# Distribution of pelagic sound scattering layers across the southwest Indian Ocean

Philipp H. Boersch-Supan[1,2,3,*], Alex D. Rogers[2], Andrew S. Brierley[1]

25th October 2016

1. Pelagic Ecology Research Group, Scottish Ocean Institute, University of St Andrews, Fife KY16 8LB, UK
2. Department of Zoology, University of Oxford, Oxford OX1 3PS, UK
3. Present address: Department of Integrative Biology, University of South Florida, Tampa, Florida 33620, USA
* philipp@boersch-supan.de

**Abstract**

Shallow and deep scattering layers (SLs) were surveyed with split-beam echosounders across the southwest Indian Ocean (SWIO) to investigate their vertical and geographical distribution. Cluster analysis was employed to objectively classify vertical backscatter profiles. Correlations between backscatter and environmental covariates were modelled using generalized additive mixed models (GAMMs) with spatial error structures. Structurally distinct SL regimes were found across the Subantarctic Front. GAMMs indicated a close relationship between sea surface temperature and mean volume backscatter, with significantly elevated backscatter in the subtropical convergence zone. The heterogeneous distribution of scattering layer biota reflects the biogeographic zonation of the survey area and is likely to have implications for predator foraging and carbon cycling in the Indian Ocean.

## 1   Introduction

Pelagic biota in the open ocean are structured at a broad range of spatial and temporal scales (Haury et al., 1978; Steele, 1978). Abundance and assemblage composition vary on scales from millimetres, to miles, to entire ocean basins. Temporal variation encompasses minutes to decades, incorporating daily patterns such as diel vertical migration (Angel, 1985), seasonal cycles (Urmy et al., 2012;

Cisewski et al., 2010) and inter-annual trends (O'Driscoll et al., 2009). Drivers of this heterogeneity are numerous physical and biological factors such as temperature, mixing, current patterns, predator density and food availability.

A substantial proportion of biomass below the photic zone is concentrated in so called sound scattering layers (SLs) which have been observed since the invention of echosounders in the mid 20$^{th}$ century (Christensen et al., 1946; Lyman, 1947). Scattering layer assemblages encompass species from diverse taxa such as myctophid and stomiiform fishes, squids, decapod shrimps and various groups of gelatinous zooplankton (Farquhar, 1977; Bradbury et al., 1970; Pakhomov and Froneman, 2000). They often exhibit high species diversity (e.g. Kemp et al., 2014; Letessier et al., 2014, this issue) and as a result species discrimination and consequently direct biomass estimation are usually not possible from acoustic data alone. Nonetheless backscatter data, particularly when using multiple acoustic frequencies, can provide rich ecological information. Already in the early decades of scattering layer research, attempts were made to link acoustic information with species abundance and diversity data. For example, investigations were undertaken to determine whether pelagic biogeographical provinces were reflected in the acoustic backscatter signature (Backus and Craddock, 1977; Tont, 1976) or to determine the effects of hydrographic features such as fronts and eddies on scattering layer structure and composition (Cole et al., 1970; Conte et al., 1986). More recently, distinct differences in the vertical structure and/or migration dynamics of pelagic SLs combined with species turnover have been observed across mesoscale oceanographic fronts (Lara-Lopez et al., 2012; Ohman et al., 2013; Cox et al., 2013) and eddies (Godø et al., 2012), confirming that biogeographic patterns are not only reflected in backscatter spectra but also the vertical organisation of scattering layers.

A large proportion of scattering layer biomass undertakes diel vertical migrations (DVM), generally moving from the mesopelagic zone (200-1000 m) into more productive near-surface waters at night, and returning to depth at dawn. This behaviour is thought to balance predation pressure with the need for foraging (Pearre, 2003). Vertical migrators play an important role in marine foodwebs as they link surface zooplankton with higher trophic levels, but also in marine biogeochemical cycles, as they actively contribute to the downward flux of nutrients and particulate matter (Hidaka

et al., 2001; Al-Mutairi and Landry, 2001; Steinberg et al., 2008). Both aspects have only recently been incorporated into in global models of the Earth system (Lehodey et al., 2010; Bianchi et al., 2013) and there is a need for a better understanding of biomass distributions and migration dynamics as inputs into these ecosystem models. This need is particularly pressing, as there is great uncertainty about the global standing stock of scattering layer biota. Recent acoustic and modelling studies suggest that the global biomass of mesopelagic fish may have been underestimated by an order of magnitude or more (Kloser et al., 2009; Kaartvedt et al., 2012; Irigoien et al., 2014).

The predators of SL organisms are a diverse set of oceanic nekton, including species of commercial interest such as tunas and billfish (Potier et al., 2007, 2005), marine mammals (Naito et al., 2013; Arranz et al., 2011) and species of conservation concern such as oceanic sharks (Howey-Jordan et al., 2013). Trophic interactions between scatterers and their predators are constrained in space and time by the vertical SL structure and SL migration dynamics on one hand, and the diving capabilities of the predators on the other. Along the continental slope and near oceanic islands and seamounts abrupt topography can mediate predator-prey relationships, as SL organisms can become trapped on these physical features by being blocked on their vertical descent (Isaacs and Schwartzlose, 1965) and/or as a result of diverse flow-topography interactions (Genin, 2004). The import of oceanic biomass through biophysical coupling has been suggested as an explanation for the occurrence of rich benthic communities (Genin et al., 1986; Rowden et al., 2010) as well as substantial aggregations of high trophic level biota around seamounts (Tseytlin, 1985; Koslow, 1997; Morato et al., 2008, 2009), of which the latter are targeted globally by commercial high-seas fisheries (Pitcher et al., 2010).

Despite the vast expanse of the mid-ocean ridge system, and the ubiquity of seamounts in the global ocean (Yesson et al., 2011; Kim and Wessel, 2011), little is known about the influence of these topographic features on pelagic scattering layers and, ultimately, oceanic food-webs and biogeochemical cycles at a global scale. While locally enriched biomass has been observed, Priede et al. (2013) suggested that, on a basin scale, ridges have a neutral effect on biomass, as the increased benthic biomass on topographic features is balanced by the displacement of pelagic biomass by the topography itself. Opdal et al. (2008) describe scattering layers along the northern Mid-Atlantic

Ridge and possible topographical, hydrographical and biological processes shaping them. They report increased pelagic backscatter above a fracture zone that coincides with the northern Subpolar Front. Cox et al. (2013) used multi-frequency acoustics to estimate epipelagic (<200 m) zooplankton biovolumes and scattering classes across the same front. Their acoustic observations are indicative of faunal turnover across the Subpolar Front, but do not detect a ridge effect.

Apart from a largely descriptive study of scattering layer structure and elephant seal foraging (Boersch-Supan et al., 2012), little research has been conducted on the large scale distribution of scattering layers in the Southern Indian Ocean or along the Southwest Indian Ridge (SWIR) and their variability with respect to latitudinal environmental gradients. This study intends to fill this gap. An understanding of the influence on scattering layers of the regional oceanography and the ridge as a whole is not just important in understanding ecological and biogeochemical processes at the basin scale, but also as a background to understanding the local-scale interactions (c. 10 km) between pelagic biota and individual seamounts.

Fisheries acoustic data possess a number of properties which make their statistical analysis challenging. These features include patchiness, scale dependency and spatio-temporal correlation. Modelling acoustic data in a statistically appropriate matter is therefore not trivial, and there is currently a lack of appropriate analytical tools (Ciannelli et al., 2008). Spatio-temporal correlation is particularly pertinent for observations of oceanic SLs, with their continuous extent over large geographic distances, and their often complex vertical migration cycles. Individual sampling units along a vessel's trajectory are highly correlated in space and time. Furthermore space and time are usually confounded in opportunistically collected data, where the movement of the vessel is largely determined by other objectives such as fishing or other surveying tasks and often no replicates are collected at any one location. Positive spatial auto-correlation violates the usual assumption of independence between data points and leads to the underestimation of standard errors, and elevated type I errors, if not accounted for (Legendre, 1993).

Fisheries acoustics are widely employed in routine stock assessments (Fernandes et al., 2002) and a variety of statistical approaches both in survey planning and data analysis are employed to estimate and minimize the effects of non-random sampling in space (e.g. Rivoirard et al., 2000;

Petitgas, 2001). Most scattering layer studies, however, have not addressed the issue of spatio-temporal correlation in a systematic fashion. A substantial proportion of the literature is largely descriptive (e.g. Opdal et al., 2008; Klevjer et al., 2012), while the more quantitative studies tend to address autocorrelation by averaging over space and/or time (e.g. Godø et al., 2012) or reducing spatially explicit observations to derived indices (e.g. Burgos and Horne, 2008; Urmy et al., 2012), resulting in the loss of information, and complicating the transfer of results into the parameterisation of spatio-temporally explicit ecosystems models. Yet other studies do not address the issue of spatial autocorrelation at all (e.g. Hazen and Johnston, 2010).

More recently, generalized additive models (GAMs; Wood 2006) have gained popularity for modelling non-linear relationships between acoustic backscatter and environmental variables (e.g. Zwolinski et al., 2009; Hazen and Johnston, 2010; Murase et al., 2013). GAMs are non-parametric extensions of generalized linear models with a linear predictor involving a sum of smooth functions of covariates. The advantage of a GAM over the explicit specification of non-linearities e.g. in a non-linear least squares framework, is that no *a priori* functional relationship between response and dependent variable has to be assumed. GAMs are data-driven and the data determine the nature and smoothness of the relationship between response and predictors. Cross-validation or likelihood-based methods are employed to prevent over-fitting (Wood, 2006). These methods, however, rely on independent observations, and will not adequately perform when confronted with non-independent observations. In this case there is a risk that an overly-complex smooth term is fitted, leading to false estimates of the functional relationship between a predictor and the dependent variable, as well as the associated confidence intervals (Ciannelli et al., 2008).

In this study correlations between environmental variables and backscattering strength are explored using a modelling framework based on linear mixed models (Pinheiro and Bates, 2000) that incorporates spatial correlation structures, while retaining the flexibilities of GAMs.

## 2 Material and Methods

### 2.1 Data collection

Basin-scale observations of acoustic backscatter were made from two research vessels (*RV Dr. Fridtjof Nansen*, DFN; *RSS James Cook*, JCO) and three fishing vessels in the period between November 2009 and April 2012. Data from the research vessels were collected for the purpose of this thesis, while the data from fishing vessels were collected as part of the Australian Integrated Marine Observing System's Bio-Acoustic Ship Of Opportunity Programme (IMOS; Ryan, 2011). IMOS data were obtained through the IMOS Ocean Portal (http://imos.aodn.org.au/imos/) as mean echointegrated volume backscattering coefficient ($s_v$ (m$^{-1}$); MacLennan et al., 2002) for cells of 1000 m distance and 10 m depth.

Data from fishing vessels were collected using Simrad ES60 echosounders (Kongsberg Maritime AS, Horten, Norway) operating at 38 kHz, while the research vessels employed Simrad EK60 echosounders operating at 38 kHz and 18 kHz. All echosounders were calibrated following the recommendations of Foote et al. (1987), and accounting for the systematic triangle wave error that is embedded in the ES60 data (Ryan and Kloser, 2004). Calibration parameters for DFN and JCO can be found in Table S1. All sampling programmes involved the collection of data while the vessels were in transit between sampling stations or fishing locations, usually at speeds around 5 ms$^{-1}$. Survey tracks are illustrated in Figure 1.

### 2.2 Acoustic data processing

Echoview software (Version 4.90, Myriax Pty Ltd, Hobart, Tasmania, Australia) was used to visualise acoustic data in the form of calibrated echograms. Transducer parameters were adjusted during this step using calibration data and environmental parameters were calculated from CTD data using algorithms of Fofonoff and Millard (1983). Processing of the DFN and JCO data followed the IMOS processing framework (Ryan, 2011), which involved filters to remove intermittent noise spikes (Anderson et al., 2005), attenuated pings, persistent intermittent noise and finally background noise (De Robertis and Higginbottom, 2007).

Integrated mean volume backscatter ($S_v$ (dB re m$^2$); MacLennan et al., 2002) was exported as georeferenced cells of 1000 m distance and 10 m depth, to a maximum depth of 1000 m. IMOS data were converted into the log-domain using the relationship

$$S_v = 10 \log_{10}(s_v).$$ (1)

Apparent solar time for every cell was calculated from midpoint longitude $lo_i$ and UTC timestamp $t_{UTC,i}$ as

$$t_{local,i} = t_{UTC,i} + (lo_i/15) \times 3600 + EoT(t_{UTC,i})$$ (2)

where $EoT$ is the equation of time (Meeus, 1998) as implemented in the insol package for R (Corripio, 2013). Sunrise and sunset times were calculated using routines provided by the maptools package (Lewin-Koh and Bivand, 2010). Cells were classified into day, night or twilight, where twilight was defined as 1.5 hours before and after sunset and sunrise (Kloser et al., 2009), and into Austral Summer or Austral Winter.

## 2.3 Environmental covariates

As no *in situ* observations for environmental parameters were available for the IMOS data, environmental parameters were sampled for each acoustic sampling unit from gridded data products based on remote-sensing observations of the sea surface. These predictors can only serve as habitat-proxies, as the environment at the sea surface is not experienced directly by the vast majority of SL organisms

Sea surface temperature (SST) was sampled from the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA; Donlon et al., 2012) at a spatial resolution of 1/20°. Surface chlorophyll concentrations were sampled from GlobColour grids at a spatial resolution of 1/24° (Anon., 2010) using monthly model-averaged observations for case 1 water (CHL$_1$ parameter of GlobColour). Water depth was sampled from the GEBCO bathymetric grid at a resolution of 30 arc seconds (GEBCO, 2010). Sampling from gridded datasets was achieved using a combination of the netCDF operators (NCO Version 4.3.0; Zender, 2008) and the raster package for R (Hijmans and van Etten, 2012).

185  Distance to the ridge and distance to the nearest coastline were calculated using GRASS (Ver-
186  sion 7.0.svn; Neteler et al., 2011) and the geosphere package for R (Hijmans et al., 2012). Spreading
187  ridge axes were taken from Coffin et al. (1998). The axis of the Ninety East Ridge, as well as all
188  shorelines were extracted from GEBCO bathymetry.

189  Remotely sensed SST and chlorophyll concentration were validated against *in situ* surface meas-
190  urements of temperature and fluorescence made using an SBE 21 SeaCat thermosalinograph (Sea-
191  Bird Electronics, Bellevue, WA, USA) during the two scientific expeditions. OSTIA SST showed
192  a consistent positive offset of about 1°C across the measured SST range, but was otherwise in very
193  close agreement (Pearson's $r = 0.994$, Fig. S1A). GlobColour monthly averaged chlorophyll con-
194  centrations correlated reasonably well with *in situ* fluorescence measurements (Pearson's $r = 0.536$,
195  Fig. S1B). The weaker correlation between the two chlorophyll measures is likely the result of
196  a strong diel fluctuation of the *in situ* fluorescence signal, which is an artefact of physiological
197  and photochemical effects rather than an indication of biomass fluctuations (Falkowski and Kolber,
198  1995).

## 2.4  Statistical analyses

200  Analyses were performed using the R environment for statistical computing (Version 2.14; R De-
201  velopment Core Team, 2010).

### 2.4.1  Cluster analysis

203  Vertical $S_v$ profiles were subjected to cluster analysis in an attempt to objectively classify scatter-
204  ing layer regimes with differing vertical structures. Unconstrained clustering was achieved using
205  $k$-medoids clustering as implemented in the partitioning around medioids algorithm (PAM; Kauf-
206  man and Rousseeuw, 1990) as implemented in the cluster package for R (Maechler et al., 2012).
207  PAM was run on distance matrices based on euclidean distances as well as a similarity coefficient
208  described by Gower (1971, see suppl. methods for details).

### 2.4.2 Statistical modelling

Generalised additive mixed models (GAMMs; Wood, 2006) were used to investigate the relationship between environmental variables and backscattering strength in different depth strata and at different frequencies. This family of models allows the inclusion of correlation structures to model the inherent autocorrelation of acoustic survey data.

Models were fitted using different combinations of explanatory variables and optimal models for each response were selected by Akaike's information criterion (AIC; Akaike, 1973). Where multiple models had similarly low AIC scores and did not show significant differences in a likelihood ratio test, the more parsimonious model was selected. Time of day was modelled using a cyclic regression spline. Autocorrelation of residuals was modelled using a first-order autoregressive error structure (AR(1); Pinheiro and Bates, 2000). The AR(1) model was nested within individual survey legs. Models were fitted to integrated backscatter between the surface and 200 m (shallow scattering layer; SSL) and between 200 m and 1000 m (deep scattering layer; DSL) using an identity link function and Gaussian errors.

# 3 Results

A total of 33 406 km of survey track were analysed in this study, largely based on 38 kHz data. 6 295 km of survey track also included observations at 18 kHz. A detailed overview of the data sources and their distribution with respect to echosounder frequency, season and time of day is provided in Table 1.

Survey coverage from JCO is sparse, as interference with geophysical and oceanographic echosounders prohibited continuous data collection. Furthermore severe bubble entrainment under the JCO's hull severely degraded data quality and resulted in the rejection of over 80% of the collected pings during initial processing.

## 3.1 Vertical structure

Variations in the vertical backscatter structure were examined using cluster analysis. Average silhouette widths indicated weak structure ($\bar{s}(k) \approx 0.25$; Nagpaul, 1999) and favoured a two cluster

solution in all cases. Partitions based on Gower's similarity coefficient had consistently higher $\bar{s}(k)$ values than those based on euclidean distances, indicating tighter clustering (Fig. S2).

### 3.1.1  38 kHz data

For the 38 kHz dataset the composition of the optimal partitions differed markedly between the two underlying dissimilarity metrics. The two clusters based on euclidean distances (E1, E2; Fig. 2A) cover the entire geographical range of the data, but their composition with respect to time-of-day is distinct. Cluster E1 contains mostly daytime samples and cluster E2 mostly nightime samples (Fig. S3A). The vertical backscatter profiles differ the most in the top 400 m of the water column (Fig. 3A), with the night time cluster E2 exhibiting increased backscatter in this depth interval. The two 38 kHz clusters based on Gower's distances (G1, G2; Fig. S3B) are largely indistinguishable with respect to time-of-day, but are well separated in geographic space (Fig. 2B). Cluster G2 occupies the latitudinal centre of the survey area (32°S-42°S), while cluster G1 is geographically discontinuous and occupies the northern (<32°S) and southern (>42°S) edges of the survey area. The largest difference in the backscatter profiles is found between 200 m and 700 m depth (Fig. 3B), with mid-latitude cluster G2 exhibiting elevated backscatter in this depth interval, as well as a slight upward shift of the shallow $S_v$ maximum in G2 to approx. 50 m. Based on the bimodal latitudinal distribution, G1 can be separated into two subclusters (G1a, G1b; Fig. 3C), both of which show less backscatter than the central cluster. The depth profile of the northern cluster G1a shows a distinct shallow and deep scattering layer, which is similar to the layer structure in G2, whereas the layer structure in the south (G1b) is quite distinct indicating several, less pronounced sub-layers.

### 3.1.2  18 kHz data

For the 18 kHz data the $k$=2 partitions for both dissimilarity measures resulted in a day-night split. In an attempt to extract a possible geographic structure the dataset was subdivided into daytime and nighttime $S_v$ profiles and subjected to separate PAM analyses based on Gower dissimilarities (Fig. S2C), as clusters based on this distance measure had higher $\bar{s}(k)$ values compared to euclidean distances. Spatially separated clusters were obtained for both time periods, with a North-South

divide near 35°S in both cases (Fig. 2C,D). Daytime $S_v$ profiles differed between the two clusters in the depth of the subsurface $S_v$ maximum (Fig. 3D). In the northern cluster D1 deep backscatter peaks around 400 m, whereas D2 exhibits a peak some 50 m shallower as well as elevated backscatter in the surface layer. Nighttime $S_v$ profiles differed in backscattering strength below 100 m, with consistently lower levels in the Northern cluster N2. A short segment of a transect near 42°S 48°E was classified as belonging to the northern cluster N2 (Fig. 2D). $S_v$ profiles in this segment differed from the remainder of cluster N2 by the presence of two deep $S_v$ maxima near 100 m and 500 m depth, respectively (Fig. 3F).

## 3.2 Environmental drivers of backscattering strength

Temperature and time of day consistently emerged as important predictors in all models, while chlorophyll concentration was the only predictor that was not included in any of the final models. The correlation coefficient $\phi$ of the AR(1) error structure was close to 1 in all cases, indicating strong positive autocorrelation among the residuals. Sample variograms showed increasing semivariance up to a range of approx. 100 km (Fig. S4).

### 3.2.1 38 kHz data

The optimal model for $S_v$ $^{\mathbf{DSL38}}$, the backscatter from the DSL at 38 kHz included smooths of SST (stratified by season), time of day and distance to the nearest shoreline ($r^2_{adj}$=0.35, AIC=38 418, $\phi$=0.985; Table 2, Fig. 4). Model output showed that variations in SST across the survey area had the largest effect on $S_v$ $^{\mathbf{DSL38}}$. During Austral Summer the deep scattering layer exhibits a pronounced maximum at approximately 18°C and then drops off substantially towards colder temperatures (Fig. 4A). Model predictions for the Winter have much larger standard errors at the extremes of the sampled temperature range, but the model exhibits a local $S_v$ $^{\mathbf{DSL38}}$ maximum at 14°C (Fig. 4B). Time of day shows a weaker effect on $S_v$ $^{\mathbf{DSL38}}$, with a minimum around 03:00 hrs local time and peaks during local daytime (Fig. 4C). Distance to shore proved to be a highly significant predictor during model selection and show moderate effect sizes and a complex relationship with $S_v$ $^{\mathbf{DSL38}}$ (Fig. 4D).

287 The optimal model for backscatter from the SSL at 38 kHz ($S_v$ $^{\textbf{SSL38}}$) included a linear depth

288 term and smooths of SST (stratified by season), time of day, and distance to the nearest ridge

289 ($r^2_{adj}$=0.66, AIC=100 150, $\phi$=0.936; Table 2, Fig. 5). Again, SST showed the largest effect on

290 backscatter, although for the SSL there is a general trend of decreasing $S_v$ $^{\textbf{SSL38}}$ with temperature

291 (Fig. 5A-C). Two local maxima appear during summer near 17°C and 8°C, while during winter there

292 is a pronounced local minimum at 13°C and a local maximum at 9°C. Time of day shows a pro-

293 nounced effect with elevated backscatter at night time, peaking just before and just after midnight

294 and a local maximum during noon (Fig. 5D). The influence of water depth is weak, with $S_v$ $^{\textbf{SSL38}}$

295 increasing linearly towards shallower depths (Fig. 5E), and the distance to ridge term shows fairly

296 stable levels of backscatter within 500 km of a ridge followed by a slight decrease (Fig. 5F).

### 3.2.2   18 kHz data

298 The optimal model for $S_v$ $^{\textbf{DSL18}}$, the backscatter from the DSL at 18 kHz, included smooths of

299 SST, time of day (stratified by location on or off SWIR) and depth, and both distance measures

300 ($r^2_{adj}$=0.73, AIC=15 380, $\phi$=0.815; Table 2, Fig. 6). Model output showed that distance to the Ridge

301 had the largest individual effect on $S_v$ $^{\textbf{DSL38}}$, with a sharp increase at distances beyond 1200 km

302 (Fig. 6E). Temperature also had a large effect with a local $S_v$ $^{\textbf{DSL18}}$ maximum at approximately

303 17°C followed by relatively high values towards colder temperatures and another rise in backscat-

304 ter at temperatures below 10°C (Fig. 6A). Time of day showed a weaker effect on $S_v$ $^{\textbf{DSL38}}$, that

305 differed significantly at locations on the SWIR relative to elsewhere. In both cases, daytime back-

306 scatter was elevated relative to night time backscatter, with afternoon backscatter on the ridge being

307 elevated relative to off the ridge (Fig. 6B,C). The influence of water depth is moderate at depths,

308 with $S_v$ $^{\textbf{SSL38}}$ increasing linearly from 1500 m towards shallower depths (Fig. 6D). No depth effect

309 was seen at depths beyond 1500 m. Distance to shore proved to be a highly significant predictor

310 during model selection and show decreasing $S_v$ $^{\textbf{DSL18}}$ towards both extremes of the predictor range

311 (Fig. 6F).

312 The optimal model for backscatter from the SSL at 18 kHz ($S_v$ $^{\textbf{SSL18}}$) included smooths of SST,

313 time of day (stratified by location on or off SWIR), and distance to the nearest ridge ($r^2_{adj}$=0.55,

314 AIC=20 389, $\phi$=0.943; Table 2, Fig. 7). Time of day showed the largest effect on backscatter. In

315 both strata $S_v$ $^{\mathbf{SSL18}}$ was elevated during the night, but while off the ridge backscatter remains low

316 during the day (Fig. 7B), there is a secondary backscatter peak around noon on the ridge (Fig. 7C).

317 There is a general trend of increasing $S_v$ $^{\mathbf{SSL18}}$ with decreasing temperatures, with a local minimum

318 at 20°C (Fig. 7A). The influence of water depth is comparably weak, with $S_v$ $^{\mathbf{SSL18}}$ decreasing

319 linearly from 3000 m towards deeper depths (Fig. 7D).

# 4 Discussion

## 4.1 Vertical structure

322 Clustering based on Gower's dissimilarities provided evidence of three structurally distinct scat-

323 tering layer regimes for both frequencies. The approximate boundaries of these regions are along

324 32°S and 42°S, which roughly corresponds to the boundaries of the Subtropical convergence zone

325 (SCZ). The northern and central regimes are structurally similar, both exhibiting a pronounced shal-

326 low and deep scattering layer, although the vertical positions of these layers are shifted upwards

327 by 50–100 m in the SCZ. The scattering layer structure in the area south of 42°S was quite differ-

328 ent, with backscatter being distributed more uniformly through the water column. These results

329 confirm an earlier, descriptive analysis of scattering layer structure across the Subantarctic Front

330 (Boersch-Supan et al., 2012), and are in line with step changes in scattering layer structure ob-

331 served across oceanographic fronts elsewhere (Cole et al., 1970; Lara-Lopez et al., 2012; Ohman

332 et al., 2013). In particular, the shallowing of the daytime 18 kHz DSL south of the convergence

333 is consistent with a faunal turnover from species with larger swimbladders in the North to species

334 with smaller swimmbladders in the South, which would shift the resonance depth upwards in the

335 watercolumn. This matches observations from trawl samples, which feature the myctophid *Cerato-*

336 *scopelus warmingii* and the hatchetfish *Argyropelecus aculeatus* as numerically dominant species

337 in the North, whereas the smaller sternoptid *Maurolicus muelleri* is dominant in the South (Kemp

338 et al., 2014, this issue).

339 The Subtropical (STF) and Subantarctic Fronts (SAF) are thought to represent major biogeo-

340 graphic boundaries, dividing distinct faunal provinces in the southern Indian Ocean (Longhurst,

1998; Vierros et al., 2009). Previous research, as well as the biological sampling programme run concurrently to the acoustic surveys analysed in this study, confirms the role of the SAF as a faunal boundary across multiple trophic levels from microbes and phytoplankton (Djurhuus et al., 2014; Bornman et al., 2014, both this issue) and macrozooplankton/micronekton (Pakhomov et al., 1994; Letessier et al., 2014, this issue) to cephalopods (Laptikovsky et al., 2014, this issue) and fishes (Kemp et al., 2014, this issue). As both acoustic properties and vertical position in the water column are species specific properties, the drastic change in the vertical structure of the SLs is not surprising.

## 4.2 Backscatter and sea surface temperature

GAMM results indicate a strong effect of SST on the backscattering strength. There is a pronounced local maximum in backscatter from the DSL at both observation frequencies in temperature range indicative of the STF (Figs. 4A,B and 6A), and this $S_v$ maximum appears to follow the seasonal shift of the STF axial temperature (Figs. 4A,B). Increased backscatter, as well as increased net-sampled macrozooplankton/micronekton biomass have also been observed in the SCZ south of Cape Agulhas (c. 20°E; Barange et al., 1998; Pakhomov et al., 1994), while no net increase in acoustic backscatter was reported across the STF east of New Zealand, where the frontal temperature gradient is much shallower (McClatchie et al., 2004). Another local maximum in DSL backscatter appears in the axial temperature range of the SAF. This more pronounced in the 18 kHz data (Fig. 6A) than in the 38 kHz data (Fig. 4B), and not evident for the Winter subset of the 38 kHz, where sparse observations lead to very large model errors (Fig. 4B). Increased 18 kHz backscatter across a subpolar front was also observed by Opdal et al. (2008) in the North Atlantic.

The pattern is not quite as clear for backscatter from the SSL. The 38 kHz backscatter generally decreases with temperature. The vertical $S_v$ profiles for these data suggest that this might be an artefact caused by the shallowing of the SSL across the SCZ (Fig. 3B), resulting in the movement of scatterers into the so called surface deadzone, i.e. the part of the water column, that is shallower than the downward-looking, hull-mounted echosounder transducer can sample (O'Driscoll et al., 2009). However, the $S_v$ profile for 38 kHz backscatter south of the SAF (Cluster G1b, Fig. 3C) further suggests an actual decrease in surface backscatter in the southern part of the survey area.

This might be reflected in the steeper decrease in $S_v^{\textbf{SSL38}}$ south of the SAF (Fig. 5A), although it is difficult to estimate the relative contributions of both effects. 18 kHz backscatter from the SSL steadily increases across both frontal zones, which for the SAF matches the observations of Opdal et al. (2008) in the North Atlantic.

The results of this study have potential implications for energy flux to the deep sea. The southern hemisphere SCZ has been estimated to provide approximately 5% of the global net primary production in the oceans (Longhurst et al., 1995). If the results of Barange et al. (1998), Pakhomov et al. (1994) and this study are indeed indicative of elevated deep micronekton biomass across the SCZ in the southern Atlantic and Indian Ocean, a larger proportion of this production than previously thought may be actively exported to the deep ocean.

## 4.3 Ridge effect on backscattering strength

No compelling evidence could be found for an effect of the ridge on backscatter. An on ridge/off ridge factor variable was discarded during model selection, as it did not improve model fit, although stratifying the time-of-day smooth by this variable did improve both 18 kHz models. However, most 18 kHz samples within the SCZ were collected on the ridge, confounding potential effects of the frontal zone with those caused by the ridge. In fact, the cluster analysis does not provide evidence for distinct vertical structures between the ridge and elsewhere within the frontal zone (Figs. 2C,D). Depth and distance to the ridge were retained as a predictor in the 38 kHz SSL model (Fig. 5E). Distance to ridge exhibited a negative effect at large distances from the ridge (>750 km), which applies to data sampled near the African continental shelf and in the wider vicinity of Kerguelen Plateau, suggesting no role of the ridges themselves in this result. Depth was the weakest predictor in this particular model and indicates slightly increased backscatter with decreasing depth. More convincing evidence for a ridge effect comes from the 18 kHz DSL model. Large effect sizes (>10 dB, equivalent to >10 fold on the linear scale) are predicted (Fig. 6E), although again, the biggest effect is predicted at ranges >1200 km from the ridge indicating effects of the African and/or Kerguelen shelf break, rather than from the SWIR. The model, does however predict a stepwise reduction of backscatter by approximately 3 dB, equivalent to a two-fold reduction on the linear scale, within

400 km of the ridge. Depth is also retained in the 18 kHz DSL model, with a strong positive effect (3 dB) at depths shallower than 1500 m. Generally, these results have to be treated with care, as both distance measures are rather crude metrics on a basin scale, potentially just capturing unexplained variation caused by unidentified factors.

Similar to Opdal et al. (2008) and Cox et al. (2013), this study could not detect an unequivocal effect of the ridge on backscattering strength. This may be because the ridge does not exert an effect on scattering layers, at the vertical and geographical scale investigated. It contrasts observations of Priede et al. (2013), who present a notable increase of deep scattering layer biomass around the Reykjanes ridge at spatial scales of hundreds of km. However, Priede et al. (2013) ascribe this to behavioural associations with topography and not locally increased pelagic productivity, maintaining that the effect of mid-ocean ridges on basin-scale total biomass is neutral. They do not present any data to support this, and no aggregating effect is apparent for the continental slopes in their dataset, casting some doubt on the aggregation versus enhanced productivity conjecture.

### 4.4 Diel variations in backscatter

Apart from temperature, time of day consistently exhibited pronounced effects on backscatter in all models. Generally model fits indicated high levels of backscatter during the day and lower levels during the night for the DSL, and the opposite pattern for the SSL. This pattern conforms to a DVM cycle with scatterers moving out of the mesopelagic zone and into the euphotic zone at night and returning to deeper water during day (Angel, 1985; O'Driscoll et al., 2009). However, the diel backscatter cycle was more complex, with the SSL exhibiting a secondary peak around local noon as well as a secondary trough at local midnight (Figs. 5D and 5B,C). The causes for this are less clear, although a possible explanation could again be scatterers migrating into and out of the surface deadzone during the extremes of the diel cycle. Some evidence supporting this is provided by the fact that the noon peak of the DVM cycle in the 18 kHz SSL is more pronounced in the "on SWIR" smooth, compared to the "off SWIR" smooth (Fig. 7B,C), which corresponds to a shallower and more intense SSL in the southern cluster D2, compared to the northern cluster D1 (Fig. 3D).

Vertical migrations within and between sampled volumes, however, may not be the only factors

causing diel backscatter variations. The acoustic target strength, i.e. the propensity of a given target to reflect acoustic energy, can be strongly dependent on the aspect at which a target is insonified. Target strength of mesopelagic fishes can vary by several orders of magnitude between extreme tilt angles, i.e. head up vs. head down (Benoit-Bird and Au, 2001; Yasuma et al., 2003, 2010), adding considerable uncertainty to acoustic biomass estimates (e.g. Demer, 2004). Observations of the *in situ* orientation of scattering layer organisms are scarce, but significant day-night differences in the orientation of myctophids and *Cyclothone* spp. have been reported from the Pacific (Barham, 1970). Target strength furthermore is not independent of depth, as migrations through the hydrostatic depth gradient can alter swimbladder volume. This can bias target strengths, in particular near the resonance frequency, leading to artifical increases of backscatter at particular depths (Godø et al., 2009).

## 4.5 Conclusions and outlook

With programs like the IMOS Bio-Acoustic Subfacility, the amount of scattering layer data are likely to increase considerably in the future, and there is an imperative to try to make the most out of these data. Fisheries acoustics are certainly not an unbiased sampling methodology for pelagic biota, but few, if any, sampling technologies can provide information across mid to high-trophic level biota with the same spatial and temporal coverage.

Overall the analysis showed a heterogeneous distribution of pelagic backscatter across the southwest Indian Ocean. Hydrographic features explained a large amount of the variation in the data, with increased backscatter at both the STF and SAF. No strong evidence of a ridge effect was found at the basin scale. Cluster analysis suggests structurally distinct scattering layer regimes north and south of the SAF, which together with differing $S_v$ intensity patterns between the two frequencies south of the SAF indicates that acoustic observations can indeed delineate biogeographical regions. Stratifying acoustic backscatter data into faunal provinces may increase the accuracy of acoustic biomass estimates on global scales, such as the recent estimate by Irigoien et al. (2014), which was based on globally averaged target strength-biomass relationships.

Tracking and biologging studies have highlighted the importance of the SWIO, and in particular

the fronts south of the STF and into the Southern Ocean, as a foraging habitat for predators of meso-pelagic fishes. This includes several million pairs of macaroni penguins *Eudyptes chrysolophus*, the single largest consumer among seabirds (Brooke, 2004; Thiebot et al., 2011), as well as southern elephant seals *Mirounga leonina* (Bailleul et al., 2007; Dragon et al., 2010). It is likely that the elevated backscatter observed in the present study reflects the prey resources exploited by pelagic predators. Tracking data have furthermore indicated differences in elephant seal foraging behaviour and/or foraging success on either side of oceanic fronts (McIntyre et al., 2011; Boersch-Supan et al., 2012; Guinet et al., 2014), which likely demonstrates adaptive foraging behaviour in distinct pelagic biomes, as reflected by distinct vertical SL profiles.

GAMMs performed adequately, although they came with a high computational cost. Individual GAMMs for the full 38 kHz dataset took between 0.3 and 2.2 hrs to fit on a high-performance workstation (MacPro 5,1 with 12 processor cores and 48 GB of memory; Apple Inc., Cupertino, CA, USA), consuming about 7 GB of working memory in the process. GAMs for the same data but without a correlation structure were usually fitted in less than a minute. Computational limitations can probably be addressed by more efficient software and better hardware, but the quest for better analytical tools remains for acoustic data. This study addresses spatial and temporal correlation along the survey track, but is limited to a set of independent models with an arbitrary depth stratum of integrated backscatter as a response variable. This reduces ecological information and leaves some uncertainty in the interpretation of model results, e.g. of the underlying causes for different time-of-day smooths. Ultimately the aim for analyses of SL data should be depth-explicit models that capture both geographical as well as vertical patterns in one unified framework. Such a model could be useful to produce spatially and temporally explicit predictions of SL density and diel vertical migration dynamics, which could in turn be used to parameterize ecosystems models (e.g. Lehodey et al., 2008).

## Acknowledgements

## References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. *In* B. Petran and F. Csaaki (eds.) Second international symposium on information theory, pp. 267–281. Akademinai Kiado, Budapest, Hungary.

Al-Mutairi, H. and M. R. Landry, 2001. Active export of carbon and nitrogen at Station ALOHA by diel migrant zooplankton. *Deep Sea Research Part II: Topical Studies in Oceanography* **48**:2083–2103.

Anderson, C., A. Brierley, and F. Armstrong, 2005. Spatio-temporal variability in the distribution of epi- and meso-pelagic acoustic backscatter in the Irminger Sea, North Atlantic, with implications for predation on *Calanus finmarchicus*. *Marine Biology* **146**:1177–1188.

Angel, M., 1985. Vertical migrations in the oceanic realm: Possible causes and probable effects. *Contributions in Marine Science.* **27**:45–70.

Anon., 2010. ESA DUE GlobColour – Global Ocean Colour for Carbon Cycle Research, Product User Guide. ACRI-ST/LOV, UoP, NIVA, BC, DLR, ICESS consortium. Reference: GC-UM-ACR-PUG-01, Version 1.4.

Arranz, P., N. de Soto, P. Madsen, A. Brito, F. Bordes, and M. Johnson, 2011. Following a foraging fish-finder: Diel habitat use of Blainville's beaked whales revealed by echolocation. *PLoS ONE* **6**:e28353.

Backus, R. and J. Craddock, 1977. Pelagic faunal provinces and sound-scattering levels in the Atlantic Ocean. *In* N. Andersen and B. Zahuranec (eds.) Oceanic sound scattering prediction, *Marine Science*, vol. 5, pp. 529–547. Plenum Press, New York.

Bailleul, F., J. Charrassin, P. Monestiez, F. Roquet, M. Biuw, and C. Guinet, 2007. Successful foraging zones of southern elephant seals from the Kerguelen Islands in relation to oceanographic conditions. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**:2169–2181.

Barange, M., E. Pakhomov, R. Perissinotto, P. Froneman, H. Verheye, J. Taunton-Clark, and M. Lucas, 1998. Pelagic community structure of the subtropical convergence region south of Africa and in the mid-Atlantic Ocean. *Deep-Sea Research Part I* **45**:1663–1687.

Barham, E. G., 1970. Deep-sea fishes lethargy and vertical orientation. *In* G. Farquhar (ed.) Proceedings of an International Symposium on Biological Sound Scattering in the Ocean, pp. 100–118. Maury Center for Ocean Science, Washington, D.C.

Belkin, I. and A. Gordon, 1996. Southern Ocean fronts from the Greenwich meridian to Tasmania. *Journal of Geophysical Research* **101**:3675–3696.

Benoit-Bird, K. and W. Au, 2001. Target strength measurements of Hawaiian mesopelagic boundary community animals. *The Journal of the Acoustical Society of America* **110**:812–819.

Bianchi, D., C. Stock, E. D. Galbraith, and J. L. Sarmiento, 2013. Diel vertical migration: Ecological controls and impacts on the biological pump in a one-dimensional ocean model. *Global Biogeochemical Cycles* .

Boersch-Supan, P., L. Boehme, J. Read, A. Rogers, and A. Brierley, 2012. Elephant seal foraging

dives track prey distributions, not temperature: Comment on McIntyre et al. (2011). *Marine Ecology Progress Series* **461**:293–298.

Bornman, T. G., M. J. Sonnekus, and E. E. Campbell, 2014. Phytoplankton and nutrient dynamics of six south indian seamounts. *Deep Sea Research II* .

Bradbury, M., D. Abbot, R. Bovbjerg, R. Mariscal, W. Fielding, R. Barber, V. Pearse, S. Proctor, J. Ogden, J. Wourms, L. Taylor Jr., J. Christofferson, J. Christofferson, R. McPhearson, M. Wynne, and P. Stromberg Jr., 1970. Studies on the fauna associated with the deep scattering layers in the equatorial Indian Ocean, conducted on R/V TE VEGA during October and November 1964. *In* G. Farquhar (ed.) Proceedings of an International Symposium on Biological Sound Scattering in the Ocean, pp. 409–452. Maury Center for Ocean Science, Washington, D.C.

Brooke, M. d. L., 2004. The food consumption of the world's seabirds. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**:S246–S248.

Burgos, J. and J. Horne, 2008. Characterization and classification of acoustically detected fish spatial distributions. *ICES Journal of Marine Science* **65**:1235.

Christensen, R., G. Duvall, and J. MW Johnson RL Ely, 1946. Stratification of sound scatterers in the sea. Tech. Rep. M397, University of California War Research Division, U.S. Navy Electronics Laboratory, San Diego, CA.

Ciannelli, L., P. Fauchald, K.-S. Chan, V. N. Agostini, and G. E. Dingsør, 2008. Spatial fisheries ecology: recent progress and future prospects. *Journal of Marine Systems* **71**:223–236.

Cisewski, B., V. Strass, M. Rhein, and S. Krägefsky, 2010. Seasonal variation of diel vertical migration of zooplankton from ADCP backscatter time series data in the Lazarev Sea, Antarctica. *Deep Sea Research Part I: Oceanographic Research Papers* **57**:78–94.

Coffin, M., L. Gahagan, and L. Lawver, 1998. Present-day plate boundary digital data compilation. Tech. Rep. 174, University of Texas Institute for Geophysics.

Cole, H., G. Bryan, and A. Gordon, 1970. The deep scattering layer: patterns across the Gulf Stream and the Sargasso Sea. *In* G. Farquhar (ed.) Proceedings of an International Symposium on Biological Sound Scattering in the Ocean, pp. 281–293. Maury Center for Ocean Science, Washington, D.C.

Conte, M., J. Bishop, and R. Backus, 1986. Nonmigratory, 12-kHz, deep scattering layers of Sargasso Sea origin in warm-core rings. *Deep-Sea Research Part A: Oceanographic Research Papers* **33**:1869–1884.

Corripio, J. G., 2013. insol: Solar Radiation. URL http://CRAN.R-project.org/package=insol. R package version 1.1.

Cox, M. J., T. B. Letessier, and A. S. Brierley, 2013. Zooplankton and micronekton biovolume at the Mid-Atlantic Ridge and Charlie-Gibbs Fracture Zone estimated by multi-frequency acoustic survey. *Deep Sea Research II* In press; http://dx.doi.org/10.1016/j.dsr2.2013.07.020.

De Robertis, A. and I. Higginbottom, 2007. A post-processing technique to estimate the signal-to-noise ratio and remove echosounder background noise. *ICES Journal of Marine Science* **64**:1282.

Demer, D., 2004. An estimate of error for the CCAMLR 2000 survey estimate of krill biomass. *Deep Sea Research II* **51**:1237–1251.

Djurhuus, A., J. F. Read, and A. D. Rogers, 2014. The spatial distribution of particulate organic carbon and microorganisms on seamounts of the south west indian ridge. *Deep Sea Research II* .

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012. The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment* **116**:140–158.

Dragon, A., P. Monestiez, A. Bar-Hen, and C. Guinet, 2010. Linking foraging behaviour to physical oceanographic structures: Southern elephant seals and mesoscale eddies east of Kerguelen Islands. *Progress in Oceanography* **87**:61–71.

Falkowski, P. and Z. Kolber, 1995. Variations in chlorophyll fluorescence yields in phytoplankton in the world oceans. *Australian Journal of Plant Physiology* **22**:341–355.

Farquhar, G., 1977. Biological sound scattering in the oceans: a review. *In* N. Andersen and B. Zahuranec (eds.) Oceanic sound scattering prediction, *Marine Science*, vol. 5, pp. 493–527. Plenum Press, New York.

Fernandes, P. G., F. Gerlotto, D. V. Holliday, O. Nakken, and E. J. Simmonds, 2002. Acoustic applications in fisheries science: the ICES contribution. *ICES Marine Science Symposia* **215**:483–492.

Fofonoff, N. and R. Millard, 1983. Algorithms for computation of fundamental properties of seawater. *Unesco Technical Papers in Marine Science* **44**:46–50.

Foote, K., H. Knudsen, G. Vestnes, D. MacLennan, and E. Simmonds, 1987. Calibration of acoustic instruments for fish density estimation: a practical guide. *ICES Cooperative Research Reports* **144**:69.

GEBCO, 2010. General Bathymetric Chart of the Oceans – The GEBCO 08 Grid. Version 20100927, http://www.gebco.net.

Genin, A., 2004. Bio-physical coupling in the formation of zooplankton and fish aggregations over abrupt topographies. *Journal of Marine Systems* **50**:3–20.

Genin, A., P. Dayton, P. Lonsdale, and F. Spiess, 1986. Corals on seamount peaks provide evidence of current acceleration over deep-sea topography. *Nature* **322**:59–61.

Godø, O., R. Patel, and G. Pedersen, 2009. Diel migration and swimbladder resonance of small fish: some implications for analyses of multifrequency echo data. *ICES Journal of Marine Science* **66**:1143.

Godø, O. R., A. Samuelsen, G. J. Macaulay, R. Patel, S. S. Hjøllo, J. Horne, S. Kaartvedt, and J. A. Johannessen, 2012. Mesoscale eddies are oases for higher trophic marine life. *PLoS ONE* **7**:e30161.

Gower, J. C., 1971. A general coefficient of similarity and some of its properties. *Biometrics* **27**:857–871.

Guinet, C., J. Vacquié-Garcia, B. Picard, G. Bessigneul, Y. Lebras, A. C. Dragon, M. Viviant, J. P. Arnould, and F. Bailleul, 2014. Southern elephant seal foraging success in relation to temperature and light conditions: insight into prey distribution. *Marine Ecology Progress Series* **499**:285–301.

Haury, L., J. McGowan, and P. Wiebe, 1978. Patterns and processes in the time-space scales of plankton distributions. *In* J. Steele (ed.) Spatial pattern in plankton communities, *NATO Conference Series in Marine Sciences*, vol. 3, pp. 277–327. Plenum Press, New York, NY.

Hazen, E. and D. Johnston, 2010. Meridional patterns in the deep scattering layers and top predator distribution in the central equatorial Pacific. *Fisheries Oceanography* **19**:427–433.

Hidaka, K., K. Kawaguchi, M. Murakami, and M. Takahashi, 2001. Downward transport of organic carbon by diel migratory micronekton in the western equatorial Pacific: its quantitative and qualitative importance. *Deep Sea Research I* **48**:1923–1939.

Hijmans, R. J. and J. van Etten, 2012. raster: Geographic data analysis and modeling. URL http://CRAN.R-project.org/package=raster. R package version 2.0-41.

Hijmans, R. J., E. Williams, and C. Vennes, 2012. geosphere: Spherical Trigonometry. URL http://CRAN.R-project.org/package=geosphere. R package version 1.2-28.

Howey-Jordan, L. A., E. J. Brooks, D. L. Abercrombie, L. K. Jordan, A. Brooks, S. Williams, E. Gospodarczyk, and D. D. Chapman, 2013. Complex movements, philopatry and expanded depth range of a severely threatened pelagic shark, the oceanic whitetip (*Carcharhinus longimanus*) in the western North Atlantic. *PLoS ONE* **8**:e56588.

Irigoien, X., T. Klevjer, A. Røstad, U. Martinez, G. Boyra, J. Acuña, A. Bode, F. Echevarria, J. Gonzalez-Gordillo, S. Hernandez-Leon, S. Agusti, D. Aksnes, C. Duarte, and S. Kaartvedt, 2014. Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nature communications* **5**.

620 Isaacs, J. and R. Schwartzlose, 1965. Migrant sound scatterers: interaction with the sea floor.
621    *Science* **150**:1810–1813.

622 Kaartvedt, S., A. Staby, and D. Aksnes, 2012. Efficient trawl avoidance by mesopelagic fishes
623    causes large underestimation of their biomass. *Marine Ecology Progress Series* **456**:1–6.

624 Kaufman, L. and P. J. Rousseeuw, 1990. Finding groups in data: an introduction to cluster analysis.
625    Wiley, New York.

626 Kemp, K. M., P. H. Boersch-Supan, O. Alvheim, D. Benivary, V. Mangar, and A. D. Rogers, 2014.
627    Pelagic fish community composition of the south west indian ocean ridge. *Deep Sea Research II*
628    .

629 Kim, S.-S. and P. Wessel, 2011. New global seamount census from altimetry-derived gravity data.
630    *Geophysical Journal International* **186**:615–631.

631 Klevjer, T. A., D. J. Torres, and S. Kaartvedt, 2012. Distribution and diel vertical movements of
632    mesopelagic scattering layers in the red sea. *Marine Biology* **159**:1833–1841.

633 Kloser, R., T. Ryan, J. Young, and M. Lewis, 2009. Acoustic observations of micronekton fish on
634    the scale of an ocean basin: potential and challenges. *ICES Journal of Marine Science* **66**:998.

635 Koslow, J., 1997. Seamounts and the ecology of deep-sea fisheries. *American Scientist* **85**:168–176.

636 Laptikovsky, V., P. Boersch-Supan, K. Kemp, T. Letessier, and A. Rogers, 2014. Cephalopods
637    of the Southwest Indian Ocean Ridge: a hotspot of extreme biological diversity and absence of
638    endemism. Unpublished manuscript.

639 Lara-Lopez, A., P. Davison, and J. Koslow, 2012. Abundance and community composition of
640    micronekton across a front off Southern California. *Journal of Plankton Research* **34**:828–848.

641 Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**:1659–1673.

Lehodey, P., R. Murtugudde, and I. Senina, 2010. Bridging the gap from ocean models to population dynamics of large marine predators: a model of mid-trophic functional groups. *Progress in Oceanography* **84**:69–84.

Lehodey, P., I. Senina, and R. Murtugudde, 2008. A spatial ecosystem and populations dynamics model (seapodym)–modeling of tuna and tuna-like populations. *Progress in Oceanography* **78**:304–318.

Letessier, T., S. De Grave, P. Boersch-Supan, K. Kemp, A. Brierley, and A. Rogers, 2014. The biogeography of pelagic shrimps (Decapoda) and Gnathophausiids (Lophogastridea) on seamounts of the South-West Indian Ocean Ridge. Unpublished manuscript.

Lewin-Koh, N. J. and R. Bivand, 2010. maptools: Tools for reading and handling spatial objects. URL http://CRAN.R-project.org/package=maptools. R package version 0.7-34.

Longhurst, A., 1998. Ecological Geography of the Sea. Academic Press, San Diego CA, USA.

Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill, 1995. An estimate of global primary production in the ocean from satellite radiometer data. *Journal of Plankton Research* **17**:1245–1271.

Lyman, J., 1947. The sea's phantom bottom. *The Scientific Monthly* **66**:87–88.

MacLennan, D. N., P. G. Fernandes, and J. Dalen, 2002. A consistent approach to definitions and symbols in fisheries acoustics. *ICES Journal of Marine Science* **59**:365–369.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik, 2012. cluster: Cluster Analysis Basics and Extensions. R package version 1.14.3.

McClatchie, S., G. Macaulay, and R. Coombs, 2004. Acoustic backscatter and copepod secondary production across the Subtropical Front to the east of New Zealand. *Journal of Geophysical Research* **109**:C03013.

McIntyre, T., I. Ansorge, H. Bornemann, J. Plötz, C. Tosh, and M. Bester, 2011. Elephant seal dive behaviour is influenced by ocean temperature: implications for climate change impacts on an ocean predator. *Marine Ecology Progress Series* **441**:257–272.

Meeus, J., 1998. Astronomical algorithms. Willmann-Bell, Richmond, Va.

Morato, T., C. Bulman, and T. Pitcher, 2009. Modelled effects of primary and secondary production enhancement by seamounts on local fish stocks. *Deep Sea Research II* **56**:2713–2719.

Morato, T., D. Varkey, C. Damaso, M. Machete, M. Santos, R. Prieto, R. Santos, and T. Pitcher, 2008. Evidence of a seamount effect on aggregating visitors. *Marine Ecology Progress Series* **357**:23–32.

Murase, H., T. Kitakado, T. Hakamada, K. Matsuoka, S. Nishiwaki, and M. Naganobu, 2013. Spatial distribution of Antarctic minke whales (*Balaenoptera bonaerensis*) in relation to spatial distributions of krill in the Ross Sea, Antarctica. *Fisheries Oceanography* **22**:154–173.

Nagpaul, P., 1999. Guide to advanced data analysis using IDAMS software. UNESCO, Paris.

Naito, Y., D. P. Costa, T. Adachi, P. W. Robinson, M. Fowler, and A. Takahashi, 2013. Unravelling the mysteries of a mesopelagic diet: a large apex predator specializes on small prey. *Functional Ecology* **27**:710–717.

Neteler, M., M. Bowman, M. Landa, and M. Metz, 2011. GRASS GIS: a multi-purpose open source GIS. *Environmental Modelling & Software* **31**:124–130.

O'Driscoll, R., S. Gauthier, and J. Devine, 2009. Acoustic estimates of mesopelagic fish: as clear as day and night? *ICES Journal of Marine Science* **66**:1310–1317.

Ohman, M., D. Rudnick, A. Chekalyuk, R. Davis, R. Feely, M. Kahru, H.-J. Kim, M. Landry, T. Martz, C. Sabine, and U. Send, 2013. Autonomous ocean measurements in the california current ecosystem. *Oceanography* **26**:18–25.

688  Opdal, A., O. Godø, O. Bergstad, and Ø. Fiksen, 2008. Distribution, identity, and possible processes
689     sustaining meso-and bathypelagic scattering layers on the northern Mid-Atlantic Ridge. *Deep-*
690     *Sea Research Part II* **55**:45–58.

691  Pakhomov, E., R. Perissinotto, and C. McQuaid, 1994. Comparative structure of the macrozo-
692     oplankton / micronekton communities of the Subtropical and Antarctic Polar Fronts. *Marine*
693     *Ecology Progress Series* **111**:155–169.

694  Pakhomov, E. A. and P. W. Froneman, 2000. Composition and spatial variability of macroplankton
695     and micronekton within the antarctic polar frontal zone of the indian ocean during austral autumn
696     1997. *Polar Biology* **23**:410–419.

697  Pearre, S., 2003. Eat and run? The hunger/satiation hypothesis in vertical migration: history,
698     evidence and consequences. *Biological Reviews* **78**:1–79.

699  Petitgas, P., 2001. Geostatistics in fisheries survey design and stock assessment: models, variances
700     and applications. *Fish and Fisheries* **2**:231–249.

701  Pinheiro, J. and D. Bates, 2000. Mixed-effects models in S and S-PLUS. Springer, New York, NY
702     u.a.

703  Pitcher, T., M. Clark, T. Morato, and R. Watson, 2010. Seamount fisheries: Do they have a future.
704     *Oceanography* **23**:134–144.

705  Potier, M., F. Marsac, Y. Cherel, V. Lucas, R. Sabatié, O. Maury, and F. Ménard, 2007. Forage fauna
706     in the diet of three large pelagic fishes (lancetfish, swordfish and yellowfin tuna) in the western
707     equatorial Indian Ocean. *Fisheries Research* **83**:60–72.

708  Potier, M., F. Marsac, V. Lucas, R. Sabatié, J. Hallier, and F. Ménard, 2005. Feeding partitioning
709     among tuna taken in surface and mid-water layers: The case of yellowfin (*Thunnus albacares*)
710     and bigeye (*T. obesus*) in the western tropical Indian Ocean. *Western Indian Ocean Journal of*
711     *Marine Science* **3**:51–62.

Priede, I. G., O. A. Bergstad, P. I. Miller, M. Vecchione, A. Gebruk, T. Falkenhaug, D. S. Billett, J. Craig, A. C. Dale, M. A. Shields, et al., 2013. Does presence of a mid-ocean ridge enhance biomass and biodiversity? *PLoS ONE* **8**:e61550.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Rivoirard, J., J. Simmonds, K. Foote, P. Fernandes, and N. Bez, 2000. Geostatistics for estimating fish abundance. Wiley-Blackwell, Oxford, UK.

Rousseeuw, P. J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**:53–65.

Rowden, A. A., T. A. Schlacher, A. Williams, M. R. Clark, R. Stewart, F. Althaus, D. A. Bowden, M. Consalvey, W. Robinson, and J. Dowdney, 2010. A test of the seamount oasis hypothesis: seamounts support higher epibenthic megafaunal biomass than adjacent slopes. *Marine Ecology* **31**:95–106.

Ryan, T., 2011. Overview of data collection, management and processing procedures of underway acoustic data - IMOS BASOOP sub-facility. CSIRO, Division of Marine and Atmospheric Research - Hobart, Australia. Version 1.0.

Ryan, T. and R. Kloser, 2004. Improving the precision of ES60 and EK60 echosounder applications. Tech. Rep. CM 2004/B:06, ICES Working Group on Fisheries Acoustic Science and Technology.

Steele, J., 1978. Some Comments on Plankton Patches. *In* J. Steele (ed.) Spatial pattern in plankton communities, *NATO Conference Series in Marine Sciences*, vol. 3, pp. 1–20. Plenum Press, New York, NY.

Steinberg, D. K., B. A. V. Mooy, K. O. Buesseler, P. W. Boyd, T. Kobari, and D. M. Karl, 2008. Bacterial vs. zooplankton control of sinking particle flux in the ocean's twilight zone. *Limnology and Oceanography* **53**:1327.

Thiebot, J.-B., Y. Cherel, P. N. Trathan, and C.-A. Bost, 2011. Inter-population segregation in the wintering areas of macaroni penguins. *Marine Ecology Progress Series* **421**:279–290.

Tont, S., 1976. Deep scattering layers: patterns in the Pacific. *CalCOFI Reports* **18**:112–117.

Tseytlin, V., 1985. The energetics of fish populations inhabiting seamounts. *Oceanology* **25**:237–239.

Urmy, S., J. Horne, and D. Barbee, 2012. Measuring the vertical distributional variability of pelagic fauna in Monterey Bay. *ICES Journal of Marine Science* **69**:184–196.

Vierros, M., I. Cresswell, E. E. Briones, J. Rice, and J. Ardron (eds.) 2009. Global Open Oceans and Deep Seabed (GOODS) - Biogeographic Classification, *IOC Technical Series*, vol. 84. UNESCO-IOC, Paris, France.

Wood, S., 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall, London, UK.

Yasuma, H., K. Sawada, T. Ohshima, K. Miyashita, and I. Aoki, 2003. Target strength of meso-pelagic lanternfishes (family Myctophidae) based on swimbladder morphology. *ICES Journal of Marine Science* **60**:584–591.

Yasuma, H., K. Sawada, Y. Takao, K. Miyashita, and I. Aoki, 2010. Swimbladder condition and target strength of myctophid fish in the temperate zone of the Northwest Pacific. *ICES Journal of Marine Science* **67**:135–144.

Yesson, C., M. R. Clark, M. L. Taylor, and A. D. Rogers, 2011. The global distribution of seamounts based on 30-second bathymetry data. *Deep Sea Research I* **58**:442–453.

Zender, C. S., 2008. Analysis of self-describing gridded geoscience data with netCDF Operators (NCO). *Environmental Modelling & Software* **23**:1338–1342.

Zwolinski, J., P. G. Fernandes, V. Marques, and Y. Stratoudakis, 2009. Estimating fish abundance from acoustic surveys: calculating variance due to acoustic backscatter and length distribution error. *Canadian Journal of Fisheries and Aquatic Sciences* **66**:2081–2095.
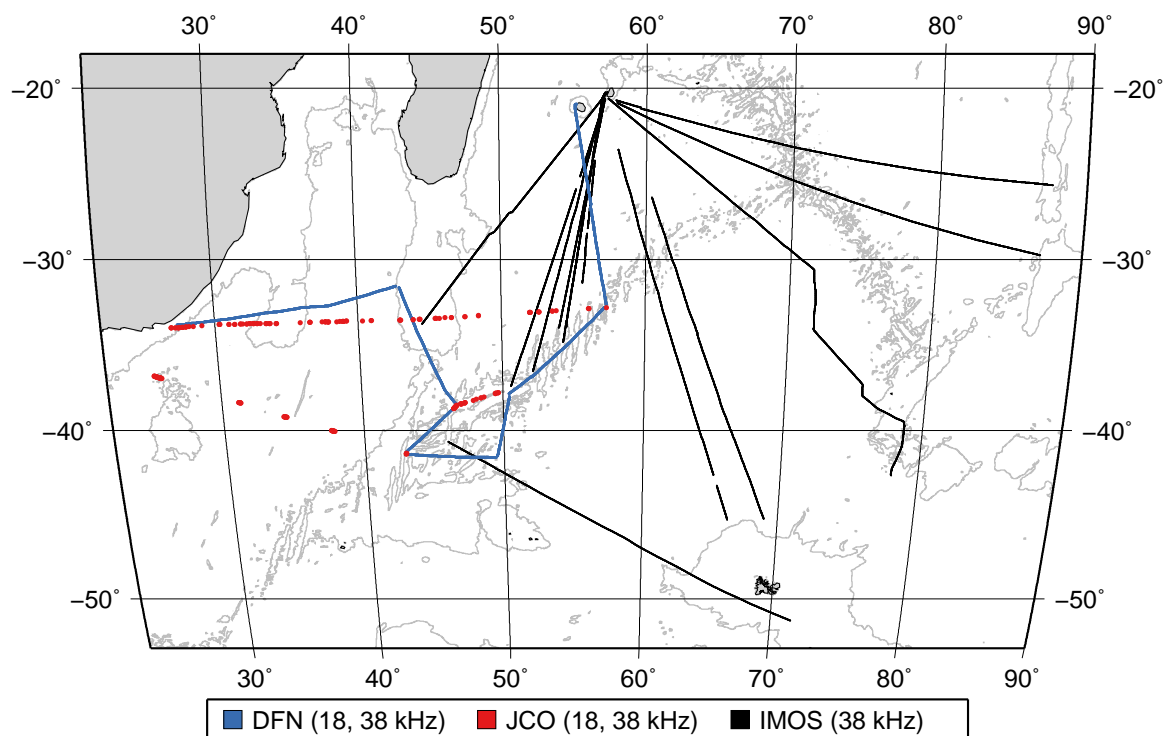
Figure 1: Sampling locations, frequencies and sources of acoustic data used in this study. Colours encode the project responsible for sampling. 3000m isobath traced from GEBCO (2010). DFN: *RV Dr. Fridtjof Nansen* Cruise 2009-410; JCO: *RRS James Cook* Cruise 66/67; IMOS: Australian Integrated Marine Observing System, Ships of Opportunity Bio-Acoustic sub-facility.
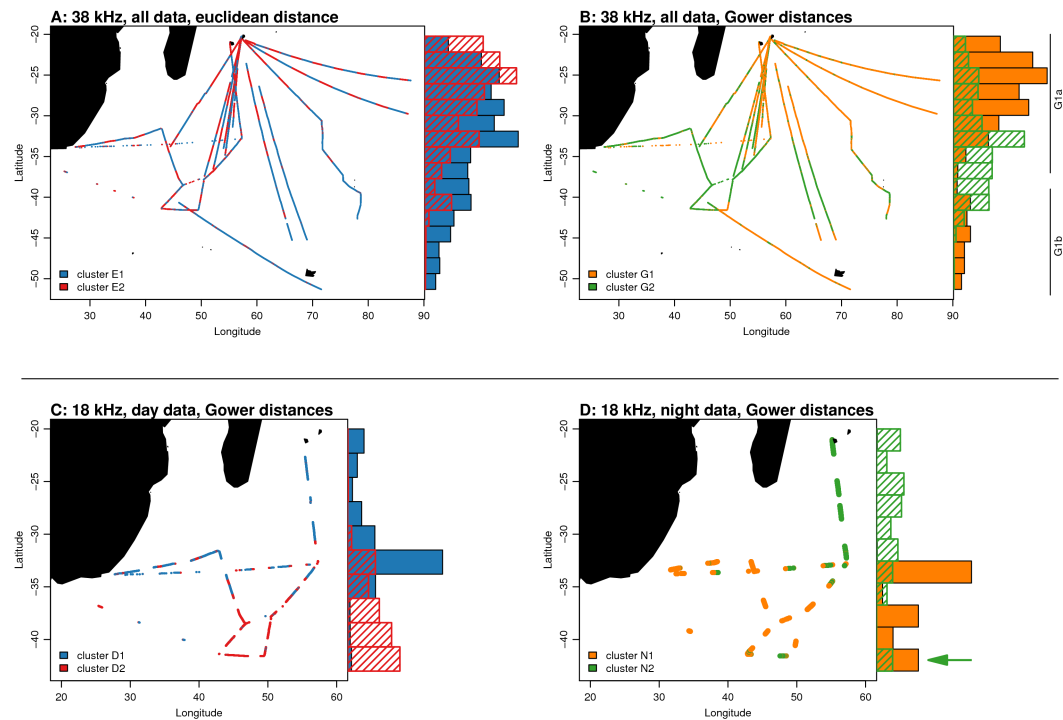
Figure 2: Maps of cluster allocations for the 38 kHz and 18 kHz datasets. Marginal histograms visualize the latitudinal composition of each cluster. In all cases $k = 2$ was the optimal partition. A: Euclidean distances partition the 38 kHz $S_v$ profiles into a daytime and a nighttime cluster (cf. Fig. S3). B: Gower's distance metric partitions the data into two clusters, of which G1 is geographically discontinuous and can be separated into two sub-clusters G1a and G1b. C: The partitioning of daytime 18 kHz $S_v$ profiles based on Gower's metric splits the dataset into a northern and southern cluster. D: The partition of nighttime $S_v$ profiles largely corresponds to that of the daytime data. The green arrow highlights a transect section near 42°S 48°E classified as belonging to the northern cluster N2.
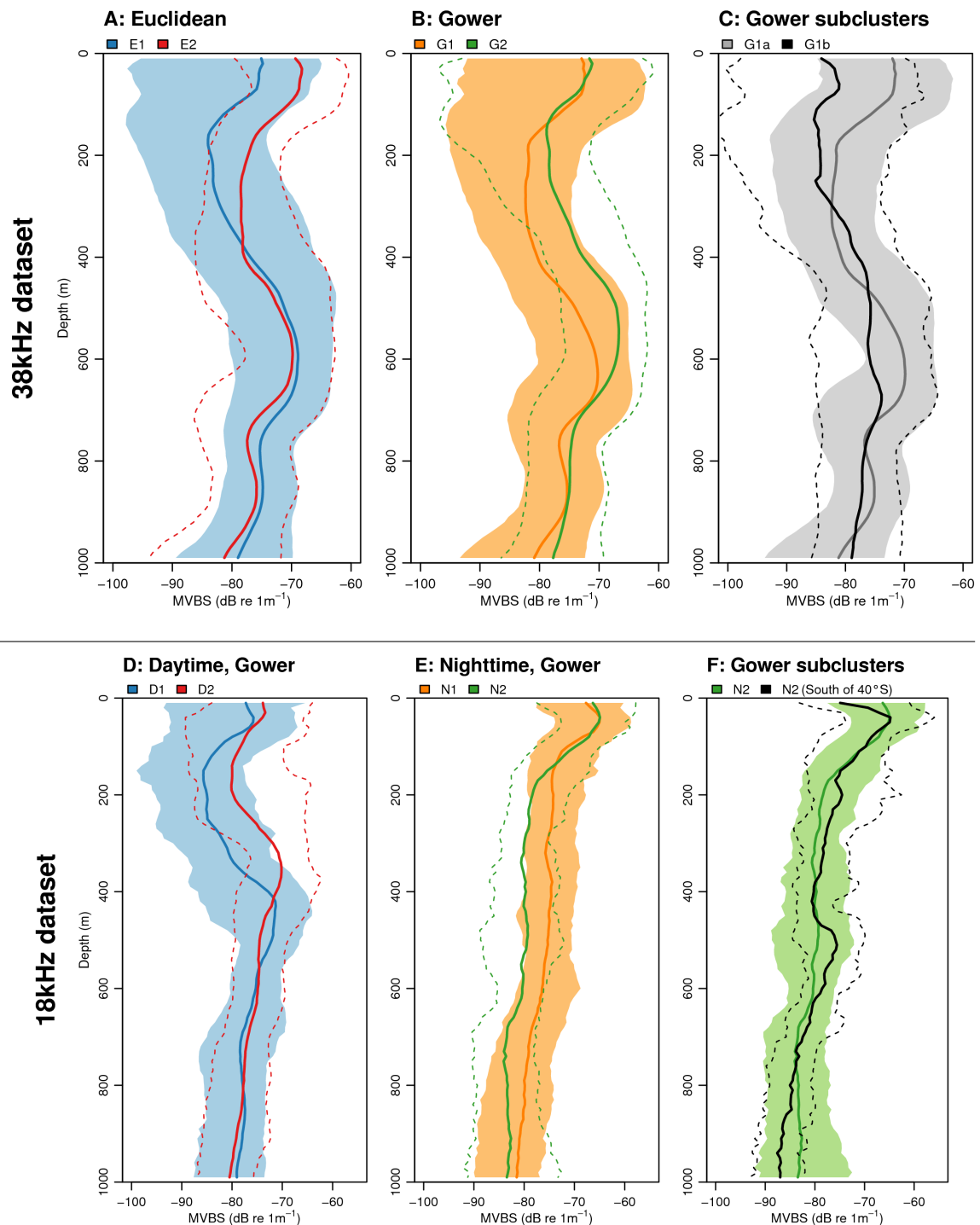
Figure 3: Vertical $S_v$ profiles ($2.5^{th}$, $50^{th}$ and $97.5^{th}$ percentile) for the clusters mapped in Figure 2. A: Day/night split achieved by euclidean distances. B: Geographic split achieved by the Gower distance metric. C: Geographic subclusters derived from G1. D: Geographic split achieved within the 18 kHz daytime data. E: Geographic split achieved within the 18 kHz nighttime data. D: Subcluster derived from N2.

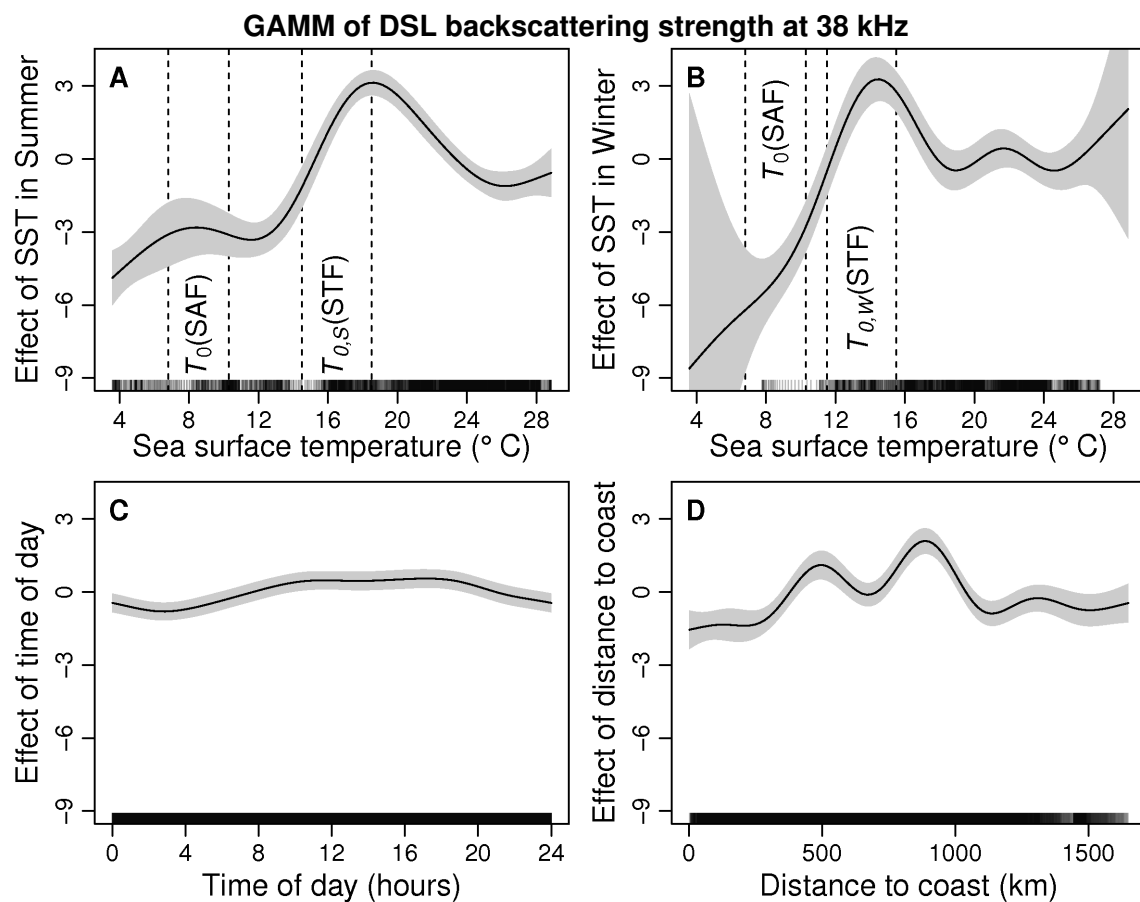**GAMM of DSL backscattering strength at 38 kHz**



Figure 4: Smooths of generalized additive model terms showing the effect of various continuous variables on DSL backscattering strength at 38 kHz. The solid lines are the estimates of the smooths, the shaded areas are standard errors of the estimated smooths, taking into account the error in the model intercept. Dashed vertical lines in panel A and B indicate the axial temperature ranges $T_0$ of the subantarctic (SAF) and subtropical fronts (STF), respectively (Belkin and Gordon, 1996). Subscripts $S$ and $W$, denote STF Summer and Winter $T_0$ ranges, respectively.
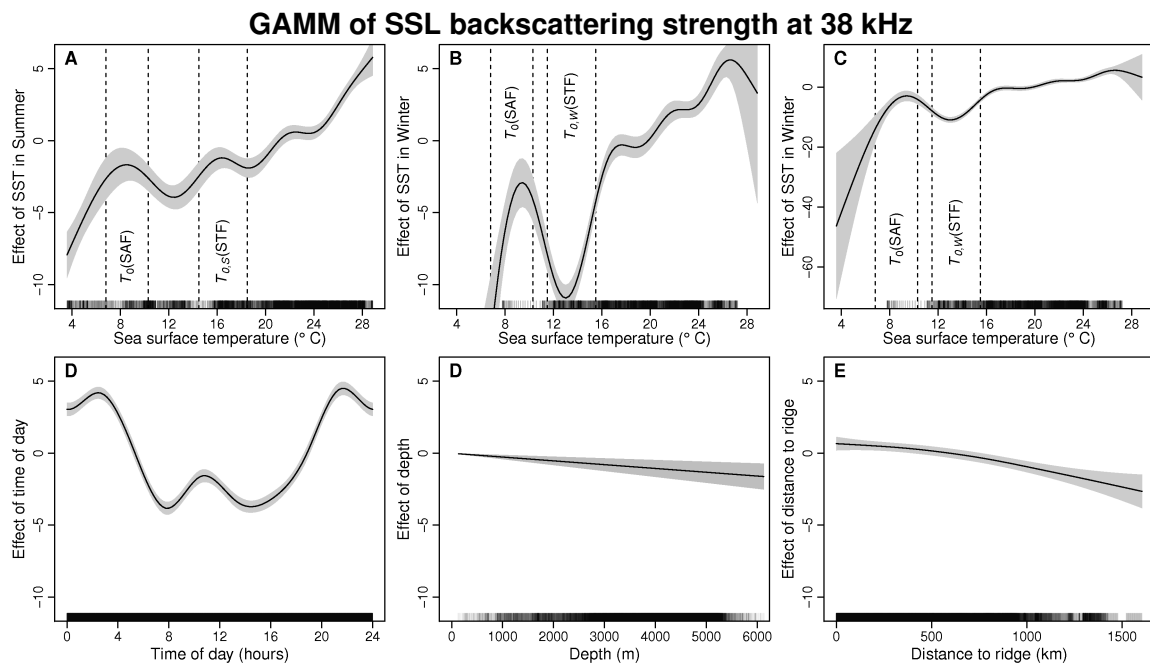
Figure 5: Smooths of generalized additive model terms showing the effect of various continuous variables on SSL backscattering strength at 38 kHz. The solid lines are the estimates of the smooths, the shaded areas are standard errors of the estimated smooths, taking into account the error in the model intercept. Panels B and C show the same smooth term, but with different y-axis scales. Dashed vertical lines in panel A and B indicate the axial temperature ranges $T_0$ of the subantarctic (SAF) and subtropical fronts (STF), respectively (Belkin and Gordon, 1996). Subscripts $S$ and $W$, denote STF Summer and Winter $T_0$ ranges, respectively.

**GAMM of DSL backscattering strength at 18 kHz**



Figure 6: Smooths of generalized additive model terms showing the effect of various continuous variables on DSL backscattering strength at 18 kHz. The solid lines are the estimates of the smooths, the shaded areas are standard errors of the estimated smooths, taking into account the error in the model intercept. The y-axis of panel D is shifted relative to the remaining panels to accommodate the range of the effect, but the scale is not altered. Dashed vertical lines in panel A and B indicate the axial temperature ranges $T_0$ of the subantarctic (SAF) and subtropical fronts (STF), respectively (Belkin and Gordon, 1996).
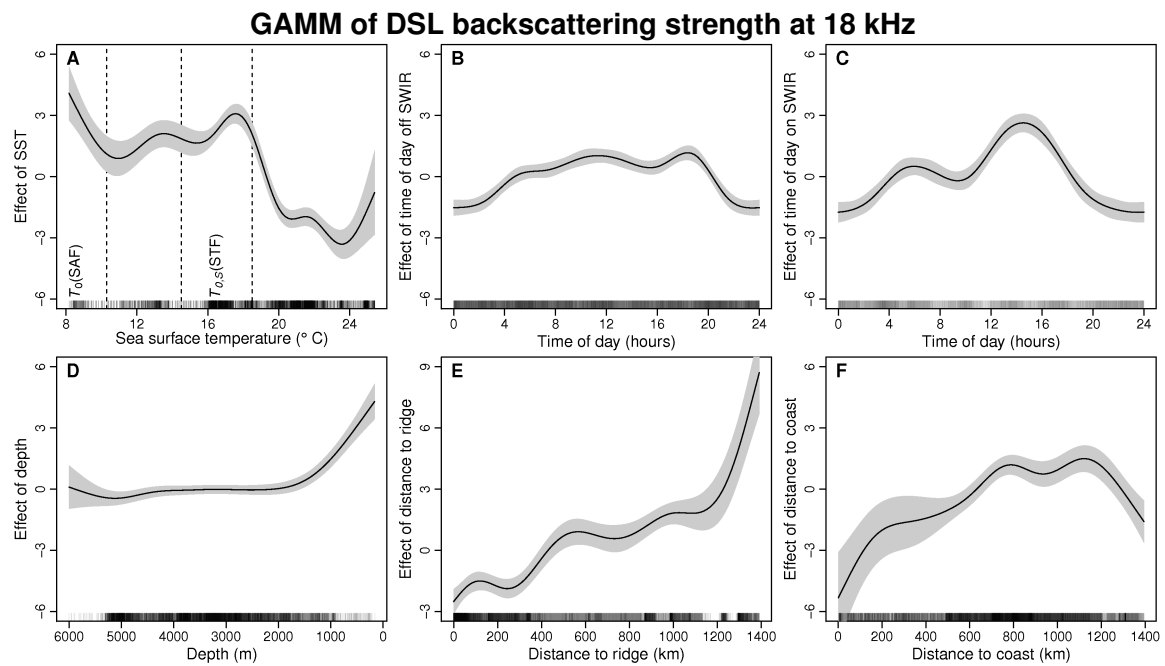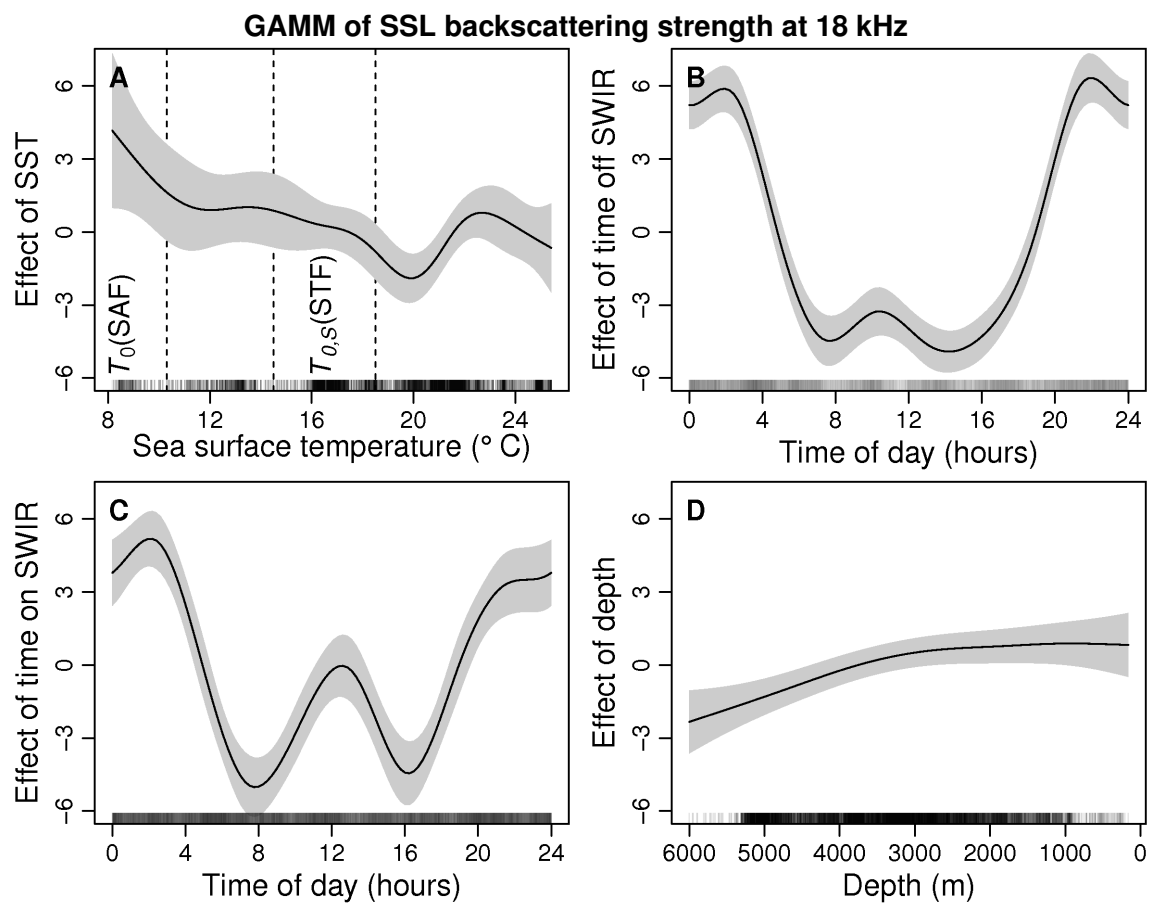
Figure 7: Smooths of generalized additive model terms showing the effect of various continuous variables on SSL backscattering strength at 18 kHz. The solid lines are the estimates of the smooths, the shaded areas are standard errors of the estimated smooths, taking into account the error in the model intercept. Dashed vertical lines in panel A and B indicate the axial temperature ranges $T_0$ of the subantarctic (SAF) and subtropical fronts (STF), respectively (Belkin and Gordon, 1996).

Table 1: Breakdown of along-track sampling units of 1 km length retained for analysis by data source, frequency, time of day and season. Numbers in brackets give the numbers of individual calender days for which data were collected.

| | | 38 kHz | | | | 18 kHz | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | IMOS | JCO | DFN | Σ | JCO | DFN | Σ |
| Time of day | Day | 10335 | 93 | 2641 | 13069 | 150 | 2641 | 2791 |
| | Night | 9980 | 70 | 1811 | 11861 | 91 | 1811 | 1902 |
| | Twilight | 6922 | 45 | 1509 | 8476 | 93 | 1509 | 1602 |
| | Σ | 27237 | 208 | 5961 | **33406** | 334 | 5961 | **6295** |
| Season | Austral Summer | 14795 | 208 | 5961 | 20964 | 334 | 5961 | 6295 |
| | | (38) | (16) | (26) | (80) | (16) | (26) | (42) |
| | Austral Winter | 12442 | 0 | 0 | 12442 | 0 | 0 | 0 |
| | | (31) | (0) | (0) | (31) | (0) | (0) | (0) |
| | Σ | 27237 | 208 | 5961 | **33406** | 334 | 5961 | **6295** |
| | | (69) | (16) | (26) | **(111)** | (16) | (26) | **(42)** |

Table 2: Explanatory variables and corresponding parameters for the GAMMs used to model mean volume backscatter $S_V$ in the deep and shallow scattering layers.

| | Dependent variable | | | |
| | 38 kHz | | 18 kHz | |
| | $S_v^{\text{DSL38}}$ | $S_v^{\text{SSL38}}$ | $S_v^{\text{DSL18}}$ | $S_v^{\text{SSL18}}$ |
|---|---|---|---|---|
| **Parametric terms[a]** | | | | |
| (Intercept) | $-76.918^{***}$ | $-73.031^{***}$ | $-75.587^{***}$ | $-73.407^{***}$ |
| | (0.154) | (0.306) | (0.0565) | (0.2691) |
| Depth | | $2.64 \times 10^{-4**}$ | | |
| | | $(7.33 \times 10^{-5})$ | | |
| **Smooth terms[b]** | | | | |
| s(SST):Summer | $7.999^{***}$ | $8.195^{***}$ | $8.565^{***}$ | $6.422^{**}$ |
| | (85.59) | (52.38) | (48.70) | (3.665) |
| s(SST):Winter[d] | $6.940^{***}$ | $8.329^{***}$ | | |
| | (24.28) | (95.48) | | |
| s(Time of day) | $5.905^{***}$ | $7.873^{***}$ | | |
| | (17.75) | (240.41) | | |
| s(Time of day):OnSWIR | | | $7.168^{***}$ | $7.610^{***}$ |
| | | | (19.66) | (75.20) |
| s(Time of day):OffSWIR | | | $6.699^{***}$ | $7.320^{***}$ |
| | | | (24.77) | (28.70) |
| s(Depth) | | | $6.743^{***}$ | $2.825^{***}$ |
| | | | (17.07) | (9.717) |
| s(DistRidge) | | $1.929^{***}$ | $8.023^{***}$ | |
| | | (24.61) | (22.00) | |
| s(DistCoast) | $8.735^{***}$ | | $7.589^{***}$ | |
| | (22.20) | | (12.87) | |
| Terms excluded from all models during model selection | | | | |
| Chlorophyll | | | | |
| **AR(1) correlation coefficient[c]** | | | | |
| $\phi$ | 0.985 | 0.936 | 0.815 | 0.943 |
| | (0.983,0.987) | (0.931,0.940) | (0.792,0.836) | (0.933,0.952) |
| **Other measures** | | | | |
| AIC | 38 418 | 100 150 | 15 380 | 20 389 |
| $r^2_{adj}$ | 0.356 | 0.663 | 0.732 | 0.545 |
| Num. obs. | 32 456 | 32 456 | 6 166 | 6 166 |

*** $p < 0.0001$, ** $p < 0.001$, * $p < 0.01$

[a] parameter estimate (std. error)

[b] effective degrees of freedom ($F$)

[c] parameter estimate (95% CI)

[d] 18 kHz data were only collected during Austral Summer

# 1 Supplementary Methods

## 2 Gower's similarity coefficient

3 The similarity coefficient described by Gower (1971) measures the similarity between two sites $i$

4 and $j$ as a weighted average of similarities for all depth strata that are available for both sites

$$S_{ij} = \sum_{k=1}^{n} s_{ijk} \delta_{ijk} / \sum_{k=1}^{n} \delta_{ijk} \tag{3}$$

5 where $n$ is the number of depth strata, $s_{ijk}$ is the similarity between $i$ and $j$ calculated on the $k$th

6 depth stratum and $\delta_{ijk}$ is equal to 0 if the value of the $k$th depth stratum is missing for one of the sites

7 and 1 it is present at both sites. For each stratum $k$ comprising $n$ sites with $S_v$ values $x_1, x_2, \ldots, x_n$

8 the inter-site difference is calculated as

$$s_{ijk} = 1 - \frac{|x_i - x_j|}{R_k} \tag{4}$$

9  where $R_k$ is the range of $S_v$ values in the stratum.

## 10 Goodness of clustering

Goodness of clustering solutions for different prespecified numbers of clusters $k$ was assessed using the overall average silhouette width ($\bar{s}(k)$; Rousseeuw, 1987). This defined as follows: Assume $n$ observations have been clustered into $k$ clusters. For each observation $n_i$, let $a(i)$ be the average dissimilarity between $n_i$ and all other points of the cluster $A$ to which $n_i$ belongs[1]. $a(i)$ can be interpreted as a measure of how well $i$ is assigned to its cluster (the smaller the value, the better the assignment). For all other clusters $C$, let $d(i, C)$ be the average dissimilarity of $n_i$ to all observations in $C$. The smallest of these $d(i, C)$ is

$$b(i) := \min_{C \neq A} d(i, C),$$

---

[1] if $n_i$ is the only observation in its cluster, $s(i) := 0$ without further calculations.

11    which can be seen as the dissimilarity between $n_i$ and the nearest cluster to which it does not belong.

12    Finally,

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$    (5)

13    Observations with a large $s(i)$ (almost 1) are very well clustered, a small $s(i)$ (around 0) means

14    that the observation lies between two clusters, and observations with a negative $s(i)$ are probably

15    placed in the wrong cluster. Thus

$$\bar{s}(k) = \frac{\sum_{1}^{n} s(i)}{n}$$    (6)

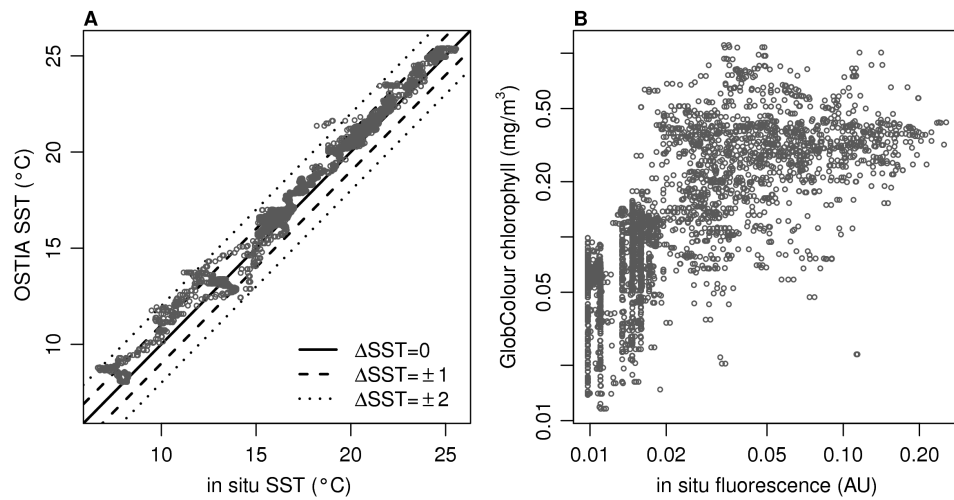16    is a measure of how appropriately the data has been clustered.

Figure S1: Correlations between *in situ* surface measurements and gridded products. **A**: *In situ* sea surface temperature (SST) versus OSTIA SST (Donlon et al., 2012); Pearson's $r$=0.99. **B**: Log-log plot of *in situ* fluorescence and GlobColour chlorophyll concentration (Anon., 2010); Pearson's $r$=0.53
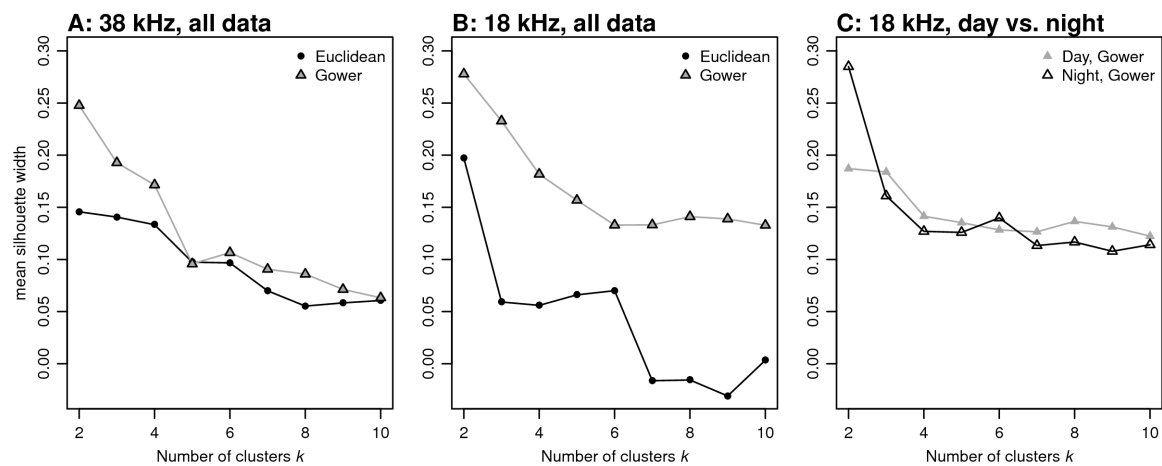
.



Figure S2: Goodness of clustering for different numbers of clusters $k$ and different distance metrics. A: Clustering of all 38 kHz $S_v$ profiles using different distance metrics. B: Clustering of all 18 kHz $S_v$ profiles using different distance metrics. C: Separate analyses for daytime and nighttime 18 kHz $S_v$ profiles based on Gower's metric.
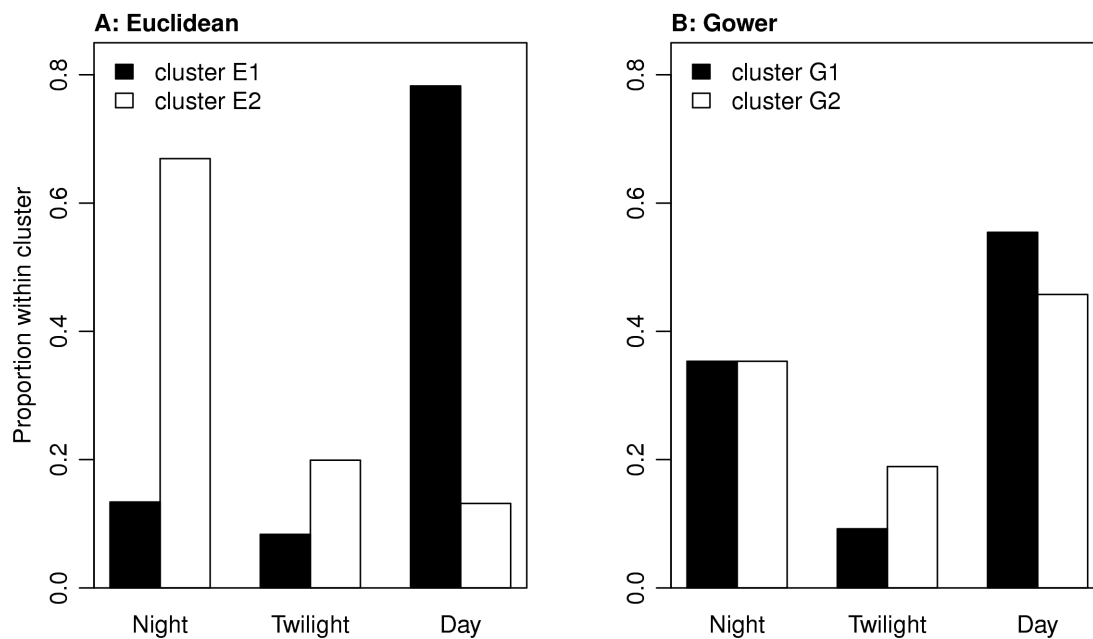
Figure S3: Compositions of 38 kHz clusters by diel phase for $k = 2$ based on two different dissimilarity measures.

Table S1: Transducer parameters and specifications for the EK60 echosounders employed in this study. All calibrations were conducted following the recommendations of Foote et al. (1987).

|  | RV Dr. F. Nansen | | RRS James Cook | |
| --- | --- | --- | --- | --- |
|  | 18kHz | 38kHz | 18kHz | 38kHz |
| Transducer model | ES18-11 | ES38-B | ES18-11 | ES38-B |
| Serial number | 593 | 489 | 2067 | 30637 |
| Absorption coefficient (dB/km) | 2.2 | 8.5 | 2.2 | 8.5 |
| Transmitted power (W) | 1000 | 2000 | 2000 | 2000 |
| Two-way beam angle (dB re 1 sr) | -17 | -20.6 | -17.3 | -21.0 |
| Transducer gain (dB) | 20.76 | 25.82 | 23.22 | 23.87 |
| $S_A$ correction (dB) | -0.62 | -0.53 | -1.19 | -0.6 |
| Transmitted pulse length ($\mu s$) | 1024 | 1024 | 1024 | 1024 |
| Alongship offset angle (degrees) | 0.08 | 0.11 | -0.07 | -0.3 |
| Athwartship offset angle (degrees) | 0.02 | 0.05 | 0.06 | -0.6 |
| Minor-axis 3dB beam angle (degrees) | 11.06 | 7.05 | 11.21 | 7.06 |
| Major-axis 3dB beam angle (degrees) | 10.98 | 7.06 | 11.26 | 7.22 |
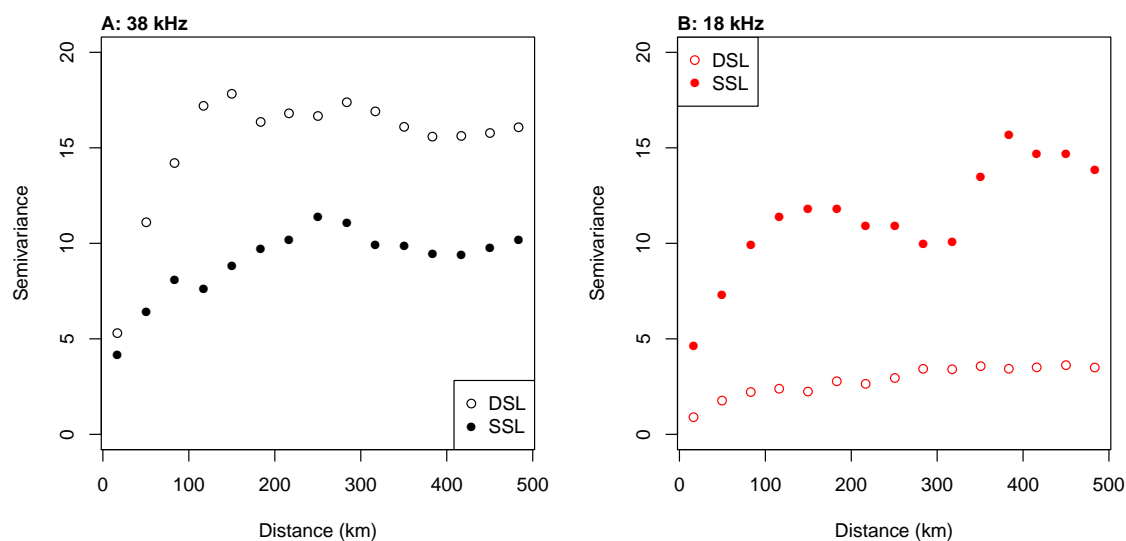
Figure S4: Sample variograms of SSL and DSL data. Semivariance increases up to a range of approx. 100 km.