of recurrent images in the digitization devices that can affect the prediction of new samples, providing poor estimates of biomass of living particles. Artefacts are images recorded by the

digitization devices that do not correspond to any particle in the sample. The characteristics of these artefacts are very specific to the process of acquisition of images of each type of automatic sampling device, and thus it is not possible to develop a general tool to remove them. The ideal situation is to minimize their appearance, but, in case this cannot be achieved, we have shown how a series of filters can be designed taking into account the characteristics of the digitization mechanism. These filters allowed us to remove the artefacts with high confidence, avoiding the need to eliminate them manually, which can be a highly time-consuming process.

The presence of detritus in the samples is, however, a common feature of plankton imaging devices. The problem has been addressed with different strategies by other authors. Some have included detritus in their calculations since they consider it part of the energetics of the ecosystem (San Martin *et al.*, 2006), while others have eliminated it by visual recognition (Barofsky *et al.*, 2010), avoided the digitization of detritus using cell staining (Alcaraz, 2003) or, as we have done, tried to classify them automatically (Irigoien *et al.*, 2009; Zarauz *et al.*, 2009). The automatic method we have developed, based on statistical differences between certain image attributes of detritus and living particles, has improved the classification. We have shown that it is possible to automate the elimination of non-target images before the classification in order to obtain acceptable classification efficiencies in a non-selective digitization method, although the uncertainty in the autoimage samples is still higher than in fluorescence mode. This demonstrated that a previous step designed to eliminate artefacts and detritus is highly recommended. It can be done recording with the images a discriminative variable such as the presence of fluorescence or staining, or using a statistical filter that applied to the description of the images permits discrimination of most of the non-target images.

The second challenge addressed, the problem with the two-dimensionality of the images, is common to several sampling methods including sedimentation techniques for microscopy (Utermöhl, 1958) or continuous sampling with sample-in-flow techniques such as FlowCAM (Sieracki *et al.*, 1998). In this work the classification scheme we adopted permits not only a taxonomic but also a morphological characterization of the particles. We used the morphological information of the particles provided by the automatic classification to calculate their biovolume, a method similar to the one used by microscopists. The unique requirement is to form groups in the training set having in mind that they must contain individuals of the same taxonomic class and the same morphological class. For instance, if the desired taxonomic resolution includes a taxon class of diatoms, we can create a group with various diatom species as long as all they can be assigned univocally to a shape class. This way, the prediction of an image to belong to a group provides taxonomic and morphological information that permits the calculation of the biovolume of each particle as a revolution volume specific for its geometric shape. With this information we can implement an automatic-based method that renders comparable results to the microscopy-based method. For nanoplankton cells, smaller than 20 μm in ESD, taxonomic information is not available, since the resolution of the images does not permit the separation of taxonomic groups. Without taxonomy, information about the third dimension cannot be inferred but it has been demonstrated that in nanoplankton cells the measurement of the third dimension can be omitted since most cells can be assumed to be spheres or ellipsoids with equal depth and width (Verity *et al.*, 1992).

SVM classification algorithms have been shown to classify plankton images obtained with FlowCAM with relatively high accuracy (Blaschko *et al.*, 2005). We have applied SVM trained with our morphologically oriented training set together with the filters described above successfully to a data set of FlowCAM plankton images obtained on a seasonal basis in the Bay of Biscay area. The extent of our training set and testing and field evaluation data ensures that the approach is robust and reliable across a range of conditions. The global accuracy we have obtained in classifying morphological and taxonomic classes is comparable to that obtained in similar studies, specifically, in those focused on the classification of microplankton that tested the accuracy over independent natural samples and with a similar number of taxonomic categories considered (Sosik and Olson, 2007). The accuracy per class differs from one class to another; it is lower for the less abundant and more heterogeneous taxonomic classes, such as crustaceans and ciliates. This is a common feature in the automatic classification of plankton due to its heterogeneity and poses, in fact, a specific challenge for the automatic classification of plankton images. The accuracy per class compensates for false negatives excluded from a class with false positives included within it, provided the error is homogeneously distributed. This parameter is a good estimator of the effectiveness in the estimate of abundance per class. On the other hand the specificity evaluates the percentage of cells that really correspond to a certain class in relation to the total cells classified in that class. The specificity must be taken into account as an evaluation parameter when the aim is not only the estimate of abundance but the estimate of another variable related with size. A high specificity is desirable because the error in biovolume estimates can not be

compensated between false assignations. At class level, in the prediction of a taxon class, a high specificity is an indicator of high accuracy in the estimate of biovolume or biomass per class. At individual level, a high specificity in the prediction of shape indicates that the probability of the correct geometrical calculation of biovolume is also high, which implies that the estimate of biovolume per class can be improved.

The shape assignment has been demonstrated to be sufficiently accurate to correct the biovolume estimates based on two-dimensional images. Thus, the discrepancies between the NASS of natural samples constructed using area-based and automatic-shape-based biovolume estimates are expected to be caused by the misestimation of biovolume associated with the area-based method. Cylinders and discuses are the geometric shapes that have more different biovolumes considering area-based or shape-based biovolume. The biovolume of discuses is underestimated and the biovolume of cylinders is overestimated when area-based biovolume is considered: the higher the aspect ratio, the higher the misestimation of biovolume. Single cell centric diatoms, which usually present high aspect ratios, do not cause large underestimation of biovolume if it is calculated with the area-based method. Chain-forming centric or pennate diatoms present lower aspect ratios and can cause large overestimation of biovolume when the area-based method is applied. Thus, in samples with a large proportion of chain-forming centric or pennate diatoms the effect can be conspicuous, although determined by the aspect ratio of the chains. For instance, in a hypothetical sample composed of a 100% of chain-forming diatoms with an aspect ratio of 0.2 or 0.1, the biomass would be, respectively, duplicated or triplicated. A similar effect occurs with the Coulter counter, which tends to underestimate the biomass of cells with a shape differing strongly from spheres (Bakker *et al.*, 1985). That is important when analysing samples, because depending on where and when they are collected and their taxonomic composition, they can be susceptible to poor estimation of biomass or even poor size-structure characterization. The misestimation of autotrophic biomass is presented as an example of the error that can be made assigning biomass to a certain trophic status. This can be important in the analysis of trophic webs or when applying a modelling approach. For instance, in samples collected during the spring bloom, when the abundance of diatoms is high and they represent the main part of the microplankton community, an accurate estimate of biomass is important to estimate, for instance, the transfer of energy and/or the sedimentation of biological carbon.

However, it does not matter how good this new approximation is if it cannot be applied easily to the analysis of plankton samples. Our approach is easy to use since it adapts the state of the art techniques of automatic classification of plankton. Only two aspects must be considered, the morphological division of the training set and the calculation of biovolume depending on the shape predicted. Both tasks do not imply necessary additional work. For instance, the classification scheme applied by Sosik and Olson (Sosik and Olson, 2007) is susceptible to use this approximation, since their 22 categories are morphologically independent. Besides, the linear dimensions of the cells needed to calculate biovolume as a revolution volume are attributes extracted routinely from plankton images. In fact they have been used recently to estimate biovolume as a revolution volume (Jakobsen and Carstensen, 2011), although without shape differentiation. Gathering both aspects permits substantial improvement of biovolume estimates in a routine work, making our approach a candidate to join the available set of tools developed for automated classification of plankton images.

In summary, an improved method to estimate the biovolume of the cells coupled with an automated image acquisition system, such as FlowCAM, presents obvious advantages. The system can deal with many images providing information at scales of variability that have been inaccessible with traditional methods, but, at the same time providing more accurate biovolume estimation resulting in this information being more reliable.

SUPPLEMENTARY DATA

Supplementary data can be found online at http://plankt.oxfordjournals.org.

FUNDING

This work was supported by the Plan de Ciencia, Tecnología e Innovación del Gobierno del Principado de Asturias [project IMAGINA ("Integración de métodos de análisis de imagen de grupos planctónicos con técnicas de inteligencia artificial") and research grant BP07-081 to E.A.], the Ministerio de Ciencia e Innovación (project PERPLAN, CTM2006-04854/MAR) and the Instituto Español de Oceanografía (projects ECOPEL and RADIALES).

ACKNOWLEDGEMENTS

This is a contribution to SCOR Working Group 130 on Automatic Visual Plankton Identification. We thank the

captain and crew in the R/V Thalassa, the R/V Francisco de Paula Navarro and the R/V José Rioja for their assistance during the cruises of the projects ECOPEL ("Ecología de pequeños pelágicos"), PERPLAN ("Efecto de las perturbaciones meteorológico-hidrográficas en la estructura de comunidad planctónica", CTM2006-04854/MAR) and RADIALES ("Programa de series temporales oceanográficas") that provided the data set for the present work. We are indebted to all participants in the cruises for their work at sea and in the laboratory.

REFERENCES

Alcaraz, M. (2003) Estimating zooplankton biomass through image analysis. *Mar. Biol.*, **143**, 307–315.

Álvarez, E., López-Urrutia, A., Nogueira, E. *et al.* (2011) How to effectively sample the plankton size spectrum? A case study using FlowCAM. *J. Plankton Res.*, **33**, 1119–1133.

Bakker, C., Prins, T. C. and Tackx, M. L. M. (1985) Interpretation of particle spectra of electronic counters by microscopical methods. *Hydrobiol. Bull.*, **19**, 49–59.

Barofsky, A., Simonelli, P., Vidoudez, C. *et al.* (2010) Growth phase of the diatom *Skeletonema marinoi* influences the metabolic profile of the cells and the selective feeding of the copepod *Calanus* spp. *J. Plankton Res.*, **32**, 263–272.

Benfield, M. C., Grosjean, P., Culverhouse, P. *et al.* (2007) RAPID: research on automated plankton identification. *Oceanography*, **20**, 12–26.

Blaschko, M. B., Holness, G., Mattar, M. A. *et al.* (2005) Automatic *in situ* identification of plankton. *Proceedings of the Seventh IEEE Workshops on Application of Computer Vision (WACV/MOTION'05)*, **1**, 79–86.

Cermeño, P. and Figueiras, F. G. (2008) Species richness and cell-size distribution: size structure of phytoplankton communities. *Mar. Ecol. Prog Ser.*, **357**, 79–85.

Chang, C.-C. and Lin, C.-J. (2009) LIBSVM: a Library for Support Vector Machines. http://www.csie.ntu.edu.tw/~cjlin/libsvm/.

Davis, C. S., Hu, Q., Gallager, S. M. *et al.* (2004) Real-time observation of taxa-specific plankton distributions: an optical sampling method. *Mar. Ecol. Prog Ser.*, **284**, 77–96.

Dodson, A. N. and Thomas, W. H. (1978) Reverse filtration. In Sournia, A. (ed.), *Phytoplankton Manual*. UNESCO, Paris, pp. 104–107.

González, P., Díez, J., Del Coz, J. J. *et al.* (2010) Predicción taxonómica de muestras de microplancton usando técnicas de Aprendizaje Automático. In Troncoso, A. and Riquelme, J. C. (eds), *Actas del V Simposio de Teoría y Aplicaciones de Minería de Datos (TAMIDA 2010)*. Ibergaceta Publicaciones, Madrid, pp. 319–328.

Gorsky, G., Ohman, M. D., Picheral, M. *et al.* (2010) Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.*, **32**, 285–303.

Grosjean, P. (2005) Analyze your plankton through digitized images. http://www.sciviews.org/zooimage/index.html.

Grosjean, P., Picheral, M., Warembourg, C. *et al.* (2004) Enumeration, measurement and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES J. Mar. Sci.*, **61**, 518–525.

Haralick, R. M., Shanmugam, K. and Deinstein, I. (1979) Textural features for image classification. *IEEE Trans. Syst. Man Cyber.*, **6**, 610–621.

Hillebrand, H., Dürselen, C.-D., Kirschtel, D. *et al.* (1999) Biovolume calculation for pelagic and benthic microalgae. *J. Phycol.*, **35**, 403–4424.

Hu, M. K. (1962) Visual pattern recognition by moment invariants. *IRE Trans. Info. Theory*, **8**, 179–187.

Irigoien, X., Fernandes, J. A., Grosjean, P. *et al.* (2009) Spring zooplankton distribution in the Bay of Biscay from 1998 to 2006 in relation with anchovy recruitment. *J. Plankton Res.*, **31**, 1–17.

Jakobsen, H. H. and Carstensen, J. (2011) FlowCAM: sizing cells and understanding the impact of size distributions on biovolume of planktonic community structure. *Aquat. Microb. Ecol.*, **65**, 75–87.

Karatzoglou, A., Meyer, D. and Hornik, K. (2006) Support Vector Machines in R. *J. Stat. Softw.*, **15**, 1–28.

Luo, T., Kramer, K., Goldgof, D. B. *et al.* (2004) Recognizing plankton images from the shadow image particle profiling evaluation recorder. *IEEE Trans. Syst. Man Cyber. B*, **34**, 1753–1762.

Menden-Deuer, S., Lessard, E. J. and Satterberg, J. (2000) Carbon to volume relationships for dinoflagellates, diatoms and other protist plankton. *Limnol. Oceanogr.*, **45**, 569–579.

Pau, G., Fuchs, F., Sklyar, O. *et al.* (2010) EBImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics*, **26**, 979–981.

Peters, A. and Hothorn, T. (2009) ipred: Improved Predictors. R Package Version 0.8-8. http://CRAN.R-project.org/package=ipred.

Quiñones, R. A., Platt, T. and Rodríguez, J. (2003) Patterns of biomass-size spectra from oligotrophic waters of the Northwest Atlantic. *Prog Oceanogr.*, **57**, 405–427.

Rodriguez, J. and Mullin, M. M. (1986) Relation between biomass and body weight of plankton in a steady state oceanic ecosystem. *Limnol. Oceanogr.*, **31**, 361–370.

San Martin, E., Irigoien, X., Harris, R. P. *et al.* (2006) Variation in the transfer of energy in marine plankton along a productivity gradient in the Atlantic Ocean. *Limnol. Oceanogr.*, **51**, 2084–2091.

Sheldon, R. W. and Parsons, T. R. (1967) *A Practical Manual on the Use of the Coulter Counter in Marine Science*. Coulter Electronics, Toronto, pp. 66.

Sheldon, R. W., Prakash, A. and Sutcliffe, W. H. Jr (1972) The size distribution of particles in the ocean. *Limnol. Oceanogr.*, **17**, 327–340.

Sieracki, C. K., Sieracki, M. E. and Yentsch, C. S. (1998) An imaging-in-flow system for automated analysis of marine microplankton. *Mar. Ecol. Prog Ser.*, **168**, 285–296.

Sklyar, O. and Huber, W. (2006) Image analysis for microscopy screens—image analysis and processing with EBImage. *The Newsletter of the R Project*, **6**, 12–16.

Solow, A., Davis, C. and Hu, Q. (2001) Estimating the taxonomic composition of a sample when individuals are classified with error. *Mar. Ecol. Prog Ser.*, **216**, 309–311.

Sosik, H. M. and Olson, R. J. (2007) Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Meth*, **5**, 204–216.

Tang, X., Lin, F., Samson, S. *et al.* (2006) Binary plankton image classification. *IEEE J. Oceanic Eng*, **31**, 728–735.

Tang, X., Stewart, W. K., Vincent, L. *et al.* (1998) Automatic plankton image recognition. *Artif. Intell. Rev.*, **12**, 177–199.

Utermöhl, H. (1958) Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. *Mitt. Int. Ver. Theor. Angew. Limnol.*, **9**, 1–38.

Zarauz, L., Irigoien, X. and Fernandes, J. A. (2009) Changes in plankton size structure and composition, during the generation of a phytoplankton bloom, in the central Cantabrian Sea. *J. Plankton Res.*, **31**, 193–207.

Zarauz, L., Irigoien, X., Urtizberea, A. *et al.* (2007) Mapping plankton distribution in the Bay of Biscay during three consecutive spring surveys. *Mar. Ecol. Prog. Ser.*, **345**, 27–39.

Zernike, F. (1934) Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode (Diffraction theory of the cut procedure and its improved form, the phase contrast method). *Physica*, **1**, 689–704.