Environ Sci Pollut Res (2013) 20:6896–6902
DOI 10.1007/s11356-012-1437-4

Continuous automated imaging-in-flow cytometry for detection and early warning of *Karenia brevis* blooms in the Gulf of Mexico

Lisa Campbell · Darren W. Henrichs · Robert J. Olson · Heidi M. Sosik

Received: 30 September 2012 / Accepted: 13 December 2012 / Published online: 11 January 2013

Abstract Monitoring programs for harmful algal blooms (HABs) typically rely on time-consuming manual methods for identification and enumeration of phytoplankton, which make it difficult to obtain results with sufficient temporal resolution for early warning. Continuous automated imaging-in-flow by the Imaging FlowCytobot (IFCB) deployed at Port Aransas, TX has provided early warnings of six HAB events. Here we describe the progress in automating this early warning system for blooms of *Karenia brevis*. In 2009, manual inspection of IFCB images in mid-August 2009 provided early warning for a *Karenia* bloom that developed in mid-September. Images from 2009 were used to develop an automated classifier that was employed in 2011. Successful implementation of automated file downloading, processing and image classification allowed results to be available within 4 h after collection and to be sent to state agency representatives by email for early warning of HABs. No human illness (neurotoxic shellfish poisoning) has resulted from these events. In contrast to the common assumption that *Karenia* blooms are near monospecific, post-bloom analysis of the time series revealed that *Karenia* cells comprised at most 60–75 % of the total microplankton.

Keywords Early warning · Gulf of Mexico · Harmful algal bloom · Imaging FlowCytobot · *Karenia* · Time series

Responsible editor: Philippe Garrigues

Electronic supplementary material The online version of this article (doi:10.1007/s11356-012-1437-4) contains supplementary material, which is available to authorized users.

L. Campbell
Department Oceanography, Texas A&M University, College Station, TX 77843, USA

L. Campbell (✉) · D. W. Henrichs
Department Biology, Texas A&M University, College Station, TX 77843, USA
e-mail: lisacampbell@tamu.edu

R. J. Olson · H. M. Sosik
Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA

Introduction

Over the last several decades, the frequency of harmful algal blooms (HABs) has increased worldwide (Anderson 2009, Hallegraeff 1993). HAB events caused by explosive growth and/or accumulation of toxic phytoplankton threaten human health, tourism, fisheries and ecosystem function. Early warning is an essential step in mitigating the effects of HABs, so rapid and species-specific methods are needed to identify and quantify HAB cells. To meet this challenge, a number of new technologies for detection and identification have been developed (Anderson et al. 2012, Paul et al. 2007, Sellner et al. 2003). Constraints that can limit these approaches include the need for continuous monitoring to detect events that can appear suddenly, the ability to identify novel species (if previously unknown at that location), and the flexibility to identify multiple species.

The Imaging FlowCytobot (IFCB) is one solution for HAB monitoring in coastal areas (Sosik and Olson 2007, Sosik et al. 2011). IFCB is capable of unattended long-duration deployments and produces high-quality images that allow many phytoplankton cells to be identified at the genus or even species level. Using a combination of flow cytometric and video technology, IFCB captures high resolution images of individual phytoplankton (~10–150 μm) and measures the

chlorophyll fluorescence associated with each image. All images are archived, so data can be re-examined to identify novel species or co-occurring species.

IFCB has been employed successfully in the Gulf of Mexico, where a novel and unexpected occurrence of *Dinophysis ovum* in 2008 produced the first shellfish closure in the US due to okadaic acid (diarrhetic shellfish poisoning; Campbell et al. 2010). Early warning from IFCB prompted targeted sampling for conventional microscopy and toxin assays. Subsequently, state agencies were alerted and regional waters were closed to oyster harvesting. IFCB provided time series data for *D. ovum* abundance at the ship channel in Port Aransas, TX throughout the 4-month bloom. In view of the large amounts of data collected during continuous monitoring (>300 million images each year), an automated image classification is essential to provide identification and determination of cell abundances. Verification with independent samples manually examined by conventional microscopy has demonstrated that the accuracy of IFCB is comparable to that of human experts (Sosik and Olson 2007; Campbell et al. 2010). Successful early detection of *D. ovum* was repeated in 2010, 2011, and 2012.

Here, we report on the development of the IFCB approach for early warning of *Karenia brevis*, the major HAB species of concern in the Gulf of Mexico. Closure of shellfish harvesting is mandated when abundance of *K. brevis* reaches 5 cells mL^{-1} (Comprehensive Shellfish Control Code 1987). Unlike other states with recurring HABs, Texas has no regular monitoring program in place. Coverage is uneven and often depends on volunteers. Furthermore, a continuous and automated monitoring program with the IFCB serves to guide field sampling for shellfish toxicity testing and the more labor intensive manual microscopy.

In 2009, early warning of a *Karenia* bloom from manual inspection of IFCB images prompted state agencies to delay the opening of oyster season, which successfully prevented human illness. Here, we describe use of an automated classifier trained with images from the 2009 bloom to evaluate a subsequent bloom in 2011 (no bloom was detected in 2010). Successful implementation of automated file downloading, processing and image classification allowed results to be available within 4 h after collection. An automated notification system, in which classification results are sent to state agency representatives by email, extended the usefulness of IFCB for early warning of HABs.

Materials and methods

Imaging FlowCytobot

The IFCB combines flow cytometry with video technology and can be used to study a wide variety of plankton cell types, ranging from <10 μm flagellates to large diatom chains (Olson and Sosik 2007, Sosik and Olson 2007, Sosik et al. 2011). The instrument utilizes a red diode laser; as each particle passes through the laser beam, light is scattered and chlorophyll-containing phytoplankton emit fluorescence. The fluorescence signal triggers a 1-μs exposure from a flash lamp and capture of a video frame from a CCD camera. High-quality images (~1 μm resolution) are recorded and permit detection of characteristic features such as flagella, spines, cilia, and internal cell structures, so that most taxa can be identified at least to genus level.

Deployment

The IFCB was deployed in the pier laboratory at the University of Texas Marine Science Institute (UTMSI) in Port Aransas, Texas (27.84° N, 97.07° W) beginning in September 2007 (Campbell et al. 2010) (Figure S1). The instrument collected and analyzed 5-mL samples every 20 min and internally stored red fluorescent standard particles (9 μm, XPR-1653, Duke Scientific Corp., Palo Alto, CA, USA) were run after every 50th seawater sample. Data files were transferred from UTMSI back to the Campbell Laboratory via the Internet for analysis and archival and images were classified within 1 day of acquisition in 2009 and within 4 h in 2011.

Automated classification

The automated classification approach required the use of a variety of image processing and feature extraction techniques together with a support vector machine (machine learning algorithm) that had been trained with example images from each category (Sosik and Olson 2007). Because the focus was accurate enumeration of *Karenia*, a series of "one-versus-all" support vector machines was employed, as described previously for *D. ovum* (Campbell et al. 2010). Training set images were taken from samples selected at random intervals from the 2007–2009 time series; because *Karenia* was not detected or present at very high abundance at other times during this interval, most of the *Karenia* images used for training were from the 2009 bloom. The "one-versus-all" approach allowed different subsets of image features to be used for each classifier and each subset was optimized to discriminate the corresponding category from all other categories combined. For this case, 12 categories (*Karenia*, Acantharia, *Brachidinium*, *Chattonella*, *Cymatosira*, centric diatom, ciliates, detritus, small dinoflagellates, *Ebria*, *Prorocentrum minimum*, and *Thalassiosira*) were used (Fig. 1); these correspond to classes that exhibited a relatively high rate of misclassification with *Karenia* if they were not explicitly included. Each image from the time series was evaluated with the resulting classifiers to determine the best match. Assignment to a category was based on the highest probability of association (if an

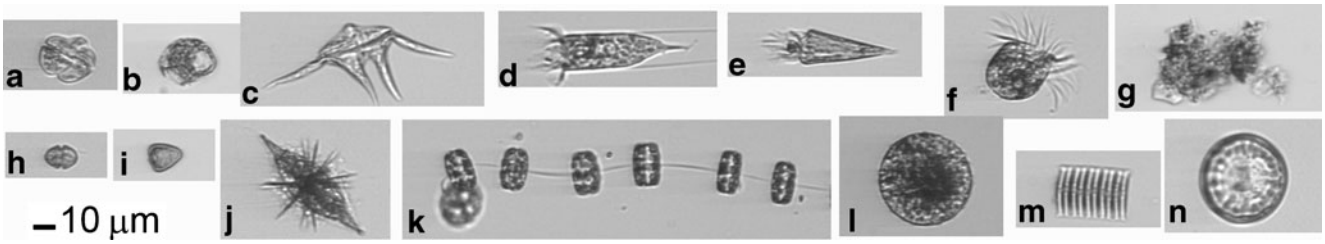

Fig. 1 Representative IFCB images for the categories used in *Karenia* classification and error correction **a** *Karenia*, **b** *Ebria*, **c** *Brachidinium*, **d**–**f** ciliates (e.g., Tintinnids, Strombidiids, Strobilidiids), **g** detritus, **h** small dinoflagellates, **i** *Prorocentrum minimum*, **j** Acantharia, **k** *Thalassiosira*, **l** *Chattonella*, **m** *Cymatosira*, **n** centric diatoms

image was positively associated with more than one category), with images that did not meet the positive decision threshold for any of the classes disregarded from further class-specific analysis (i.e., labeled "unclassified"). Several categories included a mix of genera (e.g., mixture of centric diatoms, mixture of ciliates, mixture of small dinoflagellates) which were not separated in more detail by the current classification scheme focused on detection of *Karenia*. Final cell abundance estimates for the time series were obtained after correction for the percentage of false positives and the percentage correctly identified, which were determined from classifier performance on an independent set of manually identified images.

Manual classification and *K. brevis* bloom composition

Approximately 9000 5-mL samples were measured during the 5 month *Karenia* bloom period from August–December in both 2009 and 2011. For each year, a subset (~400 samples selected at random from each day of the time series) was manually inspected and all images were annotated. All *Karenia* images were assigned to "*Karenia*" and all non-*Karenia* images classified as "other." For 2009, additional samples (~7,500) were manually inspected and corrected for use in determining the threshold for automated notification. These data were binned in 2-h intervals and manually corrected files were compared with the cell counts obtained from automated classification of the same samples. For the time series plots, data were binned into daily intervals and used to calculate corrected cell abundance (in cells mL^{-1}).

Archived images from both the 2009 and 2011 bloom time series were also examined to determine the relative contribution of *Karenia* to the total microplankton (by number) during different phases of the bloom. Images from samples over the course of the blooms (initial, peak, and termination phases)

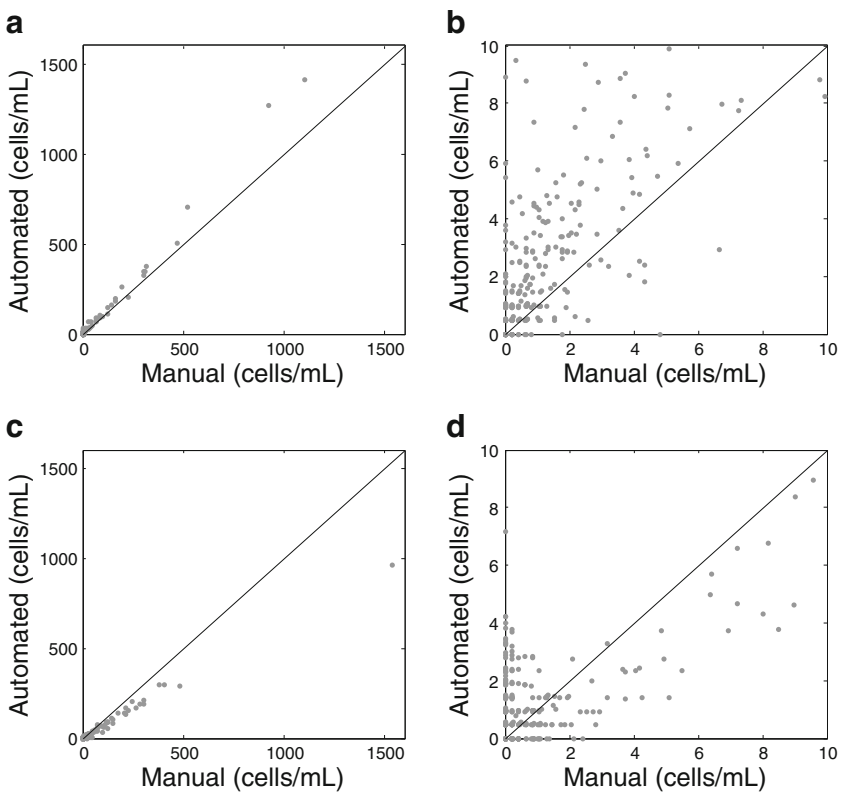

Fig. 2 Comparison between manual counts (visually inspected and corrected) and automated classifier counts (after correction for estimated % false positive and % detection) with the 1:1 line shown. **a** The 2009 *Karenia* bloom at Port Aransas, TX; r=0.995, **b** Expanded view of 0–19 cells mL^{-1} range, **c** The 2011 *Karenia* bloom at Port Aransas, TX; r=0.995, **d** Expanded view of 0–10 cells mL^{-1} range

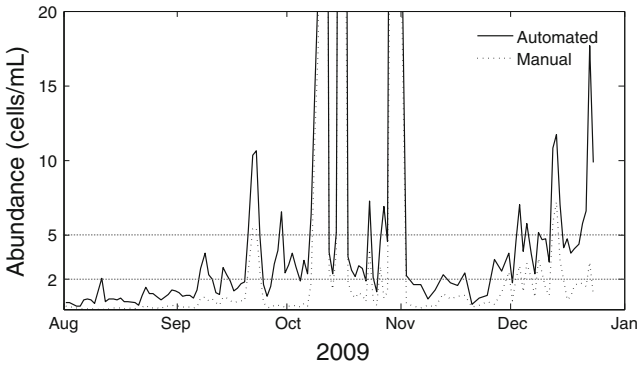

Fig. 3 Time series for daily binned files for the 2009 *Karenia* bloom from automated classification and manual classification. The regulatory threshold of 5 cells mL^{-1} and the 2 cellsmL^{-1} threshold selected for early warning notification are also plotted

were manually inspected to classify all debris images as "detritus," and this category was subtracted from the total events to obtain a "total microplankton" count for each file. For this subset of the observations, the average percentage of detritus was calculated for 2009 and 2011 (10.6 and 13.6 %, respectively) and used to estimate "total microplankton" counts over the time series during each of the two blooms and the % *Karenia* at each bloom phase. Note that although the number of detritus particles generally greatly exceeds the number of phytoplankton (Olson and Sosik 2007), most are not imaged by IFCB. Only detritus particles with sufficient chlorophyll fluorescence will trigger imaging.

Results

Automated classification

The first *K. brevis* bloom observed by the IFCB at Port Aransas, TX occurred in fall 2009. Manual inspection of images from the time series allowed detection of *Karenia* cells in early August and continued observation delivered direct observations of the bloom as it developed. This 2009 event provided the IFCB images used to create a training set from which a *Karenia* classifier could be developed.

The automated classifier produced for *Karenia* based on the 2009 bloom was used to classify all images from both the 2009 and 2011 bloom (Fig. 2). *Karenia* abundance (in cells mL^{-1}) calculated from automated classification (after correction for classification error) were strongly correlated with the results from manual inspection (Fig. 2). In 2009, the correlation was very strong (r=0.995) with automated classification producing slightly overestimated counts (Fig. 2a). The same relationship was noted when all ~7,500 data points for 2009 were plotted (data not shown). Much of the variability was evident at the lower cell abundances (Fig. 2b). In 2011, the counts were also highly correlated (r=0.995; Fig. 2c) and the automated classification results underestimated *Karenia* abundance. Again, most of the variability was observed at abundances <10 cells mL^{-1} (Fig. 2d).

Monitoring for early warning

For successful early warning, a lower limit for automated notification needed to be established so that *Karenia* occurrences were not missed, but at the same time the number of false alarms was minimized. Based on a comparison of the time series for both the automated classification and manually

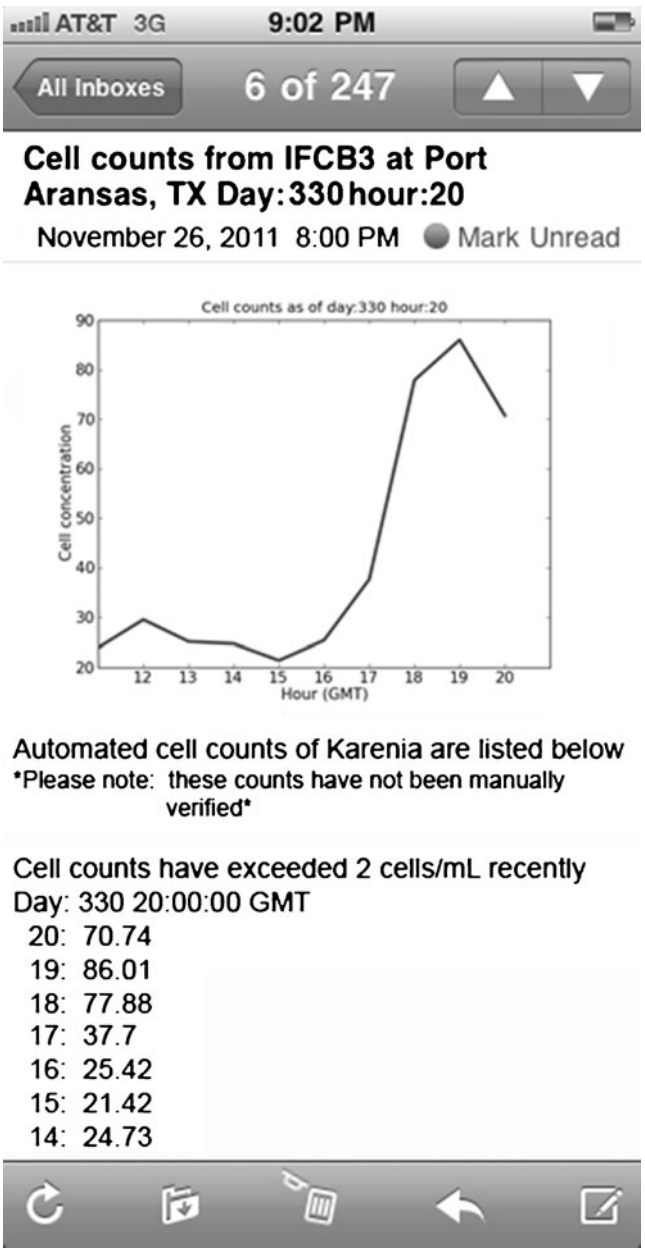

Fig. 4 Sample email message sent during the 2011 *Karenia* bloom. Automated classification results and a graph of *Karenia* abundance (not manually verified) for previous ~10 h

verified counts in 2009, a threshold of 2 cells mL^{-1} was chosen. Although 266 "real" occurrences (when 2-h binned manual counts were >2 cells mL^{-1}) were observed on 57 days during this 5 month period, for the purpose of evaluating early warning only the pre-bloom period was considered. Pre-bloom (Aug 1–Sept 20) was defined as the period before the daily average *Karenia* abundance exceeded 5 cells mL^{-1} (Fig. 3). During these 51 days, 9 "real" detections were reported (when both 2-h binned manual and automated counts were >2 cells mL^{-1}), none were missed, but 93 false alarms (when manual counts were <2 cells mL^{-1}, but automated were >2 cells mL^{-1}) were noted. The number of false alarms was 17 % of the total pre-bloom reports (554 2-hour binned automated counts). For comparison, when a threshold of 1 cell mL^{-1} was used, the number of "real" detections increased to 41, none were missed, and the number of false alarms increased to 172 (31 %). In contrast, if the threshold was increased to 3 cells mL^{-1}, although the number of false alarms decreased to 48 (9 %), only 1 "real" detection was reported, and none were missed. Manual counts suggested low, but variable abundance of *Karenia* cells throughout the early warning period.

Automated notification

Using the threshold value of >2 cells mL^{-1}, an automated notification system was implemented to send an email message that included a graph of the cell counts for the previous 8 to 10 h (Fig. 4). Image data were downloaded every 3 h from UTMSI, features were automatically extracted and presented to the automated classifier to estimate *Karenia* abundance. If automated classification results indicated abundance >2 cells mL^{-1}, an email was generated and sent to a distribution list that included the researchers and representatives for the HAB program at Texas Parks and Wildlife and the Texas State Department of Health Services (for shellfish monitoring). During 2011, more than 130 automated email notifications were sent.

Community structure

Although the number of *Karenia* was often a large percentage of the total microplankton, blooms were never "monospecific." In 2009, at the early stage of the developing bloom, *Karenia* contributed <1 % of the total phytoplankton abundance, but 3 weeks later during the peak of the bloom, the average daily percentage of *Karenia* was 34–48 %, though the range within a single day could be quite variable (26–76 %; Fig. 5). During mid-October, IFCB image data also revealed a bloom of *Asterionellopsis* co-occurring with the bloom of *Karenia*. In contrast, during 2011 there were no "initiation" phase samples in the month prior to the bloom because peak *Karenia* abundances >5 cells mL^{-1} appeared suddenly in mid-October. A gap in data also coincided with this period (Fig. 6). These few days were removed because the sample illumination was not sufficient for correct feature extraction, so automated classification was not possible. The first peak in 2011 showed a higher proportion (30 %) of *Karenia* (Fig. 6) than the early bloom samples in 2009 (Fig. 5). Samples examined from successive

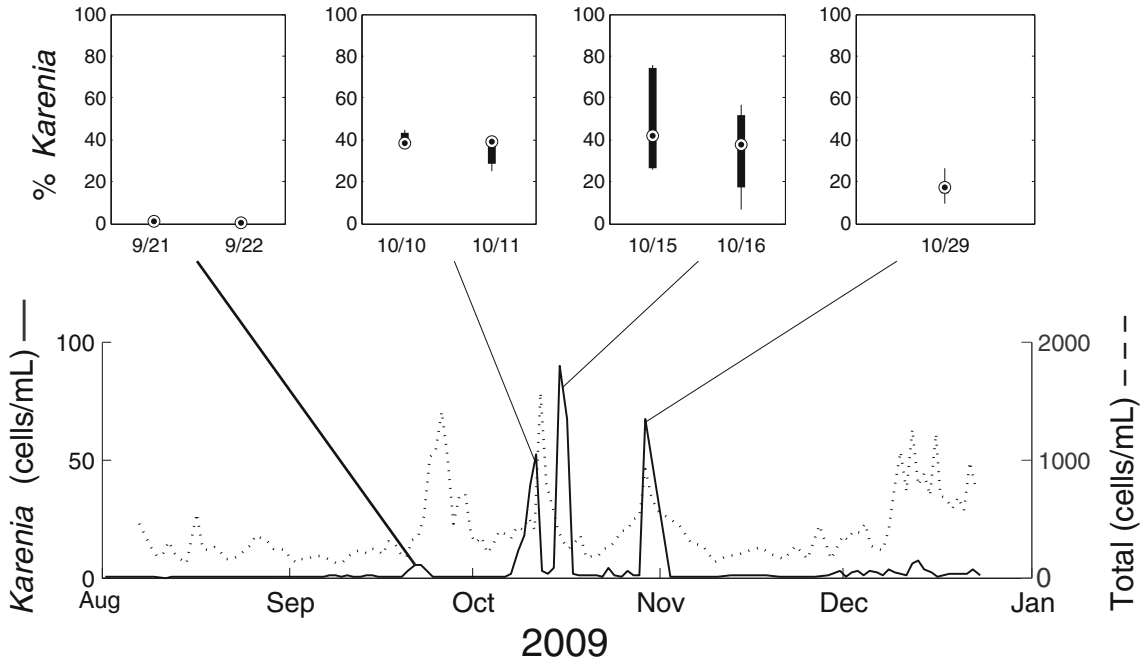

Fig. 5 The 2009 time series at Port Aransas, TX. *Top Karenia* as a percentage of the total microplankton over the bloom. Files were selected at random on 9/21 (n=3); 9/22 (n=3); 10/10 (n=3); 10/11 (n=3); 10/15 (n=6); 10/16 (n=4); 10/29 (n=5). *Points* indicate median, *error bars* indicate range of 25th percentile, and *whiskers* indicate range of all data. *Bottom* Time series of abundance for total microplankton and *Karenia*

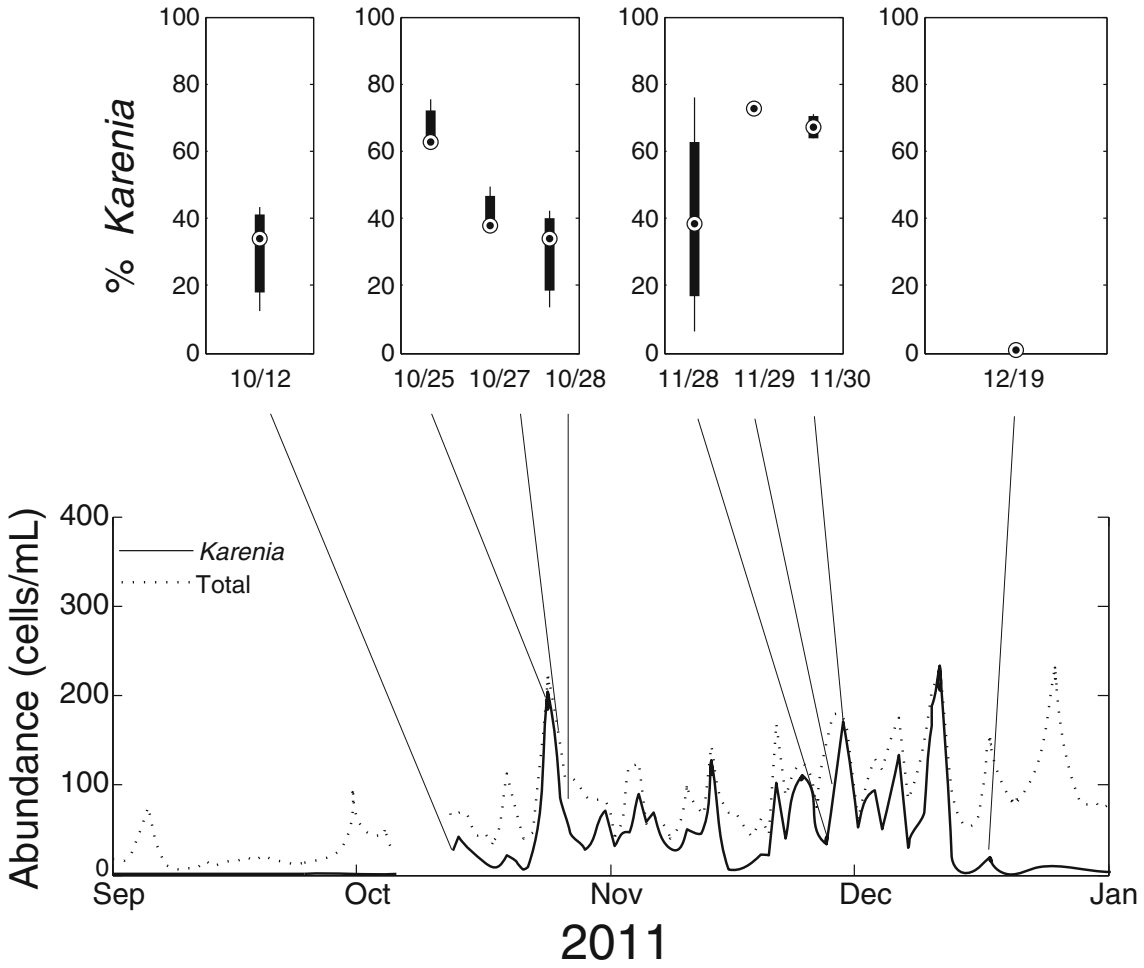

Fig. 6 The 2011 time series at Port Aransas, TX. *Top Karenia* as a percentage of the total microplankton over the bloom. Files were selected at random on 10/12 (n=3); 10/25 (n=3); 10/27 (n=3); 10/28 (n=3); 11/28 (n=4); 11/29 (n=4); 11/30 (n=4); 12/16 (n=3). *Whisker plots*, as in Fig. 5. *Bottom* Time series of abundance for total microplankton and *Karenia*

days showed that the highest percentages coincided with the peaks in abundance (60–73 %; Fig. 6). Several diatom (e.g., *Asterionellopsis*, *Chaetoceros*, *Ditylum*) and dinoflagellate (e.g., *Ceratium*, *Prorocentrum*) species were observed during the 2011 bloom period. None of these co-occurring species were observed at elevated concentrations, which is consistent with the overall higher percentage of *Karenia* in 2011.

Discussion

Mitigation of the effects of HABs is most effective if sufficient early warning is provided. To be successful, an early warning method for HAB detection should be continuous, automated, and taxon specific; the IFCB meets these criteria. It has provided a nearly continuous time series since its deployment in 2007 at Port Aransas, TX. The downloading and analysis of data has been fully automated, and the IFCB has provided images for identification and early warning for six HAB events. The 2008 *D. ovum* event was the first toxic bloom of this species ever observed in Gulf of Mexico waters (Campbell et al. 2010). Despite not being the objective of the monitoring program (because blooms had never previously been reported), the flexibility of the imaging approach and the ability to develop taxon-specific classifiers demonstrated the advantage of the IFCB approach. Following this initial discovery, images from IFCB have provided successful early warning for three additional *D. ovum* blooms in 2010, 2011, and 2012 (data not shown).

The success of the IFCB has now been repeated for *Karenia*. In 2009, initial observations of low cell abundances were reported at least 1 month in advance of the cell abundance reaching the legal limit (5 cells mL^{-1}) requiring shellfish closures. In both 2009 and 2011, *Karenia* blooms were detected with sufficient lead time to close fisheries and prevent human illness. The choice of 2 cells mL^{-1} appears to be an appropriate threshold for early warning notification. Balancing the number of false alarms vs. missing a "real" occurrence, this threshold is a conservative approach. False

alarms initiate manual verification of the images, which is easily accomplished, and field sampling. In future implementations, we plan to enhance alert emails with web links to images automatically identified as potential *Karenia*; this will allow resource managers and other users to quickly discern false alarms. Another advantage of the imaging approach by the IFCB is that all of the larger phytoplankton cells are recorded, not just *Karenia*, which provides valuable information about the structure of the phytoplankton community during different phases of a bloom. Fluctuations in *Karenia* abundance in 2009 co-occurred with a bloom of the diatom *Asterionellopsis*. These types of observations provide valuable information for ecological studies of harmful algal species.

There are some limitations to the IFCB and automated analysis approach. Training sets tailored to a local ecosystem are required, but the flexibility of developing additional classifiers makes adjusting to incorporate novel species or new occurrences possible. We also note that the IFCB image details are not always sufficient to reliably distinguish between closely related species of *Karenia*, e.g. *K. brevis*, *Karenia mikimotoi*, *Karenia papilionacea*. However, previous observations suggest that fish-killing blooms are usually dominated by *K. brevis* (Heil and Steidinger 2009). During both the 2009 and 2011 blooms, the presence of *K. brevis* was validated both by manual inspection of the IFCB images and by light microscopy of field samples collected by the Texas Parks & Wildlife Department (data not shown). In 2009, samples were genotyped to confirm species identification as *K. brevis* (Henrichs et al. 2013). Genus-level identifications provided by IFCB will be valuable for early warning of *K. brevis* because these warnings will trigger an early response and validation with field samples. The training images are very important to successful classification. Lack of 2011 *Karenia* cells in the training set may explain the underestimated abundances in 2011 compared to 2009.

Limitations associated with automated image processing are another consideration. Gaps in data (e.g., Fig. 6) can occur if image data exists but the illumination conditions are outside the range appropriate for the current image processing and feature extraction algorithms. While our current approach is robust to moderate variations in illumination, it is not adequate for the full range of conditions in the time series and we are developing refined image processing steps to provide a more robust analysis procedure.

Conclusion

The IFCB deployed on the Texas coast at Port Aransas has provided early warnings for six HAB events. Implementation of automated processing and image classification, together with an automated email notification system, has allowed state agencies to be notified within 4 h after sample collection. In 2009 and 2011, this rapid notification permitted a timely response and closure of oyster harvesting, which successfully prevented any human illness due to neurotoxic shellfish poisoning. As continuous ocean observing systems are now being built, new technologies for cell enumeration will become more automated and address the traditional problem of undersampling the ocean. Improved image processing, feature extraction and classification methods for IFCB data means the presence of HAB species can be detected much earlier than routine, and less frequent, sampling could report, which will lead to more accurate early warning notifications and continued prevention of human illness as a result of toxic shellfish.

Acknowledgments Support was provided by the National Oceanographic and Atmospheric Administration/Ecology of Harmful Algal Bloom program (NOAA/ECOHAB) to L.C., R.J.O. and H.M.S. for this project and by the Gordon and Betty Moore Foundation to R.J.O. and H.M.S. for advances in IFCB and associated data processing. We acknowledge Stephen Christopher, Laura Harred and Ian Ivy for their assistance with manual correction and data analysis. This paper is ECOHAB contribution # 721.

References

Anderson DM (2009) Approaches to monitoring, control and management of harmful algal blooms (HABs). Ocean Coast Manag 52:342–347

Anderson DM, Cembella AD, Hallegraeff GM (2012) Progress in understanding Harmful Algal Blooms: Paradigm shifts and new technologies for research, monitoring, and management. In: Carlson CA, Giovannoni SJ (Editors), Annual Review of Marine Science 4. pp. 143–176

Campbell L, Olson RJ, Sosik HM, Abraham A, Henrichs DW, Hyatt CJ, Buskey EJ (2010) First harmful *Dinophysis* (DINOPHYCEAE, DINOPHYSIALES) bloom in the US is revealed by automated imaging flow cytometry. J Phycol 46:66–75

Hallegraeff GM (1993) A review of harmful algal blooms and their apparent global increase. Phycologia 32:79–99

Heil CA, Steidinger KA (2009) Monitoring, management, and mitigation of *Karenia* blooms in the eastern Gulf of Mexico. Harmful Algae 8:611–617

Henrichs DW, Renshaw MA, Gold JR, Campbell L (2013) Population genetic structure of the toxic dinoflagellate *Karenia brevis* from the Gulf of Mexico. J. Plankt. Res. (in press)

Olson RJ, Sosik HM (2007) A submersible imaging-in-flow instrument to analyze nano- and microplankton: Imaging FlowCytobot. Limnol Oceanogr Methods 5:195–203

Paul JH, Scholin C, van den Engh G, Perry MJ (2007) In situ instrumentation. Oceanography 20:70–78

Sellner KG, Doucette GJ, Kirkpatrick GJ (2003) Harmful algal blooms: causes, impacts and detection. J Ind Microbiol Biotechnol 30:383–406

Sosik HM, Olson RJ (2007) Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. Limnol Oceanogr Methods 5:204–216

Sosik HM, Olson RJ, Armbrust EV (2011) Flow cytometry In: Suggett DJ, Prasil O , Borowitzka MA (Editors), Chlorophyll *a* fluorescence in aquatic sciences: Methods and applications. Developments in Applied Phycology, 4. Springer, pp. 171–185

The Comprehensive Shellfish Control Code (1987). Florida Administrative Code. Chapter 16R-7