VARIABILITY IN ULTRAPLANKTON AT THE PORCUPINE ABYSSAL PLAIN STUDY SITE (49º50'N 16º30'W)

Adrian P. Martin[a,*], Mikhail V. Zubkov[a], Ross J. Holland[a], Glen Tarran[b], Peter Burkill[c]

[a] National Oceanography Centre, Southampton, SO14 3ZH, UK

[b] Plymouth Marine Laboratory Prospect Place West Hoe Plymouth PL1 3DH

[c] Sir Alister Hardy Foundation for Ocean Science, The Laboratory, Citadel Hill, Plymouth, Devon, PL1 2PB

* Corresponding author. Tel.: 00 44 (0)23 80596342; fax: 00 44 (0)23 80596247; email: apm1@noc.soton.ac.uk

Abstract

Observations of the ultraplankton (<5μm) are presented from a 4 day mesoscale survey centred on the Porcupine Abyssal Plain (PAP) study site (49º00'N 16º30'W), in July 2006. The organisms enumerated include two groups of phytoplankton, *Synechococcus* cyanobacteria, heterotrophic bacteria, large viruses and two size classes of heterotrophic protist. The dataset comprises over 400 samples from the mixed layer taken over a 100km x 100km area at a spatial resolution of typically 2-3km.

For phytoplankton and heterotrophic bacteria there is a clear bimodal structure to the histograms of abundance indicative of two distinct communities in the region. Using the strong bimodality of one of the phytoplankton groups' histogram as a basis, the dataset is split into two subsets, with roughly 200 points in each, corresponding to the two histogram peaks. Doing so provides evidence that *Synechococcus* and viruses may also have a bimodal structure. Correlations between all pairings of these 5 organisms (both phytoplankton groups, *Synechococcus*, heterotrophic bacteria and viruses) are positive and quite high ($r>0.7$). The two communities can therefore be characterised as high and low abundance. Although there is a coincidence of low abundances with high temperatures in the southwest corner of the region, where there was known to be an eddy present, the spatial distributions of these organisms over the whole region is poorly predicted by temperature (or salinity or density). Furthermore, the spatial distributions of heterotrophic protists are found to differ strongly from those of the other organisms, having a unimodal structure and no obvious large scale structure. The more random structure of the heterotrophs' spatial distribution compared to their prey is consistent with previous results from the continental shelf, but is demonstrated for the open ocean here for the first time.

Spatial variability is a large potential source of error in point samples, such as those comprising time series or transect cruises, unless a sufficient number of samples are taken. This large dataset is further used to provide guidance on the number of samples that would be required to estimate the mean abundance for the organisms accurately in this spatially variable region. Even if the bimodal structure was known initially, many of the organisms would require 10 or more samples to estimate the mean with 25% accuracy.

1. Introduction

The ultraplankton (plankton <5μm in size) are the most abundant organisms in the sea and play a fundamental role in marine biogeochemical cycles. Little is currently known of their spatial variability, however, particularly at the mesoscale (1-100km). In shelf seas there is evidence of very strong mesoscale spatial variability in ultraplankton (Martin et al. 2005). Furthermore, the spatial distributions of heterotrophs is markedly different from that of their prey there (Martin et al. 2008). Although there are some mesoscale datasets from the open ocean (e.g. Zubkov & Quartly, 2003) none to date have been adequate to ascertain whether the shelf sea phenomena are more general.

In June and July of 2006, RRS Discovery conducted a month long multidisciplinary cruise centred on the Porcupine Abyssal Plain study site (49º50'N 16º30'W). Towards the end of the cruise an intensive 4 day mesoscale survey of the area was conducted. During this survey ultraplankton samples were taken from the mixed layer, with typically 2-3km between samples. This provides a high resolution map of ultraplankton in the area, at such a level of taxonomic detail that two size classes of heterotrophic protist were separately counted as well as a range of their prey including two types of phytoplankton, *Synechococcus* cyanobacteria, heterotrophic bacteria and large (>0.2μm) viruses. All the numerically dominant representatives of the ultraplankton predator-prey system were mapped.

This paper presents the results of that survey. It also uses this large dataset to offer practical guidance on the number of samples that need to be taken to avoid spatial variability leading to large errors in estimates of the mean abundance of organisms for this area. The high-resolution survey described here is one of very few yet conducted on ultraplankton (Jacquet et al., 2002; Zubkov & Quartly, 2003). Most

surveys must satisfy themselves with a much smaller number of data points. It is therefore germane to ask how many points it would be necessary to sample to obtain accurate estimates of the mean abundance of the various ultraplankton groups in a spatially variable region.

2. Methods

2.1 Flow cytometry

The mesoscale survey took 4 days, with each day devoted to surveying one of the quadrants, shown in Figure 1. Samples were taken for analysis from the non-toxic surface water supply every 12 mins giving a spatial resolution of typically 2-3 km. The intake for the non-toxic supply is at a depth of 6 m and hence well within the mixed layer (typically 20-30m throughout the cruise). Samples (2 x 2.4 ml) were automatically collected into polypropylene flow cytometry tubes (Fisher, Loughborough, UK) and fixed with 1% paraformaldehyde (PFA) using a Miniprep-60 autosampler (Tecan, Reading, UK). Absolute concentrations of microbes were determined by flow cytometry (FACSort, Becton Dickinson Biosciences, Oxford, UK). Microbes were stained with SYBR Green I DNA stain (Sigma-Aldrich, Poole, UK), 1:5000 final dilution of initial stock, in the presence of potassium citrate, 30 mM final concentration, in the dark at +4^{o}C for 2 – 4 h (Marie et al. 1997; Zubkov et al., 2007). A yellow-green 0.5 μm bead (Fluoresbrite Microparticles, Polysciences, Warrington, USA) concentration standard was added at known dilution to determine absolute cell concentrations (Zubkov & Burkill, 2006).

96 The two phytoplankton groupings were characterised by size and fluorescence. For
97 simplicity they will be referred to as large orange fluorescence phytoplankton (LOP)
98 and picophytoplankton (PPH). Regarding the former, dinoflagellates and *Mesodinium*
99 are far too rare to be counted by flow cytometry. Therefore, they are most likely to be
100 small (~ 3-8μm) Cryptophytes. Out of caution we simply refer to them as LOP. In
101 addition to phototrophs and heterotrophic bacteria, heterotrophic protists and viruses
102 were also enumerated (Zubkov et al. 2004; Zubkov et al. 2007). The two groupings of
103 heterotrophic protists will be referred to simply as small (SHE) and large (LHE)
104 heterotrophic protists. As the survey enumerated all numerically significant organisms
105 in this size range it therefore represents a snapshot of the predator-prey dynamics of
106 the ultraplankton in this location. The effect of the relatively low abundances of
107 heterotrophic protists on the accuracy of counts has been examined previously (Martin
108 et al., 2008) and is not significant enough to affect our results.
109
110 *2.2 Statistics*
111 The Central Limit Theorem provides a straightforward way to estimate the number of
112 samples that must be taken to obtain a particular level of accuracy for an estimate of
113 the mean. The one assumption we will make here is that the mean and variance
114 estimated from our full dataset are the true mean, μ, and variance, σ^2, respectively.
115 Clearly this is an approximation. However, as our dataset comprises greater than 200
116 points even if it is split into the two halves described later, the difference is likely to
117 be small. If it is not, then tackling spatial variability is clearly a very major problem.
118 Given this assumption the estimate of the mean, m, using N samples X_i where i=1:N,
119 has a normal distribution with mean μ and variance σ^2/N. Therefore, for example, if
120 we want to estimate how many points we need to sample to obtain an estimate of the

mean within P% of the true mean with 99% probability we simply need to estimate the number of points to constrain 99% of the distribution for m within (1-P/100)*m and (1+P/100)*m. Remember that the distribution of the mean narrows as we increase N since the variance is σ^2/N.

3. Results

3.1 Large-scale structure

Figure 2 shows the spatial distribution for 6 of the 7 classes of ultraplankton enumerated, with each coloured spot corresponding to a sample. For phytoplankton, *Synechococcus*, heterotrophic bacteria and viruses a quick visual analysis might suggest a strong relationship with temperature and/or physical structures. The eddy visible as a warm anomaly in the southwest corner (Painter et al., 2007) is systematically associated with low abundances for all these organisms. However, the correlations between these organisms and temperature are generally relatively weak (Figure 3), with the most significant correlation being between viruses and temperature with a correlation coefficient of -0.66. Similarly weak correlations are found with salinity and density (not shown). The cause for this weak dependence over the region as a whole is apparent at the fringes of the survey region. Particularly along the eastern and northern edges of the survey, organism abundances are as low as within the eddy.

The spatial distribution of large heterotrophic protists (strongly mimicked by that of small heterotrophic protists which is, therefore, not shown) has little large-scale structure. Though there are perhaps a few more high abundances within the

146 eddy, similarly high values can be found in all quadrants of the survey. The
147 correlation coefficient for large heterotrophic protists and temperature is just 0.20
148 with that for small heterotrophic protists and temperature being 0.38.
149 Regardless of the strength of the relationship between temperature and the
150 various organisms's abundances, it is clear in Figure 2 that, at least at large scales, the
151 distributions of phytoplankton, *Synechococcus*, heterotrophic bacteria and viruses
152 have very similar structure. It is also apparent that this structure is markedly different
153 to those of the heterotrophic protists. Further evidence for the former point can be
154 found in the correlation coefficients, shown in Figure 3. The correlation coefficient is
155 greater than 0.7 for all pairings of phytoplankton groups, *Synechococcus*,
156 heterotrophic bacteria and viruses. The largest correlation coefficient for any one of
157 these 5 organisms with either large or small heterotrophic protist groups is -0.25
158 between *Synechococcus* and large hetrotrophic protists. Note that large and small
159 heterotrophic protists are even uncorrelated with each other (correlation coefficient
160 equals zero to 2 decimal places).
161 Despite the fact that temperature does not appear to offer an easy explanation
162 for the organisms' distributions (something reinforced below), there is clear evidence
163 for a divided community in the area. Figure 4 shows histograms of organisms'
164 abundance. Phytoplankton and heterotrophic bacteria clearly have bimodal
165 distributions, strongly suggestive of two different communities in the region. Figure 4
166 also shows how the abundances are distributed when the samples are split according
167 to a criterion based on the histogram of the LOP phytoplankton group (see Figure 4
168 caption). As LOP shows one of the clearest bimodal distributions, with its two peaks
169 most clearly separated, this organism was used to split the data into two roughly equal
170 parts: S1 comprising samples where LOP abundance<8102 cells/ml (208 points) and

171 S2 comprising samples where LOP abundance ≥8102 cells/ml (211 points). Clearly
172 this is subjective but there is no objective way to split a bimodal distribution,
173 particularly in situations such as here where the component 'peaks' are both strongly
174 non-Normal. (Similar results were obtained using the PPH phytoplankton group as a
175 basis for the split instead.) Splitting in this way suggests a common bimodality, not
176 just for phytoplankton and heterotrophic bacteria, but also for *Synechococcus* and
177 viruses. The same split, however, results in no clear demarcation in the distribution of
178 temperature, once more indicating that it is not the causal factor (similar results hold
179 for salinity and density though they are not shown). Consistent with the spatial
180 distributions seen in Figure 2, the two halves of both large and small heterotrophic
181 protists under the split have near identical histograms (Figure 4). This is consistent
182 with their spatial distributions, discussed above. In particular there is no evidence for
183 a bimodal distribution in either large or small heterotrophic protists.
184 Figure 3 shows scatter plots for all pairings of organism abundances and
185 temperature, with corresponding correlation coefficients at the top of the plot for each
186 pairing. These plots generally confirm what has already been stated about the
187 organisms' spatial variability. The scatter plot of large versus small heterotrophic
188 protists clearly demonstrates that there is no evidence for two distinct populations for
189 either of these organisms. Calculating correlation coefficients for samples in each of
190 the two halves of the split dataset does not increase the magnitude of any of the
191 coefficients. Looking at some of the relationships exhibited in Figure 3 (e.g. PPH-
192 SYN or LOP-VIR) one might be tempted to infer evidence of a trajectory associated
193 with the population dynamics of the region (Srokosz et al., 2003). However, it should
194 be borne in mind that these 'trajectories' link the high abundance cluster in phase
195 space to that of the low abundances. It may therefore simply be due to mixing of the

196 waters representing the two peaks in the bimodal distribution. We return to this in the
197 Discussion.

198 Given the bimodal nature of the distributions for several of the organisms and
199 the reasonable strength of correlation between them it is worth investigating their
200 spatial distributions in a little more detail. Figure 5 shows the location of samples with
201 abundances corresponding to the peaks in the histograms for the two phytoplankton
202 groups, LOP and PPH, shown in Figure 4. For points comprising the lower abundance
203 peak, LOP and PPH have a broadly similar distribution with some focussed in the
204 eddy in the southwest corner and others scattered around the northern and eastern
205 boundaries of the survey region. However, the points comprising the peak in high
206 abundance can be found primarily in the southeast corner for LOP but primarily in the
207 northern half for PPH. Hence, though LOP and PPH both display a bimodal
208 distribution, with each 'half' comprising roughly the same sample points, and though
209 they have a correlation coefficient of 0.81, there is no strong coincidence in the
210 geographical locations of the samples comprising the two peaks in their abundance
211 histograms. If one instead focuses on the location of minimum and maximum values
212 (once more Figure 5), although there is some mismatch with the locations of samples
213 comprising the peaks in the histograms for each organism, there is still no agreement
214 between LOP and PPH.

215 The same results were obtained if the bins involved in the histogram in Figure
216 4 were shifted along by half a bin width.

217

218 *3.2 Variability*

219 Having described the spatial variability in the organisms' distributions we now turn to
220 how we may accurately estimate the mean abundance of organisms in the area. Our

221 approach to quantifying this has already been explained in the Methods section (2.2).
222 It should be noted that the following analysis is intended as illustrative. Although
223 there is no reason to suppose that other regions have less spatial variability than the
224 site studied, the results we find can only be used as guidelines for other sites. In
225 particular we are in no way claiming that a bimodal distribution is typical, as this is
226 likely to arise in only certain situations (see Discussion). However, the dataset, in
227 revealing two scales of spatial variability – the large scale manifested as the peaks
228 bimodal distribution and the smaller scale as the distribution about each of the two
229 peaks – provides an excellent illustration of the dangers of using single or few points
230 to estimate mean abundances for an area. Figure 6 shows the predicted error in the
231 mean for a range of sample sizes.
232 Although the temperature histogram shows no evidence of a bimodal
233 distribution, the predicted errors for the two subsets of data, S1 and S2, are shown
234 here for comparison – likewise for small and large heterotrophic protists. The
235 maximum error for temperature is predicted to be largest for S1. This is to be
236 expected given the longer tail of the distribution for S1 (Figure 4). The error may
237 seem small at 0.026, but it should be remembered that this is fractional error. Since
238 the mean temperature for S1 is 15.69 ^{o}C, this corresponds to an error of 0.41^{o}C. The
239 maximum error for S2 is 0.17 ^{o}C. To obtain an equally small error for S1 would
240 require at least 10 samples to be taken.
241 The error curves for S1 and S2 are very similar for both groups of
242 heterotrophic protist. This is consistent with the unimodal distributions for these
243 organisms seen in Figure 4. For large heterotrophic protists the largest potential error
244 is 55% of the true mean, with that for small heterotrophic protists potentially larger

than 80%. To bring the likely error for either down to the 25% level would require at least 10 samples to be taken.

All of the remaining organisms (phytoplankton, *Synechococcus*, heterotrophic bacteria and viruses) have significantly different errors when the mean is estimated for S1 and S2. In all cases, the wider histogram peak for S1 leads to a larger error than for S2 using the same sample size. The lowest error is associated with heterotrophic bacteria for which 10 samples would push the error down to 25% for S1 and almost to 15% for S2. The most extreme errors are associated with *Synechococcus* and viruses where for S1 errors are still at least as large as 50% even with 10 samples. Although the presence of a bimodal distribution may be debated for these organisms it should be noted that regardless of the manner in which the split was carried out the error in estimated mean associated with either S1 or S2 is still guaranteed to be less than that obtained using the undivided dataset. For the most clearly bimodally distributed organisms, corresponding to the two phytoplankton groups, errors for estimating the mean of S1 are lower than 20% when 10 samples are used. The same number of samples would give errors of order 40% however for S2.

It is worth stressing that bimodality is not the only source of significant potential error when estimating mean abundances. Even knowing which 'sub-population' a sample belongs to, so that it can be treated independently as a unimodal distribution, it is still necessary to have of order 10 samples for an accurate estimate of the mean.

Note that it is more difficult to form estimates of the likely error in estimating the variance for the organisms' abundances. It is tractable for the mean because the distribution of mean estimates is Normal. For variance estimates, however, the distribution is more complicated (Pearson Type III) and no simple guidance can be

4. Discussion

The data presented here from the D306 cruise cover just 4 days of one year. It is pertinent to ask how representative these results are of other years and times. Unfortunately, flow cytometers are still far from being a standard instrument on ships, so few datasets which enumerate the ultraplankton exist for the PAP area. The best source of data is the UK NERC-funded Atlantic Meridional Transect (AMT) programme. In particular data are available for AMT13 (September 2003), AMT 16 (June 2005) and AMT 17 (October 2005). There are also data available from the German Poseidon 300/1 cruise in July 2003. Because these cruises had other objectives and did not sample the PAP area anywhere near as intensively for ultraplankton as the D306 cruise, relatively few data are available, even when pooled from all AMT and Poseidon datasets. Specifically, only *Synechococcus* was counted on all these previous cruises so we must restrict attention to this organism. Using only data from the top 20m, to be consistent with the mixed layer data presented here from D306, we have just 23 previous observations from within our mesoscale survey area or 49 if we take data from anywhere within the Porcupine Abyssal Plain (approximately 43-50N 10-20W). If we further insist that the data must have been collected in the summer (June or July) then these two figures fall to 18 and 30 observations respectively. Hence, we cannot carry out any meaningful statistical analysis in the manner described above. However, we can at least investigate how

these earlier datasets compare to our observations. Figure 7 once more shows the histogram of abundance for *Synechococcus* but is now superimposed with data from previous cruises. Previous data from the same area as that mapped by the D306 mesoscale survey show a clear clustering with the peak in high abundances. This is true whether data are restricted to the summer months to be consistent with D306 or not. If data are allowed from the broader Porcupine Abyssal Plain area, but restricted to the summer, this clustering remains. Data from September and October, however, do contain a number of points of lower abundances. Taking the summer data as providing the best comparison to D306 data it is tempting to draw the conclusion that the peak in low abundances seen in the histograms represents a community not normally present in this area at that time. There is further evidence for this. Painter et al. (this volume) find evidence for the presence of two distinct water masses in the region over the course of the cruise. At the time of the mesoscale survey the eddy in the southwest corner represents a more northerly form of Eastern North Atlantic Central Water (ENACWp) while the rest of the region is dominated by a southerly variant (ENACWt) originating from near the Azores. Smythe-Wright et al. (this volume) find that photosynthetic pigments that can be used as markers for different types of phytoplankton also reflect this juxtaposition of significantly differing water masses. Finally, Hartman et al. (this volume) provide evidence that the dominant water mass in this area may vary from year to year, indicating that the region is one where the stirring together of markedly different water masses may occur. If the different water masses have different biogeochemical signatures and different community structures then the bimodality discussed here may be frequent. Caution is required, however. First, there are still relatively few observations. Second, it could be argued that the whole of the D306 dataset is indicative of a different community to

previous years. No significant numbers of *Prochlorococcus* cyanobacteria were detected on D306. Yet on all of the previous cruises mentioned above, *Prochlorococcus* outnumbered *Synechococcus*. There is currently no understanding of why *Prochlorococcus* should be absent during D306. Such a dramatic shift does highlight the need for care in comparing datasets though.

From a broader perspective, our data can be compared to recent results from the shelf seas (Martin et al, 2008). The analysis presented here provides the first open-ocean evidence of ultraplankton predators being spatially distributed in a significantly different, essentially more random, manner than their prey. This was exactly the phenomenon reported by Martin et al. (2008) from the Celtic Sea between UK and Ireland. Although the bimodal distribution of organisms is very clear for some of the organisms in Figure 4, this clarity of division is not carried over to their geographical distributions, and the split between the two communities carries a degree of subjectivity. Hence, we cannot apply any of the spatial statistical techniques applied by Martin et al. (2008). Nevertheless our work supports their findings: that heterotrophic protists can display a significantly different spatial distribution to their prey. Martin et al. (2008) put forward a hypothesis that top-down pressure, via strong predation of the heterotrophic protists, could explain this apparent lack of relationship even in the presence of a strong trophic link between the protists and their prey. Unfortunately, as in that study, we do not have data on higher predators' abundances to test this hypothesis. The questions of why viruses were found here to share a similar distribution to their hosts (most likely LOP) and why large and small heterotrophic protists had a radically different distribution to their prey must be addressed by future work.

As part of our analysis we make use of standard linear correlations to examine relationships between different organisms and temperature. A few points are worth making on this. First, it should be noted that we make no attempt to take into account possible dependency of neighbouring points – such a dependency can lead to incorrectly high correlation coefficients. However, to deal with this properly it is necessary to calculate spatial and temporal auto-correlation scales – only data points separated by distance and time greater than these two scales respectively should be used for the calculation of the correlation. We choose not to do this for two reasons. First, as is clear from the above discussion there is evidence of strongly differing communities with no clear geographical segregation between them, though with some evidence of a cause in the juxtaposition and mixing of two different water masses. This makes the spatial variability in the region strongly non-isotropic. The autocorrelation scales will therefore vary with position, direction and time, making such an approach unfeasible. Second, using correlation coefficients to infer an absence of relationship between heterotrophic protists and their prey is proof against this issue precisely because correlation analysis will, if anything, make the relationship look stronger than it really is, so we can have some confidence in a poor correlation.

Taking the dataset presented here in isolation once more, our analysis raises concerns for how we might approach ecological modelling of the area. It is relatively straightforward to put together an ecological model – there are a host of previous formulations that can be used off-the-shelf or combined to give a model of the required structure provided nothing novel is intended. The greater difficulty lies in finding sufficient accurate data to constrain the model. This is increasingly important as the complexity of the model increases, since the number of degrees of freedom to

constrain increases. Time-series have long been a stalwart source of data for constraining models. The provision of data throughout a year allows the model to be constrained as the ecosystem responds to seasonal changes in forcing. Historically, the errors reported on data from time-series have been restricted to those arising from methodology, they have not taken into account spatial variability. This survey has demonstrated that to obtain an accurate estimate of the mean concentration of a given organism or group of organisms, it may be necessary to obtain 10 or more samples at a range of locations. Otherwise, the 'error' arising from insufficient sampling of spatial variability may be much larger than that arising from methodological effects. This has obvious consequences for models constructed using time-series data.

5. Conclusions

We have presented high spatial resolution data from a mesoscale survey at the PAP study site enumerating a range of ultraplankton. Histograms of abundance reveal a divided community for phytoplankton, prokaryotes, heterotrophic bacteria and viruses, with two clear groupings. Heterotrophic protists are evenly, if more randomly, distributed over the whole region. From a practical point of view, we have further demonstrated that to be able to estimate the mean abundance of some of the organisms present would require in excess of 10 samples to be taken at different locations. Without due regard for the spatial variability, it can be a source of greater variability in observations than observational methodology. There is currently no reason to think that the PAP site has particularly strong spatial variability relative to any other open ocean site.

Acknowledgements

We would like to thank Robin Hankin for statistical advice and the captain and crew of RRS Discovery for providing such a superb platform for oceanographic research. This manuscript contributes to the OBE NSRD Oceans 2025 core and AMT programmes funded by the Natural Environmental Research Council, UK. Adrian Martin was funded by a NERC Advanced Research Fellowship (NER/J/S/2001/00708).

References

Burkill, P.H., 2006. RRS Discovery Cruise 306, 23 Jun - 6 Jul 2006. Pelagic biogeochemistry of the PAP site. Cruise Report No. 9. National Oceanography Centre, Southampton, UK.

Hartman, S.E., Larkin, K.E., Lampitt, R.S., Lankhorst, M. and Hydes, D.J. Seasonal and inter-annual biogeochemical variations at PAP (49^{o}N 16.5^{o}W) 2003-2005 associated with winter mixing and surface circulation. Deep-Sea Research II, this volume.

Jacquet, S., Prieur, L., Avois-Jacquet, C., Lennon, J.F., and Vaulot, D., 2002. Short-timescale variability of picophytoplankton abundance and cellular parameters in surface waters of the Alboran Sea (western Mediterranean). Journal of Plankton Research, 24, 635-651.

417 Marie, D., Partensky, F., Jacquet, S. and Vaulot, D., 1997. Enumeration and cell cycle
418 analysis of natural populations of marine picoplankton by flow cytometry using the
419 nucleic acid strain SYBR Green I. Appl. Environ. Microbiol., 63, 186-193
420
421 Martin, A.P., Zubkov, M.V., Burkill, P.H. and Holland, R.J., 2005. Extreme spatial
422 variability in marine picoplankton and its consequences for interpreting Eulerian time-
423 series. Biology Letters, 1, 366-369.
424
425 Martin, A.P., Zubkov, M.V., Burkill, P.H. and Holland, R.J., 2008.Microbial spatial
426 variability: an example from the Celtic Sea. Progress in Oceanography, in press.
427
428 Painter, S.C., Pidcock, R., Allen, J.T., 2007. Mesoscale eddies driving spatial and
429 temporal heterogeneity in the productivity of the euphotic zone of the NE Atlantic.
430 Deep-Sea Research II, this volume.
431
432 Smythe-Wright, D., Boswell, S., Kim, Y.-N., Kemp, A.E. Spatio-temporal changes in
433 the distribution of phytopigments and phytoplanktonic groups at the PAP site. Deep-
434 Sea Research II, this volume.
435
436 Srokosz, M.A., Martin, A.P. and Fasham, M.J.R., 2003. On the role of biological
437 dynamics in plankton patchiness at the mesoscale: an example from the eastern North
438 Atlantic Ocean. *Journal of Marine Research*, 61, (4), 517-537.
439
440 Zubkov, M.V., Allen, J.I., and Fuchs, B.M., 2004. Coexistence of dominant groups in
441 marine bacterioplankton community - a combination of experimental and modelling

approaches. Journal of the Marine Biological Association of the United Kingdom, 84, 519-529.

Zubkov, M.V. and Burkill, P.H., 2006. Syringe pumped high speed flow cytometry of oceanic phytoplankton. Cytometry Part A, 69A, 1010-1019.

Zubkov, M.V., Burkill, P.H., and Topping, J.N., 2007. Flow cytometric enumeration of DNA-stained oceanic planktonic protists. Journal of Plankton Research, 29, 79-86

Zubkov, M.V. and Quartly, G.D., 2003. Ultraplankton distribution in surface waters of the Mozambique Channel – flow cytometry and satellite imagery. Aquatic Microbial Ecology, 33, 155-161.

FIGURE CAPTIONS

Figure 1.

Map showing location of the mesoscale survey. The ship track during the survey is shown as a solid line. The four quadrants of the survey are indicated by the dotted line. Each of the quadrants took one day to complete and were completed in the following order: northeast, southeast, northwest, southwest. In each case the ship carried out the 'radiator' path before returning to the centre station using the diagonal 'dog-leg'.

Figure 2.

Spatial distribution of (a) PPH phytoplankton group, (b) large heterotrophic protists, (c) *Synechococcus*, (d) LOP phytoplankton group, (e) heterotrophic bacteria and (f) viruses. Each dot corresponds to a sample and the colour of each dot corresponds to log(abundance +1) where abundance is is cells/ml and log denotes natural logarithm. The organism abundances are superimposed on contours of temperature (degrees C) with specific temperatures labelled in a..

Figure 3.

Scatter plots for all pairings of log(abundance+1), with abundance in units of cells/ml, for all ultraplankton groups enumerated and temperature (T). Black and grey points correspond to the split defined in Figure 4. The X-Y labels denote the axes on which each variable is plotted. TEM corresponds to temperature, BAC stands for heterotrophic bacteria and VIR represents viruses. All other abbreviations are as in the

text. Figures at the top of each subplot denote correlation coefficients for the full dataset.

Figure 4.

Histograms of (a) temperature (degrees C) and log(abundance + 1) where abundance is in cells/ml for (b) PPH phytoplankton group, (c) large heterotrophic protists, (d) small heterotrophic protists, (e) *Synechococcus*, (f) LOP phytoplankton group, (g) heterotrophic bacteria and (h) viruses. The black sections correspond to the fraction of the histogram comprising samples when log(LOP+1)<9 (or LOP<8102 cells/ml). The grey sections are the remaining fraction. This criterion was chosen by eye to give the clearest split of the bipolar LOP distribution into 2 separate unipolar distributions. Similar results are obtained splitting with criterion based on PPH.

Figure 5.

Location of samples where abundance of (a) LOP and (b) PPH falls within the numerically dominant bins for the histograms shown in Figure 4. Black and grey denote samples within the most abundant bin for the same halves of the dataset depicted in Figure 4. For LOP this corresponds to 7.9577<log(abundance+1)<8.1185 for black and to 9.5653<log(abundance+1)<9.7261 for grey. For PPH this corresponds to 9.2082<log(abundance+1)<9.3398 for black and to 10.2617<log(abundance+1)<10.3933 for grey. Abundances are in units of cells/ml. Squares denote the 20 highest abundances and diamonds the 20 lowest. Also shown are temperature (degrees C) contours for reference.

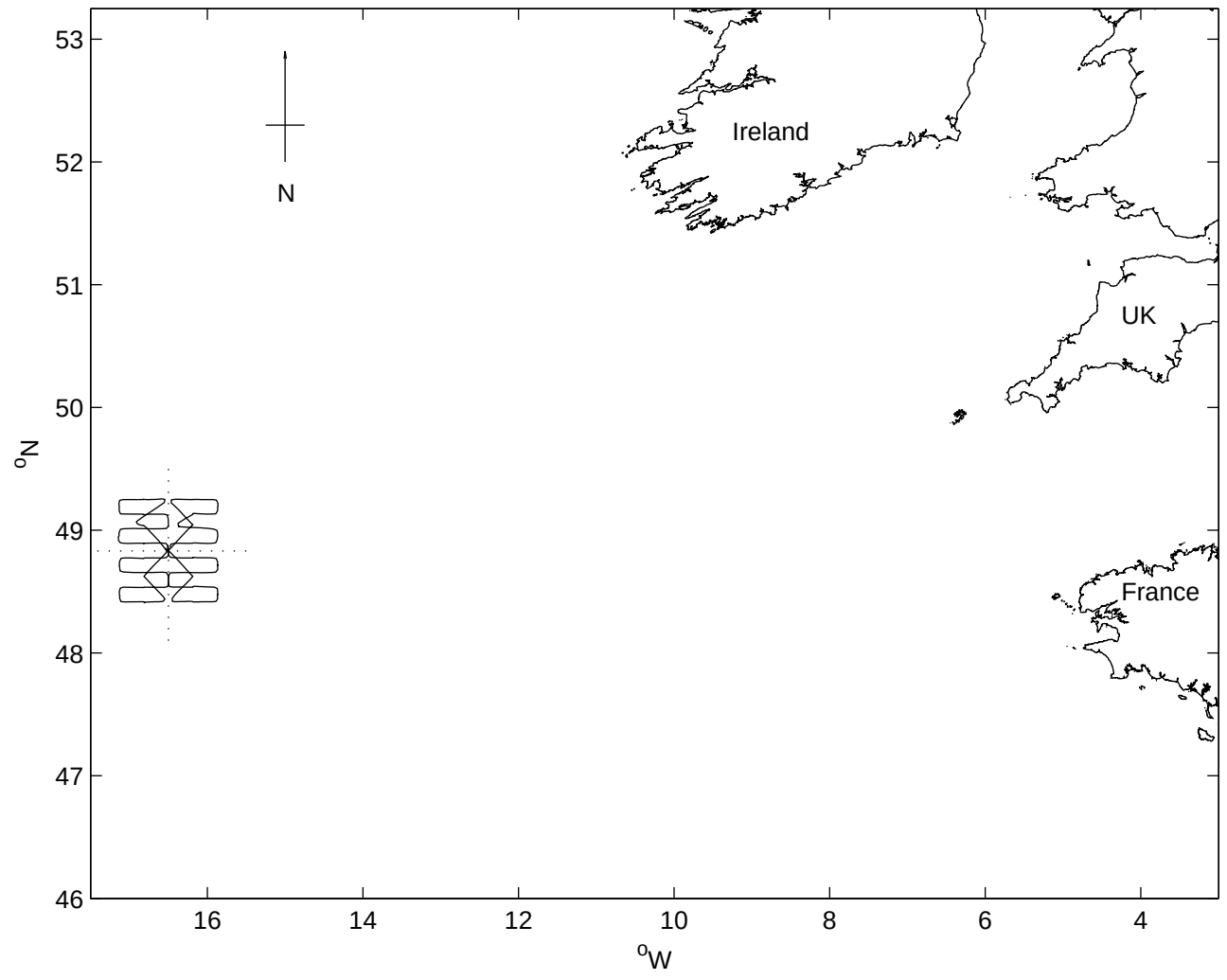
Figure 6.

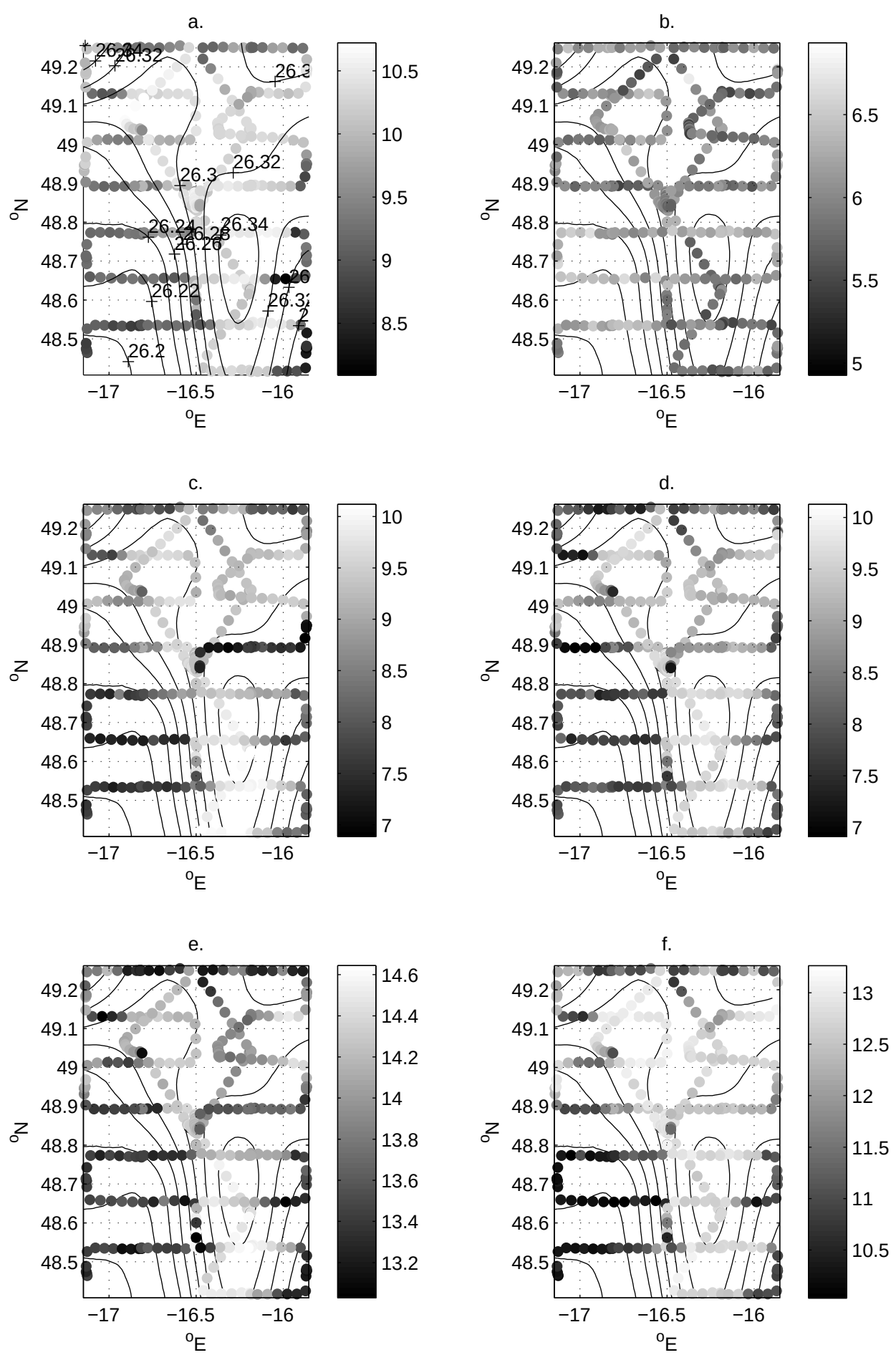
Measure of error in estimated mean as a function of the number of samples. The error is defined as follows. For a given number of samples N, the error is the value Y for which there is a 99% probability that the estimate lies between (1-Y)*μ and (1+Y)*μ where μ is the true mean. For example, a value of Y=0.7 means that the estimate will be within 70% of the true mean with 99% probability. (a) temperature (degrees C), (b) PPH phytoplankton group, (c) large heterotrophic protists, (d) small heterotrophic protists, (e) *Synechococcus*, (f) LOP phytoplankton group, (g) heterotrophic bacteria and (h) viruses. Solid (low abundances) and dashed (high abundances) lines correspond to the split defined in Figure 4.

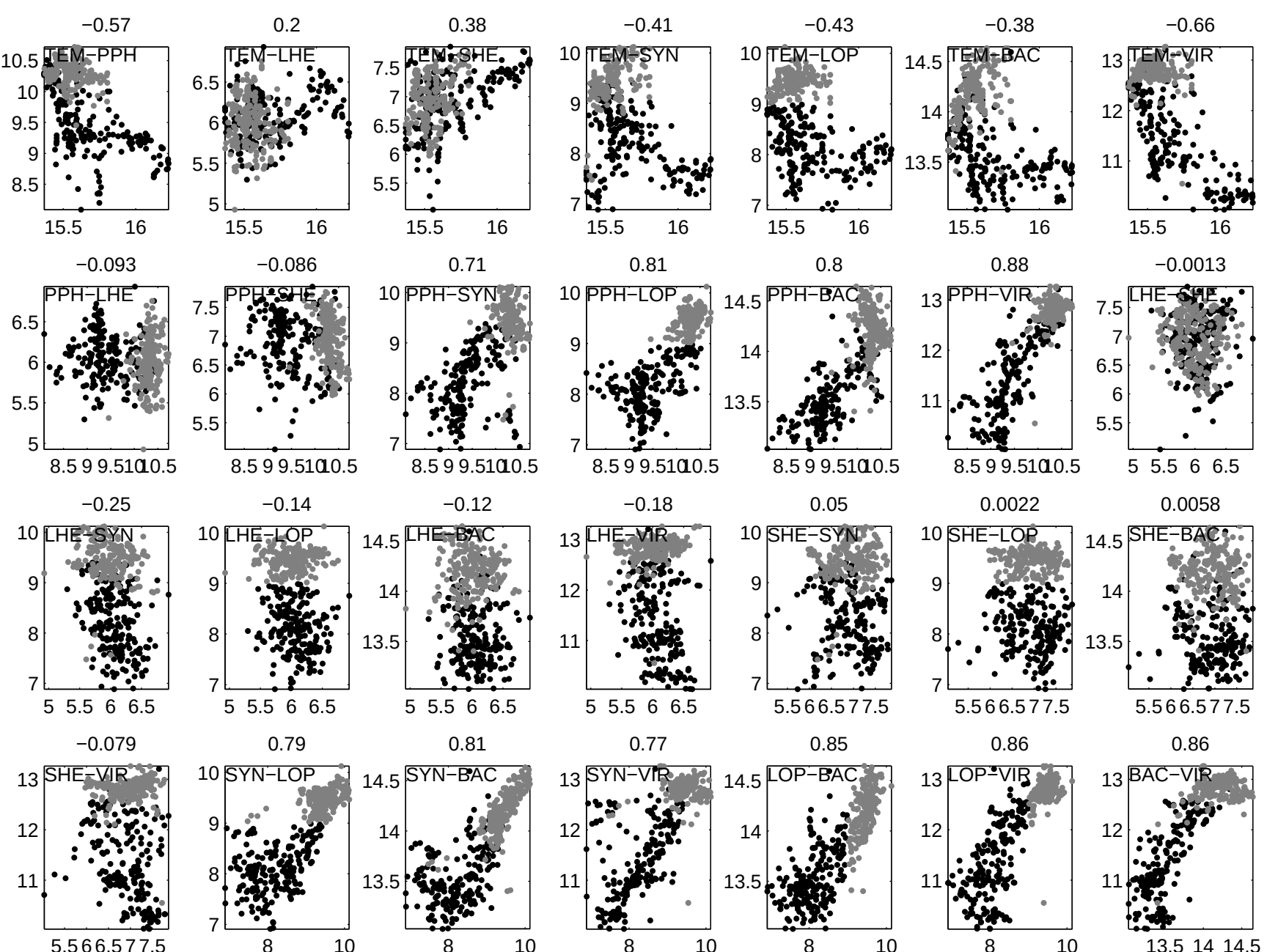
Figure 7.

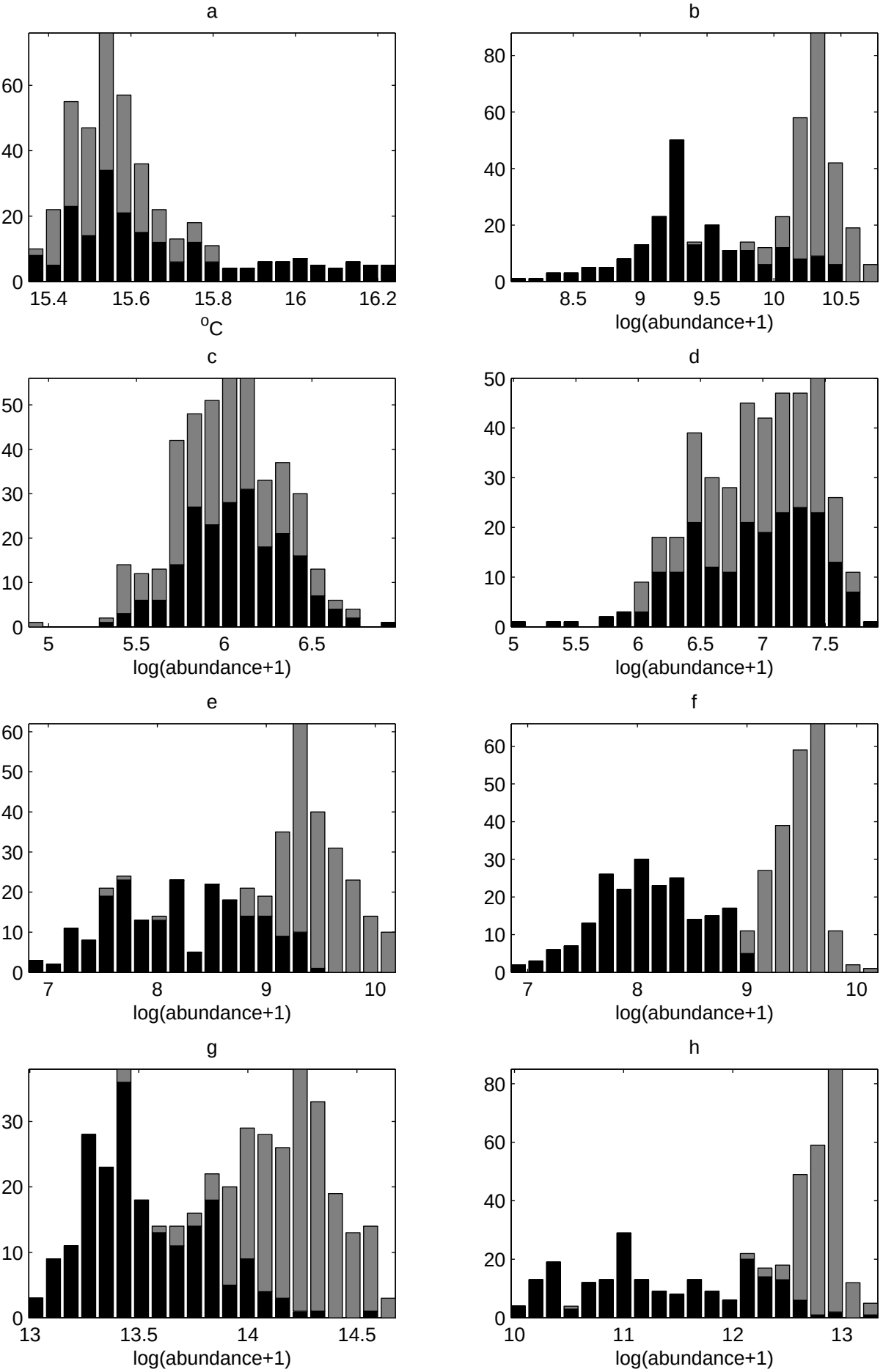
Histogram of log(abundance +1) for *Synechococcus* where abundance is in units of cells/ml. The split corresponds to that described in the caption to Figure 4, with black/grey bars corresponding to low/high abundances respectively. Superimposed are data from AMT13, AMT16, AMT17 and Poseidon 300/1. Squares denote data from any time of year anywhere within the Porcupine Abyssal Plain, whilst crosses indicate data from the same area but restricted to June and July. Circles denote previous data obtained within the boundaries of the D306 mesoscale survey with diamonds the same data restricted to the months of June and July once more.

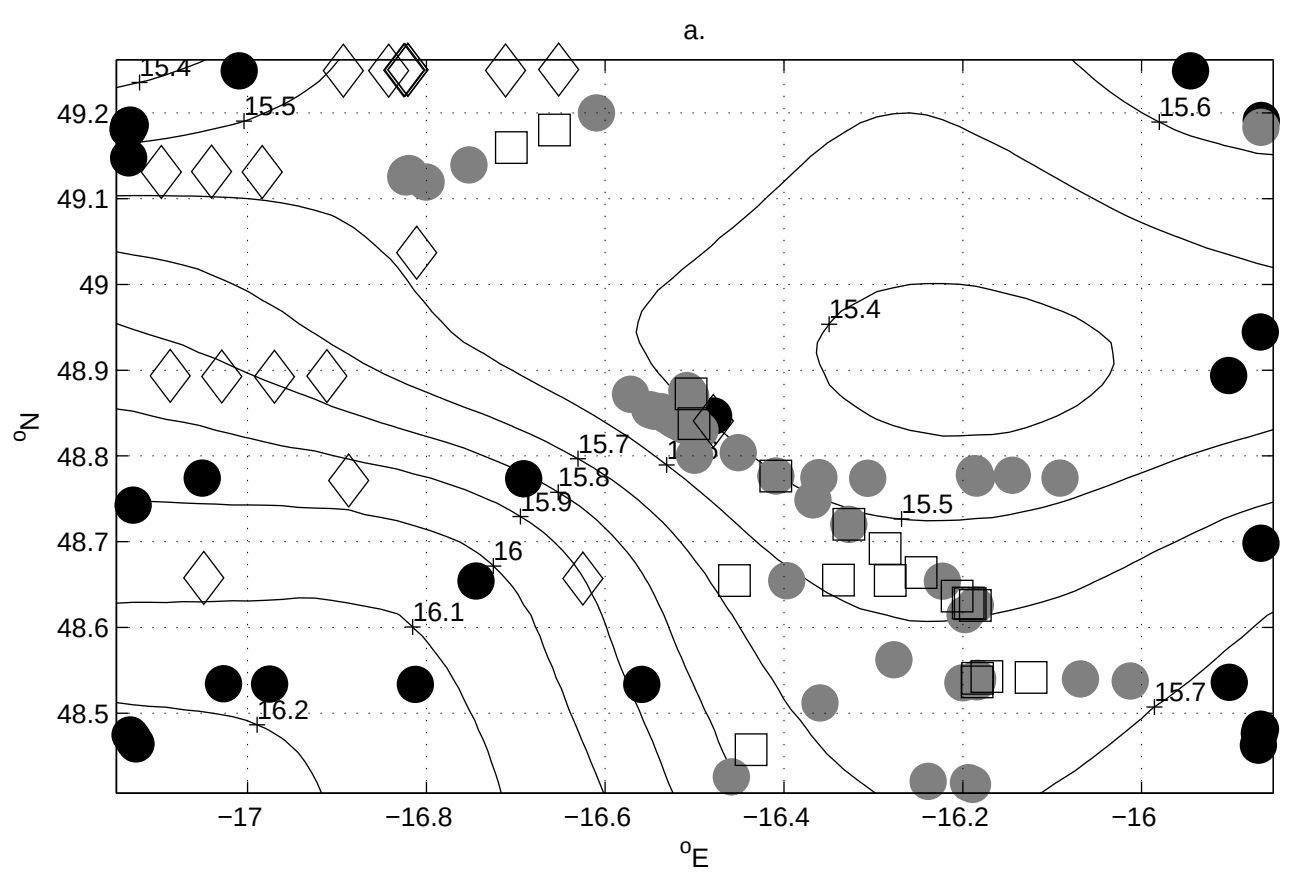
523

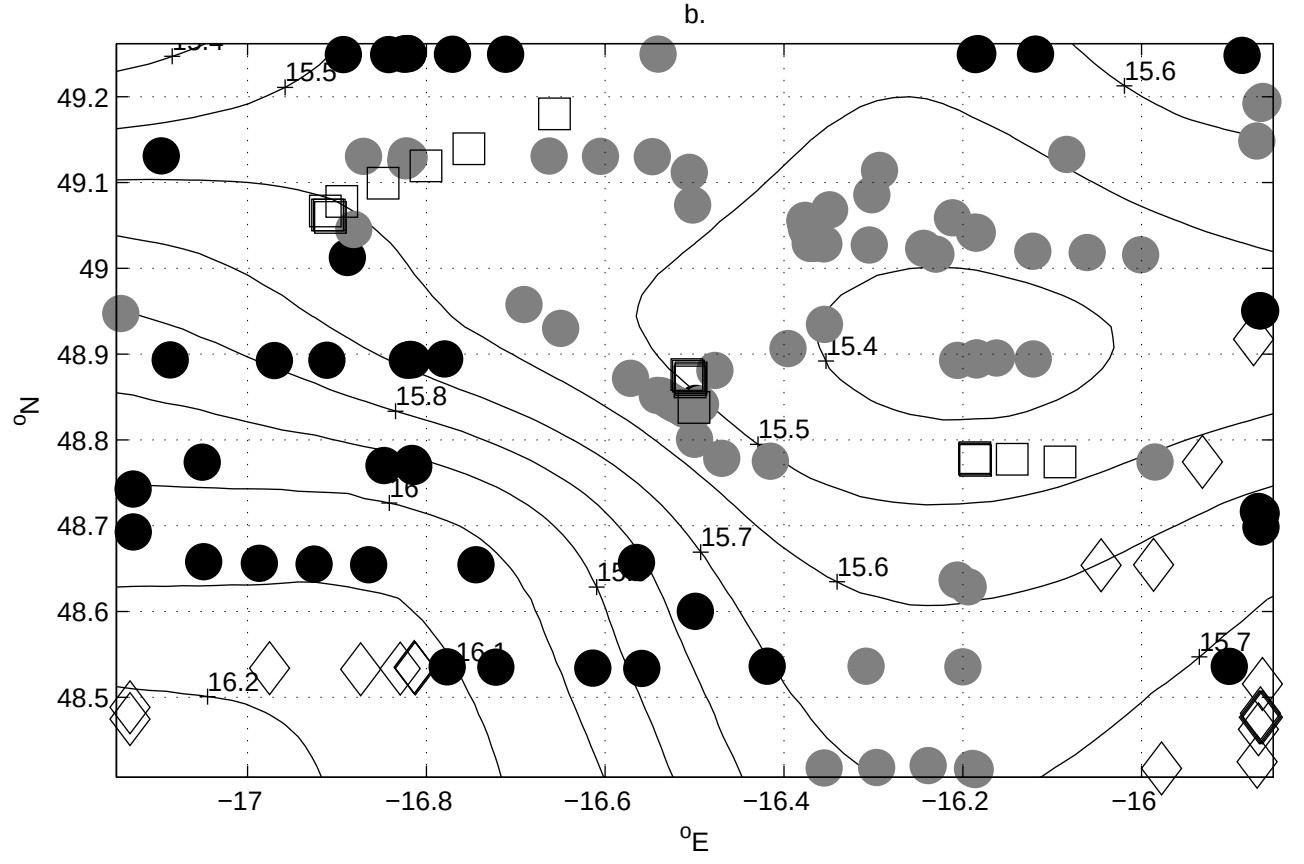

Ireland
UK
France
N
53
52
51
50
49
48
47
46
°N
16
14
12
10
8
6
4
°W

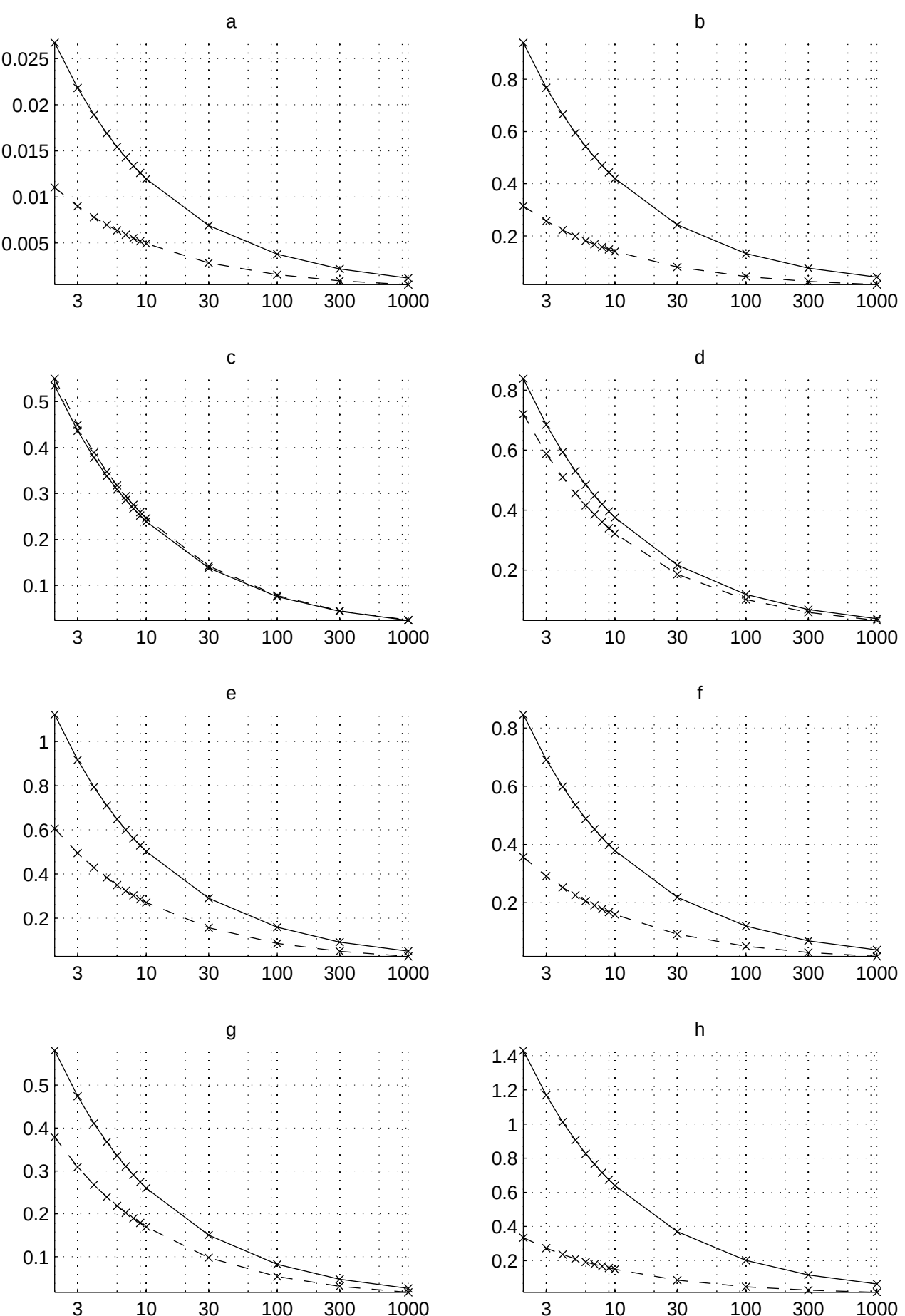

a.
b.
c.
d.
e.
f.
°N
°E
49.2
49.1
49
48.9
48.8
48.7
48.6
48.5
−17
−16.5
−16
26.32
26.3
26.32
26.3
26.24
26.34
26.26
26.22
26.2
10.5
10
9.5
9
8.5
6.5
6
5.5
5
10
9.5
9
8.5
8
7.5
7
10
9.5
9
8.5
8
7.5
7
14.6
14.4
14.2
14
13.8
13.6
13.4
13.2
13
12.5
12
11.5
11
10.5

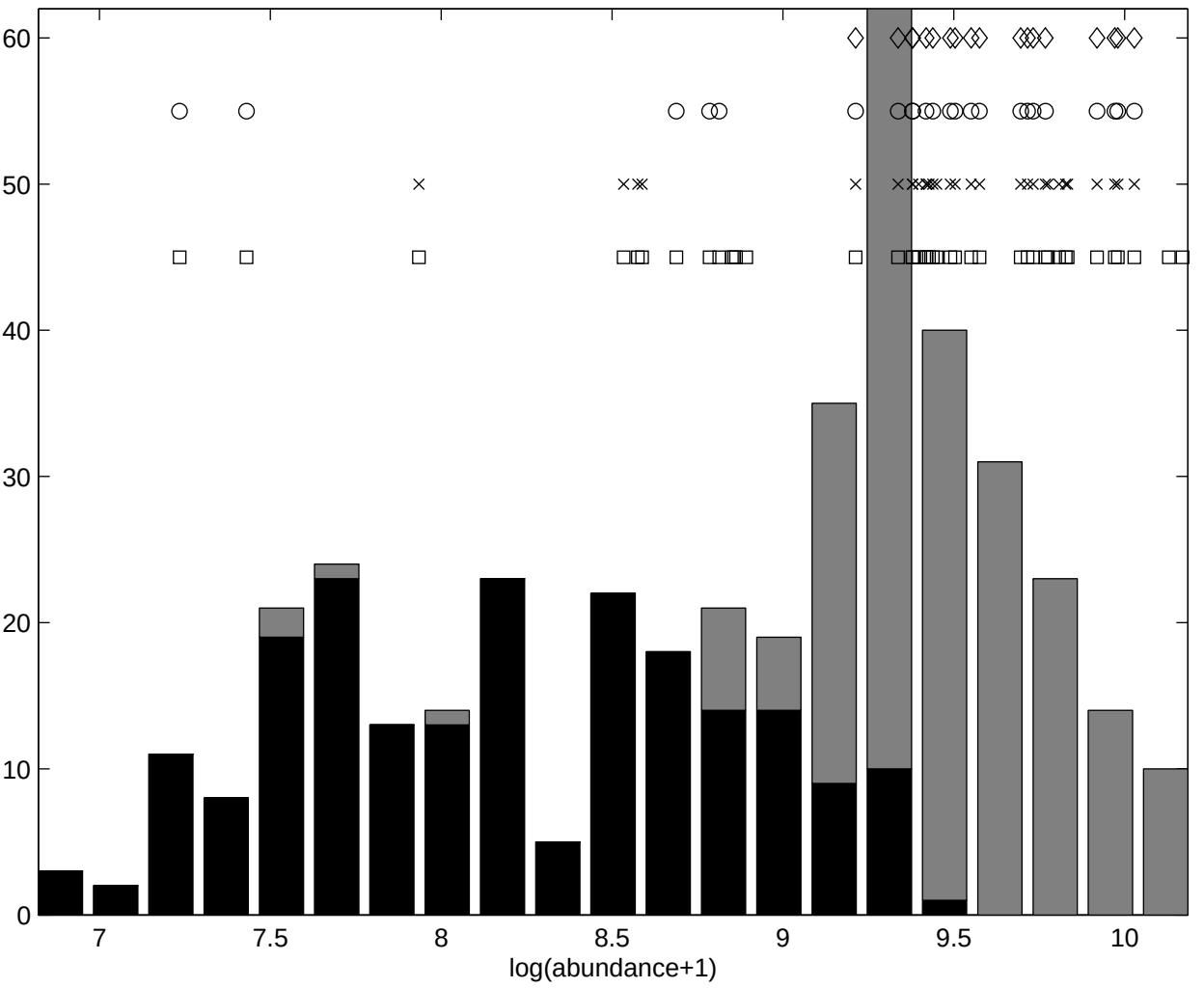

−0.57
TEM−PPH
0.2
TEM−LHE
0.38
TEM−SHE
−0.41
TEM−SYN
−0.43
TEM−LOP
−0.38
TEM−BAC
−0.66
TEM−VIR
−0.093
PPH−LHE
−0.086
PPH−SHE
0.71
PPH−SYN
0.81
PPH−LOP
0.8
PPH−BAC
0.88
PPH−VIR
−0.0013
LHE−SHE
−0.25
LHE−SYN
−0.14
LHE−LOP
−0.12
LHE−BAC
−0.18
LHE−VIR
0.05
SHE−SYN
0.0022
SHE−LOP
0.0058
SHE−BAC
−0.079
SHE−VIR
0.79
SYN−LOP
0.81
SYN−BAC
0.77
SYN−VIR
0.85
LOP−BAC
0.86
LOP−VIR
0.86
BAC−VIR


a
b
c
d
e
f
g
h
°C
log(abundance+1)
log(abundance+1)
log(abundance+1)
log(abundance+1)
log(abundance+1)
log(abundance+1)
log(abundance+1)
log(abundance+1)


a.
15.4
15.5
15.6
15.4
15.7
15.8
15.9
16
16.1
16.2
15.5
15.7
49.2
49.1
49
48.9
48.8
48.7
48.6
48.5
°N
−17
−16.8
−16.6
−16.4
−16.2
−16
°E


b.
15.5
15.6
15.4
15.8
15.5
16
15.7
15.6
16.1
16.2
15.7
49.2
49.1
49
48.9
48.8
48.7
48.6
48.5
°N
−17
−16.8
−16.6
−16.4
−16.2
−16
°E


a
b
c
d
e
f
g
h
3
10
30
100
300
1000


60
50
40
30
20
10
0
7
7.5
8
8.5
9
9.5
10
log(abundance+1)