

# New perspectives on fish movement: kernel and GAM smoothers applied to a century of tagging data on coastal Atlantic cod

Sigurd Heiberg Espeland<sup>1</sup>, Esben Moland Olsen<sup>2</sup>, Halvor Knutsen<sup>2</sup>,  
Jakob Gjørseter<sup>2</sup>, Didrik Danielssen<sup>2</sup>, Nils Christian Stenseth<sup>1,2,\*</sup>

<sup>1</sup>Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway

<sup>2</sup>Institute of Marine Research Flødevigen, 4817 His, Norway

**ABSTRACT:** Tagging–recovery studies have been used successfully on many occasions, and the data analysed in a variety of ways with respect to fish movement. Here we provide a synoptic account of useful statistical techniques for delimiting area utilization and movement patterns of a population of individuals. We focus specifically on kernel smoothing for population utilization distribution and generalized additive models for temporal movement patterns. Furthermore, we use these 2 techniques to re-analyse data from tagging studies on coastal Atlantic cod *Gadus morhua*, shedding new light on one of the underlying mechanisms determining the population structure of the coastal cod along the Norwegian Skagerrak coast. Cod were found to use small areas. On average 25% of a population used areas smaller than 10 km<sup>2</sup>, while 95% of a population used areas smaller than 160 km<sup>2</sup>. The movements displayed were fjord-limited and predominantly within the bounds of the population's home range. By re-analysing existing data on neutral genetic markers, we show that the cross-distance of the area utilized was smaller than the distance over which pair-wise genetic differences appear.

**KEY WORDS:** Atlantic cod · *Gadus morhua* · Population home range · Population density distribution · Local populations · Stock structure · Kernel model · GAM

—Resale or republication not permitted without written consent of the publisher—

## INTRODUCTION

Migration and dispersal, behaviour collectively termed 'movement,' play important roles in population structuring of marine species. Depletions, extinctions and re-colonisations all depend on the rate of movement within and between populations. Groups of individuals often undertake regular migrations between areas for the purposes of feeding and spawning (Metcalfe 2006). The degree to which individuals travel between different areas represents an important determinant of the degree to which populations are genetically structured. Movement will also determine the geographic unit of individuals responding to population processes. Knowledge of area utilization by fish is

essential, as management practices are often based on geographic areas.

A great variety of tagging–recovery studies, conducted on a range of species, have been used in stock assessment to determine movement and population structure (e.g. Pawson & Jennings 1996, Robichaud & Rose 2004). Although tagging–recovery studies provide little information on fine-scaled behaviour of individuals, they are useful for investigating broad-scale patterns such as area utilization and the structure of movement by groups and populations of individuals.

Several well-acknowledged problems arise when analysing tagging data with respect to movement (Porter & Dooley 1993, Bolle et al. 2005). The number of tags recovered from an area is determined by the

\*Email: n.c.stenseth@bio.uio.no

fishing intensity and the rate at which tags are reported (Pollock et al. 2001). Report rates may vary in time and space depending on the motivation to return tags. The fishing intensity depends on efficiency and effort (Ellis & Wang 2007), both of which may vary over temporal and spatial scales. In well-managed fisheries with well-controlled quotas, these factors may be estimated and it would be possible to apply different weighting schemes to deal with regional differences in fishery intensity (Hunt et al. 1999, Neilson et al. 2006, Stensland et al. 2006). In complex fisheries with a mix of recreational and commercial fisheries, both efficiency and effort depend on several factors such as incentives to fish (economic and recreational), geographical knowledge of the area, economic opportunities to use different gears and weather (Arlinghaus 2006). These factors are functions of both time and space, and may generate heterogeneity in fishing intensity on a fjord/bay scale. Furthermore, the spatial and temporal precision of reported tags depends on the fishers' knowledge of the species, interest in the marine environment and familiarity with the area encompassed by the fishery. All this may add error to the precision in recovery datasets, and, thus, may be difficult to adjust for.

Every tag sighting is one indicator of the density distribution of the tagged fish. Given that tagged fish are representative of the underlying population, the recovery will be a 'draw' from the population distribution at the time of recovery. When making inferences about the 'normal' spatial extent of the population and movement patterns within the area, we aim to include individuals displaying habitual behaviour as determined migration (e.g. cyclic migration between forage and feeding areas). However, we would like to exclude fish emigrating from the population as well as tagged fish that do not belong to the population and display abnormal behaviour. Consequently the area habituated by the population is not necessarily identical to the area bound by a line encompassed by all the recoveries. Overall, these problems suggest that the single observation should be given less emphasis, and more attention should be focused on the overall pattern of the recovery data.

The purpose of this study is 2-fold. First, we provide a brief methodological review on analysing tagging data with concern to movement and population structure. Then building on this we report on the analysis of a century's worth of tagging data on the coastal Atlantic cod *Gadus morhua* from the Norwegian Skagerrak coast (Dahl & Dannevig 1906, Løversen 1946, Danielssen & Gjosaeter 1994). Previous studies of the Skagerrak coastal cod indicated that there are several local populations along the coast (Smith et al. 2002), which also was suggested by Knutsen et al. (2003) who

used genetic analysis on adult cod along the same coastline. The coast receives an inflow of eggs and larvae from the North Sea; however, this influx does not erode the genetic structure of local populations (Knutsen et al. 2004, Stenseth et al. 2006). This would be because either the fish return to the North Sea before spawning, or the fish from the outer areas of the coast have limited genetic exchange with local fjord populations. We analyse the historic tagging studies both with regard to population extent calculated from population utilization distributions, and with respect to the dynamics of population movement using generalized additive models (GAMs). Lastly, we provide a direct link between the density distribution of coastal Atlantic cod estimated from the tagging studies, and the neutral genetic population structure estimated from available microsatellite DNA data (Knutsen et al. 2003, Jorde et al. 2007).

## MATERIALS AND METHODS

**Non-parametric techniques for analysing population movement.** Several methods have been proposed for analysing tagging data with respect to movement. The most common method consists of reporting spatial extent as the geographic regions of grid cells of recovery, with the corresponding number of recoveries (Shimada & Kimura 1994, Lawson & Rose 2000, Moncada et al. 2006, Neilson et al. 2006, Shepherd et al. 2006). Movement is often depicted as arrows between release and recovery positions (Shimada & Kimura 1994, Neilson et al. 2006). Here the spatial extent and resolution will be dependent on the size of the grid cell. Drawing a convex line around the outermost points at the individual level (Mohr 1947) or at the population level (Jorde et al. 2007) is not dependent on an *a priori* choice of cell size, but the resulting area is highly correlated with sample size. Both methods are, therefore, vulnerable to outliers and are poor predictors of the area occupied by the population.

The analysis of several observations from some few individuals, as, for instance, in intensive radio tracking of mammals, has a long history in terrestrial biology (Millspaugh & Marzluff 2001). The term 'Home Range' (Burt 1943) refers to the area traversed by an individual in its normal activities. Likewise the utilization distribution (UD) refers to the relative frequency distribution for the position data points of an animal over a period of time (Van Winkle 1975). These concepts, and methods for analysing them, may in turn be transferred to the population level (Ford & Krumme 1979). As observations of an individual are a sample from the individual's utilization distribution, so the recoveries are a sample from the population's density distribution.

Treating the population as analogous to the individual, we may apply the same analysis techniques at the population level. We may, for instance, estimate the population home range as the area traversed by the individuals in the population in their normal activities excluding random migration.

The most common method for creating a utilization density distribution for individuals is through kernel density estimation (Worton 1987, 1989, Seaman & Powell 1996), a method that has only recently been applied to marine species (Parra 2006, Righton et al. 2007, Svedäng et al. 2007). This method consists of substituting every recovery with a probability distribution (a kernel). By summarising the individual probability distributions we are able to create the overall probability distribution for capturing a tagged individual. Given a regional homogenous fishery, this also provides the utilization distribution for the population over the time period of the study. Different isolines will encompass different percentages of the mass of the distribution and may be described by the value  $\alpha$ , being the proportion of the total distribution. By considering the  $\alpha = 0.95$  area as the population home range, or the extent of the population, we exclude some of the dispersers as well as accounting for noise in the reported observations.

A variety of models have been used to evaluate the movement pattern of populations from tagging data (Porter & Dooley 1993, Shimada & Kimura 1994, Righton et al. 2007). Dealing with individuals moving within a population, we rarely have good *a priori* grounds for assuming a linear relationship between covariates like days at liberty and distance moved. Here we suggest applying a non-parametric additive regression method, the GAM. These methods do not assume linearity throughout the data, instead they fit  $k+1$  piecewise polynomials (with continuous first and second derivatives at the knots, the points where the polynomials are joined) to the data in  $k+1$  intervals. Given the constraint that the smoother is linear at the boundaries, we are able to represent the relationship with a natural cubic spline (Hastie & Tibshirani 1990).

Both GAMs and kernel analysis are non-parametric smoothing techniques. Consequently they are largely dependent on the degree of smoothing (Hastie et al. 2001). In kernel models the smoothing represents the width of the kernel. If this width is too narrow, the utilization distribution is reduced to

small areas around the discrete observations. If the smoother is too large, the form and extent of the utilization distribution is determined by the kernel and the positions only centre the location of the dome in space. In the case of the GAMs, the smoothing parameter relates to the number of knots linking the splines together. A high number of knots may result in overfitting, while no knots will simply give the least square linear fit to the data (Hastie et al. 2001).

**Tagging data.** The data analyzed are from 7 different tagging studies on wild-caught Atlantic cod from the Norwegian Skagerrak coast undertaken between 1905 and 2005 (Table 1, Fig. 1). All studies were performed on fish caught immediately prior to tagging, and released in approximately the same areas as caught. The recoveries were based on reports from commercial and recreational fisheries motivated by a reward. At any given time the fishing fleet would be a complex mix of a large number of recreational as well as commercial fishermen. During the periods between these studies there were considerable changes in social structure and incentives to fish, which indirectly affected fishing intensity. Due to the short duration of the studies (<6 yr; average, ~250 d), these effects were assumed to be constant within each study. Spatially, there was probably a trend towards decreased fishing intensity with distance from shore. In addition, there may have been low magnitude heterogeneity on the bay/fjord scale along the coast, but no large regional trends appeared in any study (Hallenstvedt &

Table 1. *Gadus morhua*. Data sets. Tagging of Skagerrak coastal Atlantic cod in the Risør and Flødevigen areas, showing which releases are grouped together and considered a discrete study unit (Study), the number of fish released (Rel), the number of fish recovered (Recov), the proportion recoveries (Ratio), and the length range (cm) and mean length (cm) for all recovered fish. Source indicates the references where the studies are described as follows: D&D = Dahl & Dannevig (1906), L = Løversen (1946), D&G = Danielssen & Gjøsaeter (1994), O = E. M. Olsen & S. H. Espeland (unpubl.)

Study Area	Year	Month	Rel	Recov	Ratio	Length range	Mean length	Source
1 Risør	1904	8–9	76	7	0.09	22–29	25	D&D
		4–5	45	6	0.13	38–45	38	
	1905	4	583	84	0.14	23–52	33	
2 Flødevigen	1937	1	80	43	0.54	24–43	32	L
		5	100	51	0.51	25–42	35	
		12	88	57	0.65	12–49	33	
3 Risør	1939	5	200	128	0.68	24–35	29	
4 Risør	1970	12	606	273	0.45	26.5–57	38	D&G
		5	234	51	0.22	33–46	39	
5 Risør	1988	12	1387	404	0.29	22.5–64.5	36	D&G
		12	237	51	0.22	29–52	40	
6 Flødevigen	2005	4–6	898	293	0.33	25–70.5	39	O
7 Risør			743	231	0.31	25–80	41	

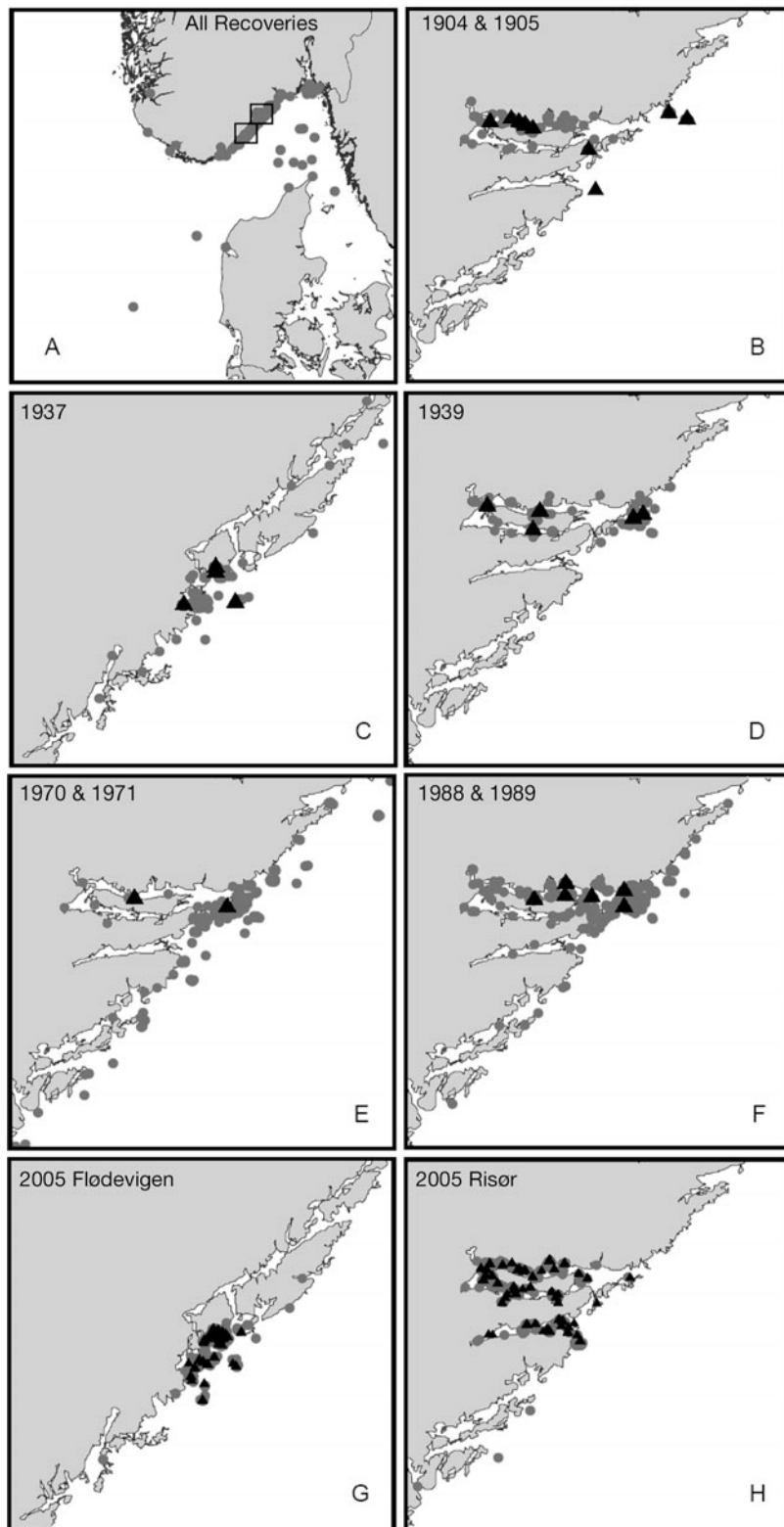


Fig. 1. *Gadus morhua*. Recoveries of tagged Skagerrak coastal Atlantic cod. (A) All recoveries throughout all studies. The upper right box in this panel is the Risør area while the lower left is the Flødevigen area. (B–H) Separate panels for each of the 7 studies. ▲: release locations; ●: recovery locations. The last 2 panels (G & H) are the 2005 studies with 1 panel for each region

Wulff 2004). Due to the poor quality, the existing data on historical fishery intensity were not used to weight the recovery data. All data were provided by Institute of Marine Research, Flødevigen, except the 1905 study, which was taken from Dahl & Dannevig (1906). Only recoveries with identifiable locations were used in the analyses (a total of 42 out of 1628 recoveries were at unidentifiable locations). Studies done in the same areas in consecutive years were pooled, but studies done in separate areas were considered as discrete studies, resulting in a total of 7 studies. The precision of the reported recapture location varied largely between fishers. When only a location name was given, a geographical position was assigned as precisely as possible. Usually this was given as the centre of the geographic area reported. In the 2005 dataset a considerable number of recoveries were live recaptures during the following tagging season (2006). This was assumed to be biased towards sedentary fish since the tagging program was undertaken in similar locations both years. Live recaptures were released and, therefore, only the last of the consecutive recaptures were used, treating them as dead recoveries. The live recaptures represented 35 of 293 recoveries in Flødevigen, and 124 of 231 recoveries in Risør. Initially, all analyses were performed using both live recaptures and dead recoveries. For comparison, additional analyses were performed on a subset of the data containing only dead recoveries.

**Tagging data analysis.** To calculate the population density distributions we used the kernel smoothing function in the statistical software R (R Development Core Team 2006) given in the *splancs* package (Rowlingson & Diggle 1993). The distributions follow the form:

$$f(x, y) = (1/nh^2) \sum_{t=1}^n K\{1/n [E(x) - x_t, E(y) - y_t]\}$$

where  $n$  is the number of recoveries, and  $x_t$  and  $y_t$  are the coordinates of the  $t$ -recovery.  $E(x)$  and  $E(y)$  are the expected average recovery positions calculated as the arithmetic means of the

sample. The kernel  $K$  is a symmetrical bivariate tri-cube probability–density function, and  $h$  is the smoothing parameter (describing the width of the kernel,  $K$ ). For each of the studies we created one density distribution from all the recoveries from one region. This produced a total of 7 distributions.

Usually the least square cross validation (LSCV) method is applied to find the unbiased smoothing parameter  $h$  (Blundell et al. 2001, Gitzen & Millspaugh 2003, Gitzen et al. 2006, Horne & Garton 2006). In these datasets the recoveries were overly aggregated, due to the assigned positions in many of the recovery locations. In this case the LSCV method converges towards an infinitely small smoothing parameter. Instead we used the simpler reference smoother  $h_{\text{ref}}$  (Worton 1995), even though it may give a higher smoothing than the optimal unbiased estimator. This smoother is given as:  $h_{\text{ref}} = \sigma n^{-1/4+d}$ , where  $n$  is the number of recoveries,  $\sigma$  is the SD, and  $d$  is the number of dimensions of the data. The smoother in this case provides a biological interpretation as it represents both the spread of fish around a recovery and the error in reported recovery positions. We assumed the same noise effect in all studies; thus, we calculated one smoothing parameter for all released fish, pooled and centred on the same release location. One distribution was then calculated for each of the 7 study units. All distributions were truncated by land.

To study the dynamics of the movement patterns through time, a simple GAM (Hastie & Tibshirani 1990) was applied to the different studies. Movement was expected to be affected by the size of the fish, the number of days at liberty and the distance from open sea. This was to control for the possibility of North Sea cod in the samples. The model is given by the equation:

$$\log(D_{ik}) = \beta_{0k} + s_{1k}(L_{ik}) + s_{2k}(T_{ik}) + \beta_{1k}(S_{ik}) + \varepsilon_{ik}$$

where  $D_{ik}$  is the distance between release and recovery of the  $i$ th fish in the  $k$ th study. The logarithm of  $D_{ik}$  was used to normalize the response, and to avoid taking  $\log(0)$ , all zero distances were given the value of the lowest observed distance (7 m).  $\beta_{0k}$  is an study specific intercept,  $L_{ik}$  is the length of the fish at release,  $T_{ik}$  is the time between release and recovery, and  $S_{ik}$  is the degree of shelter from the open ocean (calculated as the distance between release site and an offshore transect running at a defined distance from the coastline). The degree of shelter ( $S_{ik}$ ) will then give high numbers for fish released deep inside the fjords. The effect of  $S_{ik}$  was considered to be linear (slope  $\beta_{1k}$ ) due to few release locations in several of the datasets. The  $s_{1k}$  and  $s_{2k}$  parameters are specific smoothing curves (natural cubic splines), tailoring the effect of the covariates to the response in the different  $k$ -studies. The number of

knots for the smoothing curves was found using generalized cross validation in the mgcv package in R (Wood 2006). The error term,  $\varepsilon_{ik}$ , is identical and independently normally distributed for all fish and all studies. The models were also applied to the 2005 dataset with only dead recoveries.

**Genetic data analysis.** Knutsen et al. (2003, 2004) and Jorde et al. (2007) sampled 100 adult Atlantic cod, during the spawning seasons of 2000 to 2002, from a range of fjords and sheltered areas along the Skagerak coast. These fish were screened for genetic variation in 10 to 13 microsatellite DNA loci, and used for testing the hypothesis of population structure along the coast. Here we re-analysed genetic data from Knutsen et al. (2003) and Jorde et al. (2007) with a focus on comparing the Risør sample with other samples along the coast (cf. Fig. 5). Risør is the location where we have extensive tagging data, and, thus, was well suited for making comparisons (e.g. population density distribution data versus genetic findings). A population density distribution, unconditioned by land, from the Risør 2005 study was created to compare with the genetic data.

## RESULTS

### Population utilization distribution

The mean core area (area utilized by 25 % of a population,  $\alpha = 0.25$ ) was 8.4 km<sup>2</sup> (3.9 to 15.0 km<sup>2</sup>). The mean population home range area ( $\alpha = 0.95$ ) covered by the populations was 159 km<sup>2</sup> (70 to 335 km<sup>2</sup>, Table 2). The smoothing factor  $h$  calculated for all studies was  $h_{\text{ref}} = 4444$ . The population density distributions were usually unimodal, with the peak at the location of the majority of releases in the given study

Table 2. Kernel modelling. Studies are described in Table 1.  $n$  is the total number of recoveries used to create a density distribution;  $\alpha = 0.95$  is the area contained within the boundary of the 95 % of the mass of the distribution, or population home range;  $\alpha = 0.25$  is the area contained within the boundary of the 25 % of the mass of the distribution or the population core area

Study	Release year	Area	$n$	$\alpha = 0.95$ (km <sup>2</sup> )	$\alpha = 0.25$ (km <sup>2</sup> )
1	1904 & 1905	Risør	96	117	4.3
2	1937	Flødevigen	139	127	7.2
3	1939	Risør	124	75	10.6
4	1970 & 1971	Risør	312	335	15.1
5	1988 & 1989	Risør	444	158	10.7
6	2005	Flødevigen	289	70	7.1
7	2005	Risør	231	231	3.9



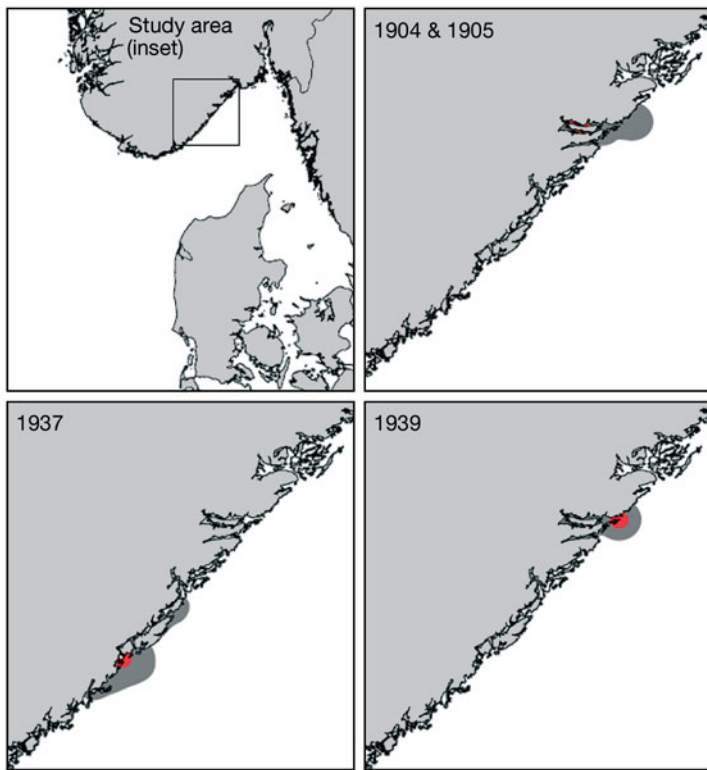


Fig. 2. *Gadus morhua*. Population home range areas from the utilization distributions for the groups of coastal Atlantic cod from 3 different study periods (1904–1905, 1937, 1939) (truncated by land). The top left panel shows the relative location of the Skagerrak area and coastline (inset) referred to in the other panels. The grey areas are the  $\alpha = 0.95$  areas, while the inner red areas represent  $\alpha = 0.25$  of the distribution

(Figs. 2 & 3). There was a non-significant weak correlation between population home range and the corresponding number of recoveries ( $r = 0.26$ ,  $p = 0.57$ ).

### Temporal movement patterns

The GAMs all had relatively low  $R^2$  (0.04 to 0.24) and large residuals appearing shortly after release (Table 3, Fig. 4). In most of the GAMs, the number of days a cod was at liberty had a significant effect on the distance between release and recovery positions (Table 3, Fig. 4). In the 1904 and 1905 datasets, the distances increased over the course of the first 70 d after release. This was followed by a period of no recoveries, then 20% of the recoveries were made after 128 d, but these were closer to the intercept distance from the point of release. In 1937 there was also an increase in distance over time, although at a somewhat

slower rate, but this increase stabilises after approximately 400 d. In 1939, the effect was nearly linear and non-significant and could be removed from the model for an improved fit (Table 3). In the studies undertaken during the 1970s, the effect of number of days at liberty is non-significant, although there is an improved fit of the model when keeping the term. The effect describes a stable situation until about 700 d after release, and then a slow increase in distances. In the 1988 and 1989 datasets, the  $T_{ik}$  effect resembles the 1970 situation, but is significant due to slowly increased distances after 600 d. In both of the 2005 studies, the fish were recovered at increasing distances with time after tagging, reaching a peak distance approximately 220 d after release. This corresponds to December through January. This peak is more pronounced in the Risør sample. A year after tagging (300 to 400 d) the recoveries were again at

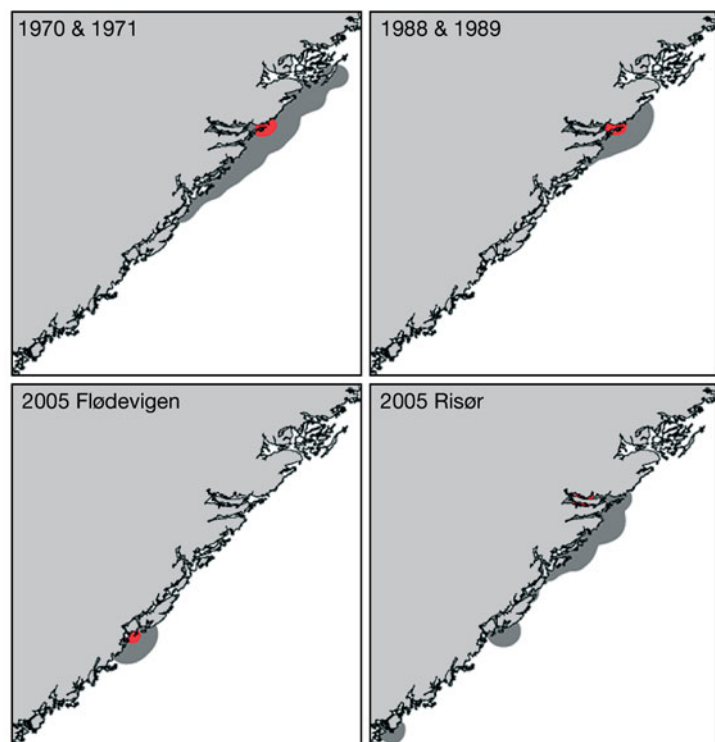


Fig. 3. *Gadus morhua*. Population home range areas for the most recent studies (1970–1971, 1988–1989, 2005; see Fig. 2 for relative location of the Skagerrak area). The 2 lower panels show the 2005 studies with 1 panel for fish released in each area. The dark grey areas are the  $\alpha = 0.95$  areas, while the inner red areas represent  $\alpha = 0.25$  of the distribution

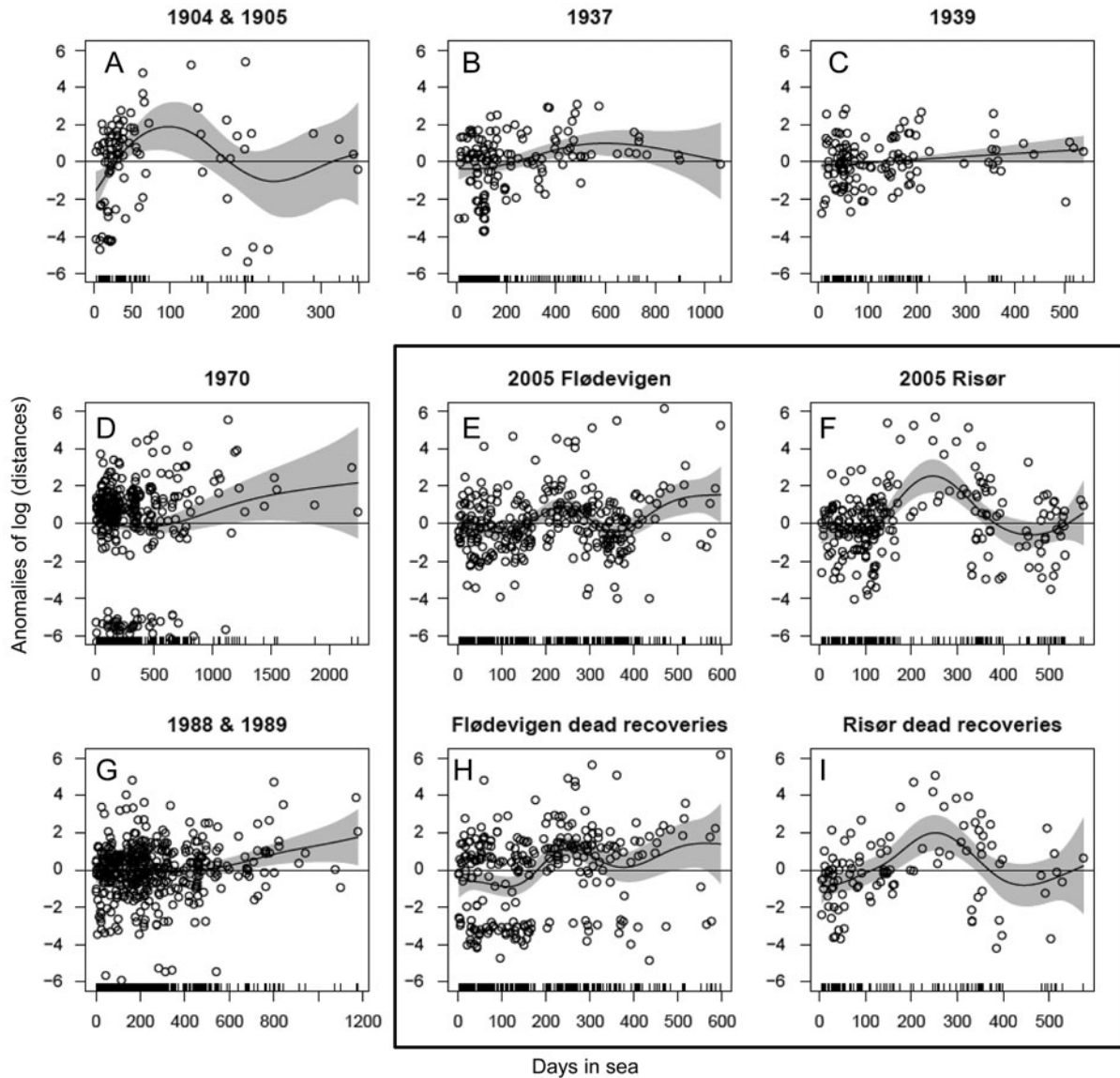


Fig. 4. *Gadus morhua*. Effect of the number of days a tagged coastal cod has been at liberty on the distance moved. The black border around panels E, F, H and I indicates the 2005 studies where the top 2 panels (E, F) depict the effect of using both live and dead recoveries, while the bottom panels (H, I) represent the effect of using only dead recoveries

median distances. This effect is apparent from both the total samples and the models containing only the dead recoveries, although the live recaptures amplify the oscillation pattern.

In all but the first 2 studies (1904 to 1905 and 1937) there is a significant effect of fish length at tagging on the distance of recapture from the point of release (Table 3). In most cases this relationship is linear or nearly linear, and represents an increase in distance with size. Therefore, large fish tend to be recovered at more distant locations, whereas fish that were small at the time of tagging tend to remain close to the point of release. Due to the poor quality of the recovery information on length, the effect is not corrected for

growth. A fish tagged as a juvenile, but recaptured much later as an adult, would enter the model with a small tagging length.

In 1904 and 1905, there is a significant positive effect on the distances moved when going towards the coast from the open sea (Table 3). This effect corresponds to a movement of  $\sim 1$  m per meter of distance from the open sea (defined as distance from the offshore transect). In the Risør area the most sheltered release locations are found  $\sim 16$  km inshore from the open sea. In the 1937 study from Flødevigen this relationship is reversed, meaning that fish from more exposed locations tended to move greater distances. In the studies undertaken during 1939, 1970 and 1971 at Risør, and

Table 3. Generalized additive models (GAMs). The columns indicate the 7 different study units. R-sq and GCV are the r-square and generalized cross validation, respectively, of the models containing the terms length, days at liberty and degree of shelter. Intercept is the log(distance) in meters moved when covariates are centred on their mean value. The p-values indicate the probability that the null hypothesis is correct (the true effect of the covariate is zero). The  $\Delta$ GCV is the change to GCV mediated when removing the term from the model (negative numbers indicate that the GCV decreases when the term is removed; thus, the term should be removed for improved fit). The Edf is the estimated degrees of freedom, or the wiggleness of the function (e.g. Edf of 1 indicates a linear relationship). The degree of shelter, as given by the distance to open sea in meters, is linear by default; the estimate for the slope is given. Significance: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

Study unit	1904 & 1905	1937	1939	1970 & 1971	1988	2005 (Flødevigen)	2005 (Risør)
R-sq	0.235	0.177	0.044	0.042	0.106	0.129	0.217
GCV	5.336	2.117	1.448	6.739	2.100	2.299	2.648
Intercept	4.377	8.548	6.573	7.779	7.192	4.703	6.026
<b>Days at liberty</b>							
p-value	**	**	0.19	0.15	**	***	***
$\Delta$ GCV	0.607	0.163	0.003	0.059	0.047	0.250	0.502
Edf	3.624	3	1.041	2.541	3.920	6.298	5.951
<b>Length</b>							
p-value	0.09	0.75	0.03*	**	***	0.03*	0.02*
$\Delta$ GCV	0.031	-0.030	0.30	0.15	0.15	0.05	0.07
Edf	1	1	1	1	3.138	1.558	4.287
<b>Shelter</b>							
p-value	**	***	0.43	0.47	0.01*	0.91	0.03*
$\Delta$ GCV	0.553	0.167	-0.016	-0.033	0.019	-0.016	0.021
Estimate	1.5e-04	-3.2e-04	1.6e-05	-2.6e-05	5.1e-05	-1.0e-05	-5.4e-05

during 2005 at Flødevigen, the effect of shelter was not significant. In 1988 there was a trend similar to that of the 1904 to 1905 study, while in the 2005 study from Risør this pattern is reversed.

### Comparing area utilization and genetics

An untruncated population utilization distribution of the 2005 recoveries from Risør was compared to the pair-wise  $F_{ST}$  values between Risør and other locations (data from Knutsen et al. 2003, Jorde et al. 2007) to observe the link between movement and genetic differences (Fig. 5). No genetic differences were recorded within a few tens of kilometers of Risør; the Tvedestrand (TV) individuals were genetically similar to those from Risør. However, at only slightly greater distances (in both direc-

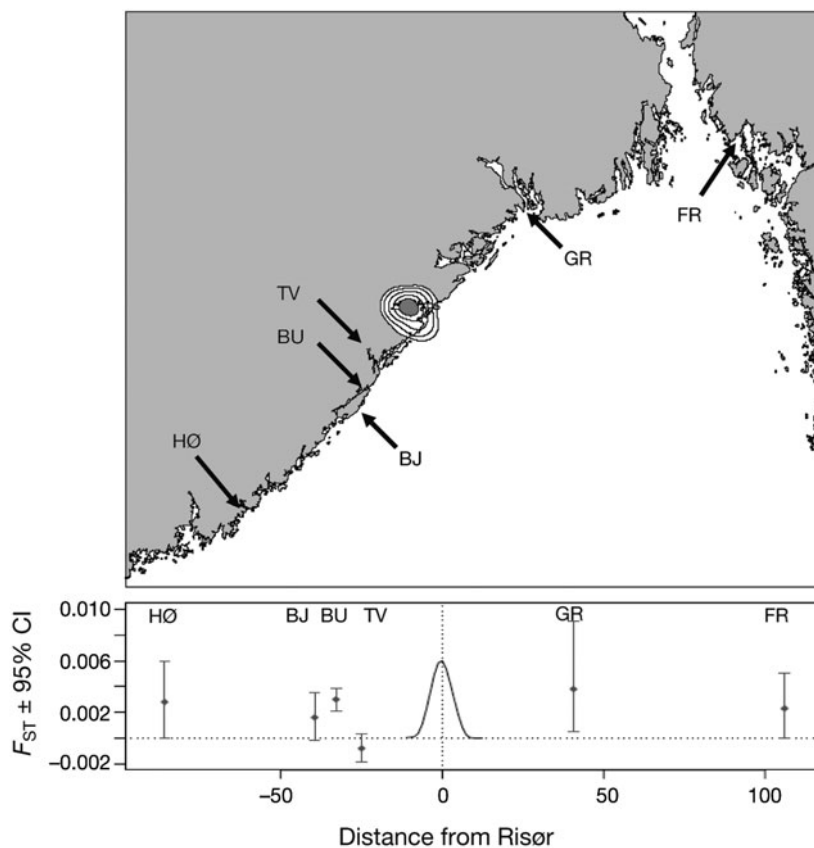


Fig. 5. *Gadus morhua*. Tag recoveries of coastal Atlantic cod compared with population genetic structure. Upper panel gives the untruncated density distribution estimated from the recoveries from 2005. The solid inner polygon represents the  $\alpha = 0.25$  area, while the lines correspond to the 0.5, 0.75, 0.9 and  $\alpha = 0.95$  iso-lines. The arrows indicate locations of genetic sampling taken from Jorde et al. (2007). Going from south to north the locations are: HØ = Høvåg, BJ = Bjelland, BU = Buøya, TV = Tvedestrand, GR = Grenland, FR = Fredrikstad. Lower panel describes the pair-wise genetic differences ( $F_{ST} \pm 95\%$  confidence interval [CI]) between the Risør sample and 6 other locations. The distance between the Risør sample and other locations is given on the x-axis (from left: Høvåg, Bjelland, Buøya, Tvedestrand, Grenland and Fredrikstad). The solid curve is a cross section of the unconditioned recovery probability distribution



tions along the coast) genetic differences were observed; Buøya (BU) and Grenland (GR) cod samples both exhibited significant positive  $F_{ST}$  values indicating differences in genetic composition (cf. Fig. 4 and Knutsen et al. 2003). The Bjelland (BJ) sample has a non-significant but positive  $F_{ST}$ ; however, individuals were not caught at a spawning location (cf. Jorde et al. 2007). Jorde et al. (2007) reported distances between genetically different populations to be 30 km or less. From the tagging data, the cross-section, along the coast, of the  $\alpha = 0.95$  area of the density distribution was approximately 18 km.

## DISCUSSION

Identifying separate populations and estimating the geographic extent of these populations is an important issue for population ecology and fisheries management. Determining population boundaries will be increasingly important, for instance, when designing marine protected areas (MPAs) for threatened populations. Genetic analysis is a powerful tool for identifying discrete populations with little connectivity. However, the power of statistical analysis of genetic data may be complicated by the effect of a small number of widely dispersing individuals, quickly reducing the signal for genetic structure (e.g. Waples 1998). Thus, applying several methods for revealing patterns of subdivision in marine species may prove useful.

We have used kernel smoothers to estimate the home range of coastal Atlantic cod. Here, we emphasize that the smoothing factor could be overestimated in the kernel density distribution estimation due to the use of the smoother  $h_{ref}$  (Worton 1995), so the real distribution would be narrower around the peak. On the other hand, if there is a significant decrease in fishing intensity when moving away from the coast, the population home range boundary would be shifted outwards towards open sea. Heterogeneity in fishing intensity along the coast could move the boundaries in either direction, but probably not much due to the low magnitude differences.

We have used GAMs to estimate coastal Atlantic cod movement patterns. The large residuals in the GAMs indicate a potential for rapid relocation within the population home range, seen as marked deviations from the expected distance moved during the first period following release. The slow increase in average distance through time, on the other hand, indicates that the individual fish either move initially but then become sedentary, or move around inside the population home range. This suggests that the populations are fairly localized but with some movement (e.g. for feeding or spawning within the population home range).

We show that cod on the Norwegian Skagerrak coast generally use small areas, stretching a few tens of kilometers along the coast. This pattern has been fairly stable through the last century, although the border delimiting the area used by the population differs somewhat between studies. The movement within the studies also seems to be bound within the population home range.

When comparing the cross-section of the population density distribution from Risør in 2005 with estimates of genetic differences, and contrasting these findings for Risør with other coastal locations in both directions along the coast, we demonstrate that the movements of coastal Atlantic cod are bounded within the distance where genetic similarities appear (Fig. 5). The distance from Risør to the untruncated  $\alpha = 0.95$  isolines (Fig. 5) is nearly 10 km while genetic differences do not appear between populations less than 30 km apart (Jorde et al. 2007). In some years, particularly 1937, 1970 and 1971, and 2005 (Risør), the areas occupied by  $\alpha = 0.95$  of the population density distributions extended farther along the coast and partially overlap. This does not necessarily mean that there was an exchange of individuals between populations since the data represent the utilization distribution for the entire period the cod were available for recapture.

The genetic disparities recorded were found amongst Atlantic cod sampled during spawning; i.e. when cod appeared to aggregate inside different fjords (Jorde et al. 2007, Knutsen et al. 2007). The areas charted in Figs. 2 & 3 were composed both of the spawning areas and area usage outside the spawning season for foraging. This indicates that the cod used the periods between spawning to migrate intermediate distances for the purposes of foraging, thus creating contiguous population areas. The areas depicted in Figs. 2 & 3 represent the regions where there is 95% probability of recovering a tagged fish. Some of the fish tagged in these studies were recovered far outside these boundaries, like the North Sea and Danish side of the Skagerrak (Fig. 1). These fish may represent North Sea cod returning after previously being transported to Skagerrak as eggs or larvae (cf. Knutsen et al. 2004). In 4 of the studies, the effect of degree of shelter in the GAM models is significant, but in opposite directions. This could be an effect of variable influx of eggs and larval recruitment from the North Sea (Knutsen et al. 2004, Stenseth et al. 2006) settling in the outermost skerries.

If large seasonal migrations were present, these could be expected to be correlated with 'day in year' scales rather than the 'days at liberty' scale. However, in most of these studies the mass of releases are concentrated in time so these 2 scales would be nearly identical. With the exception of the 1905 and 2005

datasets, the GAMs do not show any signs of cyclical movement throughout the year as would be expected for migrating populations. This was the case even when day in year was used in the models as a covariate. Since tagging and recovery data consist of only 2 positions, release and recovery, this regression technique may have low power when describing the temporal cyclic pattern of migration in only some fish. Cod would have to move in an ongoing, uniform fashion for the cyclic behavior to be detected. Fish alternating between spawning and barren seasons, as well as a few sedentary fish, would obscure the pattern. Electronic tagging, as data storage tags and acoustic tracking, may provide more detailed information on some individuals and may be combined with traditional tagging as a valuable supplement.

Our findings, using novel techniques to re-analyse past data, are important as they highlight the relationship between individual movements, populations' utilization areas and genetic spatial structure. Through this combined approach we have been able to describe how genetically distinct populations distribute themselves in space with regard to neighbouring populations. It is apparent that a stock may be structured on a smaller scale than the scale of appearance of genetic differences (Fig. 5). This will have consequences for the management of coastal Atlantic cod, and the same principles may be applied to other marine populations.

*Acknowledgements.* We acknowledge the invaluable support we received from the Institute of Marine Research. Special assistance was received from S. E. Enersen in extracting the data from previous databases. H. Gundersen and J. O. Vik have both provided valuable input regarding the analyses. K. Nash and 3 anonymous referees helped us improve our manuscript. Financial aid was received from the Institute of Marine Research and the University of Oslo (through the Centre for Ecological and Evolutionary Synthesis. The paper was finalised as part of the project 'Dynamics and genetics of oceanic — coastal cod population complexes' funded through the Norwegian Research Council.

#### LITERATURE CITED

- Arlinghaus R (2006) On the apparently striking disconnect between motivation and satisfaction in recreational fishing: the case of catch orientation of German anglers. *N Am J Fish Manag* 26:592–605
- Blundell GM, Maier JAK, Debevec EM (2001) Linear home ranges: effects of smoothing, sample size, and autocorrelation on kernel estimates. *Ecol Monogr* 71:469–489
- Bolle LJ, Hunter E, Rijnsdorp AD, Pastoors MA, Metcalfe JD, Reynolds JD (2005) Do tagging experiments tell the truth? Using electronic tags to evaluate conventional tagging data. *ICES J Mar Sci* 62:236–246
- Burt WH (1943) Territoriality and home range concepts as applied to mammals. *J Mammal* 24:346–352
- Dahl K, Dannevig GM (1906) Undersøgelse over nytten af torskeudklækning i østlandske fjorde. Betenkning. *Årsberetninger Norges Fiskerier* 1:1–121
- Danielssen DS, Gjosaeter J (1994) Release of 0-group cod, *Gadus morhua* L., on the southern coast of Norway in the years 1986–1989. *Aquac Fish Manag* 25:129–142
- Ellis N, Wang YG (2007) Effects of fish density distribution and effort distribution on catchability. *ICES J Mar Sci* 64:178–191
- Ford RG, Krumme DW (1979) Analysis of space use patterns. *J Theor Biol* 76:125–155
- Gitzen RA, Millsaugh JJ (2003) Comparison of least-squares cross-validation bandwidth options for kernel home-range estimation. *Wildl Soc Bull* 31:823–831
- Gitzen RA, Millsaugh JJ, Kernohan BJ (2006) Bandwidth selection for fixed-kernel analysis of animal utilization distributions. *J Wildl Manag* 70:1334–1344
- Hallenstvedt A, Wulff I (2004) Fritidsfiske i sjøen 2003. Norges Fiskerihøgskole/ Universitetet i Tromsø
- Hastie TJ, Tibshirani RJ (1990) Generalized additive models. Chapman & Hall, London
- Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning; data mining, inference and prediction. Springer Series in Statistics, New York
- Horne JS, Garton EO (2006) Likelihood cross-validation versus least squares cross-validation for choosing the smoothing parameter in kernel home-range analysis. *J Wildl Manag* 70:641–648
- Hunt JJ, Stobo WT, Almeida F (1999) Movement of Atlantic cod, *Gadus morhua*, tagged in the Gulf of Maine area. *Fish Bull* 97:842–860
- Jorde PE, Knutsen H, Espeland SH, Stenseth NC (2007) Spatial scale of genetic substructuring in coastal cod *Gadus morhua* and the geographic extent of local populations. *Mar Ecol Prog Ser* 343:229–237
- Knutsen H, Jorde PE, André C, Stenseth NC (2003) Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Mol Ecol* 12:385–394
- Knutsen H, Andre C, Jorde PE, Skogen MD, Thuróczy E, Stenseth NC (2004) Transport of North Sea cod larvae into the Skagerrak coastal populations. *Proc R Soc Lond B Biol Sci* 271:1337–1344
- Knutsen H, Olsen E, Cianelli L, Espeland SH and others (2007) Egg distribution, bottom topography and small-scale cod population structure in a coastal marine system. *Mar Ecol Prog Ser* 333:249–255
- Lawson GL, Rose GA (2000) Seasonal distribution and movements of coastal cod (*Gadus morhua* L.) in Placentia Bay, Newfoundland. *Fish Res* 49:61–75
- Løversen R (1946) Torskens vekst og vandring på Sørlandet. Fiskeridirektoratets skrifter, serie havundersøkelser 17: 449–471
- Metcalfe JD (2006) Fish population structuring in the North Sea: understanding processes and mechanisms from studies of the movements of adults. *J Fish Biol* 69:48–65
- Millsaugh JJ, Marzluff JM (2001) Radio tracking and animal populations. Academic Press, London
- Mohr CO (1947) Table of equivalent populations of North American small mammals. *Am Midl Nat* 37:223–249
- Moncada F, Abreu-Grobois FA, Muhlia-Melo A, Bell C and others (2006) Movement patterns of green turtles (*Chelonia mydas*) in Cuba and adjacent Caribbean waters inferred from flipper tag recaptures. *J Herpetol* 40:22–34
- Neilson JD, Stobo WT, Perley P (2006) Pollock (*Pollachius virens*) stock structure in the Canadian Maritimes inferred from mark-recapture studies. *ICES J Mar Sci* 63:749–765
- Parra GJ (2006) Resource partitioning in sympatric delphinids: space use and habitat preferences of Australian snubfin and Indo-Pacific humpback dolphins. *J Anim Ecol* 75:862–874

- Pawson MG, Jennings S (1996) A critique of methods for stock identification in marine capture fisheries. *Fish Res* 25: 203–217
- Pollock KH, Hoenig JM, Hearn WS, Calingaert B (2001) Tag reporting rate estimation: 1. An evaluation of the high-reward tagging method. *N Am J Fish Manag* 21:521–532
- Porter JH, Dooley JL (1993) Animal dispersal patterns—a reassessment of simple mathematical models. *Ecology* 74: 2436–2443
- R Development Core Team (2006) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Righton D, Quayle VA, Hetherington S, Burt G (2007) Movements and distribution of cod (*Gadus morhua*) in the southern North Sea and English Channel: results from conventional and electronic tagging experiments. *J Mar Biol Assoc UK* 87:599–613
- Robichaud D, Rose GA (2004) Migratory behaviour and range in Atlantic cod: inference from a century of tagging. *Fish Fish* 5:185–214
- Rowlingson B, Diggle P (1993) Splanx: spatial point pattern analysis code in S-Plus. *Comput Geosci* 19:627–655
- Seaman DE, Powell RA (1996) An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology* 77:2075–2085
- Shepherd GR, Moser J, Deuel D, Carlsen P (2006) The migration patterns of bluefish (*Pomatomus saltatrix*) along the Atlantic coast determined from tag recoveries. *Fish Bull* 104:559–570
- Shimada AM, Kimura DK (1994) Seasonal movements of Pacific cod, *Gadus macrocephalus*, in the eastern Bering Sea and adjacent waters based on tag-recapture data. *Fish Bull* 92:800–816
- Smith TD, Gjøvsæter J, Stenseth NC, Kittilsen MO, Danielssen DS, Tveite S (2002) A century of manipulating recruitment in coastal cod populations: the Flødevigen experience. *ICES Mar Sci Symp* 215:402–415
- Stenseth NC, Jorde PE, Chan KS, Hansen E and others (2006) Ecological and genetic impact of Atlantic cod larval drift in the Skagerrak. *Proc R Soc Lond B Biol Sci* 273: 1085–1092
- Stensland E, Carlen I, Sarnblad A, Bignert A, Berggren P (2006) Population size, distribution, and behavior of indo-pacific bottlenose (*Tursiops aduncus*) and humpback (*Sousa chinensis*) dolphins off the south coast of Zanzibar. *Mar Mamm Sci* 22:667–682
- Svedäng H, Righton D, Jonsson P (2007) Migratory behaviour of Atlantic cod *Gadus morhua*: natal homing is the prime stock-separating mechanism. *Mar Ecol Prog Ser* 345:1–12
- Van Winkle W (1975) Comparison of several probabilistic home-range models. *J Wildl Manag* 39:118–123
- Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J Hered* 89:438–450
- Wood S (2006) Generalized Additive Models: an introduction with R. Chapman & Hall/CRC, Boca Raton, FL
- Worton BJ (1987) A review of models of home range for animal movement. *Ecol Model* 38:277–298
- Worton BJ (1989) Kernel methods for estimating the utilization distribution in home range studies. *Ecology* 70: 164–168
- Worton BJ (1995) Using Monte-Carlo simulation to evaluate kernel-based home-range estimators. *J Wildl Manag* 59: 794–800

Editorial responsibility: Jon Hare,  
Narragansett, Rhode Island, USA

Submitted: October 8, 2007; Accepted: August 29, 2008  
Proofs received from author(s): December 4, 2008