

RESPONSE TO EVERSON (2008)

ESTIMATING STANDING STOCK OF KRILL USING MAXIMUM ENTROPY IMAGE RECONSTRUCTION

A.S. Brierley✉ and B.G. Heywood
Pelagic Ecology Research Group
Gatty Marine Laboratory
University of St Andrews, Scotland, UK
Email – asb4@st-and.ac.uk

S.F. Gull
Astrophysics Group
Cavendish Laboratory
University of Cambridge, UK

Using multifrequency acoustic data from the CCAMLR-2000 Survey (Hewitt et al., 2004; Watkins et al., 2004), krill target strengths (TS) determined from an SDWBA model (Demer and Conti, 2005), and the Jolly and Hampton (1990) frequentist statistical analysis approach, the biomass of Antarctic krill in the southwest Atlantic has been estimated as 109.4 million tonnes with a standard deviation (SD) of 11.38 million tonnes (Demer and Conti, 2005). Heywood et al. (2006) analysed 1 n mile along-transect krill density data from the same survey, scaled from the same echo-intensity data using the same TS values, in a Bayesian Maximum Entropy (MaxEnt)-based image-reconstruction framework (e.g. Brierley et al., 2003), and determined the most-likely krill biomass estimate for the same area to be 207.98 million tonnes (SD 10.08 million tonnes). Biomass and its associated variance (SD squared) are both crucial input parameters used by CCAMLR to set precautionary catch limits for krill (Hewitt et al., 2004), so it is important that the provenance of these different mean and SD values, and reasons for their differences, be understood.

Dr Everson has suggested that the mismatch between the Demer and Conti (2005) and Heywood et al. (2006) krill biomass estimates arises because ‘something is wrong’ with either or both methods, and has implied that since, in his opinion, the MaxEnt approach is not ‘simple and robust’ (whereas presumably he believes that the Jolly and Hampton (1990) approach is) the fault must lie with the MaxEnt approach. The two methods are however quite different, particularly in the ways in which SD is calculated, and it is essential that

these differences be considered during comparison of output. Whilst it is true that application of the MaxEnt approach is not ‘simple’ (it is computationally quite intensive, but well within the capabilities of a PC), the theory is well established, and proven in many fields, and it is certainly robust. Bayesian approaches generally are being used increasingly by ecologists (e.g. Ellison, 2004), and MaxEnt methods are recognised increasingly in many fields – including stock assessment and species distribution mapping – for the benefits they can offer over conventional frequentist approaches (Vignaux et al., 1998; Wyatt, 2003; Phillips et al., 2006).

The Jolly and Hampton (1990) approach treats individual survey transects as the sampling unit, on the assumption that mean density over each transect is representative of the mean density of the stratum in which the transect lies (Hewitt et al., 2004). A stratum is a distinct geographical area defined as an independent entity during survey design (see Trathan et al., 2001 for the CCAMLR-2000 strata). Stratum biomass is determined from the mean transect density raised to the total stratum area. For a single survey stratum, the mean density is determined with the number of transects (n) as the denominator, whilst the variance calculation uses ($n \times (n-1)$) as the denominator. The CCAMLR-2000 Survey comprised a total of 48 transects in 7 strata, with a maximum of 10 transects per stratum. The total transect number was determined pragmatically at the planning stage as a function of available effort from the amount (length) of cruise track that could be sampled by the participating vessels in the available time, and the longitudinal and latitudinal

extent of the entire survey area (Trathan et al., 2001). Thus the value of n was more a function of logistic constraint than biological or statistical preference. Aside from allocating extra effort (survey km per square km) to regions where high krill densities were expected (e.g. around island groups, in areas where fishing activity had been focused), the survey design (transect spacing and number) did not consider scales of patchiness of krill distributions. The survey was neither designed with nor, under the Jolly and Hampton (1990) scheme, analysed in recognition of the numerous spatial scales of krill aggregation (e.g. Weber and El-Sayed, 1985; Murphy et al., 1988; Miller and Hampton, 1989) that Dr Everson recognises.

The MaxEnt approach, by contrast, takes explicit account of spatial structure (spatial autocorrelation) in the survey data. Krill density per unit distance along transect (1 n mile in the case of Heywood et al., 2006) is considered in a framework of spatial correlation to adjacent units, and the spatial structure *apparent in the data* is used explicitly to estimate the density in units of the same spatial dimension (1 n mile \times 1 n mile in this case) off-transect. The Bayesian framework enables spatial structure at multiple spatial scales to be examined, the scales most contributory to the overall spatial distribution to be determined and, for these scales – that are most consistent with the observed survey data – to be used in the estimation of biomass values off-transect. The MaxEnt method therefore absolutely *does* consider spatial scale determined from the krill distribution, whether that pattern of distribution is caused by krill behaviour, predation, environmental forcing or other factors, and it is wrong for Dr Everson to suggest otherwise. Our analysis considered potential spatial structuring over scales from 1 to 512 n miles and determined quantitatively, with reference to Bayesian evidence, that structure between 1 and 128 n miles was prominent in the observed survey data. These scales were then used in the estimation of biomass values for each of the 1 n mile squares in the 1 578 \times 1 094 n mile oblong within which the CCAMLR-2000 Survey fell. The biomass estimate of 207.98 million tonnes reported by Heywood et al. (2006) is the sum of biomass in all the 449 500 1 n mile squares in the defined (Trathan et al., 2001) survey area. It would be quite straightforward to estimate distribution of biomass across the CCAMLR-2000 Survey area at a finer spatial scale by using finer spatial scale input data. However, we need to remember that the multifrequency technique used to identify echoes arising from krill requires a certain amount of averaging to overcome ping-to-ping stochastic variability (Watkins and Brierley, 2002). With this in mind, Hewitt et al. (2004) averaged the raw acoustic

survey data over 50 pings horizontally before applying the dB difference technique to identify krill. At a nominal survey speed of 10 knots and a ping rate of 0.5 Hz, this equates to an along-track resolution of approximately 0.5 km. Thus an analysis could be conducted with consensus CCAMLR data down to a 0.5 km scale, but analysis at a finer scale than this would require a reappraisal of the techniques used to delineate backscatter attributable to krill.

In the same way that the MaxEnt approach evaluates biomass for the survey as the sum of the biomasses in each n mile square, SD is determined as the sum of SDs for each individual 1 n mile square. Individual square SDs are determined from the within-square variability between the family of most-likely reconstructions (the posterior distribution – see Gull and Skilling, 1999). The survey SD reported by Heywood et al. (2006) is therefore based on an n of 449 500, whereas the Jolly and Hampton value is founded on the variation between just 48 transects. The MaxEnt approach enables the spatial distribution of variance to be studied (maps of SD by n mile square can be constructed – see Heywood et al., 2006, Figure 7), and this in turn could provide guidance for distribution of survey effort in subsequent surveys.

Jolly and Hampton (1990)-derived transect means smooth along-transect variability such that the transect-based density range (maximum value minus minimum value) will be less than the range between individual 1 n mile intervals along transect. In contrast, the MaxEnt image reconstruction technique has the capability of inferring off-transect values that are larger than the maximum along-transect sampled values (if this is warranted by structure in the sampled data). There is no reason to suppose that any of the limited number of survey transects would pass exactly through the very highest krill densities, particularly if high density swarms are rare. Thus, there is no reason for the MaxEnt approach to ‘provide an estimate of standing stock close to that of the Jolly and Hampton method’ as Dr Everson suggested, indeed the precision and accuracy of the two approaches may well be different. Since the distribution of krill densities by 1 n mile grid cell is highly skewed (not normal) – much krill biomass is located within few high-density swarms – the MaxEnt reconstructed distribution could well be expected to yield a higher biomass (essentially the integral of the map of distribution) because more of the rare high-density aggregations will be included. What is more, high inferred biomass values need not necessarily be associated with high SD if they are present in a large proportion of the family of most-probable

reconstructions (Gull and Skilling, 1999). High individual cell SD is not moderated in overall SD by division. The apparent close agreement of SD between techniques is essentially a coincidence of numbers. Both SD values are computationally valid, but are fundamentally different. We note that the generalised yield model (GYM) (Constable and de la Mare, 1996), that is used by CCAMLR in determining the precautionary catch limit for krill, does not appear to specify explicitly how the SD that it takes as input should be calculated. It may be implicit that it is the SD arising from a Jolly and Hampton (1990) analysis but, given the likely proliferation of analysis techniques, and the potential for discrepancy between them, it ought perhaps to be stated formally.

A problem inherent with any estimation process from sample data is that the 'truth', in this case the actual krill biomass/map of distribution, is not known. We have conducted simulated surveys through simulated data – for which the distribution of density is known – and have shown that the MaxEnt approach is good at recovering detail (Heywood, 2008). A related possibility for exploring the scale of agreement between Jolly and Hampton (1990) and MaxEnt analyses would be to run a number of CCAMLR-2000-like surveys through the reconstructed MaxEnt distribution and determine what the variability between surveys would be. The narrow sampling beams of standard echosounders have the potential to miss spatially small but high-biomass aggregations, so minor deviations in ship's track have the potential to yield very different biomass estimates. Progress towards use of multibeam echosounders, that provide a larger window of observation, could reduce this for patchily distributed resources such as krill (Cox et al., in press). A further development, that would get away from the need to use simulated data to compare analytical approaches, would be to run simulated surveys through distribution data generated by the emerging 'ocean acoustic waveguide remote sensing' technique (Makris et al., 2006). This provides direct snapshots of species distributions on very large scales without the need to interpolate between often widely spaced transects. We strongly encourage krill researchers to embrace the whole range of available survey and analysis techniques – including MaxEnt – to improve the precision and accuracy of krill biomass estimates, understanding of krill and its ecology, and to support robust, ecosystem-based management for the species (Nicol and Brierley, in press). Until the 'truth' is known, numerous interpretations of data will remain 'plausible'.

References

- Brierley, A.S., S.F. Gull and M.H. Wafy. 2003. A Bayesian maximum entropy reconstruction of stock distribution and inference of stock density from line-transect acoustic-survey data. *ICES J. Mar. Sci.*, 60 (3): 446–452.
- Constable, A. and W.K. de la Mare. 1996. A generalised model for evaluating yield and the long-term status of fish stocks under conditions of uncertainty. *CCAMLR Science*, 3: 31–54.
- Cox, M.J., D.A. Demer, J.D. Warren, G.R. Cutter and A.S. Brierley. In press. Three-dimensional observations of swarms of Antarctic krill (*Euphausia superba*) made using a multi-beam echosounder. *Deep-Sea Res.*, II.
- Demer, D.A. and S.G. Conti. 2005. New target-strength model indicates more krill in the Southern Ocean. *ICES J. Mar. Sci.*, 62: 25–32.
- Ellison, A.M. 2004. Bayesian inference in ecology. *Ecol. Lett.*, 7: 509–520.
- Gull, S.F. and J. Skilling. 1999. Quantified Maximum Entropy. *MemSys5: Users Manual V1.2*. www.maxent.co.uk/documents/MemSys5_manual.pdf, 99 pp.
- Hewitt, R.P., J. Watkins, M. Naganobu, V. Sushin, A.S. Brierley, D. Demer, S. Kasatkina, Y. Takao, C. Goss, A. Malysenko, M. Brandon, S. Kawaguchi, V. Siegel, P. Trathan, J. Emery, I. Everson and D. Miller. 2004. Biomass of Antarctic krill in the Scotia Sea in January/February 2000 and its use in revising an estimate of precautionary yield. *Deep-Sea Res.*, II, 51: 1215–1236.
- Heywood, B.G. 2008. Investigations into the use of quantified Bayesian Maximum Entropy methods to generate improved distribution maps and biomass estimates from fisheries acoustic survey data. M. Phil., University of St Andrews.
- Heywood, B.G., A.S. Brierley and S.F. Gull. 2006. A quantified Bayesian Maximum Entropy estimate of Antarctic krill abundance across the Scotia Sea and in small-scale management units from the CCAMLR-2000 Survey. *CCAMLR Science*, 13: 97–116.
- Jolly, G.M. and I. Hampton. 1990. A stratified random transect design for acoustic surveys of fish stocks. *Can. J. Fish Aquat. Sci.*, 47: 1282–1291.
- Makris, N.C., P. Ratilal, D.T. Symonds, S. Jagannathan, S. Lee and R.W. Nero. 2006. Fish

- population and behavior revealed by instantaneous continental shelf-scale imaging. *Science*, 311: 660–663.
- Miller, D.G.M. and I. Hampton. 1989. Krill aggregation characteristics: spatial distribution pattern from hydroacoustic observations. *Polar Biol.*, 10: 125–134.
- Murphy, E.J., D.J. Morris, J.L. Watkins and J. Priddle. 1988. Scales of interaction between Antarctic krill and the environment. In: Sahrhage, D. (Ed.). *Antarctic Ocean and Resources Variability*. Springer-Verlag, Berlin Heidelberg: 120–130.
- Nicol, S. and A.S. Brierley. In press. Through a glass less darkly – new approaches for studying the distribution, abundance and biology of Euphausiids. *Deep-Sea Res.*, II.
- Phillips, S.J., R.P. Anderson and R.E. Schapire. 2006. Maximum entropy modelling of species geographic distributions. *Ecol. Model.*, 190: 231–259.
- Trathan, P.N., J.L. Watkins, A.W.A. Murray, A.S. Brierley, I. Everson, C. Goss, J. Priddle, K. Reid, P. Ward, R. Hewitt, D. Demer, M. Naganobu, S. Kawaguchi, V. Sushin, S.M. Kasatkina, S. Hedley, S. Kim and T. Pauly. 2001. The CCAMLR-2000 Krill Synoptic Survey: a description of the rationale and design. *CCAMLR Science*, 8: 1–24.
- Vignaux, M., G.A. Vignaux, S. Lizamore and D. Gresham. 1998. Fine-scale mapping of fish distribution from commercial catch and effort data using maximum entropy tomography. *Can. J. Fish. Aquat. Sci.*, 55: 1220–1227.
- Watkins, J.I. and A. Brierley. 2002. Verification of acoustic techniques used to identify and size Antarctic krill. *ICES J. Mar. Sci.*, 59: 1326–1336.
- Watkins, J.L., R. Hewitt, M. Naganobu and V. Sushin. 2004. The CCAMLR 2000 Survey: a multinational, multi-ship biological oceanography survey of the Atlantic sector of the Southern Ocean. *Deep-Sea Res.*, II, 51: 1205–1213.
- Weber, L.H. and S. El-Sayed. 1985. Spatial variability of phytoplankton and the distribution and abundance of krill in the Indian Sector of the Southern Ocean. In: Siegfried, W.R., P.R. Condy and R.M. Laws (Eds). *Antarctic Nutrient Cycles and Food Webs*. Springer-Verlag, Berlin Heidelberg: 284–293.
- Wyatt, R.J. 2003. Mapping the abundance of riverine fish populations: integrating hierarchical Bayesian models with a geographic information system (GIS). *Can. J. Fish. Aquat. Sci.*, 60: 997–1006.