Modelling environmental influences on squid life history, distribution, and abundance

Graham J. Pierce[1], M. Begoña Santos[2], Colin D. MacLeod[1], Jianjun Wang[1], Vasilis Valavanis[3] and Alain F. Zuur[4]

[1]School of Biological Sciences, University of Aberdeen, Tillydrone
Avenue, Aberdeen AB24 2TZ, UK (g.j.pierce@abdn.ac.uk).

[2]Instituto Español de Oceanografía, Centro Costero de
Vigo, Cabo Estay, Canido, 36200 Vigo, Spain.

[3]Marine Geographic Information Systems, Hellenic Center for Marine Research,
P.O. Box 2214, 71003 Iraklion, Crete, Greece.

[4]Highland Statistics Ltd, 6 Laverock Road, Newburgh,
Aberdeenshire AB41 6FN, UK.

Following Caddy and Gulland's (1983) classification, squid populations can be described as displaying irregular or spasmodic abundance fluctuations, as opposed to steady or cyclical patterns. Squid are short-lived, fast growing marine animals, thought to be especially sensitive to environmental influences (Coelho, 1985). Because generations are essentially non-overlapping, modelling of population dynamics reduces to predicting recruitment success (Caddy, 1983; Pierce and Guerra, 1994). Environmental signals are expected to have a strong effect on spawning and hatching success and on growth and survival of early life stages. Since squid are increasingly important fishery resources in many parts of the world, much recent research has focused on understanding their distribution and abundance, to provide the basis for efficient and sustainable utilisation of these resources. The recent high abundance of jumbo squid *Dosidicus gigas* in the eastern Pacific Ocean raises questions as to why abundance and range have increased, the ecological consequences and whether or for how long its current status will persist. The aim of this brief review is to synthesise approaches to modelling the spatiotemporal patterns in squid life history, distribution, abundance and fisheries, and to identify relevant research questions in relation to *D. gigas*.

In principle, models of spatio-temporal variation can be fitted to data on fishery catches and squid distribution, abundance and life history parameters. Critical stages of the life cycle include spawning, hatching, early growth, recruitment to the fishery and movements to the spawning grounds. In many squid, the paralarval (post-hatching) period tends to be the least well known. Because there is no buffering effect of older age classes, environmental effects on the extant generation are a particularly important feature of population dynamics, which has led to a focus on empirical rather than mechanistic models of abundance. In particular, we need to understand the relationship between spawning stock size and subsequent recruitment and pre-spawning mortality of recruits.

Data on squid may arise from fisheries, surveys, predators, tagging or direct observation: all have associated limitations and biases. Relevant considerations include availability, coverage, resolution (in time and space), accuracy and precision. Relevant environmental factors include large scale phenomena such as the *El Niño*-Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO), current systems, fixed physical phenomena such as seabed depth and substrate, oceanographic parameters such as sea surface temperature (SST) and salinity (SSS), meso-scale ocean surface (and sub-surface) features and daily, lunar and seasonal cycles. The selection of relevant variables is normally a compromise between biological relevance and data availability. The most readily available oceanographic data tend to be related to surface characteristics measured by satellite-based instruments, but it is also important to consider the vertical dimension. Developments in remote sensing, geographic information systems (GIS) and statistical modelling have all facilitated current modelling applications.

Empirical modelling is not without its critics: by focusing on the data rather than the underlying mechanisms it encourages data-driven rather than hypothesis-driven research. However, these are also advantages: hypotheses can be generated about the mechanisms and functional forms of relationships. As with all models, adequate testing of predictions is needed to eliminate spurious (coincidental) relationships: as noted by Solow (2002), time-series correlations often disappear once longer series become available, and similar caveats apply to spatial models.

Several generic issues apply to time-series and spatial modelling: these include model selection, model fitting, testing predictions, decisions about scale, incorporating environmental effects that are displaced in space and/or time (time-lags and teleconnections), autocorrelation, seasonal patterns and unexplained trends, interactions between variables, identification of data distributions, variance structure and linearity of relationships.

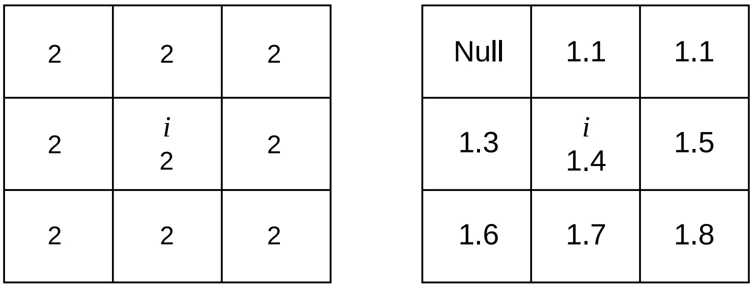
Time series can be modelled using regression, generalised linear or additive models (GLM or GAM), but such models may be invalidated by temporal auto-correlation, i.e. non-independence of adjacent values of the response variable, which can inflate apparent statistical significance by up to 400% (Zuur *et al.,* 2007). In short-lived species, the link between abundance or life history parameter values in successive generations (years) may be weak or non-existent. In the former case, inserting a term for the previous year's value as an extra explanatory variable into the model may adequately account for autocorrelation. Whether this is an issue can be confirmed by testing the model residuals for temporal autocorrelation. Otherwise, solutions include Generalised Additive Mixed Modelling (GAMM), Seasonal and Trend decomposition using Loess (STL), Autoregressive Integrated Moving Average Models (ARIMA) or Dynamic Factor Analysis (DFA, a multivariate extension of STL). On a cautionary note here, the autocorrelation, moving average, seasonal and trend terms in such models capture variation without explaining it. Environmental parameters may however, be included. Time-lagged relationships with environmental variables can also be detected using cross-correlation analysis; some authors recommend so-called "pre-whitening" (de-trending) of series prior to analysis. However, common trends thus removed may indicate a genuine causal link. Links between *Illex argentinus* abundance and ENSO events, with a 5-year time-lag were detected by Waluda *et al*. (1999).

ARIMA and DFA models of temporal trends in *Loligo forbesi* abundance have been published, and demonstrate effects of SST and the NAO index on abundance (Pierce and Boyle, 2003; Zuur and Pierce, 2004). However, these models tend to have low predictive power and, given the weak temporal structure of the squid time-series, approaches using regression, GAMs or regression trees, may be equally useful (e.g. Bellido *et al*., 2001; Waluda *et al*., 2001). Sims *et al*. (2001) used polynomial regression to demonstrate a link between timing of migration and the NAO index in *L. forbesi*. Pierce *et al*. (2005) used GAMs to extract interannual variation in size at maturity in *L. forbesi*, and then used correlation analysis to demonstrate that this residual variation could be related to the NAO index.

Spatial modelling has been greatly facilitated by GIS (Pierce *et al*., 2002). Routines have been developed to identify meso-scale ocean surface features from variability or discontinuities in temperature (Valavanis *et al*., 2005; Wang *et al*., 2007, see Figure 1), as well as from variability in temperature and chlorophyll anomalies (Valavanis *et al*., 2004a). It is also possible to incorporate temporal variation into such models, although to date, there has been little work modelling distribution in the vertical dimension. Data from tagging have allowed description of squid movements underwater, in relation to oceanographic data, and such data would be amenable to fitting models.

Essential habitat and migration corridor models may be constructed using an entirely GIS-based process, incorporating fishery and environmental data constrained by life history data on species' "preferred" living environmental conditions (Valavanis *et al*., 2002; Valavanis *et al*., 2004b). GAMs have been used to describe distribution patterns in both *L. forbesi* and *I. argentnus*, while regression trees have also been applied to data on *L. forbesi*, revealing relationships with temperature, salinity and depth (Pierce *et al*., 1998; Bellido *et al*., 2001; Sacau *et al*., 2005). Moreno *et al*. (in press) used GAMs to separate seasonal, annual and temperature effects on growth rates in *L. vulgaris*. The latter effects represent spatial variation in growth rates.

Where presence records are available (e.g. from predator samples or tags) but there are no absence records, presence only modelling techniques such as ENFA (Hirzel *et al*., 2000) may be applied. There are no current applications to data on squid.

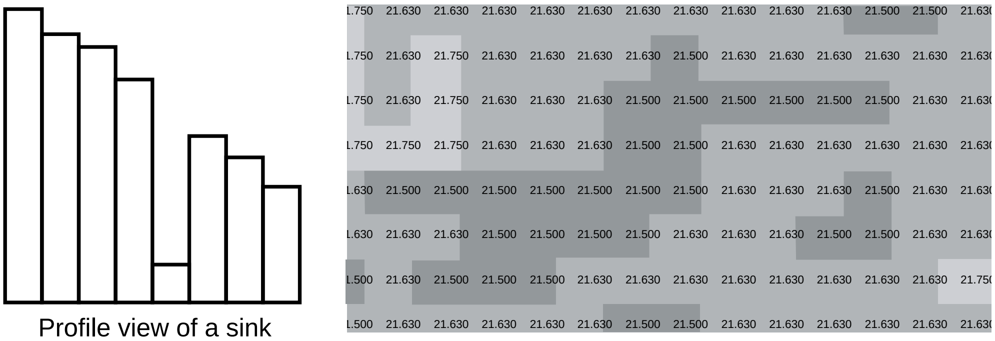


2	2	2
2	i 2	2
2	2	2

a) Local SST relative variability
$RV_i = 1/9 = 0.111$

Null	1.1	1.1
1.3	i 1.4	1.5
1.6	1.7	1.8

b) Local SST relative variability
$RV_i = 7/8 = 0.875$



Profile view of a sink

1.750	21.630	21.630	21.630	21.630	21.630	21.630	21.630	21.630	21.630	21.630	21.500	21.500	21.63
1.750	21.630	21.750	21.630	21.630	21.630	21.630	21.500	21.630	21.630	21.630	21.630	21.630	21.63
1.750	21.630	21.750	21.630	21.630	21.630	21.500	21.500	21.500	21.500	21.500	21.500	21.630	21.63
1.750	21.750	21.750	21.630	21.630	21.630	21.500	21.500	21.630	21.630	21.630	21.630	21.630	21.63
1.630	21.500	21.500	21.500	21.500	21.500	21.500	21.500	21.630	21.630	21.630	21.500	21.630	21.63
1.630	21.630	21.630	21.500	21.500	21.500	21.500	21.630	21.630	21.630	21.500	21.500	21.630	21.63
1.500	21.630	21.500	21.500	21.500	21.630	21.630	21.630	21.630	21.630	21.630	21.630	21.630	21.75
1.500	21.630	21.630	21.630	21.630	21.630	21.500	21.500	21.630	21.630	21.630	21.630	21.630	21.63

Figure 1. Use of AVHRR SST data to identify meso-scale ocean surface features. Right: calculating relative variability in SST (Wang *et al*., 2007). Below: identifying SST (and Chla) "sinks" (Valavanis *et al*., 2005).

As with time-series models, autocorrelation in spatial data is an important issue. Model residuals can be checked for spatial autocorrelation, and if none exists, spatial autocorrelation may be ignored. However, ideally, spatial autocorrelation should be included in the model (Pinheiro and Bates, 2000; Wood, 2004; Zuur *et al*., 2007) and this is possible using the software package R, among others. Latitude and/or longitude may be included in models as explanatory variables, although if these effects interact, results will be difficult to interpret. In any case, spatial trends revealed in this way are essentially unexplained trends.

Many of the published models of environmental variation in squid relate to demersal species, in which temperature, salinity and large-scale phenomena (e.g. NAO) have been shown to influence distribution, abundance and life history parameters. While the role of mesoscale ocean surface features was not analysed in these studies, generally because the relevant data were not available, recent analysis has shown that local variability in SST (presumed to be related to fronts) affects hake distribution in the south west Atlantic (Wang *et al*., 2007).

Pelagic squid such as *D. gigas* tend to occur more remotely from land and are less well known than demersal squid. Less dependent on the substrate, they produce many more eggs, with the eggs being pelagic rather than attached to the seabed, hence strongly dependent on currents for dispersal. They tend to be highly migratory and might be expected to show stronger associations with meso-scale ocean surface features and have more variable abundance patterns.

Useful goals for environmental modelling in *D. gigas* would include improved understanding of egg and paralarval distribution, recruitment success and limits to distribution. Tagging data should facilitate 4-dimensional (in space and time) modelling of individual movement patterns. Useful parallels could be drawn from comparative studies on European ommastrephid species such as *Todarodes sagittatus*.

Acknowledgements

GJP thanks the Workshop organisers for the invitation to attend the workshop and for financial support and the workshop participants for useful discussions on the topic of analysing environmental relationships. Much of the work described in this paper was funded by the European Commission under various collaborative research projects, most recently CEPHSTOCK (Q5CA-2002-00962).

References

Bellido J.M., G.J. Pierce and J. Wang. 2001. Modelling intra-annual variation in abundance of squid *Loligo forbesi* in Scottish waters using generalised additive models. Fisheries Research 52: 23-39.

Caddy J.F. 1983. The cephalopods: factors relevant to their population dynamics and to the assessment and management of stocks. In: J.F. Caddy (Ed.). Advances in assessment of world cephalopod resources. FAO Fisheries Technical Paper 231: 416-449.

Caddy J.F. and J.A. Gulland. 1983. Historical patterns of fish stocks. Marine Policy 7: 267-278.

Coelho M. 1985. Review of the influence of oceanographic factors on cephalopod distribution and life cycles. NAFO Scientific Council Studies 9: 47-57.

Hirzel A.H., V. Helfer and F. Metral. 2001. Assessing habitat-suitability models with a virtual species. Ecological Modelling 145: 111-121.

Moreno A., M. Azevedo, J. Pereira and G.J. Pierce. 2007. Growth strategies in the squid *Loligo vulgaris* from Portuguese waters. Marine Biology Research 3(1): 49-59.

Pierce G.J., N. Bailey, Y. Stratoudakis and A. Newton. 1998. Distribution and abundance of the fished population of *Loligo forbesi* in Scottish waters: analysis of research cruise data. ICES Journal of Marine Science 55: 14-33.

Pierce G.J. and P.R. Boyle. 2003. Empirical modelling of interannual trends in abundance of squid (*Loligo forbesi*) in Scottish waters. Fisheries Research 59: 305-326.

Pierce G.J. and A. Guerra. 1994. Stock assessment methods used for cephalopod fisheries. Fisheries Research 21: 255-285.

Pierce G.J., J. Wang and V.D. Valavanis. 2002. Application of GIS to cephalopod fisheries: workshop report. Bulletin of Marine Science 71: 35-46.

Pierce G.J., A.F. Zuur, J.M. Smith, M.B. Santos, N. Bailey, C.-S. Chen and P.R. Boyle. 2005. Interannual variation in life-cycle characteristics of the veined squid (*Loligo forbesi*) in Scottish (UK) waters. Aquatic Living Resources 18: 327-340.

Pinheiro J. and D.M. Bates. 2000. Mixed Effects Models in S and S-PLUS. Statistics and Computing, Springer-Verlag, New York.

Sacau M., G.J. Pierce, J. Wang, A.I. Arkhipkin, J. Portela, P. Brickle, M.B. Santos and X. Cardoso. 2005. The spatio-temporal pattern of Argentine shortfin squid *Illex argentinus* abundance in the southwest Atlantic. Aquatic Living Resources 18: 361-372.

Sims D.W., M.J. Genner, A.J. Southward and S.J. Hawkins. 2001. Timing of squid migration reflects North Atlantic climate variability. Proceedings of the Royal Society of London, Series B 268: 2607-2611.

Solow A.R. 2002. Fisheries recruitment and the North Atlantic Oscillation. Fisheries Research 54: 295-297.

Valavanis V.D., S. Georgakarakos, A. Kapantagakis, A. Palialexis and I. Katara. 2004b. A GIS environmental modelling approach to essential fish habitat designation. Ecological Modelling 178: 417-427.

Valavanis V.D., S. Georgakarakos, D. Koutsoubas, C. Arvanitidis and J. Haralabous. 2002. Development of a marine information system for cephalopod fisheries in the Greek seas (eastern Mediterranean). Bulletin of Marine Science 71: 867-882.

Valavanis V.D., A. Kapantagakis, I. Katara and A. Palialexis. 2004a. Critical regions: A GIS-based model of marine productivity hotspots. Aquatic Science 66: 139-148.

Valavanis V.D., I. Katara and A. Palialexis. 2005. Marine GIS: Identification of mesoscale oceanic thermal fronts. International Journal of Geographical Information Systems 19: 1131-1147.

Waluda C.M., P.G. Rodhouse, G.P. Podestá, P.N. Trathan and G.J. Pierce. 2001. Surface oceanography of the inferred hatching grounds of *Illex argentinus* (Cephalopoda: Ommastrephidae) and influences on recruitment variability. Marine Biology 139: 671-679.

Waluda C.M., P.N. Trathan and P.G. Rodhouse. 1999. Influence of oceanographic variability on recruitment in the *Illex argentinus* (Cephalopoda: Ommastrephidae) fishery in the South Atlantic. Marine Ecology Progress Series 183: 159-167

Wang J., G.J. Pierce, M. Sacau, J. Portela, M.B. Santos, X. Cardoso and J.M. Bellido. 2007. Remotely sensed local oceanic thermal features and their influence on the distribution of hake (*Merluccius hubbsi*) at the Patagonian Shelf edge in the SW Atlantic. Fisheries Research 83(2-3): 133-144.

Wood S.N. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association 99: 637-686.

Zuur A.F., E.N. Ieno and G.M. Smith. 2007. Analysing ecological data. Springer-Verlag.

Zuur A.F. and G.J. Pierce. 2004. Common trends in Northeast Atlantic squid time series. Netherlands Journal of Sea Research 52: 57-72.