

Theme session G

Beyond correlations: what are suitable methods for describing and testing non-linear spatio-temporal changes, patterns and relationships?**Methodology for analysing long-term variation in harbour seal diet and relationships with fish abundance**

Graham J. Pierce^{1,2}, Elena N. Ieno^{3,1}, M. Begoña Santos^{2,1}, A. Edridge¹, Anatoly A. Saveliev⁴ and Alain F. Zuur^{3,1}

1. Oceanlab, University of Aberdeen, Main Street, Newburgh, Aberdeenshire, AB41 6AA, UK;
2. Instituto Español de Oceanografía, Centro Oceanográfico de Vigo, P.O. Box 1552, 36200, Vigo, Spain;
3. Highland Statistics Ltd, 6 Laverock Road, Newburgh, Ellon, Aberdeenshire, AB41 6FN, UK
4. Faculty of Geography and Ecology, Kazan University, 18 Kremlevskaja street, Kazan, 420008, Russia

Long time series of dietary data can potentially provide information on how predator diets respond to changing prey abundance, e.g. due to overfishing and/or climate change. Here we analyse data on the summer diet of harbour seals (*Phoca vitulina*) in Orkney (NE Scotland, UK), based on opportunistic sampling of seal faeces at haul-out sites on the island of Eynhallow during 1986-2006. The main component of the diet in summer is sandeels (Ammodytidae, mainly *Ammodytes* spp.). Aside from a patchy distribution of samples and involvement of various different researchers over the 20 year time period, the data present several statistical challenges. Thus, individual otoliths in a faecal sample may not represent independent samples and the statistical distributions of prey numbers include an excess of zero values. Linear mixed effects modelling was applied to model temporal patterns in the size of sandeels eaten (using otolith length as a proxy of fish size). In addition, zero-inflated Poisson GAM was applied to model trends in the numerical importance of sandeels. We discuss the possible biological significance of the trends identified in this data set.

Keywords: long-term trends, monitoring, seal, diet, zero-inflated Poisson, GAM, mixed models

Contact author: Graham Pierce, Tel. +44 1224 272459, e-mail: g.j.pierce@abdn.ac.uk

INTRODUCTION

There are several seal haul-out sites on the small uninhabited island of Eynhallow (Orkney, see Fig. 1), used mainly by harbour seals (*Phoca vitulina*) in summer and grey seals (*Halichoerus grypus*) in winter, although the two species are often both present. The island has been the focus of a long-term study on fulmars (which started in the 1950s). The present study originated as part of a wider survey of seal diet in Scotland during 1986-89 (see Pierce *et al.*, 1990, 1991a, b, c, d; Thompson *et al.*, 1991) and a follow-up study in 1993-94 (see Pierce & Santos, 2003). Subsequent sampling has been mainly opportunistic, running along side the fulmar project and forming the basis of several MSc projects.

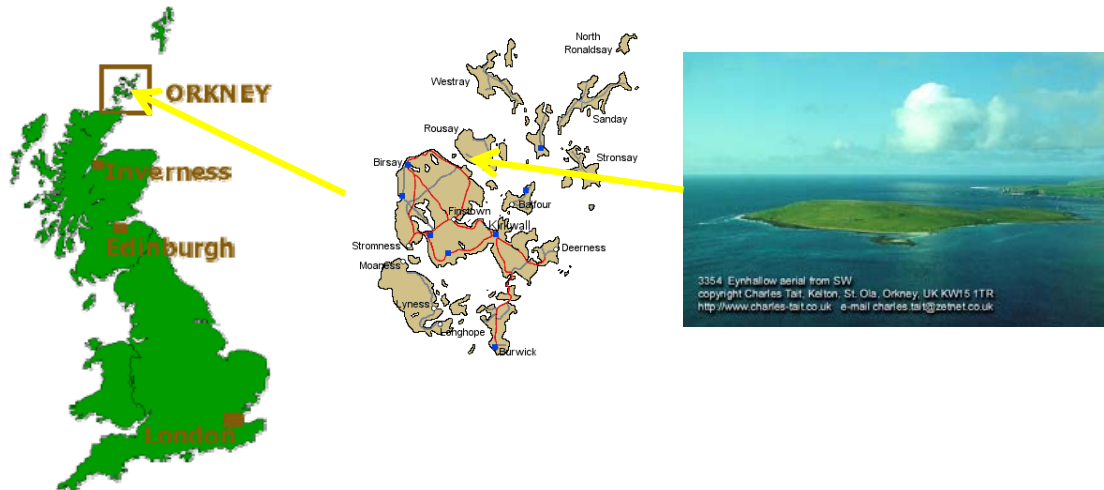


Figure 1. Map of Scotland showing the location of the sampling.

We focus on data collected during May to August in 1986-88, 1993-96, 1998, 2002-03 and 2005-06. The summer diet of harbour seals in Orkney is dominated by sandeels (Ammodytidae) (Pierce *et al.*, 1990) and preliminary exploration of the whole long-term data set confirms that this continues to be the case. Hence, the main biological focus of the present analysis is on whether long-term trends can be discerned in the size and numbers of sandeels eaten, and whether there is evidence that this is related to the abundance of sandeels, and of harbour seals, in the area.

The analysis of sandeel size is a relatively straightforward mixed modelling problem. We account for the likely non-independence of the size of individual sandeel otoliths in the same scat by treating scat ID as a random effect. The analysis of sandeel numbers is a potentially more difficult problem. In principle count data could follow a Poisson or negative binomial distribution but it turns out that the distribution is overdispersed, with an excess of zero values. These data are therefore used in an exploration of the application of zero-inflated models.

Zero-inflated models assume that zero values can be divided into “true” and “false” zeros. If only “true” zeros are present, an alternative modelling approach, hurdle models, can be used (Zuu *et al.*, 2009; Martin *et al.*, 2010). False zeros can arise for various reasons, variously categorised as structural errors, design errors, observer errors, “animal” errors and “naughty zeros” (Austin & Meyers, 1996; Kuhnert *et al.*, 2005; Martin *et al.*, 2005). In the present context, a true zero occurs when a scat did not contain sandeel otoliths and the seal had not eaten sandeels. A false zero arises if a seal ate sandeels but no sandeel otoliths were found in

its scats. This could arise, for example, because the otoliths were present only in unsampled scats, were fully digested, or were present but lost during processing or mis-identified. The latter options are unlikely as sieve mesh size is chosen to avoid loss of small otoliths, otoliths are relatively easy to pick out from other material in scats even for inexperienced observers, and sandeel otoliths are neither difficult to identify nor among the most digestible otoliths (Härkönen, 1986). However, the first option is quite plausible. If true and false zeros are present, so-called mixture models can be applied and this is what we will do here.

METHODS

Sample collection and processing

Scats were collected from the haul-out sites around the shore of Eynhallow and sometimes on offshore skerries. The schedule of sample collection and numbers of samples are shown in Table 1. If seals were present, they were counted and identified to species. Each scat (distinguished from adjacent scats according to location, colour and consistency) was bagged, labelled, transferred to the laboratory and stored at -20°C. Thawed scats were sieved (smallest mesh 0.355 mm) and all identifiable prey remains extracted (i.e. fish otoliths, jaw bones, vertebrae and other skeletal elements, crustacean exoskeleton, polychaete jaws, mollusc shells and mandibles). Remains were identified to the lowest possible taxon. Otoliths and cephalopod mandibles were counted and measured to estimate original prey size. In the present analysis, we focus on remains of sandeels recovered during May-August each year.

Table 1. Numbers of samples available for analysis, by year and month of collection

Year	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Nov	Dec
1986						1	53			
1987		16		18	58		24			
1988	3	4			15	16			19	
1993				11		90				
1994		55						98		
1995		16	1	17	9		65			23
1996				22						
1998					21		25			
2002				19			62			
2003							21			
2005							39			
2006							24			

Analysis of trends in sandeel length

Sandeel otoliths were divided into length classes approximately equivalent to the size range of 0-1 and 2-3 year old *Ammodytes marinus* (Boulcott et al., 2007). Linear mixed effects modelling (Zuur et al., 2009) was used to model otolith length (OL) as a function of time

(expressed as days since the first sampling period); scat ID was used as random intercept. Models of the following form were fitted to data on small and larger sandeels:

$$OL_{is} = \alpha + \beta \times Time_{is} + a_s + \varepsilon_{is}$$

$$a_s \sim N(0, \sigma_{scat}^2) \quad \text{and} \quad \varepsilon_{is} \sim N(0, \sigma^2)$$

Since many different assistants contributed to the data collection, in years when several assistants contributed to the data, results from each were compared. Comparisons confirmed that there were no significant observer differences.

Analysis of trends in sandeel numbers

Because fish have two otoliths we use the variable “Sandeels”, which is the number of otoliths divided by two (and rounded, if not an integer). Data exploration followed Zuur *et al.* (2010). The main explanatory variable for modelling is time. We used the variable **Day** which represents the numbers of days since 1 January 1986, the year in which the first samples were taken. We expect there to also be seasonal variation the diet, not least because in the summer mainly harbour seals are present and in the winter grey seals. Consequently we only use the May – August data, which should largely eliminate seasonal variation.

The first model fitted was a Poisson GLM. Results indicated overdispersion of the response variable. The degree of overdispersion suggested that quasi-Poisson would also be unsuitable. This leads to consideration of several alternative approaches (Fig. 2).

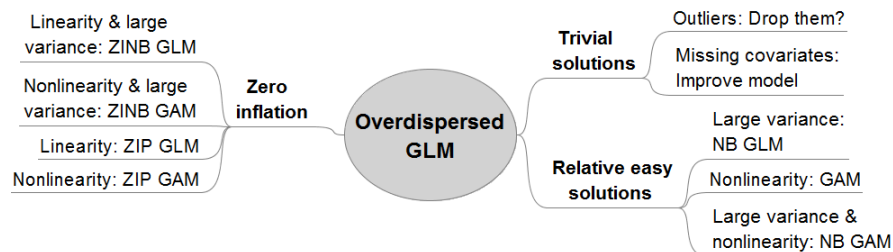


Figure 2. Flowchart showing options when a (Poisson) GLM model shows overdispersion, a quasi-Poisson model is not an option and we have a mixture of true and false zeros. NB = negative binomial distribution.

A negative binomial GLM showed only very slight overdispersion but there were trends in the residuals, indicating a need to consider GAM. Both Poisson and negative binomial GAMs indicated overdispersion, confirming the need to consider zero-inflated models. For the present data set, it can be argued that the zeros consist of true zeros and false zeros. A false zero may be due to otoliths being located in unsampled scats or (less likely) being totally digested. If a count data set consists of true and false zeros, then zero-inflated Poisson (ZIP) or zero-inflated negative binomial (ZINB) models should be applied, whereas data sets with only true zeros can be analysed with hurdle models (Zuur *et al.* 2009). Hence, the sandeel otoliths data set should be analysed with zero-inflated GLMs or zero inflated GAMs, and not with hurdle models.

The probability distribution for a ZIP model is:

$$\begin{aligned}\Pr(Sandeels_{is} = 0) &= \pi_{is} + (1 - \pi_{is}) \times e^{-\mu_{is}} \\ \Pr(Sandeels_{is} = y_{is} \mid y_{is} > 0) &= (1 - \pi_{is}) \times \frac{\mu_{is}^{y_{is}} \times e^{-\mu_{is}}}{y_{is}!}\end{aligned}$$

where: $Sandeels_{is}$ is the number of sandeels in scat i on day s , π_{is} is the probability that $Sandeels_{is}$ is a false zero, which is assumed to be binomially distributed (hence the probability that $Sandeels_{is}$ is not a false zero is $1 - \pi_{is}$). For the second part we use a Poisson distribution with mean μ_{is}

To introduce covariates, as in Poisson GLM, we model the mean μ_{is} of the positive (non-false zero) count data as:

$$\mu_{is} = e^{\alpha + \beta_1 \times X_{is1} + \dots + \beta_q \times X_{isq}}$$

Hence, covariates are used to model the positive counts. For the probability of having a false zero, π_{is} , we use a logistic regression model with only an intercept or one with covariates, i.e.:

$$\pi_{is} = \frac{e^v}{1 + e^v} \quad \text{or} \quad \pi_{is} = \frac{e^{v + \gamma_1 \times Z_{is1} + \dots + \gamma_q \times Z_{isq}}}{1 + e^{v + \gamma_1 \times Z_{is1} + \dots + \gamma_q \times Z_{isq}}}$$

where v is an intercept, Z are the covariates (which may be different to the covariates that influence the positive counts) and γ are regression coefficients.

It is now necessary to formulate the likelihood equation based on the probability functions above and use an optimisation routine to get parameter estimates and standard errors (see Cameron & Trividi, 1998; Hilbe, 2007). The only difference between a ZIP and ZINB is that the Poisson distribution for the count data is replaced by the negative binomial distribution.

We fitted ZIP GAMs using the VGAM package in R. The following ZIP GAM model was applied.) We will call it M2:

$$\begin{aligned}Sandeels_{is} &\sim ZIP(\pi_{is}, \mu_{is}) \\ E(Sandeels_{is}) &= \mu_{is}(1 - \pi_{is}) \quad \text{and} \quad \text{var}(Sandeels_{is}) = (1 - \pi_{is}) \times (\mu_{is} + \pi_{is}\mu_{is}^2) \\ \log(\mu_{is}) &= \alpha + f_1(Day_s) \\ \text{logit}(\pi_{is}) &= v + f_2(Day_s)\end{aligned}$$

We also applied a model in which no covariates were used in the logistic link function; just an intercept. We called it M1.

RESULTS

Diet composition

Sandeels (*Ammodytidae*) were the main prey of seals sampled at Eynhallow in summer in all years (forming between 94% and 100% of the diet by number). Other frequently occurring prey were whiting (*Merlangius merlangus*), haddock (*Melanogrammus aeglefinus*), saithe and pollack (*Pollachius* spp.), *Trisopterus* spp., cod (*Gadus morhua*), herring (*Clupea harengus*), octopus (*Eledone cirrhosa*) and scad (*Trachurus trachurus*).

Using the raw data, there was a non-significant downward trend in the numerical proportion of sandeels in the diet over the course of the study period (Figure 3).

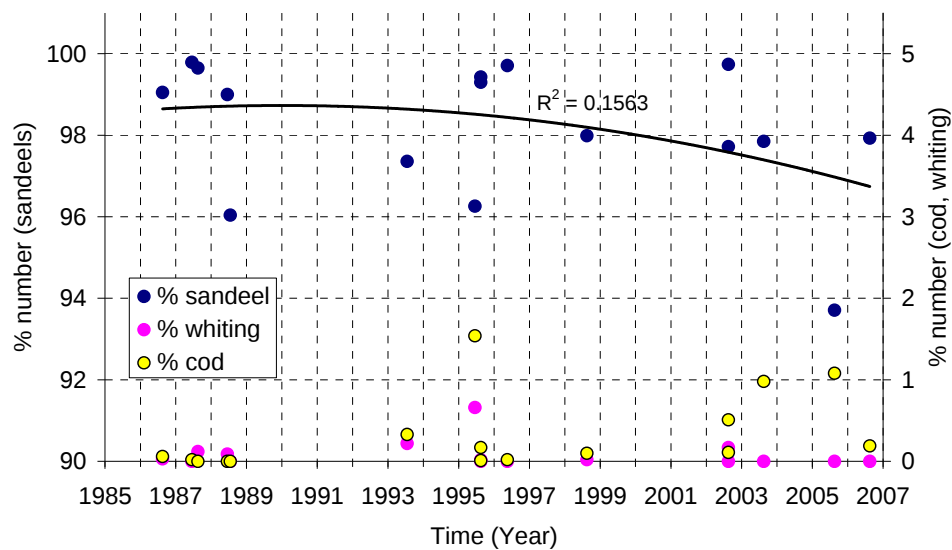


Figure 3. Trends in summer seal diet composition, 1986-2006: percentage numerical importance of sandeels (left Y axis) and fitted polynomial curve; percentage numerical importance of cod and whiting (right Y axis). Each point represents one sample set (when sample sets were available for more than one month in a year they were included separately).

Trends in sandeel length

The largest sandeel otoliths, which belong to *Hyperoplus lanceolatus*, were excluded from this analysis. Results from the linear mixed effects mode showed that there was a significant increase in the average size of sandeels eaten over the study period, an increase of around 7% for the small size classes ($p < 0.001$) and around 5% for larger classes ($p < 0.01$) (Figure 4).

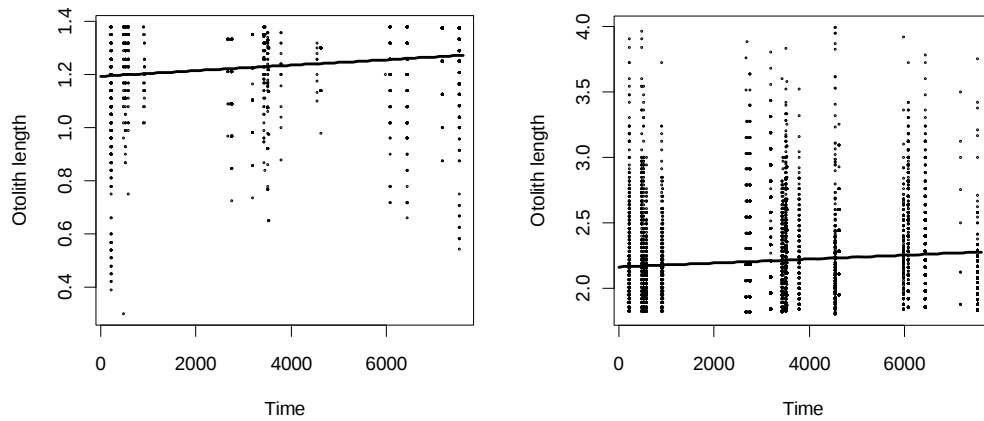


Figure 4. Observed sandeel otolith length versus time for each size class. Left panel: age class 0-1 and right panel: age class 2-3. In both cases, the line is the fitted curve from the mixed model.

Trends in sandeel numbers

The number of sandeels per scat varied from zero to over 1500, as seen in the scatterplot below (Figure 5). The plot also makes clear the patchiness of the sampling, which did not take place every year and did not always take place in the same month of sampled years.

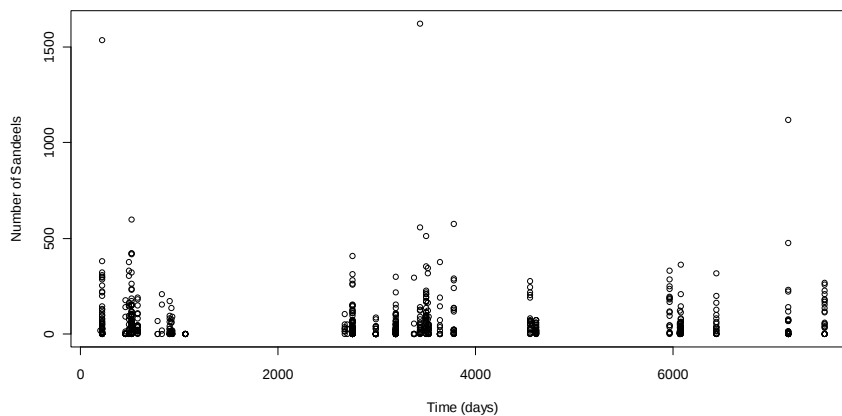


Figure 5. Number of sandeels plotted versus time in days. Day 1 corresponds to 1 January 1986.

We have fitted two ZIP models; in **M1**, only the intercept ν is used and the smoother $f_2(\text{Day})$ is omitted. In **M2** there is a smoother in the logistic link function. More complicated ZIP models may not converge, in which case using an intercept-only model for the logistic part of the ZIP is an option. In this case there were no convergence problems and the AIC values (71152.16 and 71125.02, respectively) indicate that model **M2** is better.

The overdispersion (1.374) was sufficiently close to 1 to be considered acceptable. The estimated smoothers for the ZIP GAM **M2** are presented in Figure 6.

Smoother, $f_1(\text{Day})$, represents the underlying temporal trend (on the scale of the link function) for the sandeels while filtering out the false zeros. This trend implies that the number of sandeels was highest at the start and end of the study period, and also high around 1996-98. The shape of the second smoother, $f_2(\text{Day})$, which relates to false zeros, depicts a decline in the number of false zeros between 1993 and 1998 (approx. days 3000-4500) and a subsequent increase in the most recent years.

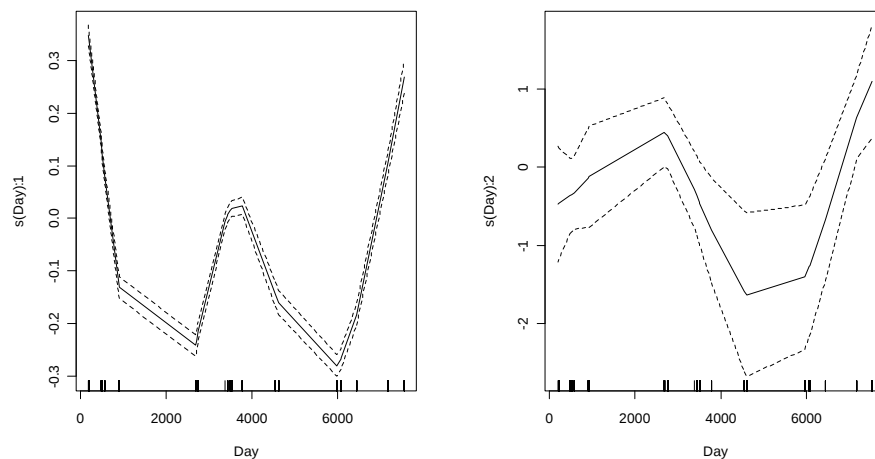


Figure 6. Estimated smoother for the ZIP GAM. Left panel: smoother $f_1(\text{Day})$ for the log link function, related to count process (counts with true zeros). Right panel: smoother $f_2(\text{Day})$ for the logistic link function, related to false zeros versus all other data.

DISCUSSION

Sandeels are clearly a prey item of major importance in harbour seal diets and their consistently high numerical abundance in the diet in Orkney implies that they could be considered as preferred prey. Results also suggest an increase in the size of sandeels eaten by seals in the summer as well as a long-term trend in the number of sandeels per scat.

One obvious causal factor to consider is prey abundance. Over the study period, sandeel (*A. marinus*) abundance in the North Sea has shown two peaks in abundance, in 1987 and in 1998 (Figure 7). Since 1998, the stock has declined markedly, although trends around Shetland (north of Orkney and treated separately by ICES) are unknown (ICES, 2009).

Variation in sandeel numbers in the scats to some extent mirrors the trend in SSB up to 2002 (compare Figures 6 and 7) but after this the incidence of sandeels in the scats rises again sharply, a trend not evident in the sandeel abundance data. We have not attempted to formally model this relationship although, in principle, annual sandeel SSB could be included as an explanatory variable in place of time - although it should be noted that we have only 12 unique values for SSB (for the twelve years out of 20 in which we have dietary data).

Another factor which may be expected to affect seal diet is competition, both intra- and inter-specific. Potential competitors of harbour seals for sandeels include grey seals, harbour porpoises, various seabirds and some predatory fish. In the UK as a whole, grey seal population numbers are probably stabilising after a long period (perhaps a century) of increasing abundance. However, grey seal pup production in Orkney has fluctuated widely in recent years (SCOS, 2007).

Harbour seal numbers in Orkney peaked at ca. 8500 in 1997 but declined to ca. 3380 in 2007 (Lonergan *et al.*, 2007; Duck *et al.*, 2008). Available evidence seems to point to elevated post-weaning pup mortality in Orkney as a reason for the decline in seals (Duck *et al.*, 2008). Poor pup condition may thus be linked to falling sandeel abundance. Lack of sandeels has also been linked with reduced breeding success and adult survival in several North Sea seabird populations (e.g. Frederiksen *et al.*, 2004).

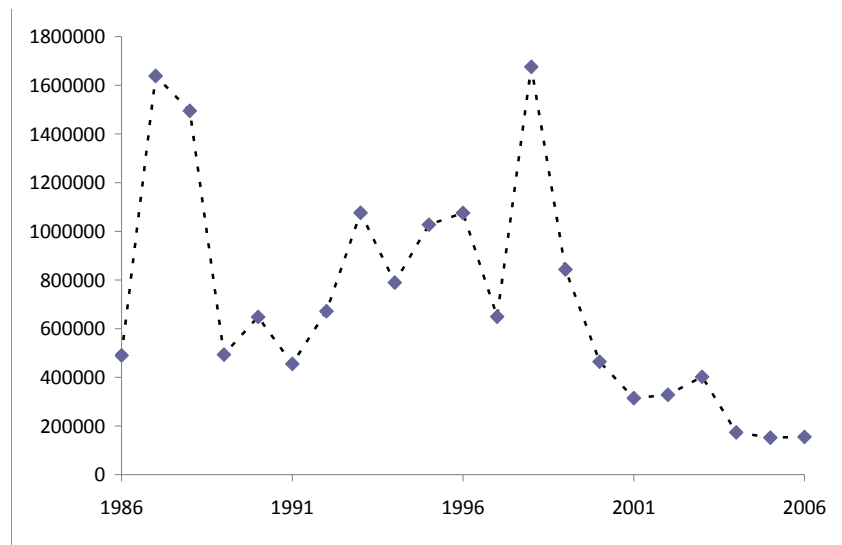


Figure 7. North Sea sandeel spawning stock biomass (tons), 1986-2006 (data from ICES, 2009).

The links between seal diet, sandeel abundance and seal abundance remain speculative. Before engaging in any such speculation, some caveats must be borne in mind. Firstly, the uneven distribution of sampling effort inevitably gives rise to potential biases. Sample size itself is probably not an issue: thus it can be shown (results not presented here) that the size of sandeels eaten could have been adequately estimated from many fewer samples. Secondly, two seal species use the haul-out sites and although there is consistent seasonality it is difficult to rule out the possibility that grey seals make a variable contribution to the sample collected. Grey seals, being larger than harbour seals as adults, would be expected to take larger sandeels. Thirdly, the number of sandeels in seal scats can relate to various aspects of feeding behaviour and diet of seals: essentially anything which affects digestibility and passage rate of food can influence the size of scats and the number of fish represented in them. Finally, the estimates of sandeel abundance cited here relate to only one of the species of sandeels eaten by seals (albeit that which is probably most commonly eaten) and are for the North Sea as a whole rather than Orkney *per se*.

The trend in numbers of sandeels in scats between 1986 and 2002, in so far as it can be reconstructed from the incomplete series, appears as though it may be related to sandeel abundance. The general increase in size of sandeels eaten over time is less easy to explain, but could relate to the lower abundance of small sandeels (recruits) in later years. In addition, the recent reduction in harbour seal numbers in Orkney could have led to reduced intraspecific competition, potentially allowing the remaining seals to take more and larger sandeels.

While these interpretations remain speculative, we can say that the linear mixed effect model and zero-inflated GAM were useful tools in the present study and represent powerful tools to reveal patterns and trends in difficult datasets.

ACKNOWLEDGEMENTS

The original 1986-89 study was funded by a grant from the (then) Department of Agricultural Fisheries for Scotland to the late Peter Boyle. Sampling during the 1990s was funded by a Stevenson Scholarship and two EU Marie Curie awards to MBS and a grant from the Carnegie Trust for the Universities of Scotland to GJP. Subsequent work has been funded from a variety of sources. GJP is currently funded by EC ANIMATE project (MEXC-CT-2006-042337). We thank the late Peter Boyle and George Dunnet, Phil Hammond, Mike Harris, Tony Hawkins, John Hislop and John Prime for support and advice during the original DAFS project, Tommy & Colin Mainland for providing transport to the island, and the numerous people who have assisted with sample collection, processing and analysis over the years, including Steve Adams, Antony Bishop, Ron Bisset, Adam Bratt, Ken Brown, Lois Calder, Avril Conacher, Tom Craig, Ian Diack, Jane Diack, Paul Doyle, Gaelle Ecobichon, Bill Edwards, Sara Frears, Beatriz Gimenez, Maxine Grisley, Mike Harding, Lee Hastie, Hassan Heba, Gema Hernandez, Steve Hoskin, Ishbel Hunter, Nick Jacob, Amanda Kear, Verena Keller, Linda Key, Kit Kovak, Steven Land, Evelyn Lauda, Sue Lewis, Andy Lucas, Gayle Mackay, Alasdair Miller, Nicola Murray, Cecile Nielsen, Fiona Read, Annette Ross, Gill Robertson, Nikki Rowberry, Angeles Santos, Laura Seaton, Arnor Sigfusson, Gabrielle Stowasser, Dominic Tollit, Sawai Wanghongsa, Jon Watt, Gerard Wijnsma, and Cynthia Yau.

REFERENCES

- Austin, M.P. & Meyers, A.J., 1996. Current approaches to modelling the environmental niche of Eucalypts: implications for management of forest biodiversity. *Forest Ecology and Management* 85, 95–106.
- Boulcott, P., Wright, P.J., Gibb, F.M., Jensen, H. & Gibb, I.M., 2007. Regional variation in maturation of sandeels in the North Sea. *ICES Journal of Marine Science*, 64: 369–376.
- Cameron, A.C. & Trivedi, P.K., 1998. *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, UK.
- Duck, C.D., Thompson D. & Mackey, B.L., 2008. The status of British common seal populations in 2007. SCOS Briefing paper 08/03. In: *Scientific Advice on Matters Related to the Management of Seal Populations: 2008*. Special Committee on Seals, Natural Environment Research Council.
- Frederiksen, M., Wanless, S., Harris, M.P., Rothery, P. & Wilson, L.J., 2004 The role of industrial fisheries and oceanographic change in the decline of North Sea black-legged kittiwakes. *Journal of Applied Ecology* 41, 1129–1139.
- Härkönen T., 1986. *Guide to the otoliths of the bony fishes of the northeast Atlantic*. Danbiu ApS, Hellerup, Denmark, 256 pp.
- Hilbe, J.M., 2007. *Negative Binomial Regression*. Cambridge University Press, Cambridge, UK.
- ICES. 2009. *ICES Advice*. International Council for the Exploration of the Sea, Copenhagen.

- Kuhnert, P.M., Martin, T.G., Mengersen, K. & Possingham, H.P., 2005. Assessing the impacts of grazing levels on bird density in woodland habitat: a Bayesian approach using expert opinion. *Environmetrics* 16, 1–31.
- Loneragan, M., Duck, C.D., Thompson, D., Mackey, B.L., Cunningham, L. & Boyd, I.L., 2007. Using sparse survey data to investigate the declining abundance of British harbour seals. *Journal of Zoology* 271, 261–269.
- Martin, T.G., Kuhnert, P.M., Mengersen, K. & Possingham, H.P., 2005. The power of expert opinion in ecological models using Bayesian methods: impact of grazing on birds. *Ecological Applications* 15, 266–280.
- Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J. & Possingham, H.P., 2010. *Ecology Letters* 8, 1235–1246/
- Pierce, G.J., Boyle, P.R. & Diack, J.S.W., 1991. Identification of fish otoliths and bones in faeces and digestive tracts of seals. *Journal of Zoology* 224, 320–328.
- Pierce, G.J., Boyle, P.R. & Diack, J.S.W., 1991. Digestive tract contents of seals in Scottish waters: comparison of samples from salmon nets and elsewhere. *Journal of Zoology* 225, 670–676.
- Pierce, G.J., Boyle, P.R. & Thompson, P.M., 1990. Diet selection by seals. In, *Trophic relations in the marine environment*, edited by M. Barnes & R.N. Gibson, Proceedings of the 24th European Marine Biology Symposium, Aberdeen University Press, Aberdeen, pp 222–238.
- Pierce, G.J., Miller, A., Thompson, P.M. & Hislop, J.R.G., 1991. Prey remains in grey seal faeces from the Moray Firth area of Scotland. *Journal of Zoology* 224, 337–341.
- Pierce, G.J. & Santos, M.B., 2003. Diet of harbour seals (*Phoca vitulina*) in Mull and Skye (Inner Hebrides, western Scotland). *Journal of the Marine Biological Association of the United Kingdom* 83, 647–650.
- Pierce, G.J., Thompson, P.M., Miller, A., Diack, J.S.W., Miller, D. & Boyle, P.R., 1991. Seasonal variation in the diet of common seals (*Phoca vitulina*) in the Moray Firth. *Journal of Zoology* 223, 641–652.
- SCOS, 2007. Scientific Advice on Matters Related to the Management of Seal Populations: 2007. Special Committee on Seals, Natural Environment Research Council.
- Thompson, P.M., Pierce, G.J., Hislop, J.R.G., Miller, D. & Diack, J.S.W., 1991. Winter foraging activity by common seals (*Phoca vitulina*) in the Inner Moray Firth. *Journal of Animal Ecology* 60, 283–294.
- Zuur, A. F., Ieno, E. N., Walker, N. J. Saveliev, A. A., and Smith, G. M. 2009. Mixed Effects Models and Extensions in Ecology with R. Springer.

Zuur, A.F., Ieno, E.N. & Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1, 3-14.