

Bachelor Thesis:  
Analysis of diatom species composition and size distribution in the  
Southern Ocean

Fabian Altvater  
University Bremen  
Matr.-Nr.: 2398258

Supervisor:

Dr. rer nat Bánk Beszteri  
Alfred-Wegener-Institute, Helmholtz Center  
for Polar- and Marine Research

**Bremen, June 2014**

## Content

<b>1</b>	<b>Index of abbreviations and technical terms .....</b>	<b>3</b>
<b>2</b>	<b>Abstract .....</b>	<b>4</b>
<b>3</b>	<b>Introduction .....</b>	<b>5</b>
<b>4</b>	<b>Material and Methods .....</b>	<b>12</b>
4.1	Sampling.....	12
4.2	Slide preparation of net samples.....	13
4.3	Automatic light microscopy .....	14
4.4	Digital Image Processing.....	16
<b>5</b>	<b>Results.....</b>	<b>20</b>
5.1	Automation yields .....	20
5.2	Species composition .....	22
5.2.1	Automatic species detection (SHERPA Software).....	22
5.2.2	Manual species detection (Metafer System).....	23
5.2.3	Species composition in total .....	23
5.3	Size distribution.....	26
<b>6</b>	<b>Discussion .....</b>	<b>33</b>
<b>7</b>	<b>Example plates of dominant Species.....</b>	<b>37</b>
<b>8</b>	<b>Index of Tables.....</b>	<b>38</b>
<b>9</b>	<b>Index of Figures .....</b>	<b>39</b>
<b>10</b>	<b>Literature .....</b>	<b>41</b>
<b>11</b>	<b>Declaration .....</b>	<b>44</b>
<b>12</b>	<b>Dedication.....</b>	<b>45</b>

## 1 Index of abbreviations and technical terms

ACC	Antarctic Circumpolar Current
Alveolus	Diatom feature. An elongated chamber or one of a series of elongated chambers running from the axial or central part of the valve to the margin, opening to the inside of the frustules by a large opening and with areolate outer layer.
APF	Antarctic Polar Front
Areolae	Diatom feature. The regularly repeated perforation through the basal siliceous layer normally occluded by a velum or rica, or both. Small areolae are often termed puncta.
F*	$F^* = \frac{length}{width} * width \text{ of } 5 \text{ costae (in } \mu\text{m)}$ (after Fenner et al. 1976)
FOV	Field of view for capturing microscopic pictures
Rica	Diatom feature. A very thin perforate layer of silica, the perforations usually circular and not more than 15 nm in diameter, often decussately arranged.
SAF	Sub Antarctic Front
Seta(e)	Diatom feature. Hollow outgrowth of valve projecting outside the valve margin, with different structure from the valve.
Stack picture	Calculated picture from different focus planes. Microscopic pictures from a single coordinate are taken in several focus planes. The focus planes are stacked together to a single extended focus depth picture.
STF	Sub Tropical Front
Striae	Diatom feature. A row of areolae or alveoli, or single alveolus where this is not part of a row.
Velum	Diatom feature. A thin perforate layer of silica across an areola, the perforations normally not less than 30 nm across and often not circular.
APF	Antarctic Polar Front

## 2 Abstract

In this work, diatoms from phytoplankton net samples from four stations from a transect southward from Cape Town and five from an eddy at the southern boundary of the Antarctic Circumpolar Current, obtained during Polarstern cruise ANTXXVIII-3, were analyzed. Material was cleaned in an acidic water bath and permanent slides, containing diatom frustules, were prepared.

The permanent slides were automatically scanned in similar areas using an automated microscopic image acquisition system. The pictures were analyzed using image analysis software developed for detecting and characterizing diatom outlines, as well as manually, in order to test the performance of the software. The test revealed a strong taxonomic bias against large-celled and thinly silicified taxa. Dominant taxa after combining automatically and manually obtained diatom counts were *Fragilariopsis kerguelensis* and the genera *Pseudonitzschia* and *Chaetoceros*.

For *Fragilariopsis kerguelensis*, statistically significant morphometric differences between the populations from the transect vs. from the eddy were found, the ecological interpretation of which will require further study.

### 3 Introduction

The polar regions play an important role in the global climate- and biological systems. The Antarctic ecosystem is, for example, influenced by ice formation or melting and therefore heat exchange is a factor this region is important for. Also this part of the ocean is a habitat for a unique variety of organisms.

The sun and the atmosphere influence almost any dynamical processes in the ocean, directly or indirectly (Steward, 2008).

The connection to the great oceans is another major role in case of the global exchange of water masses and material moving along (organic, inorganic).

The global exchange between air and water masses is driven by convergence zones, occurring due to varying abiotic factors, e.g. sun light. Highest input occurs around the equator, leading to high evaporation rates. Warm air comes up and moves to the north or south, transferring much water to heavy rain falls. These are phenomena within the subtropics, between 40°N and 40°S, divided by the inner tropic convergence zone. This zone is mostly windless and not directly laying on the equator, but a little shifted to the north (Fahrbach, Rohardt; 2009).

The Earth's rotation creates the Coriolis force. Any biotic on earth is adapted to this force and the abiotic factors are influenced or even created by it. The warm upcoming air from the inner tropic convergence zone moves to cooler regions north or south. These movements are shifted in angle around 45°, due to the Coriolis force. On the northern hemisphere the direction turns right, on the southern to the left (Steward, 2008).

The Antarctic ecosystem differs from the Arctic because of various reasons. One major difference would be the connection to the oceans which is very little for the Arctic compared to the Antarctic. Water masses are not interrupted by landscape, the major part of global landscape is placed on the northern hemisphere. In common the region around the Antarctic Continent is called Southern Ocean by many Oceanographers, even if the International Hydrography Bureau does not officially declare this name (Tomczak, Godfrey; 2001).

Moving southwards from the equator strong westerly winds occur (Medlin, Priddle; 1990) and due to constancy the region around 40° S is called "Furious Fifties" and the region around 50° S "Roaring Fifties" (Hagen, Auel; 2009).

Due to no interruption around Antarctica and the west wind drift the Antarctic Circumpolar Current (ACC) is formed. This current comprises a series of bands of eastward flow and is the largest current system on earth, carrying  $135 \times 10^6 \text{ m}^3/\text{s}$  of water from east to west over a distance of around 20,000 km, around the continent (Fahrbach, 2009).

The ACC is affected by the Drake Passage, the Scotia Ridge that connects Antarctica with South America and the outflow from the Weddel Sea gyre. These factors lead the current to the north and later eastwards to the area where the ACC leaves the Scotia Sea. This region is subscribed by high mesoscale variability, where fronts (e.g. Polar Front) could alternate over distances up to 100 km in 10 days and a significant eddy formation takes place. These variability's are likely to have strong effects on the biotic interactions in parts of the southern ocean (Medlin, Priddle; 1990).

An Eddy is a circular movement of water usually formed where currents pass obstructions or between two adjacent currents flowing in opposite directions, or along the edge of a permanent current (Garrison, 2010)

Along the Antarctic circumpolar current energetic ocean eddies are formed, which could have a horizontal scale of the order of 100 km. Thus these dynamics have not been well understood, there is strong evidence that they have an important role (Summerhayes, Thorpe; 2002)

Little is known about eddy formation, their frequency or life expectancy of individual eddies. Satellite altimeter observations indicate that eddies are not strictly distributed along the ACC but are more frequent east of the Scotia ridge (Tomczak, Godfrey; 2001).

Eddies occur clockwise or anticlockwise respective cyclonic or anticyclonic and can have up- or downwelling character. Therefore carbon and heat exchange is a point of interest for many oceanographers.

The Organisms populating the Oceans could be divided by two groups, the plankton and the nekton. The nekton are the Organisms, or Animals, that are able to move themselves against streams. Planktonic movements are, as far as possible, restricted to vertical migration (Paululat, Purschke; 2011).

Plankton could be roughly divided by two groups, the phyto- and the zooplankton. The Phytoplankton represents all autotrophic organisms, chlorophyll- or any photosynthetic-pigment carrier. They are the primary producers in any aquatic system and are mostly unicellular forms of living. Therefore they are found in the sunlit surface layer, the pho-

tic zone (Nöthig et al, 2009). The zooplankton are grazers feeding on the autotrophic organisms, e.g. Krill (*Euphasia superba/antarctica*).

The Arctic and Antarctic waters inhabit the same groups of primary producers, in descending order of biomass they are: diatoms, autotrophic nanoflagellates and dinoflagellates. Photosynthesis could also take place in heterotrophic unicellular organisms by endosymbiosis or kleptoparasitism of chloroplasts, which could be performed by ciliates, tintinnids, foraminiferans, acantharians or radularians. Mixotrophic forms, photosynthetic and feeding on unicellular organisms, are found in the group of dinoflagellates (Nöthig et al, 2009).

Diatoms are microscopic unicellular algae with a range in size from approximately 5 – 500 µm, living wherever there is moisture or even soil (Barber, Haworth 1981). In aquatic habitats they appear as pelagic (floating) and benthic (sessile) forms. These organisms possess two silica valves, the epitheca and hypotheca, fitting into each other like shoe boxes. Both valves are enveloped by at least one girdle-band attached one to another. During growth of cells new girdle bands are built. Regular cell division is performed by building new valves from the inside into the old valves respectively the latest girdle. This means new valves are fitted into the old epi- and hypotheca so the new ones must be smaller than the old ones. With each cell division the daughter cell becomes smaller. This leads to a steady decrease in size, known as the ‚Mac-Donald-Pfitzer-rule‘ (Crawford, 1980, Mac-Donald & Pfitzer 1869). At a specific size, meiosis can take place. The diatoms change to a sexual reproduction and mate. The zygote grows into an initial cell, which begins a new cycle (Bohnert, 2012). Accordingly, size distribution of a diatom population might provide information about ecological processes in relationship with life cycle (Crawford, 1980)

The Phytoplankton in marine nutrient rich water is usually dominated by diatoms (Miller, 2004).

Planktonic marine diatoms produce 25 % of the global primary production and are therefore very important to the marine food-web and to the global biogeochemistry as well (Hamm, 2009).

The Southern Ocean is a High Nutrient Low Chlorophyll (HNLC) area which is characterized by primary production that is lower than expected for available nutrients and macro-nutrients. For this circumstance three explanations have been discussed. One discusses the light intensity that may not affect phytoplankton to the optimum. Another argues about predation pressure of zooplankton and third a limitation of micro nutrients.

The limitation of micro nutrients is most commonly accepted and refers to the “Iron hypothesis” of John Martin (Cortese, Gersonde; 2007).

This Hypothesis was proved by various Iron fertilization experiments (EisenEx, Eifex, LohaFexI, LohaFexII). All these field-experiments confirmed the premise of the Iron hypothesis. An immediate increase of phytoplankton growth and a following increase in biomass. In any cases diatoms supplied the greatest part of the induced blooms, and got slightly regulated by following predation pressure. During EisenEx (2000) and EIFEX (2004) the accumulation rates of diatoms were high, even though the upper mixing layer reached 80 to 100 m. In both experiments predation pressure could not depress the phytoplankton bloom. Thereby the predation hypothesis was refuted (Smetacek, 2009).

The Southern Ocean underlies seasonal changes like many biomes. Nutrient rich water comes up from the deep, meeting receding ice providing high values of major nutrients. A phase of epontic algae is replaced by a strong bloom dominated by diatoms. These blooms follow the melting ice southwards, supported by micro nutrients like iron that accumulated in form of dust during the ice covered period. North from the winter ice zone close to the Antarctic convergence zone upwelling also provides high nutrient levels and may lead to spring blooms in November and December (Miller, 2004).

Samples observed in this study were collected during Polarstern cruise XXVIII-3 between 07.01.2012 – 11.03.2012, from Cape Town to Punta Arenas. Two main research projects were carried out, “Eddy-Pump” and “SYSTCO II” with several scientific groups working on each theme.

The “Eddy-Pump” research was an interdisciplinary study of physical, chemical and biological processes influencing the carbon exchange in the southern ocean, in particular within mesoscale eddies.

The “SYSTCO II” short form stands for “System coupling in the deep Southern Ocean II” with three major goals. First, to study the diversity, distribution and abundances of deep-sea organisms in relation to surface water productivity and sedimentation of organic material. Second goal was to study the ecology of the deep-sea fauna in general. Third part was a DNA preservation in the water column and the deep-sea sediments and its possible use to study the eukaryotes diversity in the present and past.

Both projects complemented each other and therefore took place in joint stations.

The role of the southern ocean in the global carbon cycle is undisputed but controversially debated whether it would be a source or a sink to CO<sub>2</sub> incase of changing climate nowadays (Fahrbach, 2011).

A list with Stations is shown in Table 1 and the cruise track plot is given in Figure 2.

The biological carbon pump is driven by photosynthesis in the photic zone and the export of bound organic carbon to the deep-sea by sinking (Smetacek, 1985). The physical carbon pump is often related to upwelling cold deep-water carrying up CO<sub>2</sub>. The change of the oceans quality being sink or source to carbon is related on the delicate balance between the biological, physical and chemical interactions. Latitudes between 50°S and 60°S were investigated, thus in this region upwelling deep-water masses interact with the atmosphere at the sea surface before they are sub ducted at the sub-antarctic-front. To get a more detailed picture of the carbon pumps during a significant part of the growth season, the studies were repeated in vertically coherent eddies. Eddies were chosen in a way they represent presumed different oceanographic and biogeochemical regimes in the Atlantic Sector of the Southern Ocean (Fahrbach, 2011).

Two diatom provinces are predicted for the Atlantic sector of the Southern Ocean, the genus *Chaetoceros* and *Fragilariopsis kerguelensis* provinces, shown in Figure 1.

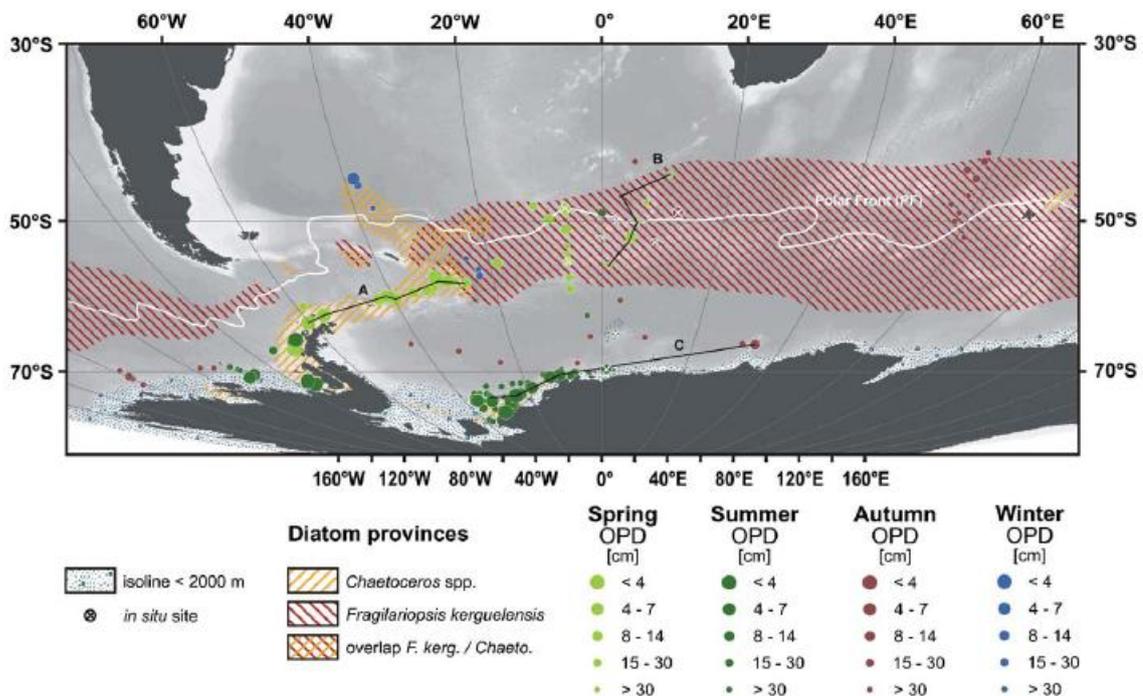


Figure 1 The orange hatched area shows enhanced *Chaetoceros* spp. spores (>20%) in surface sediments. Red hatched area shows elevated *Fragilariopsis kerguelensis* frustules (>40%). Colored dots and black lines represent various data from the origin study e.g. the spatial distribution of the oxygen penetration depth in sediments (OPD) (Figure from Sachs et al 2009)

The region south of 50°S and approximately 30°E got enhanced eddy kinetic energy, associated with a southward veering of the Polar-Front at the eastern extend of the Weddel Gyre it is expected to have many (mesoscale) upwelling events with corresponding iron supply from the deep. Despite the relatively high iron input into the photic zone this region is assumed to be dominated by *Fragilariopsis kerguelensis*.

Around 0°E at the Polar Front the Antarctic circum polar current is moving along, creating mesoscale eddies with low to moderate eddy kinetic energy. The according iron supply leads to transient phytoplankton blooms with alternating dominances of *Fragilariopsis kerguelensis* and the Genus *Chaetoceros*.

The region south of South Georgia is assumed to be dominated by the Family Chaetoceros. Close to the Antarctic Peninsula constant iron input, fostered by lateral advection from the shelf, is leading to high long term mean sea surface chlorophyll concentrations. The Drake passage above implements a strong north-south gradient of eddy kinetic energy and low chlorophyll rates at the sea surface (Fahrbach, 2011).

From the past diatoms are quite well studied and described in almost any habitat (Hustedt, Hasle, Barber, Hayworth, etc.) and many scientific cruises concerning diatoms or phytoplankton in general took place nowadays.

In this study an automated microscope was used, provided by the Company Metafer Systems (Altussheim). Any System to order is modular and can slightly be adapted to the costumers demand. Automated microscopy is not bound to a single discipline in biology. Various questions could be answered in new ways. And various studies using Metafer Systems have been made. For example scoring of human lymphocyte micronuclei, concerning mutagen genotoxin exposure, including an image cytometry system was studied (Rossnerova et al, 2010), detection and relocation of rare events in red blood cells were investigated (Hennerbichler et al, 2002) or so called 'FISH-Stained HER2/neu' cells have been analyzed automatically (Schunk, Muhammad; 2011).

The automated microscopy generates many pictures in a high resolution. In this study additional Software (SHERPA) was used to analyze these pictures. The SHERPA Software was programmed at the AWI for detecting, outlining and analyzing diatoms or other objects (Kloster et al, submitted, 2013). Objects are detected by comparison of pixels over gray-scale levels. At adjustable contrast values an outline is created for the object. Within this outline various parameters could be calculated, e.g. the longest axis through the object or the total area. Details of measurements are noted in 3.4.

Among diatomists the automated detection and identification is of great interest. In 2002 du Buf & Bayer published “Automatic Diatom Identification (ADIAC)”. The idea was to create large and specific data bases, public available. These should include classic and new feature extracting methods, e.g. for valve contours and/or ornamentation. A need for a mass analyzing Software was mentioned (du Buf, Bayer, 2002).

In 2004 Stoermer published a Paper concerning the extraction number of striae in diatoms, titled “A simple, but useful, application of image analysis”. The described Software was the base for the later occurring Image analysis software “ImageJ” or further “FIJI”.

This study analyzes the diatom species composition and the size distribution in the southern Ocean. The applied techniques and software are new to diatom investigations so various questions should be answered in this study. On the one hand these questions concern the title of the study. On the other hand questions are asked how well the new techniques support the investigation in this study and for diatoms in general.

Question 1: What kind of detection rate can be obtained by automation and what kinds of biases occur? Are all diatom taxa detected at the same rate and accordingly does the taxonomic composition obtained from the automated method reflect the real composition of the communities?

Question 2: Does the diatom composition differ between the stations or the station groups and how do relative abundances of single Taxa change?

Question 3: Does morphometry in dominant species change between the stations or station groups?

## 4 Material and Methods

### 4.1 Sampling

Samples were obtained during Polarstern Cruise XXVIII-3 from 07.01.2012 – 11.03.2012. A list with coordinates is given in Table 1.

Table 1 A list of investigated stations in this study with coordinates. Samples were collected during Polarstern cruise ANTXXVIII-3 and permanent slides containing diatoms were prepared from hand-net samples, taken from surface waters. Stations are clustered in Groups, listed in the last column

Station	PositionLat	PositionLon	Group
PS79/069-3	48° 0,17' S	9° 59,55' E	Transect
PS79/072-4	49° 0,04' S	10° 0,35' E	
PS79/078-2	50° 59,97' S	10° 0,07' E	
PS79/081-4	52° 1,38' S	10° 0,57' E	
PS79/087-2	50° 50,20' S	13° 9,25' W	Eddy 1
PS79/095-4	51° 11,46' S	11° 59,87' W	
PS79/108-2	50° 47,95' S	13° 19,72' W	
PS79/122-3	51° 36,19' S	12° 0,27' W	
PS79/133-1	50° 23,99' S	13° 20,05' W	Eddy 2
PS79/144-3	50° 47,67' S	36° 59,32' W	
PS79/155-2	50° 48,44' S	37° 35,13' W	
PS79/162-3	50° 48,01' S	39° 24,67' W	
PS79/166-2	49° 12,04' S	39° 24,21' W	

At the so called ‚deep CTD-Stations‘ hand-net samples were obtained from surface waters. The net with a mesh size of 20  $\mu\text{m}$  and a connected collection box with outlet at the end were gently pulled up from 10 m depth.

Through the outlet the liquid, sometimes greenish colored, sample was poured into a measuring cylinder. If necessary the net was rinsed with filtered seawater to obtain remained material. The gained volume in total was then fixed with Formaldehyde in a concentration of 9:1.

In this study thirteen stations were investigated, roughly divided by three groups. First group is a transect from north to south, second group in the Eddy like structure around  $^{\circ}\text{N} / ^{\circ}\text{E}$  and third in the Eddy like structure above South Georgia. A list with details is given in Table 1, the cruise plot with chosen stations is shown in Figure 2.

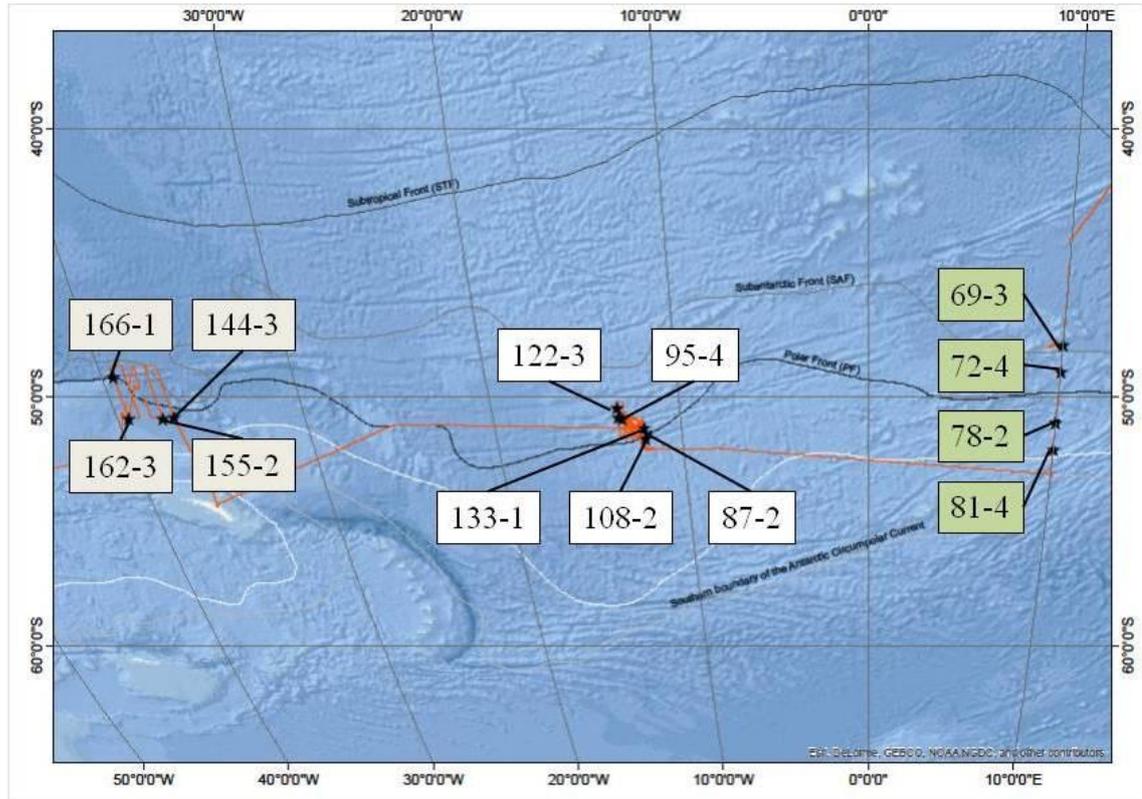


Figure 2 Track plot of Polarstern cruise ANTXXVIII-3 from Cape Town to Punta Arenas is shown as a red line. The Subtropical Front (STF), the Sub Antarctic Front (SAF), the Polar Front (PF) and the southern boundary of the Antarctic Circumpolar Current are shown. Colored boxes mark the investigated stations in this study. Green boxes are marking the station group “Transect”, white boxes show the group “Eddy 1” and grey boxes mark the “Eddy 2” group. Diatoms were prepared from hand-net samples, taken from surface waters

Fixed samples were stored at 4°C onboard and returned to the AWI in Bremerhaven.

## 4.2 Slide preparation of net samples

The fixative was washed out using de-ionized water. Samples were centrifuged in an Eppendorf centrifuge 5804 with an A-4-44 swing out rotor at 2.000 rpm for 5 minutes. The supernatant was discarded. Tubes were refilled with de-ionized water, mixing the sample material. This was repeated five or six times, depending on the pH (measured with pH-stripes). Washing was completed when the sample pH reached about 5.5.

After the final centrifuging step, most of the supernatant was discarded. To the remaining sample, the same volume of kalium-permanganate was added, tubes were filled to around one fourth of volume. This caused a dark purple color and the samples were incubated over night.

Next, the mixture of kalium-permanganate and sample was filled up with the same volume of HCl (37%) and gently cooked in a water-bath for 30 minutes. This procedure turned the sample color back into a transparent yellow. The remaining HCl was washed out in five steps, performed as described above for formaldehyde.

Two types of round coverslips, with diameters of 18 mm (for LM) and 8 mm (for SEM), were washed in pure ethanol and dried on a heating plate.

The samples were mixed with a pipette five times. Then 300  $\mu$ l of sample was transferred to a cleaned 18 mm coverslip. Two to three drops of sample were placed on the 8 mm coverslips.

Samples dried within 30 minutes.

SEM-stubs were covered with sticky foil and the 8 mm coverslips were attached to that, carrying the sample on the upper side. LM Coverslips were covered with naphrax dissolved in toluol, and left on the heating plate at 60°C over night in order to evaporate toluol.

Next, clean microscopic slides were shortly heated up in their center using a Bunsen burner for about three to four seconds. Coverslips were picked up with the hot slides. Naphrax becomes slightly liquid due to the heat of the slide and the coverslip gets attached to it. Slides containing the coverslips were reheated over night so the mountant was allowed to level out and gas bubbles could escape or be pressed out easily. Thick layers of naphrax, mostly directly surrounding the coverslip, were cut off using a razor-blade and then cleaned with butylacetate and ethanol. Slides were labeled, listed and stored in the Hustedt Collection.

The sample bottles from Station 144, 155, 162, 166 had a lot of copepods and other zooplankton in it (e.g. picture?). Slides from those Stations showed a very low yield of diatom valves so they were prepared a second time but with similar low results. For the third preparation that much organic material was taken out so the HCl treatment (described above) reacted to out spilling foam. One would expect that most material was flushed out that way and the prepared slides proved that. The fourth preparation was the final one but still with low numbers of diatom valves in these four stations (see Results).

### **4.3 Automatic light microscopy**

The Metafer System consists of a Zeiss AxioImager Microscope with an automated object table, directable in x,y and z-axis, a connected Camera and the supporting software.

This system provides various automated features in picture taking along all directable axes. Therefore a kind of scanning is possible, obtaining pictures of a whole slide, part of a slide or manual position lists. Within a scan various features can be applied, for example object detection can be programmed by various parameters. Scanning an area means that pictures are taken in a raster, but with overlapping margins. A single field in this raster is called field of view (FOV). All FOV's can be displayed as virtual slides, e.g. in Figure 3.

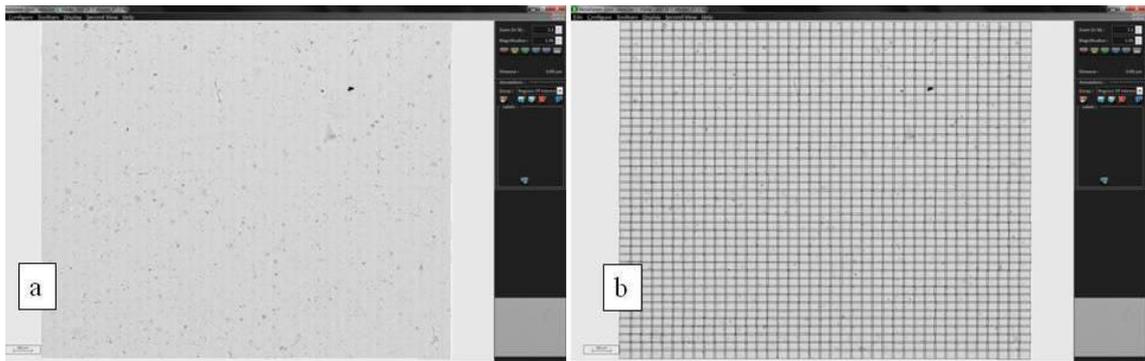


Figure 3 a) A screenshot of a virtual slide, displayed by the Software from Metafer Systems. The virtual slide consists of 1,638 single pictures called field of view (FOV). Figure a shows the whole slide. b) Figure b displays the same slide but with the raster of FOV's

To perform a scan of chosen FOV's the magnification has to be defined by/over the objective. A number of focus planes to take pictures of and the distance between these focus planes can optionally be defined. The pictures of different focus planes can then be stitched together to a stack-picture and are also part of the virtual slide. These pictures mostly provide more information than a single picture could.

Within the Virtual Slides annotations can be set manually in form of squares, circles or polygons, for example (Figure 8). These annotations could be exported as single pictures. In this study parts of permanent slides were scanned in areas consisting of ca. 2,000 pictures, positioned on the slide where densities of valves were adequate. In this case an adequate number of diatom valves were described by having enough valves on a slide to do statistical investigations and less enough to find valves singled, in valve view, not overlapped by other valves or debris. If valves or parts of them are covered by other material identifying is impeded or impossible. In this study a Plan Apochromat 63x Objective was used. For each coordinate pictures of 20 focus planes with a distance of  $1/4 \mu\text{m}$  were taken. No object detection by the Metasystem was applied. The created Images were analyzed by software named SHERPA.

#### 4.4 Digital Image Processing

The SHERPA Software is developed for detecting diatom valves in microphotographs, analyses and extracting morphometric measurements of them (Kloster et al. submitted). Outlines of objects are created over grey-scale comparison of pixels. The sum of pixels leads to an object which gets various values like an area and the maximal/minimal axis through this object for example. For pennate diatoms, correctly outlined, these axes would represent the apical- and transapical axis of a frustule precisely. Centric diatoms with a clear round shape would have the same value in both these axes.

Another feature within this software is the detection of striae and the average distance between them.

Precision of the average costae distance was not re-measured but gives a comparative idea of costae densities. It is used just for the species *Fragilariopsis kerguelensis* and should not be read in single numbers but as a value of comparison for this high frequented group of the Southern Ocean.

The software provides an image processing workflow, as an overview shows in Figure 4.

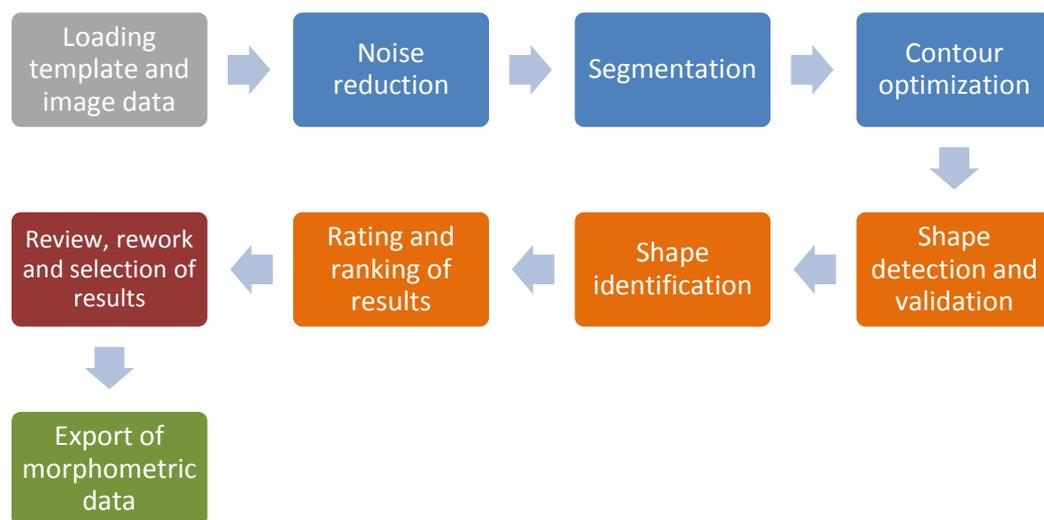


Figure 4 Basic Workflow of the computer software SHERPA (figure from Kloster et al. submitted)

Detecting contours and shape identification are two of the major features provided by SHERPA. Outlines of objects found are compared to a library of various diatom outlines called the templates. The comparison is based on elliptic Fourier descriptors, a mathematical approach for outline characterization (Claude 2008).

In this study, 465 template outlines were used, based on du Buf & Bayer (2002) Barber & Haworth (1981) and a set of templates created by the AWI. (e.g. Figure 5). The created

outline for an object is compared to the set of templates, resulting in various values matching or mismatching the best fit template (e.g. Figure 6)

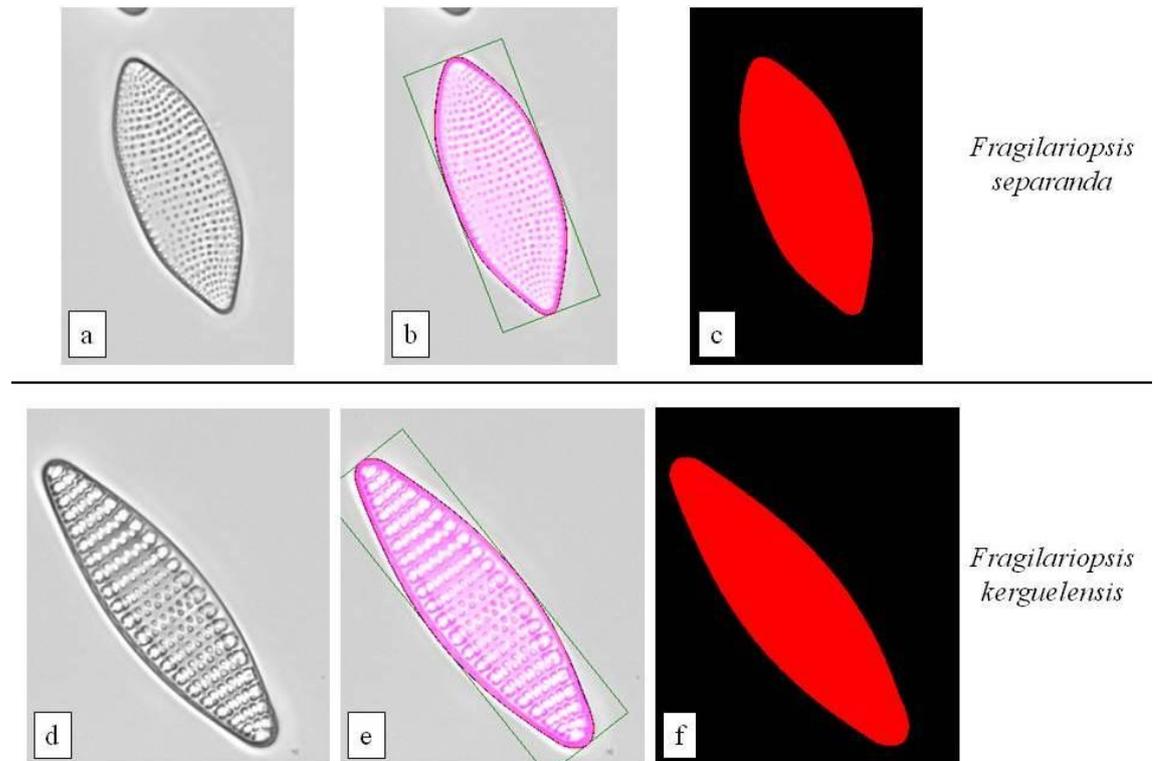


Figure 5 Two templates used by the Software SHERPA. Upper half shows a template and the related species *Fragilariopsis separanda*, the lower half shows the template for *Fragilariopsis kerguelensis*, also with an example picture, in comparison. The software SHERPA creates an outline for an object which is then compared to a set of templates, resulting in various values matching or mismatching the best fit template

SHERPA

Templates Images Objects Export Settings Help

Templates Images Objects

Validation Parameters for Contours:		Contour Optimization:		Matching, Quality and Convexity Indicators Used for Ranking:			
40	< Area/ $\mu\text{m}^2$	< 100000	<input type="checkbox"/> Apply Contour Optimization	EFD's Match with Templ	< 0.0002	<input type="checkbox"/> Use Convexity	<input type="checkbox"/> Force Convexity
20	< Perimeter/ $\mu\text{m}$	< 100000	<input type="checkbox"/> Configure Contour Optimization	Hu Match for EFD's Templ.	< 0.2500	CFR: 1.05	CDF-Match: > 0.800
	Distance to Image Border	$\geq 3$ Pixel		Std. Dev. of inner 50%	$\geq 11.00$	PCAF: < 2.00	PCAF-Match: > 0.600
	Examine only Central Area	$\leq 100$ %		Width/Height Ratio	> 1.00 < 15.00	FORMF: 20.00	Compctn-Match: > 0.800
Select Segmentation Procedures				Contour Smoothness	> 0.950 $\sigma$ 5.00	Only Show Results with Ranking Index $\leq$ 5	
Save Settings		Restore Settings		Form-Factor	> 0.150	<input type="checkbox"/> Show Heuristic Descriptors	

File	Folder	Area / $\mu\text{m}^2$	Perimeter / $\mu\text{m}$	Width / $\mu\text{m}$	Height / $\mu\text{m}$	Costas Distance / $\mu\text{m}$	Segmentation Method	Optimization Method	Template (best EFD's Match)	Best EFD's Match	Hu Match for EFD's Template	Std. Dev.	Width/Height Ratio	Contour Smoothness	Form-factor	Quality Index	Template is convex	CDF	PCAF	CHMDF
ANT28-3_72-4.000326.cropped.tif	C:\Fabi...	375.58	116.41	53.03	9.60	1.47	Otsu's Thresholding	not optimized	Fragilariopsis kerguelensis 35.tif	2.521E-5	0.195	41.327	5.525	0.994	0.348	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	1.039	2.36	60.84
ANT28-3_72-4.000325.cropped.tif	C:\Fabi...	262.69	99.03	44.46	8.60	1.58	Histogram Equalization + Otsu	not optimized	Fragilariopsis kerguelensis 22.tif	2.619E-5	0.110	56.258	5.169	0.989	0.337	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	1.049	3.01	71.18
ANT28-3_72-4.000229.cropped.tif	C:\Fabi...	174.80	73.42	33.11	7.44	1.67	Histogram Equalization + Otsu	not optimized	Fragilariopsis kerguelensis 7.tif	2.693E-5	0.409	63.267	4.449	0.994	0.407	1	<input checked="" type="checkbox"/>	1.034	1.82	2.30
ANT28-3_72-4.000243.cropped.tif	C:\Fabi...	323.53	91.68	40.51	11.00	1.80	Adaptive Thresholding	not optimized	Navicula lanceolata 549.tif	2.694E-5	0.279	43.445	3.683	0.997	0.484	1	<input checked="" type="checkbox"/>	1.033	1.10	3.09

Figure 6 Screenshot from the Software SHERPA showing various values from analyzing. Upper half shows the provided settings for parameter's, lower half shows the fitting results from comparing a set of diatom pictures to a related set of templates

Images obtained using Metafer were processed with SHERPA to detect any objects matching this extensive template set (all diatoms). From SHERPA, images cropped to

the region around each detected object were exported and these were manually sorted into species. Images for individual groups were analyzed again with SHERPA to obtain their morphometric feature values. These measurements were calibrated using the pixel scale from the Metafer system so that values on an absolute scale (in  $\mu\text{m} / \mu\text{m}^2$ ) could be obtained. Species were identified using Hasle (various), Hustedt (various), Medlin & Priddle (1990) Scott & Marchant (2005). In each sample, a number of cells could not be identified at the species level due to different reasons. In some cases, individual valves were not focused correctly. In such cases, when identification was not possible based on valve outline (which is usually visible even in not well focused images), the image was sorted into the unidentified category.

For species identification, I also used scanning electron microscopy. For this, stubs prepared as described above were sputter coated with gold and observed in a FEI Quanta FEG 200 scanning electron microscope (FEI, Munich, Gräfelfing, Germany) at an accelerating voltage of 10 kV.

Any Outline displaying a diatom valve was re-imported into the Virtual Slides (Figure 7). The source picture from the MetaSystem became coordinates as content on which the SHERPA-created outline could be related to.

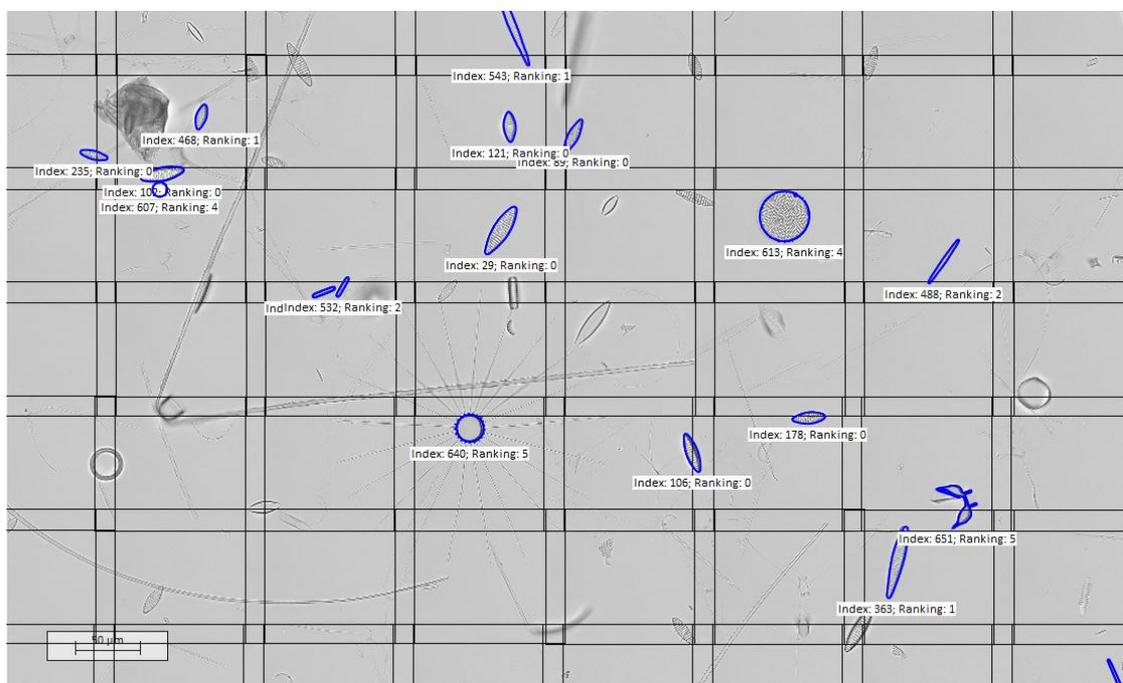


Figure 7 Screenshot of a virtual slide, displayed by the Software from Metafer Systems. Blue outlined objects represent annotations that were re-imported from the Software SHERPA. The black lined raster shows the overlaying FOV's

This Software update got possible because of a request by the AWI and a good cooperation between both parties in following. The feature of annotations was the point of intersection where the SHERPA outline could be placed. With the basis coordinates for a FOV the Object in it could be re-calculated by distance to borders of the image using the absolute pixel scale.

Six Virtual Slides (Station 69-3, 72-4, 87-2, 81-4, 87-2, 133-1) were manually screened, FOV by FOV. Cells not outlined by the re-imports from SHERPA were manual marked with annotations. Three groups were created: “not found” for cells in one FOV and good optic (Figure 8), “bad optic/angle” for cells in one FOV that were not found because of a non identifiable angle or focus (Figure 8) and the third group called “bigger two pictures” for cells laying in two or even more FOV’s (Figure 8).



Figure 8 Screenshot of a virtual slide, displayed by the Software from Metafer Systems. Blue outlined objects represent annotations that were re-imported from the Software SHERPA. Other colored marks are manual annotations in various forms. Green labels show the group “not found”, yellow labels are the group of “bad optic/angle” and the magenta colored labels are “>2pictures” (explanation see Text). The black lined raster shows the overlaying FOV’s

## 5 Results

### 5.1 Automation yields

Over all thirteen Stations 17,528 pictures were taken by the Metafer System. Nine slides were scanned automatically in an area of around 2,000 FOV's (see 3.3 Automatic Light Microscopy (Metafer System)) and displayed as virtual slides (69-3, 72-4, 78-2, 81-4, 87-2, 95-4, 108-2, 122-3, 133-1).

Four slides (144-3, 155-2, 162-3, 166-2) were analyzed manual in case of low densities of diatom valves. The coordinates from around 100 valves were marked and safed as so called manual position lists. The yields of pictures from the automation are given in Table 2. The numbers of SHERPA detected objects from these pictures are included.

Table 2 List of observed Stations during Polarstern Cruise ANTXXVIII-3 and the number of taken pictures (FOV = field of view) by automated microscopy (Metafer Systems) in permanent slides, prepared from hand-net samples taken from surface waters. The number of detected objects by the Software SHERPA and the counts of manual observed slides are listed. The scanned area was calculated and the total yield of valves is displayed. The percentage of SHERPA detected objects is given in the last column

Station	Metafer System (FOV)	SHERPA detected Objects	Scanned Area (mm2)	Total yield of valves	SHERPA Percentage
69-3	1202	120	16.407,30	232	51,72
72-4	2040	492	27.846,00	1085	45,35
78-2	2032	587	27.736,80	2305	25,47
81-4	2042	761	27.873,30	1764	43,14
87-2	1638	177	22.358,70	549	32,24
95-4	2051	498	27.996,15	2005	24,84
108-2	2014	638	27.491,10	2149	29,69
122-3	2008	514	27.409,20	1751	29,35
133-1	2023	266	27.613,95	646	41,18
144-3	128	n.a.	n.a.	n.a.	n.a.
155-2	66	n.a.	n.a.	n.a.	n.a.
162-3	186	n.a.	n.a.	n.a.	n.a.
166-2	98	n.a.	n.a.	n.a.	n.a.
	Sum	Sum	Average	Average	Average
	17528	4053	25.859,17	1.387	35,89

Manual annotations were exported as single images. These images were analyzed with SHERPA but the identification rate was poorly low. The reason for that is the exported contour by MetaSystems. The annotation, whatever form it has, is cut out and displayed on a white, squared background. In most cases SHERPA then identifies the contour of the annotation but not the object in it (Figure 9).

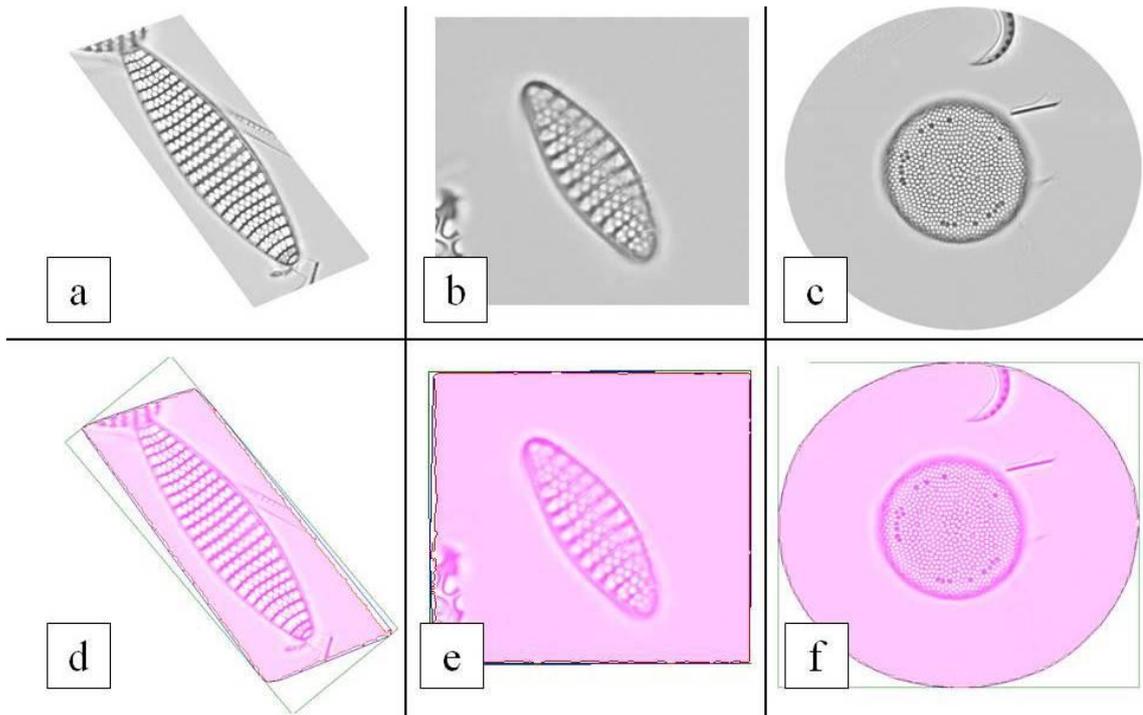


Figure 9 The first row shows exported pictures of areas that were manually marked in Virtual Slides. Picture a was marked as a polygon, b as a square and c as a circle. Picture d - f show the outline that the Software SHERPA created, analyzing the pictures a - c

The difference between the gray of background surrounding the object and the object itself is not that great as the difference of white to the background gray inside the annotation. The enhancement of this problem is far beyond the scope of this thesis.

Another problem within the virtual slides is a shift between the overlaying pictures so that correct outlining in some cases nearly is impossible (Figure 10).

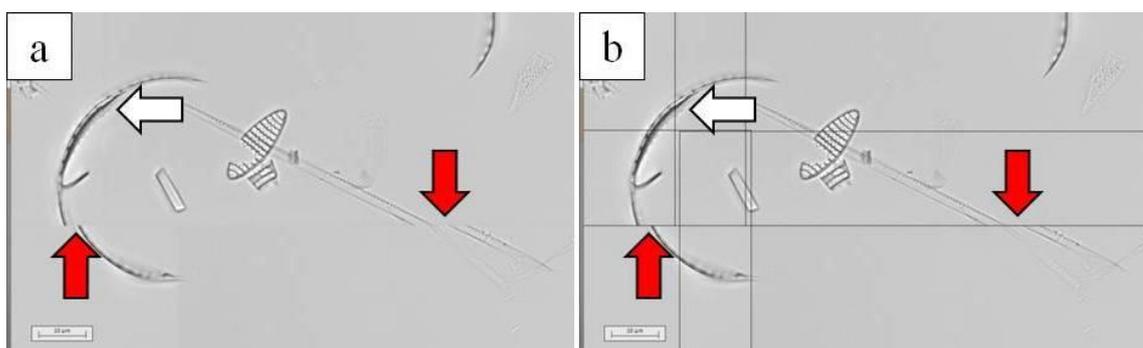


Figure 10 a) Part of a screenshot from a virtual slide showing a shift in overlaying FOV's. Red arrows mark great shifts in left-right direction, white arrow show a slightly shift in top-bottom direction b) Showing the same picture with same Arrows, including the raster of FOV's

The more and bigger gaps occurred from left to right, top to bottom seem to fit pretty good in the observed virtual slides. Valuating these shifts or correcting them is also far beyond the scope of this thesis.

## 5.2 Species composition

### 5.2.1 Automatic species detection (SHERPA Software)

Within the FOV's from the scanned slides the SHERPA Software found 3,180 objects, according to diatom valves. These objects were sorted into species or genus groups depending on visibility of features. Table 3 shows yields in total and per Station, including a small example picture.

Table 3 List of diatom species found by the Software SHERPA in single pictures from Virtual Slides. A summary per Station is given in the bottom line, a summary per species is given in the last left column.

Species	Example	69-3	72-4	78-2	81-4	87-2	95-4	108-2	122-3	133-1	Sum
<i>Asteromphalus</i> <i>sp.</i>		1	0	1	3	0	1	1	1	1	9
<i>Centric</i>		15	79	41	59	62	73	101	36	73	539
<i>Chaetoceros</i> <i>sp.</i>		0	0	2	0	1	1	2	2	1	9
<i>Corethron</i> <i>pennatum</i>		0	0	4	4	2	1	0	11	0	22
<i>Dactyliosolen</i> <i>antarcticus</i>		0	0	2	3	5	108	34	37	8	197
<i>Dictyocha</i> <i>speculum</i>		1	18	7	6	4	32	25	0	28	121
<i>Eucampia</i> <i>antarctica</i>		0	0	0	0	10	9	29	15	19	82
<i>Fragilariopsis</i> <i>keruelensis</i>		103	372	396	644	82	229	410	387	111	2734
<i>Fragilariopsis</i> <i>sp.</i>		0	5	22	22	1	3	10	17	7	87
<i>Navicula</i> <i>sp.</i>		0	0	1	6	1	0	0	0	1	9
<i>Pseudonitzschia</i> <i>sp.</i>		0	2	35	9	3	15	3	5	0	72
<i>Silicoflagellate</i>		0	3	0	0	0	4	0	0	0	7
<i>Thalassionema</i> <i>nitzschoides</i>		0	13	76	5	6	22	23	3	17	165
		120	492	587	761	177	498	638	514	266	4053

### 5.2.2 Manual species detection (Metafer System)

The valves within the manual annotations were identified and summarized, shown in Figure 11.

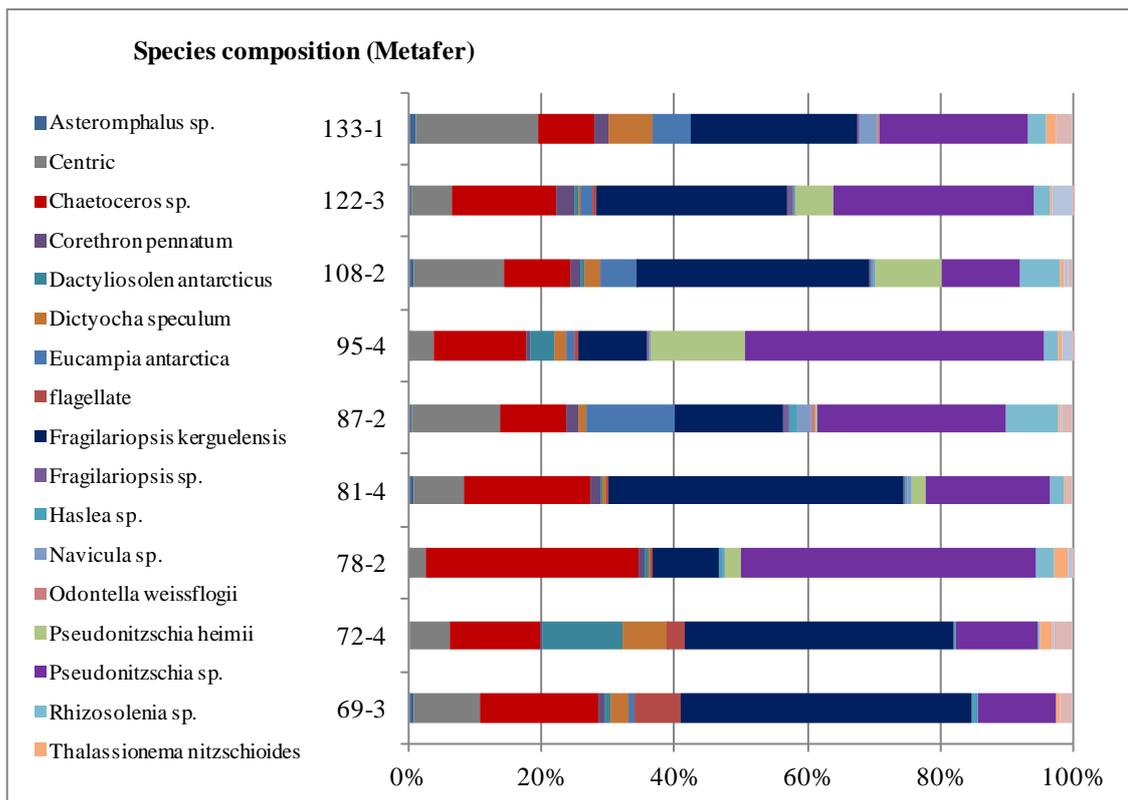


Figure 11 100 % Bars showing diatom valves that were manual marked in Virtual Slides. Cells were identified to species or genus level

### 5.2.3 Species composition in total

All diatom species or genera detected by automation (SHERPA) and manual annotations were summarized and displayed per Station as 100 % bars in Figure 12. The statistic Software Past3 was used analyzing the total composition and further results (Hammer, 2013). Past3 is available as free download from the website of the Natural History Museum, Oslo (<http://www.nhm.uio.no/english/>).

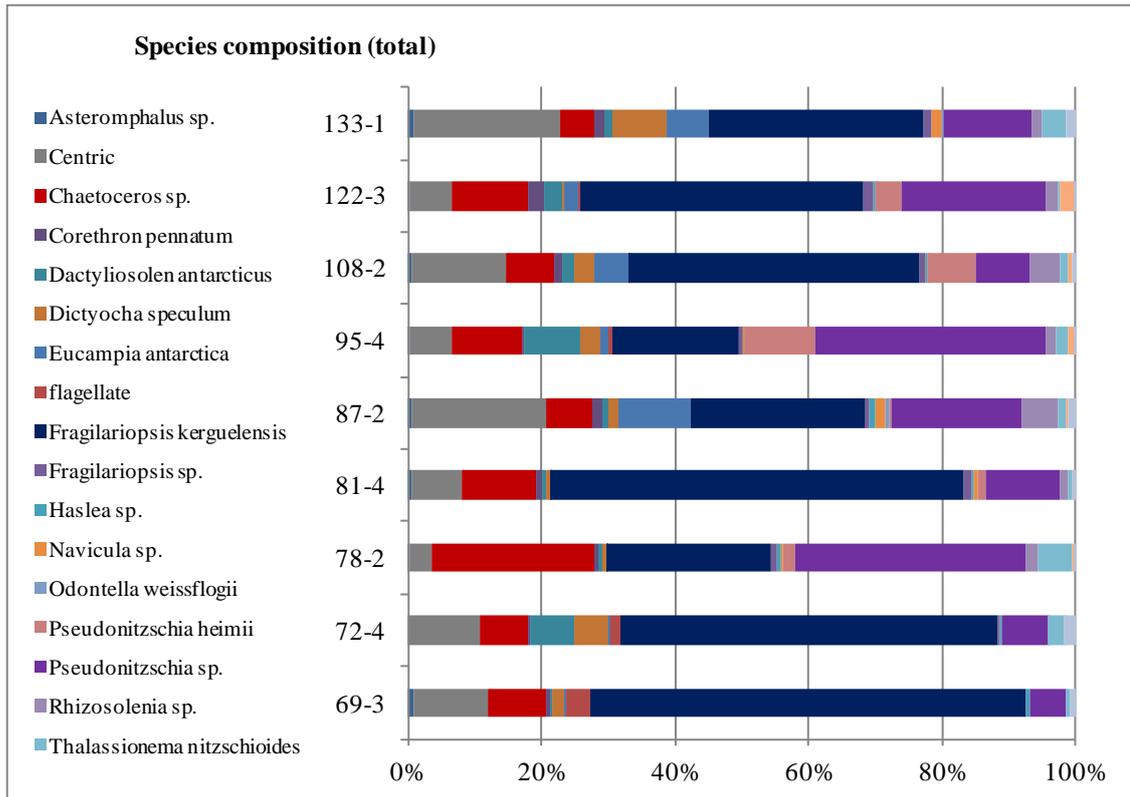


Figure 12 100 % bars showing all diatom valves found in this study. Numbers result from the summary of automated and manual found diatom valves

The detection rate for any species found in this thesis was calculated for the automation method (SHERPA) and the manual annotations (Metafer) shown in Table 4.

Table 4 Detection rates for species found in this study. The yields from automation software SHERPA and the manual annotations (Metafer) are given in percentages

Species	SHERPA (%)	Metafer (%)
<i>Asteromphalus sp.</i>	21,43	78,57
<i>Chaetoceros sp.</i>	0,61	99,39
<i>Corethron pennatum</i>	16,79	83,21
<i>Dactyliosolen antarcticus</i>	55,18	44,82
<i>Dictyocha speculum</i>	44,00	56,00
<i>Ephemera planamembranacea</i>	0	100,00
<i>Eucampia antarctica</i>	29,08	70,92
<i>flagellate</i>	0	100,00
<i>Fragilariopsis kerguelensis</i>	56,57	43,43
<i>Fragilariopsis sp.</i>	73,95	26,05
<i>Haslea sp.</i>	0	100,00
<i>Navicula sp.</i>	17,31	82,69
<i>Odontella weissflogii</i>	0	100,00
<i>Pseudonitzschia heimii</i>	0	100,00
<i>Pseudonitzschia sp.</i>	2,85	97,15
<i>Rhizosolenia sp.</i>	0	100,00
<i>Thalassionema nitzschioides</i>	68,46	31,54
<i>Thalassiothrix antarctica</i>	0	100,00
<i>Centric</i>	46,75	53,25

The number of found specimens differs between the Stations. Figure 13 shows Numbers of Taxa plotted against the number of specimens found in each station. In this case specimens are valves that were found by SHERPA or were manual mark within virtual slides.

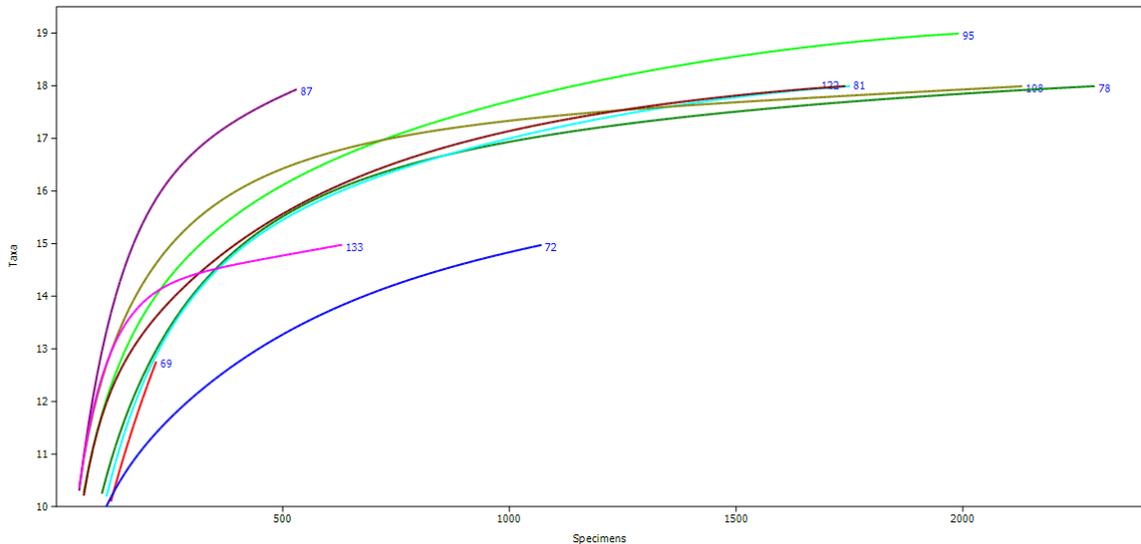


Figure 13 The number of identified Taxa are plotted against the number of specimens found in each observed Station.

To analyze the similarity between stations two hierarchical clustering were performed. Both cluster show differences between the stations that are displayed by a “dendogram”. The length of connection lines represents the dissimilarity. Figure 14 A shows a cluster using an algorithm of unweighted pair-group average where clusters are joined based on the average distance between all members in the two groups. The Bray-Curtis Index was determined.

Figure 14 B shows the same method but a reduced data set where only those taxa were included which occurred with a relative abundance  $> 1\%$  in all samples.

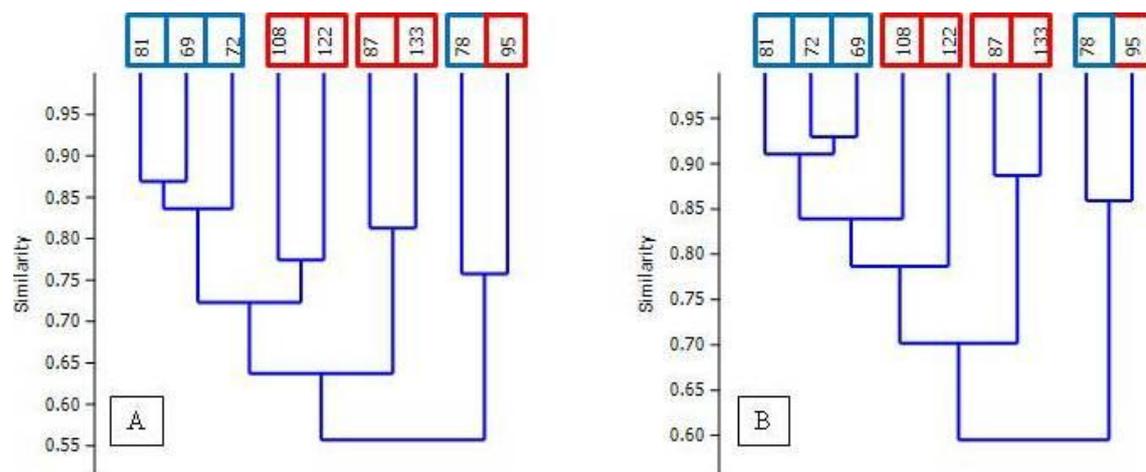


Figure 14 Similarity comparison of identified Taxa between the Stations. A) Show the “Bray-Curtis” method. B) The same method displaying a reduced data set, where species occurring less than 1 % are rejected. Stations are given in the top. Blue frames define the Station Group “Transect”, Red frames the Station Group “Eddy”

### 5.3 Size distribution

Size distribution was analyzed for the dominant species *Fragilariopsis kerguelensis* found by the SHERPA Software. The measurements include the length of apical- and transapical axis (also called length and width), the area of the object and the average distance between costae

Table 5 List of values for apical- and transapical axes described for different *Fragilariopsis* species. The number of striae and fibulae in 10  $\mu\text{m}$  is given in the last column (after ...)

Fragilariopsis species	Apical axis ( $\mu\text{m}$ )	Transapical axis ( $\mu\text{m}$ )	Striae and fibulae in 10 $\mu\text{m}$
<i>F. kerguelensis</i>	10 - 76	5 - 11	4 - 7
<i>F. ritscheri</i>	22 - 57	8 - 9	6 - 11
<i>F. atlantica</i>	20 - 43	7 - 8	18 - 21
<i>F. oceanica</i>	10 - 41	ca. 6	12 - 15
<i>F. pseudonana</i>	4 - 20	3.5 - 5	18 - 22
<i>F. rhombica</i>	8 - 53	7 - 13	8 - 16
<i>F. separanda</i>	10 - 33	8 - 13	10 - 14

The number of costae per 10  $\mu\text{m}$  is a regular identification feature and the distance of five costae could be used to calculate  $F^*$  after Fenner et al, 1976.  $F^*$  is a value where the distance of 5 striae is related to the length and width (Formula 1).

$$F^* = \frac{\text{length}}{\text{width}} * \text{width of 5 costae (in } \mu\text{m)}$$

Formula 1: Definition of the value  $F^*$ . A value for costae distance related to length and width, after Fenner et al., 1976

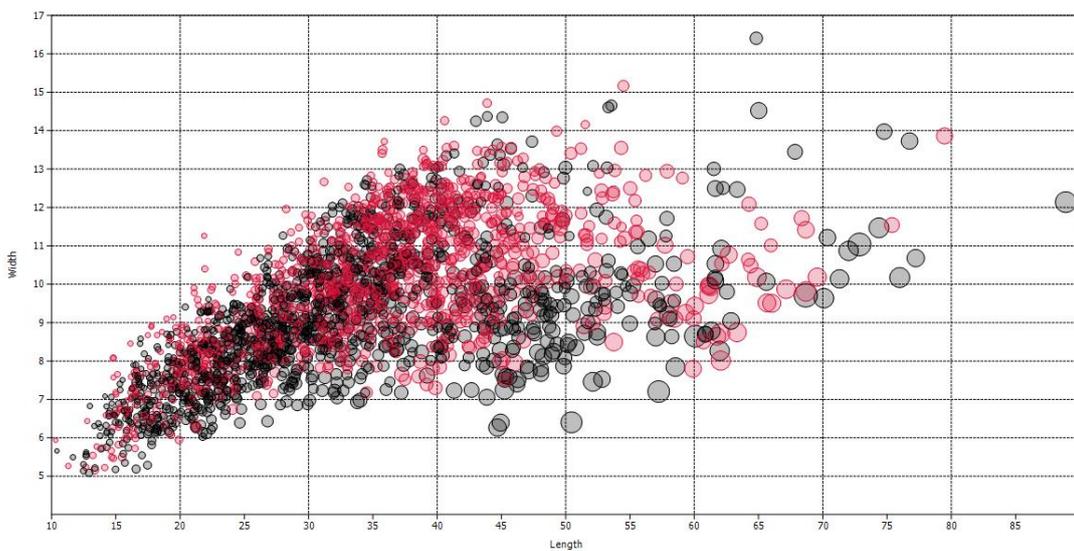


Figure 15 Scatter plot of measurements for the species *Fragilariopsis kerguelensis* obtained in all Stations. The width is plotted against the length. The circle size represents the value of  $F^*$ . Black circles show the Station Group “Transect”, red circles show the Station Group “Eddy”

The extreme values measured from *Fragilariopsis kerguelensis* over all Stations are listed in Table 6. The pictures were manually re-measured within a virtual slide. The shown numbers slightly differ from the original measurement of SHERPA but still fit in the values pretty good. The SHERPA export data does not include a scale shown at the object, therefore a screenshot from the virtual slide with manual measurement was preferred for visualization (e.g. Figure 17).

Table 6 Extreme values measured from *Fragilariopsis kerguelensis* in different Stations. Minima and maxima values, calculated by the Software SHERPA, are given for the length, width, the average distance of costae and the area. Manual re-measurements were obtained and listed, except for the area. The bottom line notes letters for pictures of the analogue cells, shown in Figure 16,

	Length		Width		Average Costae distance		Area	
	min.	max.	min.	max.	min.	max.	min.	max.
SHERPA Value	10,31	88,90	5,09	16,41	0,94	2,52	44,35	868,99
Manual measure	10,3	88,9	5,05	16,42	0,92	2,43	not measurable	
Station	133-1	81-4	78-2	81-4	81-4	81-4	72-4	122-3
according Picture	A	B	C	D	E	F	G	H

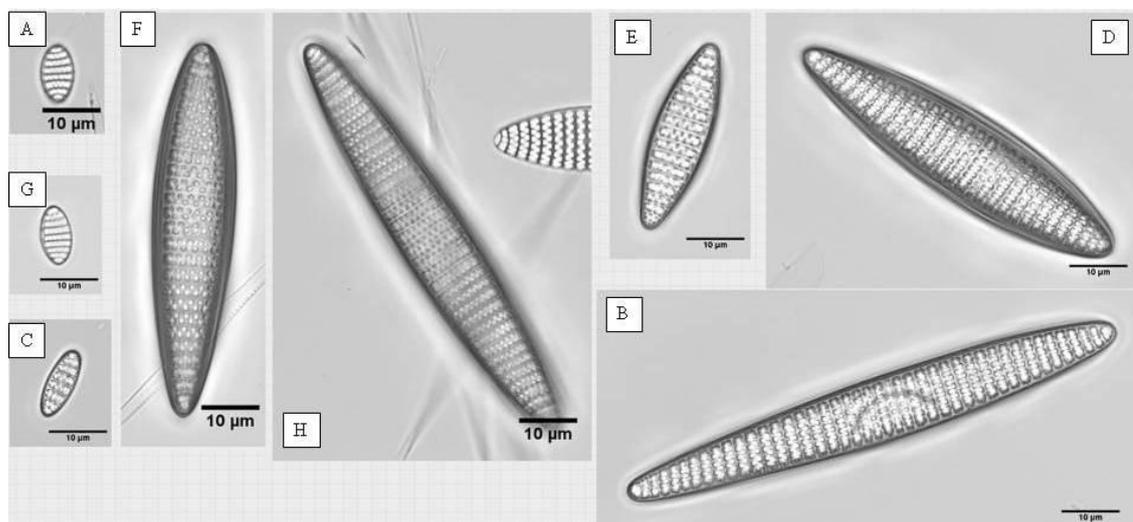


Figure 16 Pictures of *Fragilariopsis kerguelensis* having the extremists morphometric values found in all observed stations. According values are shown in Table 6. Samples origin from Polar-stern Cruise ANTXXVIII-3, taken by hand net from surface waters. Scale bar in each picture 10 µm

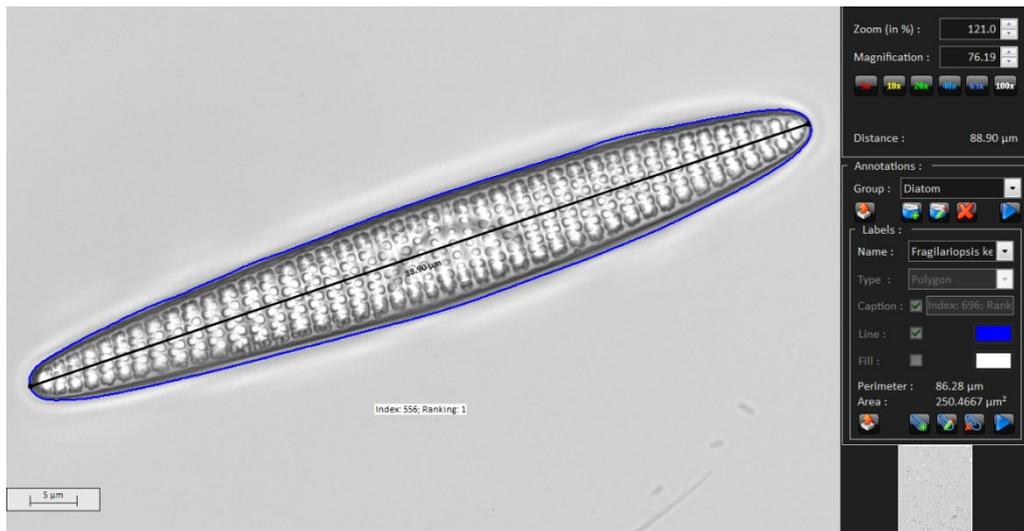


Figure 17 Example picture of a manual re-measurement in a virtual slide. The *Fragilariopsis kerguelensis* valve containing the largest length is shown, reaching 88.9  $\mu\text{m}$ . The blue outline is the imported outline, generated by the Software SHERPA, the black line over the apical axis shows the manual re-measurement

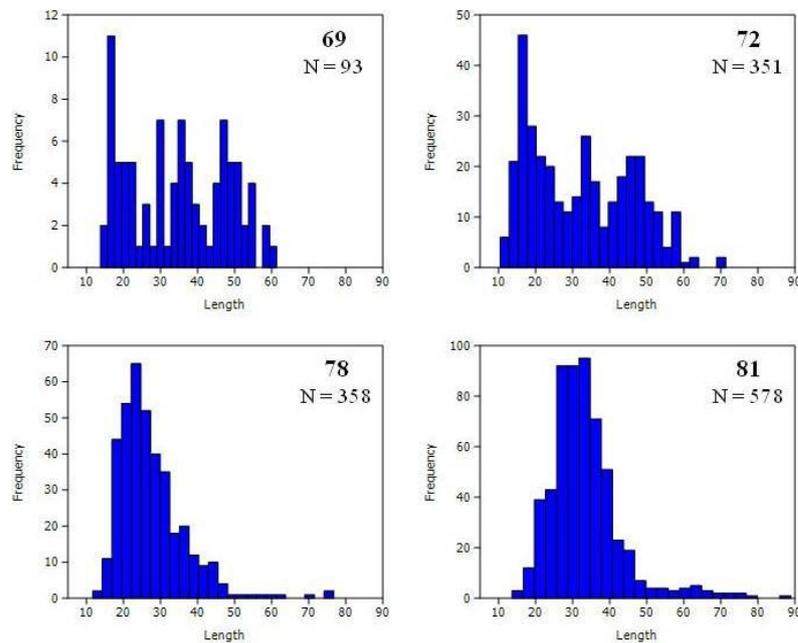


Figure 18 Four histograms from the Station Group “Transect” show the distribution of apical length measured for the species *Fragilariopsis kerguelensis*. The Station Number and found valves (n) for each are given in the top left. Equal scales on the length, varying scales on the frequency

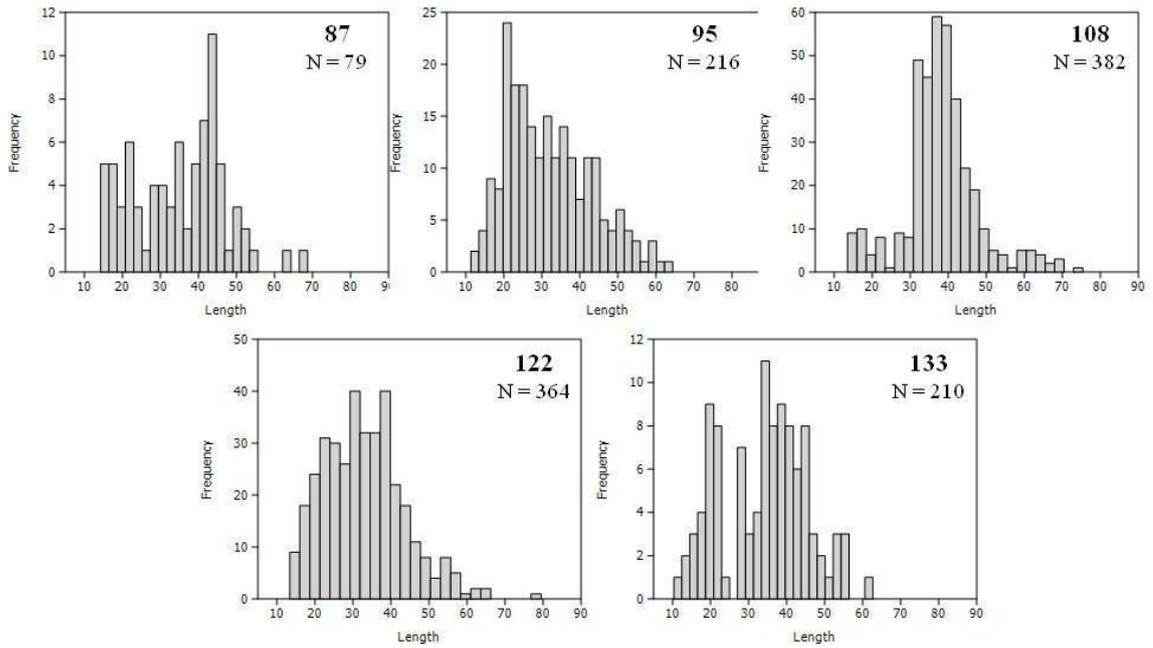


Figure 19 Five histograms from the Station Group “Eddy” show the distribution of apical length measured for the species *Fragilariopsis kerguelensis*. The Station Number and found valves (n) for each are given in the top left. Equal scales on the length, varying scales on the frequency

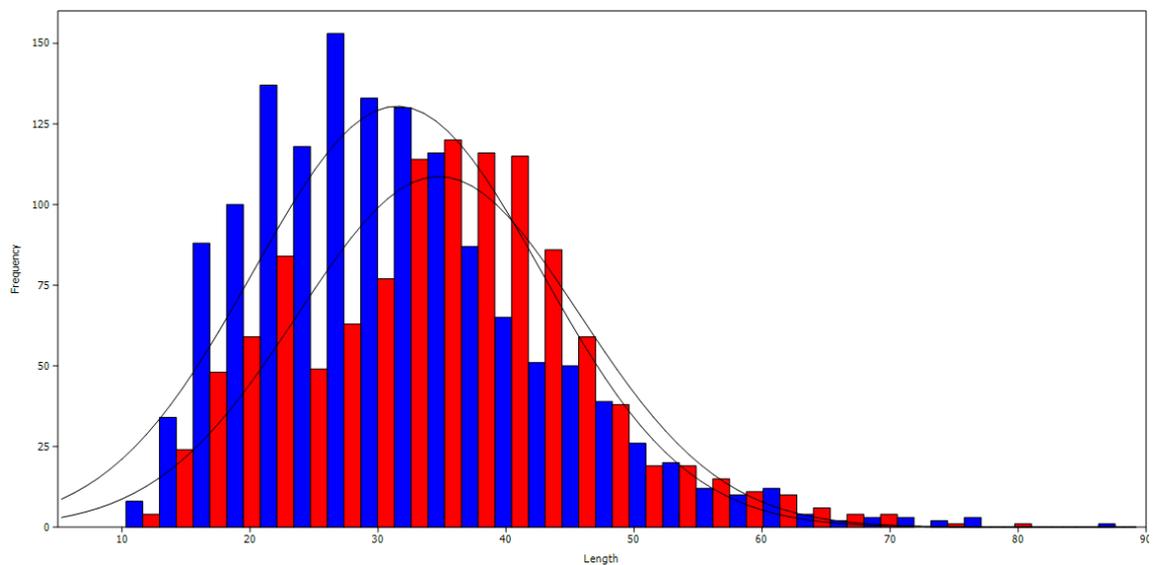


Figure 20 Histogram showing the distribution of length for the species *Fragilariopsis kerguelensis* found in all investigated stations. Blue bars represent the Station Group Transect, red bars the Station Group Eddy. Black graphs are fitted normal distributions (Parametric estimation, not least Squares)

Average values for length and  $F^*$ , calculated for the diatom species *Fragilariopsis kerguelensis* detected by the SHERPA Software are shown in Figure 21.

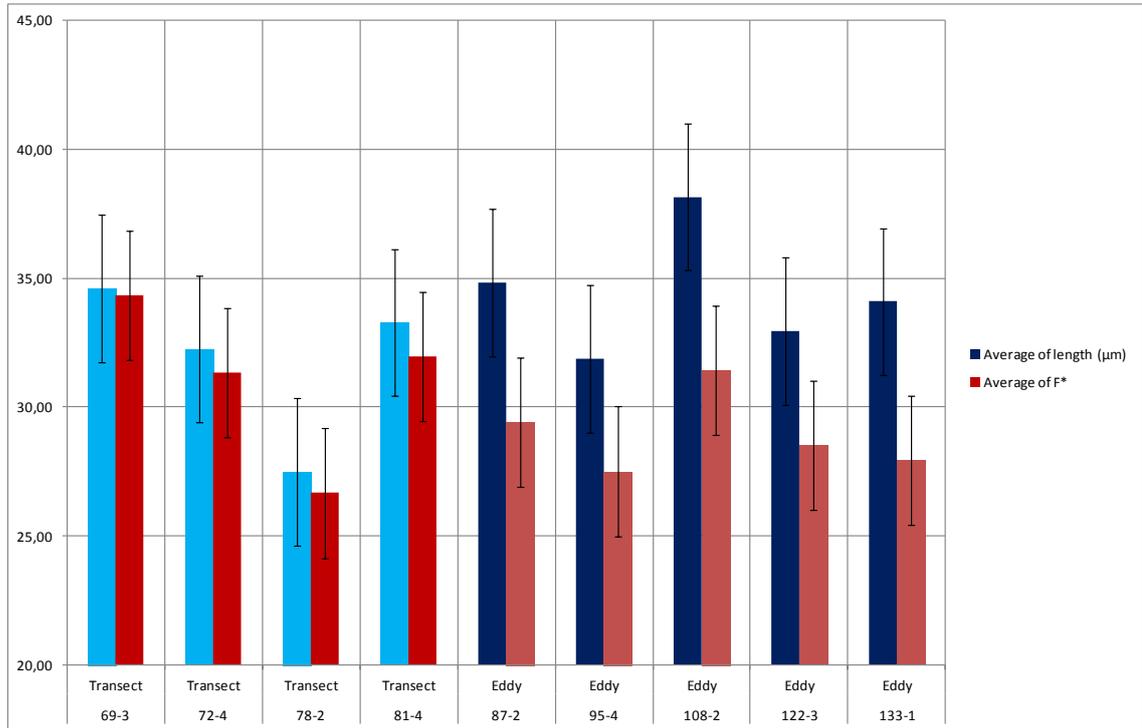


Figure 21 Averages of length and  $F^*$  measured for the diatom species *Fragilariopsis kerguelensis*. Values are shown for nine stations from Polarstern cruise ANTXXVIII-3. Samples origin from hand nets taken from surface waters. Light blue bars show the average length of Transect stations with dark red bars showing according  $F^*$  values, dark blue bars show the average length of the Eddy stations with light red bars showing the average  $F^*$  value

An analysis of co-variance was performed between the station groups containing the variables length and  $F^*$ . The resulting  $p$  – value was significant low at  $1.59E^{-52}$ .

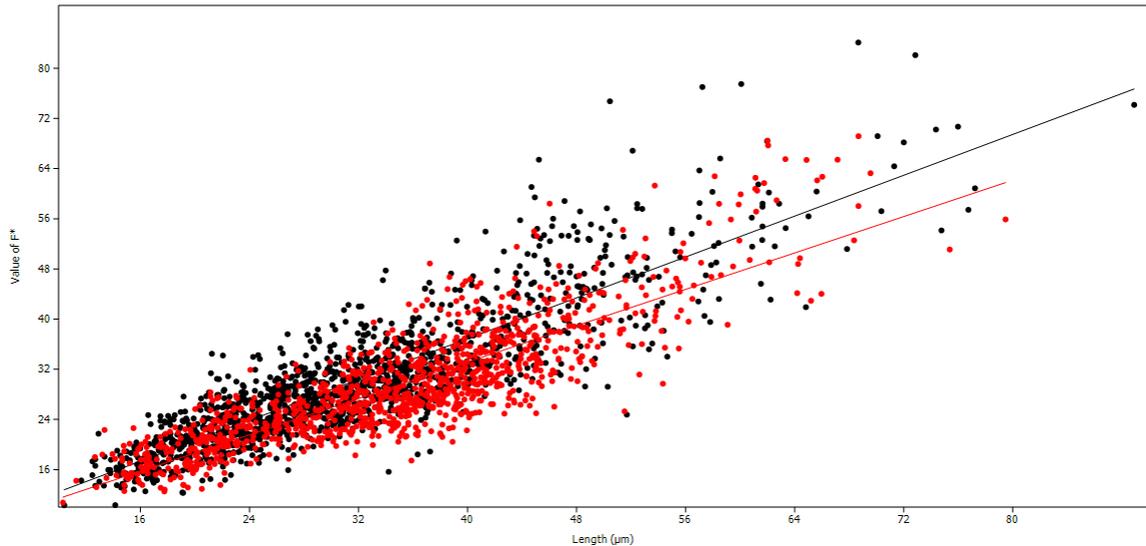


Figure 22 Scatter plot from the ANCOVA Analysis between the Station Groups “Transect” and “Eddy”. The tested values were length and  $F^*$  (see 2.) measured for the species *Fragilariopsis kerguelensis*. The line of best fit is included. Black line and dots show the Station Group “Transect”, red line and dots show the Station Group “Eddy”

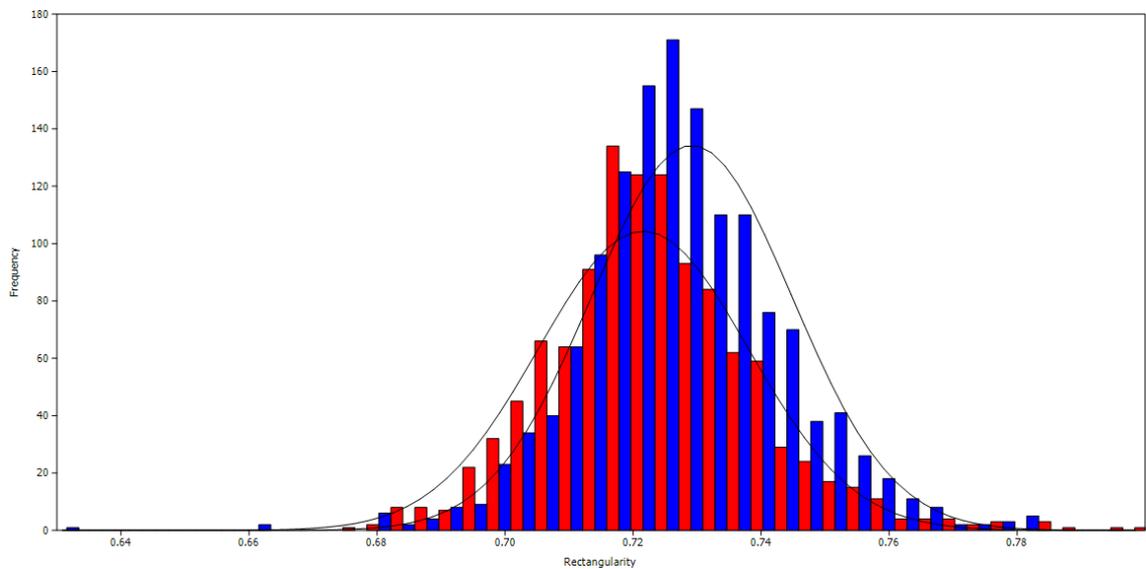


Figure 23 Histogram of a morphometric value “Rectangularity” measured for the species *Fragilariopsis kerguelensis* in all Stations. Blue bars show the Station Group “Transect”, red bars show the Station Group “Eddy”. Black graphs are fitted normal distributions (Parametric estimation, not least Squares)

Another analysis of co-variance was performed between the station groups containing the variables length and rectangularity. The resulting  $p$  – value also was significant low at  $4.72E^{-27}$ .

## 6 Discussion

Question 1: What kind of detection rate can be obtained by automation and what kinds of biases occur? Are all diatom taxa detected at the same rate and does the according taxonomic composition reflect the real composition of the communities?

The detection rate of the SHERPA software is low compared to the number of valves that were found in total. Valves in total are the summary of SHERPA detected objects and manual marked objects.

The highest percentage of SHERPA detection rate could be obtained in Station 69 reaching 51.72 % but accomplished by the lowest number of valves ( $n = 232$ ). The lowest percentage of automatic detection is reached at Station 95 with 24.84 % having the second highest number of valves ( $n = 2,005$ ). The tendency of decreasing SHERPA detections with increasing total valve numbers is not confirmed by the determined numbers (see Table 2). There are different reasons for the circumstance SHERPA does not recognize more or all valves.

First, and most affecting in this study, is cell size. Cells larger than a single FOV (ca.  $130 \times 105 \mu\text{m}$ ) cannot be detected using the methods applied here. This strongly affected counts for the genera *Chaetoceros* and *Pseudonitzschia* as they appear in low numbers detected by automation (9 : 72) (see Table 3) in spite of being abundant in the samples (see red and purple bars in Figure 11). The genus *Chaetoceros* contains long setae which mostly reach out of range of a single FOV (e.g. *Chaetoceros debilis*). The genus *Pseudonitzschia* contains species which are described by length around or above  $100 \mu\text{m}$ , like *P. antarctica*, *P. heimii*, *P. lineola* etc. (see **Fehler! Verweisquelle konnte nicht gefunden werden.**).

This detection bias led to substantially different species compositions obtained by SHERPA (see 4.2.2) when compared to the manual annotations (see 4.2.3). The combined data set including both automatically and manually found valves was analyzed to evaluate community differences between the stations or station groups.

Question 2: Does the diatom composition differ between the stations or the station groups and how do relative abundances of single Taxa change?

The species composition in total (see 4.2.3) is dominated by *Fragilariopsis kerguelensis* except for two Stations, 78 and 95. The maximum percentage per station is reached in Station 69 with 65.51 %.

The second dominant genus is *Pseudonitzschia*. In two Stations, where *Fragilariopsis kerguelensis* is not the most abundant species *Pseudonitzschia sp.* is most abundant, reaching 34 % in both Stations, 78 and 95. The first impression was that relative abundances of *Pseudonitzschia sp.* increase over the Transect but the peak in Station 78 (34 %) could not closely be reached in the following Station 81 (11 %). Within the Eddy relative abundances of *Pseudonitzschia sp.* are alternating, e.g. Station 108 (8 %) and Station 122 (21 %).

The third abundant genus is *Chaetoceros*. Their relative abundances vary around 10 % (5 – 11) with one exception in Station 78 where they reached 24 %.

Centric diatoms with a circular valve outline appeared in all stations but at different relative abundances, I did not observe a clear tendency among stations or between the transect and the eddy. Within the Transect percentages varied from 3 – 11 % and in the Eddy the minimum was 6 % and maximum 22 %.

Differences in relative abundances between the genus *Chaetoceros* and *Fragilariopsis kerguelensis* were pronounced by Sachs et al (2009) and Figure 1 shows the according distribution. Comparing these distributions with the relative abundances found in this study, similar dominances for *Fragilariopsis kerguelensis* are noted. The genus *Chaetoceros* could not compete, more likely the genus of *Pseudonitzschia* was closely equal abundant (Station 72, 81 and 108) or even more abundant than the *Chaetoceros sp.* (Station 78, 95, 122, 133) (see Figure 12). A reason for this difference to the Sachs et al (2009) distribution might be the origin of samples. Sachs is referring to sediment samples from drilling cores, where *Pseudonitzschia sp.* is not described and *Chaetoceros sp.* are not counted by valves but resting spores. The weak silicified valves of both genera might get broken on the way sinking down or are consumed more easily by predators than other genera. Samples observed in this study origin from hand net samples from surface waters so differences are reasonable.

The similarity clusters in Figure 14 show the same relations between the stations even after rejecting species occurring less than 1 %. The Station groups seem to have similar species compositions except for Station 78 from the Transect which is more equal to

Station 95 from the Eddy group. This similarity relates on the relative high abundances of *Pseudonitzschia sp.* in these stations.

Question 3: Does morphometry in dominant species change between the stations or station groups?

The morphometry for the dominant species *Fragilariopsis kerguelensis* was analyzed in size distribution, the F\* value and the rectangularity as a value for shape of outlines.

The size distribution for length and width over all stations is displayed in Figure 15. As the Station Group Transect (Black marks) includes more extreme values than the Eddy-Group (Red marks) the length was displayed for each Station, plotted together in Station Groups in Figure 18 and Figure 19. Comparing these plots no clear tendency is noticeable between the stations. Within the Transect, Station 69 and 72 are more widespread in distribution with multiple peaks below the maximum (most valves ca. 15µm) and some size ranges are missing completely. The other Transect Stations, 78 and 81, show an unimodal distribution and missing size classes are of high or extreme values.

The Station Groups were compared as a histogram in Figure 20 which displays a distribution of highest frequency between 20 and 35 µm for the Transect Group and high frequented sizes for the Eddy Group between 35 and 45 µm.

Within the extreme values found for *Fragilariopsis kerguelensis* some measurements exceed the values of identifying literature. Comparison of different identifying books showed that the values for *Fragilariopsis kerguelensis* were all equal because all referred to Grethe Hasle (19xx). Extreme values are given in Table 6 and the according pictures are given in Figure 16, clearly showing *Fragilariopsis kerguelensis*. The valve with highest length of 88.90 µm is shown in Figure 17. This measurement exceeded the literature value of Hasle (76 µm maximum) heavily. In general many valves showed a width that exceeded the maximum of 11 µm (Hasle), nearly the upper half of the plot in Figure 15. The double check of every single picture, taken of this species, assured that identification was right. The density of costae is described by 4 to 7 in 10 µm which could be recalculated to a minimum of average distance of 1.42 µm and a maximum of 2.5 µm. The minimum extreme value found in this study is 0.94 µm which is a heavy deviation.

This circumstance of exceeding and deviating to literature values, particular heavily, recommend a rework of identification literature.

In Figure 15 the circle size represents  $F^*$  and slightly differences were seen that led to further analysis. To analyze the relations between length,  $F^*$  and the Station they origin from an Analysis of Co-variance was performed and a significant difference between the Station Groups was shown (4.3.1).

The same analysis was performed for the rectangularity which also showed a significant difference (4.3.1).

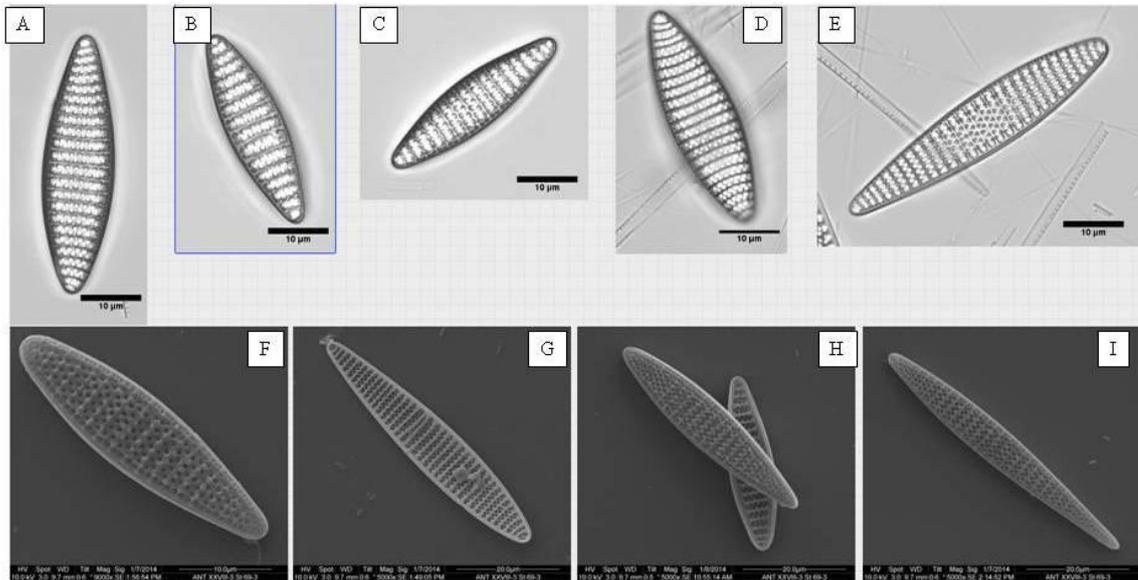
To distinguish differences in the species *Fragilariopsis kerguelensis*, concerning the striae, the average values of length and  $F^*$  were displayed per Station in Figure 21. Inside the Transect the average values of  $F^*$  are lying slightly below the average length and having the same tendency as length development. Within the Eddy this tendency is not noticed as the  $F^*$  values are clearly lower than in the Transect and even having an opposite tendency in Station 122 and 133.

This might be an evidence for different living conditions to diatoms within an Eddy or particular for *Fragilariopsis kerguelensis*. Various reasons could come into play to explain this circumstance. Different nutrient availability or different predation pressure inside an Eddy might be reasons why these Organisms change their way of building protection.

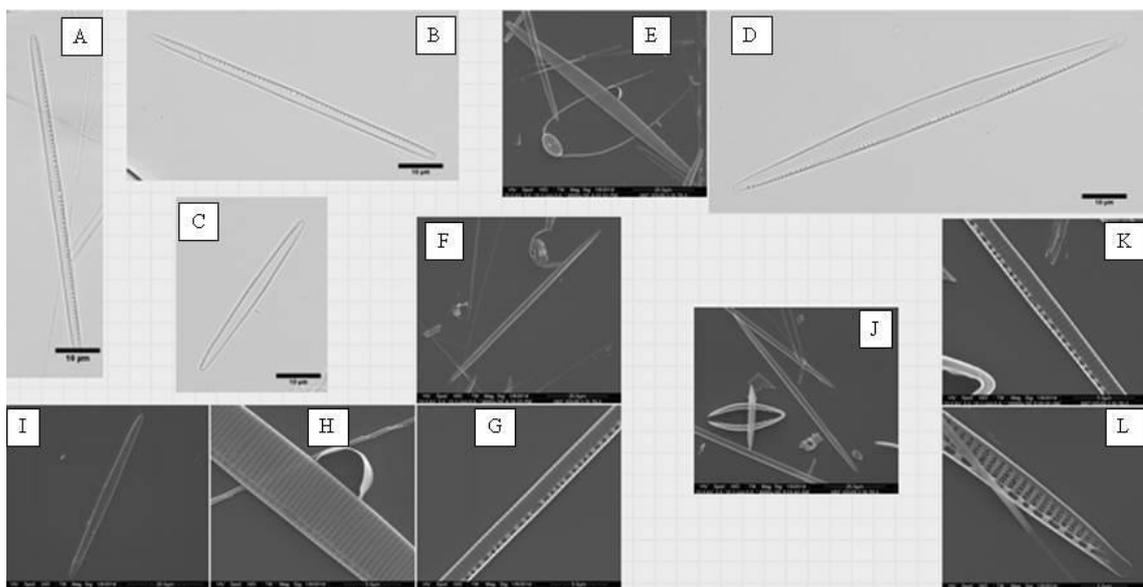
The fact that the species *Fragilariopsis kerguelensis* was reaching out of the literature values in length, width and density of costae is one of the most important facts in this study. The most important information is the difference in costae densities in the Transect and the Eddy. The morphometry changes and not even in length. Varying length distributions of diatoms could always be argued with life-cycle situation the community could be in. The found differences are clearly and are supported by the analysis of rectangularity which also shows differences.

Within the new applied techniques such morphometry patterns could be observed and it would be of greatest interest to go further on, maybe to determine new Literature Values for some species, at least for *Fragilariopsis kerguelensis*. This species seem to have various growing strategies that need to be found.

## 7 Example plates of dominant Species



Example plate for the diatom species *Fragilariopsis kerguelensis*. A – E Light microscope pictures with each scale bar 10 µm. F – I Scanning electron microscope pictures. Scale bars F = 10 µm; G – I 20 µm



Example plate for the diatom genus *Pseudonitzschia* sp. A – D Light microscope pictures with each scale bar 10 µm. E – I Scanning electron microscope pictures. K and L show details of J. Scale bars E, F, I and J 20 µm; G, H and K 5 µm; L = 2 µm

## 8 Index of Tables

Table 1	A list of investigated stations in this study with coordinates. Samples were collected during Polarstern ruise ANTXXVIII-3 and permanent slides containing diatoms were prepared from hand-net samples, taken from surface waters. Stations are clustered in Groups, listed in the last column.....	12
Table 2	List of observed Stations during Polarstern Cruise ANTXXVIII-3 and the number of taken pictures (FOV = field of view) by automated microscopy (Metafer Systems) in permanent slides, prepared from hand-net samples taken from surface waters. The number of detected objects by the Software SHERPA and the counts of manual observed slides are listed. The scanned area was calculated and the total yield of valves is displayed. The percentage of SHERPA detected objects is given in the last column.....	20
Table 3	List of diatom species found by the Software SHERPA in single pictures from Virtual Slides. A summary per Station is given in the bottom line, a summary per species is given in the last left column. ....	22
Table 4	Detection rates for species found in this study. The yields from automation Software SHERPA and the manual annotations (Metafer) are given in percentages .....	24
Table 5	List of values for apical- and transapical axes described for different <i>Fragilariopsis</i> species. The number of striae and fibulae in 10 $\mu\text{m}$ is given in the last column (after ...)	26
Table 6	Extreme values measured from <i>Fragilariopsis kerguelensis</i> in different Stations. Minima and maxima values, calculated by the Software SHERPA, are given for the length, width, the average distance of costae and the area. Manual re-measurements were obtained and listed, except for the area. The bottom line notes letters for pictures of the analogue cells, shown in Figure 17,....	28

## 9 Index of Figures

Figure 1	The orange hatched area shows enhanced <i>Chaetoceros</i> spp. spores (>20%) in surface sediments. Red hatched area shows elevated <i>Fragilariopsis kerguelensis</i> frustules (>40%). Colored dots and black lines represent various data from the origin study e.g. the spatial distribution of the oxygen penetration depth in sediments (OPD) (Figure from Sachs et al 2009).....9	9
Figure 2	Track plot of Polarstern cruise ANTXXVIII-3 from Cape Town to Punta Arenas is shown as a red line. The Subtropical Front (STF), the Sub Antarctic Front (SAF), the Polar Front (PF) and the southern boundary of the Antarctic Circumpolar Current are shown. Colored boxes mark the investigated stations in this study. Green boxes are marking the station group “Transect”, white boxes show the group “Eddy 1” and grey boxes mark the “Eddy 2” group. Diatoms were prepared from hand-net samples, taken from surface waters ..... 13	13
Figure 3	a) A screenshot of a virtual slide, displayed by the Software from Metafer Systems. The virtual slide consists of 1,638 single pictures called field of view (FOV). Figure a shows the whole slide. b) Figure b displays the same slide but with the raster of FOV’s ..... 15	15
Figure 4	Basic Workflow of the computer software SHERPA (figure from Kloster et al. submitted) ..... 16	16
Figure 5	Two templates used by the Software SHERPA. Upper half shows a template and the related species <i>Fragilariopsis separanda</i> , the lower half shows the template for <i>Fragilariopsis kerguelensis</i> , also with an example picture, in comparison. The software SHERPA creates an outline for an object which is then compared to a set of templates, resulting in various values matching or mismatching the best fit template ..... 17	17
Figure 6	Screenshot from the Software SHERPA showing various values from analyzing. Upper half shows the provided settings for parameter’s, lower half shows the fitting results from comparing a set of diatom pictures to a related set of templates ..... 17	17
Figure 7	Screenshot of a virtual slide, displayed by the Software from Metafer Systems. Blue outlined objects represent annotations that were re-imported from the Software SHERPA. The black lined raster shows the overlaying FOV’s ..... 18	18
Figure 8	Screenshot of a virtual slide, displayed by the Software from Metafer Systems. Blue outlined objects represent annotations that were re-imported from the Software SHERPA. Other colored marks are manual annotations in various forms. Green labels show the group “not found”, yellow labels are the group of “bad optic/angle” and the magenta colored labels are “>2pictures” (explanation see Text). The black lined raster shows the overlaying FOV’s ..... 19	19
Figure 9	The first row shows exported pictures of areas that were manually marked in Virtual Slides. Picture a was marked as a polygon, b as a square and c as a circle. Picture d - f show the outline that the Software SHERPA created, analyzing the pictures a - c ..... 21	21
Figure 10	a) Part of a screenshot from a virtual slide showing a shift in overlaying FOV’s. Red arrows mark great shifts in left-right direction, white arrow show a slightly shift in top-bottom direction b) Showing the same picture with same Arrows, including the raster of FOV’s ..... 21	21
Figure 11	100 % Bars showing diatom valves that were manual marked in Virtual Slides. Cells were identified to species or genus level ..... 23	23
Figure 12	100 % bars showing all diatom valves found in this study. Numbers result from the summary of automated and manual found diatom valves..... 24	24

Figure 13	The number of identified Taxa are plotted against the number of specimens found in each observed Station.....	25
Figure 14	Similarity comparison of identified Taxa between the Stations. A) Show the “Bray-Curtis” method. B) Similarity comparison with “Ward’s” method. Stations are given in the top. Blue frames define the Station Group “Transect”, Red frames the Station Group “Eddy”.....	26
Figure 15	Scatter plot of measurements for the species <i>Fragilariopsis kerguelensis</i> obtained in all Stations. The width is plotted against the length. The circle size represents the value of F*. Black circles show the Station Group “Transect”, red circles show the Station Group “Eddy”.....	27
Figure 16	Pictures of <i>Fragilariopsis kerguelensis</i> having the extremists morphometric values found in all observed stations. According values are shown in Table 5. Samples origin from Polarstern Cruise ANTXXVIII-3, taken by hand net from surface waters. Scale bar in each picture 10 $\mu\text{m}$ .....	28
Figure 17	Example picture of a manual re-measurement in a virtual slide. The <i>Fragilariopsis kerguelensis</i> valve containing the largest length is shown, reaching 88.9 $\mu\text{m}$ . The blue outline is the imported outline, generated by the Software SHERPA, the black line over the apical axis shows the manual re-measurement.....	29
Figure 18	Four histograms from the Station Group “Transect” show the distribution of apical length measured for the species <i>Fragilariopsis kerguelensis</i> . The Station Number and found valves (n) for each are given in the top left. Equal scales on the length, varying scales on the frequency.....	29
Figure 19	Five histograms from the Station Group “Eddy” show the distribution of apical length measured for the species <i>Fragilariopsis kerguelensis</i> . The Station Number and found valves (n) for each are given in the top left. Equal scales on the length, varying scales on the frequency.....	30
Figure 20	Histogram showing the distribution of length for the species <i>Fragilariopsis kerguelensis</i> found in all investigated stations. Blue bars represent the Station Group Transect, red bars the Station Group Eddy. Black lines.....	30
Figure 21	Averages of length and F* measured for the diatom species <i>Fragilariopsis kerguelensis</i> . Values are shown for nine stations from Polarstern cruise ANTXXVIII-3. Samples origin from hand nets taken from surface waters. Light blue bars show the average length of Tansect stations with dark red bars showing according F* values, dark blue bars show the average length of the Eddy stations with light red bars showing the average F* value.....	31
Figure 22	Scatter plot from the ANCOVA Analysis between the Station Groups “Transect” and “Eddy”. The tested values were length and F* (see 2.) measured for the species <i>Fragilariopsis kerguelensis</i> . The line of best fit is included. Black line and dots show the Station Group “Transect”, red line and dots show the Station Group “Eddy”.....	32
Figure 23	Histogram of a morphometric value “Rectangularity” measured for the species <i>Fragilariopsis kerguelensis</i> in all Stations. Blue bars show the Station Group “Transect”, red bars show the Station Group “Eddy”. Black line shows ... for both?.....	32

## 10 Literature

Barber HG, Haworth EY (1981): A guide to the morphology of the diatom frustule, vol. 44. Ambleside, Cumbria, UK: Freshwater Biological Association.

Claude J (2008): Morphometrics with R, New York, Springer Science+Business Media, LLC.

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

Crawford RM (1980): Some considerations of size-reduction in diatom cell walls, 6<sup>th</sup> Diatom symposium

du Buf H, Bayer MM (2002): Automatic Diatom Identification, Singapore, World Scientific Publishing Co. Pte. Ltd.

Fahrbach E (2011): Expeditionsprogramm Nr. 88, FS Polarstern, Stiftung Alfred-Wegener-Institut für Polar- und Meeresforschung, Mitglied der Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. (HGF)

Falasco E, Badino G (2010): The role of environmental factors in shaping diatom frustule, Chapter in Compton JC (2010): *Diatoms – Ecology and Life cycle*, Nova Science Publishers, Inc.

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. *Initial Res. Deep Sea Drilling Project 35: 757-813*, doi:10.2973/dsdp.proc.35.app3.1976

Garrison T (2010): *Oceanography: An Invitation to Marine Science*, 7<sup>th</sup> edition, Brooks/Cole, Cengage Learning, Belmont, USA

Klaas C, Altvater F, Kottmeier D, Kottmeier R, Rueger T, Schourup-Kristensen V (2012): Cruise Report Polarstern Cruise ANT XXVIII-3: Plankton Assemblage Composition, Chlorophyll A, Biogenic Silica, Particulate and dissolved Carbon and Nitrogen determination

Kloster M, Kauer G, Beszteri B (submitted 2013): SHERPA - An image segmentation and outline feature extraction tool for diatoms and other objects. Submitted to BMC Bioinformatics.

Medlin LK, Priddle J (1990): Polar Marine Diatoms, British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK

Miller CB (2004): Biological Oceanography, Blackwell Publishing, Malden USA, Blackwell Science Ltd.

Paululat A, Purschke G (2011): Wörterbuch der Zoologie, Spektrum Akademischer Verlag, Heidelberg

Scott FJ, Marchant HJ (2005): Antarctic Marine Protists, Australian Biological Resources Study, Canberra

Smetacek VS (1985): Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance, International Journal on Life in Oceans and Coastal Waters, Springer Verlag

Steward RH (2008): Introduction to physical Oceanography, Department of Oceanography, Texas A & M University, Online print?, September 2008 Edition

Summerhayes CP, Thorpe SA (2002): Oceanography – An illustrated guide, Manson Publishing, London

---

Tomczak M, Godfrey JS (2001): Regional Oceanography: An Introduction, The flinders University of South Australia, CSIRO Division of Oceanography, Tasmania, Australia, Online print 2001

Internet Links:

<http://mutage.oxfordjournals.org/content/26/1/169.short>

<http://www.sciencedirect.com/science/article/pii/S0165022X02000982>

[http://link.springer.com/protocol/10.1007%2F978-1-61779-055-3\\_6](http://link.springer.com/protocol/10.1007%2F978-1-61779-055-3_6)

## 11 Declaration

Hereby I declare this present Thesis as worked out by my own. In this Thesis I only used the specifically designated sources and resources. Parts in word or spirit taken from sources are identified as such.

---

Bremen, Date

---

Signature

## 12 Dedication

In the first place I have to say thanks to Christine Klaas and Dieter Wolf-Gladrow for having me as a participant of Polarstern cruise ANTXXVIII-3. I would also like to thank all the people at AWI who helped me getting this Thesis done.

I dedicate this Thesis to the Hustedt Diatom Study Center at AWI Bremerhaven. All the support and inside views of the scientific world about diatoms were amazing. Thanks to Friedel Hinz, always having the right book at hand immediately and advice where ever is needed. Thanks to Stefan Pinkernell for any ‚technical‘ support and great thanks to Michael Kloster. Your software is fantastic.

Finally I have to thank Dr. Bánk Beszteri for realizing this Thesis, patience and just being the ‚leading strand‘ throughout the Project. Thank you very much!

Thanks to my parents just for always being there and calming me.