LIMNOLOGY
and
OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 14, 2016, 305–314

doi: 10.1002/lom3.10090

FlowCAM optimization: Attaining good quality images for higher taxonomic classification resolution of natural phytoplankton samples

Marianne G. Camoying,*[1,2] Aletta T. Yñiguez[1]
[1]Marine Science Institute, University of the Philippines, Diliman, Quezon City, Philippines
[2]Institute of Environmental Science and Meteorology, University of the Philippines, Diliman, Quezon City, Philippines

Abstract

Taxonomic data on phytoplankton composition is important for ecological studies, however, such information is not easy to gather. Imaging devices and image classification software have been developed in the past decades for rapid phytoplankton assessment. Taxonomic resolution output of classification software are primarily limited by the quality of images produced by these imaging instruments. FlowCAM has been utilized in several studies for this endeavor. However, the phytoplankton categories that the instrument is currently able to discriminate are still few compared to the outputs of microscopy. This study aimed to produce high resolution FlowCAM images of fixed phytoplankton samples from natural environments without compromising sample analysis time. It was also aimed to optimize the capability of FlowCAM's *VisualSpreadsheet* software to automatically classify phytoplankton images. The use of FOV300 flow cell and 10X objective combination has proven to be effective in producing good quality images at a faster rate. The modified hardware configuration resulted to FlowCAM counts that were comparable to that of the standard microscopy method. FlowCAM was able to automatically classify images of dominant phytoplankton groups in the two key sardine fishery areas in the Philippines with relatively high accuracy values. These phytoplankton groups are represented by genera with complex morphological structures (e.g., setae) such as *Chaetoceros* and *Bacteriastrum* as well as those genera with simple shapes such as *Pseudo-nitzschia* (thin-elongate) and *Coscinodiscus* (spherical).

Phytoplankton are microscopic plants that thrive in bodies of water. They form the base of the food web and consequently, affect the distribution and abundance of organisms at higher trophic levels such as fishes. They are indicative of the overall status of aquatic ecosystems as they have species-specific growth responses to changes in environmental conditions. Some phytoplankton species also produce toxins and form harmful algal blooms (HABs), which could endanger higher consumers such as humans who eat toxin-infected shellfish. Characterization of the taxonomic composition of phytoplankton is indeed, critical for ecological studies in aquatic environments.

The most accurate and common method in studying phytoplankton is through the conventional microscopy method which, however, is laborious and time-consuming. This procedure limits the sampling frequency resolution of data, and may not be effective for long-term phytoplankton monitoring studies wherein a large number of samples are involved. Alternatively, field instruments could be used for rapid assessment of phytoplankton such as spectrofluorometers and in situ flow cytometers. However, such devices are limited only to providing bulk chlorophyll *a* measurements and the size structure of phytoplankton, respectively.

Imaging devices, both for laboratory and field use, have been developed in the past decades to provide phytoplankton data of higher taxonomic resolution at a faster rate (Benfield et al. 2007). Examples of these instruments are the FlowCytobot and FlowCAM. The submersible Imaging FlowCytobot developed by Sosik and Olson (2007a) has shown to be effective in continuously capturing images of coastal phytoplankton in the 10–100 um size range for a period of 6 months without interruption. Conversely, FlowCAM (Sieracki et al. 1998) has already been used in several studies for analyzing fixed and fresh phytoplankton samples both in the laboratory or onboard ships.

FlowCAM combines the capabilities of a flow cytometer, camera and a microscope (Alvarez et al. 2011). It automatically captures images of particles in a moving fluid. It is already equipped with a software (*VisualSpreadsheet*) that provides size measurements and other particle properties of the image which can be exported into a spreadsheet. Images

This article was published online 18 February 2016. A footnote has been added to denote the new insertion of acknowledgment section. This notice is included in the online and print versions to indicate that both have been corrected.
*Correspondence: mg.camoying@gmail.com

can be automatically classified into groups by building a statistical filter based on the particle properties of existing image libraries.

FlowCAM has a diverse application. In the field of phytoplankton research, several studies have already validated the sizing and abundance accuracies of FlowCAM utilizing the recommended instrument factory hardware and software settings (e.g., camera settings, sample flow rate, etc.). Alvarez et al. (2011, 2012, 2014) has shown that FlowCAM outputs are comparable to that of the conventional microscopy method and that the instrument is able to provide a synoptic view of natural phytoplankton assemblages.

For the past years, several improvements have been made on FlowCAM's software and hardware configurations leading to faster FlowCAM sample analysis rate. One example is the replacement of the peristaltic pump with a syringe pump for more efficient manipulation of sample flow rate. Another very significant improvement is the development of specialized FOV flow cells. Flow cells are an important component of FlowCAM as the sample being analyzed is contained in and passes through the flow cell as they are being imaged by the camera. Compared to the standard flow cell, the FOV could be described as a reinforced standard flow cell, contained in a quartz-glass prism, making it more sturdy, easier to maintain and allows imaging of the whole sample volume being run through the system. To date, two FOV flow cells are available: FOV80 and FOV300. FOV80 flow cell paired with 10X objective has a particle resolution of 20–80 um particles while FOV300 paired with 4X objective is designed for analyzing particles less than 300 um in size.

Previous phytoplankton studies utilizing FlowCAM have focused on using this instrument for obtaining overall biomass only or characterizing phytoplankton into size bins (e.g., Liu et al. 2005; Zarauz et al. 2009; Reul et al. 2014) or general groups such as diatoms, dinoflagellates, silicoflagellates, etc. (e.g., Ide et al. 2008; Alvarez et al. 2012; Romagnan et al. 2015) or in identifying a single species (Buskey and Hyatt 2006; Siswanto et al. 2013). Most of these studies have not looked into more systematically determining FlowCAM settings that can enhance phytoplankton community characterization possibly because they have been focused on obtaining other types of data and/or have just done their own tweaking of the system to obtain satisfactory data for their particular study questions. None of these studies have also shared the settings (e.g., filters used) to effectively automatically classify phytoplankton groups or they have not used FlowCAM's auto-classification at all.

Alvarez et al. (2014) has emphasized that the quantity of phytoplankton categories distinguished by FlowCAM is still very few (usually up to genus level or higher only) compared to that of the conventional microscopy method wherein identification to species level is attained. As having good quality images is critical in attaining higher taxonomic classification resolution of phytoplankton, it was the main goal of this study to optimize the FlowCAM and develop a modified sample analysis protocol that would increase the taxonomic classification resolution of the instrument without compromising sample processing and analysis time. This study is divided into two tiers: (1) the optimization of FlowCAM hardware and software settings, particularly the evaluation of the performance of the two FOV flow cells (FOV80 and FOV300) in providing high-quality images of natural phytoplankton samples in the 20–300 um size range; and (2) the development of the FlowCAM image auto-classification capability through the build-up of phytoplankton image libraries. In the first part of the study, the accuracy of the counts derived from the optimized FlowCAM hardware and software settings were evaluated by comparing it to the output of the conventional microscopy method. For the optimization of the auto-classification capability of the FlowCAM, *statistical filters* (based on the morphological properties of particles) were created for each phytoplankton image library. Accuracy tests were done to determine the most effective *statistical filter* for each phytoplankton taxon/group. To our knowledge, this is the first study to have explored in detail the efficiency of the FlowCAM's *VisualSpreadsheet* software in auto-classifying images of fixed phytoplankton samples from natural environments. Previous FlowCAM studies have utilized different image classification algorithms in classifying phytoplankton images (e.g., Alvarez et al. 2011, 2012).

Method

Field sampling

Field sampling was conducted in two major sardine fishery areas in the Philippines—Dipolog-Sindangan Bay and Butuan Bay (Fig. 1) during the SarDyn Project Research Cruise onboard M/V DA-BFAR on February 20–26, 2013. Phytoplankton samples were collected twice per station using a 20-um net deployed at a maximum depth of 50 m and were concentrated onboard with a 20-um sieve and fixed with 5% neutral formalin for later analysis in the laboratory. A total of 27 stations were sampled capturing the low to high salinity gradient in Butuan Bay and the weak to strong upwelling region in Dipolog-Sindangan Bay. Environmental parameters such as Chl *a* concentration, salinity and temperature were also measured using a conductivity-temperature-depth (CTD) profiler (Seabird 25).

Optimization of FlowCAM hardware and software settings: Acquiring good quality images

For the optimization of FlowCAM software and hardware settings, particularly the determination of optimum flow cell and objective combination, phytoplankton samples from six stations of highest Chl *a* concentrations (Fig. 1), based from CTD data, were utilized. Formalin-fixed phytoplankton samples were carefully rinsed first with particle-free seawater before sample analysis. For the general FlowCAM analysis, sieves of different pore sizes were used in the serial filtration of samples into target size-classes. Samples were analyzed in

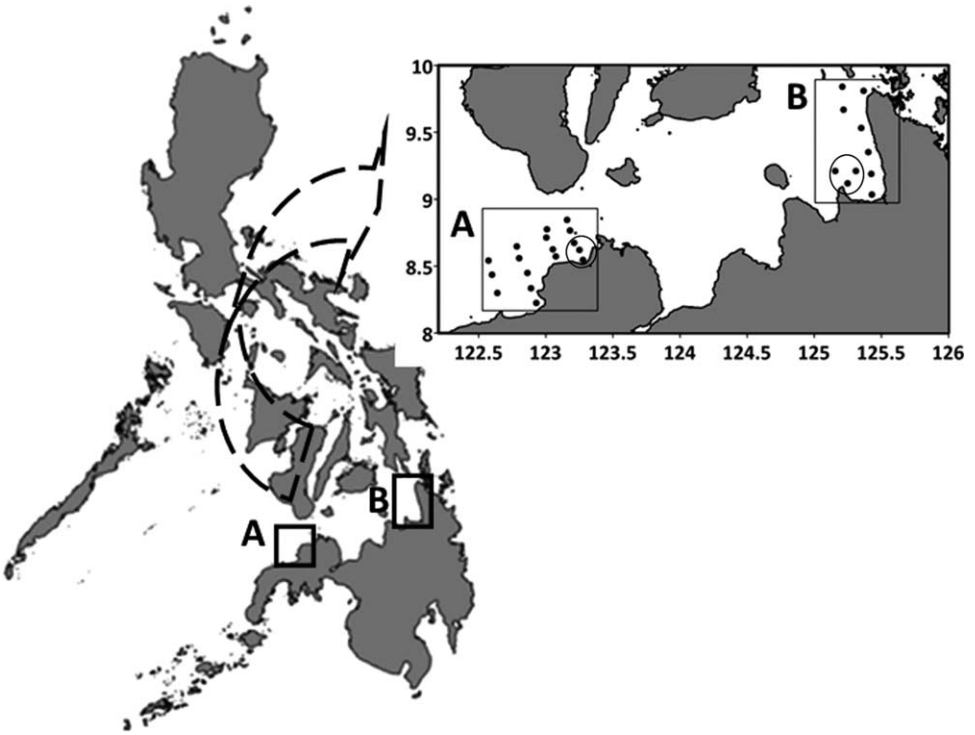

Fig. 1. Map of sampling sites. (a) Sindangan-Dipolog Bay, a coastal upwelling area; (b) Butuan Bay, an estuary receiving large nutrient inputs from Agusan River. Points encircled were stations of highest Chl *a* concentration and were used in the optimization of FlowCAM hardware and software settings.

Autoimage Mode wherein particles are imaged in a user-defined interval or in frames per second (FPS). The criteria for selecting the optimum flow cell and objective combination were the following: (1) production of good-quality images which are identifiable to at least genus level; (2) faster sample preparation and analysis time; and the (3) reduction of clogging issues.

To evaluate the counting accuracy of the determined optimum software and hardware settings (flow cell and objective combination), FlowCAM phytoplankton counts were compared to that of the microscopy method. A total of 15 phytoplankton net samples from representative stations in Dipolog (five stations), Sindangan Bay (five stations) and Butuan Bay (five stations) of variable densities were utilized. Two replicates of each sample were run in the FlowCAM and full counts of 1-mL aliquots (two replicates) were conducted under an inverted microscope (Carl Zeiss Axiovert 25) using a Sedgewick counting chamber.

Optimization of FlowCAM's *Auto-Classification* functionality: Build-up of phytoplankton library of images

For the build-up of phytoplankton image libraries, a total of 27 phytoplankton samples were run through the FlowCAM utilizing the optimized hardware and software settings. Each FlowCAM sample run generates a *List File* which contains images of target particles. Before each FlowCAM sample run, filters could be set such that images that are being captured by the camera would be limited only to those particles that fall within the target size-class. In this study, a filter was set to exclude artefacts (e.g., dirt on the flow cell wall, contaminants, etc.) with an equivalent spherical diameter (ESD) of less than 20-um to be imaged by the camera, minimizing the capture of non-target particles. Output images were then manually classified into taxonomic groups. Taxonomic references of Yamaji (1982), Tomas (1997), and Omura et al. (2012) were used in the identification of phytoplankton. Phytoplankton species captured at different orientations (valve view or girdle view) or imaged as individuals, chains or colonies, were further classified into separate subcategories (e.g., *Coscinodiscus* in valve view, *Coscinodiscus* in girdle view, etc).

Optimization of FlowCAM's *Auto-Classification* functionality: Accuracy tests of statistical filters for the different phytoplankton groups

To optimize the auto-classification functionality of FlowCAM, "*statistical filters*" were built for each phytoplankton library and were assessed for accuracy in filtering images of the target group. S*tatistical filters* are based on the morphological properties of the images that compose each phytoplankton library. These particle properties are summarized in

Table 1. Description of particle properties utilized in the set-up of *statistical filters* for each phytoplankton library.

Particle property	Description
ABD (Area Based Diameter)	The diameter based on a circle with an area that is equal to the ABD area. ABD area is the combined area of all pixels that are deemed part of the particle so the ABD Diameter is the diameter of the circle obtained by arranging these pixels in a solid circle.
ESD	The Mean value of 36 feret measurements
Length (L)	The maximum value of 36 feret measurements.
Width (W)	The minimum value of 36 feret measurements
Perimeter (P)	Total length of the edges making up a particle including the edges of any holes.
Aspect Ratio (AR)	Width/Length
Elongation (E)	A length/breadth ratio based on Perimeter and Area with the assumption that Area = length × breadth and Perimeter = 2 (length + breadth)
Circle fit (CF)	Deviation of the particle from a best-fit circle
Compactness (C)	A shape parameter derived from the perimeter and the area.
Edge Gradient (EG)	Average intensity of the pixels making up the outside border of a particle after a Sobel Edge Detect convolution filter has been applied to the raw camera image.
Roughness (R)	A measure of the unevenness or irregularity of a particle's surface—the ratio of perimeter to convex perimeter

Table 1. The accuracy of each statistical filter was assessed by classifying a set of images of mixed phytoplankton taxa, the *test set*, which refers to FlowCAM output file (*List File)* of known, manually pre-identified phytoplankton images. Thus, the *test set* consists of other phytoplankton taxa apart from the target taxa. The efficiency of each particle property statistical filter was calculated using the formula:

$$\text{Filter Accuracy (FA)} = \frac{\text{True Positive (TP)} + \text{True Negative (TN)}}{\text{False Positive (FP)} + \text{False Negative} + \text{True Positive (TP)} + \text{True Negative (TN)}}$$

Target images from the test set that were correctly classified were referred to as true positive (TP) and those nontarget images that remained unclassified were counted as true negative (TN). False positive (FP) were nontarget images that were classified while false negative (FN) were target images that were not classified by the *VisualSpreadsheet*.

Each phytoplankton library is characterized by different particle property values as phytoplankton vary in size, shape and other image attributes. The main objective here was to explore these particle properties and determine what particle property combinations best characterize each phytoplankton class. Accuracy tests were done by first (1) evaluating the efficiency of each individual particle property (Table 1) as filter; (2) simultaneously using all the twelve properties as filter and lastly; (3) combining into pairs the top three particle properties of highest accuracy.

Results

Optimum FlowCAM hardware and software settings

Several *Trial Sets* (TS) of different flow cell and objective magnification combinations were tested to assess the best setting that would give high-quality images (Table 2). In this study, the recommended factory settings have been explored first such as the use of FOV80 flow cell (FC) with 10X objective and FOV300FC with 4X objective to analyze particles in the 20–80 um size class and > 80–300 um size class, respectively (TS-1) (Table 2). A number of challenges were encountered during preliminary FlowCAM runs such as clogging of the flow cell chamber, clumping and suspension of particles, and the output of blurry and partly-imaged particles. Clogging issues were "highly" experienced in FOV80 (Table 2) while the output of unidentifiable images was common in FOV300 at 4X combination. Overall, the recommended flow cell and objective combinations (TS-1) had the lowest percentage production of good-quality images (< 20%) and was relatively time-consuming taking more than 100 min to process and analyze one sample (two replicates).

To address the issues (e.g., low-quality images) encountered with the use of the recommended factory settings (TS-1), other flow cell and objective magnification combinations were explored (Table 2). For the second trial (TS-2), the sample was divided into three size classes: (1) 20–60 um (FOV80 at 10X); (2) > 60–150 um (FOV300 at 10X); and (3) > 150–300 um (FOV300 at 10X). The production of high-quality images increased (< 50%) and clogging issues were reduced. However, TS-2 still proved to be impractical as the duration of sample preparation and analysis time increased to greater

than 120 min. For TS-3, samples were separated into two size classes: (1) 20–100 um (analyzed through FOV300 at 10X); and (2) > 100–300 um (with FOV300 at 4X). Clogging issues were minimized with the use of bigger flow cell (FOV300) and sample processing time was faster (> 80 min) compared to TS-2. Similar to TS-2, the production of good quality images was low due to the use of the 4X objective. In the last trial set (TS-4), the whole 20–300 um size class was analyzed using the FOV300 and 10X objective. Among the four trial sets, TS-4 proved to be the most efficient method as sample preparation and analysis time was significantly reduced to less than 60 min. Highest production of good-quality images (> 60%) was achieved in TS-4 (Table 2).

However, the disadvantage of using the FOV300 and 10X combination is that similar to the use of standard flow cells (not FOV), only a portion of the sample fluid is being imaged by the camera. This approach can result in underestimates in FlowCAM particle concentration calculations. To assess whether TS-4 would give accurate counts of natural phytoplankton samples, FlowCAM outputs were compared to that of the conventional microscopy method. The software parameter values utilized were within the recommended software settings stated in the FlowCAM Manual (Table 3). The actual volume sampled by the FlowCAM is computed as the (1) number of images taken during each run multiplied to the (2) volume per picture (VPP). The number of images taken is simply the length of analysis times the frame rate per second (FPS) value which is set before each sample run, while VPP is equal to the area of the field of view of the camera (Width × Height) multiplied to the depth of the flow cell in microns (FOV300 = 300 um depth). Following the said formula, running 1 mL sample volume at a flow rate (FR) of 0.1 mL/min and at 10 FPS, a total of 0.66 mL volume or 66% of the sample is imaged by the FlowCAM.

FlowCAM counts were done by manually classifying output images into lowest possible identifiable taxa through the "*Classification Menu*" of the *VisualSpreadsheet* software. Each classification file containing the list of all identified images was exported into an Excel spreadsheet. The default FlowCAM setting is to consider each image as one particle count. However, this may not always be the case as an image could contain a phytoplankton chain or several individuals of different species. Hence, images were manually checked for the presence of chains or several particles in one image and counts were adjusted accordingly.

For the 15 sets of phytoplankton samples analyzed, abundance counts (individuals/mL) were generally lower for FlowCAM compared to microscopy counts. Nevertheless, FlowCAM was able to capture the abundance trends similar to the microscopy method ($p < 0.001$, $n = 15$, $r^2 = 0.96$) (Fig. 2). FlowCAM counts for less dense phytoplankton samples of concentrations < 200 ind/mL were also comparable to that of the microscopy method.

Build-up of phytoplankton image libraries

Image libraries are important in the optimization of the auto-classification of capability of FlowCAM. Statistical filters could be built to sort out sample images that have similar characteristics to the images that comprise each library. It is recommended that when creating filters, a library should contain only 10–15 representative images to avoid introduction of error.

Output images of the FOV300 flow cell and 10X objective combination runs were manually classified into phytoplankton classes/groups through the "*Library Menu*" of the *VisualSpreadsheet*. A total of 49 phytoplankton libraries, which consist of 34 diatoms, 15 dinoflagellates, and 1 cyanobacteria species, are currently available (Table 4). As particles are being imaged in a moving fluid, phytoplankton species can be captured in different orientations. For example, centric diatoms such as *Coscinodiscus, Thalassiosira* and *Planktoniella sol* could be imaged in their valve or girdle view and some other species could also be captured as individuals or in chains and colonies (e.g., *Thalassionema, Pseudo-nitzschia, Chaetoceros, Bacteriastrum*). Hence, phytoplankton groups were further categorized into sub-groups according to the orientation of the species as captured by the camera.

The most represented groups having at least 10 high-resolution (uniform orientation), identifiable images were the diatoms *Chaetoceros*, *Bacteriastrum*, *Planktoniella sol*, *Rhizosolenia*, *Pseudo-nitzschia*, and *Thalassionema*, to name a few (Table 4). Conversely, the dinoflagellate groups such as *Ceratium* and *Protoperidinium*, being the less dominant phytoplankton in the study sites were comprised of only a few images (< 10 images). It is worth noting, however, how the FlowCAM was able to capture good images of the small-sized dinoflagellates that are less than 60-um in diameter such as *Alexandrium*, *Dinophysis acuminata, Prorocentrum cf compressum*, and some species of *Protoperidinium*).

Accuracy tests of particle property filters of dominant phytoplankton groups

The auto-classification capability of the FlowCAM was assessed by conducting accuracy tests on the efficiency of different statistical filters built for each phytoplankton library. Note however, that accuracy tests were conducted only on the seven dominant phytoplankton species encountered in the study sites.

Highest accuracy values (Filter Accuracy (FA) > 0.90) were computed for the genera with simple morphological shape such as *Coscinodiscus* (circular) and *Pseudo-nitzschia* (thin-elongate) (Fig. 3). Circle fit (CF) alone was shown to be the most effective filter in sorting out the centric diatom *Coscinodiscus* while the combination of width (W) and roughness (R) was effective for *Pseudo-nitzschia*. For those phytoplankton genera having complex structures such *Chaetoceros* and *Bacteriastrum*, lower accuracy values (FA < 0.85) were observed. The chain-forming pennate diatom

Table 2. Trial Sets (TS) for determining the optimum flow cell and objective combination.

			Criteria for the selection of optimum flow cell and objective combination		
Trial set (TS)	**Flow cell and objective combination**	**Target size class**	***Faster sample preparation and analysis time (overall duration)**	**†Reduction of flow cell clogging issues (frequency of clogging occurrence)**	**‡Production of good-quality images (percentage of good-quality images in 1 FlowCAM sample run)**
Set 1	FOV80 at 10X	20–80 um	>100 min	High clogging	<20%
	FOV300 at 4X	>80–300 um			
Set 2	FOV300 at 10X	20–60 um	>120 min	Low clogging	<50%
	FOV300 at 10X	>60–150 um			
	FOV300 at 4X	>150–300 um			
Set 3	FOV300 at 10X	20–100 um	>80 min	Moderate clogging	<50%
	FOV300 at 4X	>100–300 um			
Set 4	FOV 300 at 10X	20–300 um	<60 min	Moderate clogging	>60%

***Sample preparation time** includes the rinsing and serial sieving of samples into target size-classes, changing of flow cell and objective combination and refocusing of the camera. **Sample analysis time** refers to the overall duration of running two replicates of each (target-size class) sample.
†**High occurrence** means that flow cell clogging always happened which usually resulted to the termination of a sample run. **Moderate occurrence** means that clumping of particles occurred during sample run yet did not lead to the termination of a run. In most cases, pinching of the flow cell was able to solve the problem. **Low occurrence** means that a sample run is usually completed without interruption. Such was the case in TS-2 wherein low occurrence of clogging issues were observed which could be attributed to the low particle concentration as the sample was separated into three size-classes.
‡**Percentage of good quality images** produced was computed as the number of high resolution images (identifiable either as plankton, debris or any artefact on flowcell walls) divided by the total images (including blurry and unidentifiable images) captured in the FlowCAM output *List File*.

Thalassionema, also exhibited low accuracy. It is interesting to note that although *Chaetoceros* and *Bacteriastrum* have common structures such as the presence of external valvular outgrowths, they did not share similar filter accuracy results. For *Chaetoceros*, perimeter (P) and compactness (C) combination gave the highest accuracy while the combination of all particle properties (All) gave the lowest value. The opposite was true for *Bacteriastrum* as well as for *Meuniera membranacea* and *Rhizosolenia* wherein the simultaneous use of the 12 particle properties was the most effective filter (Fig. 3).

Discussion

Optimization of FlowCAM hardware and software settings

FlowCAM has already been used in several studies to characterize phytoplankton community structure in natural environments (e.g., Alvarez et al. 2012; Stauffer et al. 2014). However, due to the low resolution of FlowCAM images, classification of phytoplankton was usually limited only to general taxonomic groups. As data on the taxonomic composition of phytoplankton is an important indicator of the overall state of aquatic ecosystems, it was the main objective of the present study to optimize the FlowCAM to produce good quality images that would allow lower taxonomic level (e.g., species level) classification of preserved natural phytoplankton samples without comprising sample analysis time. The use of FOV300 and 10X objective magnification proved to be the most efficient in providing high-resolution images. Overall sample processing rate was faster compared to other flow cell-objective combinations considering that only a single size class (20–300 um) had to be prepared. Clogging, a common problem encountered with the use of auto-imaging devices (Sosik and Olson 2007a; Alvarez et al. 2011), impedes sample analysis time. Hence, samples should be carefully filtered such that no particles greater than the width of the flow cell will be included. In this study, clogging was significantly reduced due to the bigger size of the FOV300 flow cell. Samples should also be properly concentrated so that aggregations and overlapping of particles will be prevented. In the study of Sieracki et al. (1998), they diluted their sample to 1000 ind/mL to prevent imaging of overlapping particles. As suggested by Alvarez et al. (2011), samples should be preassessed using alternative measurements such as bulk Chl *a* from field data, to have an idea on the sample dilution or concentration to be undertaken. In the present study, as much as possible, sample flow rate (FR) was adjusted such that the "particles per used image" (PPI) did not exceed the optimum value of 1.10, meaning that a single image does not contain several particles. When PPI values were way beyond 1.10, samples were diluted accordingly.

Table 3. FlowCAM software settings employed in the study. These values are within the recommended factory setting as stated in the FlowCAM manual.

Settings	Values
Particle capture	
Distance to nearest neighbor (μm)	50
Close holes iterations	2
Fluidics	
Sample volume (mL)	1.0
Flow rate (mL/min)	0.1
Auto-image	10
Efficiency (%)	67.2
Runtime	12.5
Camera	
Shutter	8
Gain	68
White balance U	2187
White balance V	2288
Brightness	0
Gamma	0
Hue	2290
Saturation	300
Sharpness	0
Color gain	0
Red gain	0
Green gain	0
Blue gain	0
Flash/camera delay	100
Flash duration	45

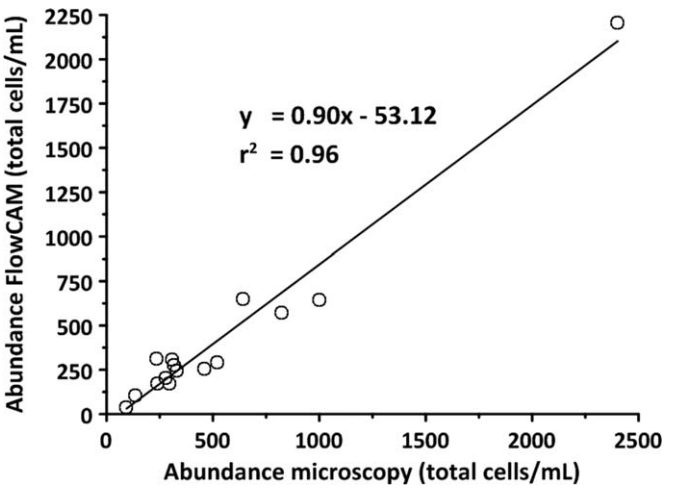

Fig. 2. Relationship between phytoplankton counts from FlowCAM and microscopy.

Previous studies have already established the reliability of the FlowCAM in providing precise estimates of the abundance and size of particles (Sieracki et al. 1998; Buskey and Hyatt 2006; Alvarez et al. 2011). As the use of FOV300 and 10X combination is beyond the recommended instrument hardware setting, FlowCAM counts in this study were compared to the standard microscopy method for validation. Similar to the results of Anjou et al. (2014) on the assessment of tephra in sediment samples, raw FlowCAM counts in this study were generally lower compared to the actual microscopy counts. This is expected since not all of the sample volume passing through the flow cell is being imaged by the camera. However, the final FlowCAM phytoplankton counts which were computed using the actual

Table 4. List of available phytoplankton libraries from FlowCAM runs of samples from Dipolog-Sindangan Bay and Butuan Bay collected during the SarDyn 2013 Cruise. Species with asterisk are the groups having at least 10 images of uniform orientation (e.g., girdle view, same number of chain, etc.) and of good resolution.

Dinoflagellates	Diatoms	Cyanobacteria
Alexandrium sp	*Actinoptychus*	*Trichodesmium thiebautii*
Ceratium furca	*Asterolamphalus*	
Ceratium fusus	*Asterolampra*	
Ceratium sp1	*Asterionellopsis*	
Ceratium sp2	**Bacteriastrum*	
Dinophysis caudata	*Bellorochea*	
Dinophysis miles	**Chaetoceros*	
Dissodinium	*Climacodium frauenfeldianum*	
Histoneis	**Coscinodiscus*	
Ornithocercus	*Coscinodiscus granii*	
Protoperidinium sp1	*Corethron criophilum*	
Protoperidinium sp2	*Ditylum brightwellii*	
Protoperidinium sp3	*Eucampia*	
Pyrocystis elegans	*Fragilaria doniopsis*	
Pyrocystis lunula	*Gossleriella*	
	**Guinardia/Dactyliosolen*	
	Helicotheca tamesis	
	Hemiaulus	
	Leptocylindicus	
	**Meuniera membranacea*	
	Odontella mobiliensis	
	Odontella sinensis	
	Odontella spp	
	Palmeria hardmaniana	
	**Planktoniella sol*	
	**Pseudo-nitzschia*	
	Pleurosigma	
	**Rhizosolenia*	
	Skeletonema	
	**Stephanopyxis*	
	Tabellaria	
	**Thalassionema nitzchoides*	
	**Thalassiosira*	

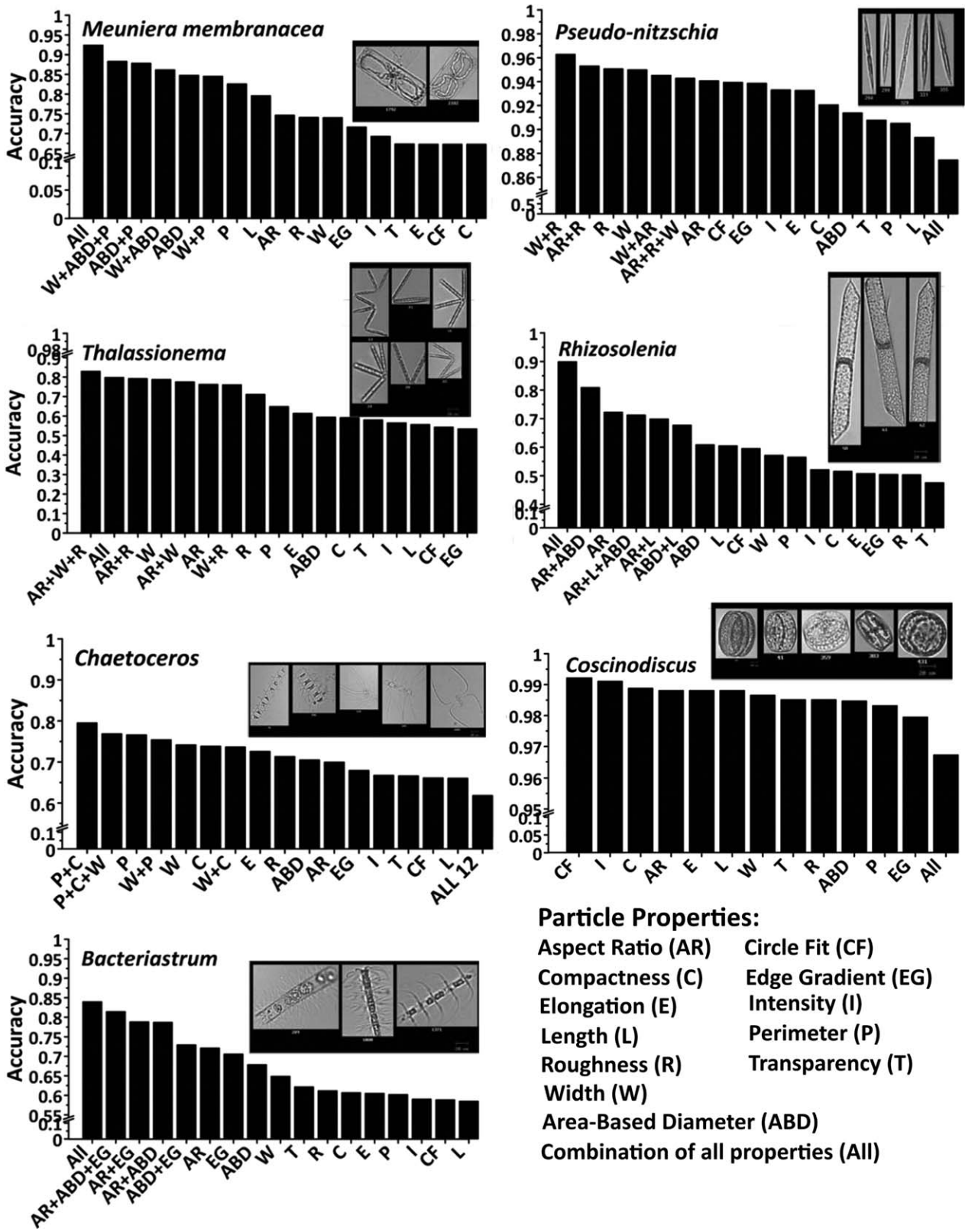

Fig. 3. Accuracy tests of filters for the different phytoplankton groups. Aspect Ratio (AR); Area-Based Diameter (ABD); Circle Fit (CF); Compactness (C); Edge Gradient (EG); Elongation (E); Intensity (I); Length (L); Perimeter (P); Roughness (R); Transparency (T); Width (W) and (All) combination of all 11 particle properties. Note that there are breaks on the *x*-axes.

volume sampled by the instrument, were highly comparable to that of the standard microscopy method. All sample counts were within the 100% coefficient of variation threshold set by Alvarez et al. (2014), meaning that the differences in count estimates between the two methods were below a factor of two. FlowCAM was also able to accurately count less dense samples with approximate concentrations of 200 ind/mL. This suggests that FlowCAM would also be effective in analyzing less dense Niskin samples provided that they are properly concentrated.

Optimization of FlowCAM's *Auto-Classification* capability

For the past decades, endeavors have been made to improve the auto-classification of phytoplankton images from different sources. Examples include the work of Gorsky et al. (1989) wherein they were able to distinguish species of phytoplankton with distinct size and shape using simple geometric properties. Culverhouse et al. (2003) employed neural network technique utilizing different texture and edge-based image properties in classifying dinoflagellate images and Verikas et al. (2014) used support vector machine (SVM) and random forest (RF) classifiers to discriminate the HAB-causing *Prorocentrum minimum*. The work of Sosik and Olson (2007b) may be the most detailed multiclass categorization of phytoplankton images from Imaging FlowCytobot. They were able to build 22 categories of phytoplankton and reported an overall accuracy of 88%. Conversely, Blaschko et al. (2005) and Alvarez et al. (2011) used SVM in classifying images from FlowCAM.

The difference of the present study from previous FlowCAM investigations is that the *VisualSpreadsheet* was fully utilized here for more detailed classification of images into taxonomic groups without further employing other software or programs. The goal was to optimize the *Auto-Classification* capability of the FlowCAM by determining the most efficient statistical filter for each phytoplankton library.

It is interesting to note that using the FOV300 and 10X objective combination, FlowCAM was able to capture good images of small-sized phytoplankton species (~ 60 μm in diameter) such as *Alexandrium*, *Dinophysis* and some *Protoperidinium* spp., which would have been difficult to capture had the recommended FOV300 and 4X objective been used. However, since not all identified phytoplankton taxa were represented by at least 10 good quality images of uniform species orientation, accuracy tests were conducted only on the dominant phytoplankton groups such as *Chaetoceros, Bacteriastrum, Coscinodiscus, Thalassionema, Meuniera membranacea, Pseudo-nitzschia*, and *Rhizosolenia*. The *VisualSpreadsheet* was shown to be most effective in classifying phytoplankton taxa with simple shapes such as the centric diatoms *Coscinodiscus* and *Thalassiosira* with accuracy values of greater than 0.90. However, as regards to phytoplankton groups having complex structures such as *Chaetoceros* and *Bacteriastrum*, its accuracy was reduced. Results of the accuracy tests conducted also suggest that the simultaneous use of all available particle properties as filter may not be effective for some phytoplankton species such as the case of *Chaetoceros*, *Coscinodiscus* and *Pseudo-nitzschia* wherein lowest accuracy values were observed. However, for the genera *Thalassionema, Bacteriastrum*, and *Rhizosolenia*, it was best to include all the particle properties in defining a statistical filter.

Summary and conclusion

The main purpose of auto-imaging devices such as FlowCAM is to provide rapid assessment of phytoplankton community structure without compromising taxonomic detail. It was the aim of the present study to optimize the FlowCAM to produce good quality images at a faster rate. The use of FOV300 and 10X objective combination was the most effective for this goal and produced phytoplankton abundance counts comparable to the conventional microscopy method. Several phytoplankton species were successfully imaged and identified even those genera with species of minute sizes (< 60 μm in diameter). It must be noted, however, that not all phytoplankton taxa encountered were sufficiently represented by good quality images. Aside from the fact that some species were rare and less abundant, producing images of uniform orientation was a challenge because particles are being captured in a moving fluid. Phytoplankton could be imaged in their valve or girdle view or as single individual or in chains or colonies. This means that a single species could be further classified into sub-categories depending on their orientation in the image. Hence, accuracy tests of statistical filters were done only for the dominant phytoplankton groups. Highest classification accuracy values were observed in phytoplankton species with simple shapes such as the spherical *Coscinodiscus* and the thin-elongate *Pseudo-nitzschia*, while lower values were generated for those groups with complex structures such as *Chaetoceros* and *Bacteriastrum*.

Previous FlowCAM studies have been limited to classifying images to general phytoplankton groups (e.g., diatoms, centric diatoms, pennate diatom, etc.) (Buskey and Hyatt 2006). However in this study, it was possible to identify phytoplankton to lower taxonomic level such as to species level, due to the high-resolution of images produced by the optimized FlowCAM configurations. As long as the hardware and software settings utilized in this study will be employed, available phytoplankton libraries may be continuously populated and their corresponding statistical filters be used for the auto-classification of images from different sets of samples. The development of these improved automated classification systems are very much needed in the field of plankton research to achieve increased taxonomic resolution at potentially finer spatial and temporal scales.

References

Alvarez, E., A. Lopez-Urrutia, E. Nogueira, and S. Fraga. 2011. How to effectively sample the plankton size spectrum? A case study using FlowCAM. J. Plankton Res. **33**: 1119–1133. doi:10.1093/plankt/fbr012

Alvarez, E., A. Lopez-Urrutia, and E. Nogueira. 2012. Improvement of plankton biovolume estimates derived from image-based automatic sampling devices: Application to FlowCAM. J. Plankton Res. **34**: 454–469. doi:10.1093/plankt/fbs017

Alvarez, E., M. Moyano, A. Lopez-Urrutia, E. Nogueira, and R. Scharek. 2014. Routine determination of plankton community composition and size structure: A comparison between FlowCAM and light microscopy. J. Plankton Res. **36**: 170. doi:10.1093/plankt/fbt069

Benfield, M., and others. 2007. RAPID: Research on automated plankton identification. Oceanography **20**: 172–187. doi:10.5670/oceanog.2007.63

Blaschko, M. B., and others. 2005. Automatic in situ identification of plankton, V. 1, p. 79–86. IEEE Workshop on Applications of Computer Vision.

Buskey, E. J., and C. J. Hyatt. 2006. Use of the FlowCAM for semi-automated recognition and enumeration of red tide cells (*Karenia brevis*) in natural plankton samples. Harmful Algae **5**: 685–692. doi:10.1016/j.hal.2006.02.003

Culverhouse, P. F., R. Williams, B. Reguera, V. Herry, and S. Gonzalez-Gil. 2003. Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. Mar. Ecol. Prog. Ser. **247**: 17–25. doi:10.3354/meps247017

D'Anjou, R. M., N. L. Balascio, and R. S. Bradley. 2014. Locating cryptotephra in lake sediments using fluid imaging technology. J. Paleolimnol. **52**: 257–264. doi:10.1007/s10933-014-9786-2

Gorsky, G., P. Guilbert, and E. Valenta. 1989. The autonomous image analyzer—enumeration, measurement and identification of marine phytoplankton. Mar. Ecol. Prog. Ser. **58**: 133–142. doi:10.3354/meps058133

Ide, K., K. Takahashi, A. Kuwata, M. Nakamachi, and H. Saito. 2008. A rapid analysis of copepod feeding using FlowCAM. J. Plankton Res. **30**: 275–281. doi:10.1093/plankt/fbm108

Liu, H., M. J. Dagg, and S. Strom. 2005. Grazing by the calanoid copepod *Neocalanus cristatus* on the microbial food web in the coastal Gulf of Alaska. J. Plankton Res. **27**: 647–662. doi:10.1093/plankt/fbi039

Omura, T., M. Iwataki, V. M. Borja, H. Takayama, and Y. Fukuyo. 2012. Marine phytoplankton of the Western Pacific, Koseishakoseikaku.

Reul, A., and others. 2014. Effect of CO_2, nutrients and light on coastal plankton. III. Trophic cascade, size structure and composition. Aquat. Biol. **22**: 59–76. doi:10.3354/ab00585

Romagnan, J.-B., and others. 2015. Comprehensive model of annual plankton succession based on the whole-plankton time series approach. PLos One **10**: 1–18. doi:10.1371/journal.pone.0119219

Sieracki, C. K., M. E. Sieracki, and C. S. Yentsch. 1998. An imaging-in-flow system for automated analysis of marine microplankton. Mar. Ecol. Prog. Ser. **168**: 11. doi:10.3354/meps168285

Siswanto, E., J. Ishizaka, S. C. Tripathy, and K. Miyamura. 2013. Detection of harmful algal blooms of Karenia mikimotoi using MODIS measurements: A case study of Seto-Inland Sea, Japan. Remote Sens. Environ. **129**: 185–196. doi:10.1016/j.rse.2012.11.003

Sosik, H. M., and R. J. Olson. 2007*a*. A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot. Limnol. Oceanogr.: Methods **5**: 195–203. doi:10.4319/lom.2007.5.195

Sosik, H. M., and R. J. Olson. 2007*b*. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. Limnol. Oceanogr.: Methods **5**: 204–216. doi:10.4319/lom.2007.5.204

Stauffer, B. A., J. I. Goes, K. T. Mckee, H. Do Rosario Gomes, and P. J. Stabeno. 2014. Comparison of spring-time phytoplankton community composition in two cold years from the western Gulf of Alaska into the southeastern Bering Sea. Deep-Sea Res. **109**: 57–70. doi:10.1016/j.dsr2.2014.03.007

Tomas, C. M. 1997. Identifying marine phytoplankton. Academic Press, San Diego.

Verikas, A., A. Gelzinis, M. Bacauskiene, I. Olenina, E. Vaiciukynas. 2014. An integrated approach to analysis of phytoplankton images. IEEE J. Oceanic Eng. **40**: 315–326. 10.1109/JOE.2014.2317955

Yamaji, I. 1982. Illustrations of the marine plankton of Japan. Hoikusha.

Zarauz, L., X. Irigoien, and J. A. Fernandes. 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. doi:10.1093/plankt/fbn107

Acknowledgments

We thank the SarDyn research team and crew of the M/V BFAR for aiding in the collection of the phytoplankton samples. We also wish to acknowledge the inputs of Joseph Dominic Palermo and Iris Bollozos in helping to determine the appropriate approach in the FlowCAM analysis. Thank you as well to Arjay Cayetano for providing useful algorithms for data analysis. We profusely thank our project and overall program leader Cesar Villanoy for providing this research opportunity, as well as for discussions on the FlowCAM outputs. This study was funded by the Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development, Department of Science and Technology under the Sardine Research Program (SarDyn). The FlowCAM was provided through an Office of the Naval Research grant to Pierre Flament (University of Hawaii).

Submitted 21 September 2015
Revised 28 December 2015
Accepted 18 January 2016

Associate editor: Paul Kemp