

1997

# THE DEVELOPMENT OF NOVEL TECHNIQUES FOR CHARACTERISATION OF MARINE ZOOPLANKTON OVER VERY LARGE SPATIAL SCALES

GALLIENNE, CHRISTOPHER PAUL

<http://hdl.handle.net/10026.1/1727>

---

<http://dx.doi.org/10.24382/3917>

University of Plymouth

---

*All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.*

**THE DEVELOPMENT OF NOVEL TECHNIQUES FOR CHARACTERISATION  
OF MARINE ZOOPLANKTON OVER VERY LARGE SPATIAL SCALES**

**by**

**CHRISTOPHER PAUL GALLIENNE**

A thesis submitted to the University of Plymouth  
in partial fulfilment for the degree of

**DOCTOR OF PHILOSOPHY**

Institute of Marine Studies

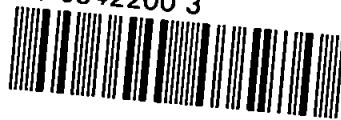
Faculty of Science

In collaboration with

Plymouth Marine Laboratory

**June 1997**

90 0342200 3



UNIVERSITY OF PLYMOUTH	
Item No.	900.3422003
Date	14 NOV 1997 5
Class No.	T574.526366A
Contl. No.	X703589406
LIBRARY SERVICES	

LIBRARY STORE

REFERENCE ONLY

## Abstract

Marine zooplankton play an important role in the transfer of CO<sub>2</sub> from the atmosphere/ocean system to deeper waters and the sediments. They also provide food for much of the world's fish stocks and in some areas of the ocean depleted of nutrients they sustain phytoplankton growth by recycling nutrients. They therefore have a profound effect on the carbon cycle and upon life in the oceans. There is a perceived lack of information about global distributions of zooplankton needed to validate ecosystems dynamics models, and the traditional methods of survey are inadequate to provide this information. There is a need to develop new technologies for the large scale survey of zooplankton, which should provide data either suitable for quick and easy subsequent processing, or better still, processed in real time.

New technologies for large scale zooplankton survey fall into three main categories: acoustic, optical and video. No single method is capable of providing continuous real time data at the level of detail required. A combination of two of the new technologies (optical and video) has the potential to provide broad scale data on abundance, size and species distributions of zooplankton routinely, reliably, rapidly and economically. Such a combined method has been developed in this study. The optical plankton counter (OPC) is a fairly well established instrument in marine and freshwater zooplankton survey. A novel application of the benchtop version of this instrument (OPC-1L) for real time data gathering at sea over ocean basin scales has been developed in this study. A new automated video zooplankton analyser (ViZA) has been designed and developed to operate together with the OPC-1L. The two devices are eventually to be deployed in tandem on the Undulating Oceanographic Recorder (UOR) for large scale ocean survey of zooplankton.

During the initial development of the system, the two devices are used in benchtop flow through mode using the ship's uncontaminated sea water supply. The devices have been deployed on four major oceanographic cruises in the North and South Atlantic, covering almost 40,000 km. of transect. Used in benchtop mode, it has been shown that the OPC can simply and reliably survey thousands of kilometres of ocean surface waters for zooplankton abundance and size distribution in the size range 250µm. to 11.314 mm. in real time. The ViZA system can add the dimension of shape to the OPC size data, and provide supporting data on size distributions and abundance. Sampling rate in oligotrophic waters, and image quality problems are two main limitations to current ViZA performance which must be addressed, but where sufficient abundance exists and good quality images are obtained, the initial version of the ViZA system is shown to be able reliably to classify zooplankton to six major groups.

The four deployments have shown that data on zooplankton distributions on oceanic scales can be obtained without the delays and prohibitive costs associated with sample analysis for traditional sampling methods. The results of these deployments are presented, together with an assessment of the performance of the system and proposals for improvements to meet the requirements specified before a full *in-situ* system is deployed.

# CONTENTS

## CHAPTER 1

### MARINE ZOOPLANKTON

<b>1.1 Oceanic zooplankton</b>	<b>1</b>
<b>1.2 The need for zooplankton survey</b>	<b>2</b>
<b>1.3 Spatial distribution of marine zooplankton</b>	<b>5</b>
<i>1.3.1 Diel vertical migration (DVM)</i>	<i>5</i>
<i>1.3.2 Geographic distributions</i>	<i>6</i>
<i>1.3.3 Patchiness</i>	<i>6</i>
<b>1.4 Temporal distribution of marine zooplankton</b>	<b>7</b>
<i>1.4.1 Diel variation</i>	<i>8</i>
<i>1.4.2 Seasonal variation</i>	<i>9</i>
<i>1.4.3 Long term variation</i>	<i>10</i>

## CHAPTER 2

### THE PROBLEM OF ZOOPLANKTON SURVEY

<b>2.1 Traditional methods of zooplankton survey</b>	<b>11</b>
<i>2.1.1 Sampling nets</i>	<i>11</i>
<i>2.1.2 Pumped systems</i>	<i>13</i>
<i>2.1.3 Towed methods - the CPR &amp; LHPR</i>	<i>13</i>

<b>2.2 Modern methods of zooplankton survey</b>	<b>15</b>
2.2.1 <i>Acoustic methods - Echo sounders</i>	15
2.2.2 <i>Acoustic methods - The acoustic Doppler current profiler (ADCP)</i>	15
2.2.3 <i>Electronic methods</i>	16
2.2.4 <i>Optical methods - The optical plankton recorder (OPC)</i>	17
2.2.5 <i>Photographic methods</i>	17
2.2.6 <i>Video methods</i>	18
 <b>2.3 The limitations of currently available methods</b>	 <b>20</b>
2.3.1 <i>Spatial and temporal resolution</i>	21
2.3.2 <i>Time and economic cost</i>	21
2.3.3 <i>Data processing demands</i>	21
2.3.4 <i>Community structure</i>	22
 <b>2.4 The proposed method</b>	 <b>22</b>

## **CHAPTER 3**

### **THE OPTICAL PLANKTON COUNTER**

<b>3.1 The towed version - OPC-1T</b>	<b>25</b>
 <b>3.2 The laboratory version - OPC-1L</b>	 <b>27</b>
 <b>3.3 Operational issues</b>	 <b>27</b>
 <b>3.4 Size distribution and community structure</b>	 <b>33</b>
 <b>3.5 Conclusion</b>	 <b>34</b>

## CHAPTER 4

### OBJECT REPRESENTATION FOR CLASSIFICATION

<b>4.1 Describing size and shape in the spatial domain</b>	<b>36</b>
<i>4.1.1 The boundary description</i>	38
<i>4.1.2 Size descriptors</i>	40
<i>4.1.3 Shape descriptors</i>	43
<b>4.2 Shape description in the frequency domain - Fourier descriptors</b>	<b>48</b>
<i>4.2.1 Fourier's theorem</i>	48
<i>4.2.2 The Fourier transform</i>	49
<i>4.2.3 Fourier shape descriptors</i>	50

## CHAPTER 5

### TAXONOMIC CLASSIFICATION PROCESSES

<b>5.1 Identifying species</b>	<b>55</b>
<i>5.1.1 The data processing problem</i>	56
<i>5.1.2 Reducing the data set</i>	56
<i>5.1.3 Reducing the operator set</i>	57
<b>5.2 Multivariate analysis</b>	<b>58</b>
<i>5.2.1 Cluster analysis</i>	59
<i>5.2.2 Principal component analysis</i>	61
<i>5.2.3 A proposed subset of size and shape descriptors</i>	68
<b>5.3 Further developments</b>	<b>69</b>

## CHAPTER 6

### THE DESIGN OF A NEW VIDEO ZOOPLANKTON ANALYSER

<b>6.1 System requirements</b>	<b>70</b>
<b>6.2 Instrument design</b>	<b>71</b>
6.2.1 <i>Cameras</i>	71
6.2.2 <i>Optics</i>	72
6.2.3 <i>Illumination</i>	72
6.2.4 <i>Image acquisition and pre-processing</i>	74
6.2.5 <i>Typical ViZA image data</i>	75
6.2.6 <i>Towed platform</i>	77
6.2.7 <i>The OPC and flow through system</i>	77

## CHAPTER 7

### SYSTEM SOFTWARE

<b>7.1 Host processor software</b>	<b>82</b>
7.1.1 <i>Image acquisition software</i>	82
7.1.2 <i>Training set procedure</i>	86
7.1.3 <i>Manual run procedure</i>	88
7.1.4 <i>Auto run procedure</i>	88
7.1.5 <i>Continuous run procedure</i>	90
7.1.6 <i>Data file handling</i>	91
<b>7.2 Principal components analysis</b>	<b>92</b>

## CHAPTER 8

### FIELD OBSERVATIONS AND RESULTS

<b>8.1 Laboratory testing</b>	<b>97</b>
<i>8.1.1 Calibration</i>	<i>97</i>
<b>8.2 Ocean trials - AMT and PRIME cruises</b>	<b>98</b>
<b>8.3 Atlantic Meridional Transect 1 (AMT1)</b>	<b>101</b>
<i>8.3.1 Methods</i>	<i>101</i>
<i>8.3.2 Data results</i>	<i>104</i>
<b>8.4 Atlantic Meridional Transect 2 (AMT2)</b>	<b>110</b>
<i>8.4.1 Methods</i>	<i>110</i>
<i>8.4.2 Data results</i>	<i>111</i>
<b>8.5 PRIME North Atlantic cruise</b>	<b>119</b>
<i>8.5.1 Methods</i>	<i>119</i>
<i>8.5.2 OPC data results</i>	<i>120</i>
<i>8.5.3 ViZA data results</i>	<i>128</i>
<b>8.6 Atlantic Meridional Transect 3 (AMT3)</b>	<b>131</b>
<i>8.6.1 Methods</i>	<i>131</i>
<i>8.6.2 OPC data results</i>	<i>133</i>
<i>8.6.3 ViZA data results</i>	<i>140</i>
<b>8.7 Data validation</b>	<b>153</b>
<i>8.7.1 AMT1</i>	<i>154</i>
<i>8.7.2 AMT2</i>	<i>159</i>
<i>8.7.3 PRIME</i>	<i>160</i>
<i>8.7.4 AMT3</i>	<i>164</i>

## **CHAPTER 9**

### **DISCUSSION AND CONCLUSION**

<b>9.1 OPC data summary</b>	<b>166</b>
<i>9.1.1 Other environmental data</i>	<i>168</i>
<b>9.2 ViZA data summary</b>	<b>172</b>
<b>9.3 Data validation</b>	<b>173</b>
<b>9.4 System performance</b>	<b>173</b>
<i>9.4.1 OPC system performance</i>	<i>175</i>
<i>9.4.2 ViZA system performance</i>	<i>176</i>
<b>9.5 Further development work</b>	<b>178</b>
<i>9.5.1 OPC development</i>	<i>178</i>
<i>9.5.2 ViZA development</i>	<i>180</i>
<b>9.6 System application</b>	<b>181</b>
<b>9.7 Conclusion</b>	<b>183</b>

<b>REFERENCES</b>	<b>185</b>
-------------------	------------

<b>APPENDIX 1</b>	1st three principal components for training set	<b>198</b>
<b>APPENDIX 2</b>	Sample images acquired by ViZA system	<b>203</b>

# LIST OF TABLES

<b>Table 4.1</b>	<b>Boundary chain code elements</b>	<b>39</b>
<b>Table 5.1</b>	<b>A sample set of size descriptors</b>	<b>63</b>
<b>Table 5.2</b>	<b>Normalised set of size descriptors</b>	<b>64</b>
<b>Table 5.3</b>	<b>Covariance matrix</b>	<b>64</b>
<b>Table 5.4a</b>	<b>Eigenvectors of the covariance matrix</b>	<b>65</b>
<b>Table 5.4b</b>	<b>Eigenvalues of the covariance matrix</b>	<b>65</b>
<b>Table 8.1</b>	<b>Size calibration for OPC and ViZA</b>	<b>97</b>
<b>Table 8.2</b>	<b>Correlation coefficients for underway OPC and ViZA data</b>	<b>140</b>
<b>Table 8.3</b>	<b>Coefficients and variance of principal components 1 to 3 for training set used on AMT3.</b>	<b>150</b>
<b>Table 8.4</b>	<b>ViZA identification success rate for test sample of animals</b>	<b>152</b>

# LIST OF ILLUSTRATIONS

<b>Fig. 3.1</b>	Schematic diagram of the principle of operation of the OPC	<b>26</b>
<b>Fig. 3.2</b>	Schematic of the debubbling device	<b>28</b>
<b>Fig. 3.3a</b>	Model prediction for 65 mm. diameter tube, bubble position as a function of time	<b>29</b>
<b>Fig. 3.3b</b>	Model prediction for 150 mm. diameter tube, bubble position as a function of time	<b>30</b>
<b>Fig. 3.4</b>	Typical size distribution from raw OPC data	<b>34</b>
<b>Fig. 4.1</b>	Major chord length and minor chord length (A+B) of an object	<b>42</b>
<b>Fig. 4.2</b>	Comparison between major chord length and fibre length of an object	<b>43</b>
<b>Fig. 4.3</b>	(a) $r(\theta)$ defined for all $\theta$ ; (b) $r(\theta)$ undefined for some values of $\theta$	<b>47</b>
<b>Fig. 4.4</b>	Slope density function, $\theta(s)$ .	<b>47</b>
<b>Fig 4.5</b>	Argand diagram of $F(u)$ represented as a vector in the complex plane	<b>50</b>
<b>Fig. 5.1</b>	Grouping of PCA example data set on first two principal components.	<b>67</b>
<b>Fig. 6.1</b>	Circuit diagram for strobe trigger	<b>73</b>
<b>Fig. 6.2</b>	Example of image captured using ViZA system during AMT3	<b>76</b>

<b>Fig. 6.3</b>	Binarised (segmented) version of the image in fig. 6.2	<b>76</b>
<b>Fig. 6.4</b>	Boundary contour extracted from the segmented image in fig. 6.3	<b>77</b>
<b>Fig. 6.5</b>	The OPC and ViZA flow through system for underway sampling using the ship's uncontaminated water supply	<b>79</b>
<b>Fig. 7.1</b>	Typical object outline as presented to the user by the acquisition system	<b>87</b>
<b>Fig. 7.2</b>	Illustration of the measure of group separation used by the PCA algorithm	<b>94</b>
<b>Fig. 7.3</b>	Typical graphical display of the results of the PCA	<b>95</b>
<b>Fig. 8.1</b>	AMT cruise plan. Ocean colour background: CZCS colour composite (courtesy of NASA)	<b>99</b>
<b>Fig. 8.2</b>	PRIME cruise track 2-22nd July 1996	<b>100</b>
<b>Fig. 8.3</b>	Transect and stations for AMT1 cruise	<b>102</b>
<b>Fig. 8.4a</b>	AMT1 raw counts along transect (underway)	<b>105</b>
<b>Fig. 8.4b</b>	AMT1 raw counts along transect (underway, cont.)	<b>106</b>
<b>Fig. 8.4c</b>	AMT1 raw counts along transect (underway, cont.)	<b>107</b>
<b>Fig. 8.5</b>	OPC biovolume & mean ESD along transect (net hauls)	<b>108</b>
<b>Fig. 8.6</b>	Size distribution of biovolume along transect (net hauls)	<b>109</b>

<b>Fig. 8.7</b>	Percentage of total biovolume in each size class along transect (net hauls)	<b>109</b>
<b>Fig. 8.8a</b>	AMT2 raw counts along transect (underway)	<b>112</b>
<b>Fig. 8.8b</b>	AMT2 raw counts along transect (underway, cont.)	<b>113</b>
<b>Fig. 8.8c</b>	AMT2 raw counts along transect (underway, cont.)	<b>114</b>
<b>Fig. 8.8d</b>	AMT2 raw counts along transect (underway, cont.)	<b>115</b>
<b>Fig. 8.9</b>	Size distribution of underway biovolume along transect for sampling period local dawn to 1000Z	<b>115</b>
<b>Fig. 8.10</b>	Size distribution of underway biovolume along transect for sampling period 1200Z to local dusk	<b>116</b>
<b>Fig. 8.11</b>	Size distribution of underway biovolume along transect for sampling period local dusk to local dawn	<b>116</b>
<b>Fig. 8.12</b>	OPC biovolume & mean ESD along transect (net hauls)	<b>117</b>
<b>Fig. 8.13</b>	Size distribution of biovolume along transect (net hauls)	<b>117</b>
<b>Fig. 8.14</b>	Percentage of total biovolume in each size class along transect (net hauls)	<b>118</b>
<b>Fig. 8.15a</b>	PRIME transect underway raw counts (days 185-188)	<b>121</b>
<b>Fig. 8.15b</b>	PRIME transect underway raw counts (days 188-192)	<b>121</b>
<b>Fig. 8.16</b>	Size distribution of underway biovolume along transect for sampling period 0200Z to local dawn	<b>122</b>

<b>Fig. 8.17</b>	Size distribution of underway biovolume along transect for sampling period local dawn to 1200Z	<b>122</b>
<b>Fig. 8.18</b>	Size distribution of underway biovolume along transect for sampling period 1400Z to local dusk	<b>123</b>
<b>Fig. 8.19</b>	PRIME homeward leg underway raw counts - integrated over 30 minute intervals	<b>123</b>
<b>Fig. 8.20</b>	Size distribution of underway biovolume along homeward leg - integrated over 30 minute intervals.	<b>124</b>
<b>Fig. 8.21</b>	Size distribution of biovolume along transect (night nets)	<b>125</b>
<b>Fig. 8.22</b>	Percentage of total biovolume in each size class along transect (night nets)	<b>125</b>
<b>Fig. 8.23</b>	Size distribution of biovolume along transect (day nets)	<b>126</b>
<b>Fig. 8.24</b>	Percentage of total biovolume in each size class along transect (day nets)	<b>126</b>
<b>Fig. 8.25</b>	Size distribution of biovolume on 8-day station (day/night nets)	<b>127</b>
<b>Fig. 8.26</b>	Percentage of total biovolume in each size class on station (day/night nets)	<b>128</b>
<b>Fig. 8.27a</b>	AMT3 raw counts along transect (underway)	<b>133</b>
<b>Fig. 8.27b</b>	AMT3 raw counts along transect (underway, cont.)	<b>134</b>
<b>Fig. 8.27c</b>	AMT3 raw counts along transect (underway, cont.)	<b>135</b>

<b>Fig. 8.27d</b>	AMT3 raw counts along transect (underway, cont.)	<b>136</b>
<b>Fig. 8.28</b>	Size distribution of underway biovolume along transect for sampling period local dawn to 1030Z	<b>136</b>
<b>Fig. 8.29</b>	Size distribution of underway biovolume along transect for sampling period 1200Z to local dusk	<b>137</b>
<b>Fig. 8.30</b>	Size distribution of underway biovolume along transect for sampling period local dusk to local dawn	<b>137</b>
<b>Fig. 8.31</b>	OPC biovolume & mean ESD along transect (net hauls)	<b>138</b>
<b>Fig. 8.32</b>	Size distribution of biovolume along transect (net hauls)	<b>139</b>
<b>Fig. 8.33</b>	Percentage of total biovolume in each size class along transect (net hauls)	<b>139</b>
<b>Fig. 8.34</b>	AMT3 morning run - total OPC biovolume less than 2000µm ESD compared to ViZA biovolume less than 2000µm ESD	<b>141</b>
<b>Fig. 8.35</b>	AMT3 morning run - total OPC biovolume 250-500µm ESD compared to ViZA biovolume 250-500µm ESD	<b>142</b>
<b>Fig. 8.36</b>	AMT3 morning run - total OPC biovolume 500-1000µm ESD compared to ViZA biovolume 500-1000µm ESD	<b>142</b>
<b>Fig. 8.37</b>	AMT3 morning run - total OPC biovolume 1000-2000µm ESD compared to ViZA biovolume 1000-2000µm ESD	<b>143</b>
<b>Fig. 8.38</b>	AMT3 afternoon run - total OPC biovolume less than 2000µm ESD compared to ViZA biovolume less than 2000µm ESD	<b>143</b>

<b>Fig. 8.39</b>	AMT3 afternoon run - total OPC biovolume 250-500µm ESD compared to ViZA biovolume 250-500µm ESD	<b>144</b>
<b>Fig. 8.40</b>	AMT3 afternoon run - total OPC biovolume 500-1000µm ESD compared to ViZA biovolume 500-1000µm ESD	<b>144</b>
<b>Fig. 8.41</b>	AMT3 afternoon run - total OPC biovolume 1000-2000µm ESD compared to ViZA biovolume 1000-2000µm ESD	<b>145</b>
<b>Fig. 8.42</b>	AMT3 overnight run - total OPC biovolume less than 2000µm ESD compared to ViZA biovolume less than 2000µm ESD	<b>145</b>
<b>Fig. 8.43</b>	AMT3 overnight run - total OPC biovolume 250-500µm ESD compared to ViZA biovolume 250-500µm ESD	<b>146</b>
<b>Fig. 8.44</b>	AMT3 overnight run - total OPC biovolume 500-1000µm ESD compared to ViZA biovolume 500-1000µm ESD	<b>146</b>
<b>Fig. 8.45</b>	AMT3 overnight run - total OPC biovolume 1000-2000µm ESD compared to ViZA biovolume 1000-2000µm ESD	<b>147</b>
<b>Fig. 8.46</b>	Euphausiid captured from training set of zooplankton from AMT3	<b>148</b>
<b>Fig. 8.47</b>	Amphipod captured from training set of zooplankton from AMT3	<b>148</b>
<b>Fig. 8.48</b>	Graphical display of the results of the PCA on the training set from AMT3. X and Y axes are principal components one and two respectively.	<b>149</b>
<b>Fig. 8.49</b>	Size distribution of sample test set from AMT3	<b>151</b>
<b>Fig. 8.50</b>	Species distribution of sample test set from AMT3	<b>152</b>

<b>Fig. 8.51</b>	AMT1 WP-2 nets - total counts from OPC and microscope analysis.	<b>153</b>
<b>Fig. 8.52</b>	AMT2 WP-2 nets - total counts from OPC and microscope analysis.	<b>154</b>
<b>Fig. 8.53</b>	AMT1 vertical net cast, day 269: biovolume in each size fraction from sample through OPC and from microscope analysis.	<b>155</b>
<b>Fig. 8.54</b>	AMT1 vertical net cast, day 270: biovolume in each size fraction from sample through OPC and from microscope analysis.	<b>155</b>
<b>Fig. 8.55</b>	AMT1 vertical net cast, day 271: biovolume in each size fraction from sample through OPC and from microscope analysis.	<b>156</b>
<b>Fig. 8.56</b>	AMT1 vertical net cast, day 272: biovolume in each size fraction from sample through OPC and from microscope analysis.	<b>156</b>
<b>Fig. 8.57</b>	AMT1 vertical net cast, days 269-272: total biovolume from sample through OPC, and from microscope analysis.	<b>157</b>
<b>Fig. 8.58</b>	AMT1 net samples - total OPC biovolume, carbon and nitrogen	<b>158</b>
<b>Fig. 8.59</b>	AMT1 net samples - carbon to nitrogen ratio and % carbon	<b>158</b>
<b>Fig. 8.60</b>	AMT2 net samples - total OPC biovolume, carbon and nitrogen	<b>159</b>
<b>Fig. 8.61</b>	AMT2 net samples - carbon to nitrogen ratio and % carbon	<b>159</b>
<b>Fig. 8.62</b>	PRIME transect daytime net samples - total OPC biovolume carbon and nitrogen	<b>161</b>
<b>Fig. 8.63</b>	PRIME transect daytime net samples - carbon to nitrogen ratio and % carbon	<b>161</b>

<b>Fig. 8.64</b>	PRIME transect nighttime net samples - total OPC biovolume carbon and nitrogen	<b>161</b>
<b>Fig. 8.65</b>	PRIME transect nighttime net samples - carbon to nitrogen ratio and % carbon	<b>162</b>
<b>Fig. 8.66</b>	PRIME station daytime net samples - total OPC biovolume carbon and nitrogen	<b>162</b>
<b>Fig. 8.67</b>	PRIME station daytime net samples - carbon to nitrogen ratio and % carbon	<b>162</b>
<b>Fig. 8.68</b>	PRIME station nighttime net samples - total OPC biovolume carbon and nitrogen	<b>163</b>
<b>Fig. 8.69</b>	PRIME station nighttime net samples - carbon to nitrogen ratio and % carbon	<b>163</b>
<b>Fig. 8.70</b>	AMT3 net samples - total OPC biovolume, carbon and nitrogen	<b>164</b>
<b>Fig. 8.71</b>	AMT3 net samples - carbon to nitrogen ratio and % carbon	<b>164</b>
<b>Fig. 9.1</b>	OPC total net biovolume comparison between AMT1 and AMT2	<b>167</b>
<b>Fig. 9.2</b>	OPC total net biovolume comparison between AMT1 and AMT3	<b>167</b>
<b>Fig. 9.3</b>	Surface sea temperature for AMT1, days 269-298 (north to south).	<b>169</b>
<b>Fig. 9.4</b>	Zooplankton size distribution, and rate of change of temperature during transect of the area of the confluence of the Brazil and Falkland currents, between Montevideo and the Falklands, AMT1.	<b>170</b>

<b>Fig. 9.5</b>	Zooplankton size distribution, and rate of change of temperature during transect of the area of the confluence of the Brazil and Falkland currents, between Montevideo and the Falklands, AMT2.	<b>171</b>
<b>Fig. 9.6</b>	Zooplankton size distribution, and rate of change of temperature during transect of the area of the confluence of the Brazil and Falkland currents, between Montevideo and the Falklands, AMT3.	<b>170</b>
<b>Fig. 9.7</b>	Example from AMT3 ViZA data set showing effect on outline of transparency to white backlighting	<b>176</b>
<b>Fig. 9.8</b>	Example from AMT3 ViZA data set showing little or no effect on outline of transparency to white backlighting	<b>177</b>

## **Acknowledgments**

I would like to take this opportunity to express my thanks to the following people for their help and support during this study.

To my supervisors, Drs. D.A. Pilgrim, D.A. Huntley (University of Plymouth, Institute of Marine Studies) and G. Ludbrook (Defence Research Agency); especially to Dr. Pilgrim, my director of studies, who has given me as much support and encouragement as I could have wished for in the initial selection of this project and throughout the period of its execution.

To Mr. D.B. Robins and Drs. R.P. Harris, J. Aiken and A.J. Bale, (Plymouth Marine Laboratory), principal investigators on the Atlantic Meridional Transect programme and PRIME project P19. The opportunity to take part in the cruises which formed part of these programmes was crucial in providing ocean scale testing of the devices under development.

To the Director of the British Antarctic Survey (BAS) and his staff for their support of the AMT programme, and to the Masters of RRS James Clark Ross and RRS Discovery and their officers and crew, for their friendliness and professionalism in support of the scientific work aboard their vessels during the AMT and PRIME cruises.

To Messrs. R.D. Jones, G. Carter and A. Prideaux, technicians in the Institute of Marine Studies, University of Plymouth, for their willing support and provision of advice, technical support, and equipment vital for the successful completion of this study.

## **Author's declaration**

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

This study was financed with the aid of a studentship from the University of Plymouth, and carried out in collaboration with Plymouth Marine Laboratory. Additional funding for equipment costs was provided by the Natural Environment Research Council (NERC) community research programme Plankton Reactivity in the Marine Environment (PRIME). The project encompassing the present research is PRIME special topic P19 - "The optical characterisation of zooplankton in relation to ocean physics; discrimination of seasonal, regional and latitudinal variations".

A programme of advanced study was undertaken, which included two modules of the MSc course in Intelligent Systems at the University of Plymouth, and a three day residential course in Optical Engineering.

Four major research cruises were undertaken in the course of this research: AMT 1, 2 & 3 and the PRIME North Atlantic cruise, a total of 133 days at sea, covering nearly 40,000 km. of transect of the North and South Atlantic Oceans.

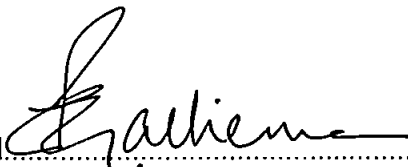
During the first year of this research programme, a contract was undertaken for the Defence Research Agency to design and construct an automated liquid crystal control shutter system for an underwater SIT camera. This project was successfully completed, and provided experience in developing automated control systems for underwater imaging devices which was very valuable in the current research project. The report to the DRA on the completion of this project was classified confidential and unpublished, but is available on a restricted basis at the University of Plymouth.

Gallienne, C.P. (1994). *An automated control system for a liquid crystal shutter*. Unpublished report, University of Plymouth.

Preliminary results from this work were presented at an AMT cruise workshop in Plymouth, a PRIME cruise workshop in Bangor, and to the Dunstaffnage Marine Laboratory in Oban. Much of the work is ongoing, and the ViZA system requires some further development, but some encouraging results have been obtained. The work with the OPC has produced two publications:

Gallienne, C.P., Robins, D.B. & Pilgrim, D.A. (1996). Measuring abundance and size distribution of zooplankton using the optical plankton counter in underway mode. *Underwater Technology* **21**(4):15-21.

Robins, D.B., Bale, A.J., Moore, G.F., Rees, N.W., Hooker, S.B., Gallienne, C.P., Westbrook, A.G., Maranon, E., Spooner, W. & Laney, S.R. (1996) AMT-1 cruise report and preliminary results. SeaWiFS Technical Report Series. NASA Goddard Space Flight Center, Greenbelt, Md. USA.

Signed  .....

Date 1st Nov 1997 .....

# **CHAPTER 1**

## **MARINE ZOOPLANKTON**

Marine zooplankton are those marine animals which, although capable of movement, are not capable of making their way against a current. They therefore drift with the currents. They are usually distinguished from phytoplankton by being heterotrophic, that is they require external organic sources of energy to synthesize body tissue, although this distinction does not always hold (Newell & Newell, 1963). Zooplankton may be classified as herbivorous, carnivorous, detritivorous or omnivorous. They may also be classified as holoplankton (those animals which remain planktonic throughout their life cycle) or meroplankton (those which are planktonic for only part of their life cycle). This study is concerned with the development of new technologies for the large scale survey of oceanic zooplankton.

### **1.1 Oceanic zooplankton**

Oceanic zooplankton are mostly holoplankton, and are usually dominated by the crustacea, amongst which the copepods are the most significant in both numbers and species. Other common crustacean types are euphausiids, amphipods, ostracods and (less significantly) cladocera and decapods. Common non-crustaceans are the chaetognaths, polychaetes, pteropods, the chordata (salps and doliolids) and the coelenterates (medusae, siphonophores and ctenophores).

Fish eggs and fish larvae can also provide a significant portion of the

zooplankton in some areas. Plankton are generally quite small, mostly less than 1 cm. in length, and are often classified according to size. Different researchers differ somewhat about the size classification, but the following is typical, representing a decadal scale:

ultraplankton	< 2 $\mu\text{m}$ .
nanoplankton	2 - 20 $\mu\text{m}$ .
microplankton	20 - 200 $\mu\text{m}$ .
mesoplankton	200 - 2000 $\mu\text{m}$ .
macroplankton	2 - 20 mm.
megaplankton	> 20 mm.

The mesoplankton and macroplankton are sometimes grouped together as mesoplankton, and it is this size category (200 $\mu\text{m}$  - 20 mm) which contains the greatest part of the zooplankton, and with which this study is concerned. The zooplankton tend to be larger at lower temperatures (higher latitudes) for any given type, but the warmer waters (lower latitudes) tend to have a much greater variety of species (Raymont, 1963).

## **1.2 The need for zooplankton survey**

The oceans of the world play a significant part in the balance of the global ecosystem and in global climate change, due to their ability to store and transport heat, and to source and sink those gases which play a part in heat transfer. The oceans contain about 92% of the total  $\text{CO}_2$  in the atmosphere/ocean system (Mann & Lazier, 1996), and ocean sediments contain many times more than this amount (Schindler, 1991). There is concern about the

possible effects of man-made changes to the earth's climate. An increase by 10% in carbon dioxide (CO<sub>2</sub>) levels in the atmosphere in the last 35 years has been linked to global warming due to the atmospheric greenhouse effect. Such increases have occurred in the past as part of a natural cycle, and the long term effects of man's impact over the last 150 years are unpredictable. Global climate models show, however, that the present rate of increase of atmospheric CO<sub>2</sub> could lead to a consistent rise in global temperatures, rising sea levels, and changes in precipitation patterns (Huntley, 1992). These climate models predict that one of the biggest changes in temperature will be in the North Atlantic, with implications for an already overpressed fisheries industry. The oceans have a significant role in any such change, through their part in carbon fluxes and the fate of CO<sub>2</sub>. Anthropogenic inputs to the atmosphere of CO<sub>2</sub> have been estimated to result in an increase of 5.3 Gt. y<sup>-1</sup>, whilst measured atmospheric concentrations are increasing only at the rate of 1.2 p.p.m. *per annum*, equivalent to about 2.1 Gt. y<sup>-1</sup> (Mann & Lazier, 1996). We need to know where the missing 3.2 Gt. are going.

Inorganic carbon is photosynthetically fixed by marine phytoplankton, which in turn provide the food energy source for zooplankton. Zooplankton play a key role in the flux of carbon within the oceans, and are one of the primary mechanisms for its transfer out of the surface waters and into the deeper waters and sediments (Harris & Duncan, 1993, Banse, 1995). Most marine animals spend some or all of their time as plankton, feeding on phytoplankton, and most fish feed on zooplankton at some stage of their development. Zooplankton

therefore play a key role in food web dynamics, representing the main food of many important fish stocks. Furthermore, in many oligotrophic areas of the world's oceans, zooplankton can be responsible for feeding the phytoplankton, by providing recycled nutrients in otherwise nutrient depleted areas (Banse, 1995)

Biogenic particles are chiefly responsible for ocean backscatter and absorption of incident radiant energy from the sun, affecting the albedo of the sea surface. They therefore play an important part in the surface heating effect of this radiation (Aebischer *et al.*, 1990, Sathyendranath *et al.*, 1991).

The spatial and temporal distribution of zooplankton therefore has a significant effect on CO<sub>2</sub> 'drawdown' through several mechanisms. Zooplankton have profound effects on the carbon cycle and life in the oceans, and have formed the focus of the US global ocean ecosystems dynamics (GLOBEC) programme (Huntley, 1992). There is a perceived lack of information on global distributions of biological properties. Traditional methods of zooplankton survey are unsuitable for gathering information on global/ocean basin scales (chapter 2). In order to measure and model these properties, there is a need for improved methods and technologies for the measurement of zooplankton abundance and community structure from small to large spatial and temporal scales. This should lead to an improved understanding of growth rates, and functional role - grazing and predator/prey relationships - and of biogenic fluxes (GLOBEC, 1992). Linked to simultaneous measurements of physical oceanographic

processes, this can enhance our understanding of, and ability to predict, population changes in response to global change. Models of biospheric carbon fluxes require information from the real world, and new sampling techniques are required to obtain these data at sufficiently fine spatial and temporal resolutions over sufficiently large spatial and temporal scales.

### **1.3 Spatial distribution of marine zooplankton**

Zooplankton are well distributed throughout the world's oceans, and throughout the water column, and have been found at depths greater than 6000 metres (Raymont, 1963), although at far lesser concentrations than the maximum which is usually found in or near the euphotic zone (0 - 300 metres), where the concentration of primary food sources (phytoplankton) are greatest.

#### *1.3.1 Diel vertical migration (DVM)*

Many species have been shown to migrate vertically through the water column on a daily basis, moving towards the surface during the night. While this migration appears to be linked to light levels, several explanations have been offered, though none has yet been proven. DVM may be caused by the conflict between the need for food and the need to hide from predators; the need to conserve energy - metabolic rates slow down in the colder, deeper water - or the need to use vertical current differentials to move to new, less depleted food resources (Longhurst, 1976). This latter effect may be enhanced by changes in phytoplankton concentrations. Higher concentrations attenuate light by a given amount over much smaller depth ranges, thereby reducing the amplitude of

DVM. This will result in less horizontal advection of the animals, allowing them to stay with rich food sources for longer (Longhurst, 1981).

Seasonal vertical migrations also occur, often linked to developmental stages in animals, copepods being an example (see below). Whatever the reason, this vertical migration enhances the transport of organic material from the surface to deeper waters, faster than would be due to sinking rates alone. This transport provides an important link in the marine food chain and in carbon flux, transporting organic carbon out of the euphotic zone.

### *1.3.2 Geographic distributions*

Zooplankton species may be endemic to particular oceanic areas, or widespread. Some species are so localised with respect to particular water masses that they are used as tracers for these water masses (Omori & Ikeda, 1984). Typical ocean biomass abundance vary between about 5 and 500 mg/m<sup>3</sup>, and are considerably higher in coastal areas.

### *1.3.3 Patchiness*

A distribution phenomenon recently much observed and studied is that of patchiness (Levin *et al.*, 1993). This is a tendency for large local gradients in concentration of zooplankton to occur over relatively short distances (sometimes of the order of a few metres - micropatchiness). It has been observed that using current methods the dominant scale of patchiness is often smaller than the sampling interval of the equipment used, and in order to

measure all such variations we need to sample at spatial scales ranging from centimetres to 10's of kilometres, and temporal scales of hours to months (Longhurst, 1981).

Patchiness is important in modelling biosystems, particularly for prediction of feeding and growth rates, population and ecosystem dynamics. There is some evidence that feeding in micropatches results in enhanced growth rates (Owen, 1989). Laboratory studies of feeding rates suggest that typical gross oceanic concentrations of food sources are insufficient to sustain the populations of zooplankton which are known to exist. Patchiness could offer some explanation for this apparent paradox. Physical determinants of patchiness have been suggested (Haury & Pieper, 1988; Genin *et al.*, 1994). Amongst these are turbulent mixing causing aggregation and dispersion; upwelling; and on a smaller scale, Langmuir circulation. Biological causes which have been proposed include predator-prey interactions, seasonal growth cycles, breeding patches and progeny swarms. Traditional methods of sampling, such as towed plankton nets, have masked this phenomenon, as samples are integrated over the length of the tow. Whatever the cause, this is an important area of study, and only the newer sampling technologies have the spatial resolution required to detect such patchiness (chapter 2).

#### **1.4 Temporal distribution of zooplankton**

Temporal and spatial variations in zooplankton abundance, as with physical parameters, can often be very difficult to separate, and variations at any point

will probably include both advective and time varying components.

Nevertheless, it is possible to identify certain influences on zooplankton abundance which are purely temporal variations.

#### *1.4.1 Diel variation*

As mentioned above, many species of zooplankton exhibit DVM, and while this produces a vertical spatial distribution, it obviously results in temporal variations in abundance at any given point. Furthermore, given the existence of current shear with depth in many ocean areas, this temporal change in vertical distribution will result in a further change in horizontal distribution with time. A phenomenon associated with DVM is the appearance of peaks in abundance at the surface at dawn and dusk, with midnight sinking apparent between these peaks. It has been suggested (Raymont, 1963) that this may be due to species seeking 'optimum' light levels. As light levels fall below this optimum, these species rise to the surface. Further reduction in light level to a point at which vertical position makes no difference results in the cessation of active swimming in response to light levels, and the natural tendency of zooplankton to sink passively takes over. As dawn approaches, and light levels again increase above this threshold, these species again rise in an attempt to reach this optimum light level. Shortly after dawn the surface light level exceeds this optimum, and the zooplankton again actively descend to deeper waters. There is some evidence of this effect in the data obtained in the current study in the sub-tropical gyre of the South Atlantic (chapter 8).

#### *1.4.2 Seasonal variation*

Large biomass of zooplankton is often associated with seasonal 'blooms' of phytoplankton. These seasonal cycles of phytoplankton abundance are thought to be caused by periods of strong mixing, bringing new nutrients into the euphotic zone, followed by a period of stratification and sufficient sunlight, which keeps phytoplankton cells in the euphotic zone long enough to synthesize new body tissue using these nutrients. In mid-latitude, these conditions typically occur in the spring after winter mixing, producing strong blooms of phytoplankton. Zooplankton peak abundance typically occurs a few weeks after the peak of phytoplankton abundance (May/June in the northern hemisphere). Secondary blooms often occur in the autumn, after equinoctial weather conditions have caused further turbulence to replenish nutrients depleted by the first phytoplankton bloom. Subsequent re-stratification produces the second bloom, often weaker than that of spring due to reduced sunlight at this time of year. The second zooplankton abundance peak then follows toward the end of autumn (October/November in the northern hemisphere).

Higher latitudes generally produce just one such bloom each year, in mid-summer, as it is only then that sufficient sunlight is available for photosynthesis to occur. In tropical waters, at lower latitudes, there is generally no such seasonal variation, with fairly constant temperature, insolation and mixed layer depth throughout the year (Blackburn, 1981). Phytoplankton and zooplankton production and abundance therefore tends to fluctuate with much less amplitude around a much lower level than at higher latitudes. Production is limited by

poor supply of nutrients from across the permanent thermocline, and from secondary sources (breakdown products).

Many species have been found to exhibit seasonal vertical migration, often associated with the developmental stage in their life cycle. Some species tend to winter at greater depths, moving into shallower water during the summer (Raymont, 1963). Seasonal variations in abundance of certain species has also been associated with spawning, breeding adults spawning at depth. The newly hatched young quickly rise to the surface waters, and throughout their developmental stages, descend to deeper waters until they too are ready to spawn (Longhurst, 1981). As discussed above, these vertical migrations also have an effect on horizontal distributions, given current shear with depth.

#### *1.4.3 Long-term variation*

Long term changes in distributions of zooplankton have been attributed to cyclical or permanent climatic changes. The detection of long-term cyclical temperature changes in the North Atlantic, and its effect on zooplankton species abundance in waters around the United Kingdom (the Russell cycle) are one example of this (Cushing, 1981).

# CHAPTER 2

## THE PROBLEM OF ZOOPLANKTON SURVEY

Conventional methods of zooplankton abundance survey (i.e. plankton nets and pumped samplers) are time consuming and have poor spatial and temporal resolution. While the introduction of continuous plankton recorders has addressed some of these problems, the spatial and temporal resolution, and sample preservation and processing methods can still be unsatisfactory. There is widespread agreement that there is a need to develop modern high data rate particle samplers to address this problem (Sprules, 1992, GLOBEC 1992, Herman, 1993, Denman *et al.*, 1985). As noted by Mann & Lazier (1996),

“Instruments for continuous measurement of zooplankton are still in the development stage. Further advances of marine ecology require major improvements in automated methods for collection of data on zooplankton.”

The latest technologies for the real time (in-situ) characterisation of zooplankton over large spatial scales have begun to address these problems, and fall into three main categories: acoustic, optical, and video techniques.

### 2.1 Traditional methods of zooplankton survey

#### 2.1.1 Sampling nets

This is still the most widely used method of collecting zooplankton samples.

The net usually consists of a cone of bolting silk whose mouth is held open by a

rigid ring to which is attached the towing bridle. At the narrow (cod) end is the collecting jar into which (most of) the plankton passes. Mesh sizes vary, and no one mesh size is suitable for all sizes of zooplankton. Various improvements have been made, such as opening and closing devices. Apart from the poor temporal and spatial resolution of nets, due to integration of the sample over the tow path, and the time consuming task of hauling in and re-deploying the net, there are several other problems associated with these nets.

Clogging can be a major problem, particularly with the finer mesh sizes. This results in sample losses as water spills from the front of the net, and also exacerbates the other main problem of resistance to towing. The latter problem can cause net damage, and also creates a pressure wave ahead of the net, which can be detected by some species, which are then capable of net avoidance. The more fragile (gelatinous) species of zooplankton can be severely damaged and/or extruded by collecting nets, resulting in an overcounting of the more sturdy crustacean species such as copepods. The latter species often make up more than 70% of net samples. Smearing of the sample can also be a problem. This phenomenon results from the time differential to travel down the length of the net for different organisms.

A recent sophistication of this type of sampling gear are the multiple opening/closing net systems, e.g. BIONESS (Sameoto *et al.*, 1980), and the MOCNESS system of the Woods Hole Oceanographic Institution. MOCNESS carries 5 to 20 nets with mouth openings between 1/4 and 20 m<sup>2</sup>, and mesh sizes of 64µm to 3 mm (Wiebe *et al.*, 1985). This system is claimed to sample from

scales of 10's of metres to 10's of kilometres, simultaneously collecting data on the physical environment. Once the net samples have been obtained, they must be preserved (usually in formaldehyde) and taken back to the laboratory, where they must be individually examined, identified, counted and classified by a trained observer using a microscope, tweezers and a counting dish. This time consuming and labour intensive task adds greatly to the cost and complexity of sampling.

### *2.1.2 Pumped systems*

Pumps may be used to sample at discrete fixed stations, perhaps at a series of depths, or continuously (underway). They are limited in depth by the complexity and length of hose required. The dynamics of flow through the hose and pump can cause turbulence, damaging fragile species, and smearing, causing a bias in the data. Sampled volumes are restricted by the hose dimensions, and avoidance at the intake can be a problem (Omori & Ikeda, 1986). Sample analysis must take account of the delay between water entering the intake, and reaching the sampling point. Most ships' non-toxic supplies, often used for pumped sampling, have filters, whose mesh size must also be taken into account.

### *2.1.3 Towed methods - the CPR and LHPR*

Many continuous plankton samplers based upon some kind of filtration mesh have been developed, the best known and most widely used being the continuous plankton recorder (CPR, Glover, 1967), and the Longhurst Hardy Plankton Recorder (LHPR, Longhurst *et al.*, 1976), which have addressed some

of the problems with traditional net sampling. Both are towed devices which collect zooplankton between two gauze strips mechanically or electrically rolled on to a common spool immersed in preserving fluid. One of the two gauze sheets passes across a channel through which the sampled sea water passes as the device is towed through the water. Particles larger than the gauze mesh size are trapped on the gauze, and then sandwiched between the two gauze layers as they are wound on to the destination spool, contained in a Formaldehyde (preservative) filled tank. The gauze may subsequently be examined and species identified and counted in the same way as for a sampling net. The CPR horizontally integrates the sample over several km at a fixed depth of 10 metres, and so also suffers from poor spatial resolution. The LHPR device is towed obliquely from a given depth to the surface. The time between discrete advances of the filtering gauze can be adjusted to vary the vertical sampling resolution.

Whilst wide areas can quickly be sampled using these devices, there are a few problems inherent in the method. The deployment and recovery of these instruments may be less problematic than with traditional nets, as tow duration may be greatly extended. Subsequent laboratory analysis of the sample presents the same problems as the traditional net methods. Organisms may be damaged by/extruded through the gauze, particularly as the two layers are pressed together. The mouths of these samplers are often quite narrow, so that avoidance by the more mobile of the plankton may still be a problem. High towing speeds may avoid this problem (Bernard *et al.*, 1973). Analysis of the

sample is just as expensive in time and labour as with conventional net sampling.

## **2.2 Modern methods of zooplankton survey**

### *2.2.1 Acoustic methods - Echo sounders*

Simple echo sounders can be used to obtain abundance and biomass estimates for zooplankton acoustic targets. With larger targets and a large number of echoes, assuming a random distribution of targets, statistical analysis of echo responses can produce an acoustic size distribution (Smith *et al.*, 1992). This method can only be used on larger targets, such as the larger macroplankton or small fish, and sophisticated statistical analysis of the data is required.

### *2.2.2 Acoustic methods - the acoustic Doppler current profiler (ADCP)*

The acoustic Doppler current profiler (ADCP) measures the strength of the backscattered acoustic signal as a function of distance from the instrument, and zooplankton biomass can be estimated from the strength of the calibrated signal. Flagg & Smith (1989) discuss the method of using the ADCP to produce estimates of zooplankton biomass. The strength of the backscattered signal from an organism increases monotonically with frequency, and the scattering is of the Rayleigh type. As the wavelength of the signal approaches the length of the scatterer, the scattering changes from Rayleigh to geometric scattering, and resonance causes fluctuations in signal strength. Using multiple frequencies, these changes are used to obtain a measure of size distribution of the scatterers (Smith *et al.*, 1992). The use of range gating allows a three-dimensional survey

to be made. The ADCP gives a broad scale, coarse resolution measure of zooplankton abundance and size distribution for sizes greater than  $\sim 1$  mm., over small and large spatial scales (Smith *et al.*, 1992). The technique is non-invasive, and can be left untended in moored configuration for long periods, but requires complex data handling and calibration, and gives no indication of community structure.

### 2.2.3 *Electronic methods*

Herman and Dauphinee (1980) describe a method of towing an electronic plankton counter on a 'Batfish' vehicle, together with fluorometer and conductivity-temperature-depth unit. The plankton counter was based on the familiar Coulter Counter principle, measuring changes in conductivity as a particle passes between two electrodes through a narrow aperture, and deriving a volume for the particle from this change. A concentrating net is placed ahead of the counter cell. The system can count up to 10 animals per second whilst keeping error due to coincidence counts down to 3-4%. Size range was  $350\text{ }\mu\text{m}$  to 3 mm. The authors found that preserved samples produced little or no response from the instrument. Counts are transferred to a computer aboard the surface vessel for on-line plotting and storage. Dimensional accuracy was stated to be between 15 and 20%. Limitations listed include the small sampling volumes dictated by the small (3 - 5 mm.) sampling aperture, and the level of electronic sophistication of the device. Towing duration was limited to 3 hours (Herman, 1993). The principal author subsequently switched his attention to optical counting methods (see OPC, below).

#### 2.2.4 Optical methods - the optical plankton counter (OPC)

The optical plankton counter (OPC) was developed by Alex Herman at the Bedford Institute, Nova Scotia, and is capable of large scale, rapid and continuous sampling of zooplankton. It uses a collimated beam of light through an enclosed volume, received by a sensor. When this beam of light is interrupted by particles of various size, the sensor signal produces a response proportional to the equivalent cross-sectional area of the particle.

The OPC is able to produce reliable abundance and size distribution indicators for zooplankton between 0.25 and 20 mm in equivalent spherical diameter (ESD), at data rates up to 200 samples  $\text{sec}^{-1}$ . Sensor response time and coincidence counts limit the densities at which the OPC can operate to 3000  $\text{m}^{-3}$ , and it cannot differentiate between more than a few species well separated in size and already known to be present (Sprules *et al.*, 1992). Bias and coincidence count problems require calibration of the instrument against plankton net tows, or some other sampling device. Laboratory calibration against spherical glass beads of known size provides the initial calibration.

#### 2.2.5 Photographic methods

Ortner *et al.* (1979) have described a technique for the rapid and accurate characterisation of zooplankton samples using high-speed silhouette photography which can be undertaken at sea without the need for sample preservation. Samples were placed in a shallow tray on top of the 8 X 10 inch film, which was exposed from above by a xenon strobe of 3  $\mu\text{s}$ . flash duration. Samples could then be counted and identified from the film, which provided a

permanent record. An in-situ towed version of the system, using concentrating nets ahead of a camera system, was subsequently developed, and reported by Ortner *et al.* (1981). While this system avoids some of the problems associated with the deployment and use of net systems, and the preservation of samples, the counting and identification of the samples must still be carried out by a human investigator. Duration and spatial resolution is also limited by the amount of film which can be carried.

#### 2.2.6 Video methods

Video devices for the study of zooplankton may be divided into two classes - those used for large scale survey of zooplankton distributions, and those used for small scale study of individual animals' behaviour, such as feeding strategies and predator-prey relationships. The present study is mainly concerned with the survey devices.

Video survey devices such as the video plankton recorder (VPR) are now beginning to be developed which aim to produce abundance and size distribution indicators of good temporal and spatial resolution, together with near real-time classification into major taxonomic groups (Berman, 1990a, Davis, *et al.*, 1992). The amount of data produced by such devices requires a great deal of processing, however, Berman (1990b) reporting that his system can process 300-500 animals per hour with a 90% accuracy using a multi-processor system, and preserved samples. The objective of real-time data gathering cannot be achieved until image processing technology and techniques are further advanced.

### *Small spatial scales*

Video systems have been used for the observation of zooplankton behaviour and distribution on small spatial scales. Devices developed over the last decade include the ecoSCOPE (Kils, 1992), Ichthyoplankton recorder (Welsch *et al.*, 1991), CritterCam, BoulderCam, and the low cost IR camera system (Schulze *et al.*, 1992). These are generally intended for behavioural studies, although the device developed by Welsch *et al.* has been used to resolve small scale patchiness at tow speeds up to 2.5 m/s (Lenz *et al.*, 1995). This device is installed in the cod end of a plankton net, and can capture images of zooplankton in the size range 0.5 to 20 mm. At the time of publication (August 1995), images were not processed concurrently, however, but recorded for later visual analysis.

### *Large spatial scales*

Davis *et al.* (1992) have produced initial results from laboratory and dockside tests of their Video Plankton Recorder (VPR). The device consists of a towed frame 3 metres long with four cameras each having concentric fields of view (5 mm. to 10 cm.) to cover the size range of interest. The imaged volume is defined by the oblique intersection of the cone defining the camera field of view, and that produced by a collimated strobed (1  $\mu$ s) 80 watt xenon light beam. This arrangement produces a kind of dark field illumination. Sampled volume varies from 1 ml. to 1 litre. Fibre optic telemetry is used to send the four video channels to the surface vessel, where the image processing system can handle each video field. Subsequent upgrading of this system is expected to

permit all pre-processing to be done in real time, and near real time production of size and species distributions. Concurrent calibration of the device is provided by an integrated LHPR system, and the image processing system is an adaptation of that of Berman (1990a). The device is able to discriminate between detrital material and zooplankton, which is not true of optical and acoustic systems. At present, video data is not analysed in real time, but recorded to tape for subsequent automated analysis (Davis *et al.*, 1996).

Huller *et al.* (1994) have reported the development of an interesting device combining the methods of the Coulter Counter, the chlorophyll fluorometer and plankton taxonomy by video image processing. The device can process animals in the size range 100  $\mu\text{m}$ . to 2 mm. This “Macro Flow Planktometer” is a laboratory system, used in flow through mode, and sampling volumes are rather too small to process the large number of samples involved in a large scale survey (typical absolute rate 4 ml/sec.).

### **2.3 The limitations of currently available methods.**

In order completely to characterise spatial patterns of zooplankton distributions in the open ocean, we must be able to measure abundance from centimetre to kilometre scales (Sprules *et al.*, 1992). Furthermore, this measurement should result either in data suitable for quick and easy subsequent processing, or better still, in results produced in real-time. This level of performance is only available from the modern instruments - acoustic, optical and video.

### *2.3.1 Spatial and temporal resolution*

Plankton nets have vertical spatial resolution to tens of metres, and horizontal spatial resolution from hundreds of metres to kilometres, depending upon the length of tow (Sprules *et al.*, 1992). This level of performance is somewhat improved in the towed continuous collection systems, and further improved by the modern acoustic, optical and video methods.

### *2.3.2 Time and economic cost*

All the non-electronic methods (plankton net systems, pumped systems, LHPR, CPR, photography) require preservation of the samples for subsequent human analysis (visual analysis of photographs in the last case). Analysis of samples produced by these methods for studies of dominant zooplankton distribution features over large areas can be prohibitively time consuming and therefore expensive (Huntley *et al.*, 1995). These authors estimate that for a recent survey of the California Current, net sample analysis for size composition alone would have required seven man-years (an Optical Plankton Counter was actually used to process the data in real time). Due to the cost of such analysis, many samples may never be analysed. The more advanced methods (ADCP, OPC, Video) avoid this problem by computer analysis of the sample, often in or near real time.

### *2.3.3 Data processing demands*

The data processing demands of the more advanced methods are variable. The OPC can sample at up to 200 animals per second, and at most produces a few

hundred bytes of information in any one second. Any modern desktop PC has more than enough power to process the resulting data stream in real time. Video devices produce huge quantities of data (~15 million bytes per second) and computationally intensive algorithms must be used to process this data. The ADCP method is intermediate between these two, producing manageable quantities of data, but fairly sophisticated statistical analyses must be performed upon the data before any meaningful information may be extracted.

#### *2.3.4 Community structure*

The non-electronic methods listed above all give full information concerning the species community structure, all being analysed by trained human operators. The acoustic and non-video optical methods give no reliable information on community structure, except that which may be inferred from what is already known of the community in terms of the relationship between its size distribution and species distribution. Only the video methods give detailed shape as well as size information, and therefore the ability to discriminate between taxonomic groups.

#### **2.4 The proposed method**

No single method is yet sufficient to fulfill the perceived objective, and it has been suggested that a combination of the available methods may be the best way forward (GLOBEC, 1992). The OPC gives reliable abundance and size distribution data, with far less raw data to process than acoustic or video devices. Species distributions may be obtained by using a video device in

combination with the OPC. A device combining the OPC and a video device has been developed and successfully deployed to sample fish eggs, using video recordings and subsequent visual analysis (Hunter, J.R., 1996). Future work on this device may include the use of video image analysis to recognise eggs of anchovy and sardine and distinguish them from other objects such as bubbles and zooplankton.

The objective of the current research is to produce an *in situ* (towed) video instrument capable of continuous high resolution counting and classification of zooplankton over large spatial and temporal scales in real time. Small scale effects, such as micro-patchiness (scale < 20 m.) should be resolved. The instrument is to be developed and used in tandem with the established technology of the OPC in order to achieve the objectives currently not attainable by any single device. Discrete net sampling and ADCP profiling will also be used for intercalibration. The OPC will provide a reliable bulk count and size distribution, whilst the video instrument will provide taxonomic information, and numerical analysis for intercomparison.

The device should be compact enough to fit with the OPC into the Plymouth Marine Laboratory's undulating oceanographic recorder (UOR), a well proven towed instrument platform (Aiken & Bellan, 1990). In order to simplify the initial problems, a benchtop version of the OPC/Video system will first be developed. It is the development and evaluation of the benchtop system which occupies the main part of the present study. The video instrument will use high

speed, high resolution CCD cameras to image the volumes of water passing through the OPC, whilst both instruments are towed at speeds of up to 5 m/s. It is desired to resolve zooplankton in the size range 200 $\mu$ m - 20 mm. This is the usual size range for mesozooplankton , and matches the size range of the OPC.

# CHAPTER 3

## THE OPTICAL PLANKTON COUNTER

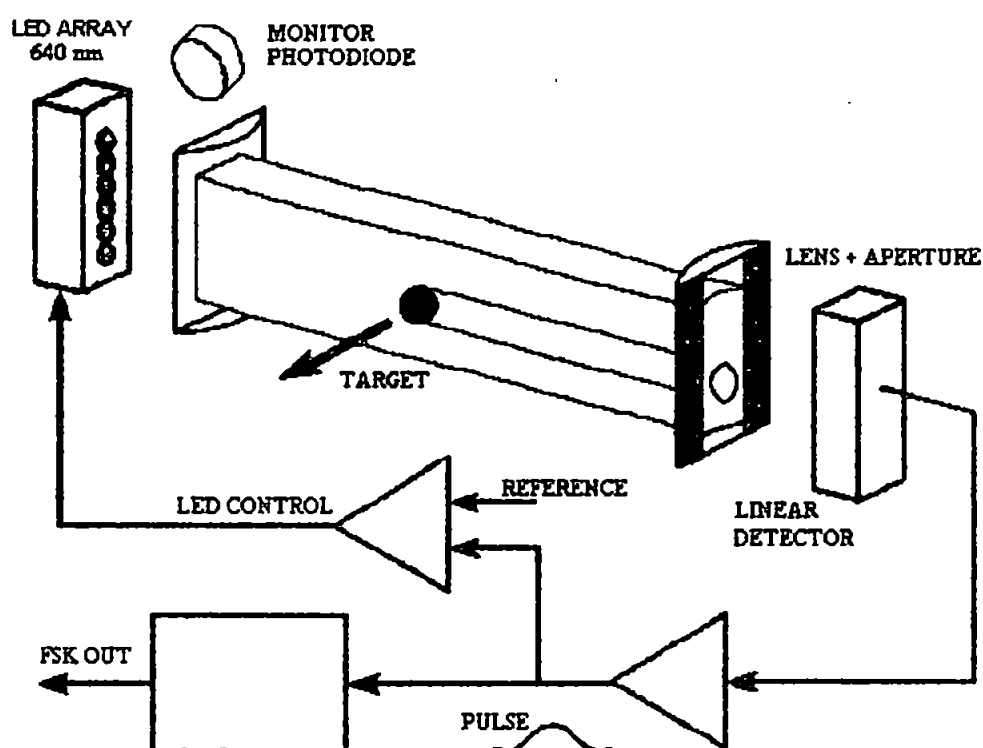
An automated video zooplankton analyser (ViZA) is to be developed to operate in tandem with the optical plankton counter (OPC), as discussed in chapter 2.

This chapter describes both versions of the OPC, and their operational use in the context of the present research programme.

### 3.1 The towed version - OPC-1T

The OPC-1T, manufactured by Focal Technologies, Inc., Dartmouth, Nova Scotia, is a towed device capable of large scale, rapid and continuous counting and sizing of zooplankton, whilst also providing an indication of phytoplankton concentrations through a measure of light attenuation (Herman, 1988). The use of this device is well established in some areas of marine and aquatic research (Sprules *et al.*, 1992). In the OPC a collimated beam of light passes through an enclosed volume and is received by a sensor. When this beam of light is interrupted by a particle, the sensor signal produces a response proportional to the equivalent cross-sectional area of the particle. This is then digitised, and converted mathematically to an equivalent spherical diameter (ESD, Herman, 1992). A feedback system senses changes in water clarity from changes in overall response from the detector, and adjusts the output of the light emitter to maintain a constant level of illumination. A second photodetector, adjacent to the emitter, monitors these changes in the intensity of the emitted light, and the

signal from this monitor is used to indicate light attenuation. The data from the OPC is transmitted to a deck unit using frequency shift keying (FSK). The deck unit is connected to a personal computer using a standard RS-232 interface. Figure 3.1 is a schematic diagram illustrating the operating principle of the OPC.



*Fig. 3.1 Schematic diagram of the principle of operation of the OPC*

The OPC can produce reliable abundance and size distribution indicators for zooplankton between 0.25 and 20 mm. in diameter (ESD), at data rates up to 200 samples per second. Sensor response time and coincidence counts limit the densities at which the OPC can operate to 3000 counts per cubic metre.

### **3.2 The laboratory version - OPC-1L**

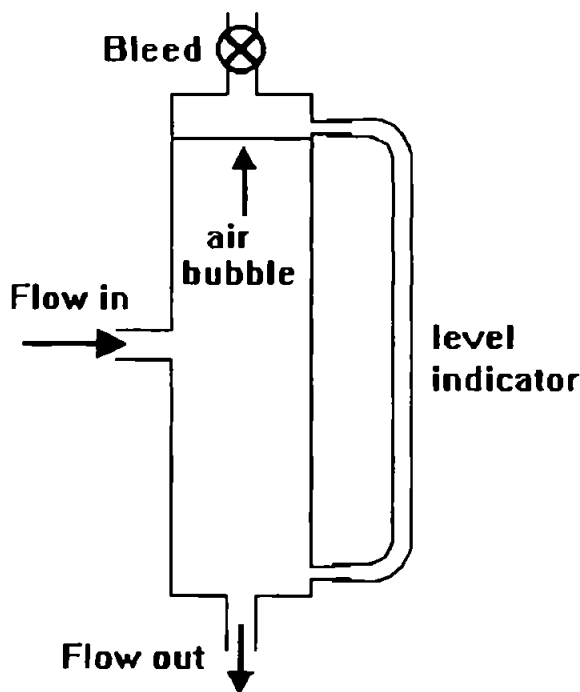
The laboratory version of the OPC (OPC-1L) is a bench-top version, similar in design and operation to the towed OPC, for processing preserved samples in pump-through mode. It was a novel application of this device which was used for underway zooplankton sampling aboard RRS James Clark Ross on the first three Atlantic Meridional Transect (AMT) cruises and the PRIME cruise (Robins *et al.*, 1996, Gallienne *et al.*, 1996, chapter 8). The only significant difference in operation is a smaller flow channel. The OPC-1L has a 20 mm. square section glass flow cell, through which water containing the sample is pumped. This may be a preserved sample, but in the present study the device is used at sea, using sea water pumped from beneath the hull of a survey vessel and the samples obtained from vertically integrated net hauls.

### **3.3 Operational issues**

A potential source of error with this system is due to the possibility of air bubbles being drawn in through the sea water intake. These will be detected by the OPC-1L as normal counts, although their transparency means that their size will be underestimated. The sample water was pumped through a de-bubbling device, before being passed through the OPC. At normal cruise speed for RRS James Clark Ross (11-12 kt., 5-6 m/s) the sea water intake was fully extended and bubbles were not present in the supply. However, in very shallow water, or at the higher speeds sometimes required by the ship's operational demands, the

intake had to be withdrawn, and air bubbles present at the hull surface layer were drawn into the supply.

At the flow rates used (*c.* 24 l min<sup>-1</sup>), the downward vertical velocity in the de-bubbling device was rather higher than the upward velocity due to buoyancy for some of the smaller bubbles (125 - 384 μm), and these passed through the OPC. This process was examined using a simple computer model, written by the author, of the de-bubbling device shown in figure 3.2. The downward velocity was calculated from the measured flow from the uncontaminated sea water supply and the cross-sectional area of the device.

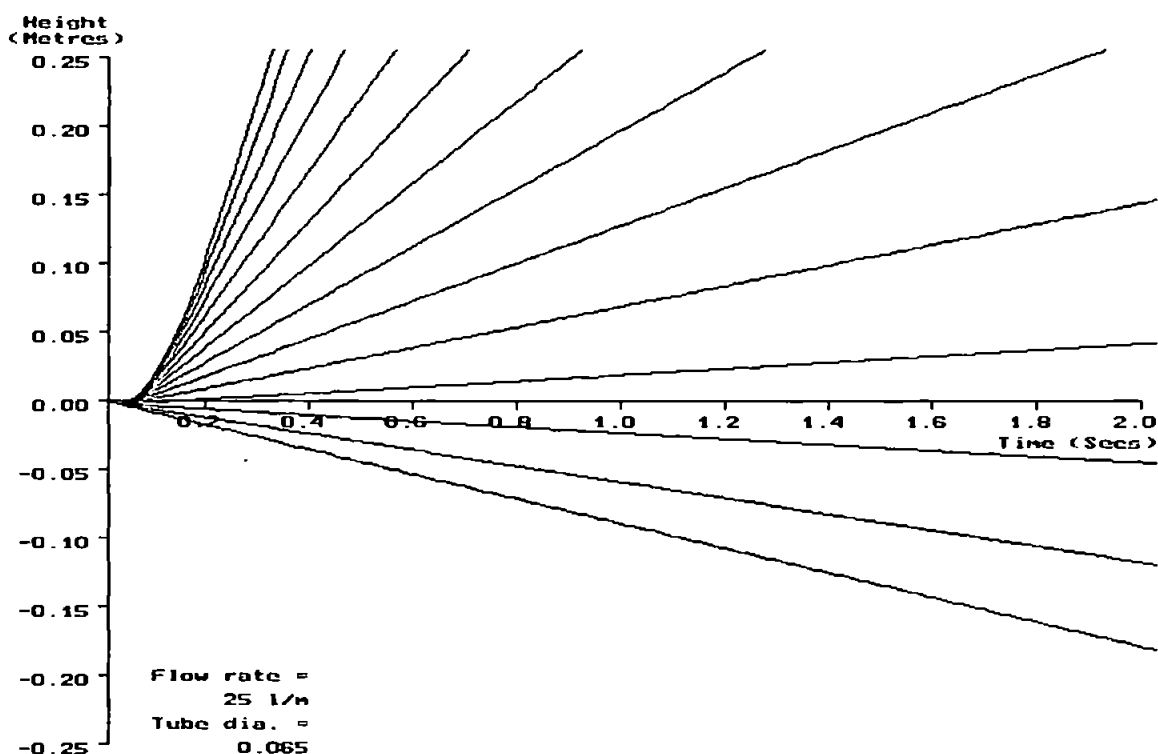


*Fig. 3.2 Schematic of de-bubbling device*

The upward velocity of air bubbles covering the size range of interest (125 μm - 11.314 mm) was calculated as a dynamic balance between the upward acceleration due to buoyancy, and the downward acceleration due to drag ( $F_D$ ).  $F_D$  is calculated from the increasing upward velocity relative to the downward flow. The drag coefficient,  $C_D$ , for a smooth sphere (0.35) was used in the formula:

$$F_D = \frac{1}{2} C_D u^2 A$$

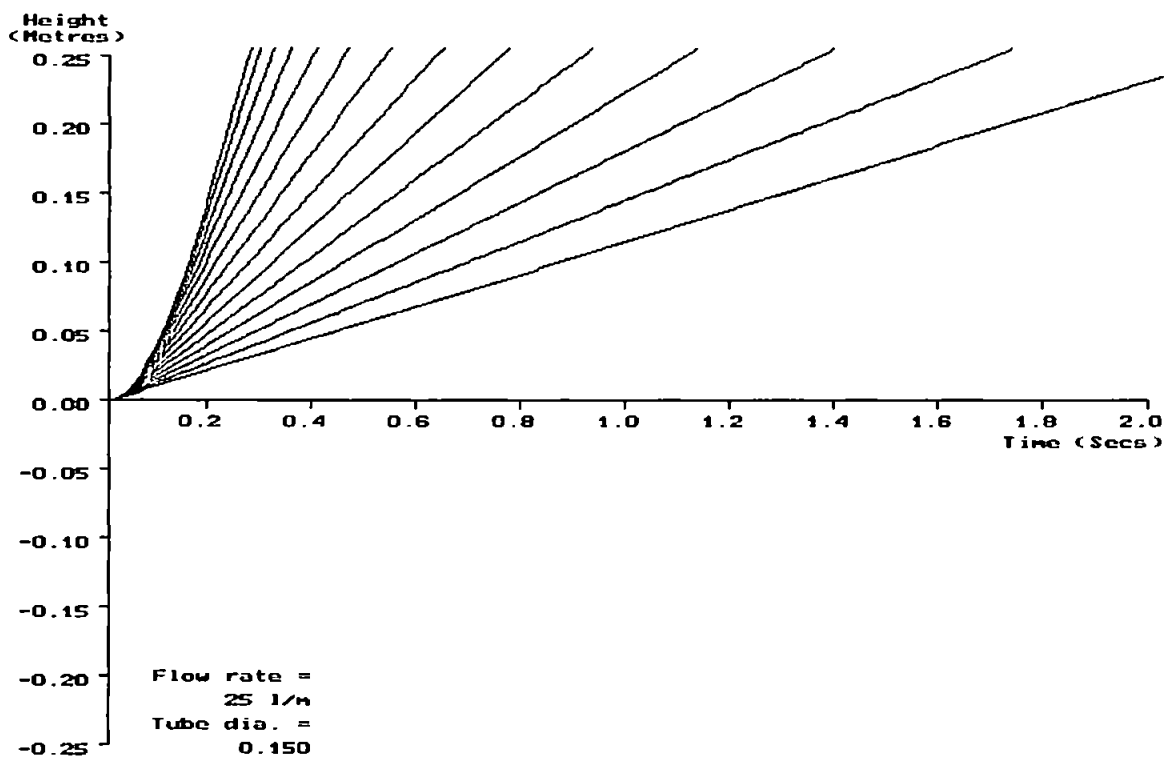
where  $A$  is the cross-sectional area of the sphere, and  $u$  is the net velocity of the bubble. The net effects of turbulence within the device were considered, for simplicity, to be zero. Figure 3.3a shows the predictions of the model for a tube of 65 mm internal diameter, as was used aboard RRS James Clark Ross on the first AMT cruise.



*Fig. 3.3 (a) Model prediction for 65 mm. diameter tube, bubble position as a function of time. Lines represent bubble size ranging from 125  $\mu\text{m}$  to 11.314 mm. in half  $\Phi$  intervals, bottom to top. Positive height represents rising bubbles, negative represents sinking.*

This predicts quite well what actually happened when bubbles were present, *i.e.* that bubbles in the three smallest size classes will be drawn down into the OPC. Figure 3.3b shows the prediction for a tube of 150 mm internal diameter, indicating that a modest increase in the cross-sectional diameter of this device

should solve this problem. Further increase in diameter above about 150 mm. resulted in no significant improvement in modelled de-bubbling performance, whilst such an increase involves a reduction in downward velocity of water below the level at which most zooplankton can swim, and therefore avoid being drawn through into the OPC. A new device of this diameter was designed and built by the author for subsequent cruises using the laboratory version of the OPC. The rate of de-bubbling for a given tube could, of course, be improved by reducing the flow rate. This option was rejected because a consequent reduction in sample volume would further reduce spatial resolution.



*Fig. 3.3 (b) Model prediction for 150 mm. diameter tube, bubble position as a function of time. Lines as in fig. 3.3(a).*

A consequence of the range of operational conditions encountered during the AMT cruises was that of condensation forming on the outside of the flow cell when sea water temperature was significantly below room temperature. This happened twice during the first AMT cruise, once as the vessel cleared the continental shelf off the Western Approaches, and again as the vessel met the colder waters at the confluence of the Brazil and Falkland currents off the coast of South America.

Whilst the optical feedback system could cope with the effect of mild 'fogging' in reducing the transmission of the optical path, the formation of beads of condensation might affect the OPC count. Particles must be in motion across the optical path at velocity between 0.5 and 4 m/s (Focal Technologies, 1992) in order to produce a response, so such condensation beads would not be counted as particles. The optical effects of introducing what are in effect small lenses into the optical path are unpredictable, however, and could degrade the performance of the instrument. The formation of condensation on the flow cell was detected by a steady rise in the value for light attenuation. If left to stabilise thermally with the sea-water running through the device, it was found that after about 5-10 minutes, the value for light attenuation would peak, and then gradually return to the previous normal value for sea water.

Apart from the two occasions mentioned, when the problem was detected and corrected, condensation was not a problem on AMT1. However, this effect could become acute in certain circumstances (*e.g.* deployment of the OPC-1L in

this mode in some parts of the North Atlantic), and it may be necessary to use an optically neutral anti-condensation compound, or some sort of air recirculation system around the flow cell.

At a flow rate of 25 litres per minute, the OPC requires 40 minutes to sample 1 cubic metre of sea water, a volume sufficient to obtain a reasonable size distribution of zooplankton in most ocean waters. At the cruising speed of *c.* 11 kt., this sample would be obtained over a distance of about 12 km. The larger size class (>2 mm) represent numerically only a very small fraction of the total count in some open ocean locations, however. In such locations, it may be necessary to sample 5 cu. m. or more, representing more than 60 km of the transect, to measure distributions of these size classes. In areas of higher biomass, where better resolution is required, numbers will be much higher, and the spatial resolution proportionally improved. The small sampling volume of the OPC-1L is a disadvantage, but this should be balanced by the simplicity and reliability with which thousands of miles of ocean sampling may be achieved using the system.

It is also important to be aware of the residual delay between the time the sample is taken at the intake, and when it reaches the OPC. For RRS James Clark Ross this delay was estimated to be insignificant at about one minute. The OPC-1L in underway mode samples at a constant depth, so that no information is available on vertical distributions. However, vertical plankton net casts once a day were used to obtain samples which were then passed through the device,

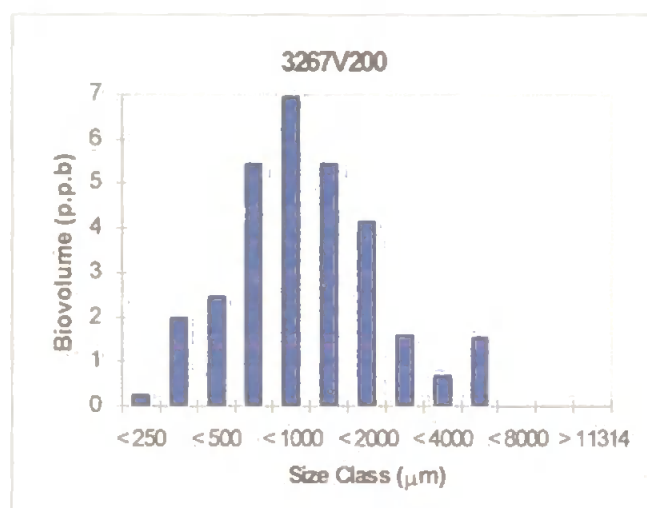
giving some measure of vertically integrated concentrations and size distributions.

### **3.4 Size distribution and community structure**

The OPC is claimed to give a good indication of the size distribution of the zooplankton community, initially binned to 4096 size classes from 250  $\mu\text{m}$  to 13 mm. Some investigators have cast doubts on the quantitative reliability of the counts produced for particles below about 500-700  $\mu\text{m}$  ESD due to the non-linearity of the relationship between ESD and digital size generated by the system at either extreme of the size range (Heath, 1995). For large numbers of particles, however, the statistical distribution seems reasonably reliable in these size classes, as shown by the calibration data given in chapter 8. The device gives no direct indication of species composition of the zooplankton passing through it. Given some prior knowledge of the species present, and species well separated in size, it may be possible to resolve some species structure (Herman, 1992). For large scale transects such as the present study, there are so many possible species at different developmental scales, that it is impossible to say anything about taxonomic structure using the OPC alone. This may be a disadvantage compared to the traditional sampling and analysis methods outlined above. The development of the video zooplankton analyser (ViZA) is intended to fill this gap, imaging the volume of water which flows through the OPC, analysing the images, and using pattern recognition techniques to give some idea of distribution by taxonomic group. The determination of size

distributions using the video device will also help in assessing the problem of reliability of size determination by the OPC at the extremes of the size range.

An important advantage of the OPC over traditional sampling methods for zooplankton is that results are available almost immediately.



*Fig. 3.4 Typical size distribution from raw OPC data.*

The OPC produces files of zooplankton data binned to 4096 size classes. These data were processed daily in spreadsheet form to produce size distributions in 14 size classes, shown in figure 3.4.

### 3.5 Conclusion

The loss of species information in the use of the OPC rather than traditional methods will be more than compensated by the advantages obtained. Thousands of kilometres of ocean can be simply and reliably surveyed for zooplankton abundance and size distribution without the need for prohibitively expensive subsequent human analysis of preserved samples. Not only are the data produced in real time, but sampling can be on the same temporal and spatial scale as the other environmental and physical parameters. The data integration period of the OPC may be chosen to match the sampling interval of the other

instruments used, down to the limits of the minimum sampling period of 0.5 seconds. This of course must be balanced against the need to maintain size spectrum resolution - *i.e.* to obtain statistically significant numbers in each of the size classes we wish to resolve. It has been shown that good correlation exists between OPC size spectra and those obtained by human analysis of net samples (Herman, 1992, Huntley *et al.*, 1995). Combined with the video device, this should allow a full description of abundance, size and taxonomic group distribution to be obtained simply and rapidly over large spatial and temporal scales.

# **CHAPTER 4**

## **OBJECT REPRESENTATION FOR CLASSIFICATION**

In order to classify objects within an image we must have some measure of the way in which they differ. The obvious characteristics from which to obtain this measure are size and shape, and these are the characteristics which are used in the present study. There are other candidate characteristics, such as colour and texture, but due to the limitations of underwater imaging and of processing constraints, these are not considered.

### **4.1 Describing size and shape in the spatial domain**

The first priority of the video imaging system is to obtain distributions of zooplankton according to size and major taxonomic group. The size parameter used in such distributions will be the 'equivalent spherical diameter' (ESD) obtained from the OPC (Herman, 1988). This measure is more realistically described as the equivalent cross-sectional diameter, as it is based upon the diameter of an equivalent circular cross section occulting the OPC light beam. The term ESD will be retained because direct comparison will be made between the results obtained from the two instruments. Other size parameters will be derived and used in the descriptions of shape described below. These size and shape descriptors may then be used to derive difference measures between taxonomic groups (classes).

The description of shape is a rather more difficult problem than that of size, and there is no generally accepted methodology (Sonka *et al.*, 1994). Many different approaches to this problem have been used, but the approach of Jeffries *et al.* (1984) has often been adopted in plankton taxonomy studies. In this work, some 41 size and shape descriptors were evaluated for their ability to discriminate between taxonomic group. Discriminant function analysis was used to identify those features most useful in the separation into groups of a set of some 265 training images of zooplankton. The smallest number of features consistent with minimal error in classification was chosen. It was found that five features were able to classify test samples under eight group headings, with an accuracy rate of over 89%. Berman (1990a), using similar methods, obtained a similar accuracy in classifying animals to 10 major groups, and required about 10 seconds per animal using a multi-processor system.

The representation of object shapes in an image can be done in one of two ways: by external characteristics (the object's boundary) or by internal characteristics (the pixels representing the internal region enclosed by the object boundary). Generally, the boundary representation is better suited to describing shape characteristics (Gonzalez and Woods, 1993), and is less demanding in terms of computer processing and storage. For a circular shape 100 pixels in diameter, the internal (region based) representation would require the processing of 7854 pixels ( $\pi r^2$ ), while the boundary representation requires the processing of only 314 pixels ( $2\pi r$ ). Furthermore, the boundary chain code representation is directly provided by the software driver of the frame processor used in the

present study. Boundary representations have therefore been used. The shape descriptors chosen should be invariant to size, translation, rotation and starting point along the boundary, as none of these parameters are predictable for freely moving objects in a three-dimensional flow.

#### *4.1.1 The boundary description*

The driver software for the frame processor card includes an applications programming interface (API) which includes a library of many image processing functions in the C programming language (see chapter 7). One such function extracts the contour (boundary) of an object in the binarised image. A contour pixel is a 'true' pixel which has at least one 'false' pixel touching it. The definition of true or false depends upon whether we are seeking a bright to dark transition (backlighting) or a dark to bright transition (front and side lighting, dark field illumination, etc.). The threshold at which the transition occurs must be determined for the device lighting conditions and object condition, and passed to the binarisation function as a parameter.

The contour is the set of such contour pixels, and the information is stored as a chain code of single bytes, each of which may take one of the values shown in table 4.1, below. These values represent the direction from the previous to the current boundary pixel in image coordinates.

This representation is an example of 8-connectivity chain code of the Freeman pattern (Freeman, 1961). Eight directions are allowed, *i.e.* the transition from

pixel  $(x_n, y_n)$  to  $(x_{n+1}, y_{n+1})$  is direct. In 4-connectivity, this transition must be made in two steps:  $(x_{n+1}, y_n)$ ,  $(x_{n+1}, y_{n+1})$ , and this boundary segment will be longer than in 8-connectivity.

**Table 4.1 Boundary chain code elements**

<i>Symbol</i>	<i>Value</i>	<i>X-Displ.</i>	<i>Y-Displ.</i>
DIR_STAY	0	0	0
DIR_E	1	1	0
DIR_NE	2	1	-1
DIR_N	3	0	-1
DIR_NW	4	-1	-1
DIR_W	5	-1	0
DIR_SW	6	-1	1
DIR_S	7	0	1
DIR_SE	8	1	1

The contour routine also allows the extraction of some basic descriptors of the object size and shape. These are: smallest enclosing box, perimeter length, area and centroid. All these features are measured in pixels, and must therefore be converted to absolute sizes by multiplication by the pixel size in the object plane.

Some parallelism in processing is allowed by the frame acquisition architecture.

The contour of the next object in the current frame may be obtained by the frame processor while the host processes the current contour, and if more than one memory frame buffer is used, the next frame capture may be initialised while the last frame is being processed.

#### 4.1.2 Size descriptors

The following size descriptors were evaluated for use in this study:

- perimeter length (P);
- area (A);
- equivalent spherical diameter (ESD);
- major chord length (MajCL);
- minor chord length (MinCL);
- fibre length (FL);
- fibre width (FW)

The fundamental size descriptor obtained from the contour routine is the boundary length. This feature is delivered as three 16 bit values: *nxdir*, *nydir* and *nxydir*. These are the number of direction values in the boundary chain code in the x (E/W), y (N/S) and xy (S/W, S/E, N/W, N/E) directions respectively. The sum of these values, adjusted for pixel size in the various directions, is the perimeter length.

The pixels of the camera used (CAM-17) are square, and the *pix\_size* variable represents the pixel size in the object plane, that is, the image sensor pixel size multiplied by the magnification of the imaging system. The perimeter length (P) may be obtained as:

$$P = (\text{pix\_size} * \text{nxdir}) + (\text{pix\_size} * \text{nydir}) + (1.41 * \text{pix\_size} * \text{nxydir})$$

The third term is necessary because the diagonal distance between adjacent square pixels is the square root of the sum of the squares of the horizontal/vertical distances.

The area is delivered simply as the number of pixels inclusively enclosed by the boundary. The absolute area is then obtained by multiplying this value by the area of a single pixel in the object plane. Note that the contour extraction algorithm extracts all boundaries in the image, including those formed by holes in the object of interest. This will be a problem with some zooplankton species with long appendages which overlap causing apparent 'holes' which the system will extract as separate entities and attempt to classify. The contour extraction algorithm offers a means to avoid this problem. If the contour extracted is such a 'hole' rather than an external boundary, the double word value representing the area of the object will be the binary negative of the number of enclosed pixels. Therefore, a simple test of the state of the most significant bit of this value (the sign bit in integer representation) will permit the rejection of all contours representing 'holes'.

The x and y values of the centroid of the object are obtained from *xarea* and *yarea*, the x and y moments about the respective axes, supplied by the contour routine.

$$\text{X-centroid} = \text{xarea} * \text{pix\_size} / \text{area}$$

$$\text{Y-centroid} = \text{yarea} * \text{pix\_size} / \text{area}$$

The fourth feature supplied by the contour routine, the smallest enclosing box, is not very useful in the present application, and is not used. This is because it is a box in image coordinates, and its dimensions are therefore sensitive to the orientation of the object. This feature is, however, useful in determining the

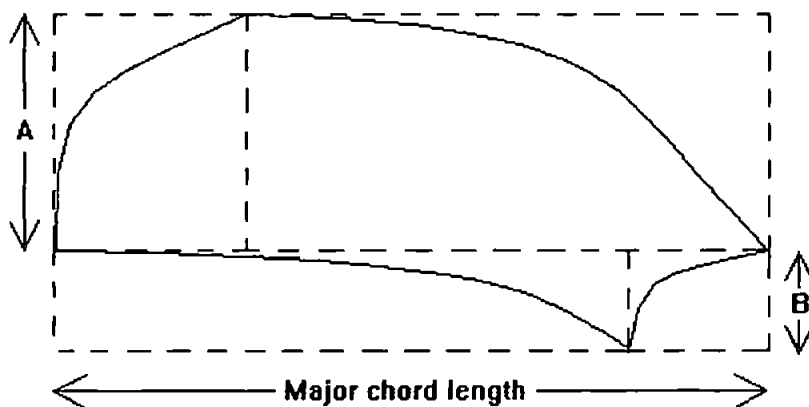
total screen area required to display the contour, and is therefore retained for use in the display section of the host program.

The equivalent spherical diameter used by Herman (1992) is simply the diameter of a circle of the same projected area as the object. It is obtained from:

$$\text{equiv. dia.} = \sqrt{[(4/\pi) * \text{area}]}.$$

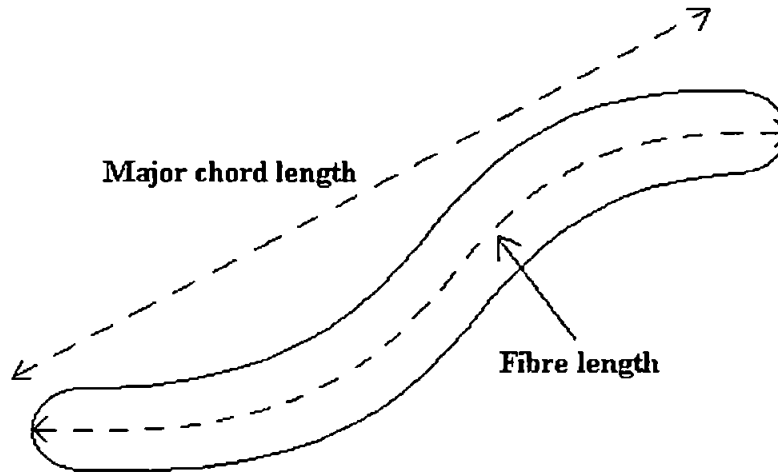
Two other size parameters are derived from the object boundary chain code: major chord length and minor chord length. The major chord length is obtained by searching for the longest distance between any two boundary pixels. This search can be considerably reduced by superimposing the circle of equivalent spherical diameter on the CG of the object, and searching only those pixels which lie on or outside this circle.

The minor chord length is usually the minor chord of the circumscribed ellipse. In the present study, however, this measure is the sum of the maximum perpendicular distances from either side of the major chord to the boundary.



*Fig. 4.1 Major chord length and minor chord length (A+B) of an object.*

The MajCL and MinCL therefore gives the dimensions of a rotationally invariant smallest enclosing rectangle. The relationship between these three variables is illustrated in fig. 4.1, above. The fibre length is the length of the 'centreline' of the object, as shown in fig 4.2. This is an important parameter for obtaining the true length of an object which may curl.



*Fig. 4.2 Comparison between major chord length and fibre length of an object.*

The fibre width is the average width of such an object. Given the assumption of uniform width along the length, the fibre length (FL) and fibre width (FW) may be obtained from:

$$FW = P - \frac{\sqrt{P^2 - 16 * \text{area}}}{4} \quad FL = \text{area} / FW$$

where P = perimeter length. (Russ, 1995)

#### *4.1.3 Shape descriptors*

Shape descriptors should be invariant to size (magnification), translation, rotation, and starting point along the boundary. For example, shape descriptors

based upon the perimeter length of a given object will be dependent upon the size/magnification of the object within the image. As magnification increases for a given object, more pixels are used to represent each feature of the boundary, more detail emerges, and noise and other imaging artefacts affect the boundary to a greater extent and the length of this boundary segment will increase - this is the 'fractal effect'. Increase of magnification will also reduce the quantisation distortion caused by the imaging system, however. There are many pre-processing procedures which may be applied to the boundary representation to compensate for size/magnification, some of which are outlined below:

1. The boundary chain code may be normalised to contain a standard number of boundary segments. The Fast Fourier transform (FFT) algorithm used to obtain the Fourier descriptors (see sect. 4.2, below) requires an input array size which is some power of two in any case, so some normalisation is required. This not only compensates for size/magnification variations, but also makes for easier signature generation (see below).
2. Gonzalez and Woods (1993) suggest resampling to a standard grid size to avoid the problem. This involves projecting the outline of the object to a standard grid, which it just fills. The solution adopted in this study is something of a combination of these two methods (see section 4.2.3).

Note that by definition, these processes remove size from the description - size must be stored separately if this descriptor is to be used in the pattern

recognition algorithms. In the present case, the ESD gives a direct measure of size.

The shape descriptors evaluated for use in this study are:

bounding rectangle (indirectly);

aspect ratio;

extent;

form factor;

roundness;

compactness;

Fourier descriptors of the signature;

Moments of area have also been used as shape descriptors for a given object (Hu, 1962, Reiss, 1991), but these are region based descriptors (based upon the whole area of an object), and the present study relies only upon boundary based descriptors. Although these moment invariants may be derived from the Fourier descriptors of the boundary signature (Russ, 1995), it is assumed that the information carried by them will therefore already be present in the Fourier descriptors and further exposition of this information is unnecessary.

As stated above, the bounding rectangle is defined by the major and minor chord lengths. The aspect ratio is found as the ratio of these two lengths:

aspect ratio = major chord length / minor chord length

Qualitatively, this feature expresses the “elongatedness” of the object.

Extent is given as:

$$\text{extent} = \text{area} / \text{bounding rectangle area}$$

Qualitatively, this feature expresses the degree to which the object fills the bounding rectangle.

Form factor is given as:

$$\frac{4 * \pi * \text{area}}{p^2}$$

Qualitatively, this feature expresses the degree of “raggedness” of the boundary.

Roundness is given as:

$$\frac{4 * \text{area}}{\pi * \text{ESD}^2}$$

Qualitatively, this feature expresses the circularity of the object - a circle will have a roundness of 1. The roundness of any other shape will be less than 1.

Compactness is defined as:

$$[\sqrt{(4 / \pi) * \text{Area}}] / \text{MajCL}$$

This feature is similar to roundness, using major chord length instead of diameter. It also shows similarities to aspect ratio, being a measure of “elongatedness” from an ideal circle.

### *Signature*

The signature of an object outline is a function of the shape of the object, as follows. In fig 4.3a an arbitrary object outline is shown.

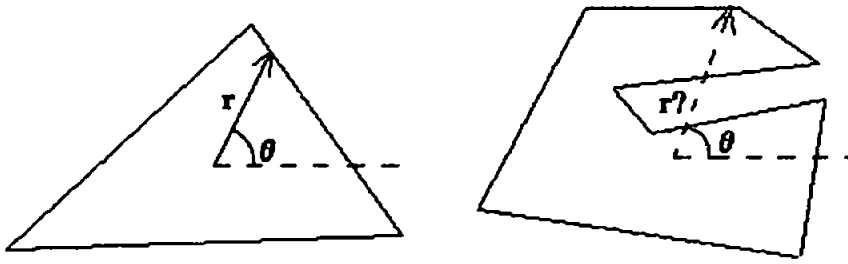


Fig. 4.3 (a)  $r(\theta)$  defined for all  $\theta$  (b)  $r(\theta)$  undefined for some values of  $\theta$

If the distance from the centre of area of the object outline to the perimeter at any point is denoted  $r$  and the angle of the line from the centre to this point is denoted  $\theta$ , then a function  $r(\theta)$  may be derived from the boundary where  $r$  is the length as a function of angle. If the boundary length is standardised to  $2\pi$ , this function may be considered to be periodic, and may be represented as a Fourier series (see section 4.2). Fig. 4.3b shows a case where the above algorithm will present a problem - if the boundary is re-entrant, the distance  $r$  will be undefined for certain angles.

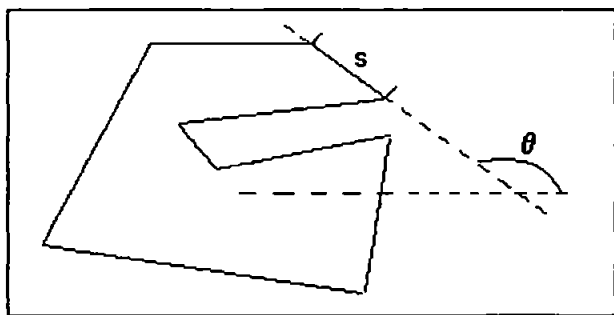


Fig. 4.4 Slope density function,  $\theta(s)$ , i.e. the slope ( $\theta$ ) of all boundary segments ( $s$ ).

A function  $\theta(s)$  which obtains the angle of the boundary segment ( $\theta$ ) at fixed boundary segment intervals ( $s$ ) will solve this problem, as shown in fig. 4.4.

The boundary chain code implicitly contains the slope versus link length information, so that the Fourier transform of the link length versus the boundary angle ( $\theta(s)$ , the slope density function) will yield the Fourier series.

#### **4.2 Shape description in the frequency domain - Fourier descriptors:**

Use of Fourier techniques in describing shape is not new (Zahn & Roskies, 1972, Richard & Hemami, 1974, Persoon & Fu, 1977). Only in the last ten to fifteen years has sufficient computing power been available in the average laboratory to make their use practical.

##### *4.2.1 Fourier's theorem*

Fourier's theorem states that any continuous periodic function can be represented as a discrete series of sine and cosine functions at the fundamental and all harmonic frequencies. For the function  $r(\theta)$  above, the Fourier series is:

$$r(\theta) = a_0 + a_1\cos(\theta) + b_1\sin(\theta) + a_2\cos(\theta) + b_2\sin(\theta) + \dots + a_n\cos(\theta) + b_n\sin(\theta). \quad (1)$$

The coefficients  $a_n$ ,  $b_n$  are the amplitudes for each harmonic  $n$  of the Fourier series. The coefficient  $a_0$  represents the zero frequency component, or the average value (the DC component in electrical network theory, and a measure of size in the current context).

#### 4.2.2 The Fourier transform

The Fourier transform produces from a discrete function the continuous frequency distribution  $F(u)$  which make up that function:

$$F(u) = \int_{-\infty}^{+\infty} f(x) e^{-i2\pi ux} dx$$

The exponential term in the equation appears in Euler's formula, which explains the representation of the Fourier components as a sine and cosine series:

$$e^{-i2\pi ux} = \cos 2\pi ux - i \sin 2\pi ux$$

The Fourier transform is reversible. Given the Fourier series of a function, the original function may be restored by the reverse Fourier transform

$$f(x) = \int_{-\infty}^{+\infty} F(u) e^{i2\pi ux} du$$

When the function being transformed is a discretely sampled function, as is usually the case, the infinite Fourier integral is limited by the sampling period and the sampling interval, and is replaced by the discrete Fourier transform (DFT) pair:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-i2\pi ux/N}$$

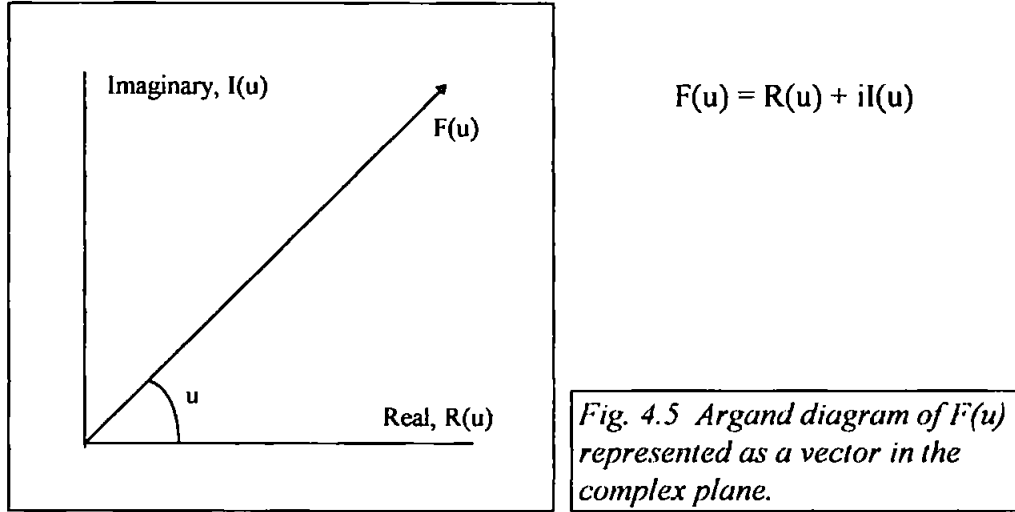
$$f(x) = \sum_{u=0}^{N-1} F(u) e^{i2\pi ux/N}$$

The DFT is usually implemented in computer systems as the fast Fourier transform (FFT) which is rather more computationally efficient.

The Fourier transform of a real function is a complex function:

$$F(u) = |F(u)| e^{i\phi(u)}$$

where  $|F(u)|$  is the magnitude part, and the exponential term represents the phase information for the series. Each value  $F(u)$  is a complex number which may be represented in an Argand diagram as shown in fig. 4.5.



$R, I$  are the real and imaginary parts, and  $i$  is the complex operator. The magnitude part of the transform  $|F(u)|$  is the frequency spectrum,

$$|F(u)| = (R^2(u) + I^2(u))^{1/2}$$

which function ignores phase. The power spectrum is simply the square of this magnitude. If the phase information is discarded then the series coefficients ( $a_n, b_n$ ) may also be represented as a vector resultant series,  $c_n$ , where:

$$c_i = \text{sqrt}(a_i^2 + b_i^2)$$

#### 4.2.3 Fourier shape descriptors

In the light of the need for standard boundary lengths and the need for some function of boundary angle versus segment length whose Fourier components are rotation, translation, magnification and starting point invariant the following representation was adopted. As a method of boundary normalisation, the

boundary was divided into 256 segments of equal length (to avoid the 'fractal effect' mentioned above). Beginning at an arbitrary point along the boundary, the values  $dx$  and  $dy$  were obtained for the segment. Given a segment length ( $s$ ) which is a constant, these values fully describe the cosine and sine respectively of the boundary angle for this segment. A one-dimensional complex array of these values was formed:

$$f(s) = x(s) + iy(s)$$

where  $i$  is the complex operator, and which function is periodic with period  $L$ . For each element of the array the  $x$  axis may be considered to be real and the  $y$  axis imaginary. The information contained by this description is the same as that contained in the functions  $r(\theta)$  and  $\theta(s)$ , but the advantage of this representation is that it reduces a two-dimensional problem to one dimension.

This function is equivalent to the familiar derivation of the sine and cosine functions from the complex exponential description of a function as  $\exp[-iu\theta]$ , using Euler's formula (Persoon & Fu, 1977). The complex FFT of this complex array returns a set of complex Fourier descriptors for the boundary.

$$f(x + iy) \Leftrightarrow F(u + iv)$$

For the present problem, consider coordinate pairs  $(x_s, y_s)$  whose sum over the boundary length  $L$  is  $(x_L, y_L)$ . Each element of this pair, taken over the boundary length, results in a repetitive function of period  $L$ . This representation is less sensitive to noise due to imaging imperfections and boundary digitisation (Richard & Hemami, 1974). The complex Fourier transform of these two

vectors, represented as a complex vector  $z(s) = x(s) + iy(s)$ , yields the Fourier coefficients  $a_n$  and  $b_n$ :

$$a_n = \frac{1}{L} \sum_{s=1}^{L-1} x_s e^{-i(2\pi/L-1)us/L}$$

$$b_n = \frac{1}{L} \sum_{s=0}^{L-1} y_s e^{-i(2\pi/L-1)us/L}$$

(Sonka *et al.*, 1993).

Rotational information is stored in the phase information of this complex series. Using polar coordinates to represent the function allows a simple illustration of the effects of rotation upon the boundary function and its Fourier transform:

$$x = r \cos \theta \quad y = r \sin \theta \quad z = r(\cos \theta + i \sin \theta)$$

$$u = r' \cos \varphi \quad v = r' \sin \varphi \quad w = r'(\cos \varphi + i \sin \varphi)$$

The exponential representations of these functions are:

$$z = re^{i\theta} \quad w = r'e^{i\varphi}$$

For a rotation of the boundary function through an angle  $\alpha$ , the boundary function becomes:

$$\begin{aligned} z &= r(\cos \theta + i \sin \theta)(\cos \alpha + i \sin \alpha) \\ &= r(\cos [\theta+\alpha] + i \sin [\theta+\alpha]) \end{aligned}$$

Similarly, the Fourier magnitudes are:

$$w = r'(\cos [\varphi+\alpha] + i \sin [\varphi+\alpha])$$

The Fourier transform pair can be represented as:

$$f(r, \theta + \alpha) \Leftrightarrow F(r', \varphi + \alpha)$$

Rotation of a point in the complex plane about the origin by an angle  $\alpha$  is equivalent to multiplication of the point by  $(\cos \alpha + i \sin \alpha)$ , or  $e^{i\alpha}$  (Stephenson,

1973). When an object outline is rotated, each point in the complex plane is multiplied by this same factor,  $f(s)$  becoming  $f(s) e^{i\alpha}$ . Rotating  $f(x + iy)$  by  $\alpha$  rotates each of the Fourier components  $F(u + iv)$  by the same amount (Gonzales & Woods, 1993). The effect on the Fourier coefficients is:

$$a_r(u) = a(u)e^{i\alpha}$$

The rotation of the boundary therefore results in a rotation of the complex Fourier components,  $F(u + iv)$ , *i.e.* a phase shift of the resultant vector  $w$  between  $v$  and  $u$ , the real and imaginary components of the series  $F(u + iv)$ .

A shift in the starting point along the boundary by  $s_0$  simply results in the new signature function:

$$f(s - s_0) = x(s - s_0) + iy(s - s_0)$$

The effect on the Fourier coefficients is:

$$a_p(u) = a(u)e^{-i2\pi s_0 u/L}$$

This is similar to the effects of rotation, being a shift in phase between the real and the imaginary axes in the complex plane, except that in this case all descriptors are affected in a different way, dependent upon the frequency ( $u$ ) of the component (Gonzales & Woods, 1993). By taking the vector resultant of the real and imaginary parts of the complex array of Fourier components

$$|F(u,v)| = (R^2(u,v) + I^2(u,v))^{1/2}$$

a set of Fourier descriptors for the boundary shape is obtained, which form the magnitude spectrum of the boundary. The phase information, which contains information about rotation and starting point, has been discarded. The magnitude spectrum of the boundary function  $f(s) = x(s) + iy(s)$  is therefore

invariant to rotation and to starting point along the boundary (Richard & Hemami, 1974, Sonka *et al.*, 1994). The normalisation of the boundary to segments of constant length means that the resultant descriptors are size invariant. The coordinate system is self referential, so they are also translation invariant.

The boundary functions  $r(\theta)$  and  $\theta(s)$  almost always contain discontinuities (Pavlidis, 1980), *e.g.* the sudden switch in boundary direction between  $-180^\circ$  and  $+180^\circ$ . This means that the Fourier coefficients  $a_n$  will decrease only slowly with  $1/n$ . For a piecewise continuous smooth function such as  $f(s)$ , the magnitude of the Fourier descriptors decrease as  $1/n^2$ , so that most of the shape information lies in the first few harmonics (Richard & Hemami, 1974). In practice it is usually found that the first 10 - 25 terms are sufficient adequately to describe the signature function, and therefore the shape of the object. These components may then be compared between samples using, for example, principal component analysis (chapter 5), to classify objects very successfully (Russ, 1995).

# CHAPTER 5

## TAXONOMIC CLASSIFICATION PROCESSES

There is no unifying theory of pattern recognition and classification; those theories extant tend to be problem specific (Therrien, 1989). Patterns are generally a function of  $n$  variables plotted in  $n$  dimensional space. Similar patterns tend to cluster together in such space, and may be classified as similar objects. The problem is then one of using some discriminant function  $F(x)$  to find  $(n - 1)$ -dimensional 'hyperplanes' which effectively partition this space into the correct classes. These functions are difference measures between classes. Often the number of classes is unknown, and the definitions of the classes is unknown. The problem is to define the classes and to classify the data.

### 5.1 Identifying species

The identification of species by image analysis in real-time is not a simple task, given the constraints of processing time and current technology available.

Descriptors of size and shape only, as obtained from a digitised image of an object in a turbid medium, are not sufficient to differentiate all species which we may come across. Biologists often have to resort to counting and examining individual legs on a given subject in order to classify it correctly. In order to make the task manageable, and in the first instance, it will be considered sufficient to classify organisms under 8-10 major taxonomic groupings, to be chosen.

### *5.1.1 The data processing problem*

The amount of data produced by a video digitiser is very large - typically around 15 megabytes per second. Furthermore, the operations required by pattern recognition algorithms are computationally intensive. In order to have any chance of processing such large amounts of information in real time, the data set and the operator set must both be stripped to the minimum consistent with achieving the declared aims. Furthermore, the algorithms and code used to implement the operator set must be as efficient as possible. Some methods of data and operator set reduction are discussed below. The problem of algorithm and code efficiency are discussed in chapter 7.

### *5.1.2 Reducing the data set*

The first step in reducing the data set is to produce from the image a set of boundary descriptors of each object in the image. If no object is present the image frame is simply discarded and the next frame acquisition initiated. Dealing with a boundary descriptor chain, as described in chapter 4, rather than a region, reduces the data set. For a typical object, say 64 pixels ESD, the data is reduced from about 3000 bytes to about 200 bytes. Boundary descriptors above or below certain maximum and minimum size thresholds will also be discarded without further processing, as being outside the size range of the pattern recognition system. (Note: objects smaller than the minimum threshold may still have the ESD logged, as they can contribute to the size distribution statistics, though not to the species distribution).

Two further methods of reducing the data set are available. Firstly, at typical oceanic concentrations, and given the small imaged volumes of each video frame, perhaps only one in twenty video frames may contain an object of interest. To save on the processing overhead of examining many such empty frames for objects, a system of multiple exposures of single frames, using the strobed illumination, may be used to reduce this overhead. At planned towing speeds of 5 m/s, and a field of view of 40 mm., the time between exposures could usefully be reduced to 8 milliseconds (the time required for the field of view to move to a completely new area). This reduces the number of frames to be processed by more than 50% (frame period = 16.67 ms). For a smaller field of view, this data reduction could be proportionally improved. The second method of data reduction is to have a line of photodetectors at the leading edge of the video field of view. If a specified minimum of these diodes were occluded (proportional to the minimum object size of interest), an interrupt to the frame processor could be generated (suitably delayed) instructing it to grab a frame. This would avoid the computational overhead of checking an empty frame for an object of interest, whilst such an object passed through the system unrecorded. Neither method has been incorporated in the system prototype, although either may later be incorporated if it should be considered necessary.

### *5.1.3 Reducing the operator set*

The size and shape descriptors listed in chapter 4 are a fairly comprehensive, and in many cases redundant set of descriptors. Many of these descriptors are describing similar features of a given object, and those which are not may be

correlated in some unknown way. For either of these reasons, there may be considerable redundancy in these descriptors, and the problem is to determine what minimum subset can be usefully retained, and what weights to assign to each descriptor to differentiate optimally between the set of species determined to be of interest. It is impossible by inspection to determine, *a priori*, which of the descriptors may be useful to discriminate between a given set of shapes. The problem is usefully approached, however, by the techniques of multivariate analysis, as described in the next section.

## **5.2 Multivariate analysis**

Standard statistical techniques use univariate analysis, measuring the variation of a single dependant variable against one or more independent variables.

Where several related dependant variables must be analysed, multivariate techniques are used. There are many techniques of multivariate analysis, and little general agreement upon which are the best to use in any given situation.

The choice is largely data dependant. Those methods of particular use to biologists, and those interested in classification, may be grouped under the heading of "ordination methods". This term describes a

"process of producing a small number of variables ... to describe the relationship between a group of objects starting ... from the values of a large number of variables measured on each object." (Manly, 1994).

The first step in most methods of multivariate analysis is to obtain some measure of the difference between objects, in terms of some function or

functions of the variables upon which those differences depend. In the current application, those variables are measures of size and shape. The most commonly used measure of distance between quantitative variables is the Euclidean distance ( $\delta$ ). An extension of Pythagoras' theorem from 2 to  $p$  dimensions shows that, given  $n$  objects and  $p$  orthogonal variables, the distance apart of any two objects  $i, j$  for variables  $k = 1..p$  is:

$$\delta_{ij} = \sqrt{\sum_{k=1..p} [x_{ik} - x_{jk}]^2}$$

If one variable is numerically much larger than the others then absolute variations in this variable will tend to dominate  $\delta_{ij}$  and swamp the influence of the other variables. The variables are usually scaled to compensate for this effect. Typically, each variable value is scaled by subtracting the mean value of the variable, and dividing by the standard deviation, producing a set of variables of zero mean and unit standard deviation. This gives equal importance in the analysis to all variables.

### *5.2.1 Cluster analysis*

Many clustering algorithms exist, mostly heuristic, and none is really objective, being data dependent. Methods are generally either hierarchical or partitional, hierarchical methods being popular in biology, where data represent a taxonomy (Jain, 1986). The following is a hierarchical method.

Given  $n$  object types  $x_1, x_2 \dots x_n$  contained in object space  $s$  we seek regions  $s_1, s_2, \dots s_k$  such that every  $x_i$  falls into one of these regions and none falls into two such regions. In this way we can find a repeatable set of groups which can be

found amongst our data, dependent upon the variables by which we choose to classify these data. The difference measures are arranged in a triangular matrix and objects are grouped together at each increasing level of difference.

*Clustering algorithm (Jain, 1986):*

1. Assign each object to a cluster of one;
2. Find smallest entry in difference matrix and merge these two clusters;
3. Update difference matrix to reflect this new grouping;
4. Repeat from step 2 until single group emerges;

The results are often displayed as a hierarchical diagram called a dendrogram.

The best grouping is then chosen (subjectively) where the difference distance between groups is sufficient clearly to demarcate them, i.e. the set of compact, convex clusters that can *reliably* be identified in a given data set.

Note that clustering algorithms *always* find clusters, even if none exist naturally in the data. The problem is then to determine whether the clustering is 'real', or merely an artefact of the algorithm used. Some external criterion must be used to validate the clustering results. Jain (1986) suggests that for higher than two-dimensional data tests are sparse - comparison of data to a random convex distribution is suggested.

The success of this grouping strategy is largely dependent upon the variables chosen to differentiate objects. The set of such variables may be chosen by many different methods (Mucciardi & Gose, 1971), such as iteration of the above procedure with different sets. Alternatively, another method may be used to find the best subset of variables, given a provisional group set obtained as

above. The clustering procedure may then be repeated using this set of variables, either confirming the groupings chosen, or suggesting a different optimal grouping strategy. This process should converge on a best fit solution. The method of principal component analysis has been used to derive such a subset of variables in this study.

### *5.2.2 Principal component analysis*

The multivariate ordination method described here is that of principal component analysis (PCA). The objective of PCA is to take  $p$  variables  $x_1, x_2, \dots, x_p$  measured on a set of objects, and find linear combinations of these variables (principal components), using coefficients  $z_1, z_2, \dots, z_p$ . These components should be uncorrelated (orthogonal) and measure different 'dimensions' in the data. These linear combinations are then used to discriminate between different classes of objects. PCA is used when quantitative data on  $p$  variables is available for *each* of the objects, as in the current application.

The first principal component,  $P_1$ , is responsible for the greatest variation in data, and so on -  $P_1(\text{var}) \gg P_2(\text{var}) \gg \dots \gg P_p(\text{var})$ .  $z_i$  are the coefficients of the principal components (PCs). If a significant proportion of the total variance in the data can be accounted for by few PCs, then there is considerable redundancy in the set of descriptor variables. Redundancy in variables which may be measuring similar features means that data are often adequately described by few such components, each containing only a weighted subset of the original  $p$  variables, permitting the operator reduction required. If the number of PCs can

be reduced to two or three then the data is considered to be 'well ordinated', and ordination axes in two or three dimensions will allow the data to be plotted for easier visual presentation.

The first PC is a linear combination of the variables, e.g.:

$$P_1 = z_{11}x_1 + z_{12}x_2 \dots + z_{1p}x_p$$

that varies as much as possible between individuals, subject to:

$$z_{11}^2 + z_{12}^2 \dots + z_{1p}^2 = 1$$

Similarly,  $P_2 = z_{21}x_1 + z_{22}x_2 \dots + z_{2p}x_p$

that varies as much as possible between individuals, subject to:

$$z_{21}^2 + z_{22}^2 \dots + z_{2p}^2 = 1$$

and  $P_1$  and  $P_2$  are uncorrelated, and so on, up to a maximum of  $p$  principal components. The coefficients  $z_{ij}$  of each PC are the eigenvectors of the covariance matrix formed from the matrix of differences between individuals. The variances due to each PC are the eigenvalues of the covariance matrix, and add up to the total variance between all individuals.

*PCA Algorithm:*

1. Normalise variables  $x_n$  to have zero mean and unit variance;
2. Calculate covariance matrix (the correlation matrix, given [1]);
3. Find eigenvalues  $l_n$  and corresponding eigenvectors  $a_n$ ;
4. The coefficients of the  $i$ th PC are  $a_i$ , variance is  $l_i$ ;
5. Discard any component whose contribution to the overall variance is considered negligible;
6. Descriptors whose coefficient in all PCs are consistently low may also be discarded in all PCs, as being of negligible effect.

### *PCA Example*

There follows an illustrative example of a principal component analysis, using a small data set obtained using the ViZA system. A set of 15 zooplankton silhouettes was imaged and processed by the system, producing the full set of size and shape descriptors. A sub-set of just three descriptors is used for the present example, for simplicity. Table 5.1 lists these descriptor values for each animal in arbitrary units (actually in pixels).

**Table 5.1 A sample set of size descriptors.**

	<b>Perimeter</b>	<b>Area</b>	<b>ESD</b>
<b>Sagitta 1</b>	1020	17628	149
<b>Sagitta 2</b>	925	11918	123
<b>Sagitta 3</b>	884	10243	114
<b>Sagitta 4</b>	913	11087	118
<b>Copepod 1</b>	277	1758	47
<b>Copepod 2</b>	346	1593	45
<b>Copepod 3</b>	306	1914	49
<b>Copepod 4</b>	470	4807	78
<b>Copepod 5</b>	613	4144	72
<b>Fish egg 1</b>	232	3953	70
<b>Fish egg 2</b>	229	3841	69
<b>Fish larva 1</b>	367	3758	69
<b>Fish larva 2</b>	331	3816	69
<b>Euphausiid 1</b>	1925	10829	117
<b>Euphausiid 2</b>	2904	22086	167
<b>Mean</b>	783	7558	90
<b>Std. Dev.</b>	742	6195	38

Values are normalised to zero mean and unit standard deviation by subtracting the mean and dividing by the standard deviation for each value. Table 5.2 shows the results of this.

**Table 5.2 Normalised set of size descriptors.**

	<b>Perimeter</b>	<b>Area</b>	<b>ESD</b>
<b>Sagitta 1</b>	0.32	1.625	1.538
<b>Sagitta 2</b>	0.192	0.704	0.855
<b>Sagitta 3</b>	0.137	0.433	0.619
<b>Sagitta 4</b>	0.176	0.569	0.724
<b>Copepod 1</b>	-0.682	-0.936	-1.139
<b>Copepod 2</b>	-0.589	-0.963	-1.191
<b>Copepod 3</b>	-0.643	-0.911	-1.086
<b>Copepod 4</b>	-0.422	-0.444	-0.325
<b>Copepod 5</b>	-0.229	-0.551	-0.483
<b>Fish egg 1</b>	-0.743	-0.582	-0.535
<b>Fish egg 2</b>	-0.747	-0.600	-0.562
<b>Fish larva 1</b>	-0.561	-0.613	-0.562
<b>Fish larva 2</b>	-0.609	-0.604	-0.562
<b>Euphausiid 1</b>	1.539	0.528	0.698
<b>Euphausiid 2</b>	2.862	2.345	2.010
<b>Mean</b>	0	0	0
<b>Std. Dev.</b>	1	1	1

From this normalised matrix, the covariance matrix was obtained, shown in table 5.3. Given this normalisation, this is also the correlation matrix for the three variables.

**Table 5.3 Covariance matrix**

	<b>Column 1</b>	<b>Column 2</b>	<b>Column 3</b>
<b>Column 1</b>	1	0.859645	0.82817
<b>Column 2</b>	0.859645	1	0.986386
<b>Column 3</b>	0.82817	0.986386	1

Obviously, the correlation between any variable and itself is 1. It may also be noted that Area is very highly correlated with ESD, not surprising since the one is derived from the other using a simple formula. Finally, the eigenvectors and eigenvalues of this covariance matrix are obtained using an algorithm based

upon the Jacobi transformation (Press *et al.*, 1992), giving us three principal components, and the variance in the data due to each. Table 5.4a and b show the results of this analysis.

**Table 5.4a Eigenvectors of the covariance matrix.**

Column 1	Column 2	Column 3
0.8271	0.0830	0.5559
-0.3230	-0.7392	0.5909
-0.4600	0.6683	0.5846

**Table 5.4b Eigenvalues of the covariance matrix.**

0.2036	0.0118	2.7846
--------	--------	--------

The columns of table 5.4a represent the coefficients of the three principal components. The total variance is 1 x the number of variables when normalised data are used - three in this case. The component which accounts for the most variance in the data is in column three. This component represents 2.7846/3 or 93% of the variance, and is given as:

$$PC1 = 0.5559 * \text{Perimeter} + 0.5909 * \text{Area} + 0.5846 * \text{ESD}$$

It is often possible to interpret components as a measure of some particular feature of the data set, although just as often it is not. In the present simple case, the first component gives approximately equal positive weights to each measure of size. It therefore represents a general measure of the size of the object.

Principal component 2 represents 0.2036/3 = 6.7% of the total variance, and is given as:

$$PC2 = 0.8271 * \text{Perimeter} - 0.323 * \text{Area} - 0.46 * \text{ESD}$$

As both area and ESD are representing a similar measure of area, and are given a negative weighting, whilst the perimeter length is given a strong positive weighting, this component represents a measure of boundary length against size, i.e. compactness or circularity. For a given size, objects with a more ragged outline will be discriminated by this component.

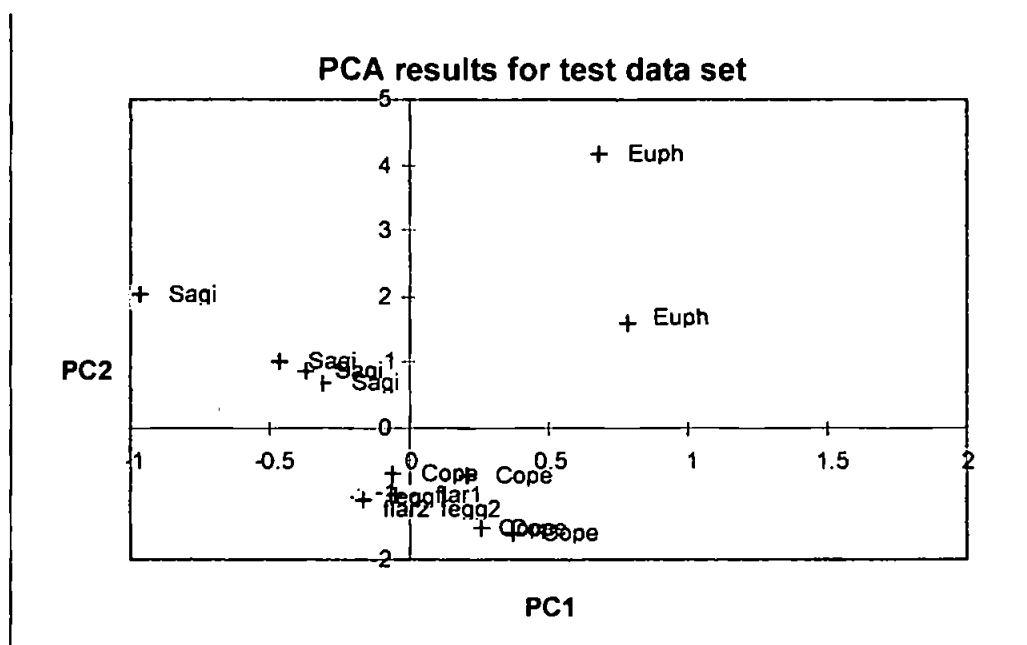
Principal component 3 represents only  $0.0118/3 = 0.3\%$  of the variance, and is given as:

$$PC3 = 0.083 * \text{Perimeter} - 0.7392 * \text{Area} + 0.6683 * \text{ESD}$$

The perimeter length is made insignificant by the weighting in this component, while area and ESD are contrasted by large weighting of opposite sign. Given that ESD is directly derived from area, this is not very meaningful, and as this component accounts for such a small portion of the variance in the original data, it should be ignored as an artefact.

The first two principal components represent 99.7% of the total variance, and indeed in this rather simplified example, principal component 1 alone accounts for 93% of the variance. For more complex and realistic data sets the first three principal components would not be expected to account for quite so much of the variance. The final decision as to how many of them to use is a trade-off between maximising the variance accounted for and minimising the number of components in the final discriminator set. Typically, for a well-ordinated data set, the first three components might be expected together to account for 80 - 90% of the total variance. Based upon this very simple data set, fig. 5.1 shows how these species group in the space defined by the first two principal

components. It can be seen that the euphausiids and sagitta can be discriminated from each other and from either the fish larvae, fish eggs or copepods, but not from all three. This is because the groups of copepods, fish larvae and fish eggs all overlap in this space. Fish larvae and fish eggs actually form groups which are separable, although this cannot be discerned on the scale of fig. 5.1. The solution to this problem is either to reduce the number of groups, or to add to the set of descriptors some measure which can separate these groups.



*Fig. 5.1 Grouping of PCA example data set on first two principal components.*

It has been noted (Hand, 1981) that error in this kind of analysis can result from the inclusion of too many descriptors as well as too few. This is because as the hyperplanes so defined become more and more complex in shape, they occupy areas of lower probability density function. Such a set of descriptors may

describe the training set more effectively, but will generalise less well to new data sets. It is therefore important to devise an optimal subset of descriptors to use in the analysis. Furthermore, Hand (1981) notes that

“To avoid spurious results which do not generalise, it is necessary to have (the number of samples) several times greater than (the number of variables measured)”

In the above example, with only three variables, this is easy, but actual training sets used to test the ViZA system with possibly 22 variables must have far more members - over one hundred at least.

### *5.2.3 A proposed subset of size and shape descriptors*

Initial experiments with real zooplankton images shows that three principal components based on five or six size and shape descriptors and two or three of the first ten Fourier descriptors can explain over 90% of the variance in the data set, and can separate seven or eight major groups so that the maximum distance of any individual from any other individual of any group to which it does not belong is significantly larger than its distance from the centre of its own group. See chapter 8 for more details of these results.

It is apparent, however, that the optimum set of weighted descriptors will be different for different communities. A standard set of discriminant principal components derived from locally obtained training sets is therefore obtained and stored with the system software. It is recommended, however, that for any location where community structure is likely to be significantly different, a new

training set is obtained, and the principal component analysis is run on this set to obtain the optimum discriminant function set for this new location. The system software, described in chapter 8, includes a program to allow the user to execute such an analysis whenever it is deemed necessary.

### **5.3 Further developments**

In order to discriminate between many more species, or between developmental stages of species, or to obtain some idea of feeding and growth rates, it is unlikely that size and shape descriptors alone will suffice. As image processing technology develops, it may be possible to introduce other types of discriminator to allow this increase in class resolution. Among such descriptors may be the use of grey scale or colour images to derive values of such parameters as gut fullness, oil storage and parasite load (Davies *et al.*, 1992).

# CHAPTER 6

## THE DESIGN OF A NEW VIDEO ZOOPLANKTON ANALYSER (ViZA).

There is a perceived lack of information on global distributions of biological properties. In order to measure and model these properties, there is a need for improved methods and technologies for the measurement of zooplankton abundance and community structure from small to large spatial and temporal scales.

### 6.1 System requirements.

The objective of the current research is to produce an *in situ* (towed) video instrument capable of continuous high resolution counting and classification of zooplankton over large spatial and temporal scales in real time. Small scale effects, such as micro-patchiness (scale < 20 m.) should be resolved. The instrument is to be developed and used in tandem with the established technology of the OPC in order to achieve the objectives currently not attainable by any single device. Discrete net samples, subjected to microscopic taxonomy and carbon analysis, and ACDP profiling will also be used for intercalibration. The OPC will provide a reliable bulk count and size distribution, whilst the video instrument will provide species identification, and numeric analysis for intercomparison. The device should be compact enough to fit with the OPC into the Plymouth Marine Laboratory's undulating oceanographic recorder (UOR), a

well proven towed instrument platform (Aiken & Bellan, 1990). The video instrument uses high speed, high resolution CCD cameras to image the volumes of water passing through the OPC, whilst both instruments are towed at speeds of up to 5 m/s. It is desired to resolve zooplankton in the size range 200µm - 20 mm.

In the first instance, a benchtop system is developed for validation of the method and algorithms. This system is used with the laboratory version of the OPC, the OPC-1L, in pump-through sampling mode using the uncontaminated sea-water supply available on most research vessels, or using live or preserved samples in the laboratory.

## **6.2 Instrument design**

### *6.2.1 Cameras*

The camera used is the EEV CAM-17 camera, which utilises a CCD imaging array 512 x 512 pixels square, each pixel being 15 µm square. The camera operates in frame transfer mode, either continuous or frame on demand, at a maximum rate of 60 frames per second (fps) using a pixel clock of 19.75 MHz. The video signal is analogue, negative synchronisation, using 13 video lines for vertical retrace (525 lines total per frame). The peak response of the sensor is at a wavelength of 750 nm., and the sensitivity 0.5 lux.

### *6.2.2 Optics*

The optics used on the first prototype consist of an Edmund Scientific HDF macro lens at a working distance of 11 cm, giving a field of view of 20 mm. and a depth of field of 28 mm. The 20 mm field of view is chosen to coincide with the 20 mm flow cell used by the OPC-1L. Assuming a minimum of 12 pixels diameter to recover useful shape information, the smallest object identifiable under such conditions will be approximately 500 $\mu$ m diameter. This figure is calculated by dividing the field of view of 20 mm. by the 512 pixels across the field of view, giving approximately 40  $\mu$ m per pixel. Twelve such pixels give a minimum diameter of approximately 500  $\mu$ m. Objects smaller than this can be stored with just diameter information for comparison with OPC size data, but no attempt to describe their shape, or classify the object will be made. In order to cover the full size range eventually required of the system (200 $\mu$ m - 20 mm.) two such systems having different magnifications, and concentric fields of view will be required. The maximum size is determined by the 20 mm. field of view, delimited by the flow cell in the OPC-1L, and the channel width in the towed version. In the development of the prototype system, this size is further limited by filtering the sea-water intake using a 6 mm. mesh aboard RRS James Clark Ross and RRS Discovery (see chapter 8).

### *6.2.3 Illumination*

In order to obtain clear images of objects moving across the field of view of the camera, the edges should not be blurred by such motion across more than one image pixel. This will require strobed illumination to freeze the image motion.

The maximum duration of the light pulse is determined, using sampling theory, as half the time taken for an object edge to cross one pixel. The effective pixel width in object coordinates has been determined above to be  $40\text{ }\mu\text{m}$ . The benchtop system uses a  $20\text{ mm}$ . square section glass flow cell, with sea water pumped through the system at 25 litres per minute. This represents a linear velocity through the cell of just over 1 metre per second. Particles will therefore cross a  $40\text{ }\mu\text{m}$ . area in  $40\text{ }\mu\text{s}$ . Half of this value gives a maximum light pulse

duration of  $20\text{ }\mu\text{s}$ .

Strobed illumination is provided by an EG&G machine vision strobe using a 60 watt xenon flash lamp with an external trigger. In order to ensure that the light pulse occurs during frame integration

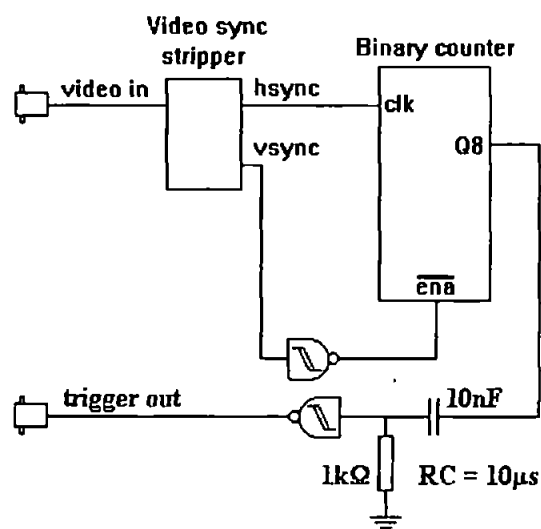


Fig. 6.1 Circuit diagram for strobe trigger

time rather than readout time when integration is disabled,

the synchronisation pulses from the camera are used to produce a trigger pulse which occurs halfway through the frame integration time. Frame transfer takes place during the vertical blanking period of the EEV camera. A digital timer is reset by this vertical synchronisation pulse. The timer is then clocked by the horizontal synchronisation pulses, effectively counting video lines. When bit 8 of the counter goes high, indicating half the integration period, this level is used

to produce a pulse 10  $\mu$ s wide, which is used to trigger the xenon flash unit.

Figure 6.1 illustrates the circuitry used for this system.

During the initial testing of the device at sea aboard the RRS Discovery, it was found that during the brief period (10  $\mu$ s) of the flash, the automatic gain control of the camera began to try to adjust the exposure of the camera to the changing light levels. The result was that the overall illumination of images was varying quite considerably between frames. A manual gain control system was therefore made and connected to the camera, allowing the adjustment of the gain of the system to obtain the best overall lighting.

#### *6.2.4 Image acquisition and pre-processing*

Image processing is provided by a Coreco Oculus F64 frame processor board. This board includes a frame grabber for image acquisition from the camera, and hardware dedicated to pre-processing and producing real-time histograms. The board also has a TMS320C40 digital signal processor (DSP) chip running at 40 MHz for image processing such as contour extraction, and sufficient memory to provide eight separate frame stores.

The F64 board is a PC ISA bus card, allowing its use in an IBM PC compatible host system. It is provided with software drivers and an applications programming interface (API) which provides 'C'-callable functions for all image processing routines required. The host PC runs the executive program, written in C++, and all speed sensitive image processing operations are performed by the fast DSP on the F64 board under control of this executive. The DSP also has

provision for a fast (80 MHz) dedicated link to separate DSP expansion cards which may be added to the host system as processing needs require.

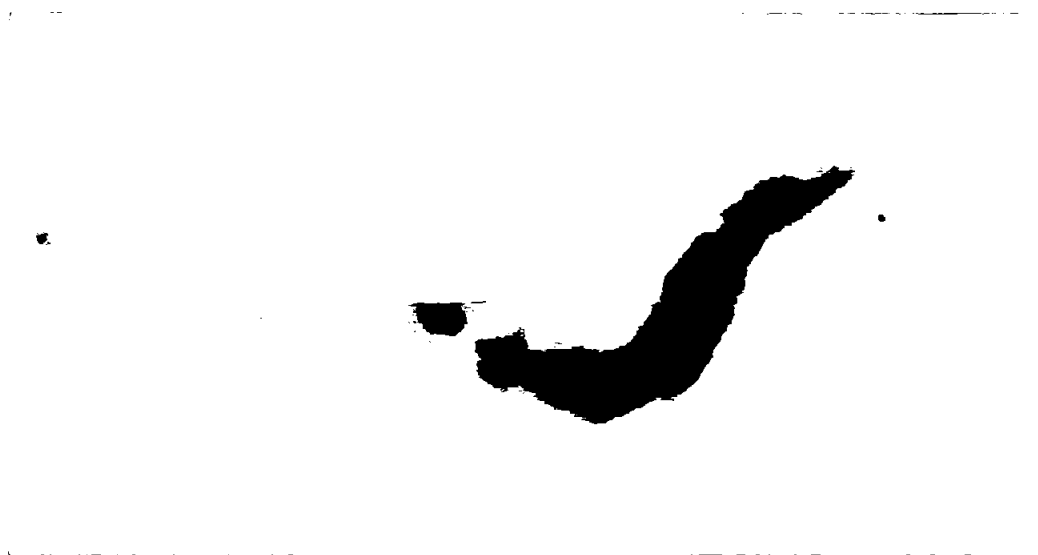
#### *6.2.5 Typical ViZA image data*

Figures 6.2 to 6.4, below, show typical images acquired by the camera, and segmented by the F64 processor card, to produce boundary chain codes which are the basic data input to the image analysis/pattern recognition algorithms of the system software, described in chapter 7. Similar images for other species are included in appendix 2. Figure 6.2 shows an image of a euphausiid, acquired during the AMT3 cruise (chapter 8). Figure 6.3 shows the image after binary segmentation, and fig. 6.4 the contour extracted from this binarised image. Note the illustration of the problem of zooplankton transparency to white light, discussed further in chapter 9. The appendages are not well imaged, due to transparency and their small width, although they can be seen in the original image. This is a positive outcome, as many zooplankton have many long appendages whose orientation is not predictable. The same animal might therefore present very different shape parameters on different occasions, and would therefore not be consistently identifiable. The shape of the main body trunk does not vary much, and it is with this shape that we therefore attempt to classify the animal. Many zooplankton species, particularly the larger among them, also exhibit transparency to white light in their main body sections, and the euphausiid shown below is an example. The body section connecting the head to the rest of the body is transparent, and when segmented, the head appears as a separate entity. It has generally been found that the size and shape

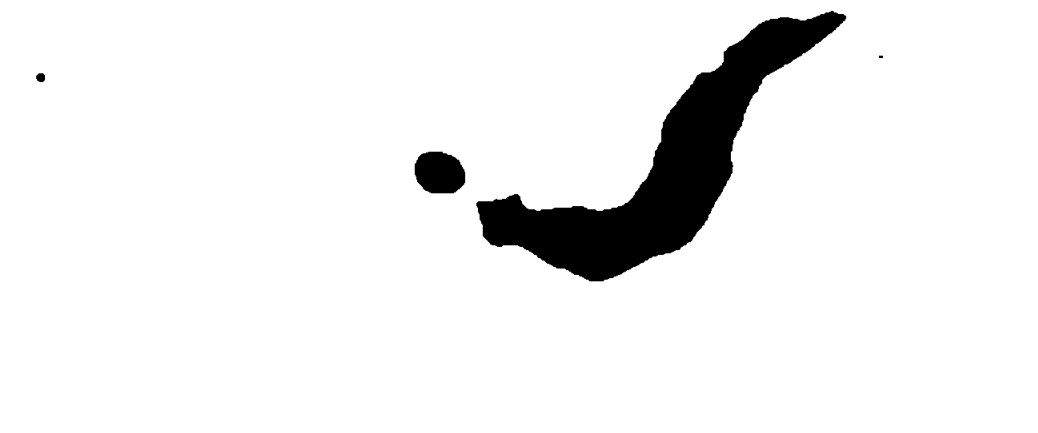
of euphausiids is sufficiently distinctive that images such as this are still identifiable as euphausiids. The loss of the head and part of the body will however lead to two inaccuracies:

- (i) The size of the animal is underestimated;
- (ii) A non existent animal (the head) is captured and mis-classified;

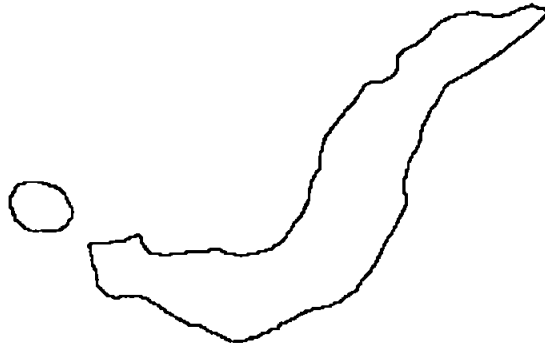
This problem, and a possible solution, is further discussed in chapter 9.



*Fig. 6.2 Example of image captured using ViZA system during AMT3.*



*Fig. 6.3 Binarised (segmented) version of the image in fig. 6.2*



*Fig. 6.4 Boundary contour extracted from the segmented image of fig. 6.3*

#### *6.2.6 Towed platform*

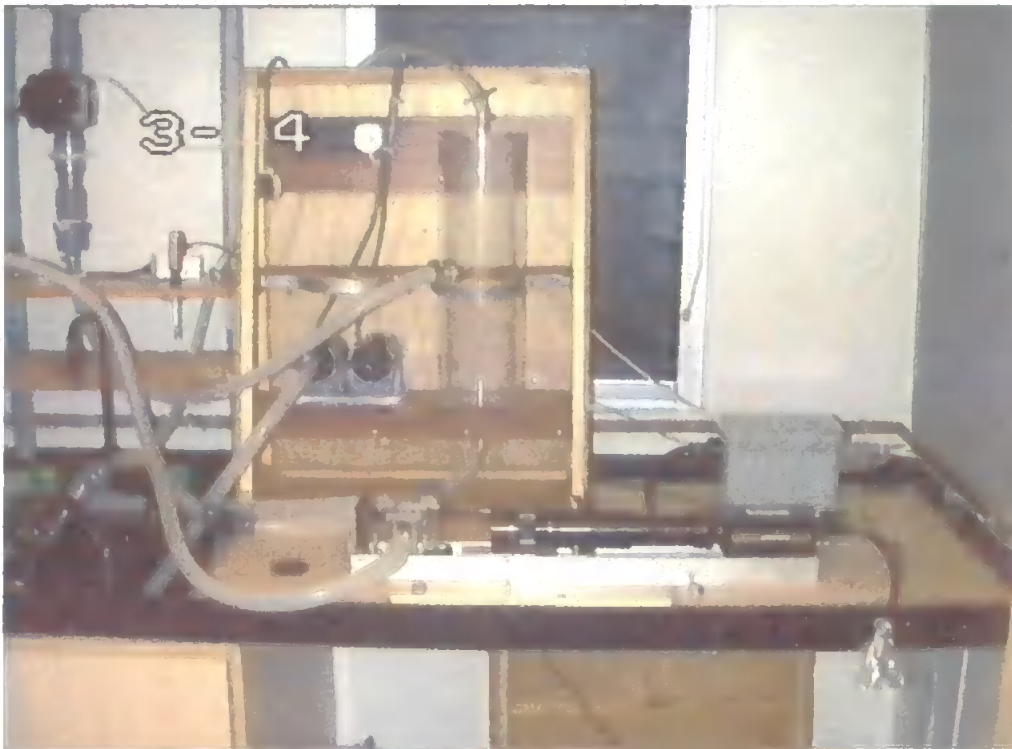
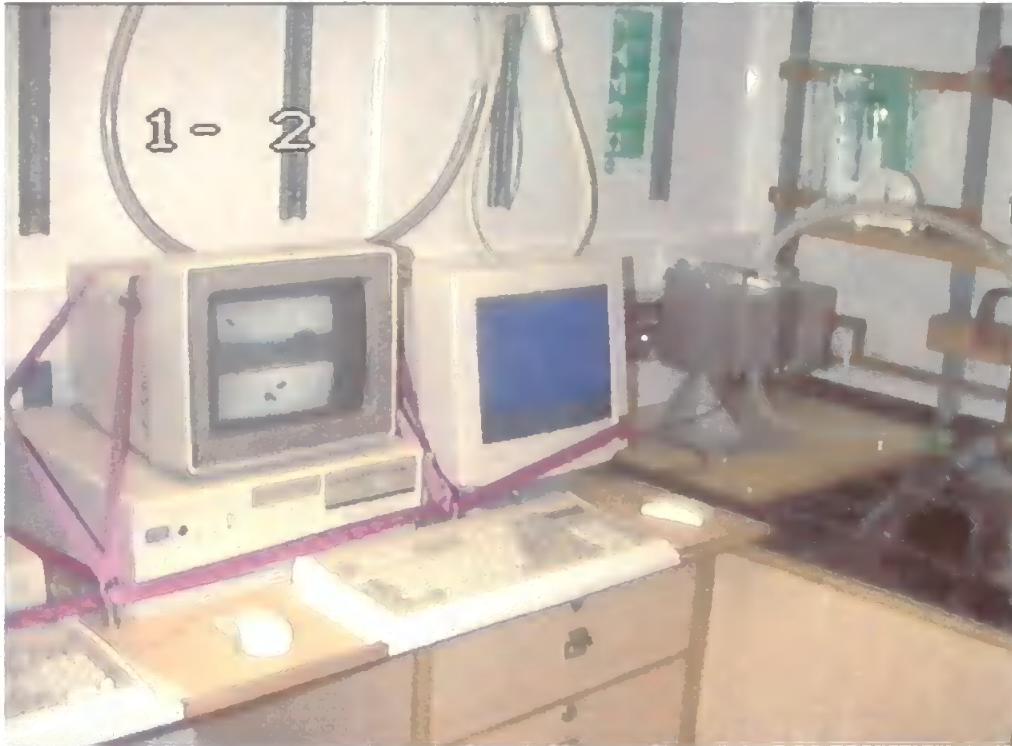
The towed platform for which the system is intended is the Plymouth Marine Laboratory's Undulating Oceanographic Recorder (UOR) (Aiken & Bellan 1990). This device will carry both the OPC and the video device, as well as optical and physical environmental sensors. As stated, the device has initially been developed as a benchtop system, and development of the towed system has yet to begin.

#### *6.2.7 The OPC and flow through system*

The OPC and the de-bubbling device are described in chapter 3. Figure 6.5 is a photographic pair taken aboard RRS James Clark Ross, showing the connection of the OPC and ViZA to the ship's uncontaminated supply via the debubbler.

The top picture shows the OPC-1L and the computer used to control the video device. The lower picture shows the wooden box holding pumps and de-bubbler, with the Xenon strobe unit to the right and the video camera in the foreground. Top left is the stopcock connected to the ship's uncontaminated sea

water supply. The video device is mounted upon a rigid aluminium beam with the backlight unit on the optical axis, and the flow cell between the two. The flow cell is the same as that fitted in the OPC, and is mounted with its axis 110 mm. from the camera lens, giving a field of view of 20 mm., which means that the internal dimension of the square sectioned flow cell just fills the field of view.



*Fig. 6.2 The OPC and ViZA flow through system for underway sampling using the ship's uncontaminated sea water supply.*

At this distance, the depth of field of the HDF optical system is 28 mm. so that all animals in the flow cell should be in focus.

The flow through pump for processing preserved or vertical net samples delivers 10 litres per minute, giving a flow speed of 0.5 m/s, the minimum for correct operation of the OPC. The water is recirculated through the OPC and ViZA using a 10 litre container as a header tank. The sample is added to the container, passes through both instruments, and collected by a filter fitted to the outlet. Samples may then be preserved for laboratory analysis.

# CHAPTER 7

## SYSTEM SOFTWARE

The system software consists of a host PC executive program, written in C++. This program controls all the activity of the frame processor boards including acquisition, segmentation and pre-processing, and the geometrical description and identification of any objects found in the video frames acquired. The program also allows the collection of acquired image objects as a training set for the recognition and classification algorithms. A second program, written in Turbo Pascal and using the Turbo Vision library, performs statistical analysis of stored data sets, including the principal component analysis used to derive the discriminant functions used in the classification procedure. This program may be called from within the host program after the collection of a training set for analysis. The principal components derived from this analysis are stored in a disk file for use by the host program in these classification procedures. The program listings for the system software are bound separately, and accompany this thesis, being rather too lengthy to be included in the same volume. This chapter simply presents a brief overview of the functionality of the programme in the context of its use. A detailed analysis of the code would be rather too lengthy, and is inappropriate in the current study.

## **7.1 Host processor software**

The main program function consists of the display of a menu offering four choices of processes, or exit. The four process options are: Training\_Set, Manual\_Run, Auto\_Run and Continuous. Selection of one of these by the user causes a call to be made to the relevant subroutine to process the selection. At the termination of each of these routines, control passes back to the main loop, which displays the menu again and prompts the user for a further selection. This continues until the user selects exit, which causes the program to terminate and return to the system prompt. Each of the main processes is described in this section, together with the image acquisition software and the data file handling.

### *7.1.1 Image acquisition software*

The acquisition of images is handled by the Oculus F64 frame processor board. The functions controlling the various activities of this board are made available by the manufacturer as a library of C language routines, which are called from the main host program. These functions access the instrument driver, a terminate and stay resident (TSR) program, which must be loaded during host computer initialisation. The functions used by the system software are outlined below:

#### *odxbind()*

This function must be called before any other frame processor function. It initialises the F64 frame processor driver, and binds it to the host program, providing software access points to the driver.

### *setdmajor()*

This function sets the F64 device to be used in function calls. The default is to assume only one major device (0) is connected, and to use that one.

### *opr\_inq(reg, val), opr\_set(reg, val)*

These two functions allow any one of the F64 device control registers to be set or read. Examples are the memory frame buffer size, the threshold at which the image is binarised, etc.

### *fbgrab(x)\**

This function initiates the acquisition of a new frame from the camera. The parameter *x* can be -1 (continuous grab), 0 (end continuous grab) or 1 (grab one frame).

### *cntrfirst()\**

This function prepares the frame buffer containing the acquired data for contour retrieval. The image is binarised using the binary threshold previously selected, and the current frame buffer position set to *x* = 0, *y* = 0.

### *y = cntrnext(feat, ldot, nmax)\**

This function acquires the contour of the next object found within the image, and stores its length in the variable specified by *y*. The additional features to be extracted are specified by the first element of array *feat*, and are basic size descriptors such as boundary length parameters, the area and centre of area of

the object. These features are stored in the other elements of the array *feat*.

From these parameters all other size and shape descriptors are calculated by the host program as described in chapter 4. The boundary chain code is stored in the buffer *ldot*. The maximum number of values which may be stored are specified by the integral value *nmax*.

#### *drwbox()*

This function draws a box around the acquired image in the gray level specified by a register previously set by *opr\_set*. This is 255, maximum white. This is to prevent unclosed contours from being retrieved from objects overlapping the edge of the field of view of the system.

#### *histo(hist)\**

This function returns a histogram of gray scale values for the image. This histogram is extracted by dedicated hardware on board the frame processor board in real time.

These functions are implemented as macro instructions in a C library module. The macro commands are all calls to the function *odx\_send(x)* where *x* is an integer representing the function required. Due to the need to improve the speed of operation of certain of these calls (marked \* above) in the automatic mode, the calls to this library were replaced in the code by direct macros. Because calls to subroutines involve computational overheads, this resulted in a considerable reduction of the time to execute these functions.

During the initial sea trials of the video system aboard the RRS Discovery, the sampling rate was too low effectively to sample the less abundant species above about 500 $\mu$ m ESD (see chapter 8). Profiling the program using Borland's Turbo Profiler utility showed that over 90% of the run time was spent waiting for the execution of the function *cntrnext*, which seeks the next object in the image. Because the vast majority of the frames acquired had no object of interest, this time was wasted. While seeking objects which were not present, it is likely that objects could pass undetected through the system. The software was therefore modified to check the frame histogram for the presence of a minimum number of pixels whose gray level was below the threshold for detection of a contour. Because the histogram is acquired by hardware in real time, and this check involved the processing of only perhaps 100 values (as compared to 262,144 values when seeking contours in an empty frame), this resulted in a considerable saving of time, and the sampling rate for empty frames was improved by a factor of 8. This makes it far more likely that a particle will be detected by the system.

After the first frame is acquired in automatic mode, and before processing any contours within the frame, the processing and display buffer pointers are set to the value of the acquisition buffer - i.e. the memory location of the frame just acquired. The acquisition buffer pointer is then incremented and a new frame grab command issued. This allows the parallel processing and acquisition of frame buffers, thus saving the time of waiting for the next frame to be acquired. Given that eight such buffers of 512 by 512 pixels are available, this method

results in the continuous storage of acquired frames to a circular series of 8 buffers, with frame processing of the buffer previous to the one currently being written to by the acquisition hardware.

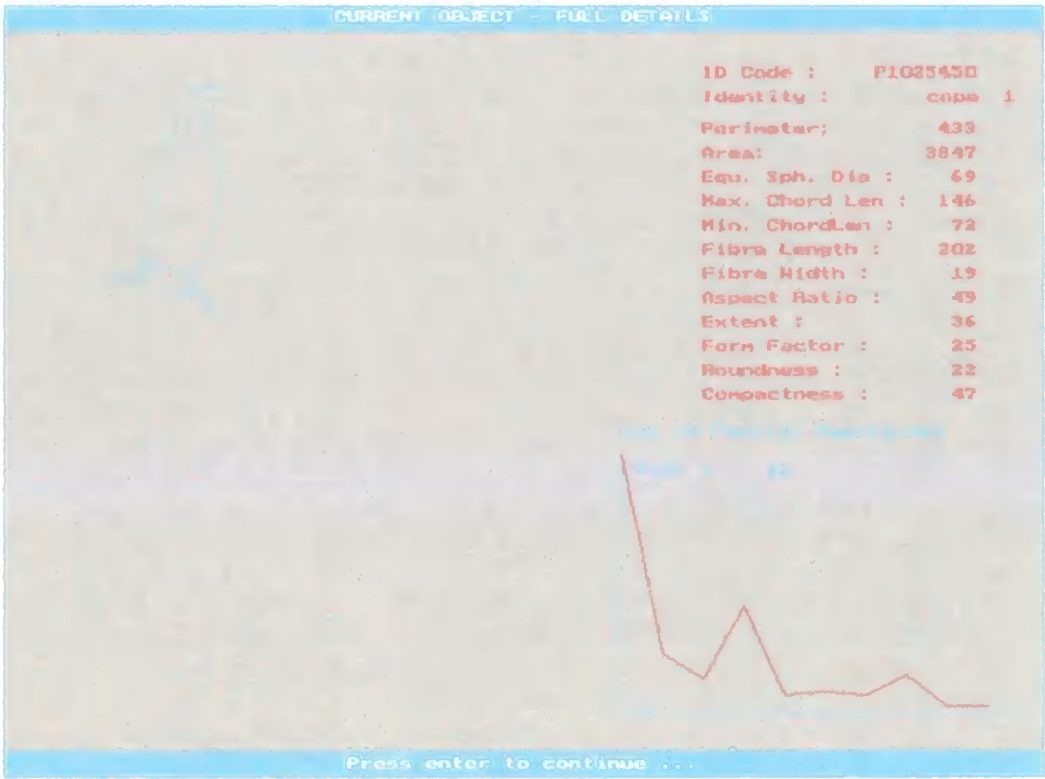
These modifications in the software resulted in an improvement in the time for processing empty frames from about 1 second to about 100 milliseconds, and the time to process a frame containing at least one object from about 1.2 seconds to about 550 milliseconds. At 12 litres per minute, and given the sampling volume of a video frame of 8 ml., this gives us a sampling ratio of 1:2.5 for empty frames. As the vast majority of the frames acquired are empty, and assuming that two particles will not appear in the same half second period, this may be taken as approximately the sampling rate of the system. This will only be true at typical concentrations in oligotrophic waters, but as concentrations increase, the empty frame problem becomes less acute.

#### *7.1.2 Training set procedure*

The purpose of this routine is to acquire a training set of object outlines, together with their size and shape descriptors. This set is then used to determine the optimum set of descriptors for classification into a chosen set of groups, using the principal components analysis program.

The routine consists of a continuous loop whereby a frame is acquired and any object outlines presented to the user one by one. A typical object, as presented to the user, is shown in fig. 7.1. The user is asked whether each outline is to be

included in the set. If the answer is yes, the user is asked to identify the group to which it belongs (copepod, euphausiid, etc.), and the object information together with this classification is stored to disk. The number of groups is inversely proportional to the efficacy of the classification algorithm in identifying membership of each group, so it is up to the user to ensure that the system is not set an impossible task. As familiarity with the system increases, the user will develop a feel for how many different groups the system can reasonably deal with. Some balance between these two conflicting requirements must be struck.



*Fig. 7.1 Typical object outline as presented to the user by the acquisition routine*

When no more objects are found in a frame, the user is asked if another frame should be acquired. If the answer is yes, the loop repeats. When the user feels he

has a full enough test set, he answers no to this question. The loop then exits, and the user is asked whether he wishes to perform the principal component analysis. If this is answered in the affirmative, the PCA program is spawned from within the host program. This program is described in section 7.2, below. Upon exit, the principal components to be used in subsequent classifications are stored in a special data file, and control is returned to the host program.

### *7.1.3 Manual run procedure*

The manual run option allows the user to acquire frames one at a time under his own control. The outline of each object found within the frame is drawn on the screen as in fig. 7.1, together with size and shape information including the first 10 Fourier descriptors. The system's classification of the object is also displayed. The purpose of this routine is really for testing the classification procedures. These procedures are described in the next section in the context of automatic processing.

### *7.1.4 Auto run procedure*

The structure of this routine is similar to the manual run routine, except that the user is not prompted between frame acquisitions, which are continuous. Before entering the main routine loop, the principal components for each group and the mean location of each group in the space so defined are retrieved from disk file. The routine loop only breaks if the user interrupts it by pressing any key on the computer keyboard. Each frame acquired is checked for objects within the specified size range (if the histogram test, discussed above, shows the frame to

be worth checking). For each object found, the derived size descriptors (ESD, major chord length, minor chord length) are calculated from the basic size descriptors (boundary chain code, perimeter length, area) as described in chapter 3. The five shape descriptors (aspect ratio, extent, form factor, roundness, compactness) are then calculated. The Fourier descriptors are also calculated using an FFT algorithm (Press *et al.*, 1992) to produce a complex Fourier transform of a complex valued boundary signature array, as discussed in chapter 4. These descriptors are then normalised using mean and standard deviation information for each variable stored with the principal components in the disk file. Principal components are formed for the object using the coefficients retrieved from the principal component analysis. For each group found in the training set, the Euclidean distance of the current object from the group in terms of its principal components is calculated. If this distance is within some pre-determined maximum, the object is classified as a member of that group, and the species distribution information is updated. This will have been determined by the PCA routine. The maximum distance from the group mean is usually two standard deviations, assuming that this distance permits all groups to be clearly separated. This should allow the successful classification of 95% of animals of each class, assuming a normal distribution of values for the discriminant functions. If no group can be found whose centre is near enough to the current object for positive classification, the object is marked as 'unknown' and stored to disk file, as above. It can later be examined for manual identification, or for investigation of the cause of classification failure. The ESD of the object is also used to update the size distribution for comparison to the

OPC data. The object information is then stored to disk file as described in the section below.

During the automatic processing of images, the results are displayed on a graphics screen containing three graph displays. The top left graph is a histogram of particle count vs. ESD, allowing direct comparison with the output of the OPC. The top right graph is a similar histogram, this time of particle count vs. species as determined by the pattern recognition system. This histogram will contain a species class for each of the groups to be discriminated, as determined by the principal component analysis of the calibration, and one additional class for unknowns, i.e. those objects which cannot reliably be classified to any such group given the parameters determined. The third graph, at the bottom of the display, gives a continuous value for particle count per cubic metre of underway analysis. This value is integrated every 0.2 litres and shows a continuously updated value for the last 25 litres imaged - about 5 minutes or 1.6 km. of cruise for the benchtop system given a flow rate of 25 litres per minute, a sampled volume of 8 ml. per frame, and empty frames sampled in 110 ms.

#### *7.1.5 Continuous grab procedure*

This routine simply initiates a continuous frame acquisition, each frame being acquired and displayed in real time. The user may terminate the continuous acquisition at any time by striking a key. On termination, the most recently acquired frame is displayed, along with a cursor. The user may move this cursor

to any point in the image using the mouse or the keyboard arrow keys. As he does so, the image co-ordinates are displayed on the program screen, along with the gray scale value at each point. This routine is useful for examining the general illumination and imaging conditions, and for finding the best gray level at the edge of any objects in the image. This gray level may then be used to determine the threshold at which images will be binarised.

#### *7.1.6 Data file handling*

The object is stored, along with its size and shape descriptors and boundary chain code, to disk file. In any one hour, 7,200 or more such objects may be stored. For long oceanic transects, many hundreds of thousands of files will be stored, and in order to make their retrieval practical, some coherent structure must be imposed on the way they are stored and identified. A system of directories and sub-directories must be established so that the number of files being dealt with is not overwhelming. The DOS operating system will in any case only allow a maximum of 512 files in any one directory.

The file names are derived from the DOS time, showing the exact time of acquisition, thus:

P1131069.DAT

where P simply indicates a plankton file, and the seven digit number represents the time in hours, minutes, seconds and tenths of a second. These files are typically 200 - 1000 bytes in size. Given the way that PC disk operating systems handle files (minimum file size 50,000 bytes for a large disk drive) this will

result in very inefficient storage. Instead of space for 1,200,000 files being available on a 1.2 Gigabyte disk drive, perhaps only space for 24,000 files will effectively be available. For a long cruise, an impractically large number of such drives would be necessary. In order to avoid this problem, object data is collected in groups of 64 within each file, which is named as above. Every 256 files so obtained are stored in separate sub-directories whose names are similarly derived from the time, this time preceded by the letter 'D', thus:

D113106

where D indicates directory and the six digit number represents the time in hours, minutes and seconds. Each of these sub-directories are grouped within a daily directory whose name is formed from the cruise identifier and the Julian day. For example, for the first Atlantic Meridional Transect (AMT1, see chapter 8), on 26th September 1995, the daily directory name would be:

AMT1J269

This representation allows the rapid retrieval of collections of files by the cruise identifier and the time and date of their acquisition. This time and date may be converted to precise position of acquisition by reference to the ship's navigation log.

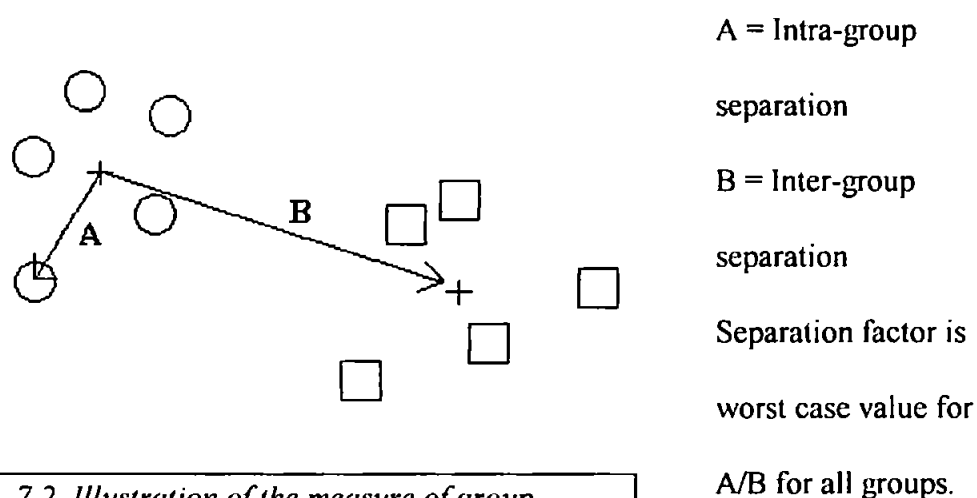
## **7.2 Principal components analysis**

A separate programme has been written for the statistical analysis of a training set in order to provide the optimum set of discriminant functions for a given community. This programme was written in Turbo Pascal using the Borland Turbo Vision programming library. This program can be called from within the

host program after the acquisition of a new training set, or independently from the system prompt. The training set (or any other data set acquired by the system) is loaded as a single collection into the program with all object outlines, size and shape descriptors. It is desired to find an optimum set of these descriptors to use in classifying the zooplankton. Whilst the principal component analysis results in a reduced operator set, the number of descriptors used in deriving the principal components will affect the outcome of that analysis. This is particularly true if, as is likely, there is considerable correlation between the descriptors. The selection of an optimal subset of these descriptors is problematic, however, as this subset will vary depending upon the species present. As these will vary from place to place, and from time to time at a single location, the *a priori* definition of an optimum set of descriptors to use in the principal components analysis is not possible. The way in which this conflict is resolved is as follows. Once the training set has been loaded, the program performs a principal components analysis on the set, using each of all possible combinations of the classifier set. (Note that given a set of 22 size and shape descriptor, this requires  $2^{22}$ , or about four million principal component analyses to be performed. On a typical 66 MHz Pentium PC, this will require about twenty hours' processing. This is not, therefore a procedure to be undertaken lightly, or more often than necessary).

For each principal component analysis, a single parameter is derived which describes the classification efficacy for this permutation of the descriptor set. This parameter is formed from the largest value for the inter-group separation

divided by the intra-group separation, from the Euclidean distance based upon the first three principal components. These first three should account for at least 80 - 90% of the data variance (see chapter 5), and usually they do. Evidently, the samples may be correctly assigned to a group if this parameter is greater than one. If it is less than one, then there is somewhere within the sample an overlap of two or more groups, and samples cannot be unambiguously allocated to these groups. This problem is illustrated in fig. 7.3. This separation parameter is presented to the user as the result of the analysis, along with the subset of descriptors identified as most effective in discriminating the samples chosen, and with the comparison with the efficacy of the default subset.



*Fig. 7.2 Illustration of the measure of group separation used in the PCA algorithm.*

A graphic display of the groups and their separation is presented to the user, where the two axes are the first two principal components. This display may help the user in his choice, and an example is shown in fig. 7.3. The three principal components are then displayed in tabular form showing the coefficients of each, and the variance accounted for by each. The user may then

opt to accept this subset of descriptors, to repeat the analysis using different groups, or to accept the default parameter set provided.

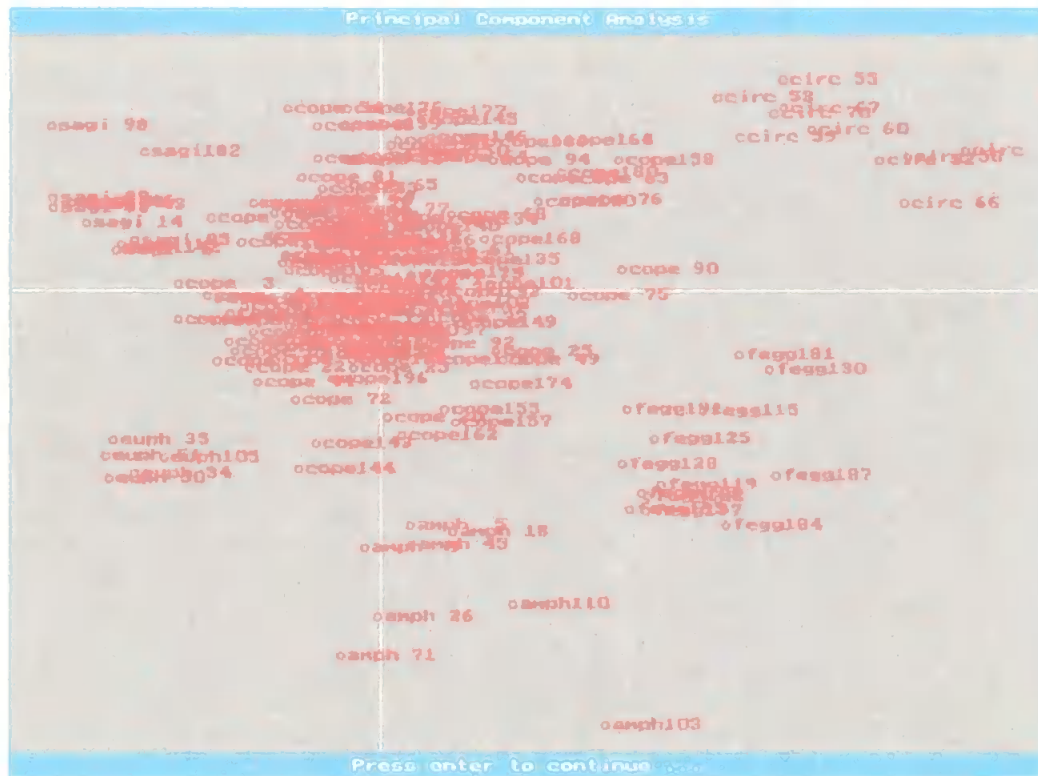


Fig. 7.3 Typical graphical display of the results of the PCA.

Once a satisfactory set of principal components has been arrived at, the coefficients of the first three components are saved to a disk file, along with information such as the mean and standard deviation of each size and shape descriptor within the training set, and the mean values (group centres) and standard deviations of the principal components for each group. This allows the descriptors for any object to be classified to be normalised, and the first three principal components calculated by application of the coefficients. The Euclidean distance of the object from each group mean may then be calculated,

based on these principal components, and compared to the standard deviations for these components in that group. The object may then be classified accordingly.

Other features which are offered by this program are:

- Loading and displaying data sets for visual inspection of identification, shape, size and shape descriptors;
- Displaying size and species distribution histograms of such data sets;
- Displaying population statistics for such data sets, i.e. mean and standard deviation for each of the size and shape descriptors used;
- Repeat identification of all members of a data set using different principal components for comparison of discriminatory effectiveness.

# CHAPTER 8

## FIELD OBSERVATIONS AND RESULTS

### 8.1 Laboratory testing

#### 8.1.1 Calibration

A joint size calibration exercise was performed using calibration beads obtained from Focal Technologies, Inc., with both the video system and the OPC. These calibration beads are of  $500 \pm 10$ ,  $1000 \pm 20$ , and  $2000 \pm 40$   $\mu\text{m}$  ESD. These were each mixed with filtered sea water and recirculated through both systems in series. The table below shows the performance of the two systems.

**Table 8.1: Size calibration for OPC and Video:**

Bead Size ( $\mu\text{m}$ )	OPC		Video	
	Mean	Std Dev	Mean	Std Dev
500	532.5	184.6	494.5	124.6
1000	1186.8	283.5	984.6	133.3
2000	2002.8	203.0	1856.6	561.9

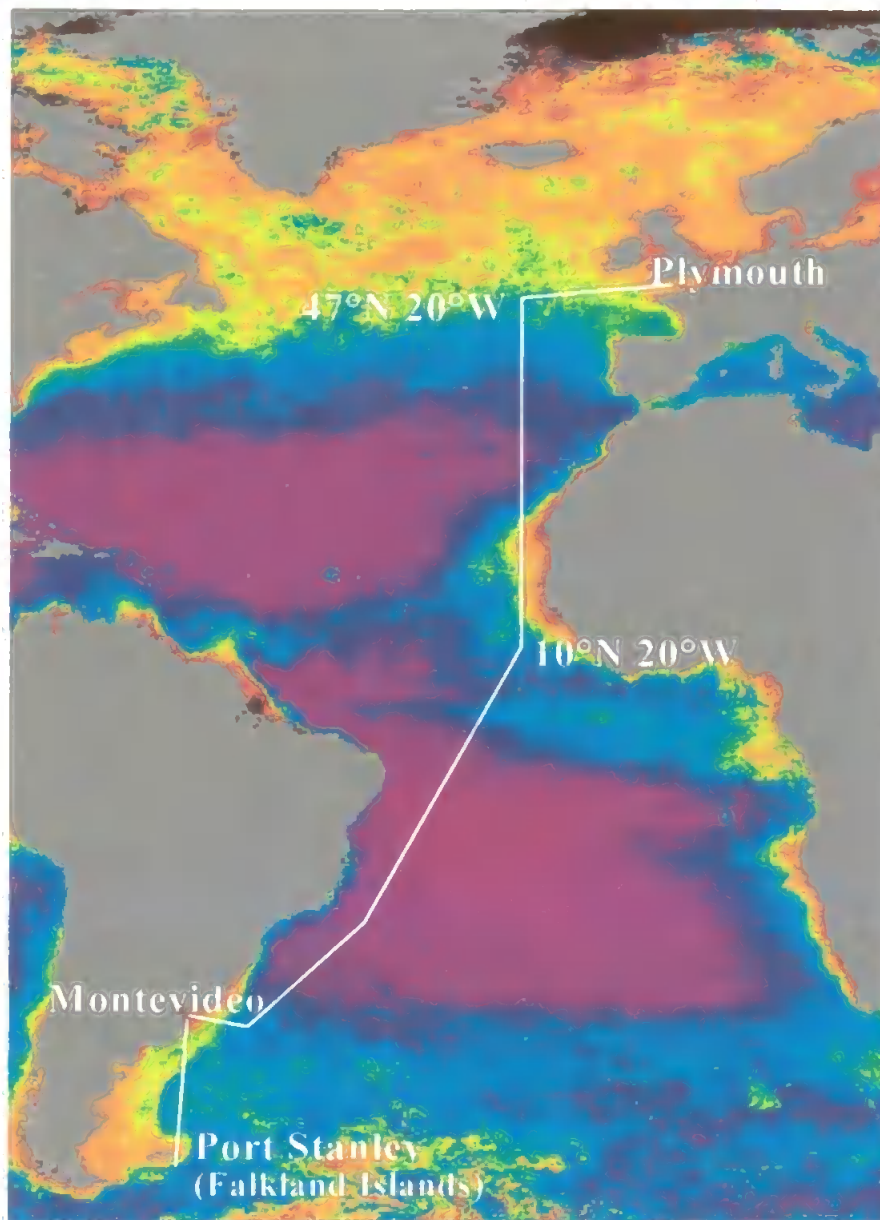
The OPC is very much more precise in air than in water, the standard deviations being an order of magnitude smaller. As the system is not intended to be used other than with sea water flow through, these results are not presented, and the ViZA system was not tested in air.

The video system is more deterministic in measuring size (measurements are made directly of the occulted area of a plane of known, fixed dimensions), and generally more accurate and precise. The 500 $\mu$ m and 1000 $\mu$ m beads are plastic, and fairly opaque, but the 2000 $\mu$ m beads are glass. The failing at 2000 $\mu$ m is due to the transparency of these glass beads. The resulting outline of many beads is incomplete, presenting a 'crescent' outline rather than a circle. This failing also applies to transparent zooplankton types such as chaetognatha and gelatinous forms. These forms are less transparent to red than to white light, and it is expected that switching to a red or near infra-red source will alleviate this problem (chapter 9). Alternatively, this fact could be used to advantage - if we are interested in carbon content, then the difference between the OPC data (which includes gelatinous species) and ViZA data, which does not, will allow us to determine size fractionated carbon directly and in real time from the two data sets, given a reliable estimate of the carbon content of gelatinous and non-gelatinous species (see below).

## **8.2 Ocean trials - AMT and PRIME cruises**

The Atlantic Meridional Transect (AMT) is a 3 year programme with the primary objective of investigating basic biological processes in the open Atlantic Ocean over very broad spatial scales (50° N to 50° S). This is fundamental for the calibration, validation and understanding of remotely sensed (satellite) observations of biological oceanography. It is also important for understanding plankton community structure over latitudinal scales, an important factor in the health of the oceans, and the role of oceans in global carbon cycles. In the longer

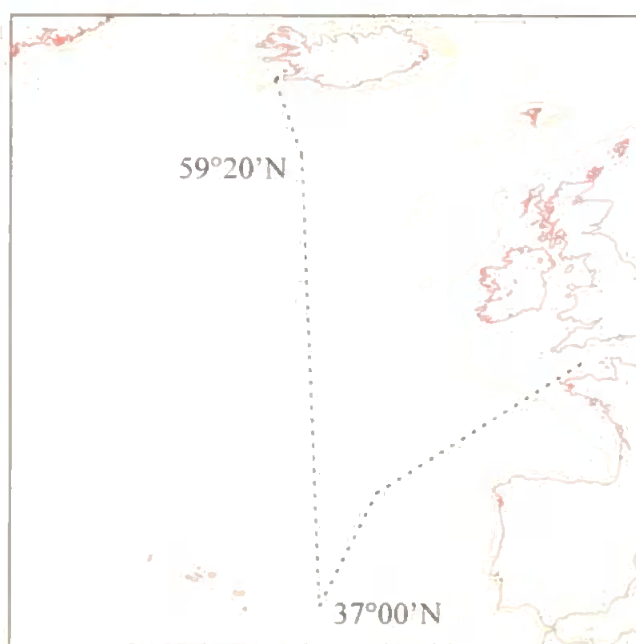
term the AMT project aims to enhance our ability to model global primary production (at basin scales) and to help develop ecosystems dynamics models which will be important in the forecast of change.



*Fig. 8.1 AMT Cruise plan. Ocean Colour Background: CZCS Colour Composite (courtesy of NASA). Red and orange represent areas of high productivity, blue and mauve much lower productivity.*

The programme is being conducted aboard the British Antarctic Survey (BAS) vessel RRS James Clark Ross on her regular annual voyage south to support the BAS Antarctic bases. The first such cruise (AMT1) departed the UK on the 24th September 1995, bound for Montevideo and Port Stanley, Falkland Islands (see chart, figure 8.1). The second cruise (AMT2) departed Port Stanley on 23rd April 1996 and followed approximately the track shown in fig. 8.1 in reverse. AMT3 departed the UK on 22nd September 1996 and followed a track similar to AMT1, reaching Port Stanley on 25th October 1996. It is hoped (subject to funding and approval) that this voyage, and the return voyage to the UK, will be repeated for the next two years.

The PRIME North Atlantic cruise forms a part of the Natural Environment Research Council (NERC) community research programme “Plankton



*Fig. 8.2 PRIME Cruise track 2-22 July 1996*

Reactivity in the Marine Environment” (PRIME). The project encompassing the present research is PRIME special topic P19 - “The optical characterisation of zooplankton in relation to ocean physics; discrimination of seasonal, regional and latitudinal variations”. The cruise, aboard RRS

Discovery, departed Reykjavik, Iceland on 2nd July 1996 and followed the 20° west meridian to 37° north, sampling continuously underway and with 200m. vertical net casts twice a day. At 37° north, an eight day station was kept, with 200 metre integrated vertical net casts twice a day (around midday and midnight). The vessel then returned to Southampton on 26th July. Some useful overlap between this and the track of the AMT cruises therefore exists.

### **8.3 Atlantic Meridional Transect 1 (AMT 1)**

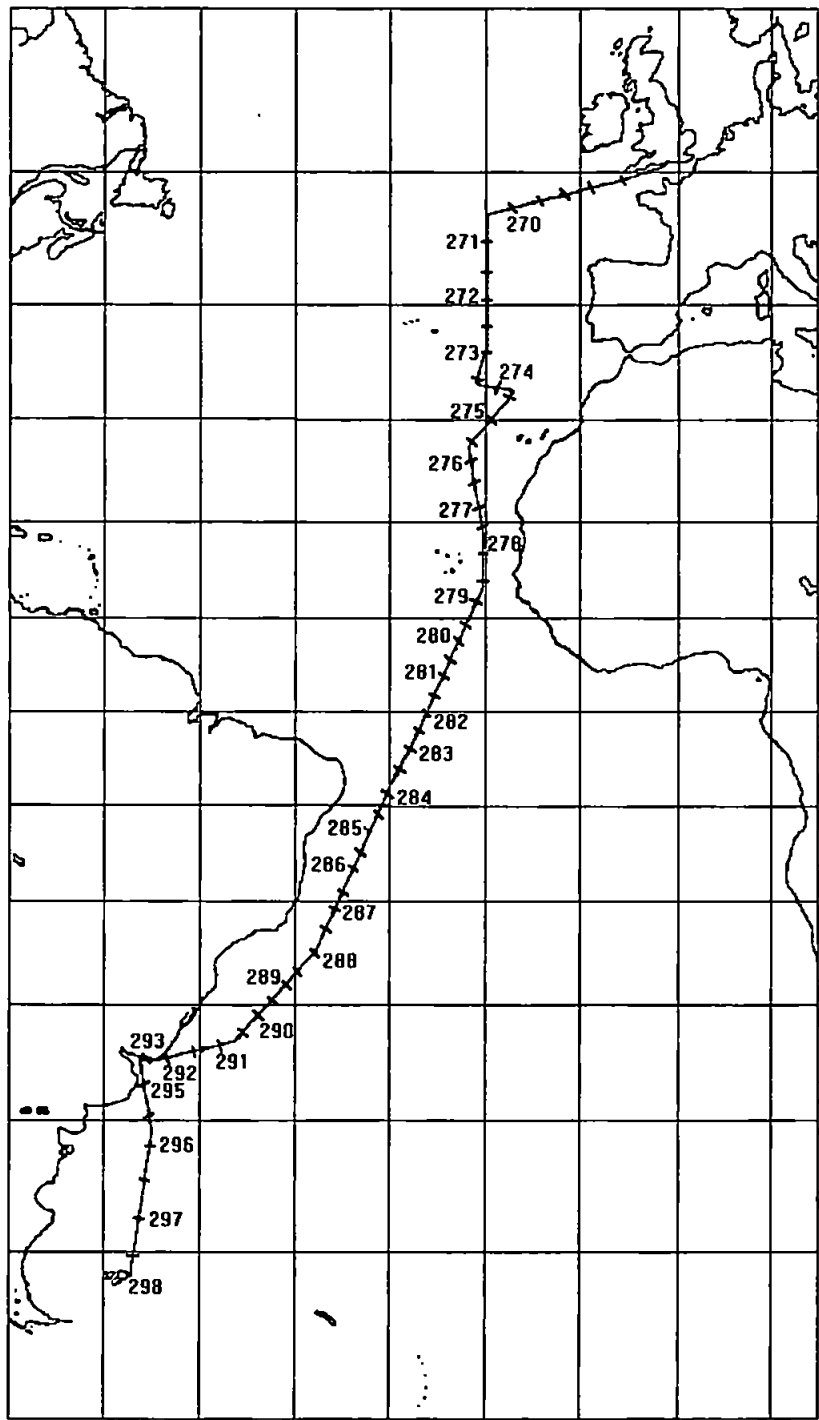
The first AMT cruise departed UK waters on September 24th 1995. The video plankton survey system was not sufficiently advanced in its development to be used on this cruise, so only the OPC was used for underway survey.

#### *8.3.1 Methods*

The laboratory and deck areas of RRS James Clark Ross are provided with a specially designed sea water supply uncontaminated by the ship's systems, for scientific sampling. The supply is taken from beneath the keel (about 6.3 metre draught) via an extensible intake. This intake extends to 30 cm. clear of the ship's keel, so that water is sampled away from the hull boundary layer which may contain bubbles created by the ship's bow wave. The normal 1 mm. mesh filters fitted to the sea water intake were replaced with 6 mm. mesh steel filters, allowing zooplankton to pass through, while removing larger objects which might damage the pumps.

There were two pumps of the centrifugal type with a capacity of about 6 cubic metres per hour each. The supply was distributed to the laboratories and working

deck areas, and some outlets were fitted with inline sensors for automatic logging of such parameters as temperature, salinity and chlorophyll fluorescence.



*Fig. 8.3 Transect and stations for AMT1 cruise.*

In addition to the underway sampling during AMT1, the vessel stopped once a day at local noon for a series of vertical profiles to determine biological and physical parameters. At these stations, a vertical zooplankton net cast was made using a 200µm mesh WP-2 plankton net, integrated from the surface to 200 m. depth. The sample obtained was split in half using a Folsom splitter. One half was used to subsample size fractionated zooplankton for carbon and nitrogen, to JGOFS (Joint Global Ocean Flux Study) protocols. The other half was run through the OPC and then preserved for taxonomic analysis. The full OPC transect therefore consisted of (more or less) continuous surface sampling interspersed with daily integrated vertical samples.

As this was the first time that the OPC had been used by the AMT team, and also the first time that the bench top OPC had been used at sea in underway sampling mode, there were some initial methodology problems to be resolved. These problems were chiefly those of de-bubbling the ship's uncontaminated sea water supply, and of condensation forming on the outside surface of the flow cell, as described in chapter 3. The methods of validation and data handling had also to be resolved during the early part of this cruise. The result of these factors was that data was not available from the OPC until Julian day 270 (the third day of the cruise), and also that the data obtained in the 36 hours after leaving Madeira were suspect. This last problem was due to the fact that the sea water intake extension was not redeployed after leaving Madeira, and this fact was unknown to the scientific team until the following day.

In spite of the above problems, a fairly comprehensive set of underway sampling data was obtained for the main part of the transect, from 47°N, 20°W to 10°N, 20°W, and thence to Port Stanley via Montevideo. The figures below show bulk data for the cruise.

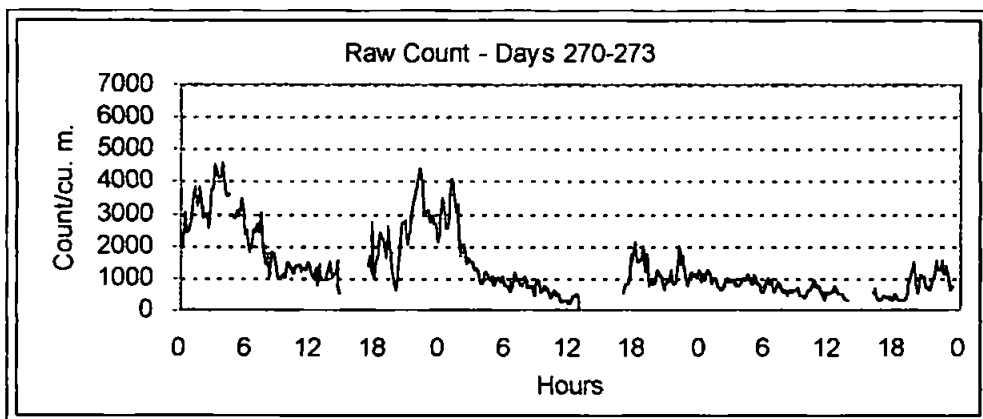
### *8.3.2 Data results*

Figure 8.4 shows the raw OPC counts (numbers of animals) for the entire underway transect between 46° N 20° W and Port Stanley, FI. The higher counts near to the UK continental shelf on days 270-271 can be seen to contrast with the lower counts in open ocean waters during days 272-273. The diurnal variation in zooplankton numbers in surface waters, due to diel vertical migration, can also be seen by the higher counts at night. Note the change of scale on the last graph of fig. 8.4. Two further points of interest are indicated by the raw counts shown on these figures:

- (i) Diurnal cycles in numbers are evident throughout the transect, but are particularly evident in the higher latitudes at either extreme. On days 270-271 and 295-297 in particular, surface counts are very high at night, and much lower during the day.
- (ii) There is also evidence of large increases in counts at dusk and dawn, particularly between 7° north and 28° south (days 280-288). This phenomenon seems to offer some support to the theory of 'midnight sinking'.

*Midnight sinking:*

A phenomenon associated with DVM is the appearance of peak abundances at the surface at dawn and dusk, with midnight sinking apparent between these peaks. It has been suggested (Raymont, 1963) that this may be due to species seeking 'optimum' light levels. As light levels fall below this optimum, these species rise to the surface. Further reduction in light level to a point at which vertical position makes no difference results in the cessation of active swimming in response to light levels, and the natural tendency of zooplankton to sink passively takes over. As dawn approaches, and light levels again increase above this threshold, these species again rise in an attempt to reach this optimum light level. Shortly after dawn the surface light level exceeds this optimum, and the zooplankton again actively descend to deeper waters.



*Fig. 8.4a AMT1 Raw counts along transect (underway).*

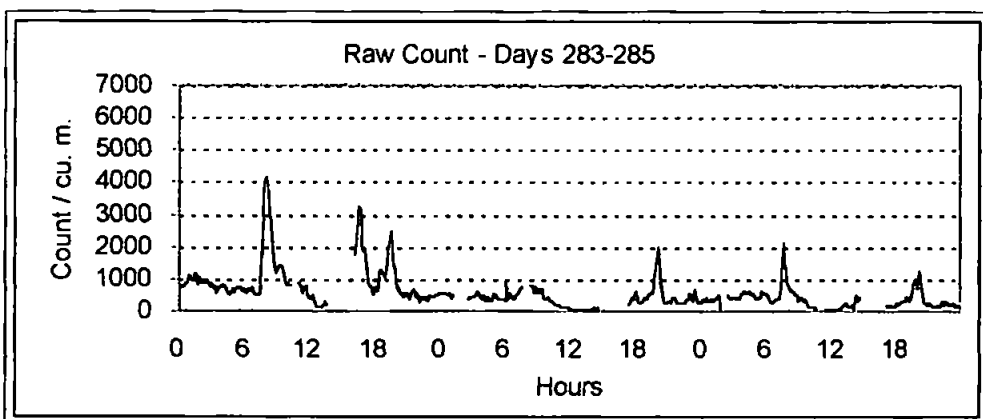
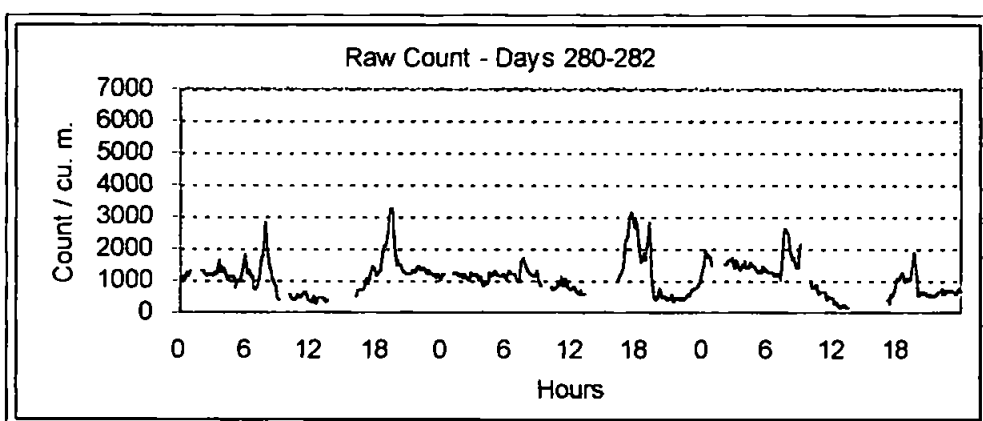
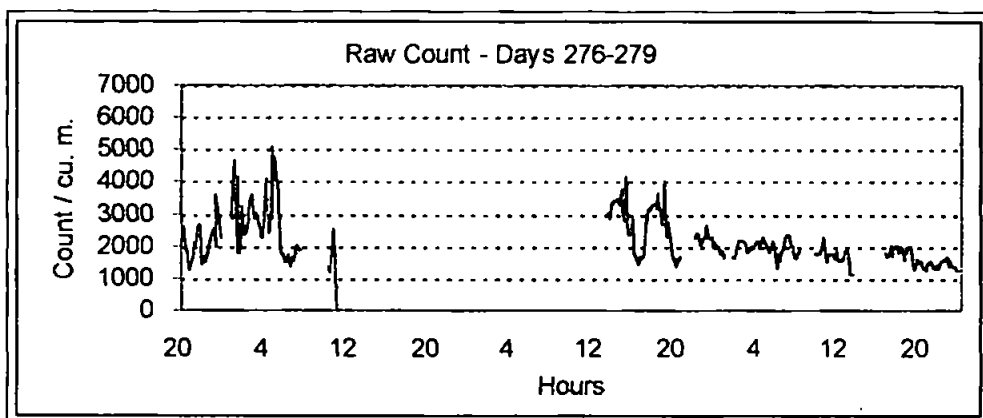


Fig. 8.4b AMT1 Raw counts along transect (underway),cont.

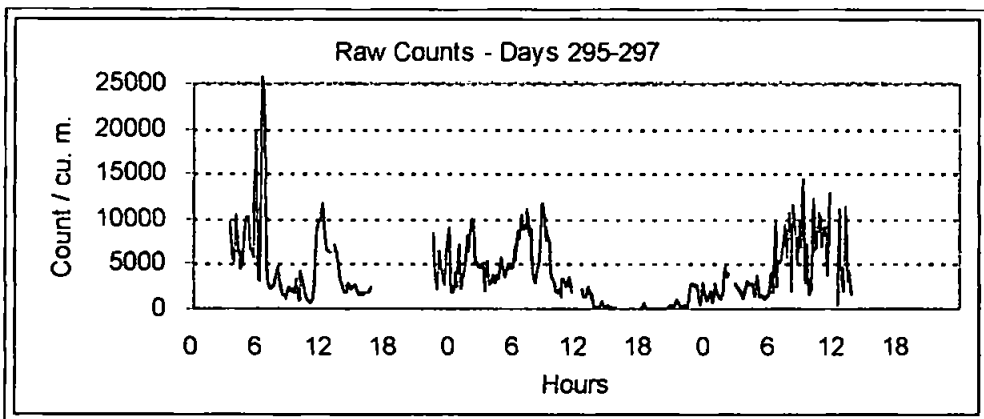
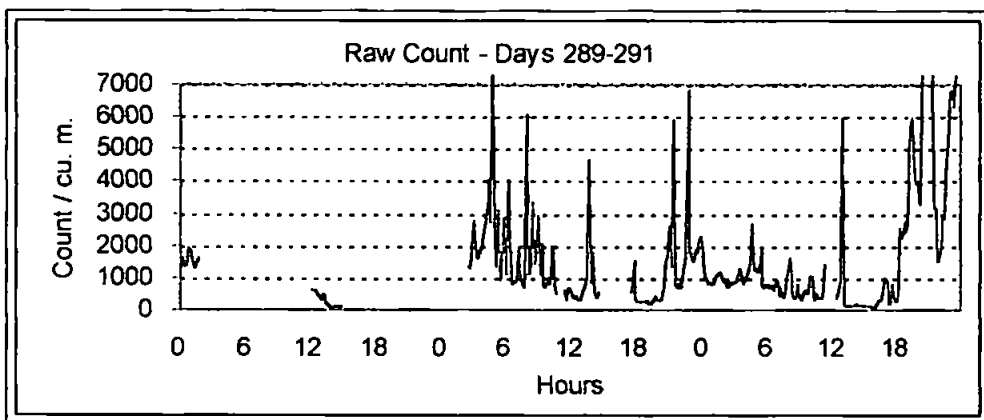
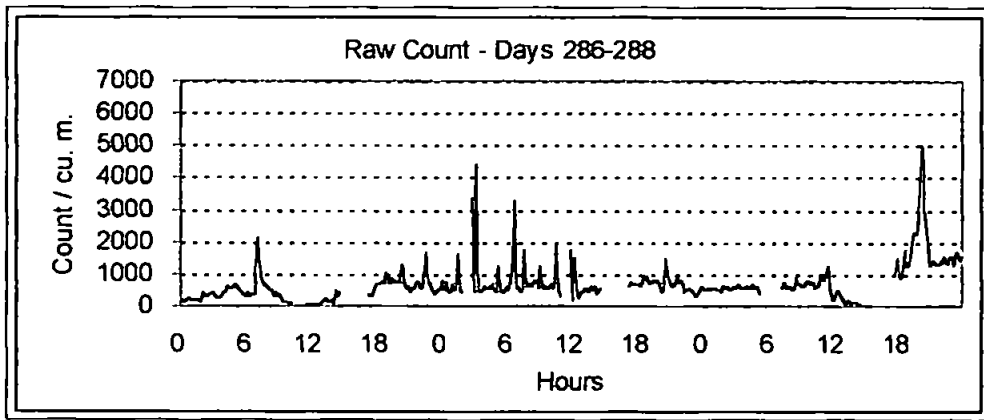
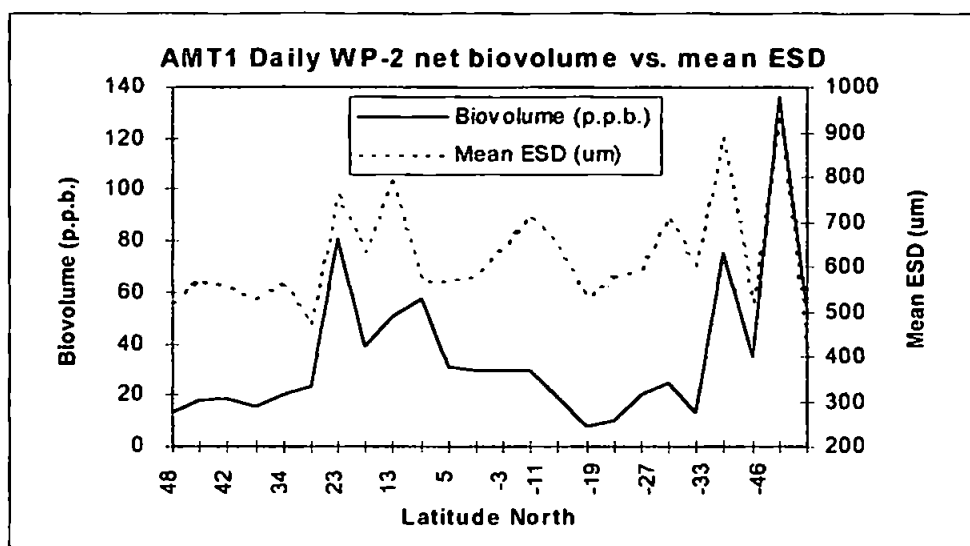


Fig. 8.4c AMT1 Raw counts along transect (underway), cont.



*Fig. 8.5 OPC Biovolume & mean ESD along transect (net hauls)*

Figure 8.5 shows the calculated daily vertical net profile biovolume for the transect, along with mean ESD. Biovolume was calculated as parts per billion (p.p.b.) by volume. The increased biomass around the West African upwelling system (c. 19° N), and the Brazil/Falkland current confluence (c. 45° S) is also clearly illustrated by this diagram. Figure 8.6 shows the net biovolume in four size classes, and fig. 8.7 shows the same distribution as percentage of total biovolume in each size class.

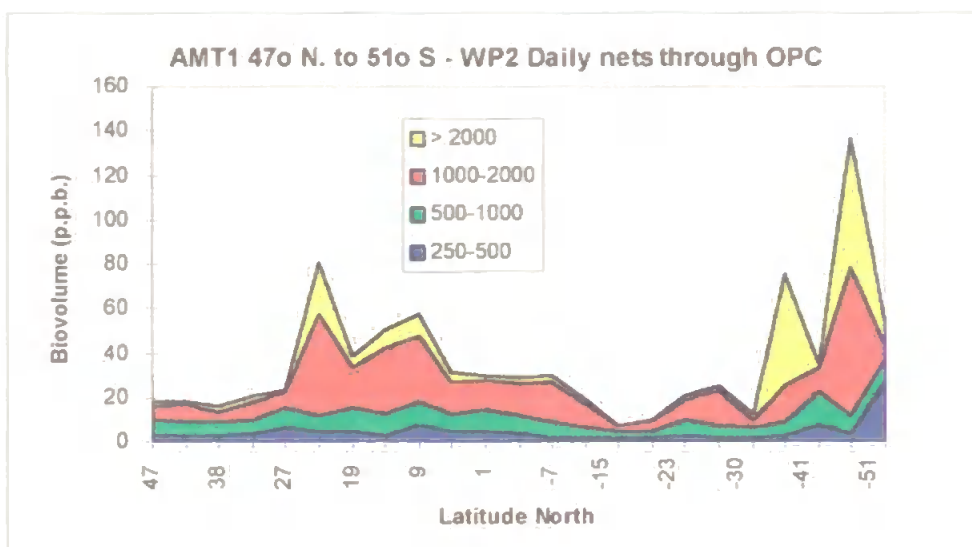


Fig. 8.6 Size distribution of biovolume along transect (net hauls) - all sizes in micrometres.

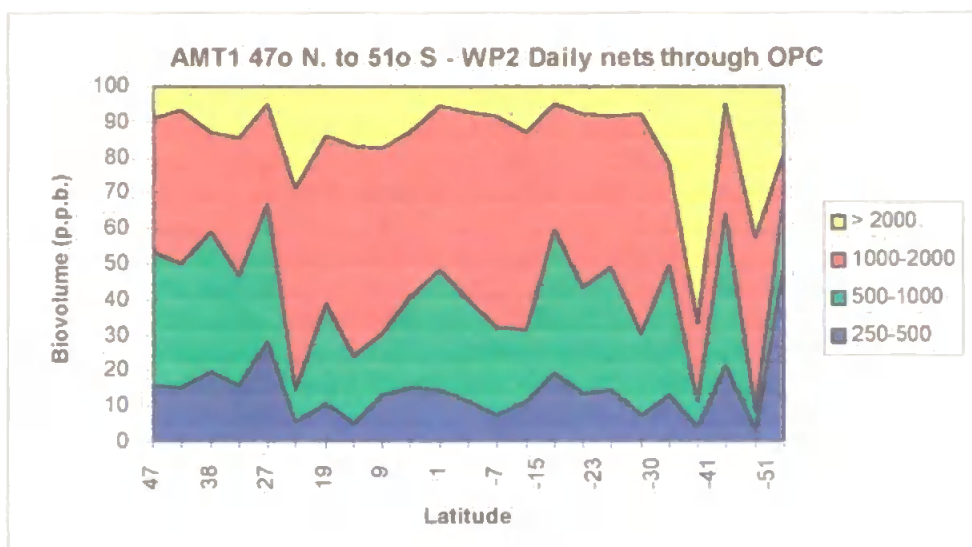


Fig. 8.7 Percentage of total biovolume in each size class along transect (net hauls) - all sizes in micrometres.

## **8.4 Atlantic Meridional Transect 2 (AMT2)**

### *8.4.1 Methods*

The OPC was deployed as for AMT1, both for continuous underway sampling from the ship's uncontaminated supply, and for processing samples from 200 m. integrated vertical net samples from daily stations. The seawater intakes had 6 mm. steel mesh filters, as on AMT1. A modified de-bubbling device with a larger cross-section was used, in the light of the limitations of the device used on AMT1.

The ViZA system was not used in its full configuration on this cruise, due to the fact that the frame processor card developed a fault just before the cruise, and had to be returned to the manufacturer for repair. The rest of the ViZA system was set up in line with the OPC and the video output recorded to tape, so that the data could be analysed subsequently. This analysis has not yet been completed, due to time constraints.

As on AMT1, at each daily station, WP-2 (200  $\mu$ m. mesh) net casts were made to 200 m. Each 200 m. sample was split using a Folsom splitter, half of the sample being passed through the OPC/ViZA system, and collected and preserved for subsequent microscopic taxonomy analysis.

The OPC was used in continuous flow-through mode during the whole cruise, using the uncontaminated seawater supply. This was interrupted only briefly at local dusk and dawn to change data files, and for about an hour once a day on

station to process the net samples. On days 126 to 134 a night net cast was also done, and the underway sampling was also interrupted for about one hour to process these samples. For about two hours each day a 200  $\mu\text{m}$ . mesh filter was connected to the flow outlet to collect the sample passed through the OPC. This was preserved for subsequent analysis to validate the OPC underway data.

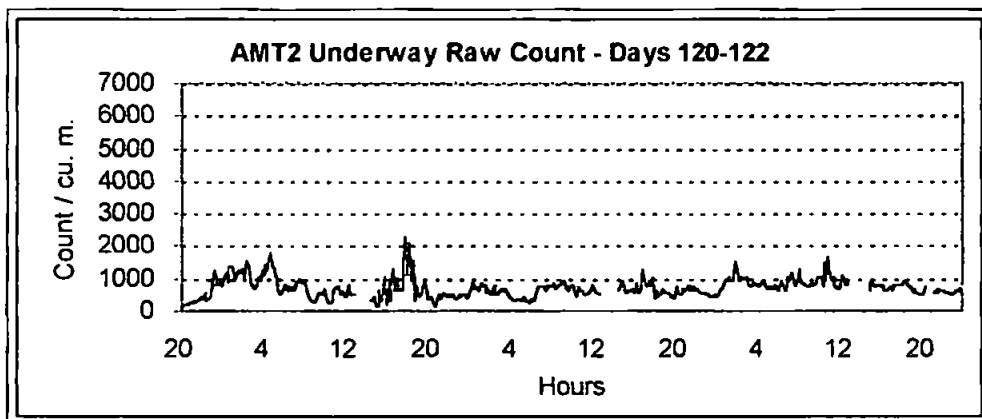
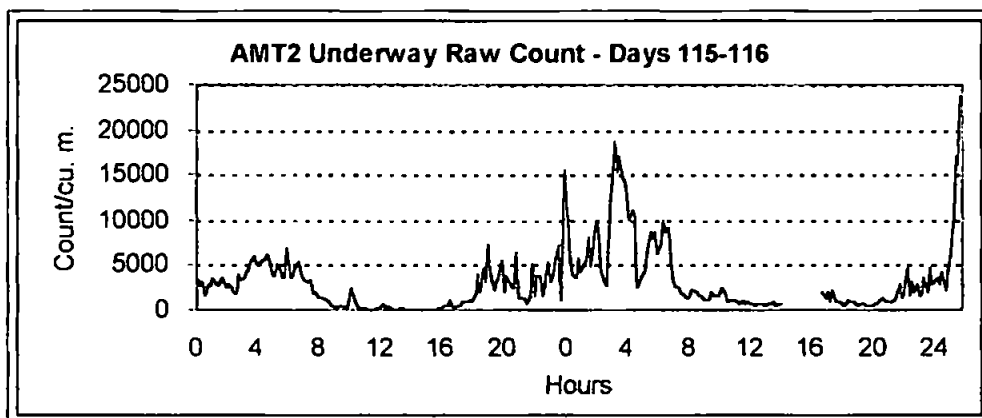
#### *8.4.2 Data results*

Figure 8.8 shows the underway raw OPC counts for the entire transect, from 50° south to the shelf edge off the southwestern approaches. The high counts during days 115/116 show the effects of frontal activity in this area, where temperature fluctuated through 6° C (Note different y axis scale on this first graph of fig. 8.8). A sharp increase is evident at about 5° north (day 131), probably showing the influence of equatorial divergence. A frontal system is indicated at about 13° north (day 133), where the temperature began falling and chlorophyll fluorescence increased. This position is coincident with the position indicated for the West African upwelling system by the net biovolume figures. Counts remain high for about 24 hours - to about 20° north, and are then lower until about 37° north. Between 37° north and 45° north counts are higher, and very much higher after this point. Two further points of interest are indicated by the raw counts shown on these figures:

- (i) Diurnal cycles in numbers are evident throughout the transect, but are particularly evident in the higher latitudes at either extreme, as was also

true of AMT1. On days 115-116 in particular, surface counts are very high at night, and relatively insignificant during the day.

- (ii) There is also evidence of large increases in counts at dusk and dawn, particularly between 20° south and 5° north (days 125 - 131). This phenomenon was also evident on AMT1, particularly between about 7° north and 16° south (see above)



*Fig. 8.8a AMT2 Raw counts along transect (underway).*

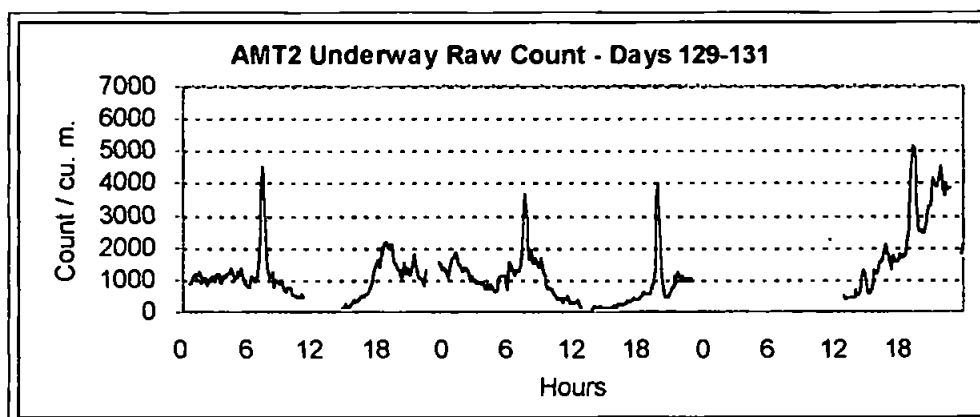
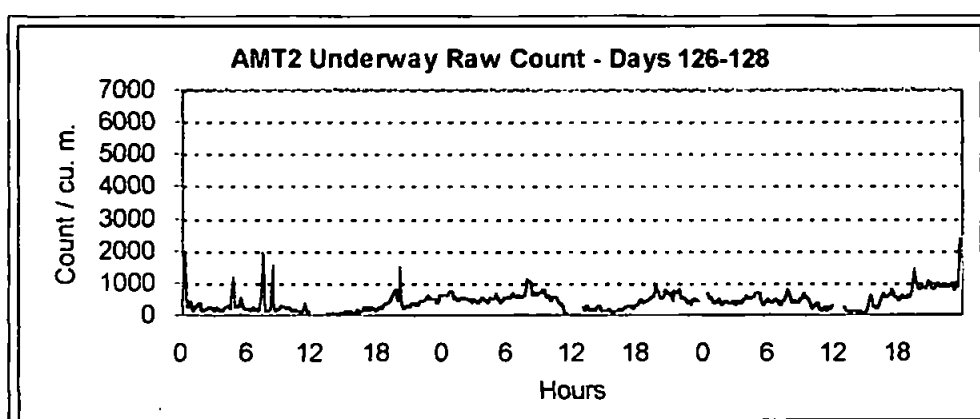
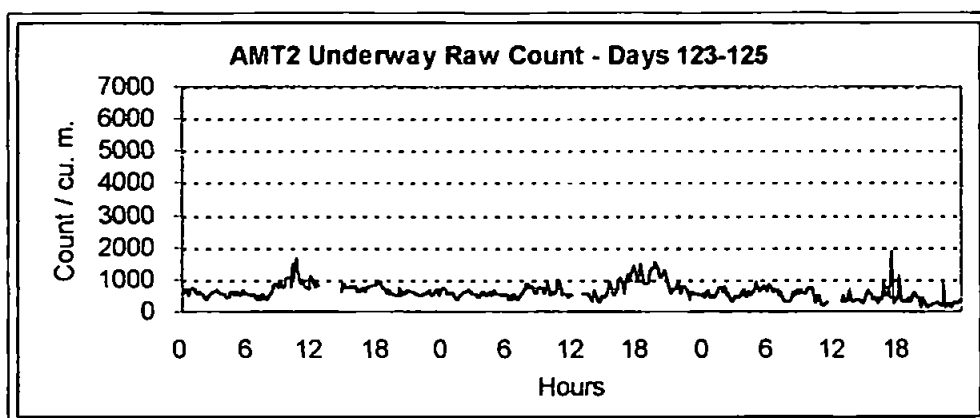
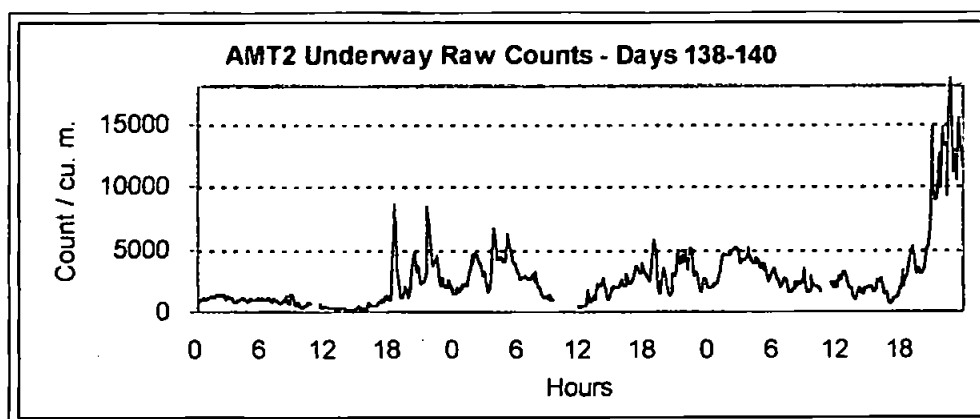
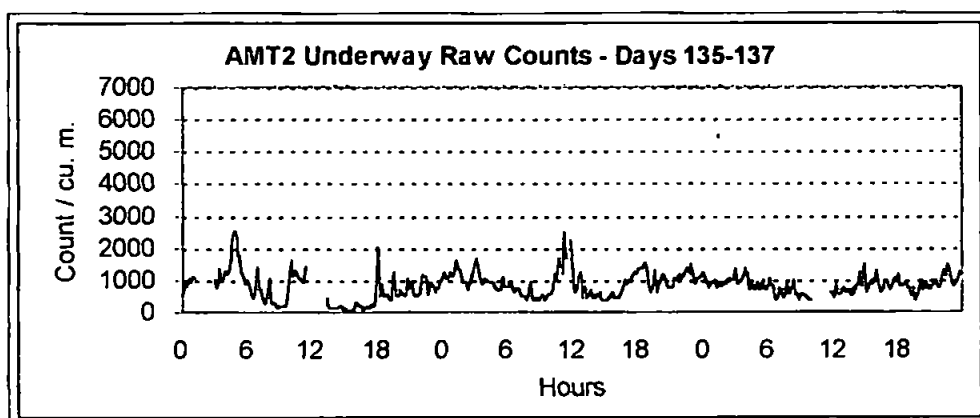
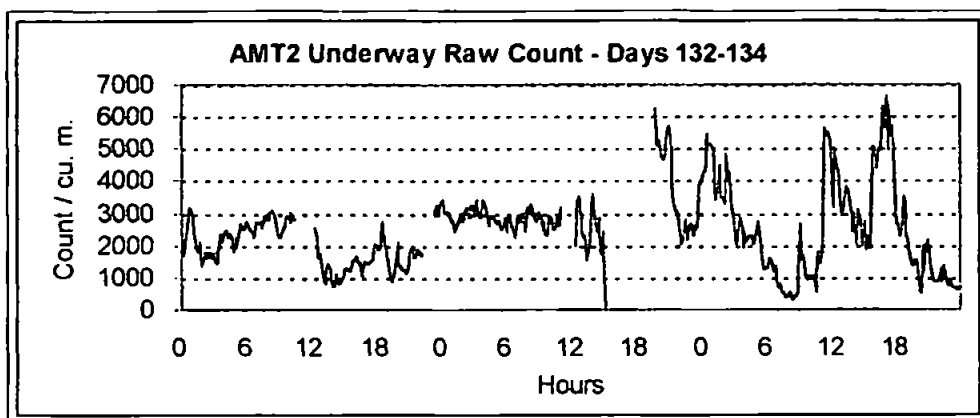
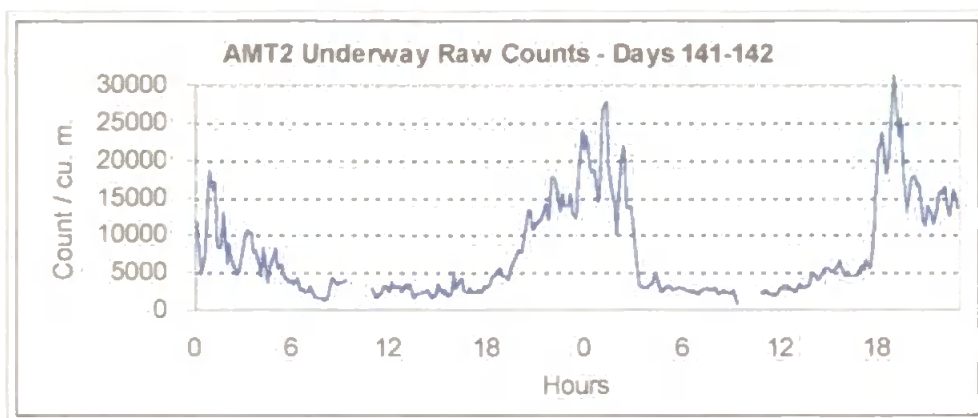


Fig. 8.8b AMT2 Raw counts along transect (underway, cont).

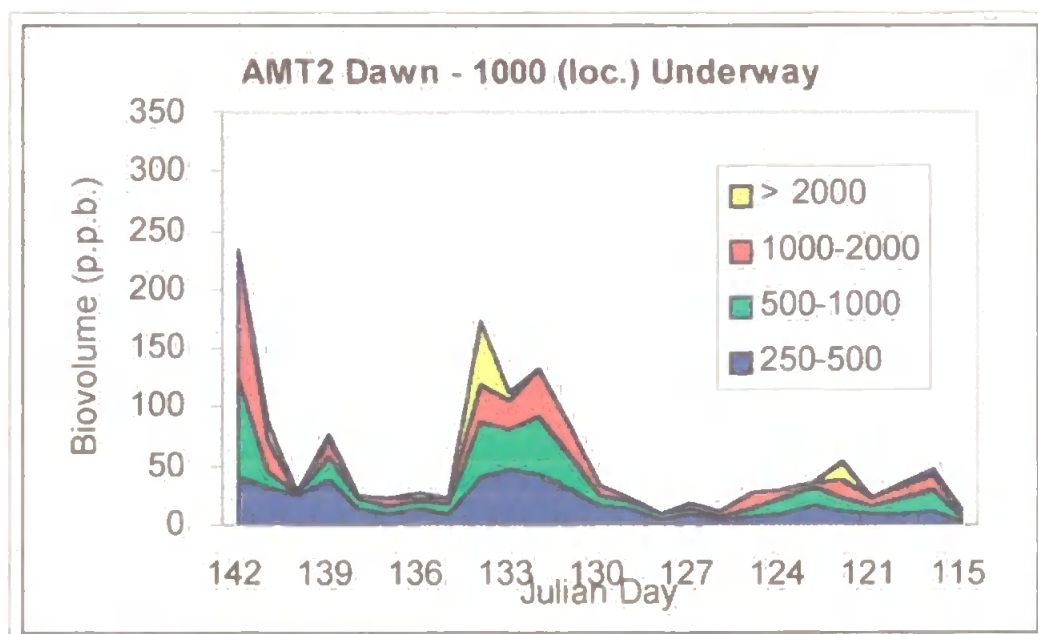


*Fig. 8.8c AMT2 Raw counts along transect (underway, cont.).*

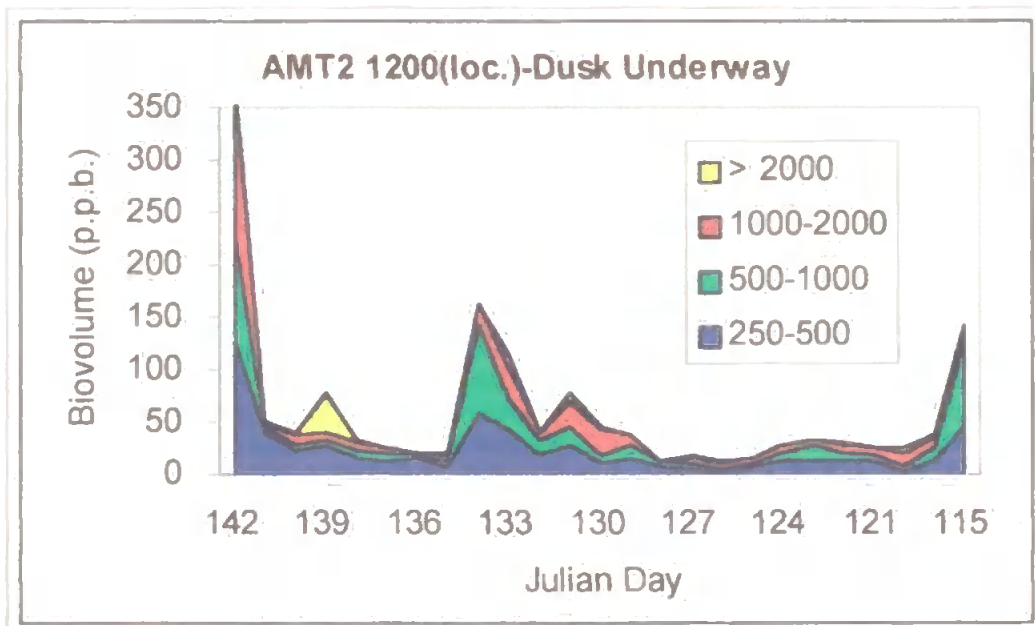


*Fig. 8.8d AMT2 Raw counts along transect (underway, cont.).*

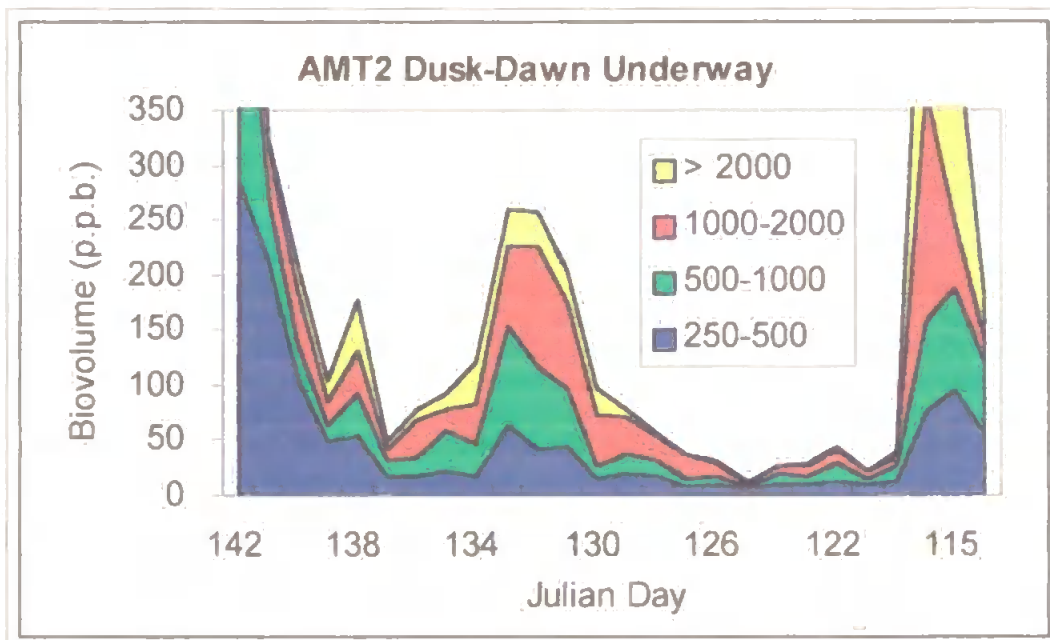
Figures 8.9 - 8.11 show the size fractionated underway OPC biovolume for the whole transect, separated into the three main daily sampling periods.



*Fig. 8.9 Size distribution of underway biovolume along transect for sampling period local dawn to 1000Z - all sizes in micrometres. .*



*Fig. 8.10 Size distribution of underway biovolume along transect for sampling period 1200Z to local dusk - all sizes in micrometres .*



*Fig. 8.11 Size distribution of underway biovolume along transect for sampling period local dusk to local dawn - all sizes in micrometres .*

Figure 8.12 shows the vertical net profile biovolume for the whole cruise, and the mean ESD for comparison. Note that as in AMT1 the mean ESD changes with biovolume except around fronts associated with upwelling zones (4° N).

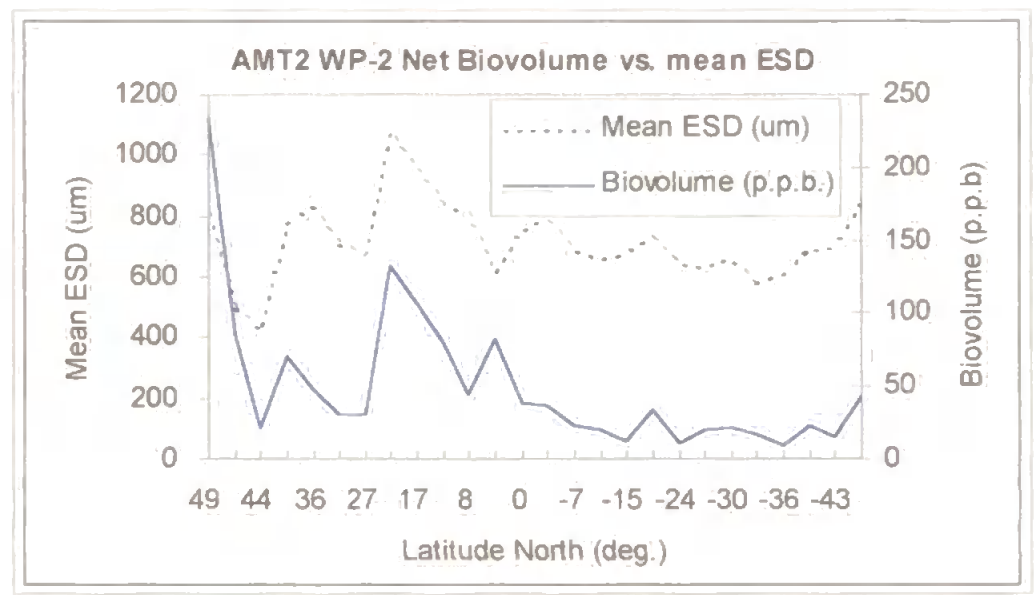


Fig. 8.12 Biovolume & mean ESD along transect (net hauls).

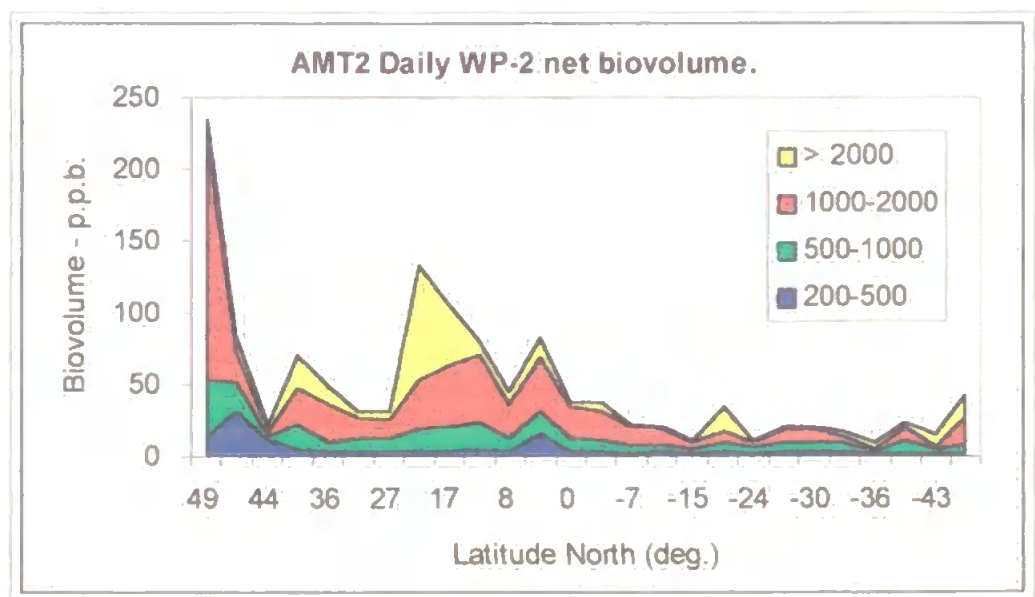
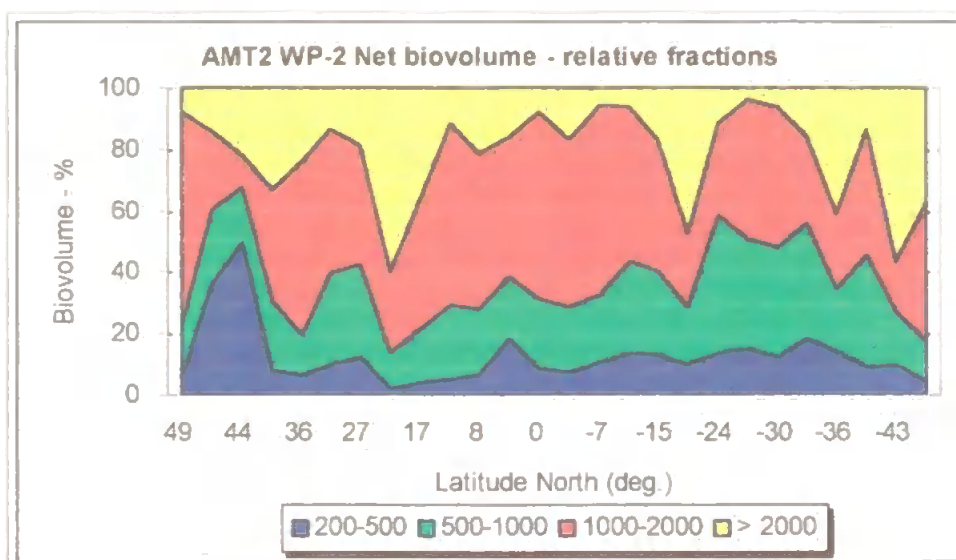


Fig. 8.13 Size distribution of biovolume along transect (net hauls) - all sizes in micrometres.



*Fig. 8.14 Percentage of total biovolume in each size class along transect (net hauls).*

Figure 8.13 shows the biovolume in parts per billion in each of the four JGOFS size classes for each daily 200 m. net cast throughout the cruise. Figure 8.14 shows the same data but with each size fraction as a percentage of total biovolume. Biovolume was generally low throughout the transect south of the equator, as might be expected in the late austral autumn. There is a general increase as the equator is approached, with a peak at 4° north. After a fall from this level, the biovolume rises again through the West African upwelling area, between about 13 and 22° north. Biovolume then falls away, but to a level generally higher than that south of the equator, until the very large increase around 45° north, 16° west as a more coastal influence is felt in what is probably the tail of the northern spring bloom.

## 8.5 PRIME North Atlantic Cruise

The OPC was deployed on the second leg of the PRIME cruise, departing Reykjavik on 2nd July 1996 and returning to Southampton on 23rd July 1996.

### 8.5.1 Methods

Two uncontaminated supplies were available aboard RRS Discovery - one direct from the sea water intake, and one de-aerated, via a header tank. The direct supply was originally chosen to avoid any effects of the header tank, but the outlet was restricted to about 6 mm. bore, and provided insufficient flow for the OPC. Ship's technicians advised that this could not easily be changed. The OPC and ViZA systems were then connected to the de-aerated supply, whose bore was about 12 mm., and whose flow was marginally better (*circa* 10 - 12 litres per minute), and just sufficient for the OPC.

During the transect from 60° north to 37° north two WP-2 net casts were made to 20 m and 200 m at night, around midnight, and the same during the day, around noon. On longer stations (60° north, 46° north and 37° north), a 100 m cast was added for compatibility with zooplankton work on the first leg of the cruise. The 100 m. sample was not passed through the OPC/Video system, however. Sampling naturally fell into three domains: the transect from 60° to 37° north, the 8 day station at 37° North, and the run home. The data set and its presentation are separated along these lines.

### *The Transect to 37° N*

Vertical net casts were made at approximately midnight and midday. Underway sampling was split into four sessions a day, i.e. the night run was split due to the midnight station. Generally the times of the four runs were:

c. 0200 - local dawn

local dawn - c. 1230

c. 1400 - local dusk

local dusk - c. 0030

### *The Station at 37° N*

Sampling was as for stations on transect. This was an 8 day station, nominally at 37° N, 19° W. Deviation from this position was less than 30 nm. for all net casts.

### *The homeward leg:*

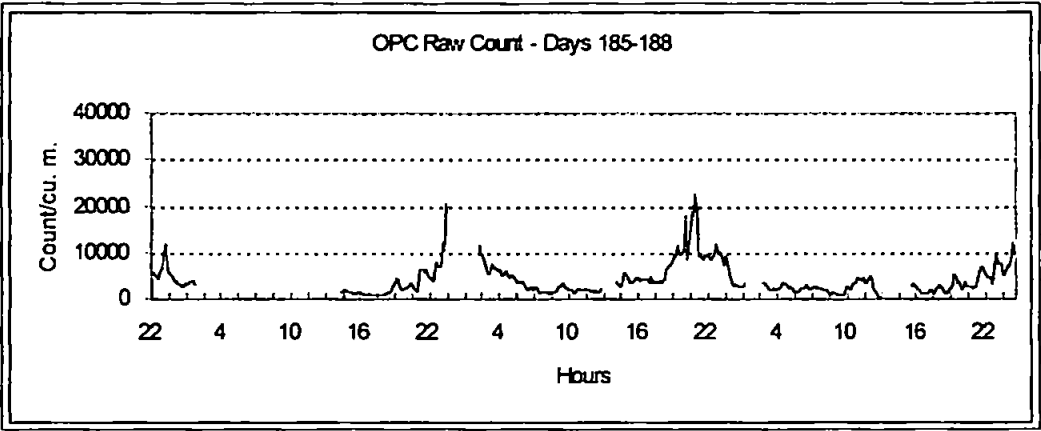
For the three days of the run home to Southampton from 37° North, the OPC and video system were operated in-line continuously. No net hauls were done on the homeward leg.

## **8.5.2 OPC Data results**

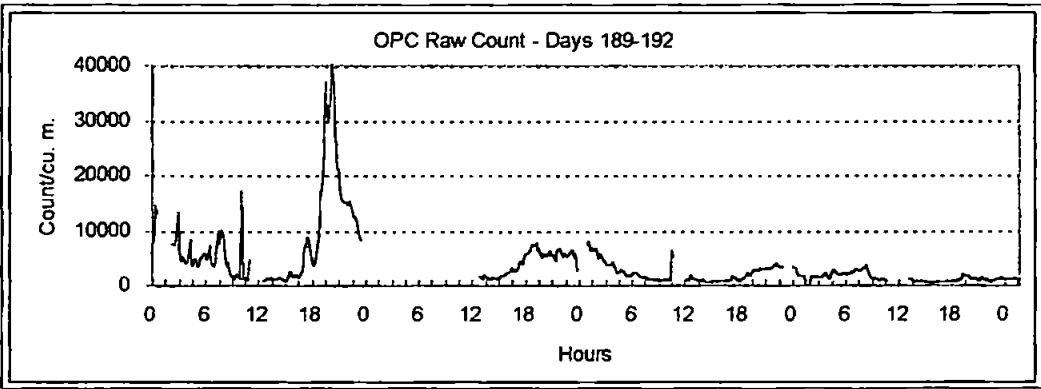
### *Transect underway sampling*

Data is presented as for AMTs in the graphs below. These comprise continuous plots of raw particle counts per cubic metre against time on the transect from 60° north to 37° north (figs. 8.15 a&b), and underway biovolume partitioned

into the four JGOFS fractions and plotted for each of three daily sampling periods against transect time from 60 North to 37 North (figs. 8.16 to 8.18).



*Fig. 8.15a PRIME transect underway raw counts - days 185-188.*



*Fig. 8.15b PRIME underway raw counts - days 188-192.*

A general trend from higher to lower biomass can be seen on each of these plots, consistently interrupted by a rise in biovolume on day 189. This can also be seen on the plot of raw counts as a very high peak of over 40,000 per cubic metre on day 189, higher than anything seen on AMT. Physical and nutrients data confirms that some kind of frontal zone was crossed the day before (c.

1145Z on day 188, 6th July, c. 52° 26' North). Generally, biovolumes and raw particle counts are greater than at the top of the AMTI transect, falling away to similar levels as we approached 46° North (1-4000 m<sup>-3</sup>).

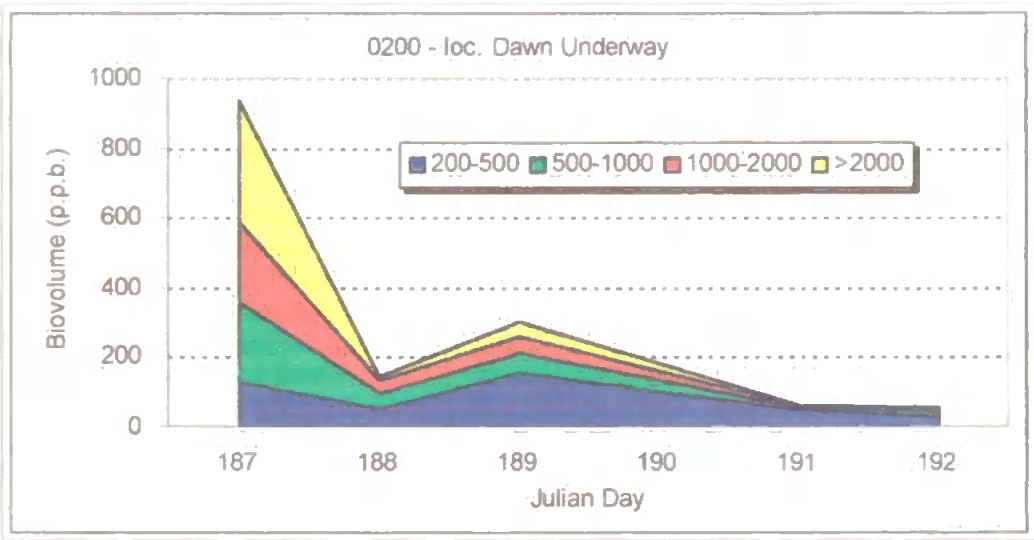


Fig. 8.16 Size distribution of underway biovolume along transect for sampling period 0200 to local dawn - all sizes in micrometres.

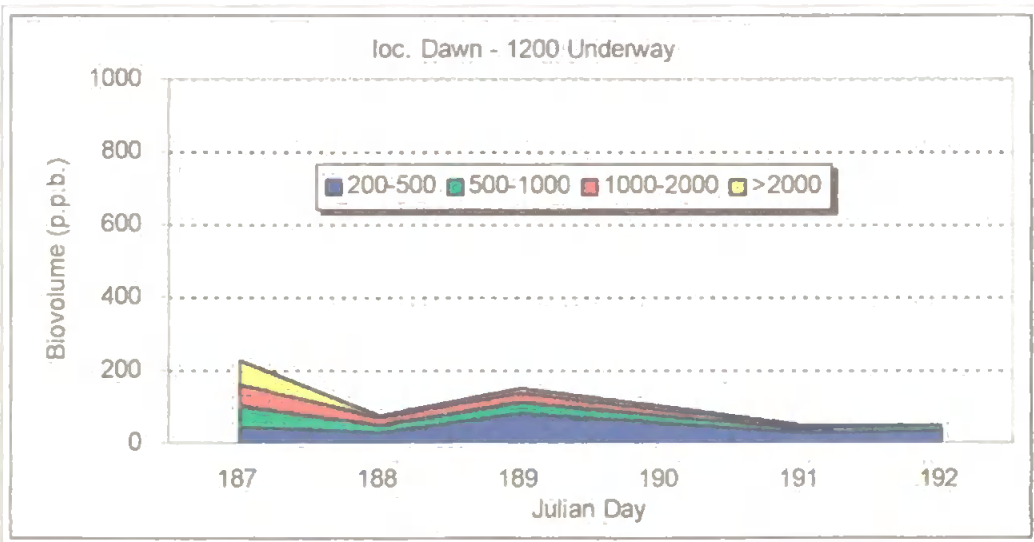


Fig. 8.17 Size distribution for underway biovolume along transect for sampling period local dawn to 1200Z - all sizes in micrometres.

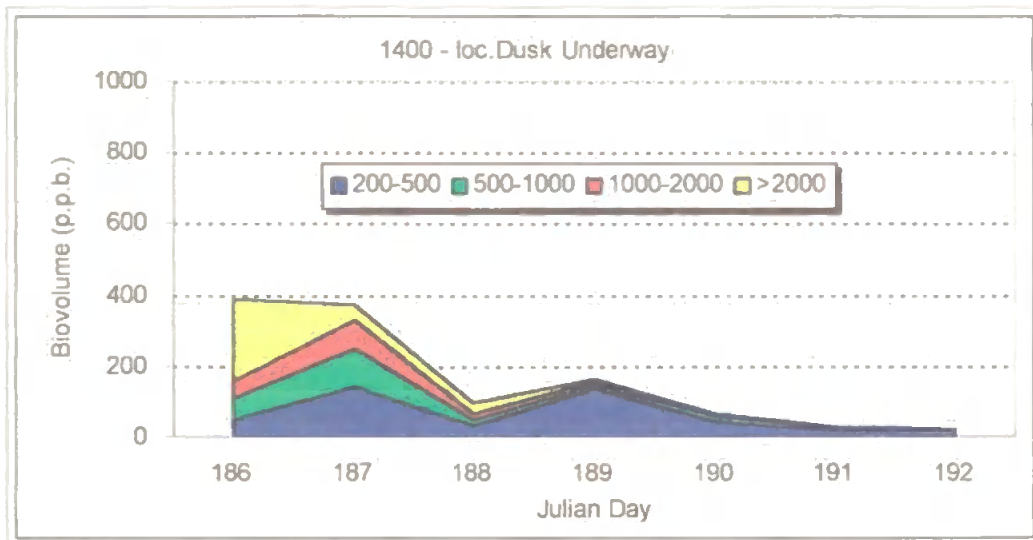


Fig. 8.18 Size distribution for underway biovolume along transect for sampling period 1400Z to local dusk - all sizes in micrometres.

#### Homeward underway sampling

The underway data for the homeward leg are presented as raw count per cubic metre versus time (fig. 8.19), and as a day/night series of biovolume in the four JGOFS size classes versus time (fig. 8.20, day/night variation apparent).

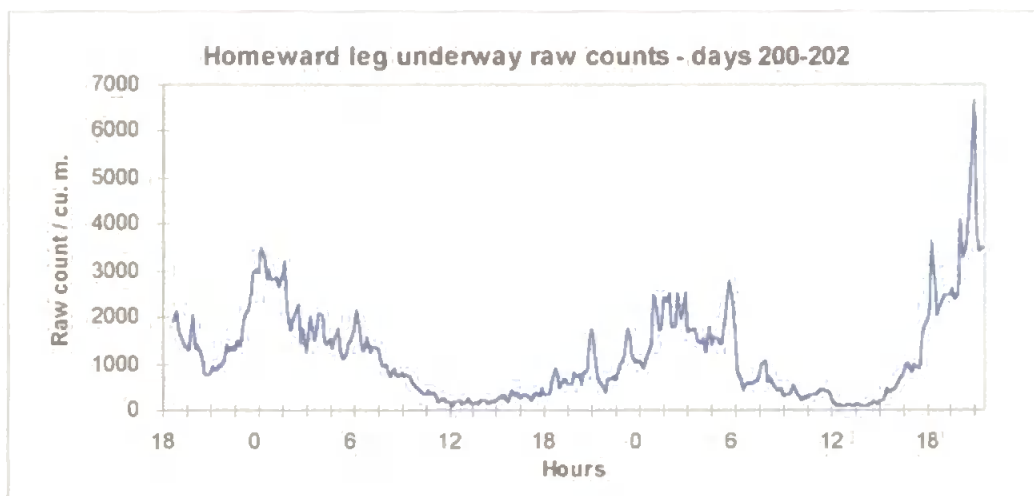
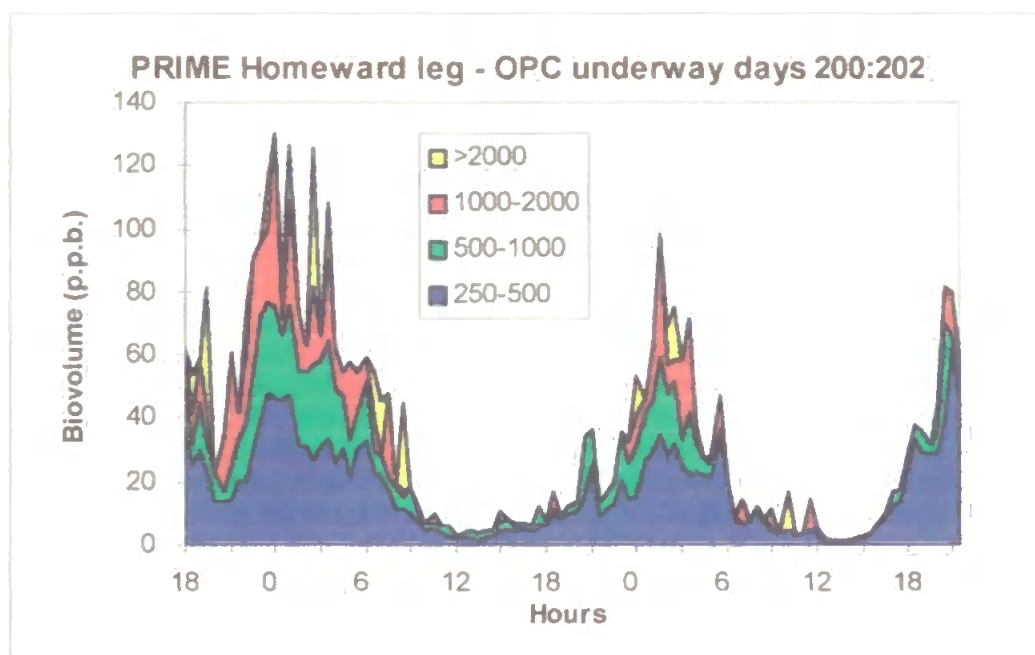


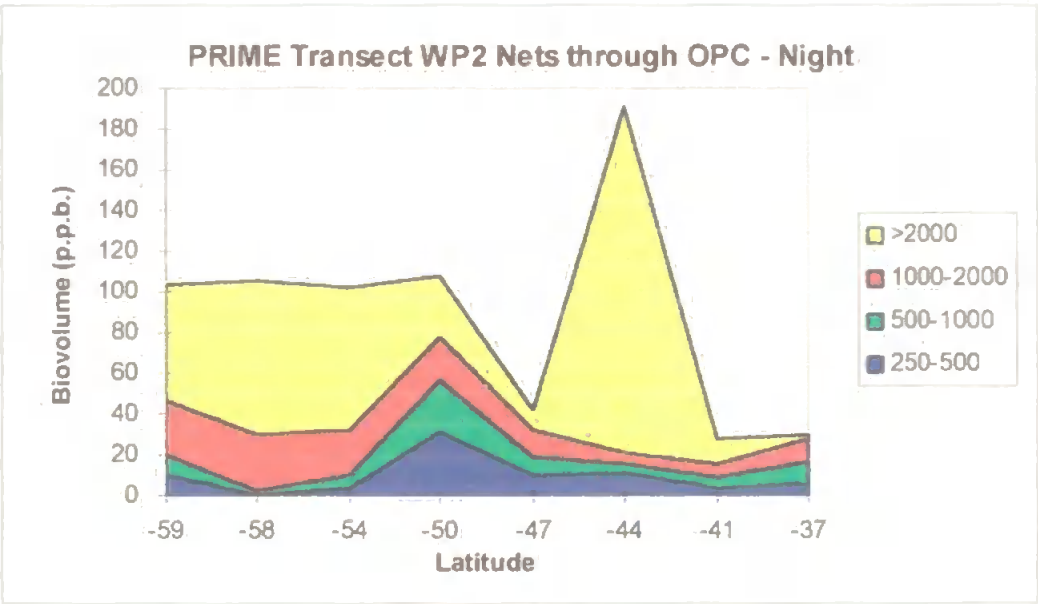
Fig. 8.19 PRIME homeward leg underway raw counts - integrated over thirty minute intervals



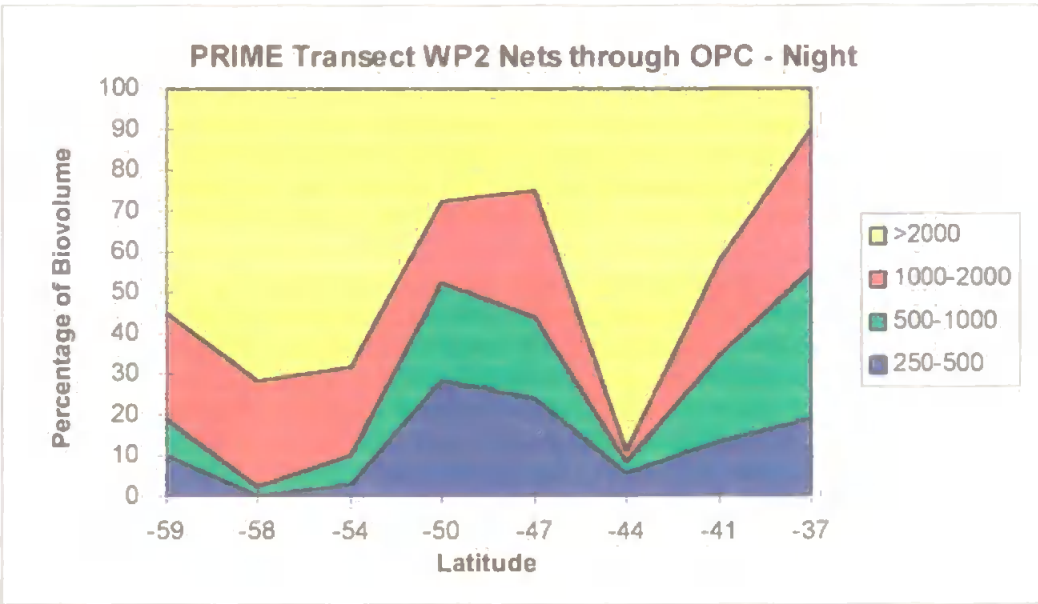
*Fig. 8.20 Size distribution of underway biovolume along homeward leg - integrated over thirty minute intervals, all sizes in micrometres*

#### *Vertical net sampling*

WP-2 net casts were made to 200 m. and 20 m. at around 0030 and 1200 hrs. during transect and on longer stations. Half of the 200 m. net and the whole 20m. net was passed through and preserved, as on AMTs. Data appear in graph form below. Vertical net samples (day) show generally rather lower values around 46° North than comparable AMT1 samples. North of the frontal zone, the net samples are dominated by the largest size fraction - very often due to gelatinous organisms. From this area to 37° north, apart from at 45° north, smaller classes dominate. Size distributions are typical of AMT1 in this area.



*Fig. 8.21 Size distribution of biovolume along transect - night nets, all sizes in micrometres.*



*Fig. 8.22 Percentage of total biovolume in each size class along transect - night nets, all sizes in micrometres.*

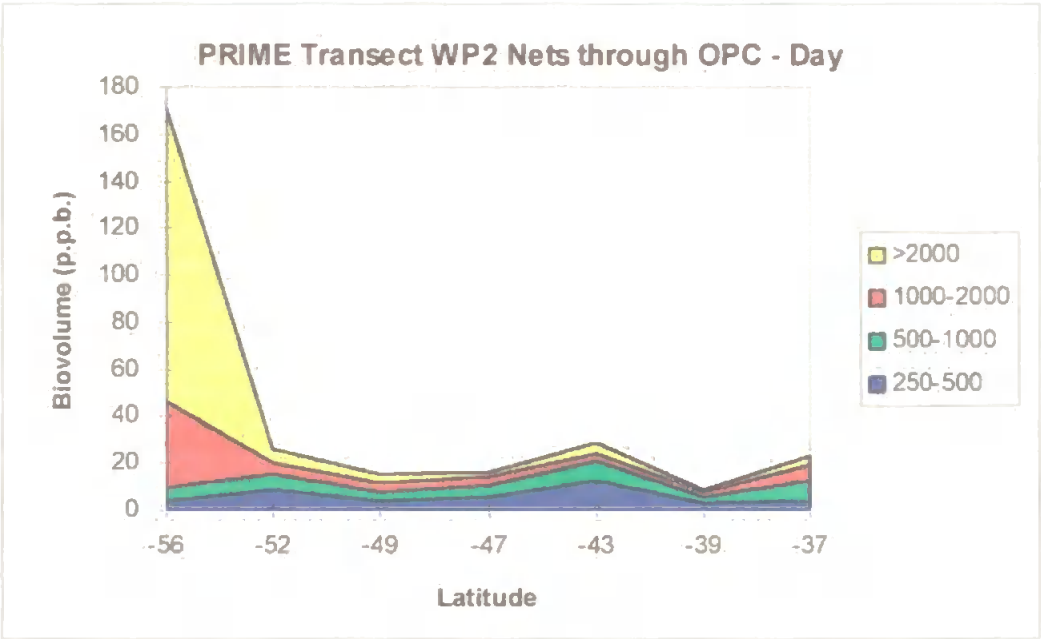


Fig. 8.23 Size distribution of biovolume along transect - day nets, all sizes in micrometres.

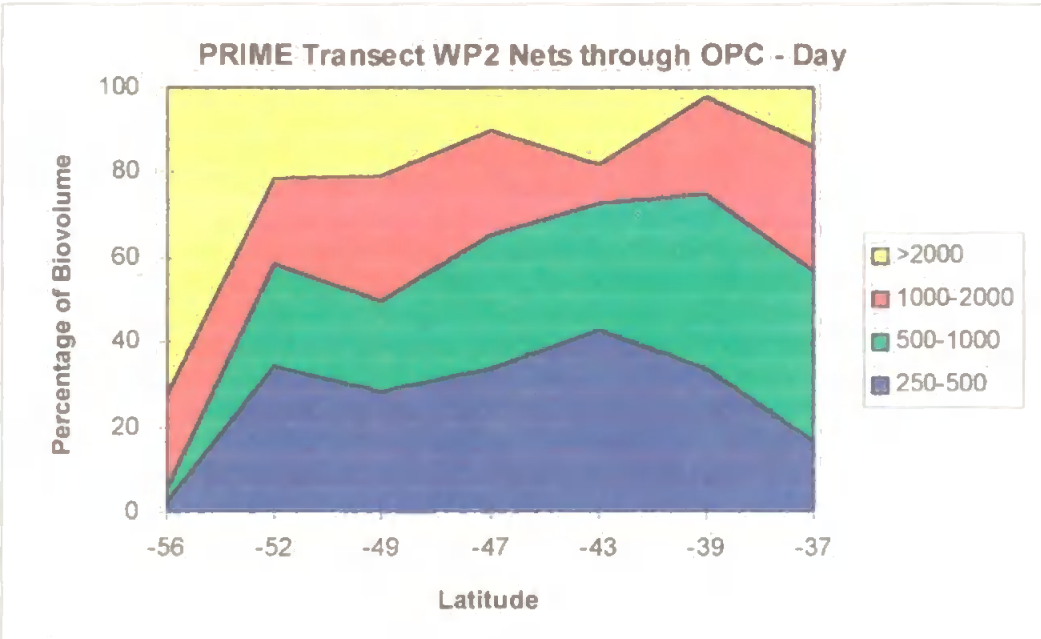
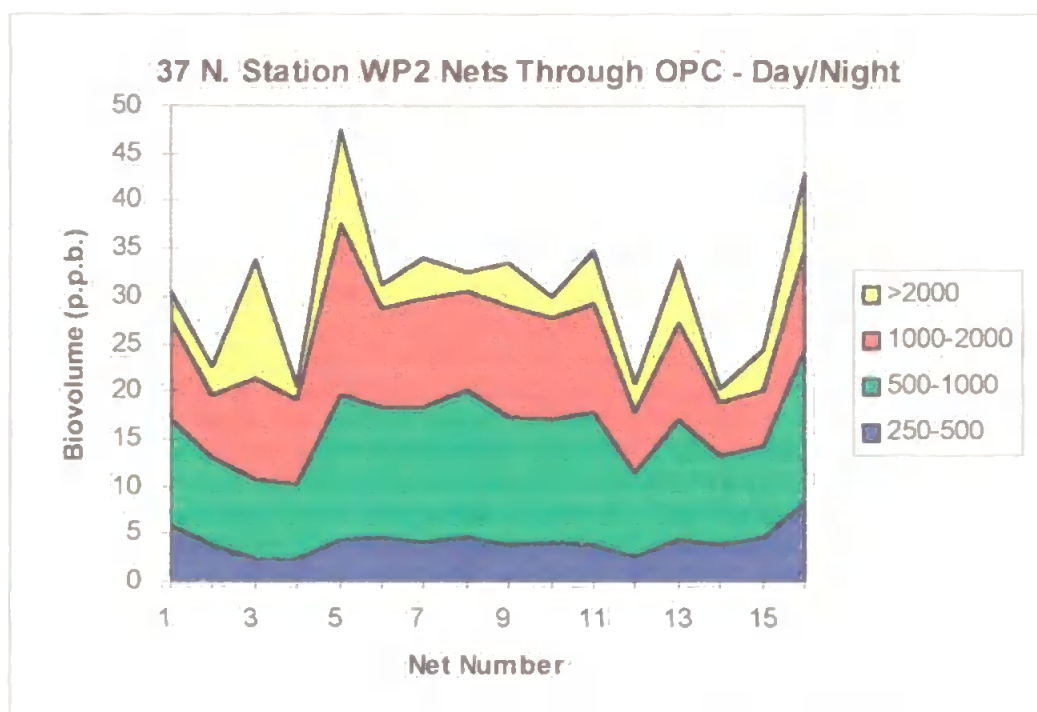


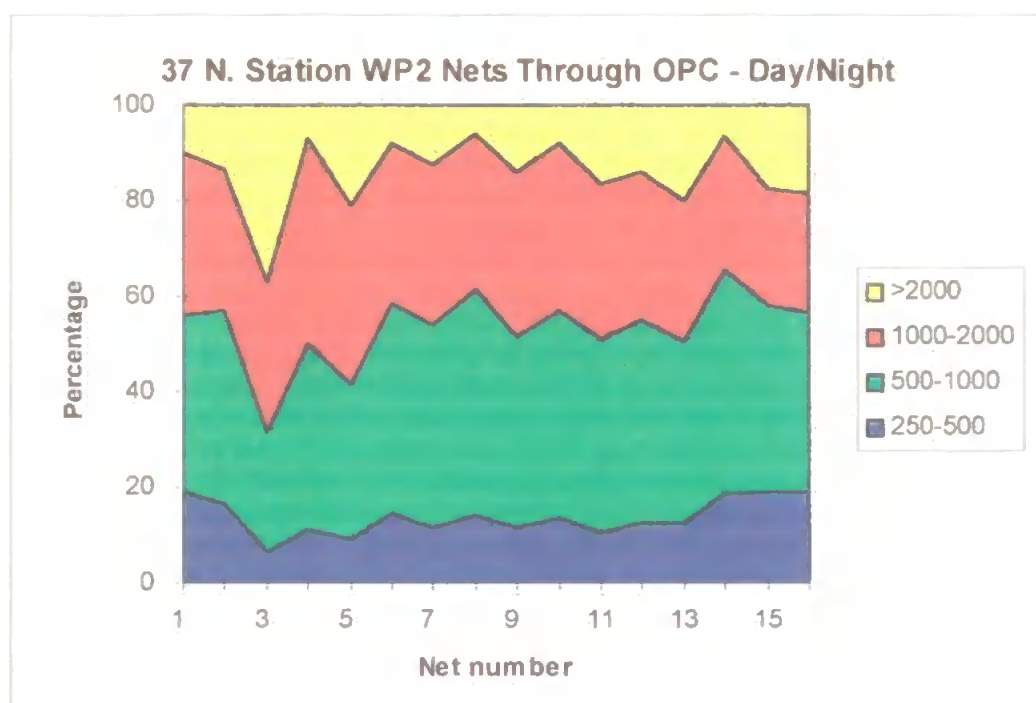
Fig. 8.24 Percentage of total biovolume in each size class along transect - day nets, all sizes in micrometres.

*Station at 37° north*

These data appear as a temporal series for the eight days on a single plot. Some variability is shown in total biovolume and in day/night comparisons for the first three and last two days (fig. 8.25). The middle three days show little variability or diurnal change. Relative size fractions (fig. 8.26) remained fairly constant throughout (apart from net 3), with small amounts of diurnal variation.



*Fig. 8.25 size distribution of biovolume on 8-day station - day/night nets, all sizes in micrometres.*



*Fig. 8.26 Percentage of biovolume in each size class on station day/night nets, all sizes in micrometres.*

### 8.5.3 ViZA Data results

This was the first occasion upon which it has been possible to use the video device in pump through mode at sea. It is not surprising that much of the early part of the cruise was spent in dealing with problems which had not arisen or been foreseen at the start of the cruise. The system was set up with the flow cell immediately downstream of the OPC, and the strobe light triggered from the video camera. Several problems quickly became apparent, as outlined below:

- (i) The magnification was too low. Set for a field of view to match the flow cell size (20 mm.), the numerical predominance of small animals (< 500

$\mu\text{m}$ ) meant that those animals which could be sampled in sufficient numbers had very poor shape definition.

*Solution:* New mounting points were made for the flow cell, bringing it closer to the camera lens, and reducing the field of view to 12 mm., and the pixel size from  $43\ \mu\text{m}$  to  $25\ \mu\text{m}$  (this has of course exacerbated the sampling rate problem, see below). Eventually two cameras of different magnifications, and concentric fields of view will solve this problem.

- (ii) At the camera frame rate of  $60\ \text{sec}^{-1}$ , the strobe was being overdriven. This resulted in overheating and thermal cut-out. It also resulted in double images of particles, showing that some kind of double trigger was present. It was found that the capacitor fitted was for 20 Hz. rather than 60 Hz. operation.

*Solution:* Temporarily, the strobe was used in 'test' mode, which means that it was triggered internally at 20 Hz. The lack of synchronism with the camera meant that some of the frames captured were not illuminated. A capacitor of correct value was subsequently fitted.

- (iii) The camera's AGC (automatic gain control - exposure) system tried to compensate for the rapidly changing light levels - this resulted in

different lighting conditions for each frame. Manual AGC setting was preset too low to gain a well lit image.

*Solution:* Strobe output increased to max., and the problem endured. Upon returning home, a manual gain adjustment system was built (see chapter 5).

- (iv) The sampling rate was too low to capture a statistically significant number of animals in underway mode. Only those animals below 250  $\mu\text{m}$ . ESD were present in sufficient numbers to be adequately sampled, and these were not large enough to obtain meaningful shape information. It is estimated that we need approximately a tenfold increase in sampling rate.

*Solution:* Profiling the program showed that most of the system's time was spent waiting for a return from the call to obtain the next contour from the current frame. As this subroutine is part of the proprietary Coreco software library, and is executed on board the F64 card, it was difficult to see how to improve it. The correct strobe rate would already improve the sampling rate by at least a factor of two. Combined with subsequent replacement of these subroutine calls by direct assembly language commands, and the histogram test for empty frames, this considerably improved the sampling rate by a factor of about 8, see chapter 7).

Despite these problems, the first deployment of the video system was encouraging. When sufficient animals were present, the system showed it could discriminate quite well between animals from 5 or 6 different taxonomic groups. Copepods were identified particularly reliably, because of the distinctive shape of the thorax. The problems identified above meant that useful data only became available towards the end of the station at 37° North. This data is therefore quite limited in extent, and biased to the smaller size fractions because of the sampling rate problem. Comparisons with the size distributions produced by the OPC would therefore not be meaningful. This deployment, however, provided a good starting point for further improvements.

## **8.6 Atlantic Meridional Transect 3 (AMT3)**

RRS James Clark Ross departed Grimsby to begin AMT3 on 19th September 1996, and arrived in the Falkland Islands on 25th October the same year.

### *8.6.1 Methods*

The OPC was deployed as for AMTs 1 & 2, both for continuous underway sampling from the ship's uncontaminated supply, and for processing samples from 200 m. integrated vertical net profiles from daily stations. The seawater intakes had 6 mm. steel mesh filters, as on AMTs 1 & 2. The ViZA system was deployed as on the PRIME cruise, after modifications in the light of the problems outlined above. ViZA was used continuously in underway mode, and also for the processing of each of the vertical net samples, the sample water passing through the OPC and ViZA flow cells in series. A burst water

connection flooded the video camera, causing some damage on day 275. The initial damage was successfully repaired by day 278, when ViZA sampling continued, but image quality gradually deteriorated until day 282 when the camera failed completely. Fourteen days of ViZA data were therefore available for analysis.

At each daily station, three WP-2 net casts were made. The first, using a double net, to 200 m., the second a single net to 200 m. and the third a single 20 m. net. The sample from the double net was used for gut evacuation experiments. The single 200m. net was split and processed as for AMT 1 & 2, using a Folsom splitter. Half of the sample was passed through the OPC/ViZA system, and collected and preserved for subsequent microscopic taxonomy analysis.

The OPC/ViZA system was used in continuous flow-through mode during the whole cruise, using the uncontaminated seawater supply. This was interrupted only briefly at local dusk and dawn to change data files, and for about two hours each day on station to process the net samples. For about an hour on each of seven days a 200  $\mu$ m. mesh filter was connected to the flow outlet to collect the sample passed through the OPC/ViZA system. This was preserved for subsequent microscopic analysis to validate the OPC/ViZA system data.

8.6.2 Data results

Figure 8.27 shows the zooplankton abundance in raw particle counts per cubic metre in underway mode for the whole cruise. High counts during days 266-269 (note difference in y axis scale in this graph) show the influence of the coastal environment, declining to much lower counts in the ‘blue’ waters of the Canary Basin on days 270-272. Counts are higher again across the West African upwelling area, and the equator (day 279 in fig. 8.27b). Counts are very low on days 280 to 287 across the Brazil Basin, and increase again in the more coastal environment approaching the River Plate estuary on days 288 - 290.

As on previous AMT cruises, there is evidence of increases in counts at dusk and dawn, particularly between 5° north and 20° south (days 278 - 284), which may to offer some support to the theory of ‘midnight sinking’.

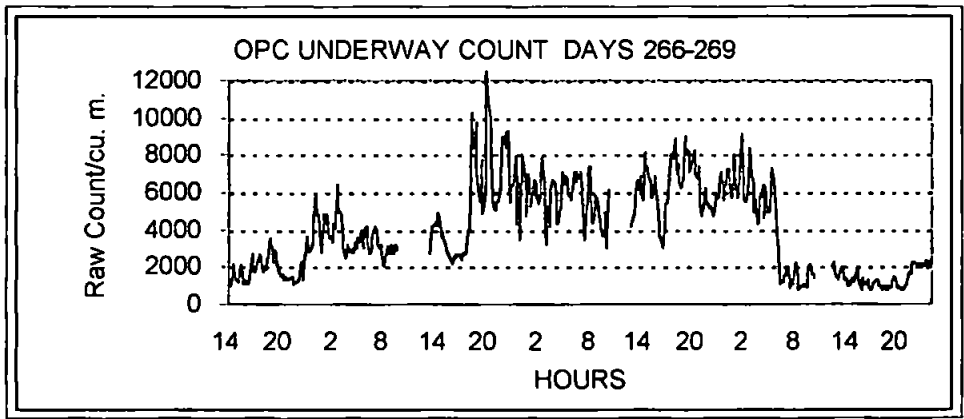


Fig. 8.27a AMT3 Raw counts along transect (underway).

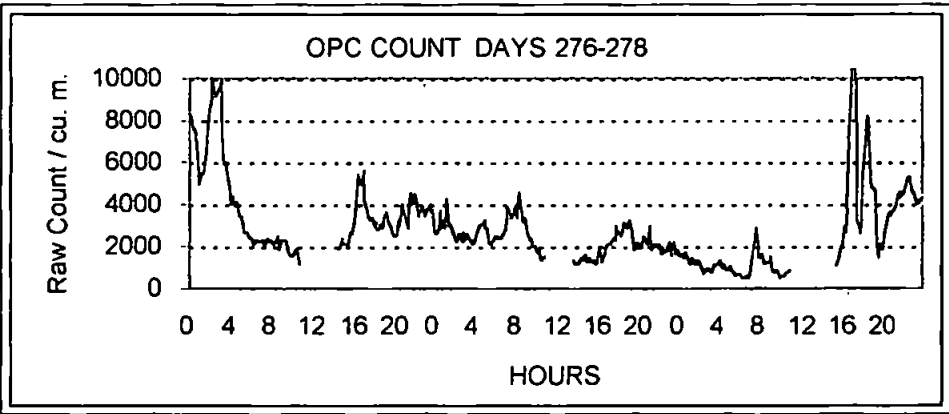
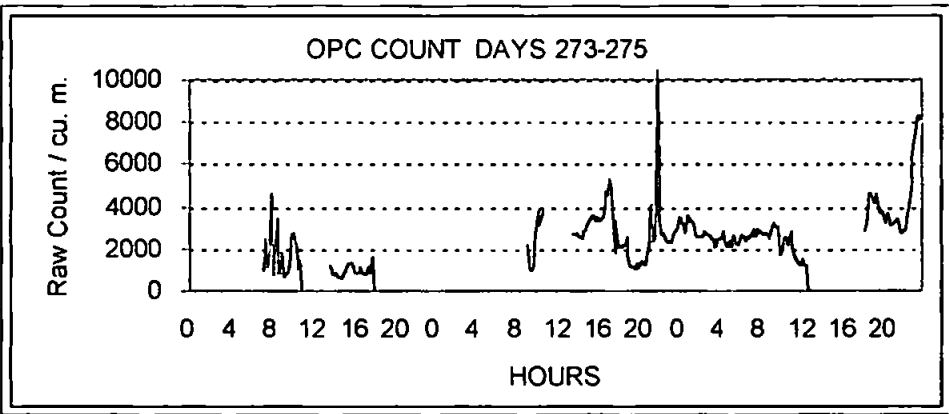
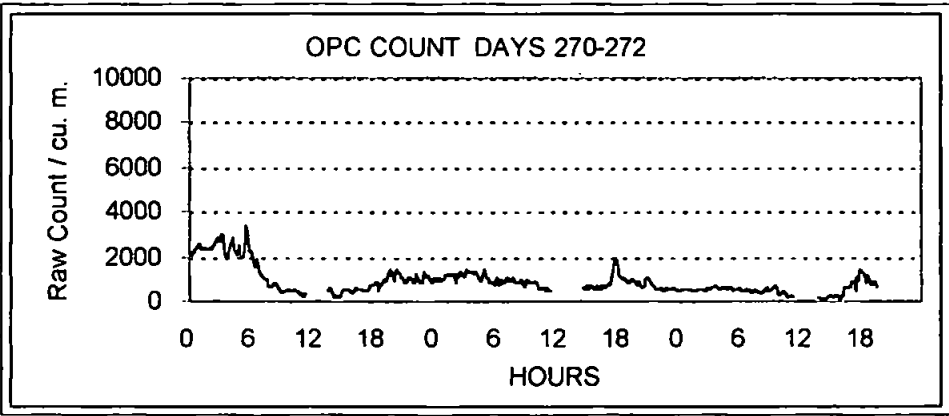


Fig. 8.27b AMT3 Raw counts along transect (underway, cont.).

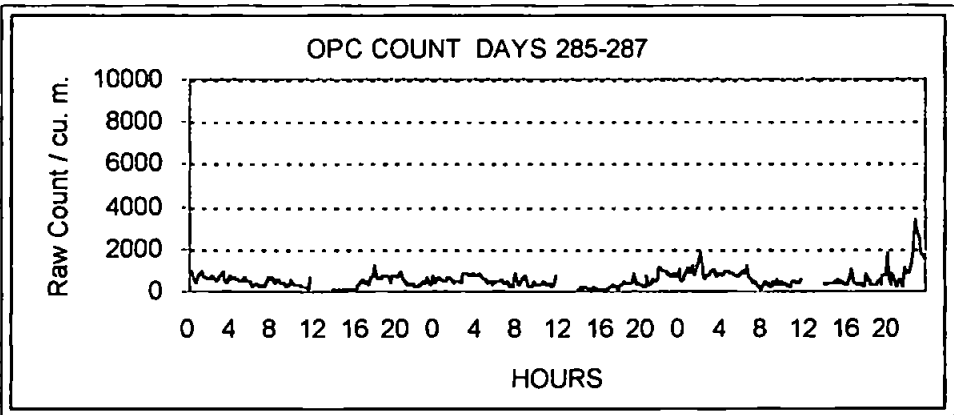
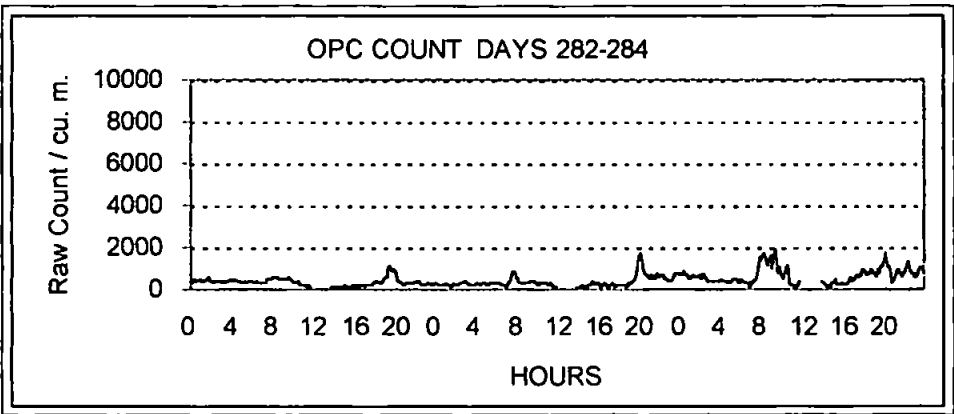
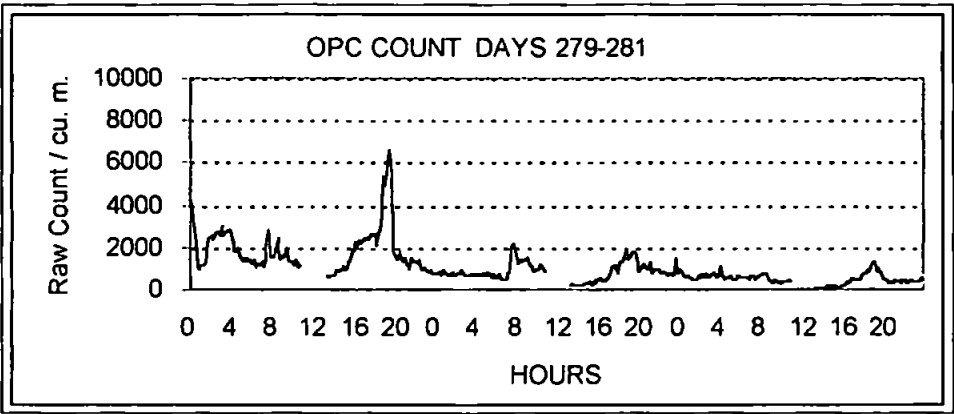


Fig. 8.27c AMT3 Raw counts along transect (underway, cont.).

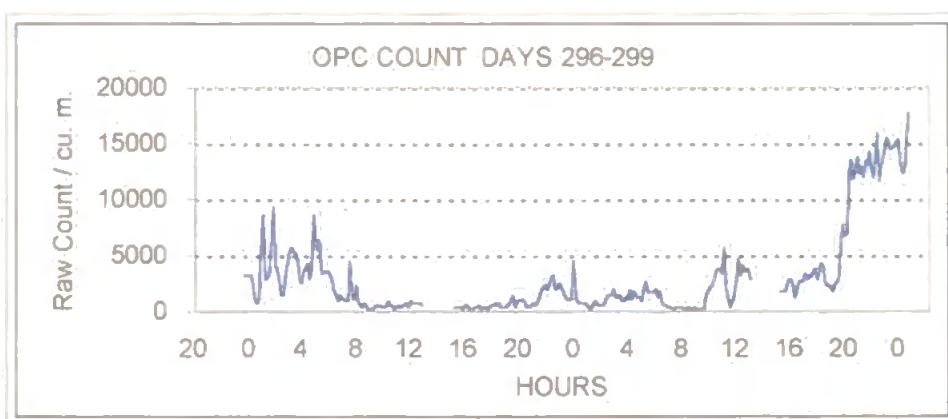
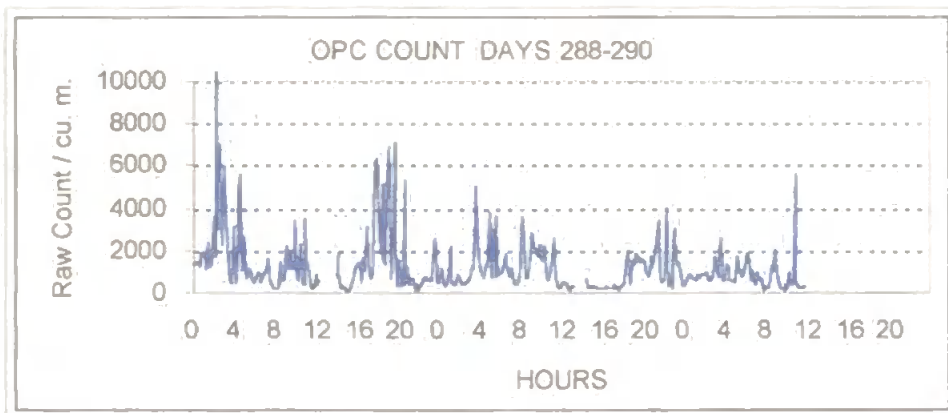


Fig. 8.27d AMT3 Raw counts along transect (underway, cont.).

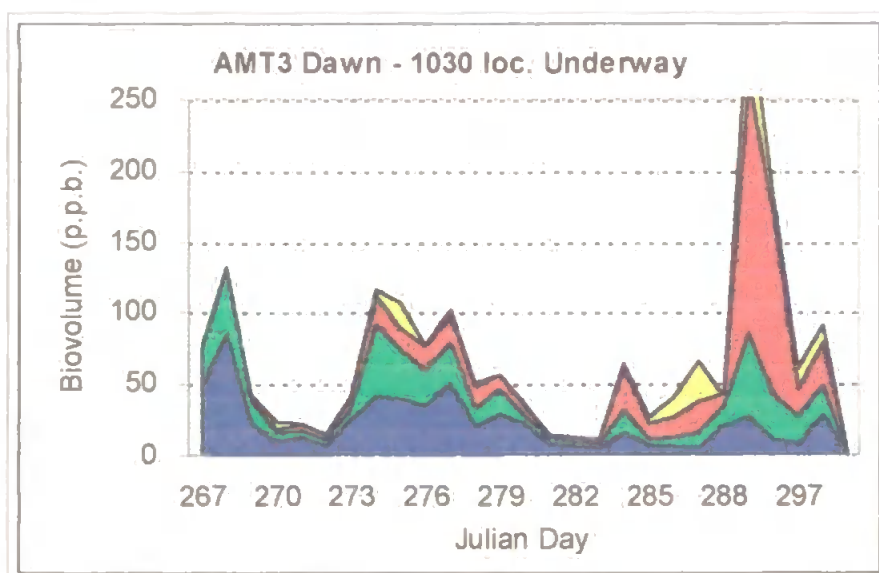
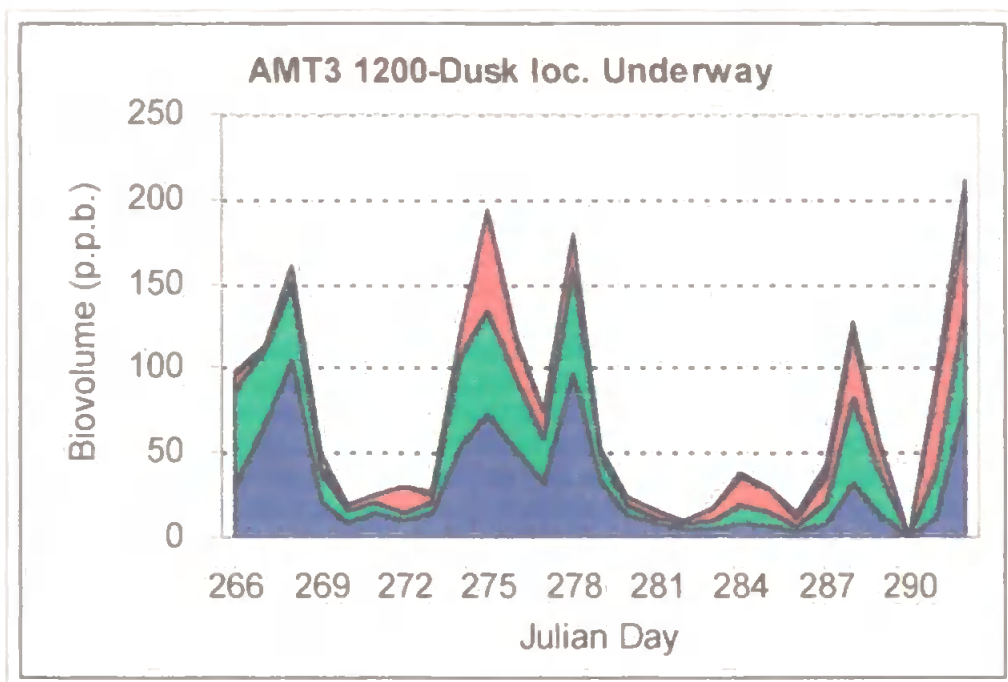
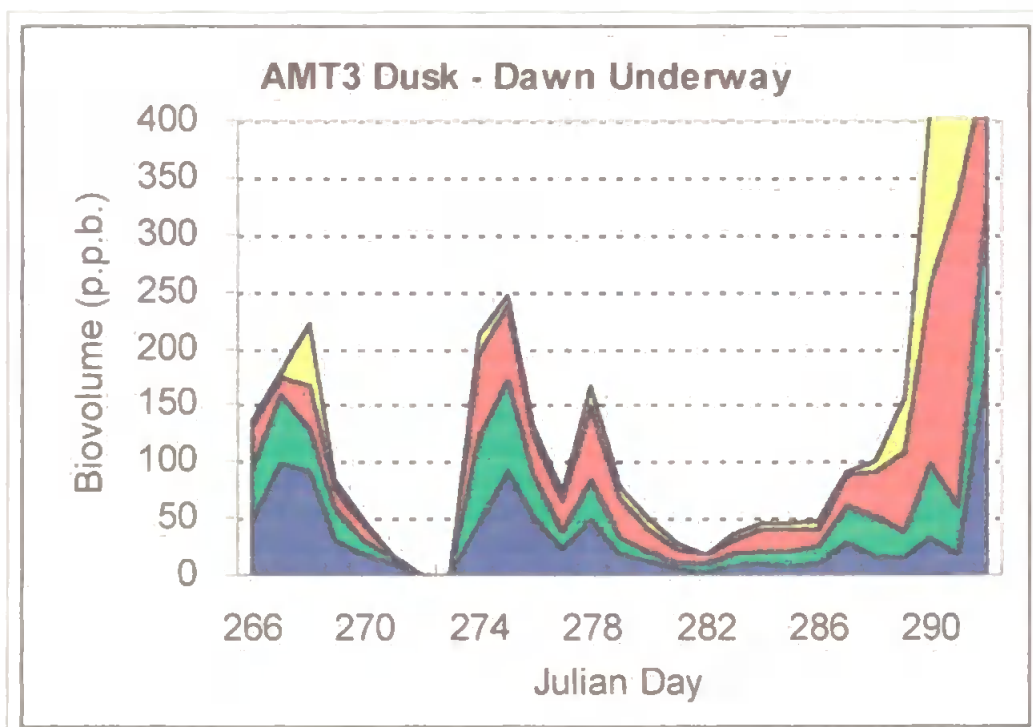


Fig. 8.28 Size distribution of underway biovolume along transect for sampling period local dawn to 1030Z.

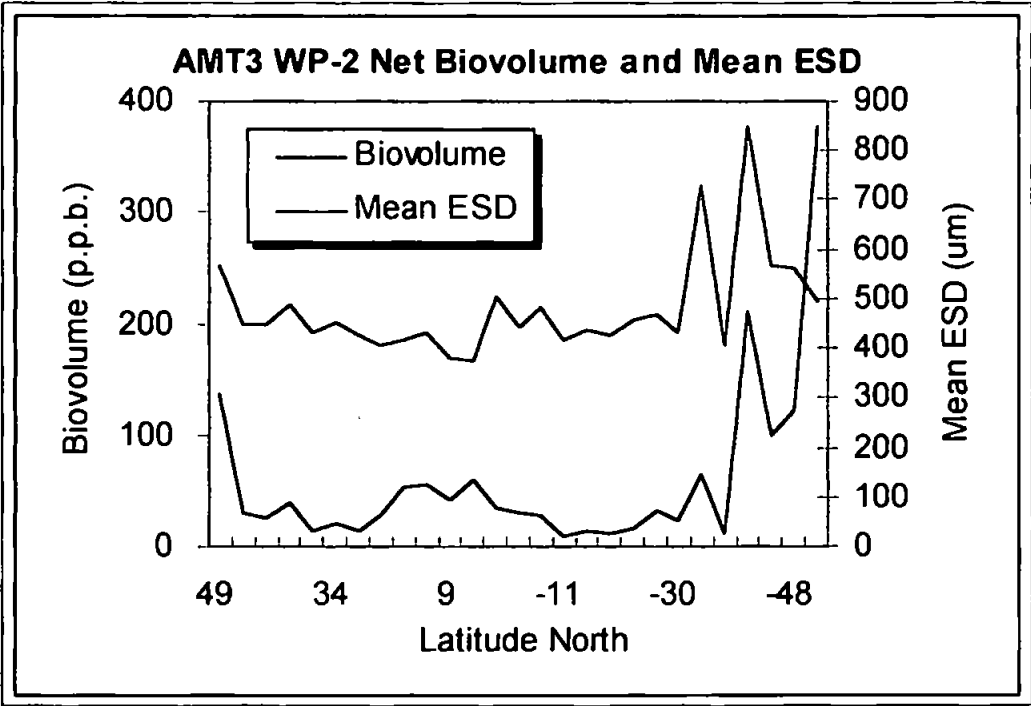


*Fig. 8.29 Size distribution of underway biovolume along transect for sampling period 1200Z to local dusk .*



*Fig. 8.30 Size distribution of underway biovolume along transect for sampling period local dusk to local dawn .*

Figures 8.28 to 8.30 show the size fractionated underway data for the three daily sampling periods for the whole cruise. Figure 8.31 shows total OPC biovolume compared to mean ESD for vertical net hauls for the whole transect. As with previous AMTs the ESD tends to follow biovolume, except at frontal zones. Figure 8.32 shows the vertical net biovolume in the four JGOFS size classes, and fig. 8.33 shows the same data as a percentage of total biovolume in each size class.



*Fig. 8.31 Biovolume & mean ESD along transect (net hauls).*

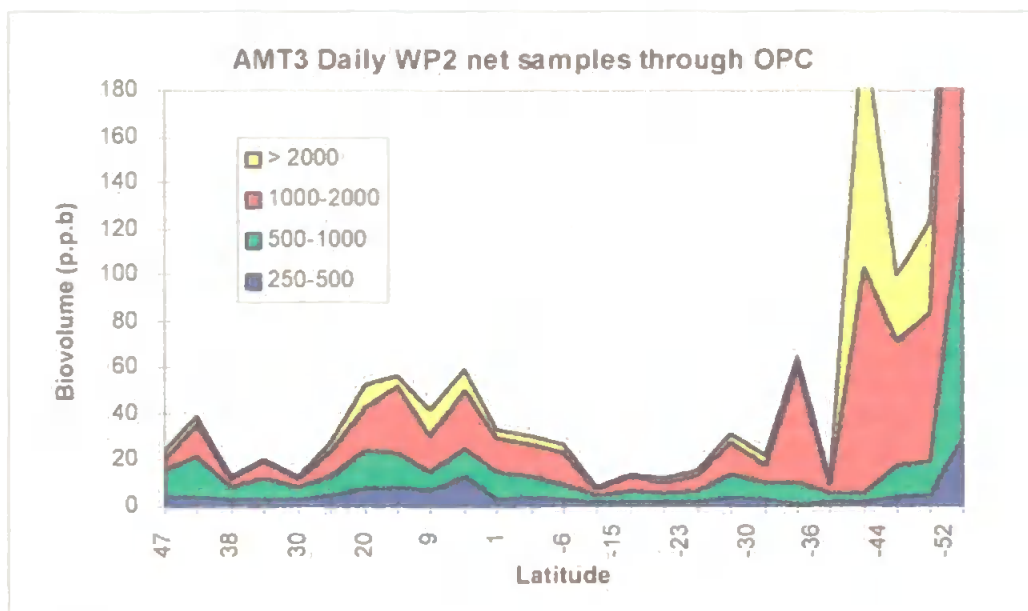


Fig. 8.32 Size distribution of biovolume along transect - net hauls, all sizes in micrometres.

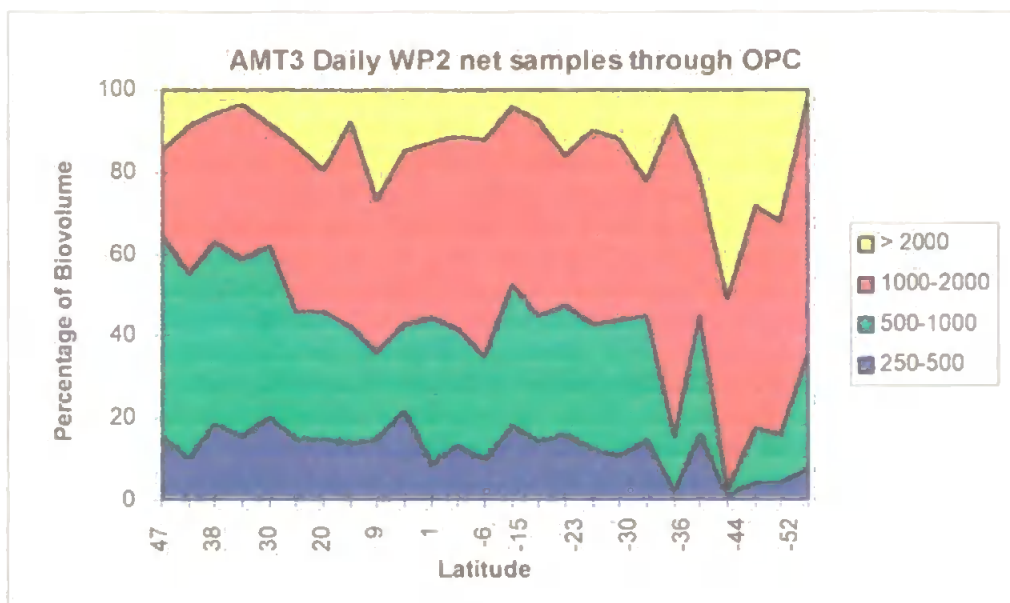


Fig. 8.33 Percentage of total biovolume in each size class along transect - net hauls, all sizes in micrometres.

### 8.6.3 ViZA data results

Figures 8.34 to 8.45, below, show the ViZA data gathered in underway mode for the cruise, and plotted as histograms, and compared to the OPC data for the same periods. Figures 8.34 to 8.37 show the data for the morning run for the whole cruise as total biovolume, and as each of the three smallest size classes. Figures 8.38 to 8.45 show the same data for the afternoon and night runs. The data for objects greater than 2000 $\mu$ m. ESD are not shown, as the undersampling problem described above is worst for this size class, and almost no data appear in this size class in the ViZA output.

ViZA data usable for this comparison exists for the periods between dusk on day 268 to midday on day 275, when the camera was flooded, and from midday on day 278 to dusk on day 282, when the camera gradually deteriorated and finally ceased to function. The ViZA data correlates quite well with the OPC data for total biovolumes and for the smaller size classes at night and in the morning (table 8.2). As size increases, the correlation becomes less good, due to the increased undersampling problem as larger, less abundant animals are sampled. For the same reason, the correlation is worst in the afternoon, when all size classes are least abundant.

**Table 8.2: Correlation coefficients for underway OPC and ViZA data.**

<b>Period:</b>	<b>Morning</b>	<b>Afternoon</b>	<b>Night</b>
<i>200-500 <math>\mu</math>m</i>	0.80	0.36	0.62
<i>500-1000<math>\mu</math>m</i>	0.93	0.54	0.86
<i>1000-2000<math>\mu</math>m</i>	0.14	0.45	0.68
<i>Total</i>	0.91	0.47	0.88

Note that the last column of table 8.2 shows a better correlation for the total biovolumes than for any individual size class. This is possible, because there may be a shift between size classes for either the OPC or the ViZA data. This also appears to occur in the comparison between OPC data and microscopic analysis for the AMT1 samples (section 8.7). The wet sieving process for the latter samples could, for example, introduce such an error. In such a case the data for each size class may correlate poorly, whilst the overall correlation is quite good.

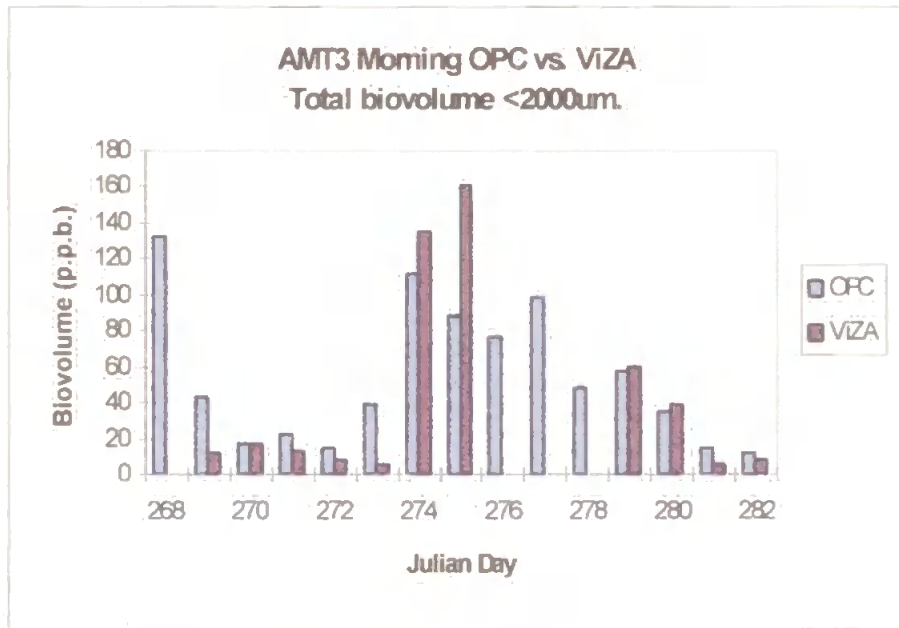


Fig 8.34 AMT3 morning run: total OPC biovolume less than 2000µm ESD compared to ViZA biovolume less than 2000µm ESD.

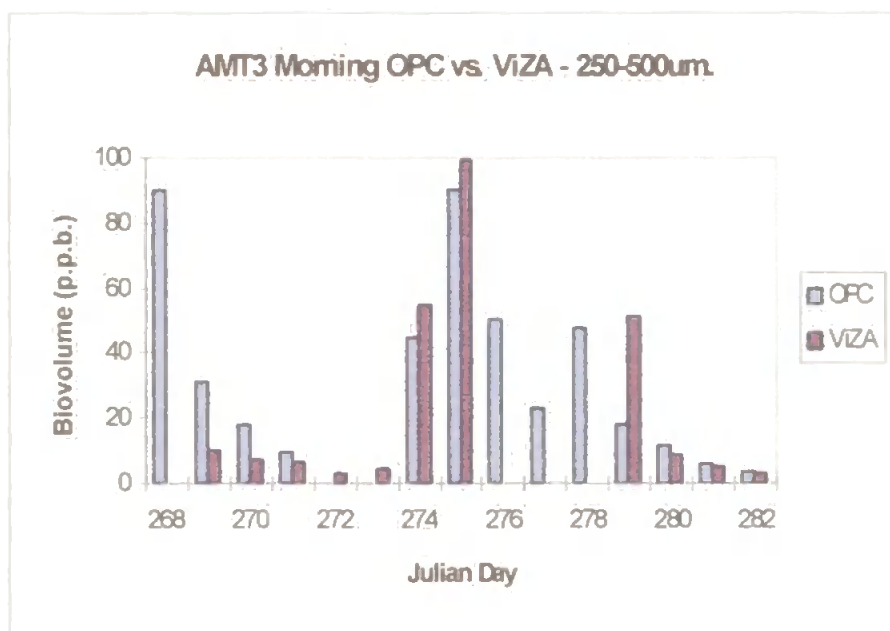


Fig.8.35 AMT3 morning run: OPC biovolume 250 - 500 $\mu$ m ESD compared to ViZA biovolume 250 - 500 $\mu$ m ESD.

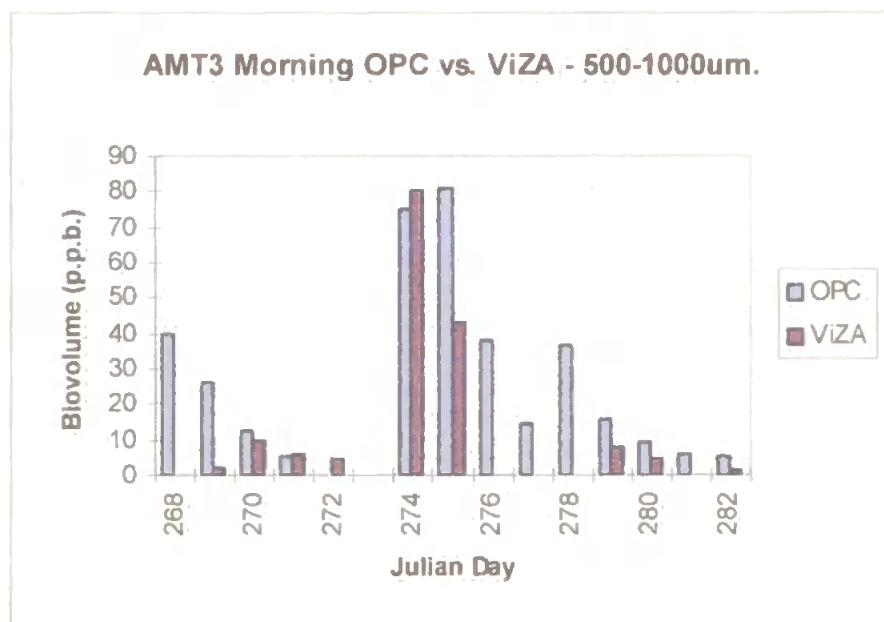


Fig 8.36 AMT3 morning run: total OPC biovolume 500 - 1000 $\mu$ m ESD compared to ViZA biovolume 500 - 1000 $\mu$ m ESD.

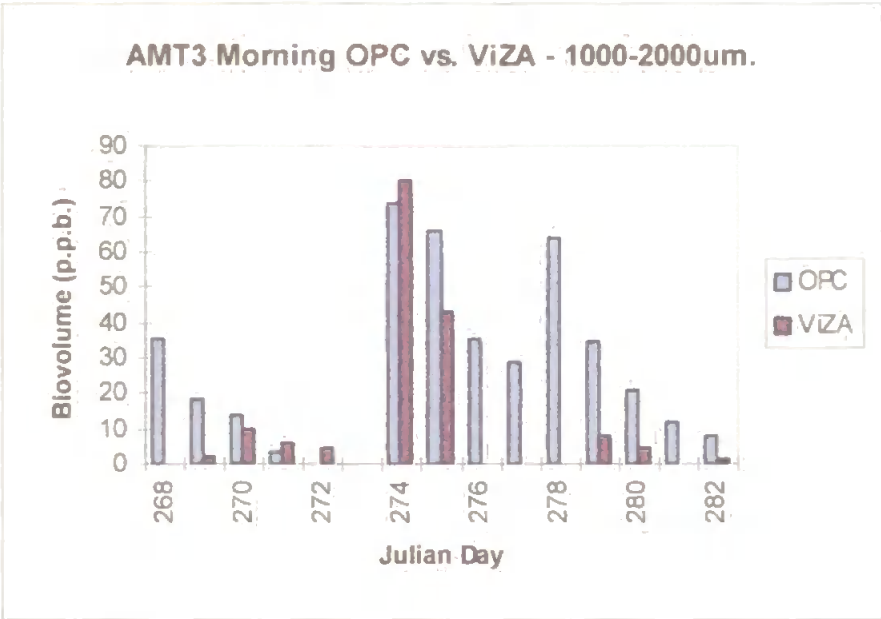


Fig 8.37 AMT3 morning run: total OPC biovolume 1000 - 2000 $\mu$ m ESD compared to ViZA biovolume 1000 - 2000 $\mu$ m ESD.

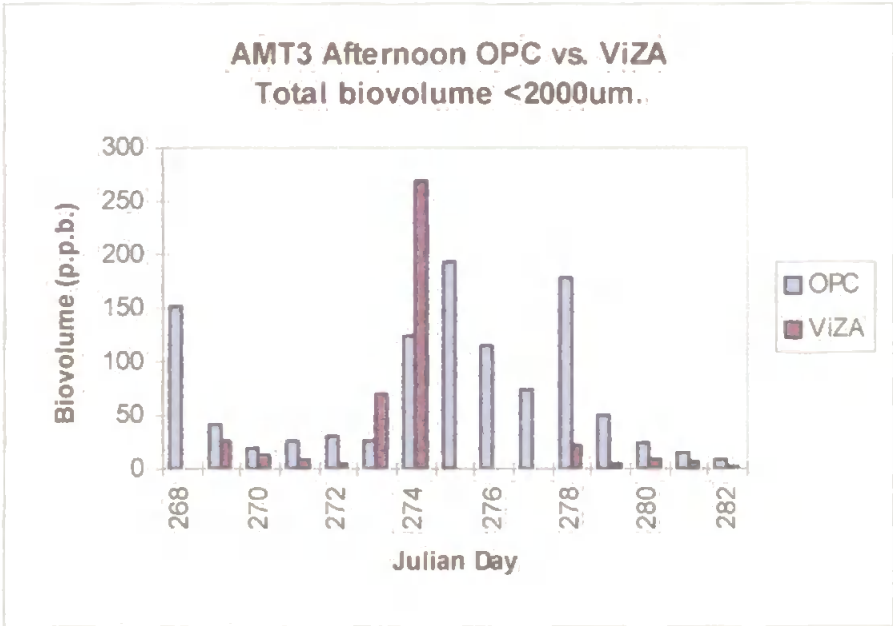


Fig 8.38 AMT3 afternoon run: total OPC biovolume less than 2000 $\mu$ m ESD compared to ViZA biovolume less than 2000 $\mu$ m ESD.

AMT3 Afternoon OPC vs. VIZA - 250-500um.

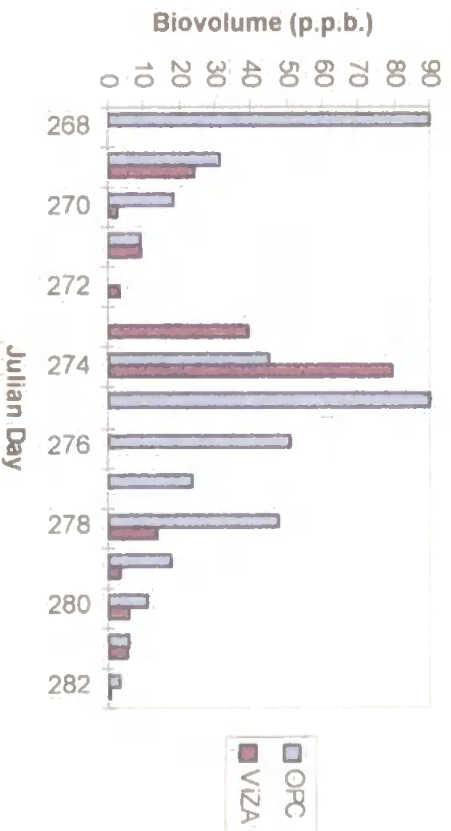


Fig 8.39 AMT3 afternoon run: total OPC biovolume 250 - 500µm ESD compared to VIZA biovolume 250 - 500µm ESD.

AMT3 Afternoon OPC vs. VIZA - 500-1000um.

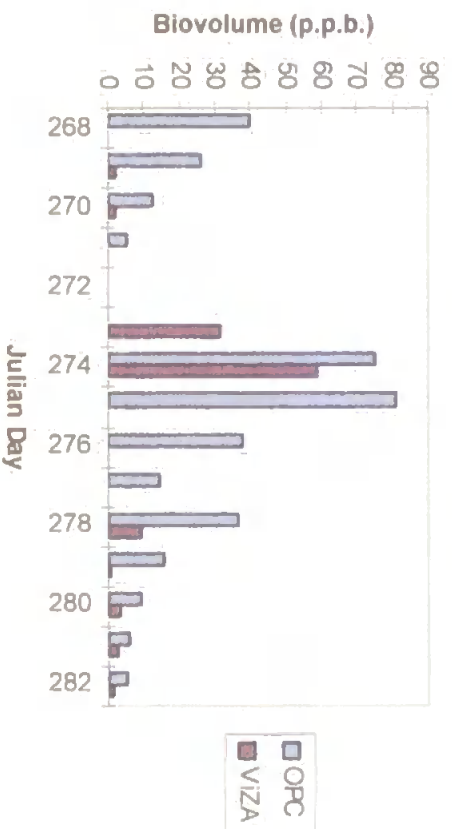
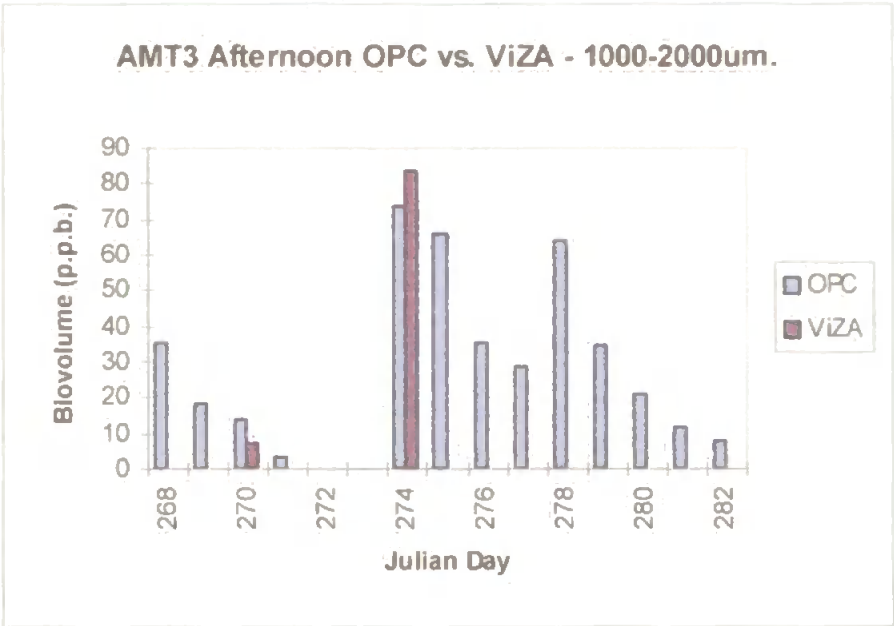
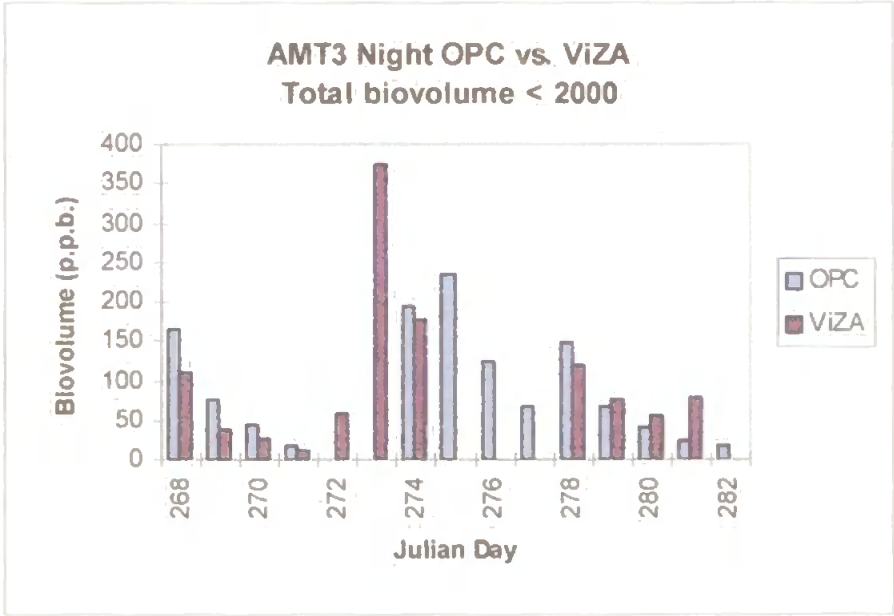


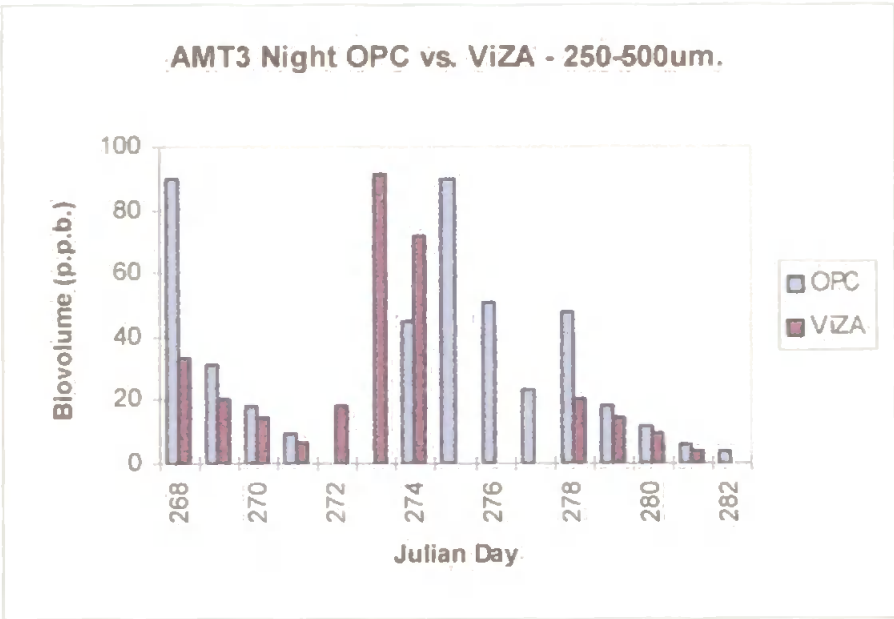
Fig 8.40 AMT3 afternoon run: total OPC biovolume 500 - 1000µm ESD compared to VIZA biovolume 500 - 1000µm ESD.



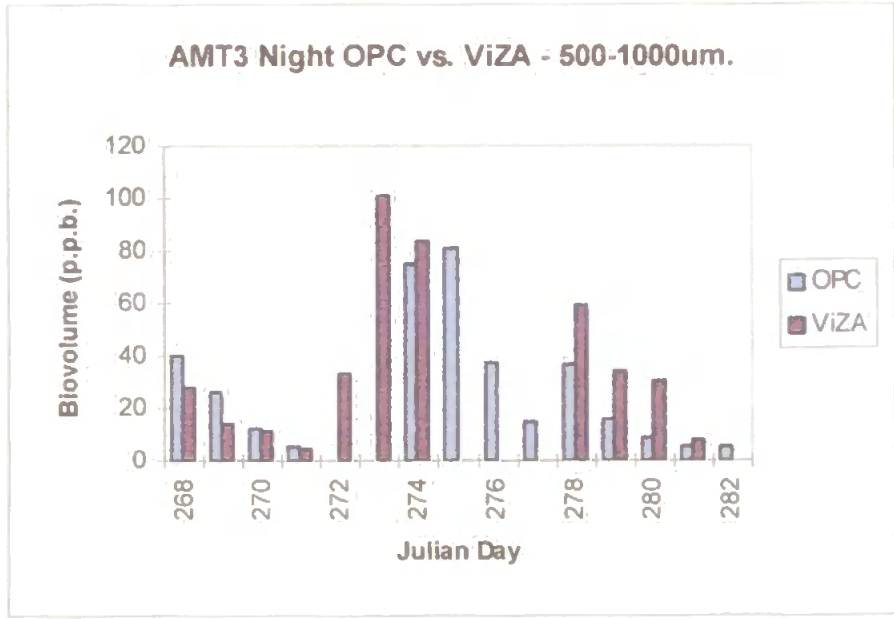
*Fig 8.41 AMT3 afternoon run: total OPC biovolume 1000 - 2000 $\mu$ m ESD compared to ViZA biovolume 1000 - 2000 $\mu$ m ESD.*



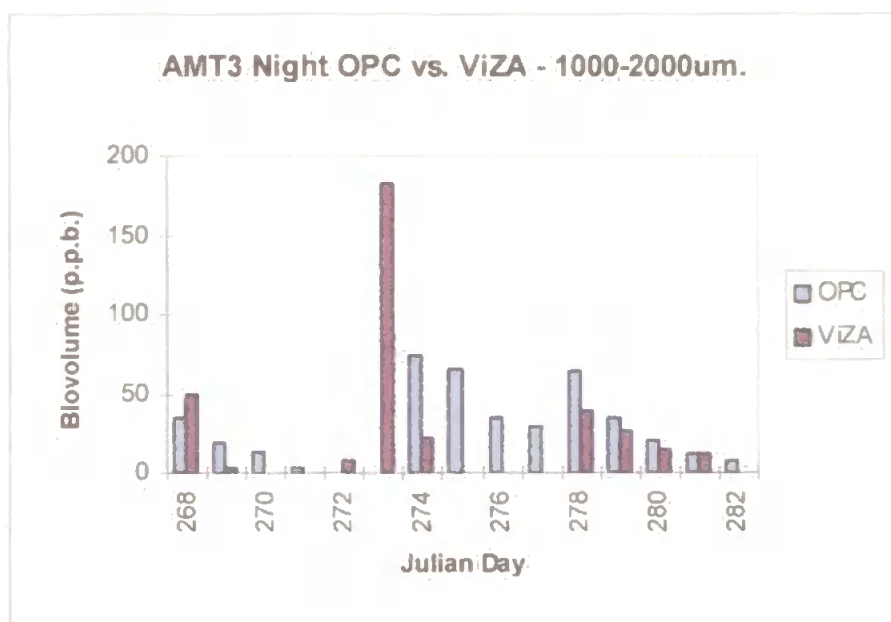
*Fig 8.42 AMT3 night run: total OPC biovolume less than 2000 $\mu$ m ESD compared to ViZA biovolume less than 2000 $\mu$ m ESD.*



*Fig 8.43 AMT3 night run: total OPC biovolume 250 - 500 $\mu$ m ESD compared to ViZA biovolume 250 - 500 $\mu$ m ESD.*



*Fig 8.44 AMT3 night run: total OPC biovolume 500 - 1000 $\mu$ m ESD compared to ViZA biovolume 500 - 1000 $\mu$ m ESD.*



*Fig 8.45 AMT3 night run: total OPC biovolume 1000 - 2000 $\mu$ m ESD compared to ViZA biovolume 1000 - 2000 $\mu$ m ESD.*

Several samples of animals passing through the OPC/ViZA system were collected on a 200 $\mu$ m mesh gauze. From these samples a training set was selected from animals in five major groups typical of AMTs, and large enough to give meaningful outlines. The copepod data shown in fig. 7.1 is one of this training set. Figures 8.46 and 8.47, below, show a euphausiid and an amphipod from the same training set. Typical size and shape descriptors for these species are given. This training set was used to derive a set of principal components for use as discriminant functions on AMT3. The display was shown in fig. 7.3 and is repeated as fig. 8.48, below, showing the results of the principal component analysis on this training set.

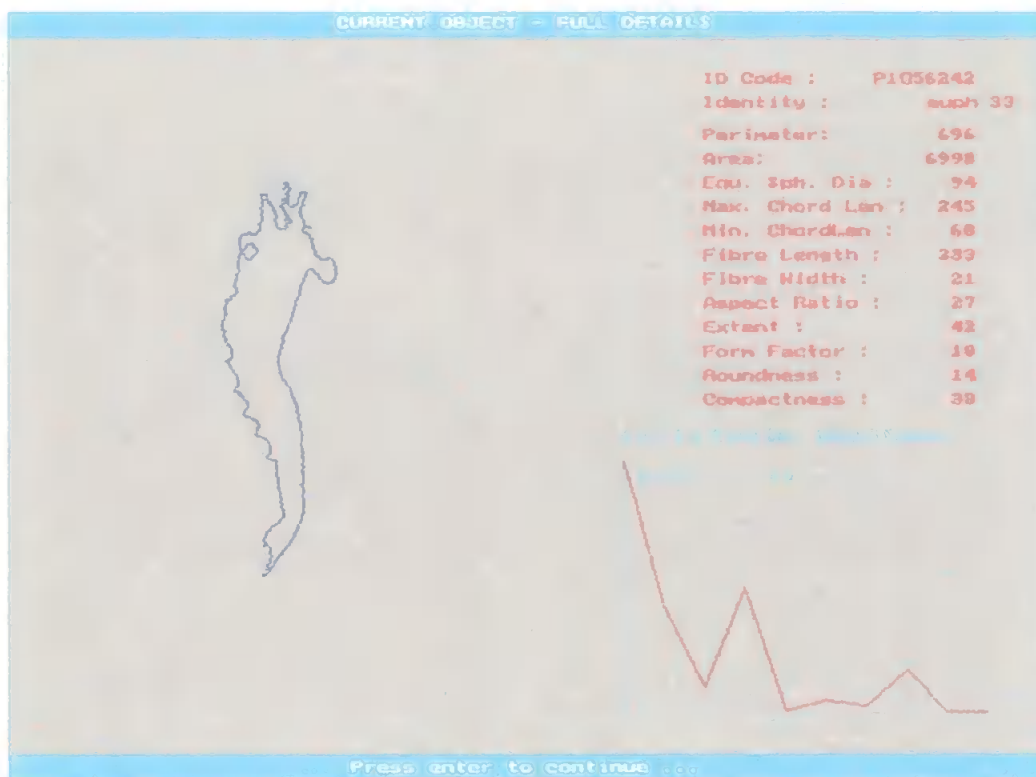


Fig. 8.46 Euphausiid captured from training set of zooplankton from AMT3.

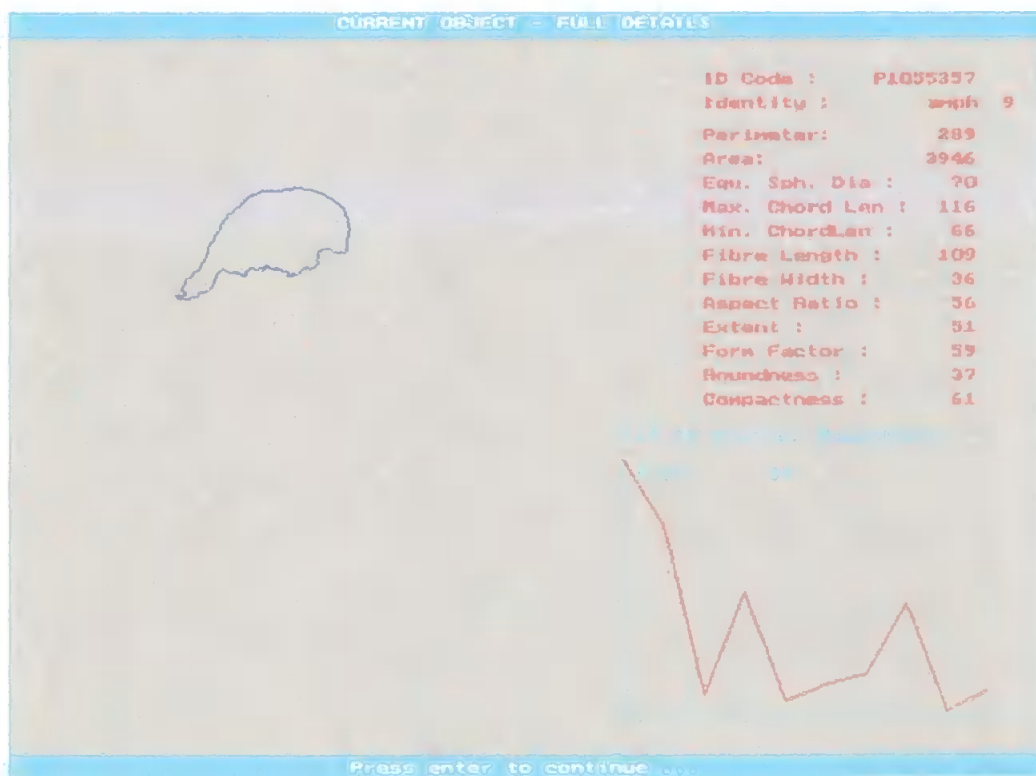


Fig. 8.47 Amphipod captured from training set of zooplankton from AMT3.

The display of fig. 8.48 shows how the objects of the training set cluster in the space defined by the first two of the best set of principal components determined by the system. All six groups are clearly separable with no overlap.

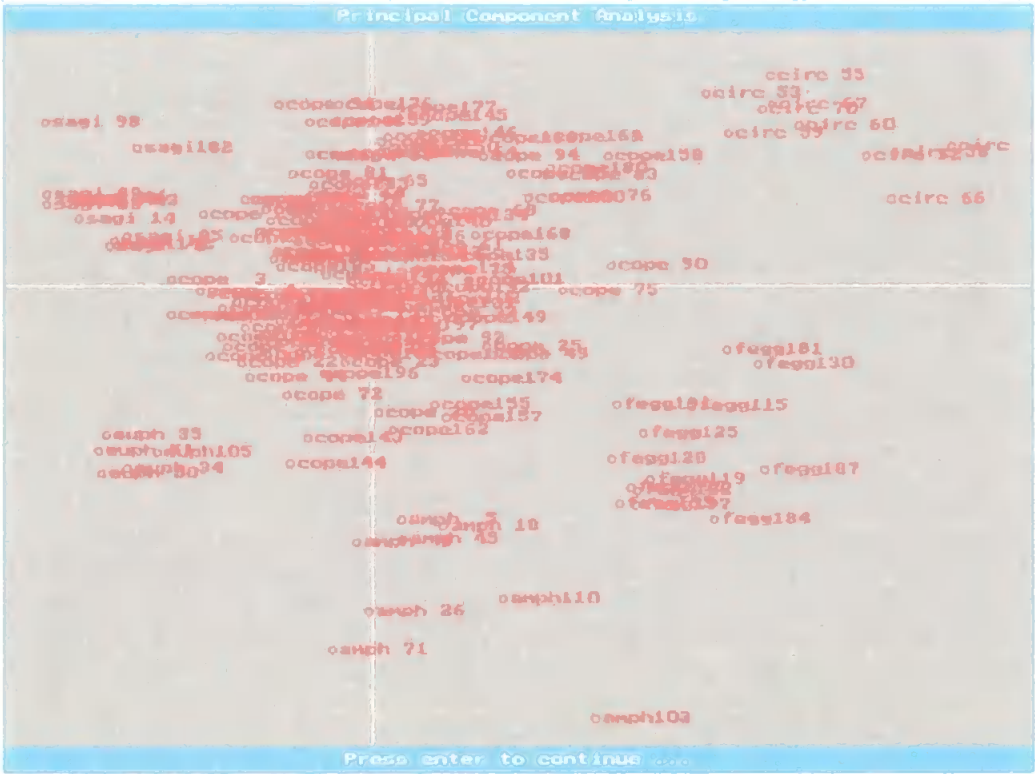


Fig. 8.48 Graphical display of the results of the PCA on the training set from AMT3. X and Y axes are principal components one and two respectively.

Table 8.3 shows the coefficients of each principal component, and the amount of variance in the data set accounted for by each of the components. The first three components account for 95% of the total variance, indicating a well ordinated data set. The separation factor is 2.49, indicating that all groups are clearly separated.

A selection of 216 objects acquired during days 273-275 on AMT3 were then assembled in a file, and presented to the analysis program for identification.

Although this represents a real data set, it has been artificially modified in two ways in order to overcome the problems discussed. First, any animals so small that pixellation effects dominate the shape information contained in the boundary were removed. Secondly, any animals from whose shape the author could not make a reasonable identification for were also rejected. This was partly to remove those animals which had been poorly imaged due to the transparency effect, or to overlapping, and partly because in order to be able to judge the effectiveness of the program in discriminating between objects, we must be able to say what they should be. The test set had 216 objects, whose size and species distribution, determined by the analysis program are shown below.

*Table 8.3 Coefficients and variance of principal components 1 to 3 for training set used on AMT3.*

<b>Parameter</b>	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>
<i>Perimeter:</i>	0.00	0.00	0.00
<i>Area :</i>	0.00	0.00	0.00
<i>ESD:</i>	0.00	0.00	0.00
<i>Maj. Chord:</i>	0.00	0.00	0.00
<i>Min. Chord:</i>	-0.10	0.68	-0.49
<i>Fibre Length:</i>	0.00	0.00	0.00
<i>Fibre Width:</i>	0.01	0.66	0.73
<i>Aspect ratio:</i>	0.51	0.29	-0.44
<i>Extent:</i>	0.00	0.00	0.00
<i>Form factor :</i>	0.00	0.00	0.00
<i>Roundness:</i>	0.60	-0.09	0.13
<i>Compactness:</i>	0.61	-0.06	0.15
<i>Variance:</i>	51.28	35.87	8.00
<i>Best separation factor:</i>		2.48	

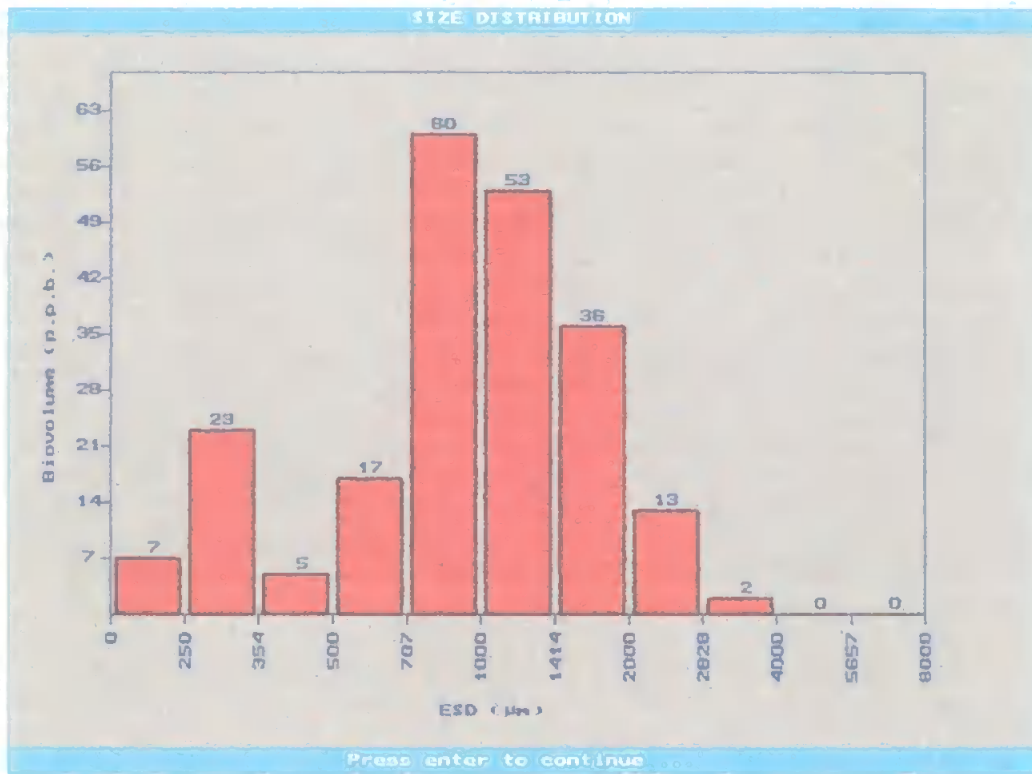


Fig. 8.49. Size distribution of sample test set from AMT3.

The training set included some objects which were air bubbles deliberately introduced into the system. Air bubbles are a problem with the OPC, as discussed above, and it would therefore be useful if the ViZA system could identify these. They are labelled as 'circ' in figs 8.48 and 8.50, being completely circular. It may be thought that fish eggs would also be very circular, but the egg case is in fact usually transparent, and all that is seen is the embryo inside, which has a distinctive, non-circular shape. Table 8.4, over, shows the efficacy of the ViZA system in classifying the 216 objects in this test set to one of the 6 classes determined by the principal components analysis. Copepods dominate this sample, as they do most samples taken on the AMT cruises, and their distinctive body shape (legs, maxillae, etc., are not imaged) allows them to be reliably identified. Euphausiids and amphipods are also reliably identified,

probably mostly due to their large size, and chaetognatha also have very distinctive shapes. The embryo shape imaged from fish eggs, however, is often mistaken for the shape of copepods, as defined by the analysis software, and circular bubbles are mistaken for fish eggs and copepods in three each of 14 cases.

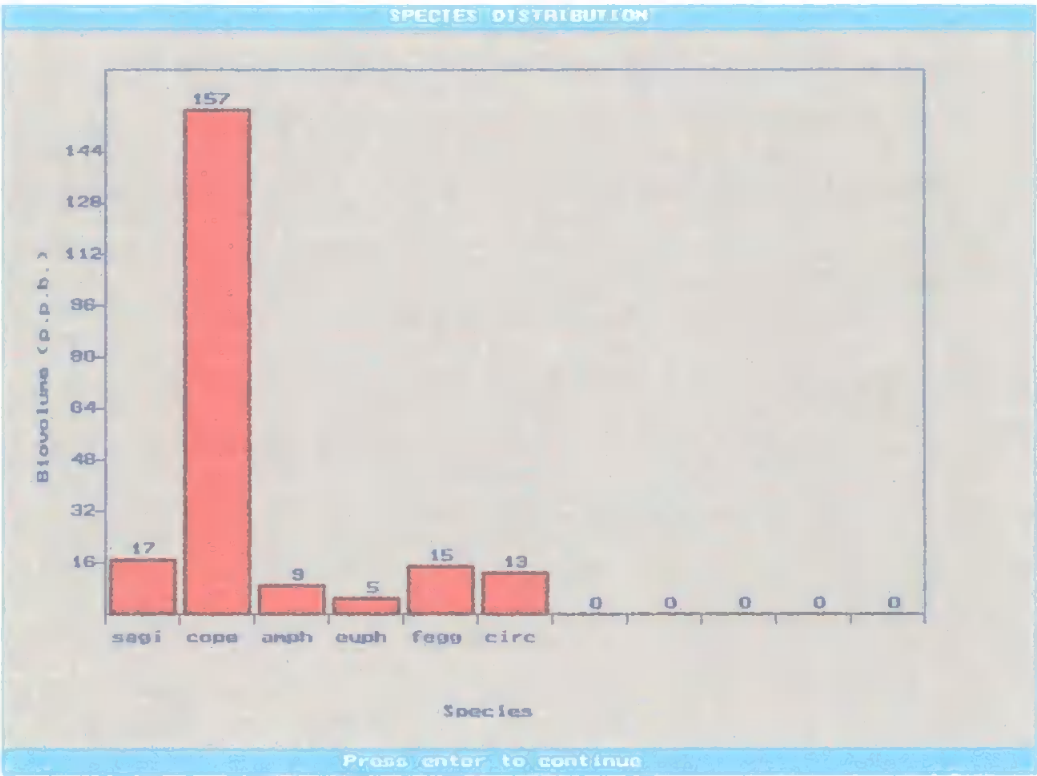


Fig. 8.50 Species distribution of sample test set from AMT3.

Table 8.4 ViZA identification success rate for trial sample of animals from North Atlantic cruise. Columns show correct class by visual identification; rows show the classification by ViZA.

	Copepod	Chaetognath	Amphipod	Euphausiid	Circ	Fish Egg
Copepod	152	1	0	0	3	4
Chaetognath	2	16	0	0	0	0
Amphipod	2	0	9	0	0	1
Euphausiid	0	0	0	5	0	0
Circ	0	0	0	0	7	0
Fish Egg	1	0	0	0	3	10

## 8.7 Data validation

Data validation for AMT and PRIME cruises was by means of preserved samples from the vertical net casts and from the inline filter. These preserved samples were to be subjected to laboratory analysis for size fractionated carbon and nitrogen, and to microscopic analysis for particle counts, size and species identification. Figures 8.51 and 8.52 show the comparison between OPC particle counts and those obtained from microscopic analysis for AMTs 1 & 2. Unfortunately, due to the cost involved in microscopic taxonomic analysis, as discussed in section 2.3.2, none of the preserved samples have been analysed for taxonomy, apart from the first 4 vertical net samples from AMT1 (fig. 8.53 - 8.57). It is hoped that these samples will be analysed, at least for major taxonomic groups, during 1997.

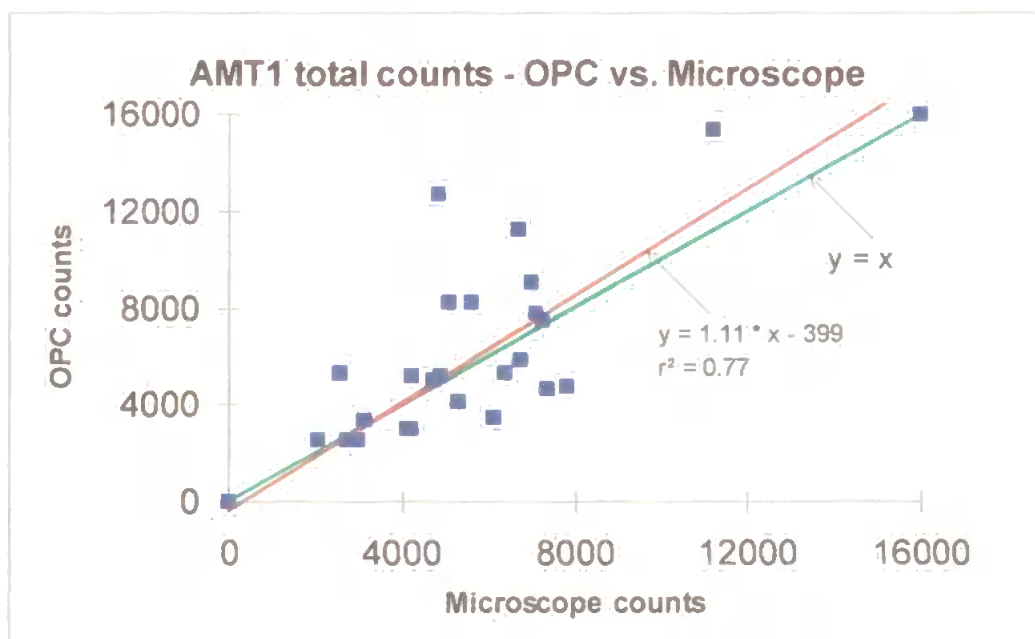


Fig. 8.51 AMT1 WP-2 nets - total counts from OPC and microscope analysis. Red line shows linear regression, green line shows ideal 1:1 relationship.

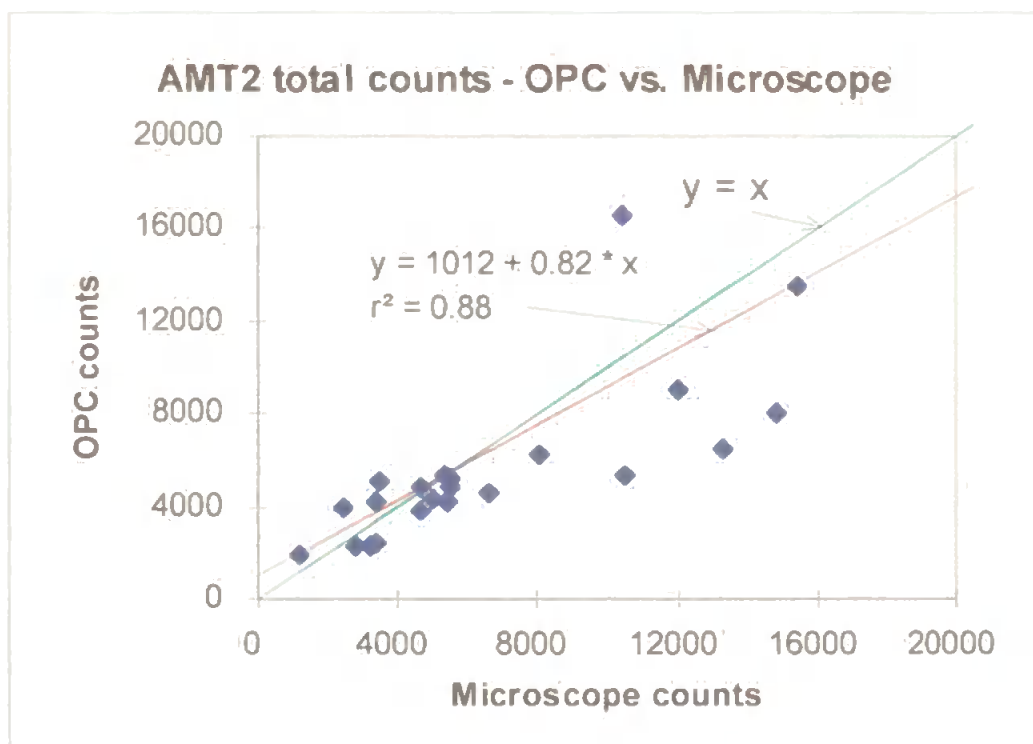
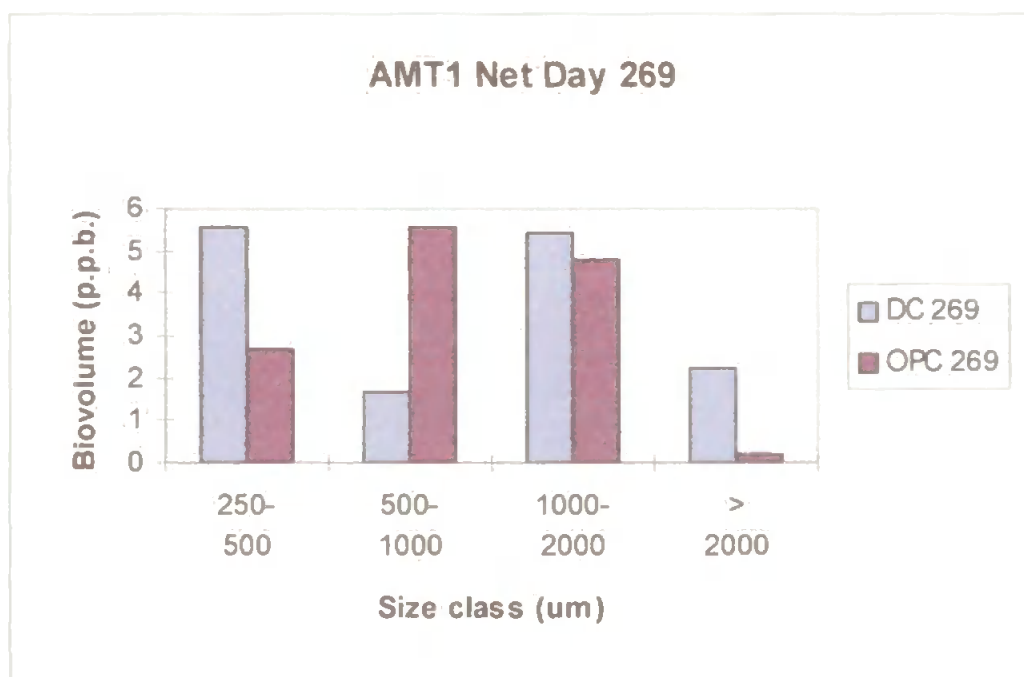


Fig. 8.52 AMT2 WP-2 nets - total counts from OPC and microscope analysis. Red line shows linear regression, green line shows ideal 1:1 relationship.

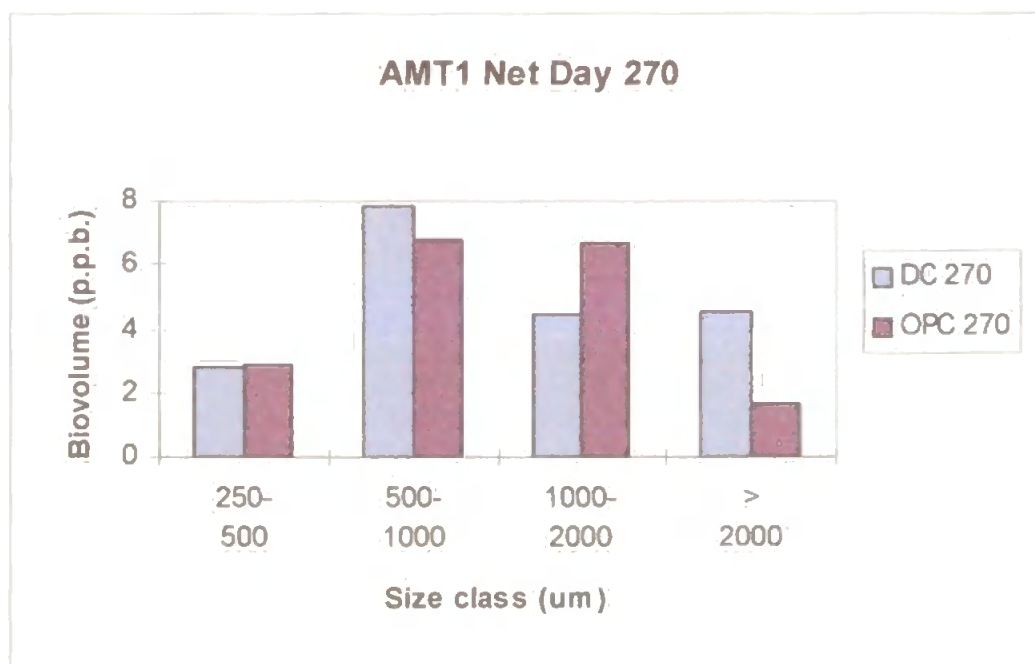
Note: In fig. 8.52, station 141 is omitted, as its high count scales other stations too close to zero for clear presentation. Nevertheless, this station fits the model quite well, having an OPC count of 43,000 and microscope count of 48,000.

### 8.7.1 AMT1

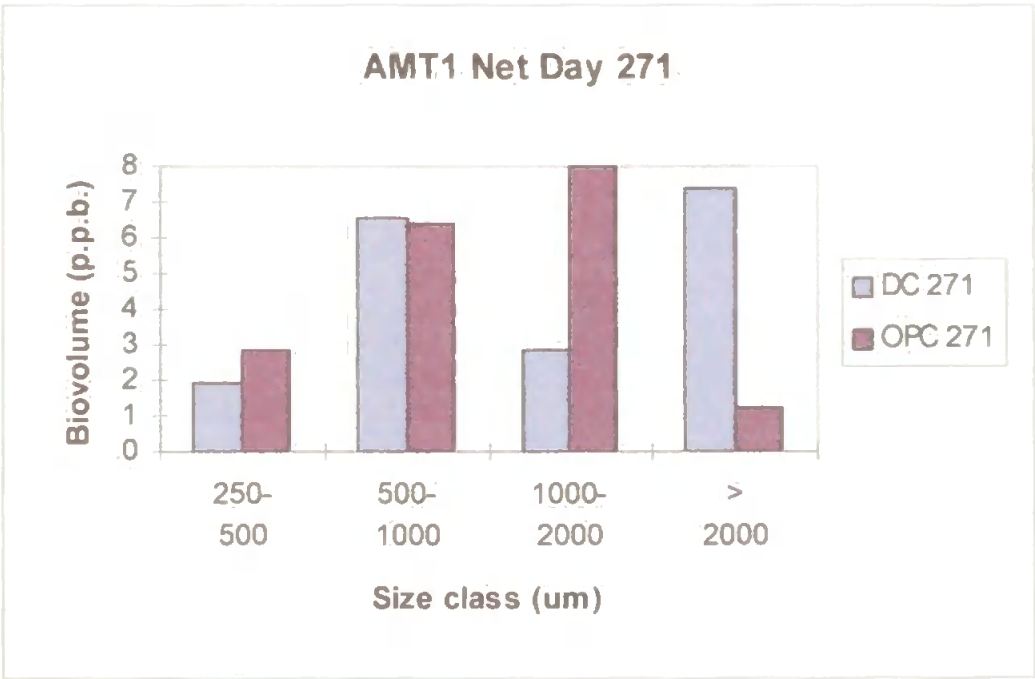
In figures 8.53 to 8.56 the correlation between OPC counts for each size class is variable, particularly at the 500µm. and 2000µm. boundaries, although the total counts correlate well as shown in fig. 8.57. As discussed above, this may be due to errors in the OPC size classification or to the wet sieving process used to separate size fractions for microscopic analysis. Rather more data is required before any consistent pattern can emerge to confirm this.



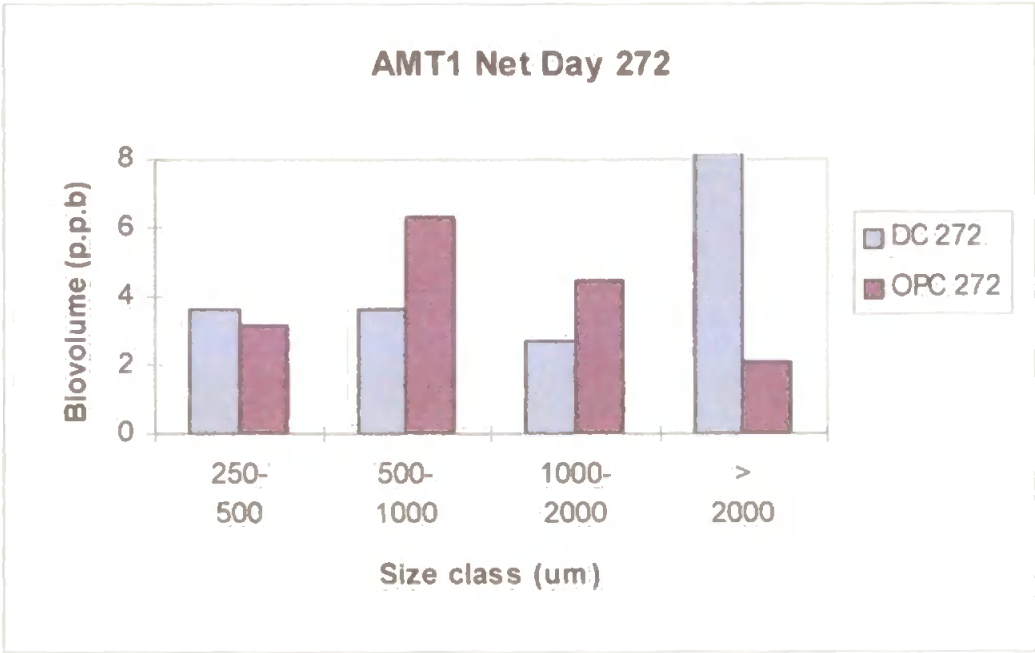
*Fig. 8.53 AMT1 vertical net cast, day 269: biovolume in each size fraction from sample through OPC, and from microscopic examination.*



*Fig. 8.54 AMT1 vertical net cast, day 270: biovolume in each size fraction from sample through OPC, and from microscopic examination.*



*Fig. 8.55 AMT1 vertical net cast, day 271: biovolume in each size fraction from sample through OPC, and from microscopic examination.*



*Fig. 8.56 AMT1 vertical net cast, day 272: biovolume in each size fraction from sample through OPC, and from microscopic examination.*

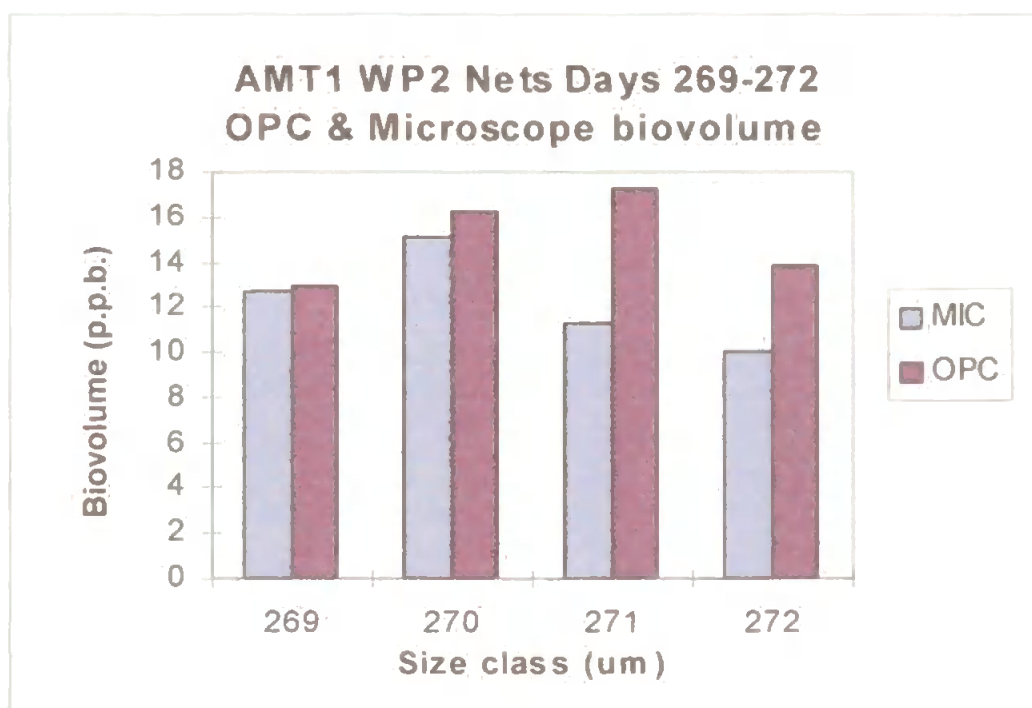


Fig. 8.57 AMT1 vertical net casts, days 269 - 272: total biovolume from sample through OPC, and from microscopic examination.

#### 8.7.1 AMT1

The following graphs show the results of the carbon and nitrogen analysis compared to the OPC biovolume data. Figures 8.54 and 8.55 show total OPC biovolume (wet weight in  $\text{mg m}^{-3}$ , assuming specific gravity of 1 for zooplankton), carbon in  $\text{mg m}^{-3} \times 25$ , and nitrogen in  $\text{mg m}^{-3} \times 100$ , and percentage of carbon by dry weight and C:N ratios. If we assume a ratio of dry weight to wet weight of 1:10, a reasonable approximation for marine zooplankton, these factors are equivalent to 40% carbon by dry weight and 10% nitrogen by dry weight. These proportions are those generally expected in zooplankton studies, with a carbon to nitrogen ratio of 4:1 (Curl, 1962, Omori, 1969). Figure 8.58 (AMT1) shows carbon and nitrogen following biovolume quite well, but at ratios a little lower than expected. Figure 8.59 shows the

carbon to nitrogen ratio consistently around 4:1 except at two points, at about 17° south, and again at about 28° south, where the ratio rises to about 8:1.

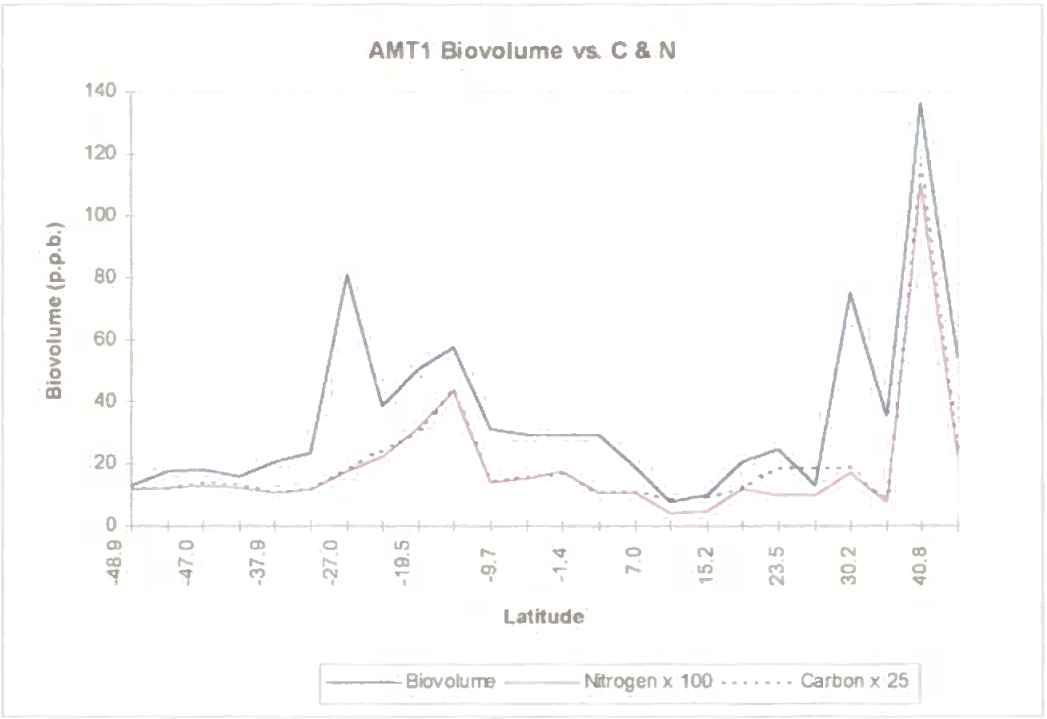


Fig. 8.58 AMT1 WP-2 net samples - total OPC biovolume, carbon and nitrogen.

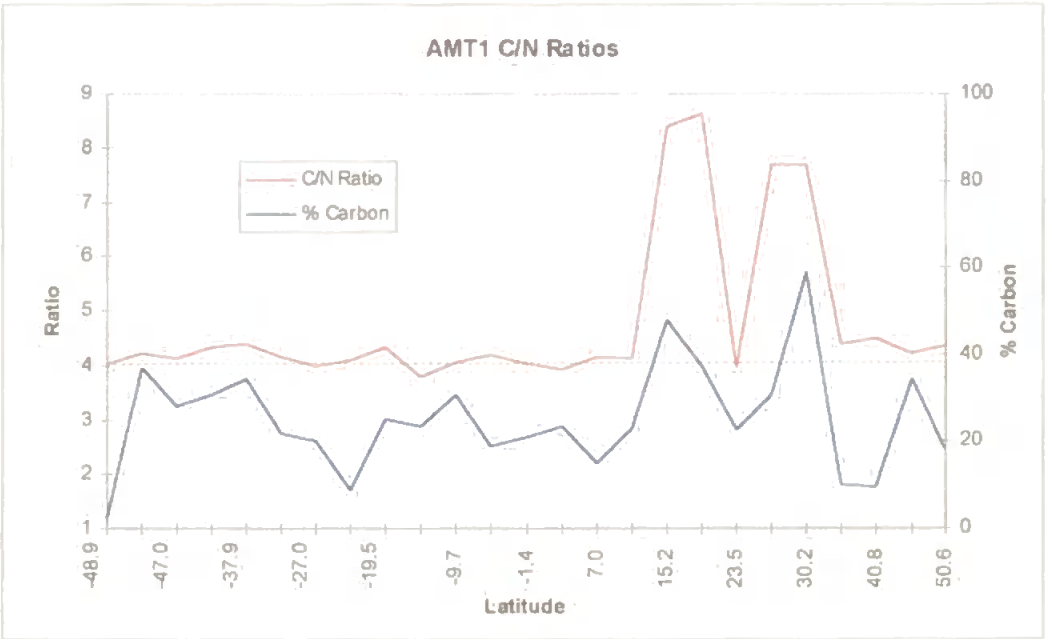


Fig. 8.59. AMT1 WP-2 net samples - carbon to nitrogen ratio and % carbon.

8.7.2 AMT2

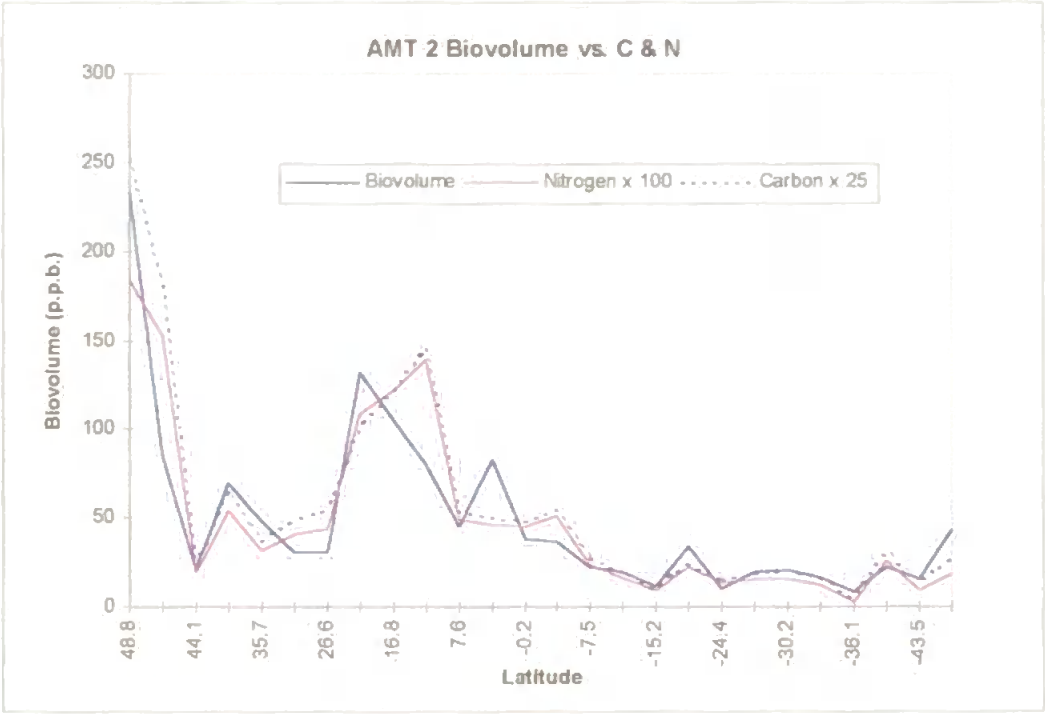


Fig. 8.60 AMT2 WP-2 net samples - total OPC biovolume, carbon and nitrogen.

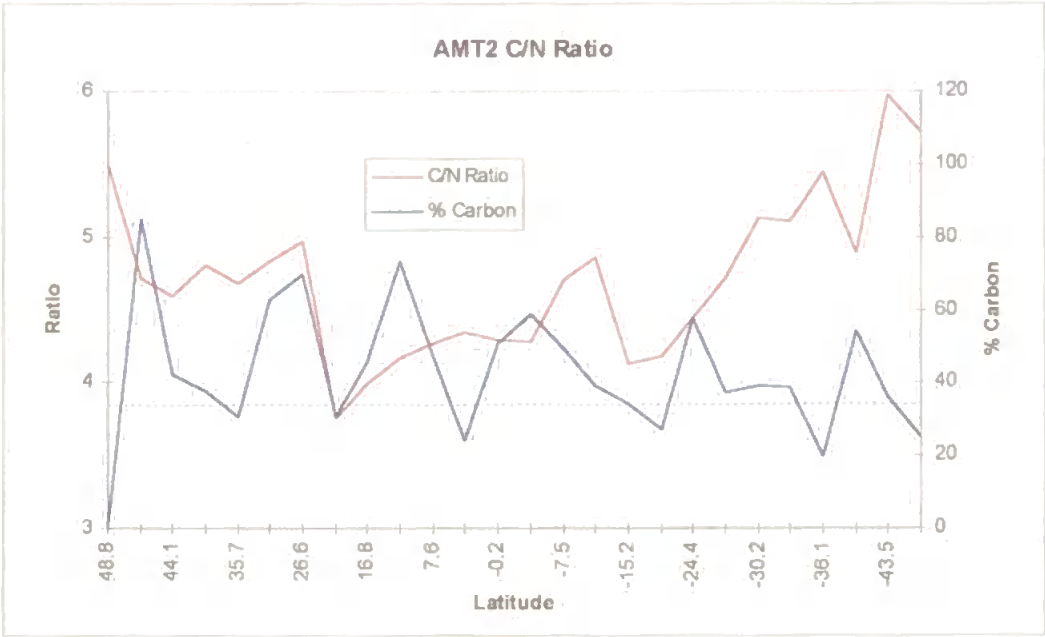


Fig. 8.61 AMT2 WP-2 net samples - carbon to nitrogen ratio and % carbon.

Figure 8.60 shows a much closer match between OPC biovolume and carbon at 40% of dry weight, and nitrogen at 10% of dry weight for AMT2. The carbon to nitrogen ratio is much more variable than for AMT1, as shown in figure 8.61. Values are close to 4:1 at low latitudes either side of the equator, but rise to between 5.5 and 6:1 at higher latitudes.

### 8.7.3 *PRIME*

The following graphs showing OPC biovolume, carbon and nitrogen for the PRIME cruise are drawn the same way as for AMTs 1&2, except that biovolume is plotted at one tenth scale (equivalent to dry weight,  $\text{mg m}^{-3}$ ). Note that all graphs show only data for size classes less than  $2000\mu\text{m}$  ESD, as the largest size fraction was not analysed for carbon and nitrogen. Carbon and nitrogen values are plotted scaled  $\times 2.5$  and  $\times 10$  respectively, equivalent to 40% and 10% of dry weight as above. For the vertical net profiles the C:N ratio is fairly steady at 4:1 throughout the transect (figs. 8.63 & 8.65). Carbon and nitrogen as a proportion of total dry weight is more variable, however. The daytime nets show higher values for carbon and nitrogen at higher latitudes, except for the first station where large gelatinous organisms dominated. The night nets show values close to those expected for carbon and nitrogen except at  $50^\circ$  north, the frontal area mentioned above, where values are higher, and at  $44^\circ$  north, where the sample was again dominated by large gelatinous organisms.

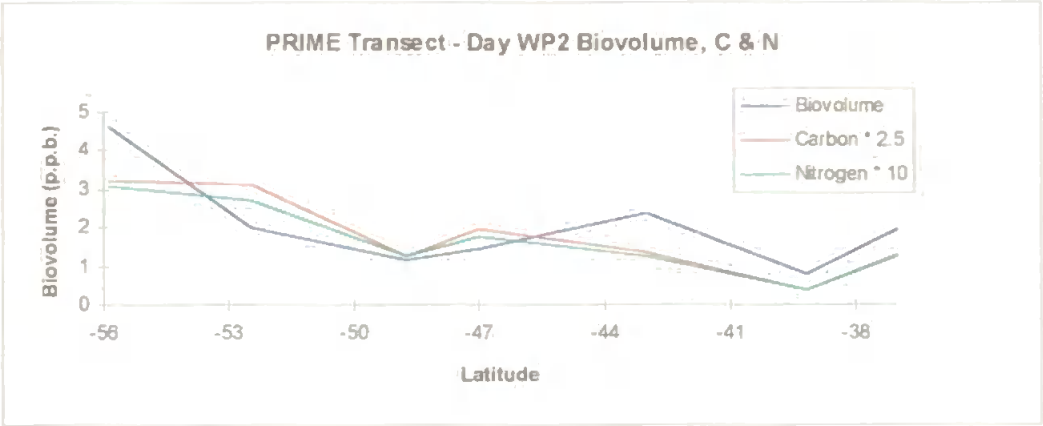


Fig. 8.62 PRIME transect daytime WP-2 net samples - total OPC biovolume, carbon and nitrogen.

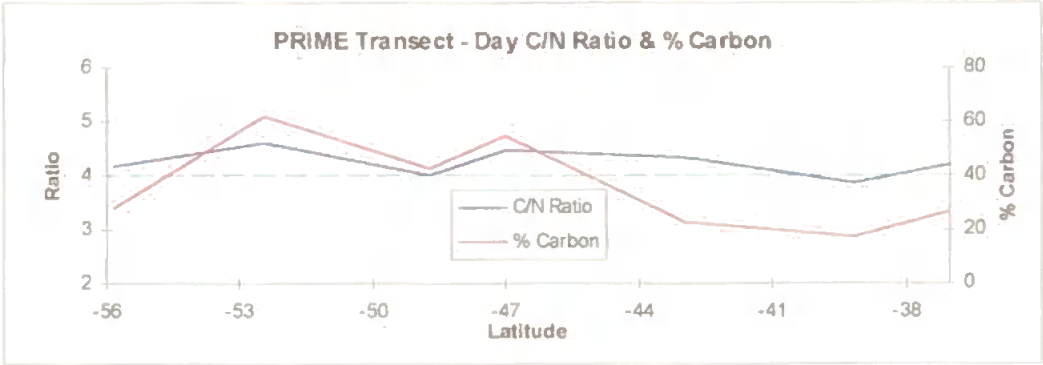


Fig. 8.63 PRIME daytime WP-2 net samples - carbon to nitrogen ratio and carbon as percentage of total biomass by dry weight.

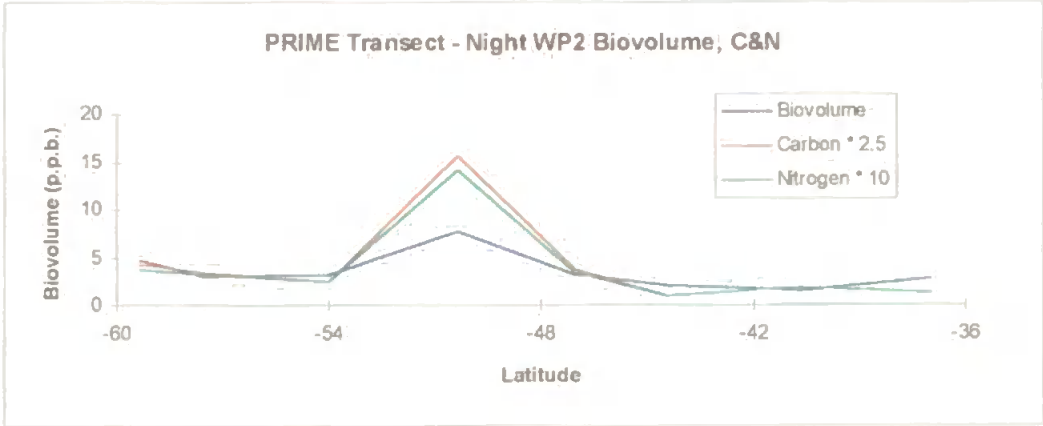


Fig. 8.64 PRIME transect nighttime WP-2 net samples - total OPC biovolume, carbon and nitrogen.

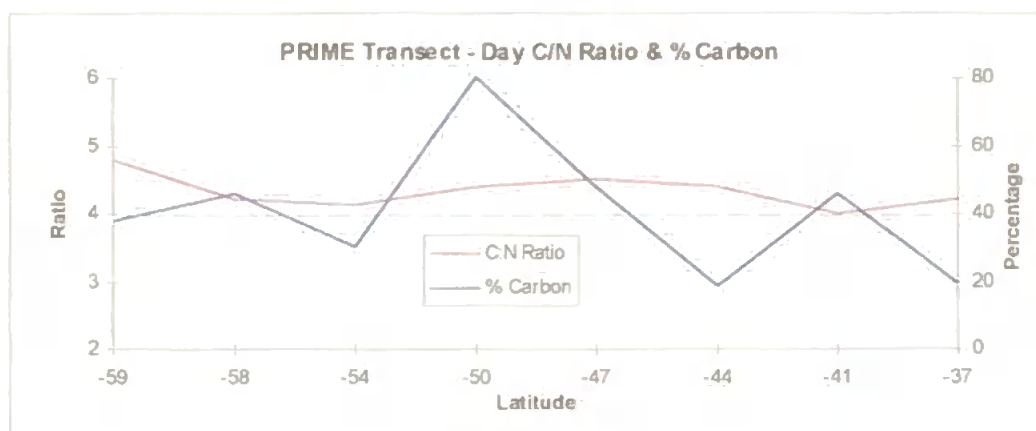


Fig. 8.65 PRIME transect nighttime WP-2 net samples - carbon to nitrogen ratio and carbon as percentage of total biomass by dry weight.

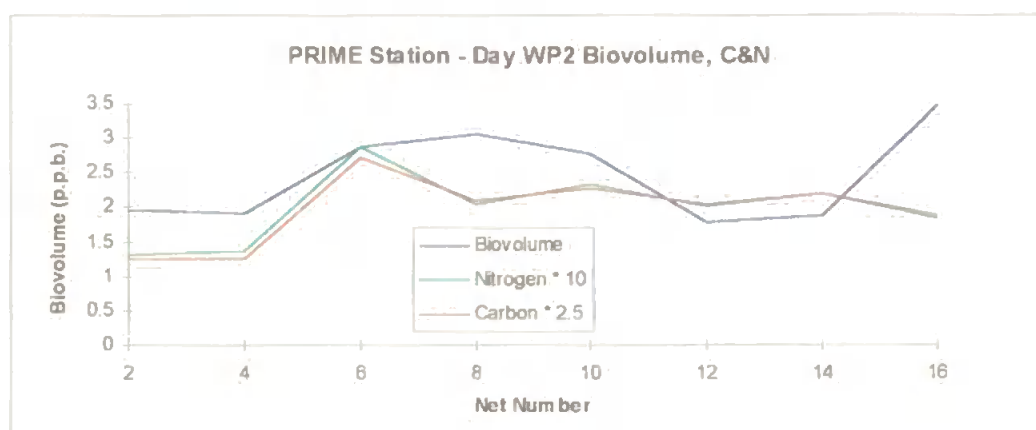


Fig. 8.66 PRIME station daytime WP-2 net samples - total OPC biovolume, carbon and nitrogen.

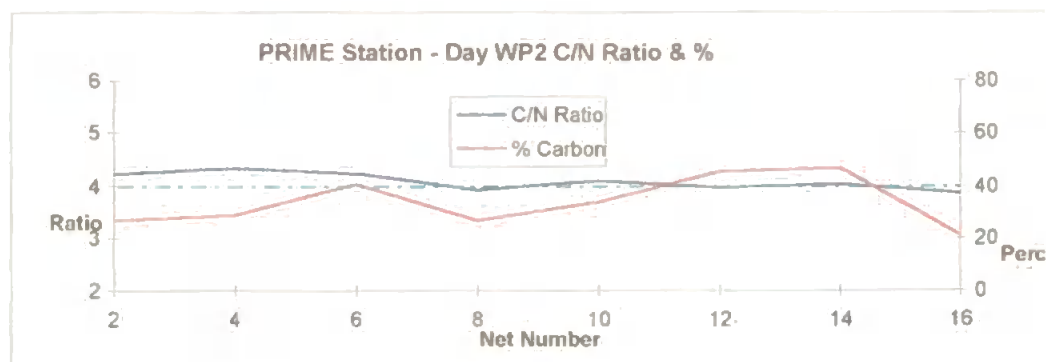


Fig. 8.67 PRIME station daytime WP-2 net samples - carbon to nitrogen ratio and carbon as percentage of total biomass by dry weight.

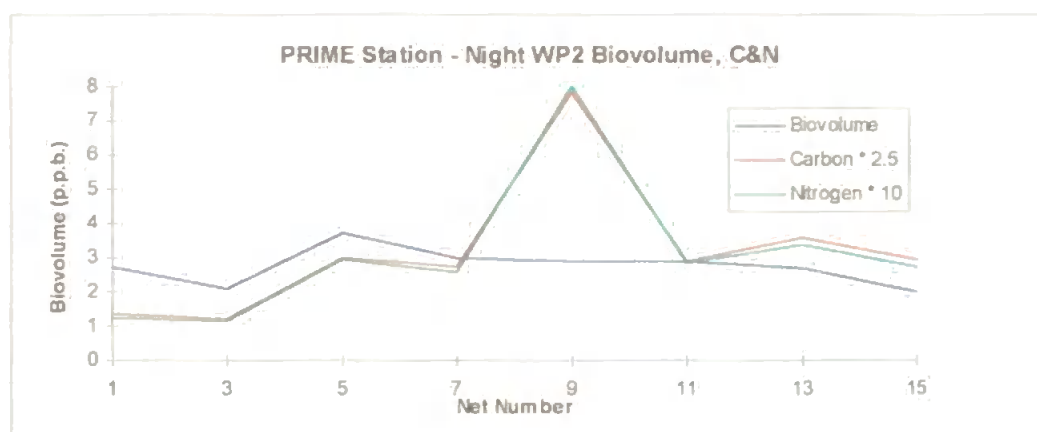


Fig. 8.68 PRIME station nighttime WP-2 net samples - total OPC biovolume, Carbon and nitrogen.

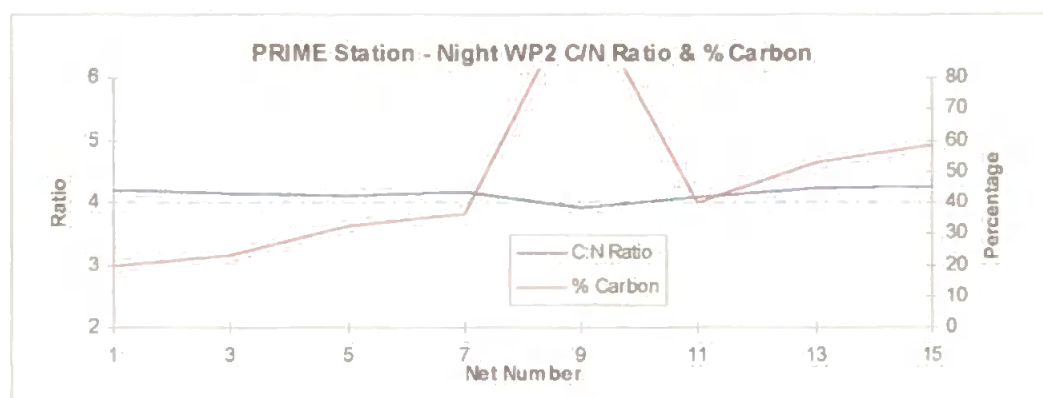


Fig. 8.69. PRIME station nighttime WP-2 net samples - Carbon to nitrogen ratio and carbon as percentage of total biomass by dry weight.

The graphs for the station at 37° north again show C:N ratios fairly steady at 4:1 (fig. 8.67). Carbon and nitrogen as a percentage of dry weight is fairly steady around the expected values during the day at this station. At night (fig. 8.69) they rise steadily from about 20% carbon on the first day to about 60% carbon on the last, apart from day 9, where there is an anomalously high value of over 100% carbon. This peak does not appear in the OPC biovolume data.

# 8.7.4 AMT3

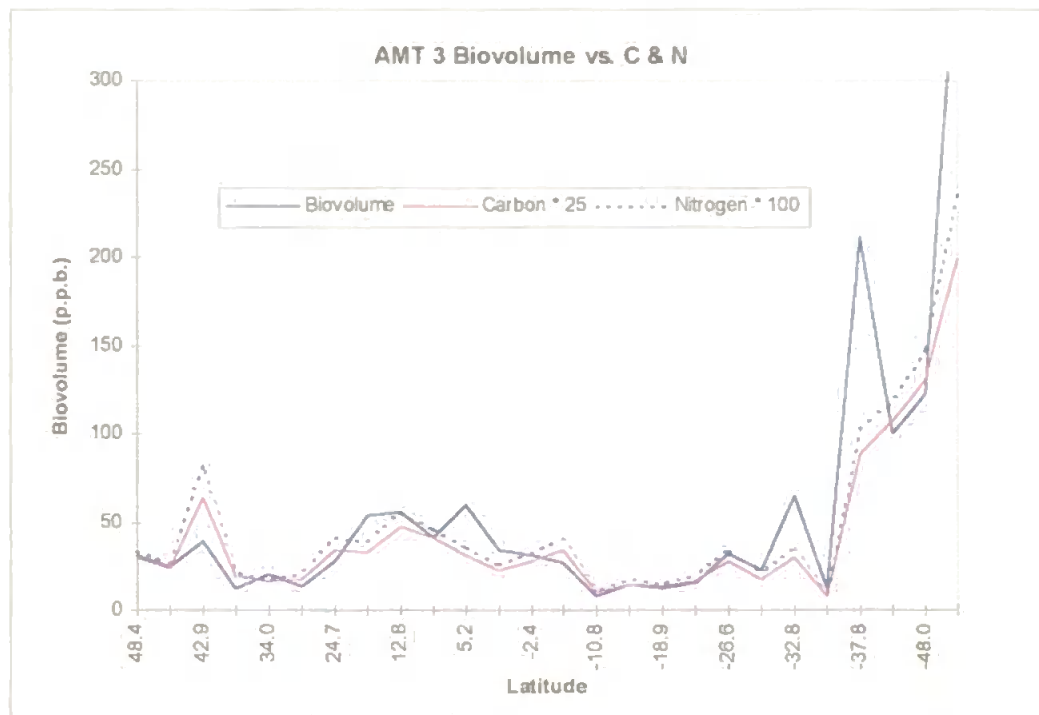


Fig. 8.70 AMT3 WP-2 net samples - total OPC biovolume, carbon and nitrogen.

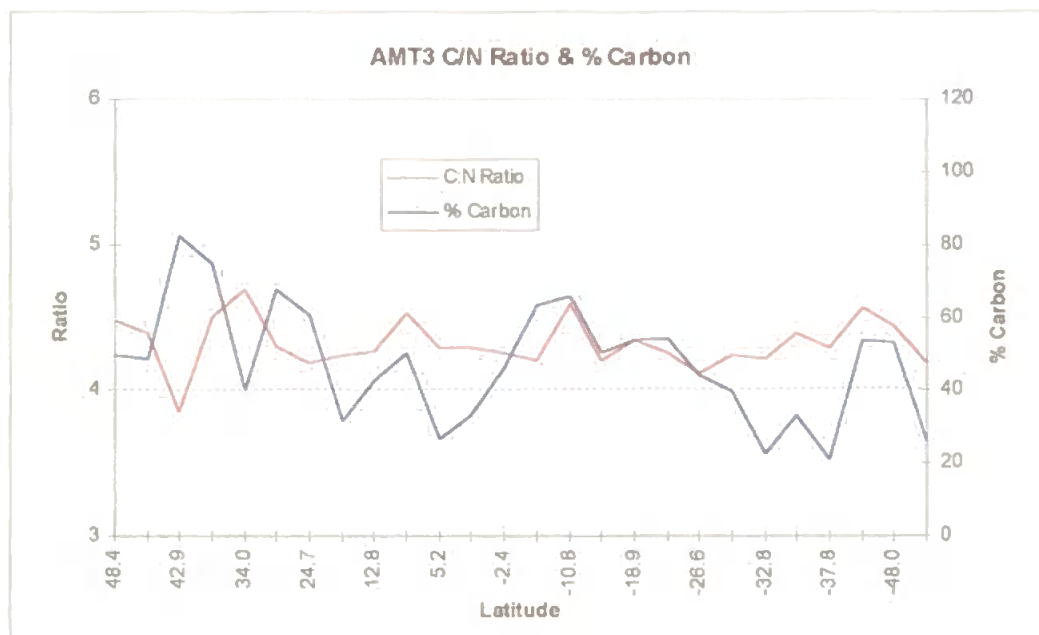


Fig. 8.71 AMT3 WP-2 net samples - carbon to nitrogen ratio & % carbon.

Figure 8.70 again shows a reasonable match between OPC biovolume and carbon at 40%, nitrogen at 10%. Carbon/nitrogen is low at 32.8° south, when the OPC log shows a large biovolume of siphonophores were identified visually in the sample. These animals typically have a carbon content of about 6-12% (Curl, 1962), accounting for the lower carbon content. High concentrations of gelatinous organisms may similarly account for the low carbon content of the net samples at 37.8 and 52° south. The total carbon as a percentage of OPC biovolume fluctuates between 20% and 80% with a mean value of about 45%, a fairly typical average value for marine zooplankton (Curl, 1962, Omori, 1969). Carbon to nitrogen ratio is fairly flat at about 4.3 over the whole transect, as for AMT1 (typical for marine zooplankton, see Omori, 1969), rather than peaking at high latitudes as for AMT2. Whether this difference is due to seasonal fluctuations or some anomaly may emerge as further transects are completed in 1997/8.

# CHAPTER 9

## DISCUSSION AND CONCLUSIONS

### 9.1 OPC Data summary

The AMT1 cruise was important, not only to validate the use of the OPC-1L for long ocean transects, but also to provide an invaluable opportunity to expose the system to a range of operational conditions. Some potential sources of error exist in the application of the instrument in this mode (see chapter 3). The success of this preliminary deployment stems partly from the development of techniques to evaluate these, and also in implementing operational protocols to minimise them.

AMT2 was complimentary to AMT1, following more or less the same track in reverse, with the seasons reversed north and south. The OPC total biovolume data for both cruises, plotted in fig. 9.1, illustrates this seasonal reversal quite well. The contrast is apparent at the two extremes of the cruises, the first of which took place during the southern spring, and the second during the northern spring. Generally similar high values can be seen in each hemisphere during spring. The West African upwelling influence is more or less coincident for each transect, at about 20° north, although AMT 2 shows higher values at the peak.

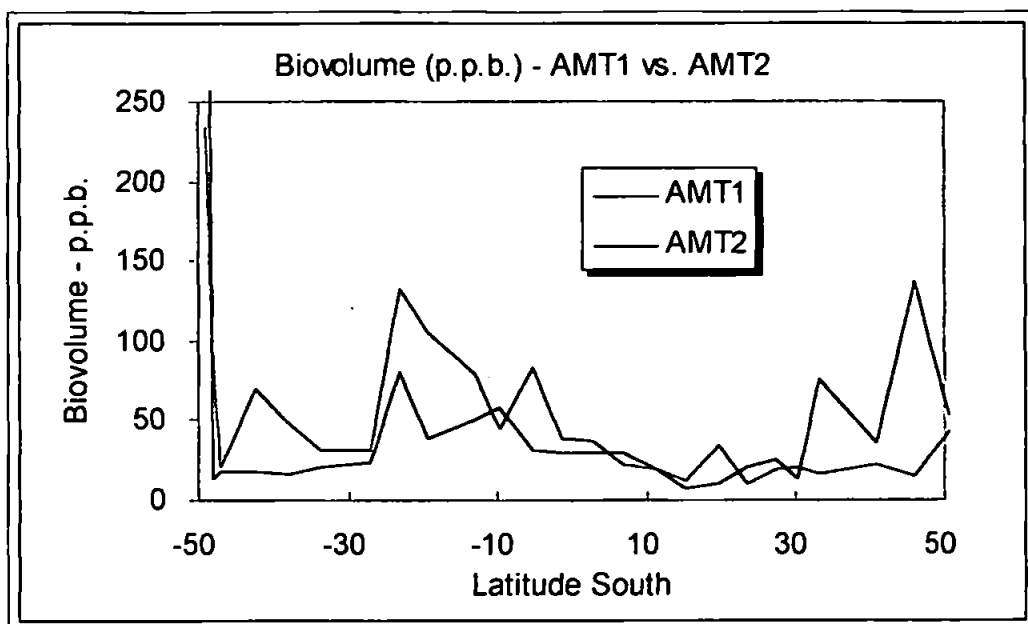


Fig. 9.1 OPC total biovolume comparison between AMT1 and AMT2.

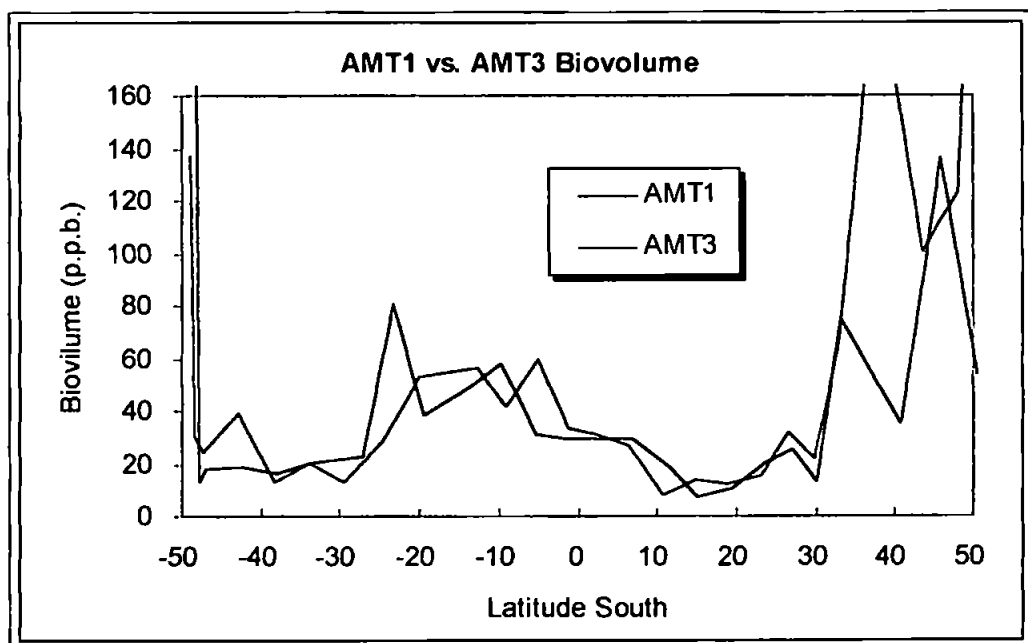


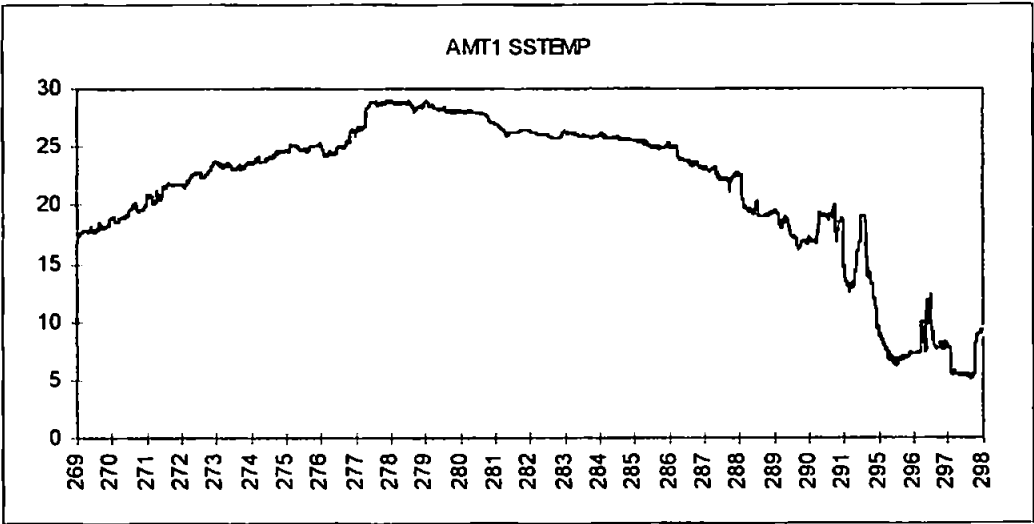
Fig. 9.2 OPC total biovolume comparison between AMT1 and AMT3.

Figure 9.2 shows the total biovolume in parts per billion for AMT3, with the same data for AMT1 plotted for comparison. These cruises took place in the same season, one year apart. Biovolume was generally distributed similarly to AMT1, with the exception of a peak of about 40 p.p.b. at 43° N., which did not appear on AMT1. The double peak around the West African upwelling system was of similar magnitude to AMT1, but centred further south at 9° N. rather than 19° N. on AMT1. As with AMT1 the plot shows a ridge of enhanced biovolume stretching from the African system across the equator to about 7° S., at about 30 p.p.b. in both cases. Biovolume minima in the Brazil basin were about 10 p.p.b. in both cases. At about 33-36° S. a value of about 70 p.p.b. also appears in both data sets. Size distributions are generally similar in both data sets, although the 1000-2000  $\mu\text{m}$ . size class in the African system was less dominant than on AMT1.

#### *9.1.1 Other environmental data*

During the examination of the zooplankton data from the four cruises, some attempt was made to seek correlations between this data and other environmental data gathered at the same time, such as temperature, salinity, chlorophyll fluorescence, carbon 14 productivity, etc. Very little correlation was evident with any of this data, apart from the obvious mild inverse relationship to sea surface temperature. This parameter, as shown for example for AMT1 in fig. 9.3, tends to be lower at high latitudes and high at low latitudes, with a latitudinal shift towards the northern hemisphere in September (after the

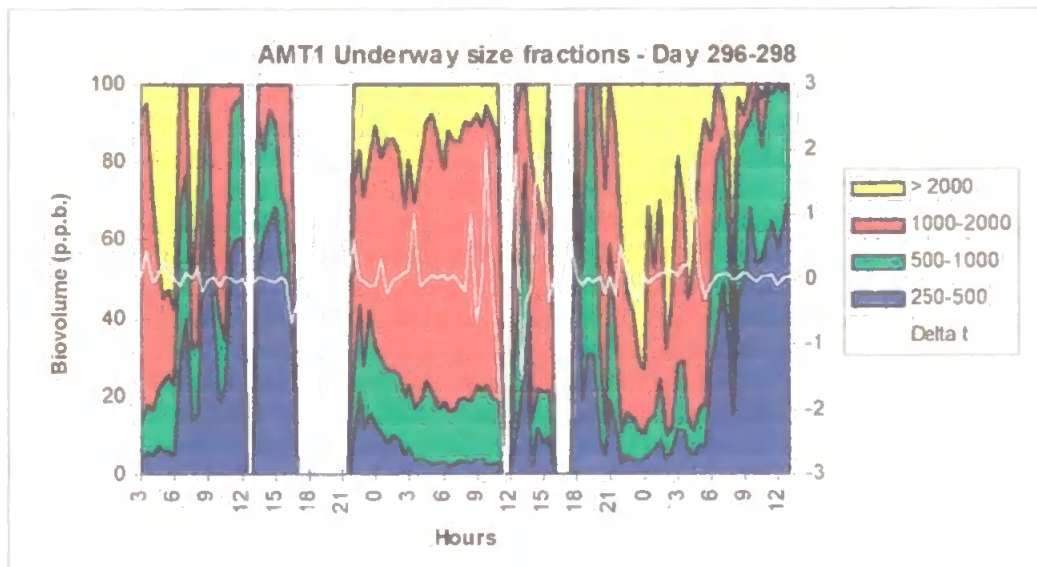
northern summer) and towards the southern hemisphere in April (after the southern summer).



*Fig. 9.3 Surface sea temperature for AMT 1, days 269-298 (north to south).*

Within this general trend, certain frontal areas exist, where temperature changes quite rapidly, and these areas are often associated with increases in zooplankton biomass, although these increases rarely occur at the fronts, but some distance either side. While it might be expected that there would be a stronger correlation between zooplankton biomass and some of the other environmental parameters, it must be remembered that while phytoplankton populations can respond to changing environmental conditions in a matter of hours, populations of zooplankton require days or weeks to develop in response to such changes. Given steady currents, it is likely that changes in zooplankton biomass may therefore be found removed both in time and space from the conditions initially responsible for these changes.

It was noticed during some parts of the AMT cruises that large changes in both biomass and the size distribution of the zooplankton population occurred either side of frontal areas, as indicated by changes in physical and other biological parameters. The area of the confluence of the Brazil and Falkland currents is an area which shows strong fluctuations of such parameters. Figures 9.4 to 9.6 show the size distribution of the zooplankton plotted with the parameter  $\Delta T$ , the rate of change of surface temperature with distance, integrated over half hour periods for this area for all three AMT cruises.



*Fig. 9.4 Zooplankton size distribution, and rate of change of temperature during transect of the area of the confluence of the Brazil and Falkland currents, between Montevideo and the Falkland Islands, AMT1.*

Figure 9.4 shows a strong fluctuation in the parameter  $\Delta T$  at around noon on day 297, followed by a change in zooplankton biomass and size distribution about 9 hours later. Figure 9.5 shows similar fluctuations in the rate of change of surface temperature of  $+4^{\circ}\text{C}$  at around noon on day 115, three hours after a strong change in zooplankton size distribution, and  $-6^{\circ}\text{C}$  at around 6.00 a.m. on

day 116, about 6 hours after such a change. Similar, although perhaps less well defined features can be seen on fig. 9.6 for AMT3.

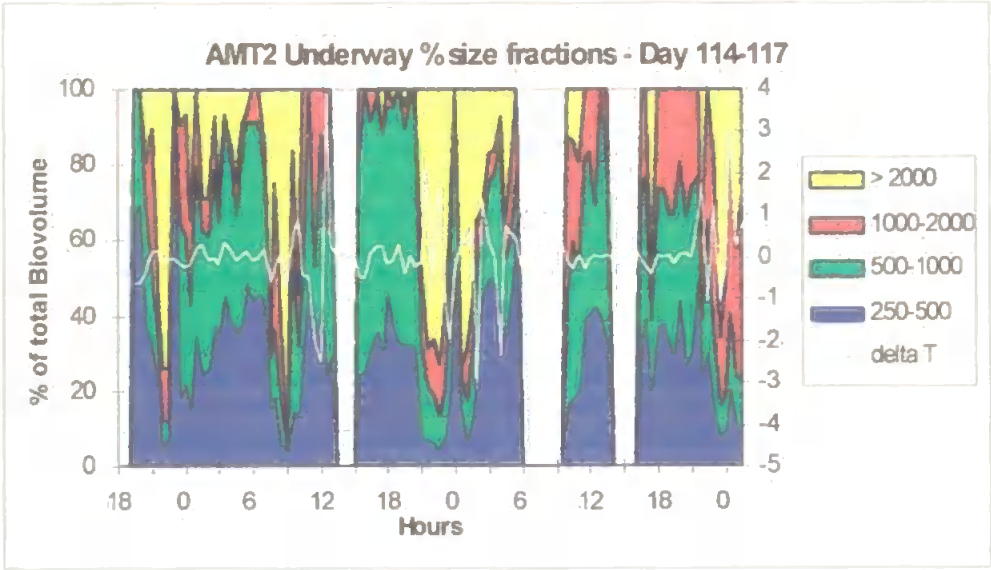


Fig. 9.5 Zooplankton size distribution, and rate of change of temperature during transect of the area of the confluence of the Brazil and Falkland currents, between Montevideo and the Falkland Islands, AMT2.

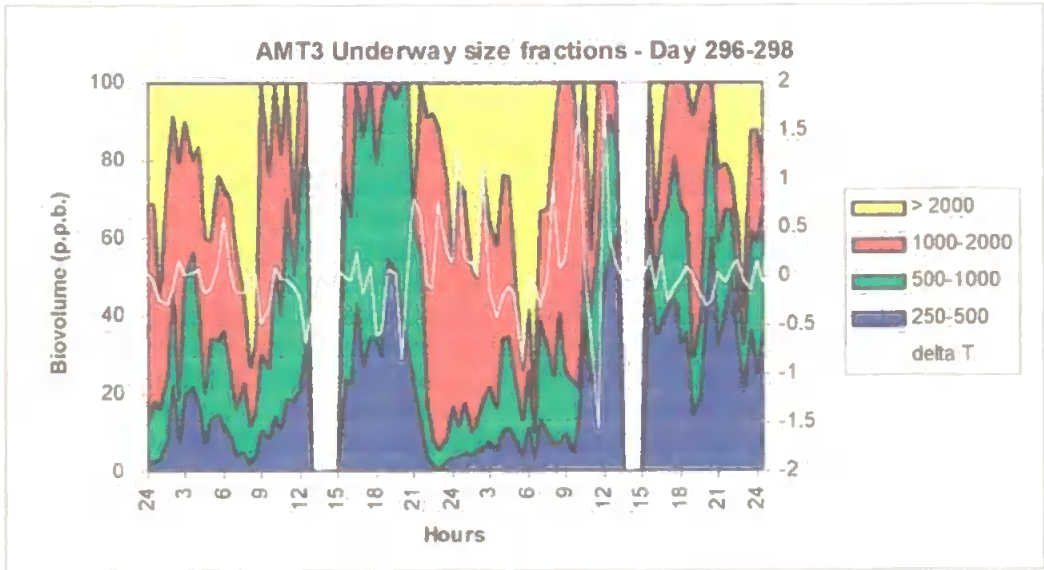


Fig. 9.5 Zooplankton size distribution, and rate of change of temperature during transect of the area of the confluence of the Brazil and Falkland currents, between Montevideo and the Falkland Islands, AMT3.

While rather more research would be required to confirm such a hypothesis, it is interesting to speculate that the real-time data on zooplankton biomass and size distribution obtained from the systems described in this report might in certain circumstances be used to predict the approach of areas of frontal activity.

## **9.2 ViZA data summary**

The ViZA system is still under development, and the data shown in chapter 8 indicate that further improvements in the performance of the system are required before it can be used operationally. The problems with the ViZA system are two-fold. Firstly, the system is undersampling, particularly when animals are least abundant (i.e. animals of ESD greater than 2 mm., and for all animals during the day in oligotrophic waters). This is due to limitations in the data processing rate, and particularly the residual delay in response of the frame processor card to requests for the next object outline. Secondly, many of the more transparent objects are poorly silhouetted by backlighting with white light. Proposed solutions to these two problems are given below. As illustrated by the results presented in chapter 8, where data are relatively free of these problems, the system can classify organisms to one of six major taxonomic groups quite effectively. Correlation between biovolume data from ViZA and the OPC is quite good, although only a limited number of data points are available.

It would be expected that ViZA biovolumes would be a little less than those determined by the OPC, due to the transparency problem with some

zooplankton species. This would mean that the size determined by ViZA would be less than the actual size of the organism. This was generally true up to day 275 - see for example fig. 8.42. After this day, however, there was a trend for ViZA biovolumes to be greater than OPC biovolumes, particularly on the last day. This was probably due to the fact that after the camera was temporarily repaired, the picture deteriorated over the next three days, growing darker. The effect of this would be for the system to detect the threshold of light to dark, indicating an object of interest, appearing further from the centre of the object. The determined area and boundary would therefore be larger.

### **9.3 Data validation**

Validation for the OPC biovolume data is only available from carbon and nitrogen analysis at this time, due to the cost of taxonomic analysis of the preserved samples. It is hoped that limited taxonomic analysis of samples from AMTs 1-3 will be performed during 1997. Carbon and nitrogen analysis of samples from AMT and PRIME cruises, based upon empirically derived average values for these components of marine zooplankton, offer some verification of the figures for biovolume obtained from the OPC.

### **9.4 System performance**

Both the OPC and the ViZA system, as used so far in this study, were operated in benchtop mode. This means that they share the same spatial resolution, defined by the volume flow rate through the flow cells (which are identical for both systems) and the speed of the vessel through the water. In order to make an

assessment of the spatial resolution of the system we must first determine what kind of sample we need. In order to obtain a reasonable size distribution we need to sample a volume of water containing a statistically significant number of animals. Given that copepods may make up 80-90% of some pelagic samples, the minimum number of animals for a reasonable distribution might be a few hundred. Such a sample may not include animals in all size classes, but if we are to insist on all size classes being present it might sometimes be necessary to sample many cubic metres and many thousands of animals, which seems unreasonable. For the purpose of this argument we shall settle on a figure of 200 animals. During most of the AMT and PRIME cruises this number of animals at worst would appear in two cubic metres. At the flow rates used ( $25 \text{ litres min}^{-1}$ ) this volume of water would pass through the system in 80 minutes, and at normal cruising speed for RRS James Clark Ross of 11 kts. this sample would be taken over about 25 km., and this would be the spatial resolution of the system. For all but the most oligotrophic waters, concentrations are at least twice this level, and spatial resolution proportionally better. While this order of spatial resolution is acceptable for gross studies for the purposes of modelling global processes, it is rather too coarse to meet all the objectives set out in chapter 2. Using the towed system, whose flow channel is ten times wider, the worst case spatial resolution will fall to about 2.5 km., rather better, but still not sufficient to resolve very small patterns of spatial distribution (small-scale patchiness). In many cases, where better resolution is required (coastal areas, areas of frontal activity, etc.), abundance is higher than this worst case.

#### *9.4.1 OPC system performance*

The towed OPC is already a well proven device in small and medium scale zooplankton survey (Sprules *et al.*, 1992, Heath, 1995, Herman, 1992, Huntley *et al.*, 1995). It is believed that the AMT project is the first time that the laboratory device has been used in 'real-time' zooplankton survey, and the first time either device has been used to sample over such large spatial scales. The device offers the means to sample at such large scales very simply and reliably, delivering data in near real time. For the first time on such large cruises, biological data can now be made available in the same time scale as physical data, using an instrument which, once set up, requires little or no intervention while sampling over thousands of kilometres of ocean transect. The OPC operated in this mode offers opportunity of large scale zooplankton survey of ocean basins using ships of opportunity (Robins & Aiken, 1996). Such an opportunity would allow comprehensive survey for zooplankton distributions of large areas of the world ocean for which up to now little or no data is available. The South Atlantic has been particularly poorly sampled, and the data from these instruments gathered during the AMT cruises could be invaluable for large scale ecosystems modellers. It is essential to obtain such data on these scales in order to understand and predict the biological effects on and responses to global changes in climate through fluxes of carbon in the atmosphere/ocean system. The provision of unattended sampling platforms will play an important role in the acquisition of such data (Denman *et al.*, 1985).

9.4.2 ViZA system performance

It was found that the problem of poor imaging due to object transparency did not occur with samples preserved in formaldehyde, as the transparent organisms quickly become opaque after preservation. The copepod shown in fig. 7.1 is a preserved sample, and the outline is clearly defined. Some copepods sampled live are rather more transparent, and fig. 9.7 shows an example of this.

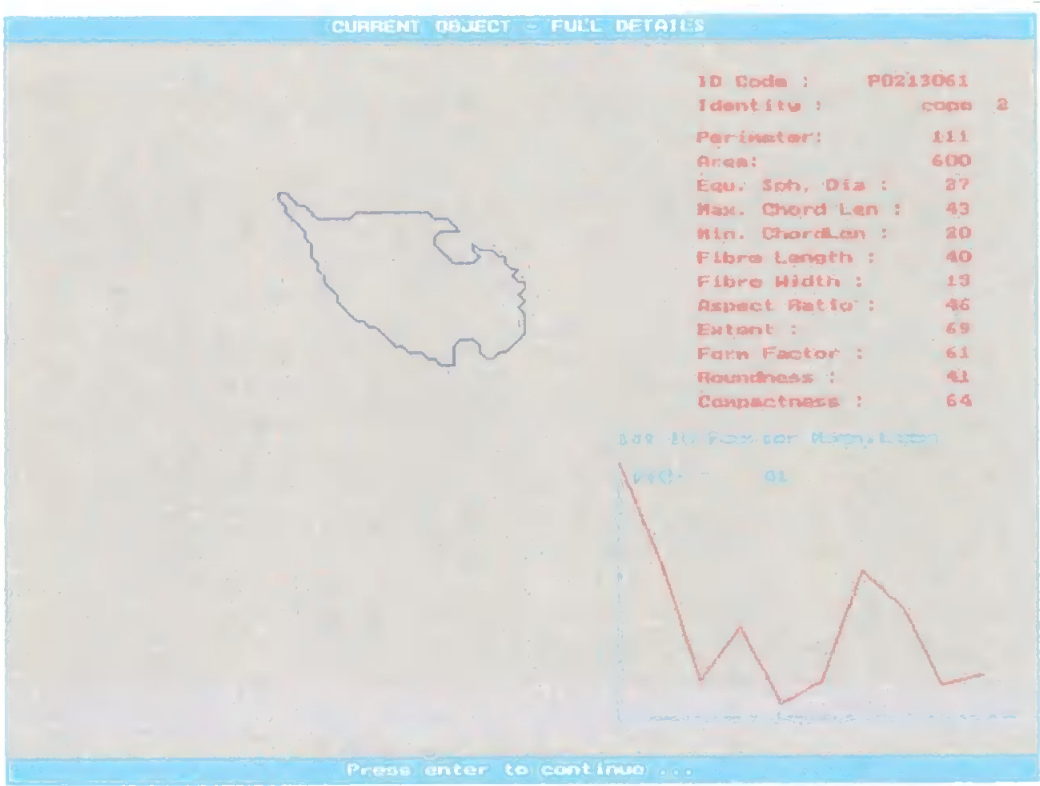


Fig. 9.7 Example from AMT3 ViZA data set showing effect on outline of transparency to white backlighting.

This example has been chosen because, while transparency effects are visible, the animal is still visibly identifiable as a copepod. There are far worse examples of this problem. Because of this, the training set of 200 animals used for AMT3 samples was acquired by hand picking suitable examples of each organism group to be determined, from a sample filtered from the seawater

supply, and preserving these in formaldehyde. This sample was repeatedly imaged until a sufficient number of good images of all organisms had been obtained for use as a training set.

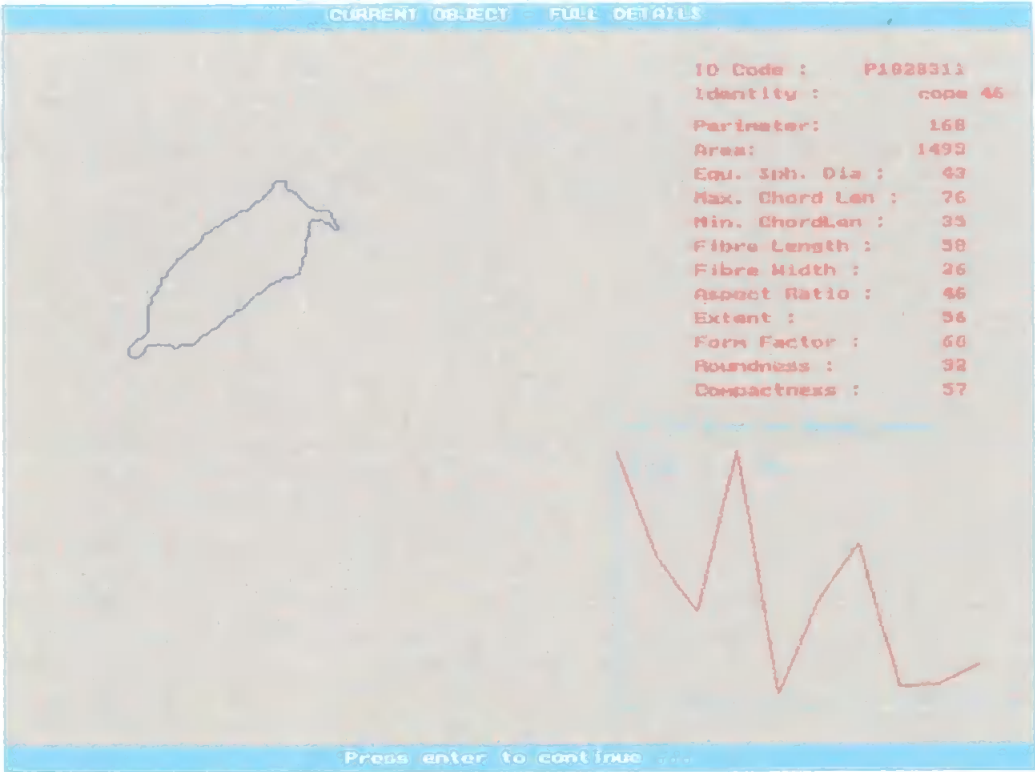


Fig. 9.8 Example from the AMT3 ViZA data set as fig 9.3, showing little or no effect on outline of transparency to white backlighting.

Figure 9.8 shows another coeoid from the same period during AMT3 whose image outline does not suffer from this effect. Note that both these examples were successfully identified as coeoids, as were examples which were much more degraded. This is due to the fact that the basic size and shape descriptors derived from typical coeoid body shape are very distinctive. Other animals are not quite so easily discriminated. The addition of more groups to be discriminated, so that the groups cluster closer together in the space defined by the principal components, means that even coeoid outlines degraded in this

way become less identifiable. Given that we require to differentiate rather more than five or six major groups, this problem must be addressed.

In spite of the problems discussed above, when objects are well imaged, and when abundance is sufficient for the ViZA system to acquire a sufficient number of object outlines, the system is able to gather data whose size distribution correlates very well with that of the OPC, and is able to classify these animals to one of six taxonomic groups with no overlap of groups.

## **9.5 Further development work**

### *9.5.1 OPC Development*

After deployment on four major oceanographic cruises during 1995-6, covering almost 40,000 km. of transect in the North Atlantic, protocols for the deployment of the OPC-1L are well established. Further AMT cruises planned for 1997-8 will use the same protocols, providing a four year time series of synoptic data on zooplankton abundance and size distribution in the upper mixed layer of the North and South Atlantic Oceans.

Because the vertical net samples are integrated over 200 metres we are producing artificial community aggregates which do not really exist at any given level within the water column (Platt, 1985). In order to address this problem, the author is presently engaged at Plymouth Marine Laboratory in the development of a vertically profiling OPC system, using the OPC-1T (towed OPC) and a dedicated *in-situ* data logger. This unit will also detect and log

conductivity, temperature, depth and chlorophyll fluorescence. Such a system will offer much more detailed information on vertical distributions of zooplankton, and of diel vertical migration of zooplankton within and into and out of the mixed layer.

It is intended that the OPC-IT will also be used on future AMT cruises in towed mode mounted on the UOR. The logger system currently being developed for the vertical profiling system will be equally applicable to the towed system, offering the possibility of avoiding the complexity of conductor core cable and associated specialist winch gear for obtaining data 'up the wire'.

It is also intended to conduct trials with the OPC (either the benchtop or towed version) with the CPR. CPR surveys along regular shipping routes in the North Atlantic have been made continuously over the last 60 years (Glover, 1967, SAHFOS, 1995). This represents an unusual and valuable time series. The CPR sampling system relies upon human analysis of preserved samples for taxonomy and size distributions, with all associated costs outlined in chapter 2. If modern systems such as the OPC/ViZA system can be shown to produce data comparable in scope and reliability to the CPR, it could offer a more cost effective means of continuing and extending this time series. The applicability of the OPC-IL to ship of opportunity survey is particularly promising in this regard.

### 9.5.2 ViZA development

The sampling rate and image quality problems discussed above prevent the ViZA system in its current form from meeting the requirements specified in chapter 2. The sampling rate problem might be overcome by the use of a more powerful frame acquisition and processing system. This would be an expensive solution, the cost of such systems being well into the tens of thousands of pounds. Filtering of the white strobe light to use only the red has been tried, but insufficient light passes through the optical system and small aperture (for high depth of field) to produce good contrast. In any case, the present strobe system is already too bulky and demanding upon power supplies to be applied to an *in-situ* system. The short light pulse duration (  $<10\ \mu\text{s}$  ) means that it is difficult to get sufficient light from red LEDs for similar reasons.

It is proposed to overcome both these problems by switching from a two-dimensional area scan camera to a one-dimensional line scanning array. This will result in continuous imaging 512 pixels wide of the flow past the sensor without the electronic and computational overhead involved in the formation, acquisition and processing of two-dimensional images. ‘Vertical ‘ resolution in a line scanning system at a given line rate would clearly be a function of flow rate past the sensor. The horizontal pixel size is  $40\ \mu\text{m}$ . In order to obtain this same pixel size along the flow, at present flow rates of  $1\ \text{m/s}$ , we should need a line scanning period of  $40\ \mu\text{s}$ , or a line rate of  $25,000\ \text{sec}^{-1}$ . For a towed system, flow would be more like  $5\ \text{m/s}$ , and the camera should be capable of operation at  $125,000\ \text{lines sec}^{-1}$ . This order of performance has recently become

available in line scan sensors of 512 pixel width, and continuous scanning of the flow through the OPC would be feasible, overcoming the present sampling rate problem.

Strobe lighting would be unnecessary with such a system, and a laser line projector could be used, producing a thin line of non-divergent light at the line sensor, back illuminating any object passing the sensor. Red/IR laser line projectors are available cheaply, and with a simple cylindrical lens to collimate the divergent line from the laser onto the linear sensor, an afocal system could be used. This will obviate the need for the complex, small aperture optics of the present system, reducing size and weight considerably. The Xenon strobe system could also be discarded, a further saving on bulk and weight, and on power requirements. Such a system would be much easier to transfer to an underwater towed platform. This afocal system could be used with the present frame processor card and system software (slightly modified). An extra processor card connected to the F64 on a fast 80 MHz bus could be added to help with pre-processing. Subject to funding it is proposed to take this route in the future development of the ViZA system.

## **9.6 System application**

As stated in chapter 2, there is a perceived need to measure zooplankton abundance over all spatial scales, resulting in data suitable for quick and easy subsequent processing, or better still, in results produced in real time. Due largely to the time and economic cost of skilled human analysis of samples

obtained, especially on the very large spatial scales considered in this study, only automated acoustic, optical and video methods are capable of this level of performance, and only the latter two can cover the entire range of sizes required.

There has been some disagreement over the last ten to fifteen years over whether a description of abundance and size distribution alone is sufficient to characterise oceanic zooplankton to meet the needs of those who seek to use such information to predict the influence of these distributions on global changes, and the effects of such global changes on these distributions. Initial classification may be aggregated - e.g. as trophic levels - rather than detailed analysis to species level. Ecosystems are often represented as flows of energy between these groupings (Platt, 1985).

Many researchers in this area have sought to demonstrate that metabolic rates in zooplankton are a function only of body size (Platt, 1985, Sprules, 1980). Very often even those who argue this case concede that some categorisation other than size is required. Silvert & Platt (1980) require also classification as plant, animal or detritus. Cousins (1980) states that information on size alone is insufficient, requiring additional categorisation as autotrophs, heterotrophs and detritus. Frost (1980) however, states that processes other than metabolic rates (ingestion rates, filtration efficiency, etc.) are dependent upon size related functions which vary between species.

Even those researchers who demonstrate methods of calculating rates based only upon size admit that the reasons are largely pragmatic, based upon the possible rather than what is desirable. Platt (1985) states that “....conventional taxonomic description of a typical pelagic sample is highly demanding on time”, and that “.. the beauty of such an approach...(*based on size only*)...is that it is realisable in operational terms”, and Sprules (1980) in defending the size based approach makes the aside that size distributions are easier to obtain than full taxonomies. Furthermore, the need for a fuller taxonomic description of pelagic zooplankton distributions may be tempered by the ability to incorporate such full descriptions in subsequent models. Cousins (1985) states that “...while the aim of the trophic continuum model is to define the ecosystem with the minimum of taxonomic distinction, this is controlled by the precision of the output which is required.”

## **9.7 Conclusion**

Most researchers seem to agree that *some* information in addition to size is required, and even the fundamental distinction between living and non-living particles cannot be made by optical particle samplers. The description of shape which is required to make such distinctions, among the current modern methods being evaluated, is only available from video devices.

No single one of these methods is therefore currently capable of characterising zooplankton in real time fully enough to meet even the minimum requirements in the characterisation of pelagic zooplankton communities. It has been the

intention of this study to show that a combination of two of these methods - optical and video - can provide such a capability. While it is not claimed that this capability has yet been achieved, it is believed that the results presented are encouraging enough to indicate that further development of this system can provide such a capability.

To conclude, and in summary:

- It is generally agreed that new techniques are required for gathering data on the abundance, species, size and spatial distribution of zooplankton autonomously, and in real time over large temporal and spatial scales.
- Of the novel technologies currently available or under development, none is alone capable of fulfilling this requirement.
- It has been demonstrated that a combination of two of these technologies - optical and video - has the potential to fulfill this requirement.
- Novel methods in the deployment of the OPC-IL show that this instrument could usefully deployed as an autonomous zooplankton sampler in ship of opportunity mode.
- The new video instrument developed in this study shows sufficient promise to warrant further development.

# REFERENCES

Aebischer, N.J., Coulson, J.C. & Colebrook, J.M. (1990). Parallel long term trends across four marine trophic levels and weather. *Nature* **347**:753-755.

Aiken J. & Bellan, I. (1990). Optical oceanography: An assessment of a towed method. In: P.J. Herring, A.K. Campbell, M. Whitfield & L. Maddock (Eds.), *Light and life in the sea*. Cambridge University Press, London.

Banse, K. (1995). Zooplankton: pivotal role in the control of ocean production. *ICES Journal of Marine Science* **52**(3/4): 265-278.

Berman, M.S. (1990a) Enhanced zooplankton processing with image analysis technology. *ICES C.M.* L:20.

Berman, M.S. (1990b) Application of image analysis in demographic studies of marine zooplankton in large marine ecosystems. In: Sherman, K., Alexander, L.M. & Gold, B.D. (Ed.): *Large marine ecosystems: patterns, processes and yields*. pp. 122-131, AAAS, Washington.

Bernard, M., Möller, F., Nassogne, A., & Zettera, A. (1973). Influence of pore size of plankton nets and towing speed on the sampling performance of two

high-speed samplers (Delfino I & II) and its consequences for the assessment of plankton populations. *Marine Biology*, **20**:109-136.

Blackburn, M. (1981). Low latitude gyral regions. In: Longhurst, A.R. (Ed.) *Analysis of marine ecosystems*. Academic Press, London.

Bracewell, R.N. (1986) *The Fourier transform and its application*. McGraw-Hill, New York.

Cousins, S.H. (1985). The trophic continuum in marine ecosystems: Structure and equations for a predictive model. In: *Ecosystem theory for biological oceanography*. Ulanowicz, R.E. and Platt, T. (Eds.). Ottawa:Department of Fisheries and Oceans, p. 76-93.

Curl, H. (1962). Analyses of carbon in marine plankton organisms. *Journal of Marine Research* **20**(3):181-188.

Cushing, D.H., (1981). Temporal variability in production systems. In: Longhurst, A.R. (Ed.) *Analysis of marine ecosystems*. Academic Press, London.

Davis, C.S., Gallagher, S.M., Berman, M.S., Haury, L.R. & Strickler, J.R. (1992) The Video Plankton Recorder (VPR): design, and initial results. *Arch. Hydrobiol./Beih. Ergebn. Limnol.* **36**:67-81.

Davis, C.S., Gallager, S.M., Marra, M. & Stewart, W.K. (1996). Rapid visualisation of plankton abundance and taxonomic composition using the Video Plankton Recorder. *Deep Sea Research (II)* 43(7-8):1947-1976.

Denman, K., Calder, W., Davis, C., Demers, S., Estrada, M., Lewis, M., Smith, D., Wangersky, P., Yentsch, C., and Zotin, A. (1985). Technological developments to implement theory into biological oceanography. In: *Ecosystem theory for biological oceanography*. Ulanowicz, R.E. and Platt, T. (Eds.). Department of Fisheries and Oceans, Ottawa. p. 254-258.

Flagg, C. N. & Smith, S. L. (1989). On the use of the RDI acoustic Doppler current profiler to measure zooplankton abundance. *Deep-Sea Research* 36(3):455-474.

Focal Technologies, Inc. (1992) *Optical Plankton Counter User's Guide*. Focal Technologies, Inc., Dartmouth, Nova Scotia.

Frost, B.W. (1980). The inadequacy of body size as an indicator of niches in the zooplankton. In: *Evolution and ecology of zooplankton communities*. Kerfoot, W.C. (Ed.). N.H. University Press of New England, Hanover, N.H., p. 742-753.

Freeman, H. (1962) On the encoding of arbitrary geometric configurations. *IRE Transactions on Electronics and Computing*, EC-10:260-268.

Gallienne, C.P. (1994). *An automated control system for a liquid crystal shutter*. Unpublished report, University of Plymouth.

Gallienne, C.P., Robins, D.B. & Pilgrim, D.A. (1996). Measuring abundance and size distribution of zooplankton using the optical plankton counter in underway mode. *Underwater Technology* 21(4):15-21.

Genin, A., Greene, C., Haury, L., Wiebe, P., Gal, G., Kaartvedt, S, Meir, E., Fey, C. & Dawson, J. (1994). Zooplankton patch dynamics: daily gap formation over abrupt topography. *Deep-Sea Research* 41(5/6):941-951.

Glover, R.S. (1967) The continuous plankton recorder survey of the North Atlantic. *Symp. Zool. Soc. Lond.* 17:189-210.

GLOBEC (1992) *GLOBEC workshop on optics technology*. Draft report. Skidway Institute of Oceanography, Savannah, Ga.

Gonzalez, R. C. & Woods, R. E. (1993) *Digital Image Processing*. Addison - Wesley, Reading, Ma.

Hand, D.J. (1981). *Discrimination and classification*. John Wiley & Sons, Chichester.

- Haury, L. R. & Pieper, R. E. (1988). Zooplankton: scales of biological and physical events. In: *Marine organisms as indicators*. D. F. Soule & G. S. Kleppel, eds., Springer-Verlag, New York.
- Heath, M.R. (1995). Size spectrum dynamics and the planktonic ecosystem of Loch Linnhe. *ICES Journal of Marine Science* **52**(3/4):627-642.
- Herman, A.W. & Dauphinee, T.D. (1980) Continuous and rapid profiling of zooplankton with an electronic counter mounted on a 'Batfish' vehicle. *Deep-Sea Research*, **27A**:79-96.
- Herman, A.W. (1988) Simultaneous measurement of zooplankton and light attenuation with a new optical plankton counter. *Continental Shelf Research*, **8**(2): 205-221.
- Herman, A.W. (1992) Design and calibration of a new optical plankton counter capable of sizing small zooplankton. *Deep Sea Research*, **39**(3/4): 395-415.
- Herman, A.W. (1993) Emerging technologies in biological sampling. *UNESCO technical papers in marine science*, No. 66, UNESCO, Paris. 48 pp.
- Hu, M.-K. (1962) Visual pattern recognition by moment invariants. *IRE Trans. Inform. Theory*, **IT-8**:179-187.

Huller, R., Glossner, E., Schaub, S., Weingartner, J. & Kachel, V. (1994). The Macro Flow Planktometer: A new device for volume and fluorescence analysis of macro plankton including triggered video imaging in flow. *Cytometry* 17:109-118.

Hunter, J.R. (1996). Continuous Underway Fish Egg Sampler (CUFES).  
Technical note, Southwest Fishery Science Center, La Jolla, California.

Huntley, M.E. (1992). GLOBEC: Global ocean ecosystems dynamics. *Oceanus*, 35(3): 94-99.

Huntley, M.E., Zhou, M. & Nordhausen, W. (1995). Mesoscale distribution of zooplankton in the California Current in late spring, observed by Optical Plankton Counter. *Journal of Marine Research* 53:647-674.

Jain, A. K. (1986) Cluster analysis. In: *Handbook of pattern recognition and image processing*. Academic Press, New York.

Jeffries, H. P., Berman, M.S., Poularikas, A.D., Katsinis, C., Melas, I., Sherman, K. & Bivins, L. (1984) Automated sizing, counting and identification of zooplankton by pattern recognition. *Marine Biology* 78:329-334.

Kils, U. (1992). The ecoSCOPE and dynIMAGE: Microscale tools for in-situ studies of predator-prey interactions. *Arch. Hydrobiol. Beih. Ergebn. Limnol.* **36**:83-96.

Lenz, J., Mees, S., Schnack, D. & Waller, U. (1991). A high speed video recording system for in-situ studies of small scale distribution of zooplankton and ichthyoplankton: preliminary results on the distribution of plankton in the Bornholm Basin (Central Baltic). *ICES C. M.* L:82.

Lenz, J., Schnack, D., Petersen, D., Kreikemeier, J., Hermann, B., Mees, S. & Wieland, K. (1995). The Ichthyoplankton Recorder: a video recording system for *in situ* studies of small-scale distribution patterns. *ICES Journal of Marine Science* **52**(3/4):409-418.

Levin, S. A., Powell, T. M. & Steele, J. H., editors (1994). *Patch dynamics*. Springer-Verlag, Berlin.

Longhurst, A.R. (1976). Vertical migration. In: D.H. Cushing & J.J. Walsh (Eds.) *The ecology of the seas*. Blackwell Scientific Publications, Oxford.

Longhurst, A.R. (1981). Significance of spatial variability. In: Longhurst, A.R. (Ed.) *Analysis of marine ecosystems*. Academic Press, London.

Longhurst, A.R. & Williams, R. (1976) Improved filtration systems for multiple-serial plankton samplers and their deployment. *Deep-Sea Research*, **23**:1067-1073.

Manly, B. F. J. (1994) *Multivariate statistical methods*. (2nd ed.) Chapman & Hall, London.

Mann, K.H. & Lazier, J.R.N. (1996) *Dynamics of marine ecosystems*. 2nd Edn. Blackwell Scientific Publications, Boston.

Mucciardi, A. N. & Gose, E. E. (1971). A comparison of seven techniques for choosing subsets of pattern recognition properties. *IEEE Trans. Comput.* **C-20**:1023-1031.

Newell, G.E. & Newell, R.C. (1963). *Marine plankton*. Hutchinson Educational, London.

Omori, M (1969). Weight and chemical composition of some important oceanic zooplankton in the North Pacific Ocean. *Marine Biology* **3**(1):4-10.

Omori, M. & Ikeda, T. (1986). *Methods in marine zooplankton ecology*. John Wiley & Sons, New York.

Ortner, P.B., Cummings, S.R., Aftiring, R.P. & Edgerton, H.E. (1979) Silhouette photography of oceanic zooplankton. *Nature* **277**:50-51.

Ortner, P.B., Hill, L.C. & Edgerton, H.E. (1981) *In-situ* silhouette photography of Gulf Stream zooplankton. *Deep-Sea Research* **28A**(12):1569-1576.

Ortner, P.B. (1993). Integration of video imaging with optical plankton counting: "Smart Sampling". US GLOBEC News, **5**:9-11.

Owen, R. W. (1989). Microscale and finescale variations in small plankton in coastal and pelagic environments. *J. Mar. Res.* **47**:197-240.

Pavlidis, T. (1980). Algorithms for shape analysis of contours and waveforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2**(4):301-312.

Persoon, E. & Fu, K-S. (1977). Shape discrimination using Fourier descriptors. *IEEE Transactions on Systems, Man and Cybernetics*, **SMC-7** (3):170-179.

Platt, T. (1985). Structure of marine ecosystems: Its allometric basis. In: *Ecosystem theory for biological oceanography*. Ulanowicz, R.E. and Platt, T. (Eds.). Department of Fisheries and Oceans, Ottawa. p. 55-64.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. (1992). *Numerical recipes in C*. 2nd Ed. Cambridge University Press, Cambridge.

Raymont, J.E.G. (1963). *Plankton and productivity in the oceans*. Pergamon Press, Oxford, 660 pp.

Reiss, T.H. (1991). The revised fundamental theorem of moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**(8):830-834.

Richard, C. W. & Hemami, H. (1974). Identification of three-dimensional objects using Fourier descriptors of the boundary curve. *IEEE Transactions on Systems, Man and Cybernetics*, **SMC-4** (4):371-38.

Robins, D.B., Bale, A.J., Moore, G.F., Rees, N.W., Hooker, S.B., Gallienne, C.P., Westbrook, A.G., Maranon, E., Spooner, W. & Laney, S.R. (1996) AMT-1 cruise report and preliminary results. SeaWIFS Technical Report Series. NASA Goddard Space Flight Center, Greenbelt, Md. USA.

Robins, D.B. & Aiken, J. (1996). The Atlantic Meridional Transect: an oceanographic research programme to investigate physical, chemical, biological and optical variables of the Atlantic Ocean. *Underwater Technology* **21**(4):8-14.

Russ, J. C. (1995) *The image processing handbook* 2nd Edn. CRC Press, Inc., Ann Arbor.

SAHFOS (1995). *Monitoring planktonic systems*. The Sir Alister Hardy Foundation for Ocean Science, Plymouth.

Sameoto, D.D., Jaroszynski, L.O. & Fraser, W.B. (1980) BIONESS, a new design in multiple net zooplankton samplers. *Canadian Journal of Fisheries and Aquatic Sciences*, **37**:722-724.

Sathyendranath, S., Platt, T., Horne, E.P.W., Harison, W.G., Ulloa, O., Outerbridge, R. & Hoepffner, N. (1991). Estimation of new production in the ocean by compound remote sensing. *Nature* **353**:129-133.

Schindler, D.W. (1991). Aquatic ecosystems and global ecology. In: *Fundamentals of aquatic ecology*. Barnes, R.S.K. & Mann, K.H., editors. Blackwell Scientific Publications, London.

Silvert, W. & Platt, T. (1980) Dynamic energy-flow model of the particle size distribution in pelagic ecosystems. In: *Evolution and ecology of zooplankton communities*. Kerfoot, W.C. (Ed.). University Press of New England, Hanover, N.H. p. 754-763.

Smith, S. L., Pieper, R. E., Moore, M. V., Rudstam, L. G., Greene, C. H. Zamon, J. E., Flagg, C. N. & Williamson, C. E. (1992). Acoustic techniques for the in situ observation of zooplankton. *Arch. Hydrobiol. Beih. Ergebn. Limnol.* **36**:23-43.

Sonka, M., Hlavac, V. & Boyle, R. (1994) *Image processing, analysis and machine vision*. Chapman & Hall, London.

Sprules, W.G. (1980). Zoogeographic patterns in the size structure of zooplankton communities with possible applications to lake ecosystem modeling and management. In: *Evolution and ecology of zooplankton communities*. Kerfoot, W.C. (Ed.). University Press, New England, Hanover, N.H., p. 642-656.

Sprules, W.G., Bergstrom, B., Cyr, H., Hargreaves, B.R., Kilham, S.S., MacIsaac, H., Matsushita, K., Stemberger, R.S., & Williams, R. (1992) Non-video optical instruments for studying zooplankton distribution and abundance. *Arch. Hydrobiol./Beih. Ergebn. Limnol.* **36**:45-58.

Stephenson, G. (1973) *Mathematical methods for science students*. 2nd edn. Longman Scientific & Technical, Harlow.

Therrien, G. W. (1989) *Decision, estimation and classification*. John Wiley & Sons, New York.

Welsch, W., Barthel, K.-G., Froese, R., Hermann, B., Lenz, J., Mees, J., Schnack, D. & Waller, U. (1991). A high speed video recording system for in situ studies on small-scale distribution of zooplankton and ichthyoplankton:

preliminary results on the distribution of plankton in the Bornholm Basin  
(Central Baltic). *ICES CM 1991/L:82*

Wiebe, P. H., Morton, A. W., Bradley, A. M., Backus, R. H., Craddock, J. E.,  
Barber, V., Cowles, T. J. & Flierl, G. R. (1985). New developments in the  
MOCNESS, an apparatus for sampling zooplankton and micronekton. *Marine  
Biology* **87**:313-323.

Williams, R. (1983). The double LHPR system: a high speed micro and  
macroplankton sampler. *Deep-Sea Research* **30**(3A):331-342.

Zahn, C.T. & Roskies, R.Z. (1972) Fourier descriptors for plane closed curves.  
*IEEE Transactions on Systems, Man and Cybernetics*, **SMC-7**:269-281.

# APPENDIX 1

1st three principal components for training set.

	Identity	PC1	PC2	PC3
1	sagi	-2.4617	-0.4546	0.1734
2	cope	-0.8994	0.4055	0.4002
3	cope	-1.8959	-0.0615	-0.1963
4	cope	-1.6388	0.0571	0.3453
5	amph	0.2318	2.7011	-0.1796
6	cope	-1.0481	0.2935	-0.5331
7	cope	-0.5963	-0.8369	0.0566
8	cope	-0.0116	-0.8473	0.3716
9	amph	-0.1740	2.9685	0.0458
10	cope	-0.6280	0.2297	-0.7280
11	cope	-1.5215	0.0948	-0.6126
12	cope	-0.4826	0.3456	-1.0466
13	sagi	-2.4241	-0.5176	0.1409
14	sagi	-2.8158	-1.0108	-0.0115
15	cope	-0.9090	0.0561	-0.4985
16	amph	0.6258	2.7749	-0.2572
17	cope	-1.5381	0.8162	-0.2932
18	cope	0.0195	1.4540	1.3868
19	cope	-1.3791	0.7045	-0.2502
20	cope	-1.2407	0.8912	-0.2156
21	cope	-0.2872	0.8768	-0.5533
22	cope	-1.6155	-0.8373	-0.5748
23	cope	1.0072	0.6754	1.1608
24	amph	-0.0583	3.7666	1.2418
25	cope	0.1282	-0.0251	0.5701
26	cope	-1.0840	-0.6277	-0.5113
27	cope	-0.6157	0.2218	-0.8616
28	cope	-0.8504	-0.3440	-0.5165
29	cope	-0.9214	-0.4179	-0.7501
30	cope	0.4163	0.3041	-0.4765
31	cope	-1.1989	0.1146	-0.4340
32	euph	-2.2844	2.1373	0.4141
33	euph	-2.4906	1.7400	0.9238
34	cope	-1.0069	0.7945	0.0037
35	cope	-0.8094	0.2986	-0.1004
36	cope	-0.8070	0.2276	-0.0571
37	cope	-0.7421	0.4232	0.4230
38	cope	-0.5668	-0.7509	0.0028
39	cope	-0.4228	0.0979	-0.5614
40	cope	-1.2948	0.1932	-0.7518

41	sagi	-2.7291	-0.9649	0.0252
42	cope	-1.1693	1.0554	-0.4569
43	amph	0.2429	2.9285	0.8690
44	cope	-1.4196	0.2662	-0.1712
45	cope	-0.3366	0.4264	-0.4040
46	cope	-1.6944	0.3480	0.0505
47	cope	1.0717	0.7716	-0.4177
48	euph	-2.5268	2.1992	1.5863
49	euph	-2.5588	1.9410	1.1079
50	circ	4.4747	-1.5608	0.5449
51	circ	2.9800	-2.2441	-0.0601
52	cope	-0.9340	-2.1076	0.0439
53	circ	3.5706	-2.4460	-1.0291
54	circ	4.7201	-1.5952	0.4889
55	circ	5.2509	-1.6721	0.3699
56	cope	-0.3738	-1.5031	-0.1921
57	circ	3.1902	-1.7912	-0.0192
58	circ	3.8497	-1.8854	-0.1683
59	cope	0.2778	-0.4775	0.8999
60	cope	-0.4038	0.0602	0.9875
61	cope	1.6748	-1.3129	0.5140
62	cope	0.3891	-1.5580	0.1514
63	cope	-0.4141	-1.2153	0.0367
64	circ	4.6904	-1.0573	0.4450
65	circ	3.5925	-2.1256	-0.0994
66	cope	0.5812	-0.8900	-0.1332
67	circ	4.4391	0.0595	0.2007
68	cope	-0.6178	-1.0719	0.1317
69	circ	3.5013	-2.0491	-1.0749
70	amph	-0.3945	4.2174	1.1907
71	cope	-0.8153	1.2588	-1.1380
72	cope	0.3568	-0.1966	-0.1523
73	cope	-0.0510	-0.3843	-0.2178
74	cope	1.6946	0.0354	-0.0314
75	cope	1.6197	-1.0737	0.2272
76	cope	-0.3210	-0.9409	0.7940
77	cope	-0.3187	0.7084	-1.1443
78	cope	-0.6367	-1.5313	-0.3515
79	cope	-0.2175	0.4486	-0.3718
80	cope	-0.7804	-1.2984	-0.5825
81	cope	-0.7102	0.1952	-0.0766
82	cope	-0.5180	0.7341	0.1198
83	cope	-0.3397	0.6771	-0.2120
84	cope	-0.4608	0.4391	-1.0010
85	cope	-0.8470	0.4906	-0.5488
86	cope	-0.8131	0.1062	-0.2996
87	sagi	-3.0507	-0.9295	0.3328

88	sagi	-3.0478	-1.0393	0.3822
89	cope	2.1379	-0.2789	1.0259
90	sagi	-2.8833	-0.9902	0.2994
91	cope	0.2826	0.6032	0.4327
92	cope	1.2052	-1.3316	0.2734
93	cope	0.9534	-1.5213	0.2016
94	sagi	-2.3072	-0.5659	0.2506
95	cope	-0.6614	-1.0207	0.2989
96	cope	0.0261	0.4431	1.7082
97	sagi	-3.0726	-1.8777	0.2563
98	cope	-0.6091	0.4330	-0.7319
99	cope	1.3744	-1.0393	0.2833
100	cope	0.8346	-0.0961	0.2383
101	cope	0.4398	0.1396	-0.2919
102	amph	2.0476	4.9903	-2.0145
103	cope	-1.4406	0.6034	-0.1858
104	euph	-2.0281	1.9543	0.4437
105	cope	-0.5962	-0.5933	0.1834
106	cope	-0.6383	-0.5299	0.2501
107	cope	-0.6773	-0.7221	-0.1059
108	cope	-0.3201	-0.5220	0.1027
109	fegg	1.7319	1.3485	-1.7264
110	amph	1.1558	3.6034	2.0810
111	cope	-1.9039	0.3425	0.9937
112	cope	-0.5869	-0.7350	-0.1264
113	cope	-1.0485	0.5754	-0.0503
114	cope	0.3816	-0.2252	-0.4176
115	fegg	2.8844	1.3477	-0.2058
116	cope	-0.2935	-0.5699	-0.1274
117	cope	-1.0938	0.6464	-0.2805
118	cope	-0.3187	-0.5148	-0.2284
119	fegg	2.5187	2.2031	-0.4127
120	cope	-0.8882	-0.3832	-0.7315
121	cope	-0.1490	0.1110	-0.2045
122	fegg	2.3991	2.3299	-0.4361
123	cope	0.0263	-1.6752	0.1954
124	cope	-0.2432	-0.0685	-0.3832
125	fegg	2.4487	1.6782	0.0822
126	cope	-0.2232	-1.5507	0.1729
127	cope	-0.4750	-0.1325	-0.4466
128	fegg	2.1587	1.9713	-0.5499
129	cope	-0.8211	-0.3559	-0.2251
130	fegg	3.4875	0.8620	1.1351
131	cope	0.3890	0.1767	0.5717
132	cope	-0.2028	-0.6140	-0.0122
133	cope	-0.6014	-1.1699	0.0766
134	cope	0.4811	-0.8375	-0.0504

135	cope	0.7064	-0.3736	-0.2126
136	cope	-1.0123	-0.9730	-0.0471
137	cope	-0.5870	-0.7226	0.2338
138	cope	-0.6556	-0.8674	0.0435
139	cope	-1.0595	-0.6316	-0.0561
140	cope	0.1495	-0.7593	-0.1438
141	cope	-1.3207	-0.5504	-0.2991
142	sagi	-2.3987	-0.4705	0.4895
143	cope	-0.6288	1.7612	-1.5275
144	cope	-0.7782	2.0732	-2.4608
145	cope	0.2980	-1.9888	0.1428
146	cope	0.3827	-1.7821	0.0235
147	cope	-0.2624	-0.1184	0.2223
148	cope	-0.6458	-1.9033	0.0710
149	cope	0.6745	0.3499	0.3080
150	cope	0.2153	-1.6112	0.0719
151	cope	0.0730	-1.7390	0.0945
152	cope	0.0763	-0.4089	0.0833
153	cope	-0.7504	-0.6229	-0.1352
154	cope	-0.0607	-0.3447	-0.1555
155	cope	0.5227	1.3584	1.6371
156	cope	-0.0777	-0.5801	-0.1146
157	cope	0.6404	1.5025	1.5195
158	cope	2.0863	-1.5379	0.3805
159	cope	-0.3992	-1.9180	0.0441
160	cope	-1.1406	-0.9993	-0.2416
161	cope	1.5332	-1.7437	0.2883
162	cope	0.1502	1.6487	1.1586
163	cope	-1.2323	-1.0060	-0.2502
164	cope	1.5332	-1.7437	0.2883
165	cope	-1.0317	-0.8986	-0.1532
166	cope	0.9331	-1.7366	0.3058
167	cope	-0.6105	-0.5581	0.0492
168	cope	0.8884	-0.6109	0.0089
169	cope	-0.5247	0.0528	0.3273
170	cope	-1.2127	0.4832	-0.5541
171	cope	-0.3968	0.7172	-0.3025
172	cope	-0.7319	-0.5218	-0.3865
173	cope	-0.9890	-0.7577	0.2752
174	cope	0.8126	1.0611	1.6764
175	cope	-0.6425	0.6489	-0.1319
176	cope	-0.3770	-2.1070	-0.0310
177	cope	0.1873	-2.0703	0.0192
178	cope	0.4363	0.7895	-0.8203
179	cope	-0.9609	-0.3115	0.3298
180	cope	1.5792	-1.3834	0.1874
181	fegg	3.2016	0.7153	1.0325

182	sagi	-2.2226	-1.5862	0.0923
183	cope	-0.3418	0.4780	-0.0157
184	fegg	3.1144	2.6725	-1.2273
185	cope	-0.8917	-0.2179	0.1660
186	cope	-0.3181	0.7830	0.1938
187	fegg	3.5657	2.0812	-0.3920
188	cope	-0.0771	0.2358	-0.0775
189	cope	-0.5841	0.1509	0.1954
190	cope	-0.4099	0.5326	0.0584
191	fegg	2.1852	1.3285	0.6310
192	cope	0.5333	0.0045	-0.6232
193	fegg	2.3444	2.2990	-0.4833
194	cope	-0.2913	0.5942	-0.2776
195	fegg	2.2457	2.4922	-0.6198
196	cope	-0.4923	1.0124	0.2350
197	fegg	2.3909	2.5102	-0.5931

## APPENDIX 2

### Sample images acquired and pre-processed by the ViZA system

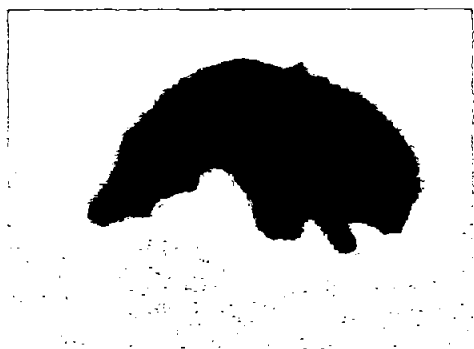


Fig A1: Image of Amphipod Shrimp, segmented image, and extracted contour.

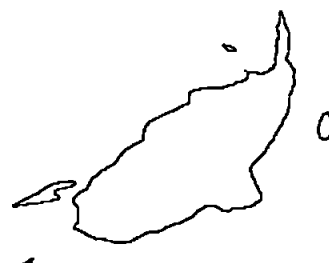


Fig A2: Image of copepod, segmented image, and extracted contour.

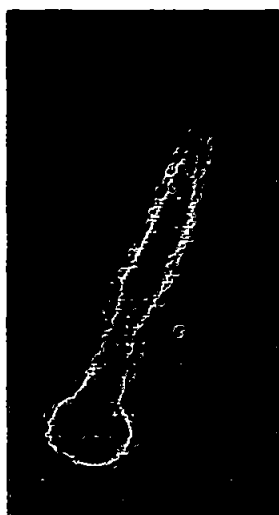


Fig A3: Image of fish larvae, segmented image, and extracted contour.