LIMNOLOGY
and
OCEANOGRAPHY: METHODS

Limnol. Oceanogr.: Methods 9, 2011, 466–477

SeaFlow: A novel underway flow-cytometer for continuous observations of phytoplankton in the ocean

*Jarred E Swalwell, Francois Ribalet, E. Virginia Armbrust**
School of Oceanography, University of Washington, Box 357940, Seattle, Washington, 98195, USA

Abstract

Current automated flow cytometers primarily target the dynamics of phytoplankton communities in coastal environments and lack the sensitivity to resolve the spatial distribution of *Prochlorococcus*, the smallest phytoplankton that dominates the open ocean. Here we present SeaFlow, a novel flow cytometer developed for continuous real-time observations of natural assemblages of small phytoplankton cells, including *Prochlorococcus*. Unlike other flow cytometers, SeaFlow does not use sheath fluid. Instead, a virtual-core is used to determine the position of a particle in the stream of seawater. By eliminating sheath fluid, SeaFlow can continuously sample the seawater stream directly from a ship's intake system. Image analysis is used to automatically align the laser with the optical system and then monitor and correct for drift. SeaFlow performs rapid quantification (up to 24,000 cells per second) of multidimensional characteristics of phytoplankton cells in the pico- to nanophytoplankton size range (0.5-20 µm) to analyze the equivalent of 480 traditional flow cytometry samples per day while on board a research vessel. Data analysis tools have been created to automatically cluster and count phytoplankton populations with geo-referenced data visualization. SeaFlow makes it possible to explore surface phytoplankton dynamics at a spatial scale ranging from a few meters to thousands of kilometers. An example data set is presented from a 450-km long transect near the Hawaiian Islands with the continuous data aggregated in 3 min intervals.

The rapidity and sensitivity with which a flow cytometer can measure individual cells has made it an integral oceanographic tool for the last three decades (Sosik et al. 2010), enabling significant progress toward addressing the patterns of phytoplankton distributions in the oceans (Li 2002). Flow cytometry measures light scatter and fluorescence emission of individual phytoplankton cells at rates of up to several thousand cells per second. Light scattering is roughly proportional to the cell size, and fluorescence is unique to the emission spectra of pigments. These parameters can be used to identify populations of phytoplankton with similar characteristics (Olson et al. 1989). For example, flow cytometery at sea made possible the discovery of *Prochlorococcus*, the smallest and most abundant photosynthetic organism in the oceans. These cyanobacteria represent 40% to 50% of the phytoplankton biomass in oligotrophic regions (Vaulot et al. 1995). Despite their importance, relatively little is known about their temporal and spatial variability because of a lack of observational tools for large-scale, high resolution mapping of phytoplankton distributions.

Use of conventional flow cytometers in the field is limited by their large size and complexity, and their operation commonly depends upon a dedicated instrument operator. Accurate measurements require alignment of a laser to a small sample core carried by sheath fluid that must be ultra clean to minimize false signals; emitted light is collected along a sensitively aligned optical path. Over the last decade, some of these complications have been overcome through development of automated in-situ flow cytometers such as the CytoBuoy (Dubelaar et al. 1999) and FlowCytobot (Olson et al. 2003). These instruments eliminate the requirement for an operator and sample preparation by automatically running samples that have been collected directly from the environment. Sheath fluid is typically filtered and recirculated with bio-fouling controlled by periodic rinsing with a bio-hazardous chemical. These instruments have been deployed primarily on

*Corresponding author: E-mail: armbrust@u.washington.edu

Acknowledgments

We thank the officers and crew aboard the research vessel *Thomas G. Thompson* for their assistance. We are grateful to D. Schruth for database and data repository design, and implementation of the analytical tools, C. D. Fox for initial work in creating the SeaFlow web interface, A. Djunaedi and S. Mau Ven for running the instrument during the cruise. The research was supported by a National Science Foundation grant and a Gordon and Betty Moore Foundation Marine Microbiology Investigator Award to EVA.

DOI 10.4319/lom.2011.9.466

mooring systems in coastal regions and are providing unprecedented temporal observation of coastal phytoplankton communities (Sosik et al. 2003; Thyssen et al. 2008). Only a few of these instruments have been used to sample while underway (Cunningham et al. 2003; Thyssen et al. 2011), and none of them can detect *Prochlorococcus*.

We describe here development of a new generation of automated flow cytometer, known as SeaFlow, that allows continuous, real-time shipboard observations of phytoplankton cells 0.5-20 µm in size. SeaFlow does not require the use of sheath fluid or alignment of sample and optics. Instead, SeaFlow uses the virtual-core technology (Swalwell et al. 2009) to determine the position of a properly focused particle in a stream of 100 µm prefiltered seawater. With this technology, measurements from particles that pass through the ideal focal position of the collection optics can be differentiated from improperly positioned particles. Measurements from the optimally positioned particles (OPP) are recovered from the data stream to produce a measurement equivalent to that obtained with a conventional cytometer. The ratio of particles within the virtual-core to the total detectable particles is used to retrieve the volumetric flow rate of optimally positioned particles, allowing accurate estimation of cell concentrations. SeaFlow also addresses measurement errors associated with light source instability. A novel imaging-based control loop continuously corrects for measurement error caused by laser pointing instability. For initial instrument setup, the same control loop is also used to automatically align the laser and sample stream to a reference position. The instrument writes the continuous data to a file every 3 min and has mapped 12000 km in 2010, collecting the equivalent of 30000 flow cytometry samples across the oceans. Automated computational methods have been developed to determine cell abundances of the different phytoplankton populations present in each sample (Ribalet et al. 2011). SeaFlow provides multivariate measurements of phytoplankton dynamics at a spatial scale from a few hundred meters to thousands of kilometers that for the first time more nearly matches underway measurements of chemical and physical parameters.

Materials and procedures

Instrument overview

SeaFlow uses virtual-core sheathless flow cytometry optics to enable accurate single particle measurements and cell concentration estimations within a continuous flow of 100 µm prefiltered seawater taken from a ship's flow-through seawater system. A pressure sensor and variable speed gear pump form a control loop to maintain a constant flow rate that enables accurate determinations of cell concentration. Laser drift, which normally hampers the long-term stability of a cytometer, is corrected through an image analysis-based control loop and automated motion stage. The measurement occurs in a stream-in-air to avoid bio-fouling in the measurement region. The 7 analysis parameters include two position-sensitive detectors (PSD), forward scatter in orthogonal and perpendicular polarization states, chlorophyll, a second chlorophyll detector for large particles, and phycoerythrin. All operation aspects of the instrument are controllable through software so can be monitored and adjusted as necessary through a remote connection to the control PC. SeaFlow can be operated remotely via satellite connection, which allows a permanent shipboard deployment and use of the instrument.

Mechanical/electrical

The SeaFlow cytometer is contained within a 5-sided rectangular housing (68 × 49 × 33 cm) constructed from 3/8" aluminum plate (Fig. 1). Electronics and power supplies are

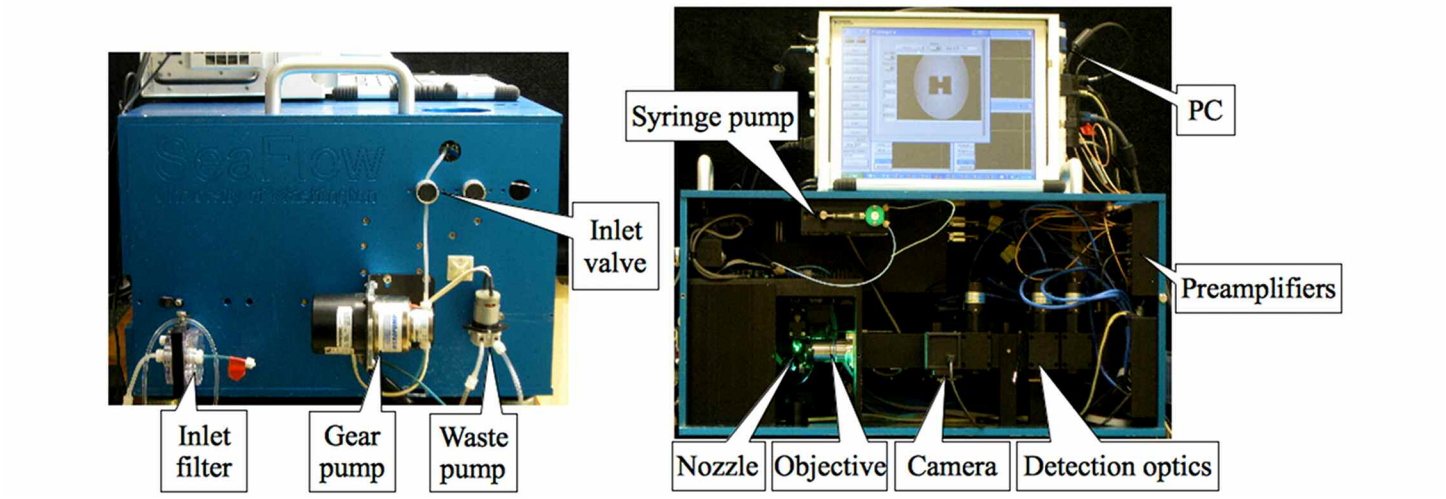

Fig. 1. SeaFlow left side and front views. The inlet filter, gear pump, waste pump, and inlet valve are located on the outside the left side of the aluminum housing. Inside the housing, the laser focusing lens, nozzle, and objective along with positioning stages (hidden) are separated by aluminum walls from the collection optics and electronics. The laser and steering mirror (hidden) are located on the inside of the back and left housing walls, respectively.

mounted to the sides and top within the housing above and away from fluidics. The single optical path and measurement chamber are mounted on the forward edge of the bottom of the housing to provide access to the stream fluidics for setup procedures.

A single switch on the outside the instrument feeds AC power through a bus bar to the two DC power supplies and laser. One of the DC supplies is an ultra low-noise linear supply (Acopian VTD15-160MF) used to power the pressure sensor, the photomultiplier tube (PMT) high voltage supplies (Hamamatsu C6270), and the analog electronics for amplification of PMT signals. A compact modular switching power supply (Exelsys XRB-00, XG3, XG4, XG7) is used to power the gear pump (Micropump GA-T23PFTA), solenoid activated inlet pinch valve (Cole Parmer 98302-54), and motion control components. Total power consumption for the cytometer and data acquisition computer while in operation is 230 watts.

The laser and stream position are each controlled by custom-designed 2 axis motion stages. These stages use miniature stepper motors with integrated lead screws (Haydon Switch, E21H4) to move the custom-designed platforms through two orthogonal directions. The motors are each driven by 1/8-step micro-stepping motor drivers (Allegro A3967) that accept logic level step and direction inputs. Single step resolution with micro-stepping enabled is 0.075 µm. Limit switches are located at the extent of each axis. For the laser-steering prism and lens aggregate, the motion is either side-to-side or up-and-down with respect to the forward scatter collection optics. Longitudinal position of the laser focus is set and permanently fixed by a lens holder composed of threaded barrel and locking ring (Thorlabs, SM1V10). For stream positioning, the two axes control the side-to-side alignment and longitudinal focal position of the stream with respect to the forward scatter optics.

Fluidics

The raw seawater to be analyzed is supplied by a ship's seawater system (Fig. 2). The normally high pressure of the seawater system is stepped down to 5-8 psi by adjustment of a stainless steel ball valve. The water passes through a 100 µm stainless steel mesh filter before entering a gear pump (Micropump GA-T23PFTA). The pump provides a 12 psi input to a 200 µm nozzle (Coulter ML05430) to form the measurement stream. The nozzle is held in a manifold that is, in turn, attached to a custom-designed automated motion stage used for stream alignment. A solenoid actuated pinch valve (Cole Parmer 98302-54) is located between the pump and nozzle to control flow to the nozzle.

The 200 µm nozzle was chosen both for the long-term stability of a larger stream and also to mitigate the likelihood of two cells arriving in the laser beam path at the same time. *Prochlorococcus* is the most abundant phytoplankton in the ocean and can reach cell densities of 10^9 cells L^{-1}. At this concentration, the cells are roughly 200 µm apart assuming they

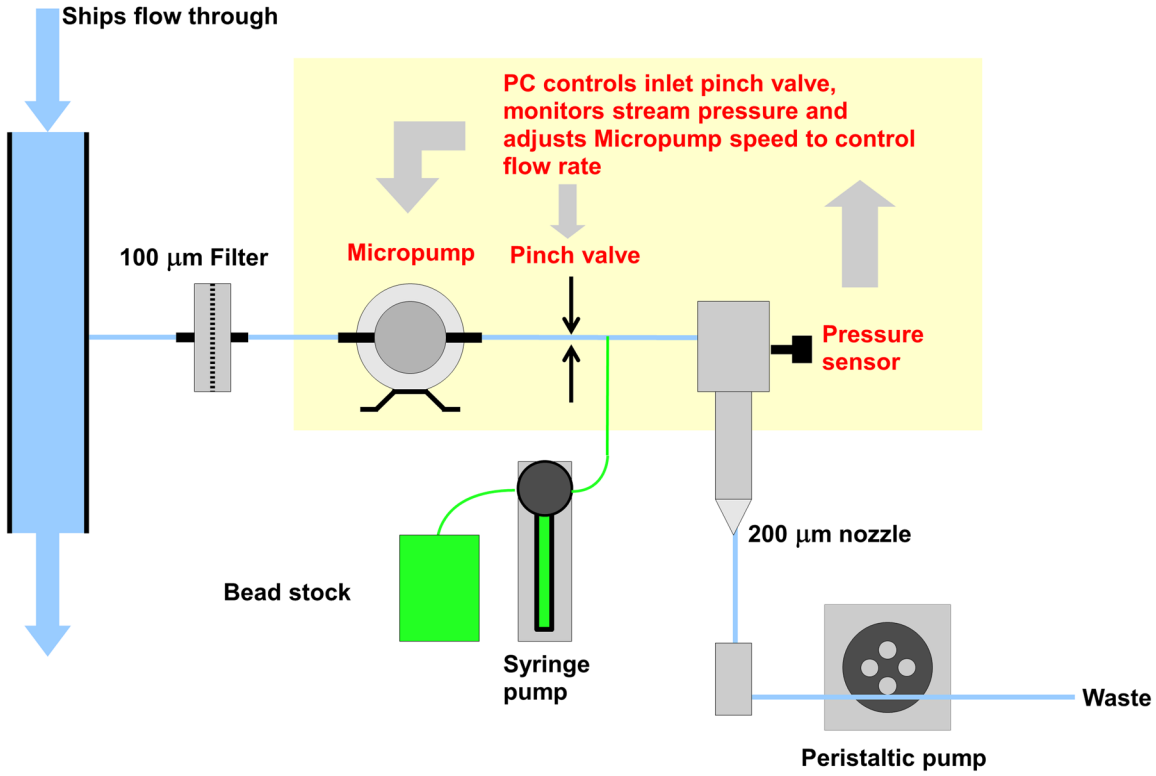

Fig. 2. SeaFlow fluidic system. Seawater is collected from a ship's flow-through system through a 100 µm pre-filter. Stream pressure preceding the 200 µm nozzle is monitored by a PC and held constant by varying the speed of a gear pump. Calibration beads are constantly injected into the sample stream.

are evenly distributed in space. The large nozzle, in combination with the shallow 0.9 µm depth of focus of the 50× objective, produces an optical system with minimal opportunity for coincidence and low sensitivity to particles passing outside the focal plane.

A pressure sensor (Honeywell 26PCDFA6D) monitors pressure between the pump and nozzle to supply a feedback signal to the pump control software and electronics to maintain a constant flow rate. The flow rate of the water stream is usually set at 15 mL min^{-1} to obtain a stable water stream. The pressure sensor in combination with the pump speed control setting is used to monitor a system leak or clog.

To assess drift in optical alignment, a syringe pump (Cavro XP3000) continuously injects 1 µm calibration beads (Invitrogen F8823) into the seawater flow ahead of the nozzle. The measurement stream is captured by a cone-shaped cup and pumped to waste by a peristaltic pump (Thomas SR10).

Optics

A 300 mW 457 nm diode pumped solid-state laser (Melles-Griot 85-BLS-601) is used for fluorescence excitation and light scattering (Fig. 3). The 1.16 mm diameter laser beam is diverted through a U-shaped path and focused in the center of the measurement stream first by a mirror followed by an aggregate prism lens assembly attached to a custom-designed automated motion stage used for laser alignment and laser drift compensation. The focal length of the spherical laser focusing lens is 75 mm (Edmund Optics), which for the Gaussian (TEM00) laser beam intensity profile yields a theoretical minimum spot diameter (1/e^2) of 32 µm.

Forward-scattered light and fluorescence emitted in the forward direction are both collected through a 50× long working distance objective (Mitutoyo 378-805-3) with a shallow 0.9 µm depth of focus; the diffraction limit of the optical system is 0.5 µm. The parallel rays are then focused by a 150 mm tube lens (Edmund Optic 49-362) onto the PSD system located at the image plane. The PSD consists of two diverting prisms mounted to the back side of a mirrored glass surface. At the center of the mirrored surface, there is an H-shaped aperture that functions as both a field stop for stray light and the active regions for the position detection and forward direction light collection. Light cast onto either of the two position sensitive regions of the H aperture is diverted by prisms to one of two photomultiplier tubes (Hamamatsu R9110). Signals detected from the position sensitive regions are used to determine particle position within the stream of seawater and to filter out improperly positioned particles from the data stream. Light rays focused onto the H aperture not covered by the position sensitive regions diverge beyond the field stop and are brought parallel by a 75 mm spherical lens (Edmund Optic 47-639). The parallel light rays are diverted through a system of

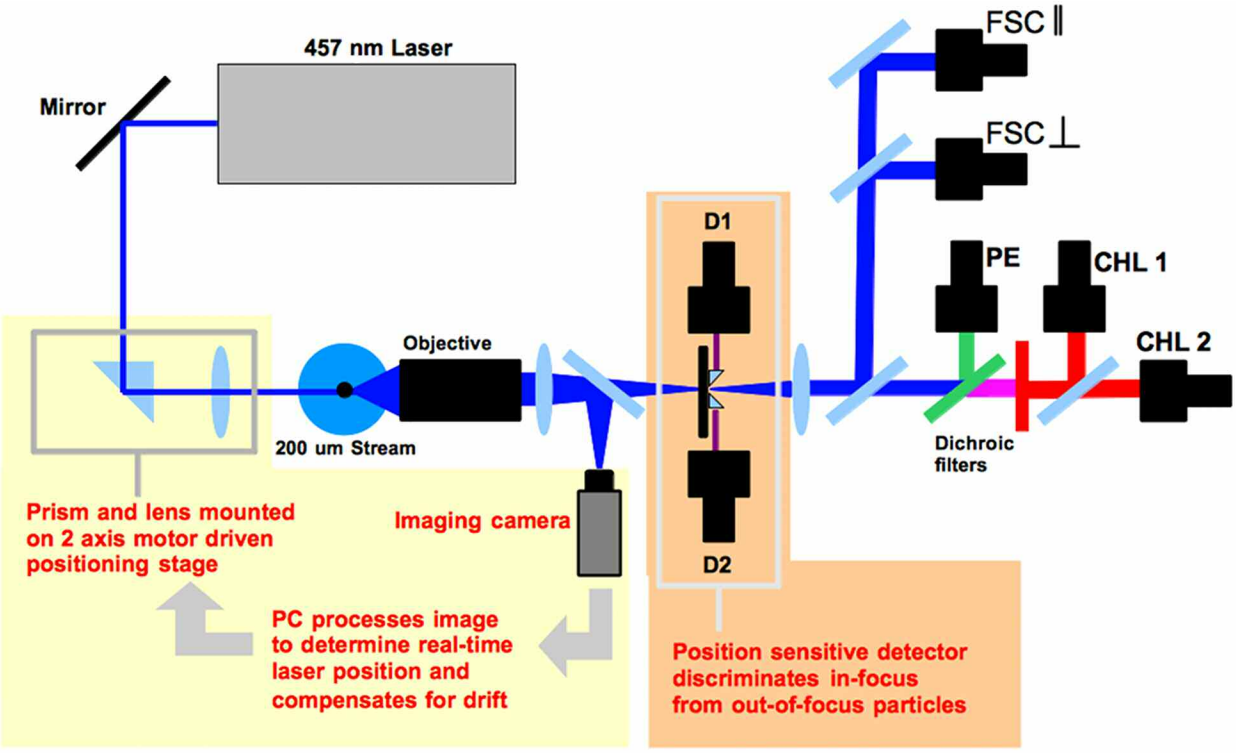

Fig. 3. SeaFlow optical system. The laser is steered by a prism and focusing lens. These elements are mounted together and positioned by an automated 2-axis motion stage. PC control of this motion stage, along with image processing of the laser position on the stream, allows for auto-alignment of the system and subsequent laser drift tracking and compensation. The position of particles passing in the stream is measured by the position sensitive detector. Scattered and fluorescent light passing through the center of the H aperture is measured by a system of beam splitters and dichroic filters.

dichroic filters and beam splitters to photomultiplier tubes (Hamamatsu R9110). The chlorophyll fluorescence signal for large and small cells is collected using a 692-40 bandpass filter (Semrock FF01-692/40-25) and the phycoerythrin fluorescence signal is collected using a 572-28 bandpass filter (Semrock FF01-572/28-25). The 457 nm scatter signal is diverted by a 45° dichroic long pass filter (Semrock Di01-R442) to a band pass filter (Semrock FF01-457/50-25). The parallel and perpendicularly polarized scatter signals are picked-off by microscope cover slips oriented at Brewster's angle. The two orthogonal polarization detectors are used to detect the polarization rotation of scattered light caused by the calcite cell walls of coccolithophores (Olson et al. 1989).

A color USB CCD camera (Imaging Source DBK31BU03) is set to a fixed exposure value and used to produce an image of the measurement stream. The camera images a microscope cover slip that is oriented at a 45° angle to the converging rays of the tube lens. The acquired color image is separated into its red, blue, and green color channels with each used for different analysis. When the stream is illuminated by the green LED back light, the green image channel is used to view and perform initial alignment of the stream to the optical system. During normal operation, the blue image channel is used to measure the location of the blue laser on the stream for automated laser drift correction. Particles excited by the blue laser that emit red fluorescence will be observed in the red image channel. This red image is a summation over the shutter opening of all passing red-light emitting particles and is therefore a measure of bulk red fluorescence within the seawater. The bulk red fluorescence is calculated by averaging the red pixel values across the horizontal and vertical to produce a proxy of bulk chlorophyll fluorescence.

Instrument control

The SeaFlow user-control software was written in the Labview programming language, which facilitates rapid and flexible programming of data acquisition systems for PC-based instrumentation. The custom written SeaFlow software is accessed through a single front panel. From this, the user can open sub panels for control of conventional cytometer settings including photomultiplier tube gain, measurement channel naming, and data storage options as well as access to the setup and control features unique to SeaFlow. The details of these unique features are discussed in this section.

The pump and stream pressure control panel allows the user to enable a proportional integral differential (PID) control loop that continually analyzes the stream pressure and will perform a correction to the pump voltage control signal to maintain a constant pressure set point (Fig. 2). When the PID control loop is enabled, the software monitors for two potentially serious fault scenarios in the fluidics system—a clogged nozzle orifice or a leak in the system. In both cases, the pump is shut down and the inlet valve is closed to protect the instrument. A clog is detected because the pump voltage set point will eventually fall below a predetermined level while the pressure set point is being maintained by the PID loop. A leak in the system is detected because the pump voltage set point will rise above a predetermined limit, but the pressure will remain below the set point.

Images from the stream visualization camera are presented to the user through a software interface. The software allows the user to enable a custom-built LED back light and control circuit to aid in alignment of the stream. A homing routine assists the user in initial setup of the stream and laser: each positioning stage is moved to the extent of its axis until the limit switch is tripped and then the axis is moved back to a calibrated position. With the back light off and the laser on, the user can use the motion control software panel to move the laser into position. So long as the laser intersects the stream and is in view of the imaging camera, an automated focusing routine can be enabled to position the laser into an ideal alignment with the collection optics.

After the stream and laser have been aligned, the software will enable the automatic laser drift correction. Initially, the laser drift correction algorithm computes two arrays that are the average of the horizontal and the vertical pixels acquired from the USB camera. These arrays are then averaged with the arrays computed from the next eight video frames. The averaged arrays are used as the reference arrays for drift correction. A running average of eight image frames is continuously cross-correlated to the reference arrays. The output of the cross correlations is an array of numbers that describes both the similarity of the two inputs and the likelihood of a shift existing between the input arrays. Generally, a white image against a dark background is formed when the laser intersects the stream. Differences in particle content in the seawater change only the intensity of the image, so the reference array and the running average arrays are always comparable. The direction of the array shift calculated by the cross correlation is used in a software-implemented proportional loop to command the laser motion stage to move a small amount one direction or the other, vertically or horizontally. In this manner, the laser drift algorithm constantly detects and corrects for the pointing drift of the laser.

Data acquisition

SeaFlow uses a modular National Instruments (NI) PXI bus based data acquisition system for all acquisition and control tasks. The PXI chassis (NI 1042Q) contains a Windows-based controller (NI PXI-8106), a simultaneously sampling 8-channel 12-bit programmable 1-60 MHz digitizer (NI PXI-5105) for acquisition of PMT signals, an 8-channel analog output module (NI PXI-6704) for control of PMT voltage settings and Micropump speed, and a multipurpose analog and digital input/output module (NI PXI-6229) for control of the positioning systems, valves, and pressure sensor voltage acquisition.

Light pulses produced by particles passing through the interrogation area are detected by the photomultiplier tubes and converted to voltage signals by custom built logarithmic preamplifiers similar in design to those described by Shapiro

(2003). Acquisition is started when a trigger channel, normally the PMT used for chlorophyll fluorescence detection, exceeds a threshold value. When this occurs, data from all detection channels are recorded by the digitizer. The measurement from all detection channels initiated by a single trigger is collectively referred to as an "event." The digitizer is programmed for a 5 MHz sampling rate with 100 point acquisition of all 8 input channels. A 3-point smoothing filter is applied to the digitized pulses before determining the maximum value of the pulse. While the instrument is under full operation, the maximum measured event rate achieved before over-writing the on-card memory buffer is ~24,000 events per second.

Pulse height measurements for each channel of an event are appended with a system time stamp and stored in a RAM buffer. This data are written from the buffer to a binary file and stored on hard disc in 3-min intervals. The files are numbered sequentially and given the .evt file extension. The sequence of files, numbered 1-480 for a full day of data, is kept in a directory named by year and Julian day. At midnight GMT a new Julian day file is written and the new day's .evt files are written to this file.

An ASCII text log file is written in each Julian day directory. GPS position, the universal time constant (UTC), and any other data collected through serial NMEA or from a ship's network (salinity, bulk chlorophyll, temperature, etc.) is appended to this file in 3-min intervals along with the current .evt file number. Because research vessels do not maintain a common format or communication protocol for broadcasting data, the SeaFlow software component used to acquire this information is normally customized on-site before departure. The log file is used in post processing to geo-reference the cytometry and ship's data.

Data analysis

The Seaflow instrument control software panel allows a user to display four bivariate dot plots to view the data as soon as it is acquired and processed. A function performs OPP gating in real time (see below for more details) and may be enabled or disabled through the dot plot panel for set-up purposes. Because the cytometry data are collected in 3 min intervals, there is also a function that will allow a user to "play" those files sequentially to create an animation of the data contained within a day's data folder.

To provide rapid post-cruise analysis, software tools have been developed for automated analysis and visualization of measured phytoplankton populations. A software package called flowPhyto (Ribalet et al. 2011), written in the statistical computing environment R, provides a collection of functions that turn the high volume of raw SeaFlow files into a summary data table and plots. Multivariate gating and statistical clustering methods (Lo et al. 2009; Aghaeepour et al. 2011) automate identification of phytoplankton populations, reducing the significant subjectivity and human-time cost that would otherwise be associated with the analysis of SeaFlow data. For each population, aggregate statistics (e.g., cell count, mean and standard deviation of light scattering and fluorescence) are performed if a sufficient number of cells (300 by default) are present within a file. Otherwise, a per-population concatenation of consecutive files is performed to meet the 300-cell requirement. The concatenation of files reduces the temporal resolution but improves the robustness of aggregate statistics.

The instrument generates the equivalent of 6700 samples after a typical 2-week long oceanographic cruise, representing a data set 35 to 135 GB depending on ambient cell concentrations. About 3-6 h of computation is typically required to process the raw data and produce phytoplankton distribution maps. Cruise data are stored in a database where the user can browse and visualize the data via a web interface available at seaflow.ocean.washington.edu. Aggregate data can be downloaded as a table or directly imported into Google Earth (Google Inc.) for 3-D visualization.

Assessment

Flow cytometric measurements

The challenge when using a raw stream of seawater is to discriminate between particles that are aligned and focused (Optimally Positioned Particles) and those that are not properly positioned. Focused particles lie within the image plane of the objective while aligned particles are located along the axis of the objective and laser. Aligned particles will scatter the light equally on the two position detectors (D1, D2) and therefore lie along the 1:1 diagonal axis of a bivariate plot of D1 versus D2 when the gains of D1 and D2 have been calibrated to respond equally to an aligned particle. Scattered light from focused particles will arrive at the center of the H aperture in which case a maximal signal is detected by the forward scatter detector (FSC) and a minimum by both position detectors. When the FSC and position detector gains are adjusted to respond equally to in-focus calibration particles, the FSC signal of unfocused particles will be lower than the signal on the position detectors. This is because the intensity profile formed on the pinhole image plane for an unfocused particle is less peaked than the diffraction-limited image of a focused particle. By normalizing the position detector signals by the FSC, unfocused particles can therefore be discriminated. It follows that two criteria define the OPP filtration:

1) the light scatter signal must be equal on D1 and D2

2) the light scatter on the FSC detector cannot be lower than on the two position detectors.

In traditional flow cytometry, the volume of the sample core is machine- and setup-dependent, determined by hydrodynamic forces within the injection system and nozzle. In virtual-core cytometry, the volume of the sample core is defined by the range of positions that are contained within the OPP filter. The first OPP criterion is overly restrictive in that it requires all particles to be perfectly aligned. This perfect alignment is never achieved in traditional flow cytometry where the accuracy of the sample injection system determines the

volume containing all sample trajectories. The OPP criteria were modified to include an acceptable error rate.

In practice, the cytometric gate for OPP discrimination is defined on a bi-variate dot plot of the two position detectors scaled by the forward scatter detector. The width of the gate to the sides of the 1:1 equal detector response defines the allowed error in particle trajectories across the width of the stream. The limit of FSC/D ≤ 1 is a filter for particles that lie within the depth of focus of the objective and depends on proper calibration of FSC and position detectors. The width of the gate can be adjusted to maximize the signal of the light scatter and fluorescence while minimizing the coefficient of variation of standard beads. Gate width can be minimized to the theoretical 1:1 equal detector response with the result of increased measurement accuracy but also a reduction in volume analyzed. This is analogous to traditional flow cytometery where the differential pressure between sample and sheath can be minimized to improve injection accuracy but also reduces the sample flow rate. The volume of the virtual-core defined by OPP gating is therefore flexible and may be adjusted based on the study being carried out. Decreasing the volume of the virtual-core size will increase measurement accuracy with a possible trade-off in acquisition of enough events to characterize rare populations. Unlike traditional flow cytometry where this decision is made by an operator at the time of the measurement, the trade-off between core size and accuracy is made in post-processing of the data set so is inherently reversible.

Examples of cytometric measurements before and after the virtual-core gating are shown in Fig. 4. The gate discriminates between all detectable cells and the ~1% that are correctly positioned for light-scattering measurements. The OPP gating recovers the general characteristics of the uniform beads, with coefficient of variations of fluorescence signals typically < 2%. Similar results are obtained using the virtual-core gating to measure more complex, environmental samples. A sample of seawater taken at 5 m depth in the subtropical Pacific Ocean revealed a dense population of *Prochlorococcus* whose chlorophyll fluorescence characteristics were well above detection limit.

Calibration of the Virtual Core to determine cell concentrations

Only a fraction of the 200 µm wide stream is analyzed by the instrument. This detectable region is defined as the cross-sectional area within the stream where particles are detected and has a fixed width set by the field of view of the optical system (Fig. 5). This width is dependent on the magnification of the objective and the width of the field stop (56 µm for the SeaFlow). The volume of the detectable region is calibrated by analyzing known concentrations of beads and comparing the SeaFlowbead count to the known concentration, Vd = Pd/Cs, where Vd = volume of detectable region, Pd = bead count, and Cs = known concentration of beads. This calibration step need only be done once. For an unknown sample, cell concentration is determined by dividing the cell count by the volume of the detectable region, Ct = Pd/Vd, where Ct = total cell concentration.

To determine cell concentrations of different populations, the volume of the OPP region (called the Virtual Core) is first calculated by multiplying the volume of the detectable region by the ratio of OPP particles to total detectable particles, Vvc = Vd × Pvc/Pd, where Vvc = volume of the virtual core and Pvc = particle count within the virtual core.

The number of cells within a given population divided by the OPP volume is the concentration (cells/mL), Cn = Pn/Vvc, where Cn = concentration of a cell population and Pn = cell count within a population.

The flow rate of the stream is usually set up at ~15 mL per min. Under proper laser and stream alignment, the flow rate of the detectable particles is ~1/10 of the flow rate of the stream while the flow rate of OPP particles within the virtual-core is ~1/100 of the flow rate of the detectable particles. Therefore, the flow rate of the virtual-core is ~15 µL per min, which is consistent with conventional flow cytometers. The virtual-core is then ~6 µm diameter for the 200 µm stream (assuming the virtual-core has a circular cross-section). The virtual-core is not a bounding volume for particles but an allowed error in position. Therefore, alignment of OPP particles deviates by ± 3 µm from the ideal trajectory.

Light scatter is a complex phenomenon and a question arises as to whether the sensitivity of the PSD is consistent across all anticipated particle sizes and shapes. To address this, we analyzed replicate samples of various size range particles with SeaFlow and a BD Influx (BD Biosciences), a conventional instrument that uses hydrodynamic sample focusing. We analyzed beads of different sizes (1, 6, and 10 µm in diameter) and various laboratory cultures, namely *Synechococcus* (2 µm diamter), *Skeletonemacostatum* (4-6 µm width), *Phaeodactylumtricornutum* (3 µm width), *Thalassosirapseudonana* (4-5 µm width), *Chaetocerossocialis* (10 µm width), because cells in this size range can reliably be analyzed by both instruments. Using the concentration calibration method described above, particle concentrations obtained from SeaFlow were in good agreement (Pearson's coefficient of correlation $r = 0.93$) with those from the BD Influx (Fig. 6a).

We evaluated the performance of the instrument in underway ship-board use on three cruises: the Subarctic Pacific Ocean on 1-9 Jun; in Puget Sound, Washington, USA on 7-9 Nov 2009; and near the Hawaiian Islands (U.S.A) on 28-30 Dec 2010. Samples were collected for analysis by a BD Influx and compared with samples analyzed with SeaFlow. The Pearson's coefficient of correlation, $r = 0.84$, indicates a good agreement between counts obtained with the two instruments (Fig. 6b). Analysis of natural samples indicates that cell concentrations from SeaFlow were reliable for different sizes of cells and cell concentrations up to at least 2.5×10^8 cells L^{-1} (Fig. 6b), a very high concentration for marine phytoplankton, which corresponds to a counting rate of ~10000 cells per second. Along

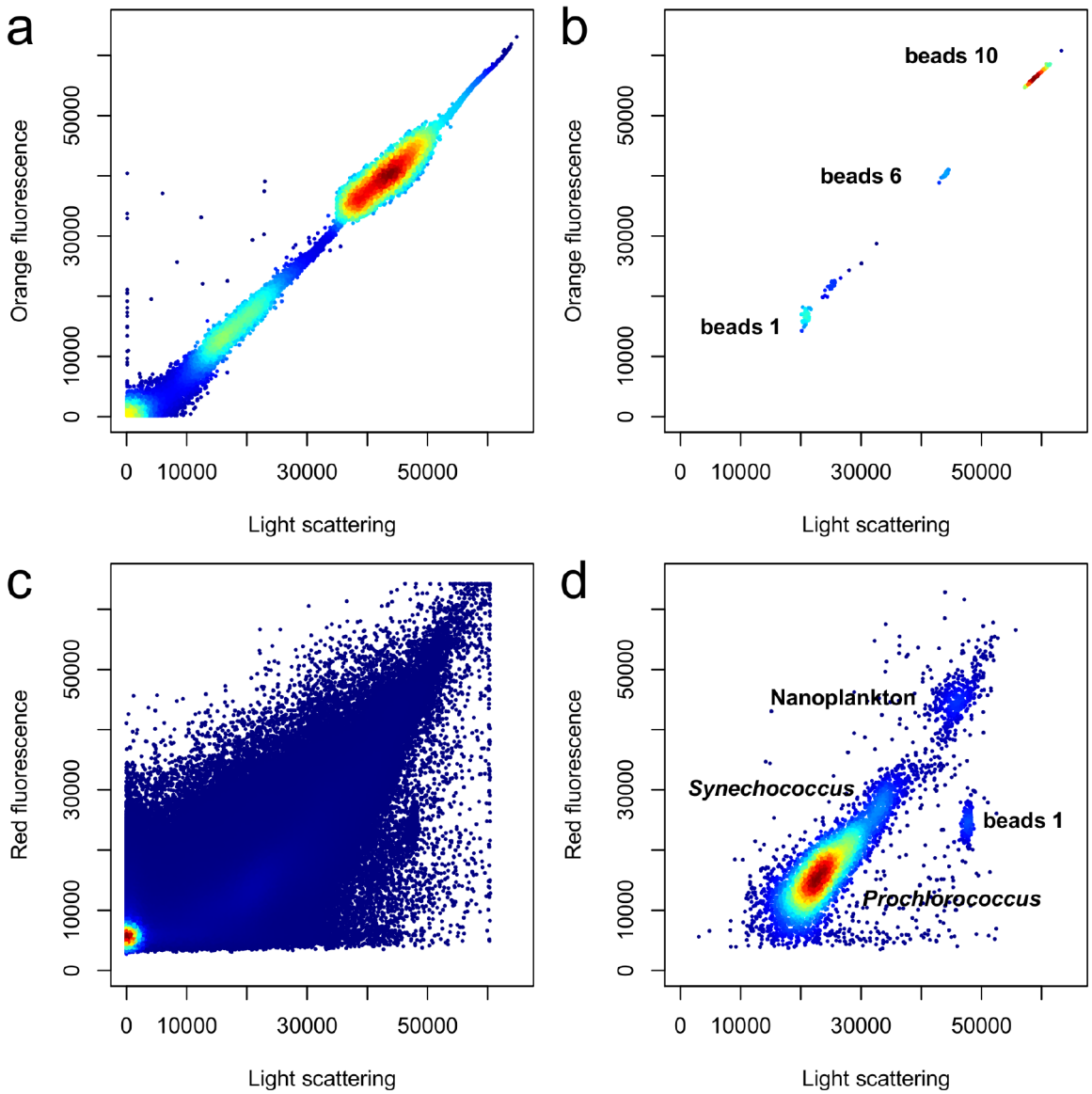

Fig. 4. Flow cytometric measurements of light scattering and fluorescence of 1, 6, and 10 µm beads (a,b) and seawater sample from the subtropical Pacific Ocean (c,d). Detectable particles (a,c) are used to calculate cell concentrations, whereas Optimally Positioned Particles (b,d) are used to identify the different cell populations. Warmer color represents higher abundance.

the 450 km transect near the Hawaiian Islands, the 0.5-20 µm size fraction of the phytoplankton communities in surface waters was composed mainly of *Prochlorococcus*, *Synechococcus*, and small nanoplankton populations (4 µm nominal mean cell size) that showed distinct light scatter and fluorescence characteristics. *Synechococcus* and nanoplankton populations displayed changes in cell concentrations that varied over ~ 1 order of magnitude (Fig. 7). Fine-scale patterns of phytoplankton distribution were observed across the transect, such as the steep increase in cell concentrations of *Synechococcus* observed over a 10-km-wide region (Fig. 7b, left panel). A frequency of one discrete sample every 15 min would have been required to detect this steep increase, which would have been challenging to capture with operator dependent sampling.

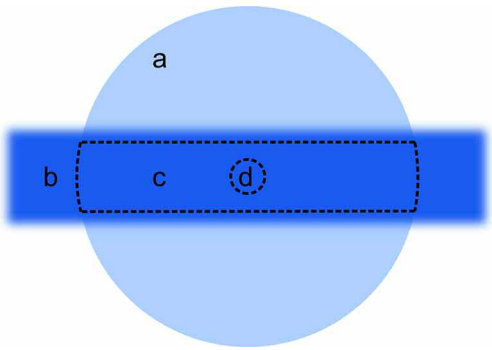

Fig. 5. Cross-section view of 200 µm stream (a) with representations of the intersection with laser beam (b), the detectable region (c), and the Virtual Core (d).

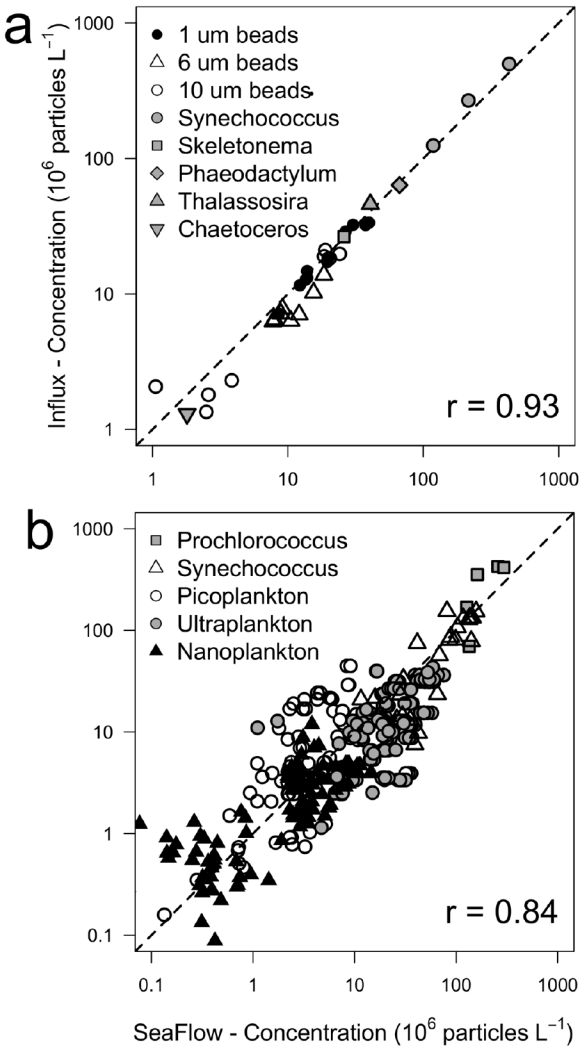

Fig. 6. Comparison of cell counts. (a) Concentrations of 1, 6, and 10 µm beads, *Synechococcus* (2 µm diameter), *Skeletonemacostatum* (4-6 µm width), *Phaeodactylumtricornutum*(3 µm width), *Thalassosirapseudonana* (4-5 µm width), and *Chaetocerossocialis* (10 µm width) measured with SeaFlow were compared with those measured with a BD Influx flow cytometer. Samples were taken directly from the SeaFlow outflow and analyzed immediately with the Influx. (b) Concentrations of *Prochlorococcus* (0.5 µm nominal mean cell size), *Synechococcus* (2 µm nominal mean cell size), picoplankton (2 µm nominal mean cell size), ultraplankton (3 µm nominal mean cell size), nanoplankton (4 µm nominal mean cell size) in seawater. Discrete samples (n = 275) were taken directly from the SeaFlow outflow and fixed with 1% paraformaldehyde and 0.01% glutaraldehyde, stored in liquid nitrogen, and analyzed with the Influx 2 months later. The Pearson's coefficient of correlation r = 0.93 and r = 0.84 in both cultures and environmental samples indicates a good agreement between counts obtained with the two instruments. The dash line represents the 1:1 slope.

Discussion

SeaFlow is able to continuously map the distribution and abundance of phytoplankton communities in the open ocean at scales that match underway chemical and physical measurements such as temperature and salinity. SeaFlow's ability to provide high-resolution spatial and temporal distributions of *Prochlorococcus* in surface waters is unprecedented. Abundances and dynamics of *Prochlorococcus* populations have a substantial impact on open ocean ecosystems and global biochemical cycles. Continuously expanding SeaFlow data sets provide a platform for the development and testing of hypotheses about the selective forces that shape phytoplankton abundance and distributions and how these may change over different temporal and spatial scales.

A benefit of SeaFlow is that accurate cell concentrations can also be determined for populations at low cell concentrations (such as 10^6 cells L^{-1}). With a conventional flow cytometer, statistically robust quantification of these populations requires analysis over relatively long periods of relatively large volumes, during which time some large cells and calibration particles may sink to the bottom of the sample tube skewing cell counts. With SeaFlow, samples are analyzed continuously, and signals can be integrated over as long an integral as necessary to provide optimal aggregate statistics, with a subsequent reduction in the spatial resolution. This fundamental difference in analysis likely explains the observed differences in the SeaFlow and Influx count data (Fig. 6a and 6b). In these comparisons, samples on the Influx were analyzed at a flow rate of 50 µL per min for 5 min, which means that only 250 cells from rare populations were counted, which is not enough to provide unambiguous clustering of cell populations. SeaFlow also uses a 457 nm excitation wavelength to optimize detection of *Prochlorococcus*, rather than the more conventional 488 nm excitation wavelength as on the Influx, which may have also contributed to differences in clustering patterns and cell counts.

Eventual installment of SeaFlow on multiple ships will provide fine-scale resolution across broad expanses of the ocean. This data can be used to both complement and fill gaps in satellite images of chlorophyll *a,* which is commonly aggregated, potentially obscuring interesting community features. In addition, SeaFlow provides greater details of phytoplankton community composition than provided by satellite images.

A key feature of SeaFlow is its ability to operate semi-autonomously. This has allowed us to operate the instrument on the R/V *Thomas G. Thompson* with minimal training for the on-board operator. A remote connection is made to the SeaFlow once or twice a day to view the current data and make adjustments as necessary, or suggest to the on-board technician more invasive intervention when necessary. To date, this instrument has provided measurements over more than 12,000 kilometers.

Comments and recommendations

The stream-in-air measurement has inherently good optical properties and is impossible to biofoul. However, the stream is sensitive to bubbles in the inlet line and rapid fluctuations in the ship's flow-through system that can occur in rough seas.

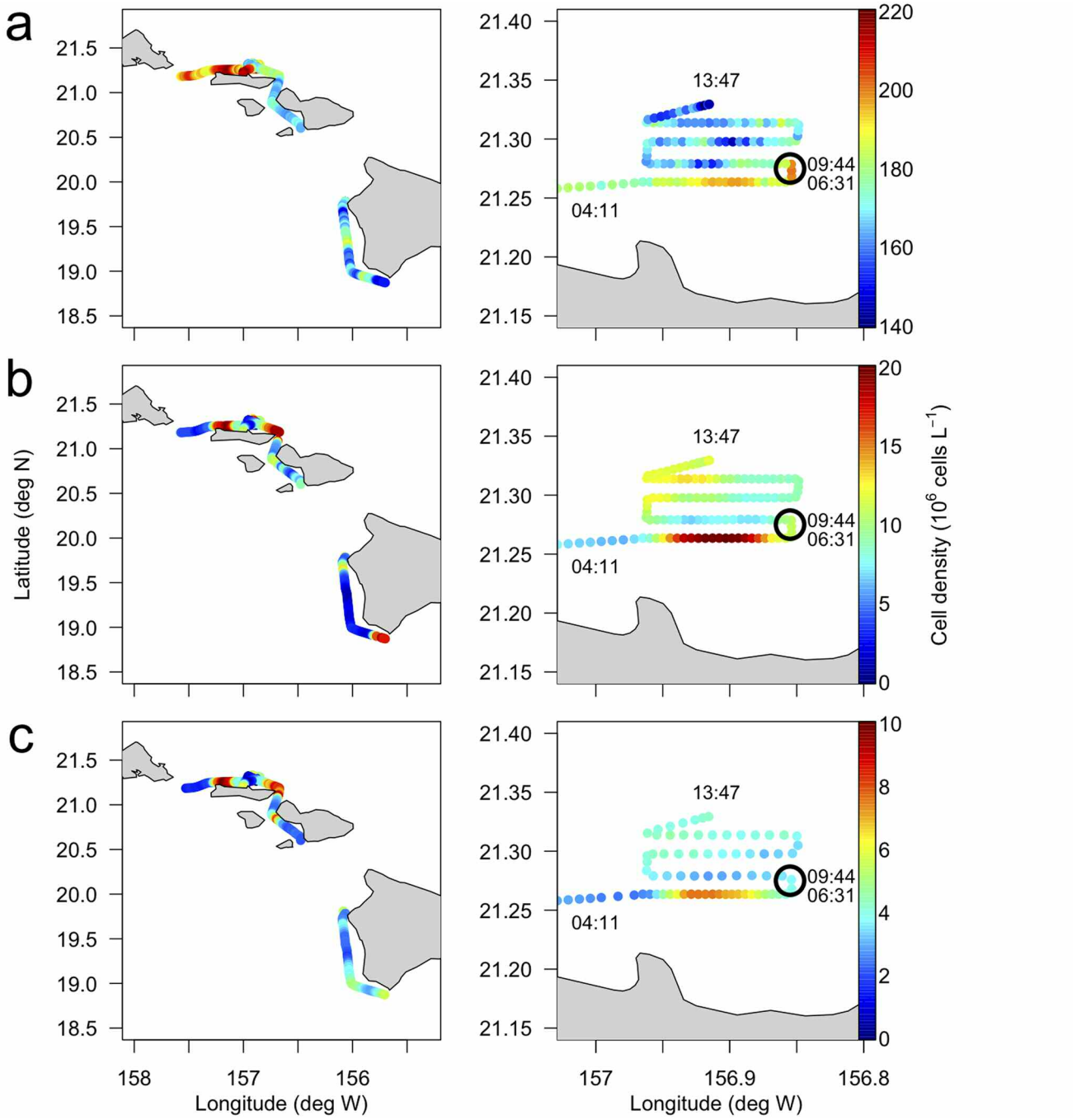

Fig. 7. Distribution of (a) *Prochlorococcus* (0.5 µm nominal mean cell size), (b) *Synechococcus* (2 µm nominal mean cell size), and (c) nanoplankton (4 µm nominal mean cell size) populations near Hawaiian Islands, USA, 27-31 Dec 2010, at large spatial scale (left panels) and small spatial scale (right panels). The black circle indicates the location of a 3-h long station.

In addition, long periods of use causes salt build-up on the catch funnel and objective face necessitating periodic maintenance from shipboard crew. For more autonomous deployments such as shore-powered moorings, the instrument will need to be adapted to use a closed plumbing system that includes a flow cell in the interrogation region similar to other autonomous cytometers.

The maximum event rate is 24,000 events per second, which allows us to study very dense phytoplankton populations such as *Prochlorococcus*. Currently, the instrument digitizes waveforms of PMT pulses and processes those pulses on board the PC. To maintain high throughput, processing has been limited to smoothing and peak height analysis only. An alternative approach would employ analog electronics by

means of peak hold or integrating amplifiers, the method typically used in most commercially available instruments. This modification would increase the bandwidth 100 fold but close the door to using pulse processing analysis for large cells (Dubelaar and Gerritzen 2000; Zilmer et al. 1995). A hybrid system of pulse processing of large cells and analog peak detection for the smaller fraction would be ideal.

The bulk fluorescence measurement produced by the red channel of the USB camera could be calibrated to produce a more accurate estimate of chlorophyll concentrations. Because the SeaFlow measurement is not able to quantify cells whose scatter signal is large enough to saturate the PMTs used in the PSD, a bulk chlorophyll measurement could be used to estimate the chlorophyll fraction that SeaFlow was not able to measure. This would entail summing the event values recorded by SeaFlow's chlorophyll fluorescence channel and comparing this to the measured bulk value. A significant deviation in the correlation of these two measurements could indicate a missed size fraction from SeaFlow's chlorophyll fluorescence channel.

As SeaFlow is further refined to become increasingly autonomous, adaptive sampling based on observed changes in community structure will be possible. The ability to sample at positions along a transect where the most significant biological gradients are observed will provide new insights into microbial structure and functions. The tools required to interpret a real-time shift in community structure will need to be developed and tested. Simple and obvious cytometry metrics such as rise or fall in event rate, or the event rate associated with a particular population have been tested. A more complex analysis based on the cytometric diversity measurement developed by Li (1997) has also been run on the SeaFlow data in real-time. Regardless of the measure, determination of a significant biological signal "in the moment" is challenging primarily because the candidate signaling algorithm must be predictive. At present, useful interpretations of measures based on high-resolution cytometry, alone or correlated to other shipboard data streams, do not exist. Data mining of current and future SeaFlow data sets will help in development of these predictive tools.

References

Aghaeepour, N., R. Nikolic, H. H. H. Hoos, and R. R. Brinkman. 2011. Rapid cell population identification in flow cytometry data. Cytometry Part A 79A:6-13 [doi:10.1002/cyto.a.21007].

Cunningham, A., McKee, D., Craig, S., Tarran, G. and Widdicombe, C. 2003. Fine-scale variability in phytoplankton community structure and inherent optical properties measured from an autonomous underwater vehicle. Journal of Marine Systems 43, 51–59 [doi:10.1016/S0924-7963(03)00088-5].

Dubelaar, G. B. J., P. Gerritzen, A. E. Beeker, R. Jonker, and K. Tangen. 1999. Design and first results of CytoBuoy: A wireless flow cytometer for in situ analysis of marine and fresh waters. Cytometry A 37:247-254 [doi:10.1002/(SICI)1097-0320(19991201)37:4<247::AID-CYTO1>3.0.CO;2-9].

———, and P. L. Gerritzen. 2000. CytoBuoy: a step forward towards using flow cytometry in operational oceanography. Sci. Mar. 64:255-265.

Li, W. K. W. 1997. Cytometric diversity in marine ultraphytoplankton. Limnol. Oceanogr. 42:874-880 [doi:10.4319/lo.1997.42.5.0874].

———. 2002. Macroecological patterns of phytoplankton in the northwestern North Atlantic Ocean. Nature 419:154-157.

Lo, K., F. Hahne, R. Brinkman, and R. Gottardo. 2009. flowClust: a Bioconductor package for automated gating of flow cytometry data. BMC Bioinformatics 10:145 [doi:10.1186/1471-2105-10-145].

Olson, R. J., E. R. Zettler, and O. K. Anderson. 1989. Discrimination of eukaryotic phytoplankton cell types from light scatter and autofluorescence properties measured by flow cytometry. Cytometry 10:636-643 [doi:10.1002/cyto.990100520].

———, A. Shalapyonok, and H. M. Sosik. 2003. An automated submersible flow cytometer for analyzing pico- and nanophytoplankton: FlowCytobot. Deep Sea Res. I 50:301-315 [doi:10.1016/S0967-0637(03)00003-7].

Ribalet, F., D. M. Schruth, and E. V. Armbrust. 2011. flowPhyto: enabling automated analysis of microscopic algae from continuous flow cytometric data. Bioinformatics 27 (5): 732-733 [doi:10.1093/bioinformatics/btr003].

Shapiro, H. 2003. Practical flow cytometry. Wiley Online Library [doi:10.1002/0471722731].

Sosik, H. M., R. J. Olson, M. G. Neubert, A. Shalapyonok, and A. R. Solow. 2003. Growth rates of coastal phytoplankton from time-series measurements with a submersible flow cytometer. Limnol. Oceanogr. 48:1756-1765 [doi:10.4319/lo.2003.48.5.1756].

———, R. Olson, and E. Armbrust. 2010. Flow cytometry in phytoplankton research. chlorophyll a fluorescence in aquatic sciences: methods and applications. 4:171-185.

Swalwell, J. E., T. W. Petersen, and G. Van Den Engh. 2009. Virtual-core flow cytometry. Cytometry A 75A: 960-965 [doi:10.1002/cyto.a.20792].

Thyssen, M., D. Mathieu, N. Garcia, and M. Denis. 2008. Short-term variation of phytoplankton assemblages in Mediterranean coastal waters recorded with an automated submerged flow cytometer. J. Plank. Res. 30:1027-1040 [doi:10.1093/plankt/fbn054].

———, and M. Denis. 2011. Temporal and spatial high-frequency monitoring of phytoplankton by automated flow cytometry and pulse-shape analysis. Global Change: Mankind-Marine Environment Interactions 5:293-298

Vaulot, D., D. Marie, R. J. Olson, and S. W. Chisholm. 1995. Growth of *Prochlorococcus*, a photosynthetic prokaryote, in the equatorial Pacific Ocean. Science 268:1480-1482

[doi:10.1126/science.268.5216.1480].
Zilmer, N., M. Godavarti, J. Rodriguez, T. Yopp, G. Lambert, and D. Galbraith. 1995. Flow cytometric analysis using digital signal processing. Cytometry 20:102-117 [doi:10.1002/cyto.990200203].

Submitted 16 March 2011
Revised 16 August 2011
Accepted 30 August 2011