2017

Assessment of Fluid Imaging for Determination of Diatom Assemblage Composition and Biometrics of Southern Ocean Sediment

Neil Redmond
University of Rhode Island, neil_redmond@uri.edu

ASSESSMENT OF FLUID IMAGING FOR DETERMINATION OF DIATOM ASSEMBLAGE COMPOSITION AND BIOMETRICS OF SOUTHERN OCEAN SEDIMENT

BY

NEIL REDMOND

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

OCEANOGRAPHY

UNIVERSITY OF RHODE ISLAND

2017

MASTER OF SCIENCE

OF

NEIL REDMOND

APPROVED:

Thesis Committee:

Major Professor Rebecca Graham

John King

Simon Engelhart

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2017

ABSTRACT

Diatoms are important ecologic indicators whose assemblage, chemistry, and valve features are reflections of their original environmental conditions. Fossil diatom biometrics are an emerging measurement introduced to supplement our understanding of the hydrographic history of the Southern Ocean. Here, we present a novel method to simultaneously measure fossil diatom assemblage and biometrics using a FlowCam, an instrument combining features from a flow cytometer and microscopic camera. It offers, computerized automatic identification to supplement manual, visual identifications, leading to increased counts and biometric measurements. To assess the viability of the FlowCam as a paleoceanographic tool, a FlowCam measured data set was compared to previously published diatom assemblage and biometric data generated by traditional microscopic methods from a Southern Ocean sediment core. Diatom assemblages and the biometric lengths of *Fragilariopsis kerguelensis* measured with the FlowCam showed similar trends to those produced by traditional microscopy. The biggest difference was the relative occurrence of *Eucampia antarctica,* which was observed more frequently using the FlowCam. The high biometric data output from the FlowCam was used to determine an empirically derived, minimum sample count and confidence intervals for future best practices.

ACKNOWLEDGMENTS

For all of their love and support, I would most like to thank my mother Linda, my father Tim, and my brother Kevin. I can never repay the kindness, motivation, and love I have received while working hard for this project. I could never have done this without the backing of my full roster of family and friends. I thank each one of you in my heart, even if I cannot find room on paper. Thanks to my soon to be wife Gianna, to whom the end of this journey is the beginning of a greater one shared.

I would like to thank everybody in my GSO experience. Thanks to my office mates Colin and Victoria for 2.5 years of everyday encouragement, friendship, and tea. Thanks to John King and all of his lab for taking me in when I was a struggling undergrad and giving me the confidence to continue in the Master's program. I would especially like to thank my adviser Rebecca Robinson whose hard work and persistent mentorship have never ceased to amaze me. Her passion for her students is embedded in every action she does, shaping myself to be a better student, scientist, and person. For that, I will always be grateful.

I would also like to thank my various sources of funding throughout my master's research including the National Science Foundation and Rhode Island EPSCoR.

PREFACE

This document is in Manuscript Format and is intended for publication in Paleoceanography.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

MANUSCRIPT

The following manuscript is prepared for submission to Paleoceanography.

INTRODUCTION

Diatoms are a ubiquitous and diverse group of phytoplankton that play a key role in marine carbon, nitrogen, and iron biogeochemical cycles, serving both as a source and loss term (Sarthou et al., 2005; Buesseler, 1998). South of the Antarctic Polar Front, diatoms are the main primary producer and important carriers of carbon and dissolved silica into the deep ocean (Singer and Shemesh, 1995; Cortese and Gersonde, 2007). The biological exchange of nutrients is an important factor in diatom growth dynamics and ultimately, sequestration of atmospheric carbon by the ocean (Cortese and Gersonde, 2007). In turn, consumption of nutrients by diatoms in the Southern Ocean has widespread impacts on the global nutrient distribution related to subsurface water masses formed in the Southern Ocean (Sarmiento et al., 2004; Cortese and Gersonde, 2007). While some diatoms are cosmopolitan in nature, many diatom species are endemic to specific ranges of environmental conditions including temperature, salinity, sea ice presence, and nutrient availability. As a consequence of these narrow ecological preferences, diatom assemblages are used to characterize the physical properties in which the community lived (Cermeño and Falkowski, 2009; Baas-Becking, 1934; Zielinski and Gersonde, 1997). The diatom fossil record of the Southern Ocean is an excellent environmental archive of major climatic and physical changes over time (Burckle and Cooke, 1983).

Across the Southern Ocean, the biogeographic distribution of many polar species are restricted by the temperature and salinity gradients of oceanographic fronts, the sea ice edge, and regions of heavy mixing or relatively stratified water (Zielinski and

Gersonde, 1997). Over glacial time scales, fluctuating limits on species extent are manifested as variations in the diatom assemblage. Despite the fact that assemblage composition is altered during deposition in the sediment, through dissolution of specific groups of diatoms, the remaining fossil assemblage and its chemistry are representative of surface hydrology (Pichon et al., 1992; Crosta et al., 2005; Zielinski and Gersonde, 1997). Fossil diatom assemblages provide an excellent first order assessment of environmental conditions and when combined with geochemical methods, they are likely to improve the quantitative nature of reconstructions related to nutrient dynamics as well. Emerging biogeochemical studies of nitrogen, carbon, and silicon are using diatom frustule associated stable isotopes to understand the degree of nutrient consumption over time (Singer and Shemesh., 1995; Sigman et al., 1999; Robinson et al., 2004; De la Rocha et al., 1997; Popp et al., 1999). In the case of carbon and nitrogen, the organic matrix is thought to be naturally protected from diagenetic processes by the siliceous biomineral surrounding it (Singer and Shemesh., 1995; Sigman et al., 1999; Robinson et al., 2004; De la Rocha et al., 1997; Popp et al., 1999). Assemblage is needed for these studies to account for difference in how individual species contribution record the isotopic signature of the water in which they grew (i.e. degree of fractionation) (De la Rocha et al., 1997; Des Combes et al., 2008; Horn et al., 2011A; Sutton et al., 2011; Sutton et al., 2013; Studer et al., 2015). Because size is an important factor in explaining variations in the biogeochemical parameters of diatoms (Sarthou et al. 2005), it should be beneficial to compare the volumetric contributions, as opposed to simple counts, to evaluate the biogeochemical contribution of diatom species.

The measurement of a diatom's shape and size, known as biometrics, is a recently developed tool for environmental reconstruction. Morphology appears to be directly related to the growth conditions and productivity of the diatom community (Crosta, 2009; Burckle and McLaughlin, 1977). The size of diatoms, specifically their relative volume and surface area, can affect their internal chemical composition and their ability to uptake nutrients, making it an important consideration for understanding growth dynamics (Sarthou et al. 2005; Wilken et al., 2011). Diatoms have a two-phase reproductive cycle where asexual reproduction leads to a reduction of cell size and sexual reproduction restores size to its initial condition (Edlund & Stoermer, 1997). The initiation of sexual reproduction is thought to be related to primary production, occurring earlier when production rates are high (Burckle and McLaughlin, 1977; Assmy et al., 2006). A positive relationship has been found between diatom abundances and size (Cortese and Gersonde, 2007; Crosta, 2009), consistent with the idea that growth conditions affect both. Diatom size, in turn, becomes an important factor in predatory protection and interspecies competition both of which have effects on assemblage and bloom dynamics (Wilken et al., 2011). Biometric studies of Southern Ocean diatoms have been shown to complement assemblage information by improving estimates of frontal position, as a secondary stratigraphic indicator of glacial terminations, and as an additional constraint on nutrient characteristics of the opal belt (Cortese and Gersonde, 2007; Burcke and Cooke, 1983; Jouse et al., 1962). The biggest factor preventing biometrics from becoming a more commonly utilized parameter is the labor intensive nature of measuring biometrics using traditional microscopic methods.

The traditional method for categorizing fossil diatom assemblages, slide-based light microscopy, provides an excellent standard for the identification of diatom species despite some known biases (Moore, 1973; Law, 1983). The largest of these is not related to the microscopic method itself but rather from slide creation and covering (Battarbee, 1973; Moore, 1973; Drooger, 1978; Schrader and Gersonde, 1978; Law, 1983). Counting bias can be influenced by size, depending on the conventions for counting fragmented particles, or by counting area (Law, 1983). By counting 300-400 diatoms, species making up less than 5% of total may have very high error in their relative count (Drooger, 1978; Schrader and Gersonde, 1978). In addition to these biases, traditional microscopic identification is labor-intensive, limiting the quantity of data that can reasonably be collected during an experiment's time frame. While this factor does not significantly impact studies designed to look at assemblage changes, where counts of 300 diatoms per sample are meaningful and relatively rapid (Schrader and Gersonde, 1978; Law, 1983; Zielinski and Gersonde, 1997), time consuming supplemental analyses, like diatom biometrics, are generally restricted for practicality. A significant drawback to using traditional microscopy for biometric analysis is that most studies limit measurements to only spatial changes, temporal changes, or a single species to accommodate the time requirement. In addition, as frustule based geochemical proxies are increasingly common, estimates of relative volumetric species contributions, rather than simple counts, become more important, increasing the need for comprehensive biometric data (Shukla et al., 2013).

To address these limitations, we present a potentially complementary method to microscopy for assessing first order changes in a sedimentary diatom assemblage that

provides robust biometric data as well. This method is not meant to replace microscopy in evaluation of diatom assemblages but rather to provide a tool to estimate changes in the relative contributions of major groups and provide biometric data for each valve counted.

STUDY LOCATION

TN057-13 PC4, from a 1996 cruise of the R/V Thomas Thompson, is an Atlantic sector Southern Ocean sediment core that spans to the last glacial period. It was recovered from 53.1728°S, 5.1275°E, about 2° N of the modern sea-ice edge and near the southern boundary of the Antarctic Polar Front, an area where diatoms can provide critical paleoenvironmental data (Orsi et al, 1995; Gersonde et al., 1999). I chose TN057-13 PC4 for its established diatom stratigraphy (Crosta, personal communication, 2010; Nielson and Hodell, 2005), high resolution glacial/ interglacial age control and opal record (Anderson et al., 2009), available biometric data for two diatom species (Shukla et al., 2013; Shukla et al., 2016), and ancillary geochemical data (Horn et al., 2011A) which I compare to the FlowCam results.

METHODS

A FlowCam is a fluid based imaging system akin to a flow cytometer developed to assist with identification and quantification of living plankton communities (Fluid Imaging Technologies, Inc., 2011). Individual particles pass in front of a magnifying lens and are photographed. The magnification is limited such that the primary benefits of the FlowCam are the high numbers of particles counted and the spatial dimensions recorded in each photograph. The distinct forms are then categorized based on user defined machine learning. This approach gives us the tools to 1. identify large number of diatoms in a short time. 2. automatically categorize diatom groups and 3. perform biometric assessment synchronously with assemblage. I sought to create an operating procedure for FlowCam analysis of fossil diatoms that will prioritize high diatom counts, automatic identification, and biometric parameterization. I established a protocol for analyzing fossil material on the FlowCam, after which, I generated a downcore record of diatom assemblage and biometric information from sediment core TN057-13 PC4 (Figure 1).

This methodology was developed using a FlowCam VS (Fluid Imaging Technologies, Inc.). A flowcell, composed of an optically clear glass sleeve fit with tubing on both ends, takes the place of the traditional glass slide. The flowcell is fixed vertically in the flow chamber, connecting to a sample funnel above and a syringe pump at it base. The shape of the flowcell restricts flow so that particles are forced along a single plane. As sample is pumped across the flowcell, the FlowCam, fitted with an optical magnifying lens, images its field of view at an assigned rate. The

software then isolates pictures of each “individual” particle it encounters. The FlowCam can distinguish particles in one of two ways; using a fluorescence sensor or as the brightness difference between trial images and calibrated background (Auto-image mode). The FlowCam takes field of view pictures of the sample and the software crops particles based on their differences from the background. The FlowCam exports the cropped particles into two collage types; one of raw images and a second, binary collage which is used to calculate spatial measurements of each particle (Figure 2). Particle images are processed in VisualSpreadsheet, a FlowCam native program, where they are assigned spatial characteristics and can be organized into libraries. Particles are sorted into classifications based on associated spatial data. I used two spatial quantities, maximum and minimum feret diameter, to measure biometric properties of the diatoms. Many diatom species are flat along one axis and thereby, tend to orient in the flowcell so that images are taken perpendicular to the diatom’s valvar axis. Maximum and minimum feret distance measures length (apical axis) and width (transapical axis) respectively which are important parameters for defining the biometrics of pennate and centric diatoms (Figure 3).

The FlowCam was operated with a X10 objective lens and illuminator with a 100 µm flowcell and 1 ml syringe pump. This allowed an effective particle range of 5 – 100 µm, though sometimes particles greater than 100 µm appeared because the flowcell’s opening size. Trials were performed under the default settings of Auto-Image mode for this flowcell and pump size (which determines rate of flow and imaging rate) (Fluid Imaging Technologies, 2011) except for segmentation threshold (Supplementary C), which was set to a dark pixel value of 10, and flash duration, set

so that mean pixel intensity is between 160 and 180. Around 1 – 10 mg of disaggregated sample, in this case diatom frustules, physically isolated and chemically cleaned following Horn et al. (2011B), was placed in ~ 10 ml of Milli-Q water to form a liquid slurry which was continuously mixed to ensure homogeneity. Before introducing sample into the FlowCam, at least 0.5 ml of deionized water was introduced in the pipette tip holder. For each sample, a test trial was conducted to confirm adequate focus and particle concentration. If particles physically blocked the flowcell or if the particle density was such that the FlowCam could not digital isolate all particles, samples were diluted until these conditions were met. At the start of each sample trial, between 0.1 and 0.5 ml of sample slurry (depending on estimated concentration) was quickly pipetted into the FlowCam after imaging had begun. As fluid levels began to drop in the pipette holder, deionized water was layered on top of sample to allow the total volume of sample to be imaged. Trials were completed when particle images were no longer being captured. Flowcells were rinsed between trials to ensure no cross contamination. Samples were analyzed in replicate with a target of 10,000 particles per sample collage.

Initially, diatoms were sorted into three major groups: Centrics, *Fragilariopsis spp.* and *Eucampia antarctica* Identifications were largely based on Scott and Marchant's *Antarctic Marine Protists* (2005). Ultimately, the assemblage counts and identifications were a combination of machine and operator effort. From test trials, libraries were compiled to teach the FlowCam the filter values needed to automatically distinguish diatom groups from each other. Libraries were made using at least 60 particles for each group. Centrics and *Fragilariopsis spp.* were chosen as good

candidates for automatic analysis because both had characteristic shapes and size ranges and were different from each other (Fischer, 2002). Identifications are first made by VisualSpreadsheet and then by an operator who looks at both the machine identifications and the rest of the collage for unmarked diatoms. This method ensured that machine identifications are accurate and that whatever could not be identified by machine alone was counted. Diatom counts were defined as being identifiable to one of the three diatom groups and having greater than half of its frustule. In cases where silicoflagellates, *Rhizosolenia*, or radiolarians could be identified, they were placed in unique libraries but not used for subsequent assessments. After these initial identifications were made, each group was further broken down into species identifications when possible. For example, *Fragilariopsis kerguelensis* was separated from the rest of the *Fragilariopsis spp.*. After the assemblage classification was made, two biometric classifications were made using the Centrics group and *Fragilariopsis kerguelensis*, the largest and most abundant of the *Fragilariopsis* species. *F. kerguelensis* was chosen specifically from the *Fragilariopsis spp.* group in order to be a point of comparison with Shukla et al. (2013). Particles were only placed in the biometric classification if the diatom was unbroken and correctly outlined by the FlowCam. Quartiles, skewness and kurtosis were estimated for each sample distribution.

RESULTS

Of the 27 replicate samples analyzed from TN057-13 PC4 imaged by the FlowCam, a total of almost 650,000 individual particles were measured, including over 39,700 identifiable diatoms. From these diatoms, over 28,000 could be measured for biometric parameters. The average trial produced around 750 identifiable diatoms, 530 of which could be measured. Only 2 - 13% of FlowCam measured particles were identifiable as one of the three diatom taxon, likely because the samples were cleaned for geochemistry, a process which likely broke frustules significantly. The remainder of the particles were diatomaceous in nature, but were broken and technically unidentifiable. Because the amount of sample material entering the FlowCam was not directly measured, absolute diatom count is not meaningful as a metric to compare trials. Values were normalized by dividing the absolute diatom count of each group by the total number of diatoms counted (Figure 4). Some recognizable species, like *Thalassiothrix antarctica*, were not included in this assessment because it was difficult to judge valve endings. It was often possible to identify the diatoms to the species level, but problems prevented this from occurring in all cases. Centric identifications were made for *Thalassiosira lentiginosa*, *Thalassiosira oliveriana*, *Thalassiosira gracilis*, *Porosira glacialis*, and *Asteromphalus spp.* based on features in the valve face for some but not all countable centrics. Our ability to identify to the species level is hampered by (1) the low magnification of the lens; and (2) a lack of focus for nearly half of the centrics imaged. Blurry images stem from efforts to maintain a good outline, for spatial information, which with the often somewhat concave or convex

valves resulted in a loss of detail on the valve face. Species identification for *Fragilariopsis spp.* was possible for 3 species; *F. kerguelensis, F. rhombica* and *F. curta*/*F. cylindrus* (*F. curta*/*F. cylindrus* are considered difficult to tell apart with general microscopy). Although *F. kerguelensis* and *F. curta/ F. cylindrus* often imply opposite environmental conditions, *Fragilariopsis spp.* was kept as a single group for assemblage comparison because species other than *F. kerguelensis* were not consistently present in the assemblage.

In all cases, *Thalassiosira lentiginosa* and *Fragilariopsis kerguelensis* dominated and were the major component of assemblage change. Using the three groups *Fragilariopsis spp.*, centrics, and *E. antarctica*, the downcore assemblage changes showed two unique regimes: 1. *Fragilariopsis spp.* dominated (60% or greater), between 100 to 800 cm and 2. *Fragilariopsis*-poor, 800 cm and deeper, which contains lower *Fragilariopsis spp.* contribution (40 – 70%), higher centric contribution, and a peak in *Eucampia antarctica* (Figure 5).

The number of diatoms available for biometric measurements is a function of the composition of the assemblage. While the average trial measured over 500 diatoms for biometric data, most measurements are from *F. kerguelensis* because it dominated the downcore abundance. Centric abundance contributed to 17% of total diatoms identified and its biometric measurements were available for ~15% of diatoms identified. By contrast, *F. kerguelensis* biometric measurements were available for 56% of diatoms identified meaning that it has more measurements than centrics. It is noteworthy that the depths between 800 - 900 cm, the inferred glacial period, had the fewest identifiable diatoms in its ~10000-particle count, which limited the number of

diatoms that could be measured for biometrics (Figure 4). 71% of all diatoms identified could be measured for biometrics during this study, ranging from 42 – 81% in individual samples. The unmeasured diatoms were essentially large fragments, identifiable and countable but would give incorrect length/diameter measurements.

Apical length was measured on *F. kerguelensis* individuals. Large variations in apical length occur, with the full range encompassing 11 – 92 µm. Despite this large range, the average sample apical length varied between 31 - 35 µm across the last glacial transition. Temporal trends of average *F. kerguelensis* length shows bimodal peaks at 10 and 15 kya (Figure 7). Before that, this size trend was relatively stable between 27.5 kya and 17.5 kya, around 34 µm. Centrics also showed a large range between 14 – 121 µm with its average diameter varying between 28 - 41 µm. The temporal trends in centrics seem to be opposite of *F. kerguelensis*, peaking when *F. kerguelensis'* sizes are smallest and decreasing/ leveling off as *F. kerguelensis* size increase (Figure 8).

DISCUSSION

ASSEMBLAGE COMPARISON ACROSS METHODOLOGIES

In order to show that the FlowCam captures robust relative assemblage data and both accurate and precise biometric measurements, I compare the Crosta (personal communication, 2010)/ Nielson and Hodell (2005) and Shukla et al.'s (2013) microscopic data of species composition (Figure 5) and biometric measurements (Figure 7), respectively to the FlowCam data. One of the major features observed from previous studies of TN057-13 PC4 is a shift in the diatom assemblage across the last deglacial period, where *T. lentigenosa* gives way to *F. kerguelensis* upon deglaciation (Nielson and Hodell, 2005; Shukla et al. 2013; Shukla et al., 2016). Our results show that the relative diatom abundance determined with the FlowCam shifts similarly at this time (Figure 5). Both counting methods show that before the glacial termination, centrics and *Fragilariopsis spp.* contribute around 50% each, or near parity. After the termination, *Fragilariopsis spp.* dominates the assemblage, contributing upwards of 65% throughout the Holocene. The FlowCam shows an overall bias towards higher relative abundance of *Fragilariopsis spp.*, with a greater proportional contribution than determined by microscopic counts. This bias is likely because the FlowCam's 5 - 100 µm range might not encompass the full natural range of centrics at this location; it may exclude some large centrics that could have been counted in traditional microscopy. Despite this bias, the FlowCam provides a good first order assessment of the relationship between centric and pennate dominated assemblages.

A key difference in relative abundances between the FlowCam and the microscope counts is the much higher abundance of *Eucampia antarctica* recorded by the FlowCam relative to the other diatom taxon. While E. *antarctica* made up less than 5% of the assemblage with traditional microscopy, its proportion in the FlowCam was as high as an order of magnitude greater, making up 36% of the assemblage during the deglacial transition (Figure 5). The trends of relative abundance are the same for both methods, but the overall contribution of *E. antarctica* is greater with the FlowCam. The difference in *E. antarctica* counts is probably a result of differences in the technique used to separate and concentrate diatom frustules. *E. antarctica* is more resistant to dissolution and breakage (Pichon et al., 1992) than many other diatom species and it would also appear more frequent if other species' frustules were preferentially broken during the cleaning process, which includes gentle sonification. This FlowCam bias may potentially be exploitable because *E. antarctica* is an important ice edge (Zielinski and Gersonde, 1997; Burckle and Cooke, 1983; Jouse et al., 1962; Fryxell & Prasad, 1990), ice rafted debris (Zielinski and Gersonde, 1997), and stratigraphic indicator for glacial conditions (Kaczmarska et al., 1993; Burckle and Cooke, 1983; Jouse et al., 1962). The FlowCam could potentially be a robust, quick method to gauge *E. antarctica* contributions for stratigraphic purposes.

COMPARISON OF BIOMETRICS

Existing measurements of *F. kerguelensis* average length from Shukla et al., (2013) allow for a comparison of FlowCam based measurements with those from

microscopy. Biometric data for *Thalassiosira lentiginosa* from TN057-13 PC4 is also available (Shukla et al. 2016). *F. kerguelensis* length measured with the FlowCam shows approximately the same absolute value and downcore variation as length measured by microscopy from Shukla et al. (2013) (Figure 7). The two datasets show the same general bimodal peaks in mean length centered at 10 kya and 15 kya during the last deglacial that correspond to episodes of peak opal flux (Anderson et al., 2009; Shukla et al., 2013). While the FlowCam initially looks to have slightly lower average length than that of Shukla et al. (2013) (0.5 µm or less in most cases), the difference is not a significant as the confidence intervals for Shukla et al. (2013) are greater than 1.5 µm (Table 1).

The interpretation of centric average diameter is more complicated than *F. kerguelensis* because it contains multiple genera that may have different natural size ranges and contributions over time and the flowcell caps sizes at ~100 µm. Although the majority of identifiable centric diatoms, by both methods, are *T. lentiginosa*, differences in mean diameter between Shukla et al.'s (2016) *T. lentiginosa* data and this study's centric diameter measurements suggest that other species may have significant contribution to the average diameter for centrics as a whole (Figure 8). Centrics show a unique size trend with a maximum just before the major deglacial increase of *F. kerguelensis* and decreases toward the present. There was no major change in size range for either diatom group over time or in the size distributions. The trends in both interquartile ranges roughly mimicked that of its average length/diameter, but the quartile changes were almost equally proportional it changes in length. These interquartile changes were not statistically different than what is to be

expected from size changes alone. Similarly, there was no distinguishable difference between skewness and kurtosis between glacial and interglacial assemblages. While mean is not normal a meaningful metric in a skewed distribution, because skewness and kurtosis do not change significantly with time, changes in mean are actually show shifts in the whole distribution.

ENVIRONMENTAL INTERPRETATION OF BIOMETRICS

Supplementing assemblage data with the biometric data from multiple groups allows for a more specific interpretation of past events than only having a single group, and in this case, helps to look at growth conditions over the stages of the last deglaciation. In comparing the biometric data from *F. kerguelensis* and centrics, I observe differences in the timing of their peaks, likely resulting from different environmental preferences, which are in phase with different stages of the deglacial. My data suggests that two major features of deglaciation at TN057-13 PC4, the retreat of sea ice and upwelling, are not synchronous events and that biometrics helps explain the environmental conditions that existed between them. Deglaciation began and sea ice influence waned at 20 kya where *E. antarctica* peaks and then sharply declines (Figure 8). The trend of *F. kerguelensis* size has been discussed in previous literature and is in phase with opal flux from a time of inferred peak nutrient supply as deglaciation took place (Shukla et al., 2013; Anderson et al., 2009) (Figure 7). This upwelling begins at roughly 17.5 kya where opal flux rapidly begins to increase. My data suggest an interval between these events where there was lower sea ice influence

but no indication of nutrient input from the upwelling (Figure 8; Horn et al., 2011; Anderson et al., 2009).

During this 2.5 ky interval, centric average diameter and diatom nitrogen isotope values (as $\delta^{15}N$) increase while dust flux stays relatively high. The increase of centric average diameter implies better growth conditions for centric diatoms, which was possibly due to alleviation of light limitation, caused by the retreat of sea ice, or perhaps related to the persistent iron supply (EPICA et al., 2004) (Figure 8). The peak in centric average diameter seems to occur with the transient peak in the diatom bound $\delta^{15}N$ record (Figure 8). Higher $\delta^{15}N$ values suggest that the demand for nutrients was greater than the supply, consistent with relief of either light or micronutrient limitation, without a significant change in the supply of major nutrients.

It has been suggested that the assemblage changes alone may influence the $\delta^{15}N$ record through species specific differences in the isotopic fractionation of nitrogen during incorporation into the diatom frustules (Horn et al., 2011B). The timing of the observed $\delta^{15}N$ shifts is not quite synchronized with the assemblage changes, making attribution to species changes difficult. Moreover, there is disagreement between the Horn et al., (2011a) culture data and data from the fossil record (Studer et al., 2015). Studer et al., (2015) found that centrics appeared to record higher $\delta^{15}N$ values than the total assemblage. The observed increase in $\delta^{15}N$ values approaching 17.5 ka could potentially be related to assemblage changes, related to a peak in volumetric contributions from centrics, consistent with a peak in their size rather than simply the result of changing growing conditions related to nutrient supply and demand. This relationship should be a line of inquiry in future studies.

ASSESSMENT OF INCREASED BIOMETRIC COUNTS

Biometrics is emerging as metric of growth conditions of diatom communities despite the difficulty in acquiring large datasets. The FlowCam makes biometric measurements rapidly, allow for high quantities of data. If robust, it could potentially increase the accessibility of measuring biometrics and provide robust statistical analysis. I used the surplus of biometric data gathered from TN057-13 PC4 to look at the potential implications of having increased biometric data and how this affects our ability to resolve changes in average frustule size in fossil records. The available studies of Southern Ocean diatoms with biometric data limited the number of biometric measurements to 100 or less for a single species. Shukla et al., (2013) measured apical valve length and calculated area of *Fragilariopsis kerguelensis* on 100 individuals per sample (Shukla et al., 2013). Using the FlowCam, I imaged up to 1200 biometric eligible *F. kerguelensis* per trial from TN057-13PC4 samples in replicate. The associated biometric data has the potential to provide a more representative assessment of average valve length and will allow us to define an optimal sample size. Here, I will compare these two biometric datasets.

There are two goals when looking to improve the biometric assessments of *F. kerguelensis*: (1) Ensure the number of counts is large enough to provide a measured average valve length and standard deviation that is representative of the population's true value; and to (2) have small enough confidence intervals that I can distinguish environmentally relevant size changes. *F. kerguelensis* shows a large variation in size; its normal ecological range is between 8 and 92 µm (Shukla et al., 2013; van der Spoel

et al., 1973; Fenner et al., 1976; Assmy et al., 2006). The distribution of *F. kerguelensis* is right skewed with mean length typically in the mid 30's. The associated high deviation makes it difficult to determine when a representative dataset has been collected.

We suggested above that the FlowCam biometric measurements are statistically indistinguishable to results from traditional microscopy for *F. kerguelensis*. The FlowCam's true benefit over the traditional method stems from its higher counts and due to the law of large numbers, increased precision of the mean. The ultimate goal of large numbers is to capture the precision needed to differentiate mean values and minimize the effects of outliers. To have a first order estimate of when these goals are met, *F. kerguelensis* length was plotted as a running average for multiple FlowCam trials (Figure 9). At low counts, these graphs show a random walk that eventually becomes centered around the true population mean. The point at which this random walk begins to show low fluctuation is generally a function of the population's variance and one of the functions used when calculating minimum sample size. I interpret that when fluctuation of the value falls to a minimum, it is closest to the populations' true values and the effects of outliers was minimum. The consensus from these graphs was that this uncertainty often occurred after 100 counts, which means that 100 counts may not adequately define these *F. kerguelensis* distribution.

To address my second goal, I must first determine what amplitude of *F. kerguelensis* size changes are required to infer actual environmental changes. Previous studies have documented a range of changes in *F. kerguelensis* average length across different environmental shifts including ~9 μm decrease across the last glacial

maximum (Nair et al., 2015), ~3 µm increase between the Hypsithermal and the Neoglacial (Crosta, 2009), ~4-6 (with a maximum of 10) µm between the last glacial and the Holocene (Gersonde and Cortese, 2007), and at total range of ~10 µm, attributed to geographic position, from samples spread across the Southern Ocean (Gersonde and Cortese, 2007). Therefore, the inferred environmentally influenced range of average valve size changes is between 3 - 10 µm. The question remains whether these changes are related to environmental conditions or random variability. To test this, I assumed that real changes were reflected in a 3 µm mean size shift and then asked how many biometric measurements are needed to narrow confidence intervals sufficiently to see such a change given the variance in the data. In order to make this estimate, I use the standard equation for minimum sample size (Equation 1; Charan and Biswas, 2013; Bennett et al., 2017) and assumed a normal distribution for the size data. This assumption allows me to make quick, first order estimations of minimum sample size, but are likely underestimated because we are dealing with a skewed distribution. To test that my standard deviations capture the same variability as traditional microscopy, I compared the range of values to a known core. Because standard deviations were not provided in Shukla et al.'s (2013) assessment of TN057-13 PC4, data from an East Antarctic core, published by Crosta (2009), is used to estimate the range of standard deviation of average *F. kerguelensis* length in the Southern Ocean. Crosta's (2009) standard deviation of average *F. kerguelensis* length ranged between 7.4 - 15.5 µm while this study showed a range of 8 – 12.5 µm, showing that much of the variance is preserved across methods. Confidence intervals will be reported as E (Equation 1) which represents the distance between the true

population value and its upper or lower limit. These confidence intervals would ideally need to be half of the environmentally significant value, in this case 1.5 µm, or less in order to capture a 3 µm change (the smallest of the recorded changes). Using Crosta's (2009) extremes and assuming I want the measurement to be within 1.5 µm at 95% confidence, you would need to measure 93 *F. kerguelensis* in the best scenario (standard deviation = 7.4) and over 400 in the worst scenario (standard deviation = 15.5). Because of the 100 counts used, Crosta's (2009) 95% intervals really reflect a range of 1.5 – 3 µm, which implies their data cannot be used to resolve a change of 3 µm between points without application of other statistical techniques. This assertion also applies to other biometric studies where the small sample size may not have the confidence intervals to resolve the environmental changes inferred to be related to their data (Crosta, 2009; Cortese and Gersonde, 2007; Nair et al., 2015, Shukla et al., 2013; Shuka et al., 2016). 100 counts with the FlowCam yields a similar 95% confidence interval of 1.5 – 2.3 µm, showing that this inability to resolve small average length changes is not necessarily based on method, but instead, a result of the high variance of the measurements. This method of resolution only applies when the difference between two points (or series of points) is needed. In many cases, a trend is still significant even in cases where points cannot be individually resolved as different.

Equation 1. Standard equation for sample size

$$n = [\frac{za_{/2}\sigma}{E}]^2$$

n = necessary sample size (count)

$Za_{/2}$ = Z score for confidence intervals (1.96 at 95%)

σ = standard deviation

E = maximum difference between average measurement and true population average

It appears that, on average, the 100-count standard cannot always precisely represent average lengths and would not likely have small enough confidence intervals to resolve size changes that have been purported to be environmentally significant. Given an abundant and easily identifiable organism, such as *F. kerguelensis*, the FlowCam is a good tool for this problem because of the lower time costs than traditional microscopy. Ideally, future studies would like to spend the least amount of time per FlowCam sample, i.e. measuring the least counts, and still capture all of the information needed, including length and standard deviation, with small enough confidence intervals to resolve environmental changes. While the values in Table 1 are slightly underestimated due to the skewed nature of the *F. kerguelensis* length distribution, they provide a first order assessment of how many counts are needed to decrease our confidence interval. Given the standard deviations calculated in this study, a 300-count per each sample would likely be the lowest count needed to provide confidence intervals less than 1.5 µm (Table 1) in order to resolve 3 µm changes in average frustule length.

POTENTIAL IMPACTS AND FUTURE WORK

The FlowCam seems to be a useful paleoceanographic tool because it can capture first order diatom assemblage information and provide a large number of biometric data for a statistically more robust result. The FlowCam also has the potential to pick out indicator species, such as *E. antarctica*, for stratigraphic assessment. The ability of the FlowCam to measure first order changes in diatom assemblage can be useful when paired with biogeochemical data, especially in scenarios like nitrate utilization where major diatoms groups are shown to have significant impacts on the total $\delta^{15}N$ (Horn et al., 2011A; Studer et al., 2015). Future studies could use the high quantity of biometric measurements from the FlowCam to calculate the volumetric contributions of different diatom groups (calculated as a function of length and width) and use them for the basis for comparing biogeochemical records rather than counts, which should be a better representation of the relative contributions to chemistry by a given organism.

This methodology could be improved in the future through analysis of changes within the *Fragilariopsis spp.* group. *F. curta* and *F. cylindrus*, seasonal sea ice zone diatoms, differ from *F. kerguelensis*, an open ocean diatom, by having opposite environmental and ecological requirements for growth (Crosta, 2009). The distinction between these species could lead to assessment of yearly sea ice cover and would be keystone in any study in which *F. curta* and *F. cylindrus* are particularly abundant. While I was able to distinguish between these species within the *Fragilariopsis spp.* group with FlowCam images, identifications for *F. curta* and *F. cylindrus* had much lower abundance than is reported in literature. This discrepancy is likely because our

FlowCam methodology did not look at particles smaller than 5 µm, so in turn, it does not fully encompass the natural size range of species at this site. While there appeared to be more *F. curta*/*cylindrus* in glacial samples, the abundance was so low that it is unlikely statistically robust. It would be, however, possible to capture this range in future experiments by sieving samples into a small size fraction. The smaller size fraction could then be run in the FlowCam with the 50 µm flowcell under the X20 lens to increase magnification needed to view this size range. This method may also be useful in identifying *Chaetoceros* resting spores which is a proxy for spring ice melting (Crosta, 2009). Similarly, a shift to a larger flow cell, could be implemented for capturing the full spectrum of centric sizes, above 100 µm. There may be issues with using the same sample on two different flowcell sizes because it would involve changing the flowcell between trials or saving sample for later. A valid option for the use of two FlowCams could involve simultaneously inputting the same sample through different FlowCam flowcells in attempts to create a method for mixed distributions and the standardization of volume/material.

The FlowCam can also benefit from additional uses of structural deep learning. While this VisualSpreadsheet uses the spatial parameters of diatom particles, this is not the only bioinformatic method to quantify and identify diatoms. Other techniques such as symmetry contouring, Fourier descriptors, texture analysis, and striation descriptors can be used as digital data which quantifies the physical properties of a diatom (Fischer, 2004) and many of them can be improved by machine learning. Because the FlowCam has a large output of diatom images, it would be a useful tool

for creating training sets for other image based methods to potential improve accuracy for automatic detection of diatom taxon.

CONCLUSION

Fossil diatom assemblages evaluated with a new FlowCam method show agreement with the trends measured by traditional microscopy. Because the FlowCam could not identify every diatom to the species level, order based groups were used for centrics and pennates, indicating that the FlowCam is suited for first order assessments in assemblage only. Microscopy clearly remains as the gold standard for species identifications. Comparison of average length data generated by both microscopy and FlowCam suggests that the FlowCam can accurately measure average biometric length of *Fragilariopsis kerguelensis*. Moreover, the FlowCam may be better suited for assessing downcore changes in biometric parameters because it can measure more particles and do more groups in less time than traditional microscopic assessments. The overabundance of biometrics from *F. kerguelensis* allowed empirical calculation of minimum sample number, confidence intervals, and environmental resolution. At the standard count of 100 for average length of *F. kerguelensis*, the FlowCam had 95% confidence intervals between 1.5 and 2.3 µm, suggesting a resolution of 3-4.6 µm size differences at best. Future best practices should employ a count of 300 *F. kerguelensis* or more to resolve inferred environmentally relevant size changes of 3 µm or greater.

FIGURES

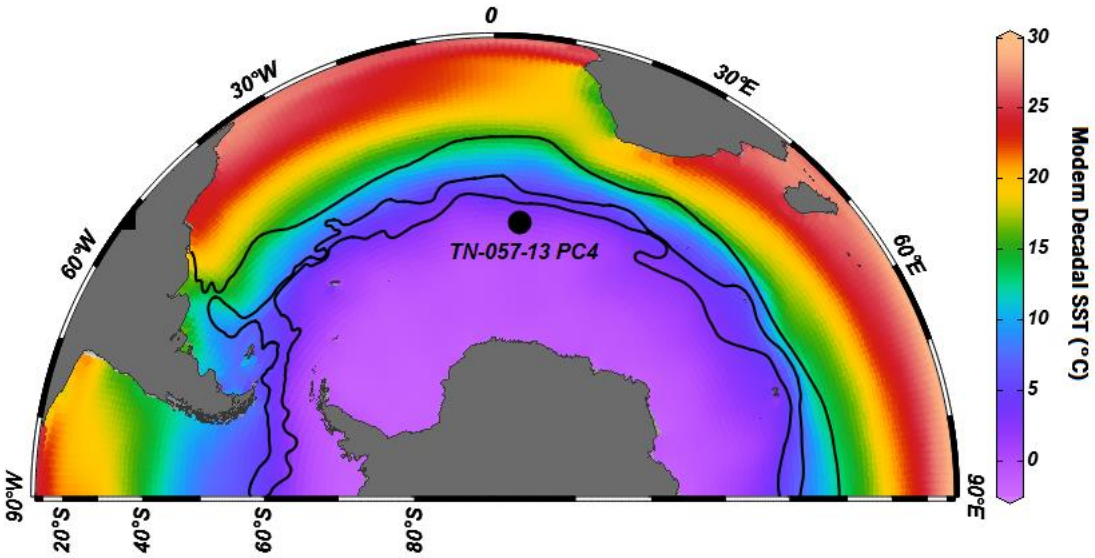

Fig. 1. A sea surface temperature map (decadal average (2000-2010)) showing the location of the studied core site, TN-057-13 PC4. The black lines reflect the Southern Ocean fronts, Antarctic Polar Front, Sub-Antarctic Front, and Sub-Tropical Front (inside to out). Winter sea ice in the Atlantic sector extends to the Antarctic Polar Front in the modern and evidence for summer sea ice in this area exists from diatom assemblage reconstructions (Gersonde et al., 2005). This core should be sensitive to changes in the positions of the Polar Front and thus temperature changes as well as variations in the extent of sea ice.

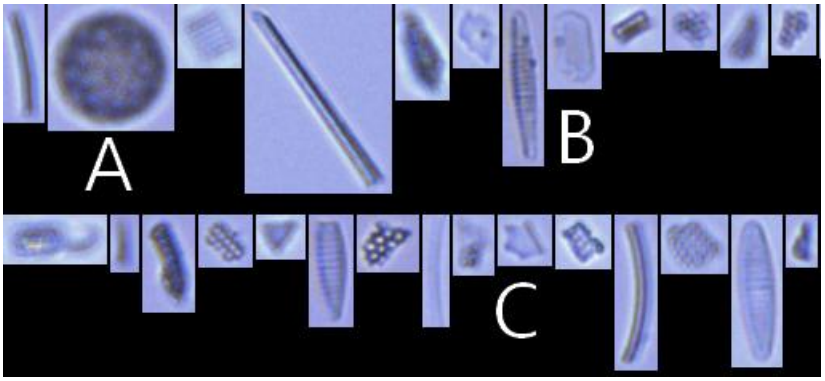

Fig. 2. Example section of a FlowCam image collage. In this example, (a). *T. lentiginosa*; (b). three *F. kerguelensis*, and; (c). broken pieces of other diatomaceous material can be seen.

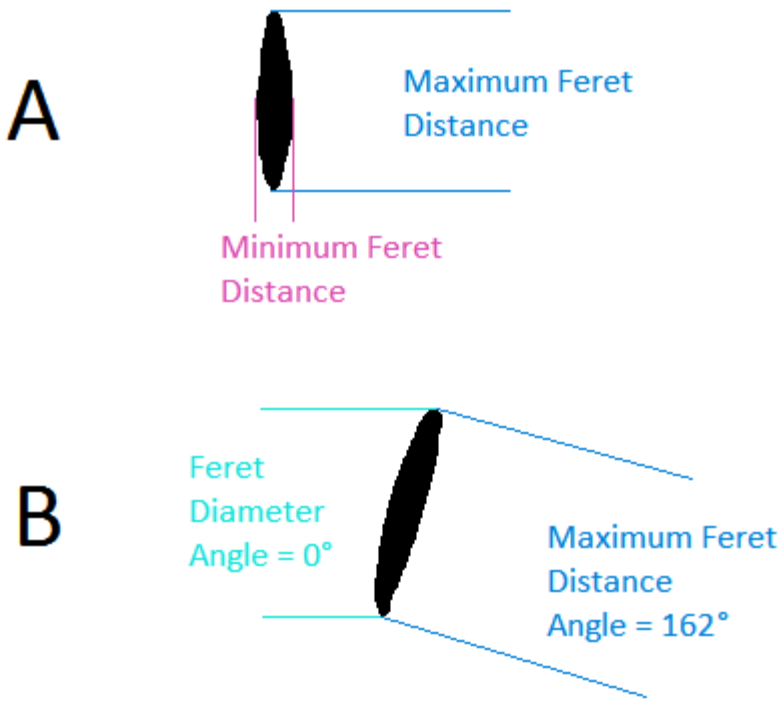

Fig. 3. Visual example of how feret/ caliper distances, from the binary data of two *F. kerguelensis*, are used to calculate apical and transapical valves respectively. (A). This *F. kerguelensis* is oriented vertically and has its largest feret distance along it length (apical) and its shortest feret distance across its width (transapical). (B). Even when diatoms are not in perfect orientation (this *F. kerguelensis* is 18° clockwise), feret distances are preserved, still measuring the longest and shortest axis.

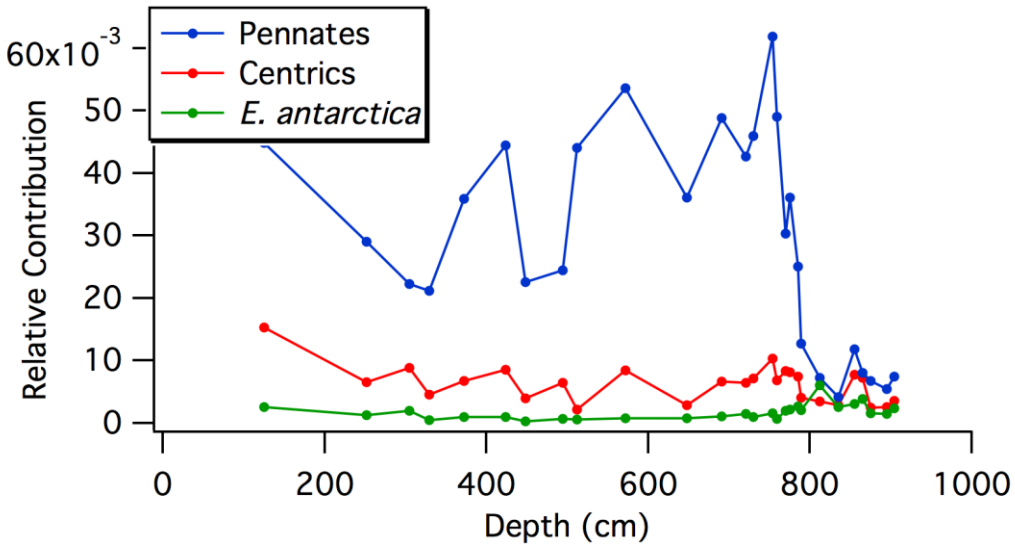

Fig. 4. A downcore comparison of the contributions of three taxon normalized to total diatoms identified.

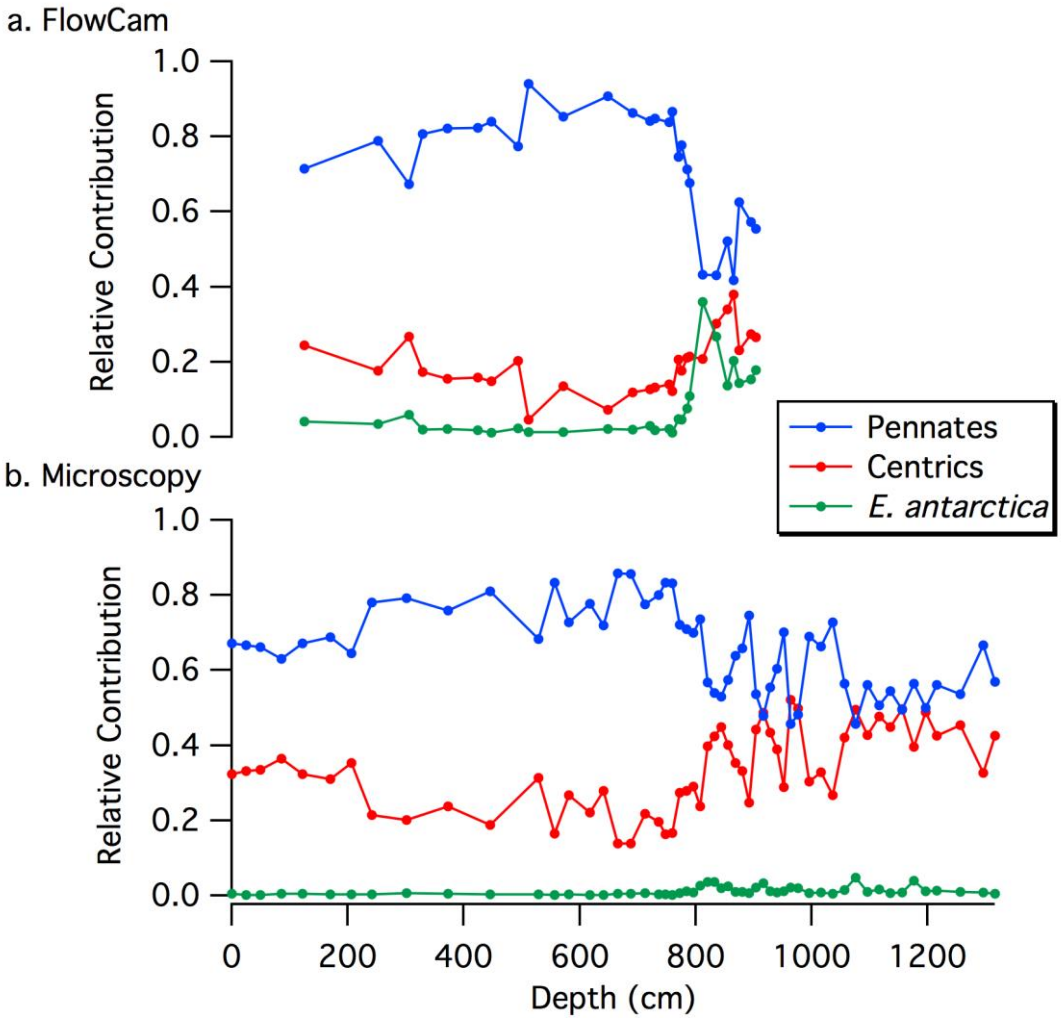

Fig. 5. A downcore comparison of the relative contributions of three taxon to total assemblage between (a). FlowCam estimates and (b). microscopy determinations (Crosta, personal communication 2010).

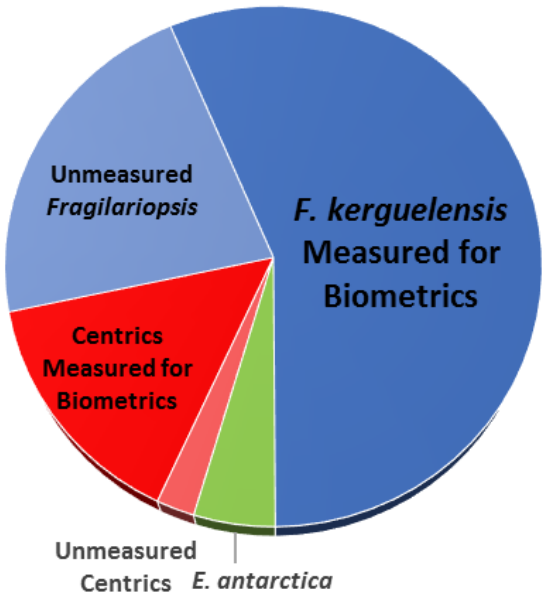

Fig. 6. Chart of the relative proportion of diatom taxon and the amount of each that could be measured for biometrics. Blue – *Fragilariopsis spp.*, Red – Centrics, and Green – *E. antarctica*. Darker red and blue show the proportion of its respective category that biometric data could be taken. About 71% of all identifiable diatoms were measured for biometric data.

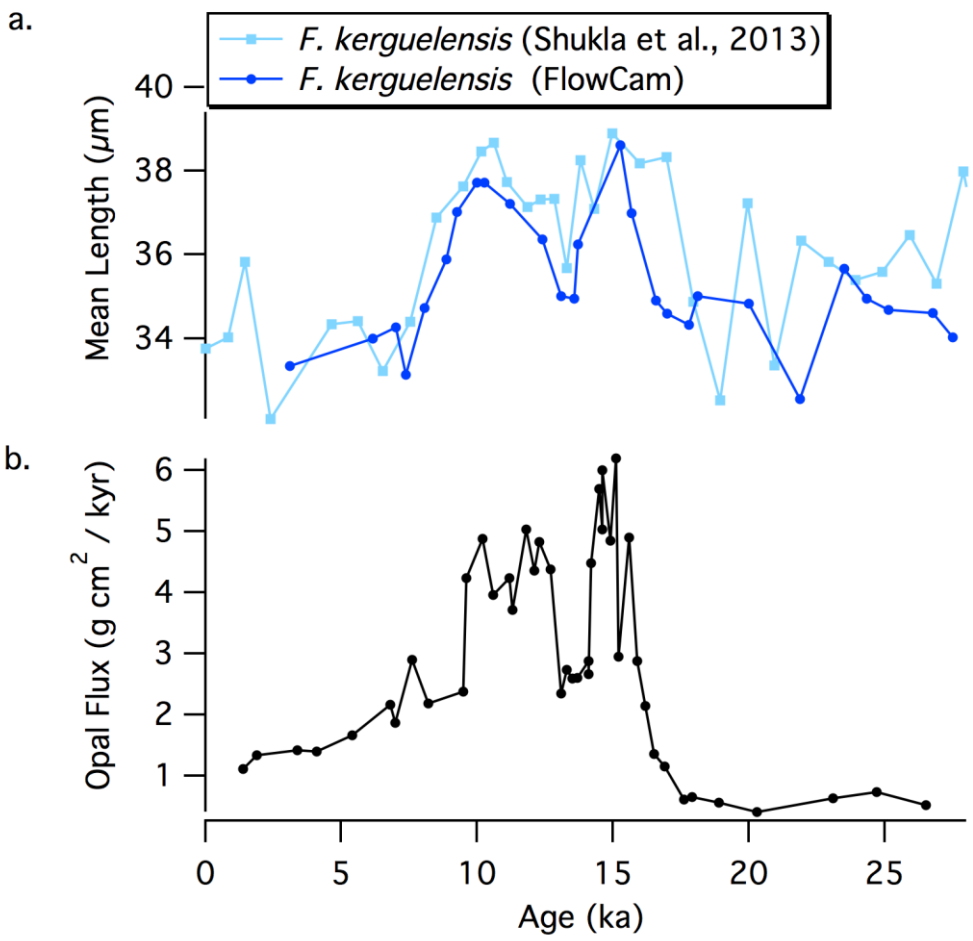

Fig. 7. A downcore comparison of average apical length of *F. kerguelensis* from FlowCam and traditional microscopy (from Shukla et al., 2013) (a). Opal flux from Anderson et al., 2009 (b). The relationship between *F. kerguelensis* size and opal flux was demonstrated in Shukla et al., 2013 and thought to be a result from increased nutrients from upwelling (Anderson et al., 2009).

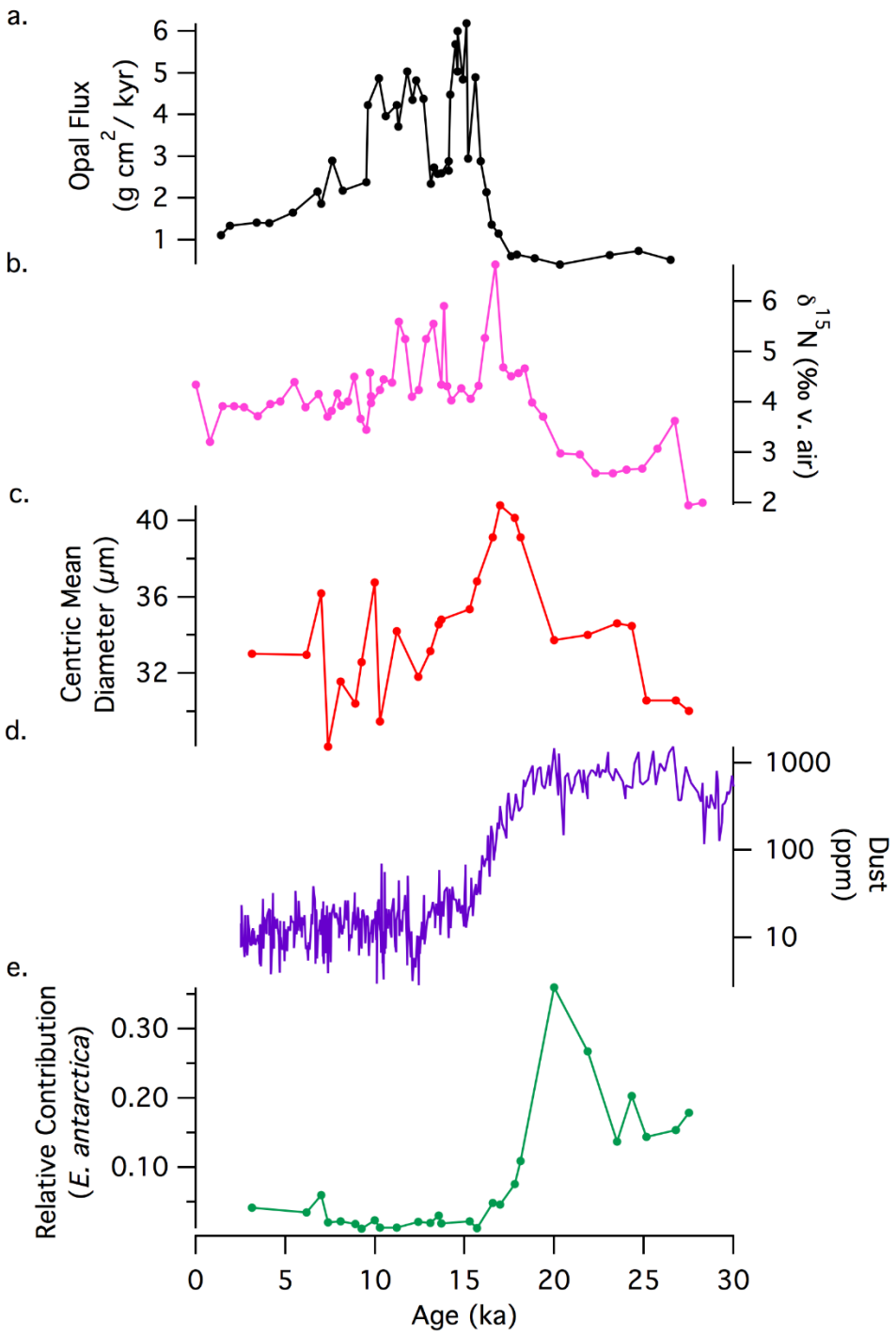

Fig. 8. A downcore comparison of (a). opal flux (Anderson et al., 2009), (b). diatom bound $\delta^{15}N$ (Horn et al., 2011A), (c). mean diameter of centrics FlowCam, (d) dust inputs from Dome C, and (e). relative contribution of *E. antarctica* measured with the FlowCam.

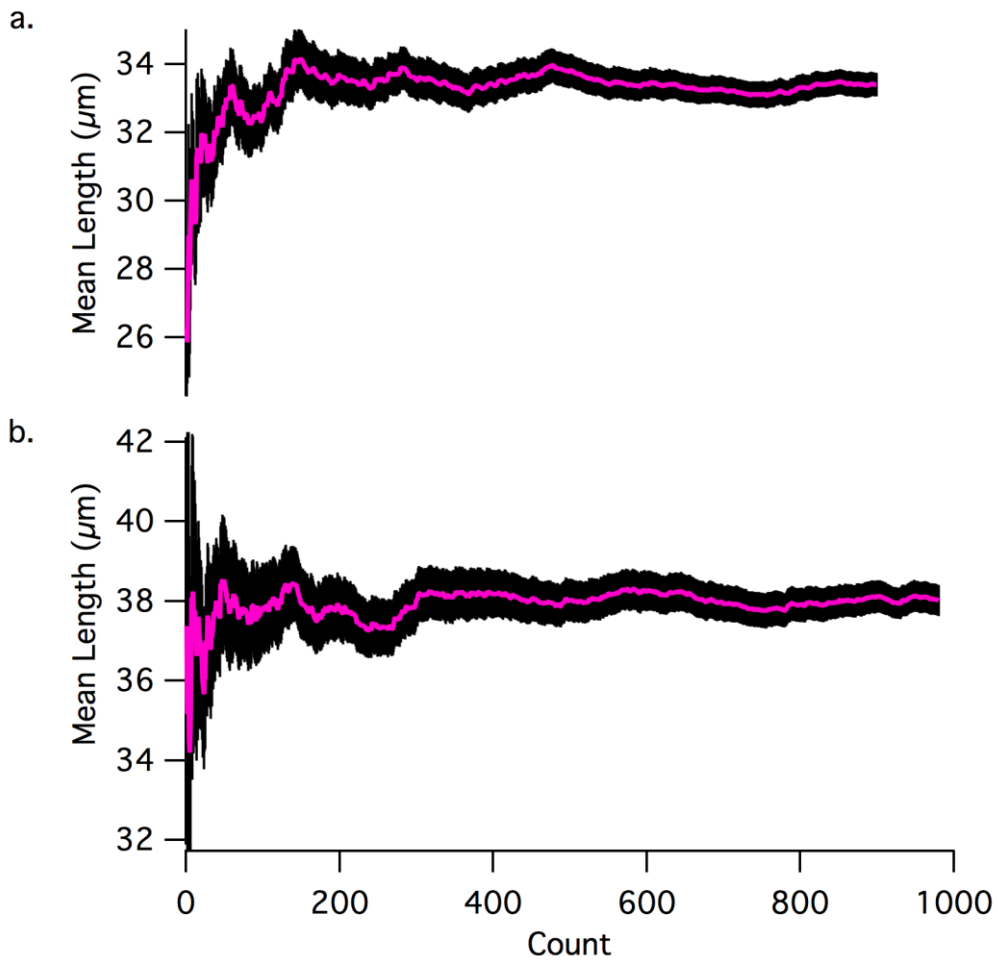

Fig. 9. Running averages of apical length as more *F. kerguelensis* were captured in a FlowCam trial. These examples show that on any given trial, variance (both fluctuation and error) is large with lower count. These represent opposite scenarios: A. Fluctuations are minimal after 150 counts. B. Fluctuations occur at a small magnitude, but do not stabilize until later. After 300 counts, both scenarios exhibit the least amount of fluctuation.

TABLES

Table 1. First order estimate of the theoretical decrease in 95% confidence intervals as biometric counts are increased based on the assumption that *F. kerguelensis* length distribution is normal. Scenarios are based on changes in the standard deviation. Standard deviations used are 8, 10.7, and 12 as the lowest, average, and highest values observed in this study respectively. In order to resolve a change in average length for one of these scenarios, the change would need to be greater than twice that value.

95% confidence interval of average *F. kerguelensis* apical valve length (µm) Calculation of E for distance between upper or lower limit of confidence interval and its true value.			
Biometric Counts (n)	Best Scenario	Average Scenario	Worst Scenario
100	1.57	2.10	2.35
200	1.11	1.48	1.66
300	0.91	1.21	1.36
400	0.78	1.05	1.18
500	0.70	0.94	1.05
600	0.64	0.86	0.96
700	0.59	0.79	0.89
800	0.55	0.74	0.83
900	0.52	0.70	0.78
1000	0.50	0.66	0.74
1100	0.47	0.63	0.71
1200	0.45	0.61	0.68

Table 2. Estimated increase in certainty of observing average *F. kerguelensis* apical length to within 1 µm of its true value as biometric counts increase. Values are overestimated as the distribution of *F. kerguelensis* length is assumed to be normal. Scenarios are based on changes in the standard deviation. Standard deviations used are 8, 10.7, and 12 as the lowest, average, and highest values observed in this study respectively.

Confidence that measured average *F. kerguelensis* apical length is within 1 µm of the actual population average			
Biometric Counts (n)	Best Scenario	Average Scenario	Worst Scenario
100	0.79	0.65	0.59
200	0.92	0.81	0.76
300	0.97	0.89	0.85
400	0.99	0.94	0.90
500	0.99	0.96	0.94
600	0.99	0.98	0.96
700	0.99	0.99	0.97
800	0.99	0.99	0.98
900	0.99	0.99	0.99
1000	0.99	0.99	0.99
1100	0.99	0.99	0.99
1200	0.99	0.99	0.99

APPENDIX A – CORE TN057-13 PC4

Table S1. Absolute counts of three diatom taxon and total particles measured in the FlowCam for the first of two replicate trials.

	Trial A			
Depth (cm)	Centric Count	Pennate Count	*E. antarctica* count	Total Particles Counted
125	395	1136	66	24378
252	117	601	30	12897
305	218	512	43	10439
329	78	314	5	6510
372	137	742	14	14721
424	137	833	19	17776
448	73	444	6	7470
494	144	506	13	8970
512	44	958	12	19615
572	160	1095	18	19712
648	57	739	19	12744
691	129	1028	18	13884
721	142	933	30	14028
730	124	846	18	11953
754	235	1373	28	13189
759	128	840	14	7806
770	148	642	39	11219
775	179	704	44	14371
785	168	542	60	17048
789	86	240	29	9606
812	68	144	132	14004
835	50	87	46	7891
855	140	260	53	12842
865	145	151	77	10372
875	47	119	35	4608
895	49	93	33	4608
904	75	147	40	7381

Table S2. Absolute counts of three diatom taxon and total particles measured in the FlowCam for the second of two replicate trials.

	Trial B			
Depth (cm)	Centric Count	Pennate Count	*E. antarctica* count	Total Particles Counted
125	215	644	38	12789
252	142	552	21	13888
305	133	374	36	8227
329	103	529	16	10391
372	133	685	24	12343
424	202	933	20	21091
448	86	451	6	7681
494	111	466	16	8983
512	41	790	12	16489
572	176	1034	14	17956
648	57	693	14	11870
691	137	913	26	13535
721	115	761	30	12019
730	161	979	22	13769
754	176	1084	36	12674
759	145	1108	13	10063
770	184	561	39	10880
775	146	731	41	14422
785	127	453	46	13830
789	75	267	53	10548
812	71	144	108	11755
835	65	77	56	8269
855	168	212	71	12258
865	145	167	78	12486
875	53	151	27	5756
895	54	123	25	5756
904	68	151	56	8291

Table S3. Number of diatoms measured for biometrics, average size, and standard deviation for centrics and *F. kerguelensis*

Depth	Centrics Measured	Centric Average Diameter	Centric Standard Deviation	*F. kerg* Measured	*F. kerg* Average Length	*F. kerg* Standard Deviation
125	518	35.02	11.19	1392	33.33	10.29
252	197	34.13	9.83	784	34.00	11.29
305	301	37.79	11.26	667	34.26	10.86
329	160	29.34	8.01	670	33.13	9.86
372	220	33.13	9.43	977	34.72	10.88
424	301	32.67	8.99	1163	35.89	11.15
448	142	35.16	11.99	604	37.02	11.32
494	240	39.20	13.18	651	37.71	11.59
512	74	29.97	9.64	1107	37.71	10.86
572	290	38.55	15.34	1394	37.21	10.58
648	102	34.26	10.46	1034	36.37	11.86
691	231	34.76	8.68	1246	35.01	9.85
721	236	35.34	9.15	1283	34.95	10.19
730	254	35.39	8.50	1368	36.24	10.86
754	368	36.53	9.98	1861	38.61	12.46
759	257	36.58	9.39	1573	36.98	10.87
770	298	40.08	10.82	951	34.90	10.86
775	290	40.79	10.29	1149	34.59	10.64
785	257	40.82	10.14	791	34.32	9.96
789	138	39.69	11.55	372	35.01	10.19
812	118	35.17	9.41	231	34.83	9.68
835	101	34.76	8.10	109	32.55	7.90
855	272	35.42	9.32	333	35.66	10.55
865	247	35.84	9.67	210	34.95	10.25
875	84	31.57	7.89	152	34.68	11.54
895	90	31.67	7.83	156	34.61	11.31
904	88	31.43	8.21	155	34.02	10.89

APPENDIX B – CORE ODP 1090

Table S4. Assemblage of three diatom taxon at ODP 1090 (42°154.50'S, 8°154.00'E)

Sample	Centrics	*Fragilariopsis spp.*	*E. antarctica*
B3,1 137-139	400	80	13
B3,2 19-21	464	286	25
B3,2 49-51	604	983	2
B3,2 79-81	562	746	37
B3,2 109-111	175	242	15
B3,3 18-20	71	126	14
B3,3 49-51	507	860	50
B3,3 79-81	235	383	42
B3,3 109-111	581	500	32
B3,3 139-141	227	335	114
B3,4 19-21	788	455	108
B3,4 49-51	359	472	60
B3,4 79-81	516	603	51
B3,4 109-111	332	91	5
B3,4 139-141	359	259	12
B3,5 109-111	179	66	3
B3,5 139-141	163	78	7
D3,2 74-76	279	103	1
D3,2 103-105	94	46	0
D3,2 133-135	353	118	3
D3,3 14-16	112	177	13
D3,3 44-46	76	100	21
D3,3 74-76	280	481	76
D3,3 104.5-106.5	305	839	14
D3,3 134-136	184	415	14
D3,4 104-106	280	80	7
D3,5 44-46	78	68	12
E3,2 70-72	264	122	0
E3,3 40-42	534	766	2
E3,3 100-102	33	11	1
E3,3 130-132	293	68	1
E3,4 10-12	51	44	0
E3,4 70-72	456	247	3

Table S5. Centric biometrics at ODP 1090 (42°154.50'S, 8°154.00'E)

Sample	Centrics Measured	Centric Average Diameter	Centric Diameter Standard Deviation
B3,1 137-139	400	31.13	8.72
B3,2 19-21	358	28.08	8.19
B3,2 49-51	409	23.86	6.23
B3,2 79-81	369	27.37	9.38
B3,2 109-111	115	28.64	8.91
B3,3 18-20	35	27.96	6.00
B3,3 49-51	293	35.67	9.49
B3,3 79-81	167	29.20	6.26
B3,3 109-111	423	34.37	8.38
B3,3 139-141	139	33.02	7.47
B3,4 19-21	523	34.27	8.27
B3,4 49-51	249	35.81	10.64
B3,4 79-81	333	29.73	10.41
B3,4 109-111	279	27.87	8.14
B3,4 139-141	313	32.05	10.31
B3,5 109-111	132	30.05	8.19
B3,5 139-141	121	33.48	8.06
D3,2 74-76	199	28.33	7.93
D3,2 103-105	78	26.26	7.73
D3,2 133-135	266	34.16	7.77
D3,3 14-16	82	28.85	7.36
D3,3 44-46	61	32.37	8.82
D3,3 74-76	173	32.89	8.94
D3,3 104.5-106.5	255	33.36	11.35
D3,3 134-136	158	24.49	7.38
D3,4 104-106	187	30.97	7.08
D3,5 44-46	60	31.35	9.83
E3,2 70-72	122	23.08	6.39
E3,3 40-42	299	27.96	8.26
E3,3 100-102	28	35.32	14.59
E3,3 130-132	234	30.33	8.80
E3,4 10-12	24	27.28	8.78
E3,4 70-72	345	28.69	8.52

Table S6. *F. kerguelensis* biometrics at ODP 1090 (42°154.50'S, 8°154.00'E)

Sample	*F. kerguelensis* Measured	*F. kerguelensis* Average length	*F. kerguelensis* length Standard Deviation
B3,1 137-139	36	31.03	7.64
B3,2 19-21	155	30.21	7.71
B3,2 49-51	497	25.61	6.89
B3,2 79-81	394	27.66	7.80
B3,2 109-111	128	28.48	9.21
B3,3 18-20	40	28.50	6.66
B3,3 49-51	330	29.72	8.73
B3,3 79-81	212	30.97	9.09
B3,3 109-111	263	30.52	7.83
B3,3 139-141	152	28.74	8.44
B3,4 19-21	245	29.40	8.82
B3,4 49-51	241	29.49	8.31
B3,4 79-81	347	30.18	8.72
B3,4 109-111	52	28.04	7.72
B3,4 139-141	154	29.94	7.61
B3,5 109-111	34	28.22	7.64
B3,5 139-141	37	33.79	12.78
D3,2 74-76	62	28.50	9.72
D3,2 103-105	19	25.08	5.77
D3,2 133-135	55	36.87	21.64
D3,3 14-16	99	28.43	9.60
D3,3 44-46	51	32.22	10.81
D3,3 74-76	219	30.52	8.97
D3,3 104.5-106.5	490	32.01	10.99
D3,3 134-136	223	31.86	10.87
D3,4 104-106	48	29.25	9.67
D3,5 44-46	41	32.74	12.60
E3,2 70-72	46	29.05	10.06
E3,3 40-42	344	31.10	7.70
E3,3 100-102	6	31.57	5.87
E3,3 130-132	30	32.01	9.18
E3,4 10-12	11	25.94	3.93
E3,4 70-72	134	28.36	11.62

APPENDIX C – STANDARD OPERATING PROCEDURE FOR IMAGING FOSSIL DIATOMS USING A FLOWCAM

Introduction

The Auto-image mode of the FlowCam is well suited to imaging fossil diatoms. Here, we outline a method for imaging fossil diatoms based on the use of relatively clean sedimentary diatoms, where much or all of the clay and other lithogenic and biogenic materials have been removed. The samples were also chemically cleaned to remove external organic matter.

Instrument Setup

Flow Cell Selection – Flow cell size should be larger than base large particle size (in microns). However, the illuminator, magnification lens, and syringe are based on the flowcell size and magnification is inversely related to the cell size. This means that the larger size flow cells result in imaging of a broader range particles but at weaker magnification.

Sample Preparation

Place ~10 mg of dried sample in ~10 ml of deionized water and mix to ensure homogeneity. Use a representative sample to focus the camera. The sample may need to be diluted further if the flowcell becomes jammed or the FlowCam does not capture individual particles.

Camera Settings

See Important Parameters of the FlowCam for settings regarding image quality. In the camera settings, change Shutter Duration until the mean intensity value is between 170 and 180. Change the value of the segmentation threshold to “Only Dark Threshold: 10” (see Segmentation Threshold Test below). Other parameters may be changed, but they are appropriate under the default setting (see Parameterization).

Analyzing a sample

The FlowCam should be prefilled with water, allowing all of the flowcell and half of the sample funnel to be filled. Auto-Image is the mode for running samples without fluorescence. As soon as the calibrations are over and the FlowCam begins to image, quickly pipette < 1 ml of sample into the sample funnel. As the level of water begins to drop, layer milli-q water into the sample funnel. The end point for a trial is adjustable (particle count, sample amount, time, etc...) but can be continuously run and stopped manually. Trials should end when no particles have been captured for around 30 seconds.

Parameterization

Segmentation Threshold, Light and Dark Pixels – This setting, with two parameters, is used to define when the gradient in color between the calibrated background and a potential particle constitutes the outline of a particle. Higher values denote more contrast needed to fulfill criteria. Objects within a sample can differ considerably on how this parameter affects them but changes in range can be

identified after a test trial. Objects will have a dark or light "halo" around them. The goal for best biometric parameterization would be to have outline on the edge of particles, inside of the halos. Recommended values: Only Dark Threshold – 10.

Particle Capture, Distance to Nearest Neighbor – This parameter sets the distance between two outlines, below which the images are considered to be from one particle. Recommended value - 5.

Collage Image Border Padding – This value influences the ability of the FlowCam to “fill in” the particle outline in order to calculate area. When this value is underestimated, it will leave gaps in its fill and underestimate area and when it is overestimated, it will overshoot and artificially inflate the particle outline. Recommended value - 4.

Segmentation Threshold Test

The segmentation threshold distinguishes particles from its calibrated background. When this threshold is too high, the FlowCam will not distinguish all particles that need to be identified. When this threshold is too low, the FlowCam will capture and crop images where no particles exist. The goal for this methodology is to capture all identifiable diatoms while minimizing the number of extraneous (not real) particles. In order to find the best settings for capturing diatoms and measuring biometrics with the FlowCam, a single sample was run in the FlowCam and the data analyzed using different settings. The images from a single FlowCam trial were saved as .RAW files types. This allowed for that trial to be digitally recreated in the FlowCam, allowing processing of particles with different settings. For each setting,

diatoms were identified and counted. The setting of “Only Dark Threshold: 10” was chosen for use in this study because it seemed to capture the most identifiable diatoms and did not create erroneous particles. Under this setting, artifacts that affected the outline of particles seemed to be generally less than under settings with higher thresholds. This benefits biometric measurements as the outline seemed to capture the organism’s actual edge in most cases.

Table S7. Change in total particle count and countable diatoms with changes in FlowCam settings

Dark Threshold	Light Threshold	Total Particles	Countable Diatoms	Notes
50	50	1464	134	
30	30	3215	226	
20	20	4459	256	Outlines are distorted
10	10	N/A	N/A	Does not pick out particles
5	5	N/A	N/A	Does not pick out particles
50	N/A	1364	140	
30	N/A	2592	218	
20	N/A	3425	255	
10	N/A	4322	287	
5	N/A	4775	287	Some images were blank background

BIBLIOGRAPHY

Altabet, M. A., & Francois R., (1994) *Sedimentary nitrogen isotopic ratio as a recorder for surface ocean nitrate utilization*. Global Biogeochemical Cycles 8, 1, 103-116

Assmy, P., Henjes J., Smetacek V., & Montresor M., (2006). *Auxospore formation by the silica-sinking, oceanic diatom Fragilariopsis kerguelensis (Bacillariophyceae)*. Journal of Phycology 42, 1002-1006

Anderson, R. F., Ali S., Bradtmiller L. I., Nielson S. H. H., Fleisher M. Q., Anderson B. E., & Burckle L.H., (2009) *Wind-driven upwelling in the Southern Ocean and the deglacial rise is atmospheric CO2*. Science 323, 1443-1448

Baas-Becking, L. G. M. (1934) *Geobiologie of Inleiding Tot de Milieukunde*. Van Stockum & Zoon, The Hague, Netherlands

Battarbee, R. W. (1973) *A new method for the estimation of absolute microfossil numbers, with reference especially to diatoms*. Limnol. Oceanogr. 18, 647-653

Bennet, J. R., Ruhland K. M., & Smol J. P., (2017) *No magic number; determining cost-effective sample size and enumeration effort for diatom-based environmental assessment analyses.* Can. J. Fish. Aquat. Sci. 74. 208-215

Burckle, L.H., & McLaughlin, R.B., (1977) *Size change in the marine diatom Coscinodiscus nodulifer A. Schmidt in the Equatorial Pacific.* Micropaleontology 23 (2), 216–222

Burckle, L. H., & Cooke, D. W., (1983) *Late Pleistocene Eucampia antarctica Abundance Stratigraphy in the Atlantic Sector of the Southern Ocean.* Micropaleontology 29, 6-10

Buesseler, K. O., (1998) *The decoupling of production and particulate export in the surface ocean*. Global Biogeochem. Cycles 12, 297–310

Cermeño, P., & Falkowski P. G. (2009) *Controls on diatom biogeography in the ocean*. Science 325, issue 5947. 1539-1541

Charan, J., & Biswas T., (2013) *How to calculate sample size for different study designs in medical research*. Indian J Psychol Med 35(2), 121-126.

Cortese, G., & Gersonde R., (2007) *Morphometric variability in the diatom Fragilariopsis kerguelensis: Implications for Southern Ocean paleoceanography*. Earth and Planetary Science Letters 257, 526–544

Crosta, X., Romero O., Armand L. K., & Pichon J. J., (2005) *The biogeography of major diatom taxa in Southern Ocean sediments: 2. Open ocean related species*. Palaeogeography, Palaeoclimatology, Palaeoecology 223, 66–92

Crosta, X., (2009) *Holocene size variations in two diatom species off East Antarctica: Productivity vs environmental conditions*. Deep-Sea Research I 56, 1983–1993

De la Rocha, C. L., Brzenzski M. A., & DeNiro M. J., 1997 *Fractionation of silicon isotopes by marine diatoms during biogenic silica formation*. Geochimica et Cosmochimica Acta 61, 23, 5051-5056

Drooger, C. M. (1978). *Diatoms and silicoflagellates. Micropaleontological Counting Methods and Techniques: An Exercise of an Eight Metres Section of the Lower Pliocene of Cap Rossello, Sicily Zachariasse WJ*. Utrecht Micropaleontology Bulletin 17. 19–46.

EPICA, C.M., Augustin L., Barbante C., Barnes P. R. F., Barnola J. M., Bigler M., Castellano E., Cattani O., Chappellaz J., Dahl-Jensen D., Delmonte B., Dreyfus G., Durand G., Falourd S., Fischer H., Flückiger J., Hansson M.E., Huybrechts P., Jugie G., Johnsen S. J., Jouzel J., Kaufmann P., Kipfstuhl J., Lambert F., Lipenkov V. Y., Littot G. C., Longinelli A., Lorrain R., Maggi V., Masson-Delmotte V., Miller H., Mulvaney R., Oerlemans J., Oerter H., Orombelli G., Parrenin F., Peel D. A., Petit J. -R., Raynaud D., Ritz C., Ruth U., Schwander J., Siegenthaler U., Souchez R., Stauffer B., Steffensen J. P., Stenni B., Stocker T. F., Tabacco I. E., Udisti R., van deWal R.S.W., van den Broeke M., Weiss J., Wilhelms F., Winther J. -G., Wolff E. W., & Zucchelli M. (2004.) *Eight glacial cycles from an Antarctic Ice Core.* Nature 429, 623–628.

Edlund, M.B., & Stoermer E.F., (1997) *Ecological, evolutionary and systematic significance of diatom life history*. Journal of Phycology 33, 6,897–918

Fenner, J., Schrader H. J., & Wienigk H., (1976). *Diatom phytoplankton studies in the Southern Pacific Ocean, composition and correlation to the Antarctic Convergence and its paleoecological significance*. In: Hollister, C.D., Craddock, C. (Eds.), 1976. Initial Reports of the Deep-Sea Drilling Project, U.S. Govt. Printing Office, vol. 35, 757-813.

Fischer, S. (2002). *Automatic Identification of Diatoms*. Dissertation at Universität Bern.

Fluid Imaging Technologies, Inc. (2011) *Zooplankton Analysis using FlowCam*. FlowCam Application Note#108. www.fluidimaging.com

Fryxell, G. A. & Prasad A. K. S. K. (1990). *Eucampia antarctica (Castracane) Mangin (Biddulphiaceae, Bacillariophyceae): Life stages at the Weddell Sea ice edge*. Phycologia 29:27-38

Gersonde, R., Hodell D.A., Blum P., & Shipboard Scientific Party (1999) Proceedings of the Ocean Drilling Program, Initial Reports 177, ch 9, 1-73 http://www-odp.tamu.edu/publications/177_IR/VOLUME/CHAPTERS/CHAP_09.PDF

Gersonde, R., Crosta X., Abelmann A., & Armand L. (2005) *Sea-surface temperature and sea ice distribution of the Southern Ocean at the EPILOG Last Glacial Maximum – a circum-Antarctic view based on siliceous microfossil record*s. Quaternary Science Reviews 24, issue 7-9, 869-896

Horn, M. G., Robinson R. S., Rynearson T. A., & Sigman D. M., (2011A). *Nitrogen isotopic relationship between diatom-bound and bulk organic matter of cultured polar diatoms*. Paleoceanography 26, PA3208

Horn, M. G., Beaucher C. P., Robinson R. S., & Brezezinski M. A., (2011B) *Southern Ocean nitrogen and silicon dynamics during the last deglaciation*. Earth and Planetary Science Letters 310, 334-339

Kaczmarska, I., Barbrick N. E., Ehrman J. M., & Cant G. P., (1993) *Eucampia Index as an indicator of the Late Pleistocene oscillations of the winter sea-ice extent at the ODP Leg 119 Site 745B at the Kerguelen Plateau*. Hydrobiologia 269/270, 103-112

Law, R. A., (1983) *Preparing Strewn Slides for Quantitative Microscopical Analysis: A Test Using Calibrated Microspheres*. Micropaleontology 29, 1, 60-65

Marchetti, A., & Cassar N., (2009) *Diatom elemental and morphological changes in response to iron limitation: a brief review with potential paleoceanographic applications*. Geobiology 7, 419–431

Moore, T., (1973) *Method of randomly distributing grains for microscopic examination*. Journal of Sedimentary Petrology 43. v 4. 3. 904-906

Nelson, D. M., Tréguer P., Brzezinski M. A., Leynaert A., & Quéguiner B. (1995) *Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation*. Global Biogeochemical Cycles 9, 359-372.

Nielson, S. H. H., & Hodell D. A., (2005) *Antarctic ice-rafted detritus (IRD) in the South Atlantic: Indicators of iceshelf dynamics or ocean surface conditions?*. U.S. Geological Survey and The National Academies; USGS OF-2007-1047, Short Research Paper 02

Orsi, A. H., Withworth T., & Nowlin W. D., (1995) *On the meridional extent and fronts of the Antarctic Circumpolar Current*. Deep-Sea Research 42, 5, 64-673

Pichon, J. J., Labracherie M., Labayrie L. D., & Duprat J., (1987) *Transfer functions between diatom assemblages and surface hydrology in the Southern Ocean*. Palaeogeography, Palaeoclimatology, Palaeoecology 61, 79-95

Pichon, J. J., Bareille G., Labracherie M., Labeyrie L. D., Baudrimont A., & Turon J. L. (1992) *Quantification of the biogenic silica dissolution in Southern Ocean sediment*. Quaternary Research 37, 361-378

Popp, B. N., Trull T., Fabien K., Wakeham S. G., Rust T. M., Tilbrook B., Griffiths B., Wright S. W., Marchant H. J., Bridigare R. R., & Laws E. A. (1999) *Controls on the carbon isotopic composition of South Ocean phytoplankton*. Global Biogeochemical Cycles 13. No 4. 827-843

Robinson, R. S., Brunelle B. G., & Sigman D. M., (2004) Revisiting nutrient utilization in the Glacial Antarctic: Evidence from a new method for diatom-bound N isotopic analysis. Paleoceanography 19, PA3001

Sarmiento, J. L, Slater R., Barber R., Bopp L., Doney S. C. Doney, Hirst A. C., Kleypas J., Mateur R., Mikolajewicz U., Monfray P., Soldatov V., Spall S. A., & Stouffer R. (2004) *Response of ocean ecosystems to climate warming*. Global Biogeochemical Cycles 18, 1-23

Sarthou, G., Timmermans K. R., Blain S., & Tréguer P. (2005) *Growth physiology and fate of diatoms in the ocean: a review*. Journal of Sea Research 53,25–42

Schmid, A., (1995). *Sexual reproduction in Coscinodiscus granii Gough in culture: A preliminary report.* In: Proceedings of the 13th International Diatom Symposium (Marino, D. & Montresor, M., editors), 139–159

Schrader, H. J., & Gersonde R. (1978). *Diatoms and silicoflagellates. Micropaleontological Counting Methods and Techniques: An Exercise of an Eight Metres Section of the Lower Pliocene of Cap Rossello, Sicily Zachariasse WJ*. Utrecht Micropaleontology Bulletin 17. 129–176.

Scott, F. J, & Marchant H. J. (2005) *Antarctica marine protists*. Australian Biological Resources Study, Canberra, and Australian Antarctic Division, Hobart

Shemesh, A., Macko S. A., Charles C. D., & Rau G. H. (1993). *Isotopic evidence from reduced productivity in the glacial Southern Ocean*. Science 262. 407-409

Shukla, S. K., Crosta X., Cortese G., & Nayak G. N., (2013) *Climate mediated size variability of diatom Fragilariopsis kerguelensis in the Southern Ocean*. Quaternary Science Reviews 69, 49-55

Shukla, S. K., Crespin J., & Crosta X., (2016) *Thalassiosira lentiginosa size variation and associated biogenic silica burial in the Southern Ocean over the last 42 kyrs*. Marine Micropaleontology 127, 74-85

Shukla, S.K., & Crosta X., (2017) *Fragilariopsis kerguelensis size variability from the Indian subtropical Southern Ocean over the last 42 000 years*. Antarctic Science 29(2), 139-146

Singer, A. J., & Shemesh A. (1995) *Climatically linked carbon-isotope variation during the past 430,000 years in Southern-Ocean sediments*, Paleoceanography 10, 171–177

Sigman, D. M., Alabet M. A., Francois R., McCorkle D. C., & Gaillard J. F., (1999) *The isotopic composition of diatom-bound nitrogen in Southern Ocean sediment*. Paleoceanography 14, 2, 118-134

Studer, A. S., Sigman D. M., Martínez-García A., Benz V., Winckler G., Kuhn G., Esper O., Lamy F., Jaccard S. L., Wacker L., Oleynik S., Gersonde R., & Haug G. H., (2015) *Antarctic Zone nutrient conditions during the last two glacial cycles*. Paleoceanography 30, doi:10.1002/2014PA002745

Sutton, J., Ellwood M. J., Maher W. A., & Croot P. L. (2011) *Oceanographic distribution of inorganic germanium relative to silicon: germanium discrimination by diatoms.* Global Biogeochemical Cycles 24, 2

Sutton, J. N., Varela D. E., Brzezinski M. A., & Beucher C. P. (2013). *Species-dependent silicon isotope fractionation by marine diatoms*. Geochimica et Cosomchimica Acta 104. 300-309

van der Spoel, S., Hallegraeff G. M., & Soest R. W. M., (1973*). Notes on variation of diatoms and silicoflagellates in the South Atlantic Ocean*. Netherlands Journal of Sea Research 6 (4), 518-541.

Wilken, S., Hoffmann B., Hersch N., Kirchgessner N., Dieluweit S., Rubner W., Hoffman L. J., Merkel R., & Peeken I., (2011) *Diatom frustules show increased mechanical strength and altered valve morphology under iron limitation*. Limnology and Oceanography 56(4), 1399–1410

Zielinski, U., & Gersonde R., (1997) *Diatom distribution in Southern Ocean surface sediments (Atlantic sector): Implications for paleoenvironmental reconstructions*. Palaeogeography, Palaeoclimatology, Palaeoecology 129, 213-250