



Distribution, density and relative abundance of Antarctic krill estimated by maximum likelihood geostatistics on acoustic data collected during commercial fishing operations



E.J. Niklitschek^{a,*}, G. Skaret^b

^a Universidad de Los Lagos, Centro i-mar, Camino a Chiquihue Km. 6, Puerto Montt 5502764, Region de Los Lagos, Chile

^b Institute of Marine Research, PO Box 1870 Nordnes, NO-5817 Bergen, Norway

ARTICLE INFO

Article history:

Received 15 March 2015

Received in revised form 9 September 2015

Accepted 22 September 2015

Available online 9 October 2015

Keywords:

Euphausia superba

Acoustics

Surveys

Geostatistics

Fishing vessels

ABSTRACT

There is a substantial harvest for Antarctic krill in the Southern Ocean, but little regular scientific monitoring of the resource. Recently, however, the Commission for the Conservation of Marine Living Resources (CCAMLR) has initialised a process to make use of acoustic data from commercial fisheries to increase the amount of relevant information available for making management decisions. We here provide an example where 34 days of acoustic data, collected during commercial krill fishing operations on the vessel 'Saga Sea' were processed to produce probability of presence, conditional density and relative abundance estimates on monthly, weekly and daily basis. Data were analyzed using a maximum likelihood time-series and geostatistical approaches, selected to account for the lack of sampling design, and likely correlation in space and time. The applied method showed low sensitivity of monthly estimates to different repeated measure criteria and location sub-settings. Most weekly estimates, but the last one, were also consistent with the full data (monthly) estimate. Highly variable and lower estimates were obtained, however, from daily data sets. Although our results suggest the method had provided an adequate treatment for time and space correlation, we were not able to evaluate potential bias due to preferential sampling of high density krill aggregations and/or limited area coverage within short time periods. The results suggest that this method, combined with some additional design based coverage by the fishing vessels, can be useful to obtain quantitative evaluations of krill density and distribution for management purposes.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Proper evaluations of distribution, abundance and/or biomass of exploited marine organisms are often keys to a successful resource management. Typically, such evaluations are based on scientific surveys characterized by the use of full-time dedicated vessels and pre-planned sampling designs (Gunderson, 1993). In some areas, however, routine scientific monitoring is limited by factors such as high costs, low accessibility and the mere size of the monitoring area of interest. The Southern Ocean is an example of such an area.

The Southern Ocean is one of the remotest fishing areas, but contains one of the most abundant marine resources on earth, namely the Antarctic krill (*Euphausia superba*), hereafter krill. The krill is exploited commercially, and at present the total annual krill catch is in the range of 300,000 t per year. Even though the krill has a

circumpolar distribution, the fishery is concentrated to the Scotia Sea, and above all to small regions at the shelf breaks of the South Georgia Islands, the South Orkney Islands and the west coast of the Antarctic peninsula (Atkinson et al., 2008; Everson, 2008). The harvesting is managed by The Commission for the Conservation of Marine Living Resources (CCAMLR) which regulates all exploitation of marine resources in the Antarctic waters. The precautionary catch level for krill based on a synoptic acoustic survey estimate from 2000 is at 5.61 million t (CCAMLR, 2010). Although the regulation of the krill harvest is carried out according to the explicit aim that the harvest shall not negatively impact krill-dependent predators (Hewitt et al., 2002), there is presently not enough information available to properly evaluate the effect of the fisheries on the predators, at the scale the fishery occurs. As a consequence, the trigger level that instigates management actions, which presently sits at 620,000 is static and based on historical catches. For several years, however, CCAMLR has been working towards a feedback management system, aiming for a more flexible resource management, where up-to-date information about the state of the krill and

* Corresponding author.

E-mail address: edwin.niklitschek@ulagos.cl (E.J. Niklitschek).

krill dependent predators can be used as a basis for management decisions (Constable, 2011; Constable et al., 2000).

As part of the feedback management approach, alternative cost-effective methods to obtain information on krill are being evaluated. One approach is to obtain information directly from the fishing vessels, which operate in the Southern Ocean waters through most of the year. As part of this approach, the CCAMLR Subgroup on Acoustic Survey and Analysis Methods (SG-ASAM) has been working with a proof of concept for using acoustic data from krill fishing vessels to inform managers (Watkins et al., 2016). Such data are typically of a poorer quality than data from scientific vessels, and even though commercial vessels may be required to carry out systematic surveys, most of the data collection will not follow any sampling design. Moreover, data tends to be clustered (Menezes et al., 2008) and preferentially collected (Diggle et al., 2010) as commercial vessels use historical records, personal knowledge and technological means to find, follow and fish the densest krill aggregations, producing observations that are very intense and repetitive in time, but highly irregular and limited in spatial coverage. Thus, underlying assumptions implicit in design-based estimation methods, such as those used by Jolly and Hampton (1990) or Bez (2002) are not valid. Even more liberal assumptions, such as the existence of random searching patterns (Aubry and Debouzie, 2000) are most probably violated.

A possible approach, suitable to face time and spatial correlation, and lack of sampling design problems, is the utilization of model-based time-series and geostatistical methods, which do not require of probabilistic sampling designs (Aubry and Debouzie, 2000; Diggle and Ribeiro, 2007). Several of these model-based methods exist, including a number of non-parametric and parametric ones, whose application may present relative advantages or disadvantages, depending on the data (Cressie, 1993; Henley, 2012). Within parametric methods, likelihood-based geostatistics (Diggle and Ribeiro, 2007; Roa-Ureta and Niklitschek, 2007) present several properties that are highly relevant for the analysis of acoustic data collected from commercial fishing operations: (i) flexibility to accommodate additional sources of sampling correlation, e.g. repeated acoustic surveys carried out over the same stock and/or data collected from several vessels operating simultaneously; (ii) flexibility to accommodate non-Gaussian distribution functions for the regionalized variable; and (iii) straightforward methods to calculate measures of precision (for geostatistical parameters) and model selection criteria, such as Akaike's (1973) Information Criterion, from likelihood profiles.

Within the likelihood-based geostatistical framework, Roa-Ureta & Niklitschek (2007) presented a method designed explicitly for fishery resource surveys. It uses a delta approach to analyse zero-inflated data (Aitchison and Brown, 1957; Lambert, 1992; Lo et al., 1992; Pennington, 1983), where distribution (presence/absence) and conditional density are treated as two independent (and multiplicative) responses. Presence/absence is modelled as a binomial process, while conditional density (i.e. positive values, $NASC > 0$) is treated as a realization of a Gaussian, gamma, poisson or any other process belonging to the exponential family. This delta likelihood-based geostatistical method has been used to evaluate distribution and abundance of several fish and shellfish stocks, using data collected through repeated surveys and/or by multiple scientific or commercial vessels operating simultaneously (Arkhipkin et al., 2013; Molinet et al., 2010, 2015; Niklitschek and Roa, 2006; Roa-Ureta and Niklitschek, 2007).

A practical issue, that may affect the magnitude of preferential sampling effects upon abundance estimates based on data from sonar-assisted commercial operations, is the time interval chosen to discretise the total period of observation. As the length of this interval increases, locations tend to be sampled only when the stock is present, but avoided otherwise, leading to overestimating the

mean and to underestimating the variance of the binomial process. As the length of sampling time-interval decreases, a larger proportion of locations will not be sampled, but extrapolated using mean values from observed locations

In the present paper, we apply a delta likelihood geostatistical method to an existing data set containing 34 days of acoustic data collected by a krill fishing vessel around the South Orkney's Islands, and use it to investigate the sensitivity and precision of abundance estimates under different sub-sampling strategies and inference time intervals.

2. Material and methods

2.1. Vessel and area of operation

The F/V 'Saga Sea' is one out of typically 10–12 vessels fishing krill in Antarctic waters. It is owned by the Norwegian company Aker ASA and operates in the CCAMLR statistical areas 48.1, 48.2 and 48.3, usually from December to August. The vessel is equipped with a continuous trawling system by which the krill is continuously pumped on board from the cod end through rigid hoses. During operation, a trawl haul at ca. 2 knots speed can last for days or even weeks in a stretch within limited areas.

2.2. Acquisition of acoustic data

The possibilities, utilities and limitations of collecting acoustic data from krill fishing vessels have been investigated in CCAMLR through a proof of concept process, which is described in detail in Watkins et al. (2016), SC-CCAMLR (2012) and SC-CCAMLR (2014). In the specific case of the 'Saga Sea', the vessel was equipped with a Simrad ES60 echo sounder system running the two frequencies 38 and 120 kHz. As part of an agreement between Aker ASA and scientists at the Institute of Marine Research, Norway, the data from the vessel's echosounder system have periodically been logged while the vessel has been carrying out regular fishing operations. The data presented here were collected from 28 January to 2 March 2009 when scientists were on board while the vessel was conducting regular fishing operations, and krill from the catch were sampled regularly. Fishing occurred along the northern shelf edge off the South Orkney Islands. The echosounder system was not calibrated so the logged data represent relative levels of acoustic backscatter. However, a post-calibration of the system was carried out off Brin-disi, Italy later that year using standard sphere calibration with a 38.1 mm tungsten carbide sphere (Foote et al., 1987), and showed that the transducers worked according to specifications, with no malfunctions.

2.3. Processing of the acoustic data

The acoustic raw data collected are typically flawed by noise from interference with other acoustic instruments, false bottom detection and surface bubbles. Regions with bad data were removed partly manually, and partly using noise removal algorithms incorporated in the software Large Scale Survey System (LSSS) (Korneliusson et al., 2006). The noise filtered data were then stored in sample bins of 50 pings horizontal \times 5 m vertical resolution, and the dB-difference method for target discrimination following the CCAMLR protocol with modifications was applied. This method takes advantage of the predictable frequency dependent volume backscattering strength (S_V ; dB rem^{-1}) for krill within a specified range of body lengths. The range of ΔS_V -values ($S_{V,120} - S_{V,38}$) is used to discriminate krill from other targets. We used the krill length distribution based on the collated trawl samples to calculate the values of ΔS_V (Reiss et al., 2008). The dB-difference method was applied to the sample bins of 50 pings

Table 1
Krill presence, density and relative abundance in the South Orkney Islands (January 28–March 2) estimated under different subsetting scenarios which selected observations considering the whole 34-days study period. Monthly-full subset includes all acoustic observations gathered during the 3 period; Monthly-2, Monthly-3 and Monthly-4 subsets were limited to locations observed, repeatedly, in any two, three or four different weeks, respectively, along the whole study period. n = number of observed locations (4-km² cells); \hat{p} = mean probability of stock presence; \hat{d}^* = mean conditional density (NASC); Φ = geostatistical range; RAI = relative abundance index (total back-scattering). Superscript letters t and g indicate time-series and/or geostatistical \hat{p} and/or \hat{d}^* models, respectively, were more informative (lower AIC) than alternative null ones.

Subset	n	Distribution			Conditional density			Relative abundance		
		\hat{p}	$SE_{\hat{p}}$	$\Phi_{\hat{p}}$	\hat{d}^*	$SE_{\hat{d}^*}$	$\Phi_{\hat{d}^*}$	RAI	SE_{RAI}	CV_{RAI}
Monthly-full	375	0.165 ^g	0.0069	4.2	1,423 ^{tg}	159	2.4	110,653	13,170	375
Monthly-2	194	0.191 ^g	0.0053	4.4	1,282 ^{tg}	141	1.7	115,682	13,093	194
Monthly-3	112	0.193 ^g	0.0074	5.0	1,441 ^{tg}	187	1.8	131,165	17,774	112
Monthly-4	38	0.200 ^g	0.0112	4.4	1,456 ^{tg}	484	5.1	137,418	46,371	38

horizontal*5 m vertical resolution, and if ΔS_v fell within the range estimated for krill targets, it was included as krill.

The target strength predictions of krill applied to calculate values of S_v at both frequencies were done using the simplified Stochastic Distorted Wave Born Approximation (SDWBA) package (Conti and Demer, 2006). However, the parameters of the simplified SDWBA were derived from an updated version of the package (Calise and Skaret, 2011), parametrized with the imaginary parts of the complex numbers included. The ΔS_v finally applied was based on a krill length range calculated in 10 mm bins based on krill TS predictions from a 95% PDF of krill length distribution based on the catches (SC-CCAMLR, 2010). After the discrimination, the retained values of Nautical Area Scattering Coefficient (NASC; m²/nmi²) for the 120 kHz data were averaged for each nautical mile.

2.4. Data analysis

The most important challenges we needed to address, regarding the analysis of the Saga Sea data, were the absence of sampling design and the lack of independence among sampling units due to two main sources of correlation: (1) correlation in time, as multiple observations were collected sequentially at each location during the study period, and (2) correlation in space, as sampling units were distributed along a continuous overlapping searching path (Fig. 1).

We defined and divided the 1616 km² (~471.15 nm²) study area into 404 squared cells of 4-km² each. A total of 375 out of these 404 cells were observed at least once during the study period. The remaining 29 cells were not observed, but located within the envelope defined by the 375 observed cells (Fig. 1). Following a delta approach (Aitchison, 1955; Pennington, 1983), we treated stock presence/absence and conditional density as separate response variables. Stock presence/absence was defined using a practical threshold of $NASC_{KRILL} > 10$ for presence. Therefore, the estimated mean for this binomial process corresponded to the probability \hat{p} of stock presence at a given location. Conditional krill density, hereafter \hat{d}^* , was instead treated as a Gaussian process, which represented all NASC values attributed to krill ($NASC_{KRILL}$) in sampled cells where the stock was present ($NASC_{KRILL} > 10$). Conditional density values were Box-Cox transformed to fulfil Gaussian distribution assumptions, using a Jacobian re-scaled function (Box et al., 1978) and transformation parameters (λ) estimated specifically, by maximum likelihood, for each data set.

Using a generalized linear mixed modelling approach (GLMM), we assumed that time and space covariances were stationary and separable (Cressie and Huang, 1999). Thus, we first produced time-wise means for each location xy (“location means” \hat{p}_{xy} and \hat{d}_{xy}^*). Location means were computed after pooling all repeated observations available for a given location and subset (see Section 2.5), under both: (i) a null model that assumed no autocorrelation in the time series, and (ii) an alternative model that assumed autoregressive error correlation of order 1 (Pinheiro and Bates, 2000).

Penalized quasi-likelihood (Venables and Ripley, 2002) and generalized least squares (Pinheiro and Bates, 2000) were used to fit presence/absence and density data, respectively.

Study area means, \hat{p} and \hat{d}^* , were computed from location means, \hat{p}_{xy} and \hat{d}_{xy}^* , under two spatial correlation scenarios: (i) a null (pure-nugget) model with no spatial correlation; and (ii) a matern correlation model, both fit through maximum likelihood procedures (Diggle and Ribeiro, 2007), using the R packages geoRglm (Christensen and Ribeiro, 2002) for presence/absence data and geoR (Ribeiro and Diggle, 2001) for box-cox transformed conditional density data.

Temporal and spatial correlation models were compared to their corresponding null models using Akaike (1973)'s Information Criterion. This comparison was, however, not possible for temporal correlation models of presence/absence data, since penalized quasi-likelihood methods do not produce true likelihood estimates. Reported means (Tables 1 and 2) correspond to those estimated using the null or alternative model shown to be the most informative one each variable and data subset. Example R scripts and additional methodological details are available as an online supplement to this paper.

As a final step, \hat{d}^* and \hat{p} estimates were combined into a relative abundance index (RAI), which was computed for each data subset and represented the total back-scattering (m²) attributed to krill in the study area ($A = 471.15$ nm²) during a certain time period. Thus,

$$R\hat{A}I = \hat{d}^* \times \hat{p} \times A$$

with nominal variance

$$V(R\hat{A}I) = \left[V(\hat{d}^*) \times \hat{p}^2 + (\hat{d}^*)^2 \times V(\hat{p}) \right] \times A^2$$

Conditional density means \hat{d}^* and their “nominal” variances $V(\hat{d}^*)$ were computed from geostatistical model intercepts, following a monte-carlo procedure. Thus, 5000 values were randomly drawn from the probability distribution function corresponding to each intercept $\hat{\beta}_{\hat{d}^*}$ and back-transformed applying the inverse Box-Cox function. Mean probability of presence \hat{p} was computed directly as the inverse logit of its geostatistical model intercept $\hat{\beta}_{\hat{p}}$, while its nominal variance was derived from its Taylor series expansion as,

$$V(\hat{p}) = \frac{e^{2\hat{\beta}_p}}{(1 + e^{\hat{\beta}_p})^4} \times V(\hat{\beta}_p)$$

We have used here the term “nominal variance of the mean” to highlight the fact that these measure of precision, computed directly from model parameter variances, only represents the variability expected if multiples means were obtained from multiple surveys, under identical circumstances. Therefore, it ignores spatial-temporal dynamics of both the target population and the observing platform. Since such assumptions are highly unrealistic in our particular sampling scenario, we also estimated “empirical”

Table 2

Krill presence, density and relative abundance in the South Orkney Islands (January 28–March 2) estimated for different data subsets, which only included observations collected by the F/V Saga Sea during a single week or day, as listed. n = number of observed locations (4-km² cells); \hat{p} = mean probability of stock presence; \hat{d}^* = mean conditional density (NASC); Φ = geostatistical range; RAI = relative abundance index (total back-scattering). Superscript letters t and g indicate time-series and/or geostatistical \hat{p} and/or \hat{d}^* models, respectively, were more informative (lower AIC) than alternative null ones. Average standard errors and CVs are provided as an empirical measure of precision.

	n	Distribution			Conditional density			Relative abundance		
		\hat{p}	$SE_{\hat{p}}$	$\Phi_{\hat{p}}$	\hat{d}^*	$SE_{\hat{d}^*}$	$\Phi_{\hat{d}^*}$	RAI	SE_{RAI}	CV_{RAI}
Weekly estimates										
Week 1	101	0.204 ^g	0.0165	1.6	926 ^{tg}	227	6.9	88,843	22,935	0.26%
Week 2	137	0.164 ^g	0.0072	4.8	1,237 ^{tg}	268	1.8	95,712	21,189	0.22%
Week 3	86	0.141 ^g	0.0142	2.0	1,780 ^{tg}	492	1.7	117,891	34,733	0.29%
Week 4	180	0.167 ^g	0.0080	3.7	938 ^{tg}	196	2.0	73,650	15,822	0.21%
Week 5	189	0.067 ^g	0.0067	5.4	868 ^g	117	2.4	27,332	4,594	0.17%
Average		0.148	0.0509	3.5	1,150	380	3.0	80,686	33,811	0.42%
Daily estimates										
Day 1	45	0.236	0.0254	0.4	948	238	2.7	105,620	28,884	0.27%
Day 6	51	0.200 ^g	0.0223	10.2	862 ^g	324	1.7	81,228	31,878	0.39%
Day 20	45	0.163 ^g	0.0167	5.4	458 ^{tg}	164	1.4	35,120	13,050	0.37%
Day 21	59	0.191 ^g	0.0210	9.7	968 ^{tg}	884	30.6	86,947	80,002	0.92%
Day 26	46	0.147 ^g	0.0254	2.5	841 ^t	255	0.6	58,325	20,342	0.35%
Day 27	49	0.123 ^g	0.0107	5.2	584 ^g	239	2.0	33,809	14,145	0.42%
Day 28	78	0.029 ^g	0.0040	6.1	537	164	1.0	7,364	2,469	0.34%
Average		0.156	0.0670	5.7	743	210	6.0	58,345	34,922	0.60%

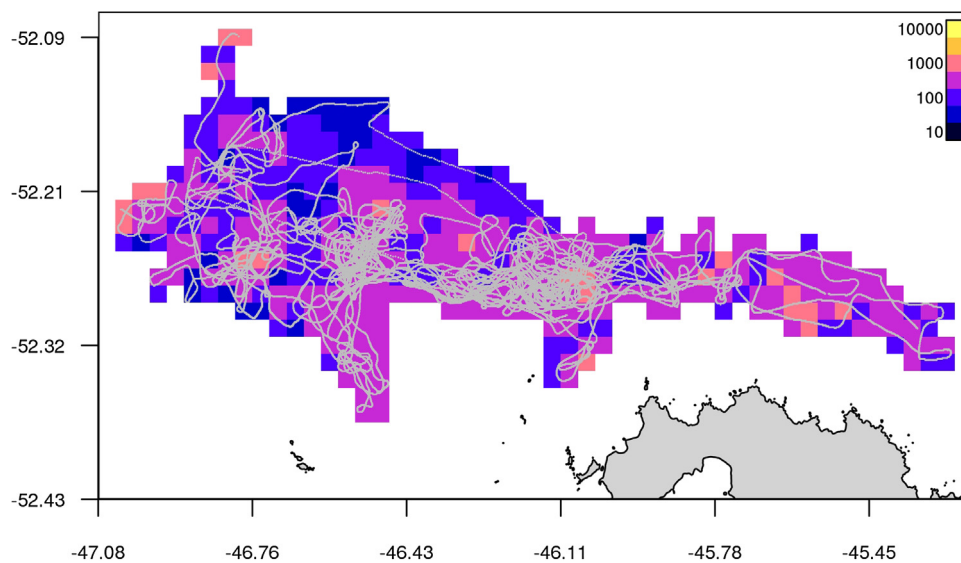


Fig. 1. Monthly (34-day) estimates of the mean nautical area scattering coefficient (NASC, m² mn⁻²) by spatial cell. Grey tracks show observations used to produce these estimates. NASC values at non observed locations predicted using simple kriging.

variances from the observed variability among daily and weekly mean estimates, as detailed in the following subsection.

2.5. Data sub-setting and empirical standard error estimates

To learn about sensitivity and precision of abundance estimates, under different sub-setting scenarios and time-intervals used for inference, we produced and compared RAI mean and standard error values obtained using the full 34-days and 375-locations data set, hereafter “monthly-full data set” and alternative monthly, weekly and daily sub-sets. As a first comparison, aimed to evaluate the sensitivity of abundance estimates to unbalanced coverage of different locations across time, we compared the full data set (“monthly-full”) with three different “monthly” subsets, which included data gathered throughout the whole study period, but from three selected sub-sets of locations. The first one (“monthly-2”) included monthly data for 193 locations that were observed, repeatedly, in two or more weeks, along the study period; the second one (“monthly-3”) included 104 locations that were observed, repeat-

edly, in at least three of the five weeks of the study period. The third one (“monthly-4”) included all data collected from 35 locations that were observed, repeatedly, in at least four of these five weeks.

To evaluate sensitivity to the time-interval used for inference, and to provide empirical standard errors for estimated means, we produced and compared RAI values considering two alternative sub-sampling intervals: (1) weekly abundance estimates, which included all data observations collected within each of the five weeks of the study period, and (2) daily abundance estimates, obtained using all observations collected at each of seven days, selected as those that exhibited observations on a minimum of 45 different locations.

3. Results

Krill distribution and vessel tracks suggested the presence of two main hotspots, located near the eastern and central parts of the study area (Fig. 1 and 2). An apparent westward displacement of both aggregations was apparent in the series of weekly estimates

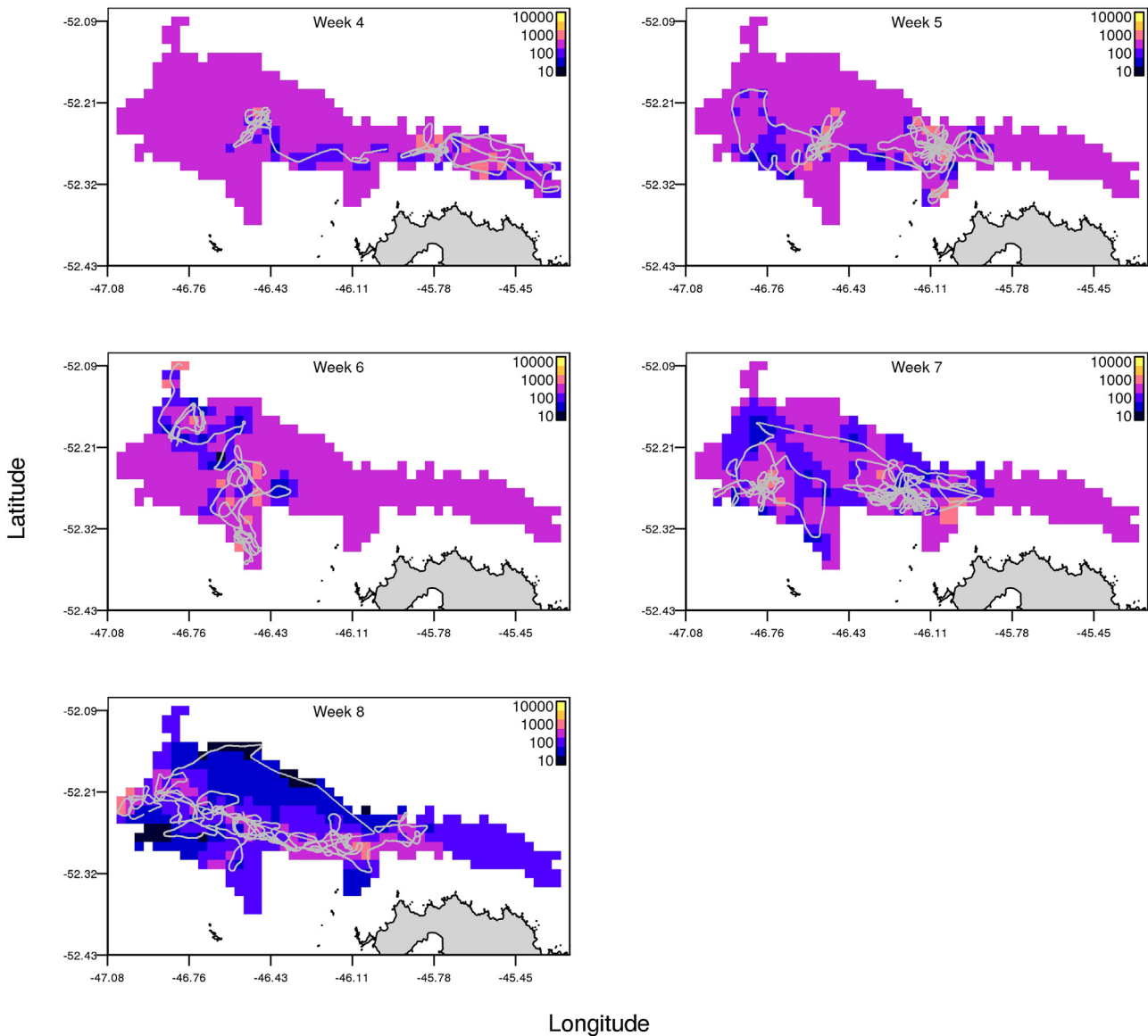


Fig. 2. Weekly estimates of the mean nautical area scattering coefficient (NASC, $\text{m}^2 \text{mn}^{-2}$) by spatial cell, for the five weeks that last the study period. Grey tracks show observations used to produce these estimates. NASC values at non observed locations predicted using simple kriging.

(Fig. 2). Time auto-regressive models for conditional density, within locations, were more informative (lower AIC) than null models for most data subsets (Tables 1 and 2). Geostatistical models for presence/absence were more informative than non-spatial models, for all but one data subset (Day 1). Geostatistical models for conditional density also tended to be more informative than non-spatial models. The latest were more informative than the first in all four monthly data subsets, all five weekly sub-sets, and four of the seven daily subsets (Tables 1 and 2). The estimated geostatistical range for presence/absence averaged 4.6 km, exhibiting a range between 2.0 and 10.2 km, while for conditional density it ranged between 0.6 and 31 km, across all data sets (Tables 1 and 2).

Monthly estimates from the full dataset (Fig. 1, Table 1) yielded RAI values between 110,700 and 137,400 m^2 , which represented a 24% difference between these two extreme values. While we did not find an evident relationship between the degree of temporal coverage (i.e. number of weeks at which a location was observed) and mean RAI estimates, nominal standard errors tended to increase with temporal coverage, as the number of locations exhibiting higher coverage was lower.

RAI estimates from weekly subsets (Fig. 2, Table 2) produced an average RAI of 80,700 m^2 , which was 27% lower than the mean RAI estimated using the full data set. Weekly estimates, however, were quite variable, ranging between 27,300 (week 5) and 117,900 (week 2), yielding nominal coefficients of variation of the mean below 30% for single week estimates, but an empirical pooled value as high as 42% (Table 2). This high coefficient of variation was largely affected by the last weekly estimate, when the estimated IAR dropped to 27,300 m^2 .

RAI estimates produced using selected daily subsets (Fig. 3) tended to be lower and more variable than monthly and weekly estimates, averaging 58,300 m^2 and ranging between 7,400 and 105,600 m^2 (Table 2). While nominal coefficients of variation of RAI means ranged between 27 and 92% for each daily estimate, the empirical value computed from observed variability among all daily estimates reached 60%.

4. Discussion

The acoustic dataset provided by the F/V 'saga Sea' showed enough quality to produce relatively precise monthly estimates of

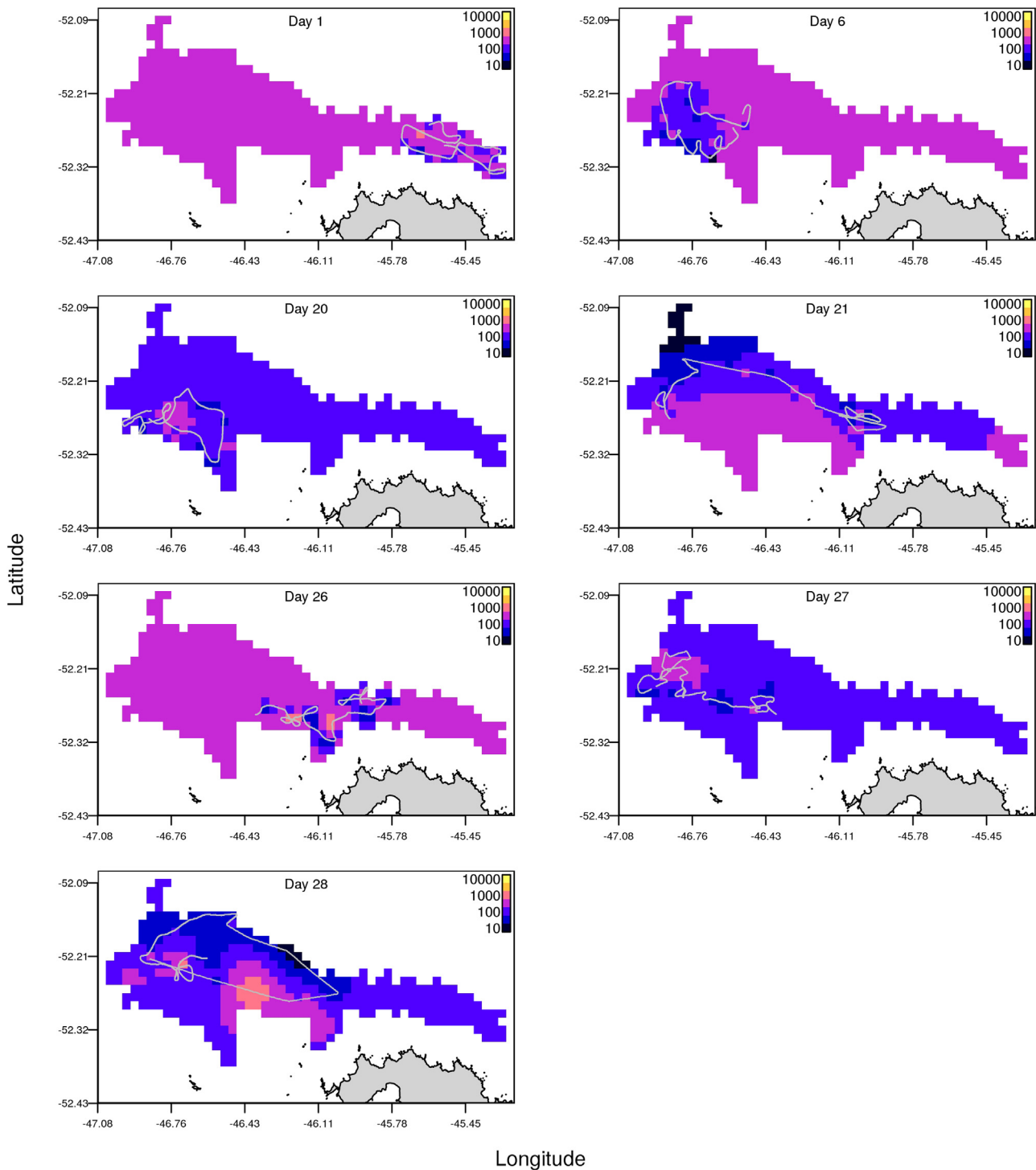


Fig. 3. Daily estimates of the mean nautical area scattering coefficient (NASC, $\text{m}^2 \text{mn}^{-2}$) by spatial cell, for seven days selected as the exhibited observations for ≥ 45 locations. Grey tracks show observations used to produce these estimates. NASC values at non observed locations predicted using simple kriging.

krill relative abundance at the South Orkney fishing ground. There was a limited sensitivity of monthly estimates to large differences in the number of locations (38–375) and the number of repeated observations per location (1–4 different weeks, 1–728 point observations) used to produce such estimates. Such stability suggests the methodological approach being used provided an acceptable way to account for time and space correlations.

Weekly estimates tended to be similar to monthly ones during the first four weeks of the study period, whose average ($94,024 \text{ m}^2$)

was 15% lower than the monthly-full estimate. Although we have no clear explanation for the very low value corresponding to the last weekly estimate ($27,332 \text{ m}^2$), it is tempting to consider it as a potential outlier, whose exclusion drops the empirical coefficient of variation from 36 to 19%. Such a low abundance estimate could result from a greater coverage of the area (189 locations) which had included a larger proportion of null and lower abundance cells, reflected in both lower \bar{P} and \bar{d}^* estimates (Table 2). On the other hand, it could reflect an actual decrease in krill abundance within

the study area, which could reflect effects from migration, predation and/or fishing removals. In fact, a decreasing trend with time is suggested in daily and weekly estimates between weeks 2 and 5.

The large variability observed in daily estimates was very likely related to the uneven coverage of the study area. Observing daily vessel tracks (Fig. 3), it is possible to conclude that its coverage was not so reduced in terms of the number of locations (45–78 per day), but highly concentrated in a small fraction of the study area. Thus, while our weekly and monthly estimates may contain some overestimation biases, attributable to preferential and clustered sampling (Diggle et al., 2010; Menezes et al., 2008), our daily estimates, may be affected by underestimation bias, due to insufficient coverage of the study area within daily operations, which tended to miss at least one of the two most apparent hotspots (Fig. 3).

We found an important potential for underestimating the variance of the mean and derived precision indices (standard errors and coefficients of variation) when they were computed directly from model parameter estimates. Hence, our empirical standard error estimates were up to 8-fold larger than nominal values calculated for conditional density means (weekly estimates). Although using such nominal estimations are soundly based upon central limit theorem, there are two main independence-related issues that limit the applicability of such theorem to krill acoustic data: (i) the selected sampling platform and gear produced inherent time and spatial correlation among most collected data, and (ii) the statistical population (krill available in a given area) is highly dynamic and may experience very large changes in short periods of time (days or weeks). While addressing the first issue may need of combining improved sampling designs and modelling approaches, addressing the second one may require of using a repeated surveys approach, on a regular basis.

It must be acknowledged that acoustic data from krill fishing vessels may differ substantially from other scenarios where maximum likelihood geostatistics of commercial acoustic data have been considered to produce relatively unbiased abundance estimates. In a SW Indian Ocean example, reported by Niklitschek and Roa (2006), data collected during orange roughly commercial operations was complemented with semi-systematic surveys conducted by the same fishing vessel to provide a broad coverage of each fishing ground right before its exploitation. This commitment was facilitated by the fact that most SW Indian Ocean seamounts are much smaller than the South Orkney fishing ground. In another successful example, Rubilar et al. (2006) managed to combine data from 5 vessels that searched intensively and simultaneously for hoki *Macruronus magellanicus* spawning aggregations within a single 40-km² canyon, off Western Patagonia. Learning from these two examples and from present results, we anticipate that the accuracy of krill relative abundance estimates from commercial acoustic data in Antarctic waters would be largely enhanced by: (1) implementing a minimum quota of mandatory design-based coverage of selected fishing grounds to all vessels, (2) constraining the area under scrutiny to habitats being used by krill as this information becomes available, and (3) retrieving and combining acoustic data from all vessels operating in a single area. As an example, for the 2013/14 fishing season there were 12 krill fishing vessels operating in the Southern Ocean, of which 8 were equipped with Simrad echo sounders (Watkins et al., 2016). An interesting alternative to reduce the operational cost of conducting design-based observations is to use processing times as done in New Zealand by hoki fishing vessels (O'Driscoll and Macaulay, 2005). This option, however, would make no difference for vessels such as the 'Saga Sea' fitted with continuous trawling systems.

For the full data set and all subsets, there was strong evidence of spatial correlation in krill distribution (stock presence/absence) at a relatively small spatial scale, with estimated geostatistical ranges of 4–10 km. Although conditional krill density did not always show

evidence of spatial correlation, when it did, it suggested a similar range of 2–15 km. Due to strong temporal correlation and the time-space synchrony inherent to the acoustic transect, it is possible that modelling time auto-correlation first had obscured spatial effects upon density. While these geostatistical range estimates may provide useful information about the probable size of krill biological patches (Perry et al., 2002; Sokal and Oden, 1978), the highly aggregated spatial pattern found in this stock must be considered in order to optimize future survey designs and to choose adequate data analysis tools.

In summary, we believe that maximum likelihood time-series and geostatistical analysis provide a suitable framework to be considered for the quantitative analysis of acoustic data from Antarctic krill fishing vessels. Nonetheless, and depending on the data properties, generalized linear mixed model approaches for zero-inflated data (Zuur et al., 2012), and non-parametric geostatistical methods (Henley, 2012) should be also considered and compared as important alternatives for this kind of data. The first ones may offer greater flexibility to accommodate non-linear covariates, and the second ones become more efficient when the much more stringent assumptions of the parametric methods can not be sustained. The geostatistical range may provide useful information about the probable size of a biological patch.

Acknowledgement

We extend our gratitude to Aker ASA and officers on board F/V 'Saga Sea' for making their acoustic equipment available for logging and allowing the acquired data to be analysed and presented in a scientific context.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.fishres.2015.09.017>.

References

- Aitchison, J., 1955. On the distribution of a positive random variable having a discrete probability mass at the origin. *J. Am. Stat. Assoc.* 50, 901–908.
- Aitchison, J., Brown, J.A., 1957. The lognormal distribution with special reference to its uses in economics. The lognormal distribution with special reference to its uses in economics. Cambridge University Press, London.
- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Caski, F. (Eds.), *Second International Symposium on Information Theory Presented at the Second International Symposium on Information Theory*. Budapest, pp. 267–281.
- Arkipkin, A., Barton, J., Wallace, S., Winter, A., 2013. Close cooperation between science, management and industry benefits sustainable exploitation of the Falkland Islands squid fisheries. *J. Fish Biol.* 83, 905–920.
- Atkinson, A., Siegel, V., Pakhomov, E., Rothery, P., Loeb, V., Ross, R., Quetin, L., Schmidt, K., Fretwell, P., Murphy, E., 2008. Oceanic circumpolar habitats of Antarctic krill. *Mar. Ecol. Prog. Ser.* 362, 1–23, others.
- Aubry, P., Debouzie, D., 2000. Geostatistical estimation variance for the spatial mean in two-dimensional systematic sampling. *Ecology* 81, 543–553.
- Bez, N., 2002. Global fish abundance estimation from regular sampling: the geostatistical transitive method. *Can. J. Fish. Aquat. Sci.* 59, 1921–1931.
- Box, G.E., Hunter, W.G., Hunter, J.S., 1978. *Statistics For Experimenters*. John Wiley & Sons, New York.
- Calise, L., Skaret, G., 2011. Sensitivity investigation of the SDWBA Antarctic krill target strength model to fatness: material contrasts and orientation. *CCAMLR Sci.* 18, 97–122.
- CCAMLR, 2010. Report of the twenty-ninth meeting of the Commission for the Conservation of Antarctic Marine Living Resources. CCAMLR-XXIX. Hobart, Australia.
- Christensen, O.F., Ribeiro, P.J., 2002. *geoRglm—a package for generalised linear spatial models*. *R News* 2, 26–28.
- Constable, A.J., 2011. Lessons from CCAMLR on the implementation of the ecosystem approach to managing fisheries. *Fish Fish* 12, 138–151.
- Constable, A.J., William, K., Agnew, D.J., Everson, I., Miller, D., 2000. Managing fisheries to conserve the Antarctic marine ecosystem: practical implementation of the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR). *ICES J. Mar. Sci. J. Cons.* 57, 778–791.

- Conti, S.G., Demer, D.A., 2006. Improved parameterization of the SDWBA for estimating krill target strength. *ICES J. Mar. Sci. J. Cons.* 63, 928–935.
- Cressie, N.A.C., 1993. *Statistics for Spatial Data*. John Wiley and Sons, Inc., New York.
- Cressie, N., Huang, H.-C., 1999. Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Am. Stat. Assoc.* 94, 1330–1339.
- Diggle, P.J., Menezes, R., Su, T., 2010. Geostatistical inference under preferential sampling. *J. R. Stat. Soc. Ser. C Appl. Stat.* 59, 191–232.
- Diggle, P., Ribeiro, P.J., 2007. *Model-based geostatistics*. Springer.
- Everson, I., 2008. *Krill: Biology, Ecology and Fisheries*. John Wiley & Sons, New York.
- Foot, K.G., Knudsen, H.P., Vestnes, G., MacLennan, D.N., Simmonds, E.J., 1987. Calibration of acoustic instruments for fish density estimation: a practical guide. In: *ICES cooperative research report N° 144, International Council for the Exploration of the Sea, Copenhagen, Denmark*.
- Gunderson, D.R., 1993. *Surveys of Fisheries Resources*. Wiley, New York.
- Henley, S., 2012. *Nonparametric Geostatistics*. Springer Science & Business Media.
- Hewitt, R.P., Watkins, J.L., Naganobu, M., Tshernyshkov, P., Brierley, A.S., Demer, D.A., Kasatkina, S., Takao, Y., Goss, C., Malyshko, A., Brandon, M.A., Kawaguchi, S., Siegel, V., Trathan, P.N., Emery, J.H., Everson, I., Miller, D.G.M., 2002. Setting a precautionary catch limit for Antarctic krill. *Oceanography* 15, 26–33.
- Jolly, G.M., Hampton, I., 1990. A stratified random transect design for acoustic surveys of fish stocks. *Can. J. Fish. Aquat. Sci.* 47, 1282–1291.
- Korneliusson, R.J., Ona, E., Eliassen, I., Heggelund, Y., Patel, R., Godø, O.R., Giertsen, C., Patel, D., Nornes, E., Bekkvik, T., others, 2006. The large scale survey system—LSSS, in: *Proceedings of the 29th Scandinavian Symposium on Physical Acoustics, Ustaoset*.
- Lambert, D., 1992. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 34, 1–14.
- Lo, N.C.H., Jacobson, L.D., Squire, J.L., 1992. Indexes of relative abundance from fish spotter data based on delta-lognormal models. *Can. J. Fish. Aquat. Sci.* 49, 2515–2526.
- Menezes, R., García-Soidán, P., Febrero-Bande, M., 2008. A kernel variogram estimator for clustered data. *Scand. J. Stat.* 35, 18–37.
- Molinet, C.A., Díaz, M.A., Muñoz, C.A., Leny, E., unise, C., ares, S., andra, L., orena, M., arin, M., arcela, P., atricia, A., storga, E., dwin, J., uan, E., lias, N., iklitschek, 2015. Spatial distribution pattern of *Mytilus chilensis* beds in the Reloncaví fjord: hypothesis on associated processes. *Rev. Chil. Hist. Nat.* 88 (1), 1, <http://dx.doi.org/10.1186/s40693-015-0041-7>.
- Molinet, C., Niklitschek, E.J., Arévalo, A., Almanza, V., Codjambassiss, J., Diaz, P., Diaz, M., González, M.T., Matamala, M., 2010. Patch structure of benthic resources exploited in Chilean management areas: the shellfish bed concept under a territorial use rights for fishers framework. *Bull. Mar. Sci.* 86, 555–569.
- Niklitschek, E., Roa, R., 2006. Likelihood-based geostatistical biomass estimation using acoustic data collected by industry vessels in the South West Indian Ocean. In: *Shotton, R. (Ed.), Management of Demersal Fisheries Resources of the Southern Indian Ocean*. FAO Fisheries, Rome, Italy, pp. 75–79.
- O'Driscoll, R.L., Macaulay, G.J., 2005. Using fish-processing time to carry out acoustic surveys from commercial vessels. *ICES J. Mar. Sci.* 62, 295–305.
- Pennington, M., 1983. Efficient estimators of abundance, for fish and plankton surveys. *Biometrics* 39, 281–286.
- Perry, J.N., Liebholt, A.M., Rosenberg, M.S., Dungan, J., Miriti, M., Jakomulska, A., Citron-Pousty, S., 2002. Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography* 25, 578–600, <http://dx.doi.org/10.1034/j.1600-0587.2002.250507.x>.
- Pinheiro, J.C., Bates, D.M., 2000. *Mixed Effects Models in S and S-PLUS*. Springer Verlag, New York USA.
- Reiss, C.S., Cossio, A.M., Loeb, V., Demer, D.A., 2008. Variations in the biomass of Antarctic krill (*Euphausia superba*) around the South Shetland Islands, 1996–2006. *ICES J. Mar. Sci. J. Cons.* 65, 497–508.
- Ribeiro, P.J., Diggle, P.J., 2001. *geoR: a package for geostatistical analysis*. *R-News* 1, 15–18.
- Roa-Ureta, R., Niklitschek, E., 2007. Biomass estimation from surveys with likelihood-based geostatistics. *ICES J. Mar. Sci.* 64, 1723–1734.
- Rubilar, P., Niklitschek, E., Zuleta, A., Gili, R., Rosales, S., Cornejo, J., Merino, R., Hernández, E., Toledo, P., 2006. Distribución y abundancia de las agregaciones reproductivas de merluza de cola. Universidad Austral de Chile-Centro de Estudios Pesqueros (CEPES)-Centro Trapananda, Coyhaique, Chile.
- SC-CCAMLR, 2014. Report of the meeting of the Subgroup on Acoustic Survey and Analysis methods. Scientific Committee, Commission for the Conservation of Antarctic Marine Living Resources (SC-CCAMLR-XXXIII), Annex 4. Hobart Australia.
- SC-CCAMLR, 2012. Report of the sixth meeting of the Subgroup on Acoustic Survey and Analysis methods. Scientific Committee, Commission for the Conservation of Antarctic Marine Living Resources (SC-CCAMLR-XXXI), Annex 4. Hobart Australia.
- SC-CCAMLR, 2010. Report of the fifth meeting of the Subgroup on Acoustic Survey and Analysis methods. Scientific Committee, Commission for the Conservation of Antarctic Marine Living Resources (SC-CCAMLR-XXIX/6), Annex 4. Hobart Australia.
- Sokal, R.R., Oden, N.L., 1978. Spatial autocorrelation in biology: 1. Methodology. *Biol. J. Linn. Soc.* 10, 199–228, <http://dx.doi.org/10.1111/j.1095-8312.1978.tb00013.x>.
- Venables, W.N., Ripley, B.D., 2002. *Modern Applied Statistics with S*, 4th edition. New York, Springer.
- Watkins, J., Reid, K., Ramm, D., Zhao, X.Y., Cox, M., Skaret, G., Fielding, S., Wang, X.L., Niklitschek, E.J., 2016. The use of fishing vessels to provide acoustic data on the distribution and abundance of Antarctic krill and other pelagic species. *Fish. Res.* 178, 93–100.
- Zuur, A.F., Saveliev, A.A., Ieno, E.N., 2012. *Zero inflated models and generalized linear mixed models with R*. Highland Statistics Limited, Newburgh.