Potential contact statistics for measuring scale-dependent spatial pattern and association: an example of northern cod (*Gadus morhua*) and capelin (*Mallotus villosus*)

Richard L. O’Driscoll, David C. Schneider, George A. Rose, and George R. Lilly

Abstract: Analysis of simulated data showed that potential contact statistics could be used to describe spatial pattern in sample density data. Potential contact is a new method, analogous to Ripley’s $K$ function for mapped point pattern analysis. Potential contact can be used to describe spatial pattern and association over a range of scales without grouping data and is robust against the presence of zeros. The statistical output is ecologically interpretable, as a measure of the degree of contact between individuals. This new technique was applied to examine changes in the spatial distribution of Atlantic cod (*Gadus morhua*) off Newfoundland, Canada, from 1985 to 1994, a period that encompassed a collapse of the cod stock. Sample data from bottom-trawl surveys indicated that cod were aggregated in patches with dimensions of 100–250 km. During the period of cod decline in the 1990s, spatial structure changed in three ways: the number of patches decreased, patch size shrank, and contact with conspecifics at small (10–20 km) scales fell. Cod were broadly associated with capelin (*Mallotus villosus*), a major prey species. Spatial distribution of capelin changed over the same time period as changes in cod distribution, and there was no evidence that contact between cod and capelin decreased.

Résumé : L’analyse de données simulées a montré que le calcul du contact potentiel pouvait servir à décrire le régime spatial dans des données sur la densité dans des échantillons. Le contact potentiel est une nouvelle méthode, analogue à la fonction $K$ de Ripley utilisée pour l’analyse de voisinage. Le contact potentiel peut servir à décrire le régime spatial et l’association sur une vaste gamme d’échelles sans grouper les données, et cette méthode est robuste en présence de zéros. Le résultat statistique est interprétable sur le plan écologique, comme mesure du degré de contact entre individus. Nous avons appliqué cette nouvelle technique à l’examen des changements dans la distribution spatiale de la morue (*Gadus morhua*) au large de Terre-Neuve (Canada) de 1985 à 1994, période qui a vu un effondrement du stock de morue. Les données d’échantillonnage fournies par des relevés au chalut de fond ont indiqué que les morues se rassemblaient en bancs dont les dimensions étaient de l’ordre de 100–250 km. Pendant la période de déclin de la morue des années 1990, la structure spatiale a changé de trois façons : le nombre des bancs a baissé; la taille des bancs a rétrécie; enfin, le contact avec les congénères à de petites échelles (10–20 km) a chuté. La morue est fortement associée au capelan (*Mallotus villosus*), qui est sa principale proie. La distribution spatiale du capelan a changé pendant la période où la distribution de la morue s’est modifiée, et rien n’indique que le contact entre la morue et le capelan ait baissé.

[Traduit par la Rédaction]

Introduction

The concept of scale-dependent spatial pattern is central to many aspects of fisheries ecology. The distributions of aquatic organisms are almost always patchy, and the degree of spatial heterogeneity usually varies with scale (Schneider 1994). Consequently, most fisheries ecologists are faced at some time with the problem of quantifying scale-dependent patterns. Description of observed spatial pattern over a range of scales is the important first step towards determining, understanding, and modelling the underlying pattern-generating processes (Levin 1992) and also in defining appropriate sampling strategy for hypothesis testing (Andrew and Mapstone 1987).

Methods for describing scale-dependent spatial pattern range from simple graphical representation to techniques involving extensive mathematical computation. Legendre and
Fortin (1989) distinguished between point pattern analysis, where the location of all individuals is known, and surface pattern analysis, where the distribution is spatially continuous and observations are only available at discrete sample points. In general, techniques for point pattern analysis are based on the distance separation between individuals (e.g., nearest neighbour, Ripley’s K; Ripley 1981; Cressie 1991), while techniques for surface pattern analysis are based on estimation of the lagged covariance function (e.g., spectral analysis; Plait and Denman 1975; autocorrelation: Legendre and Fortin 1989; geostatistics: Rossi et al. 1992) or grouped covariance structure (e.g., variance to mean ratio: Schneider and Plait 1986).

Existing point pattern analysis and surface pattern analysis methods have disadvantages when applied to ecological data. It is only feasible to map the positions of all individuals in an area at small spatial scales, especially when these individuals move. As a result, most ecological applications of point pattern analysis techniques have been in studies of vegetation (e.g., Andersen 1992) or of relatively sedentary invertebrates (e.g., Muotka and Penttinen 1994) at small scales. Surface pattern analysis methods require a data set where a large number (>30) of sampling points are spaced over the entire area, preferably in a regular grid. Because many surface pattern analysis techniques are based on a measure of lagged covariance, they are sensitive to the presence of zeros and outlying values in the data set. Points with zero values will autocorrelate, overstating covariance (Legendre and Fortin 1989), and outliers will also bias the results. Individual methods have further limitations. For example, spectral analysis may not detect nonperiodic patterns (Logerwell et al. 1998). Other surface pattern analysis methods require the data to be grouped to describe covariance at larger spatial scales. This introduces problems with defining group boundaries and reduced sample size at large scales (Malatesa et al. 1992).

In this paper, we present a new method to describe spatial structure in sampled data from continuous two-dimensional distribution surface patterns. Potential contact is a distance-based method, analogous to Ripley’s K function for point pattern analysis. This new technique allows description of spatial pattern across a wide range of scales, is robust to the presence of zeros, and can be used in an area with irregular boundaries and sampling. An added advantage of this method is that the output is easily interpreted ecologically. Our test statistic is a direct measure of the contact of an individual with neighbouring individuals and is defined as the average number of neighbouring individuals within a distance, of any individual in the distribution. Contact is of interest in many ecological processes, including social interaction, resource utilisation, and genetic recombination. Potential contact is a generalisation of mean crowding (Lloyd 1967). However, unlike Lloyd’s (1967) index of mean crowding, potential contact may be calculated at increasing spatial scales (increasing ) without grouping data. The concept of potential contact was introduced by Schneider et al. (1987) and described by Schneider (1994), but we are aware of only one other ecological study in which contact has been used directly as a measure of spatial structure (Bult et al. 1998). Our study differs from that of Bult et al. (1998) because we apply the method to sampled surface data rather than to a mapped point process.

We illustrate our statistical methodology using simulated data and then apply these new techniques to describe large-scale (tens to hundreds of kilometres) spatial pattern in the distribution of Atlantic cod (Gadus morhua) using data from bottom-trawl research surveys in Newfoundland, Canada, in 1985–1994. This data set provides an example of an ecological application of our techniques. Previous analysis of these and other data from this region indicates that the spatial distribution of cod changed during the 1990s (review in Rose et al. 2000). These changes, notably a southward contraction in the area occupied by the cod stock, may have made the fish more vulnerable to overfishing (Rose et al. 2000), leading to the collapse and closure of the northern cod fishery off Newfoundland in 1992. We were interested in how the spatial pattern of the cod, and specifically their potential contact with other cod, changed during this period of stock collapse. Maintenance of potential contact as cod numbers declined could account for the failure of catch-per-unit-effort indices to track stock size (Rose and Kulka 1999).

In addition to describing spatial pattern and aggregation of individuals, Ripley’s K may be used to analyse bivariate patterns and association in point patterns (Lotwick and Silverman 1982). Similarly, our related measure of potential contact may be used to analyse interspecific associations in sampled data from continuous surface patterns. We present examples based on simulated data and then use the bottom-trawl data set to examine scale-dependent spatial association between cod and capelin (Malatous villosus) off Newfoundland in 1985–1994. Capelin are the major prey of cod in this region (review in Lilly 1987).

Methods

Statistical methods

We define potential contact as the number of neighbours potentially encountered by an individual within an ambit of given size. For sample density data from a continuous surface pattern, potential contact may be expressed as

\[ PC(t) = \pi t^2 E|d(t)|. \]

The \( E|d(t)| \) is the expected density of neighbours within a circular ambit of radius \( t \) and \( \pi t^2 \) is the area of ambit radius \( t \). The \( E|d(t)| \) is calculated empirically as a weighted average. For each sample point \( i = 1, 2, ... , n, \) density estimates \( d_j \) from all other sample points within distance \( t \) (indicator function \( I_j = 1 \) where distance \( t < t \) and 0 otherwise) are averaged and scaled by the density at point \( i \) \( d(i) \). The result is then summed and divided by the sum of the point density estimates to estimate the expected density of neighbours within distance \( t \):

\[ E|d(t)| = \frac{\sum_{i=1}^{n} d_i |d(i)|}{\sum_{i=1}^{n} d_i} \]

where

\[ |d(i)| \]

\[ (d(i)| \]
O’Driscoll et al.

\[
(3) \quad \bar{d}_t = \frac{1}{n} \sum_{j=1}^{n} \frac{(d_{ij})_d}{I_{ij}}.
\]

Ambit size (and spatial scale) is increased by incrementally increasing ambit radius \( t \).

The PC(\( t \)) is analogous to \( E(N(0)) \), the expectation term of Ripley’s \( K \) function for a spatial point pattern (Ripley 1981). The \( E(N(0)) \) is the Lagrangian equivalent of potential contact, where cooccurrence is measured as a function of distance from an individual and is not constrained by sampling points.

There is a sampling uncertainty associated with PC(\( t \)) because \( E(d(0)) \) is calculated from sample estimates of density. We estimated sampling uncertainty using a jackknife approach (Manly 1997). Individual sample points (\( i = 1, 2, \ldots, n \)) were removed sequentially and PC(\( t \)) was calculated for each of the reduced data sets. The variability among the \( n \) jackknifed values of PC(\( t \)) was used to estimate the standard error (Manly 1997).

To test whether an observed spatial distribution was significantly different from a random arrangement, we randomised the density pattern by randomly reallocating increments of density to the sample points. The increment of density chosen was the minimum resolution of the data. For example, if density was measured to two decimal places, we randomly reallocoted density values of 0.01 until the sum of the densities across all sample points was equal to the sum of the densities in the original pattern. This randomisation method is equivalent to randomly reallocating individuals in a spatial point pattern. The randomisation process was repeated 100 times and PC(\( t \)) was calculated for each realisation. From these realisations, we calculated the average potential contact for a random arrangement, PC(\( t \)ran), and the values of potential contact equivalent to the 5th percentile, PC(\( t \)95), and the 95th percentile, PC(\( t \)5). We express our results as the “extra contact” at distance \( t \), XC(\( t \)): (4) \[ XC(t) = PC(t) - PC(t)_{\text{ran}}. \]

The XC(\( t \)) is the average number of “extra” neighbours within distance \( t \). The extra neighbours are those that are not expected if the individuals were distributed randomly throughout the study area. The XC(\( t \)) is analogous to \( L(t) = (K(t)/\pi) - t \), commonly used as a test statistic in applications of Ripley’s \( K \) function to spatial point data (Cressie 1991). The XC(\( t \)) has an expected value of zero under a null hypothesis of complete spatial randomness. When individuals are aggregated, XC(\( t \)) > 0. The upper and lower bounds for a random spatial distribution are given by (5) \[ XC(t)_{95} = PC(t)_{95} - PC(t)_{\text{ran}}. \]

(6) \[ XC(t)_{5} = PC(t)_{5} - PC(t)_{\text{ran}}. \]

The XC(\( t \)) represents the significance level of \( p = 0.05 \) for a one-sided test or \( p = 0.10 \) for a two-sided test. Because we were not considering the alternative hypothesis of uniformity, when XC(\( t \)) > XC(\( t \))95, the aggregation was considered statistically significant.

No correction for edge bias (Haase 1995) was necessary in our analysis because the randomisation scheme used to generate the null model and calculate XC(\( t \)) used the same sampling points and makes the same assumptions about the distribution of organisms outside the study area as the data.

We used a Matlab® routine to calculate XC(\( t \)), XC(\( t \))3, and XC(\( t \))95 at distances \( t = s, 2s, 3s, T \), where \( s \) is the sampling scale, related to the separation between adjacent sampling points, and \( T \) is the maximum separation between sampling points. The Matlab® code is available on request from the first author. Results are presented as plots of XC(\( t \)) as a function of \( t \). Such plots allow comparison of association over a range of scales (Haase 1995). In general, interpretation should be limited to values of \( t < T/2 \) (Cressie 1991). The XC(\( t \)) drops back to zero as \( t \) approaches \( T \) because all data become encompassed by the calculation. We do not include estimates of sampling uncertainty for XC(\( t \)) values themselves. Description of spatial pattern was based on relative levels of contact from observed and randomised arrangements of fish which were not affected by sampling uncertainty.

**Bivariate analysis**

Our measure of contact may also be used to examine associations between different species. By analogy with eqs. 2 and 3, the expected density of type B individuals within distance \( t \) of a type A individual, \( E[b(t)] \), is

\[
\sum_{i=1}^{n} a_i (\bar{d}_i) - \sum_{i=1}^{n} d_i.
\]

The \( a_i \) and \( b_j \) are the measured densities of type A and type B individuals at sample point \( i \) and \( j \), respectively. The PC(\( t \)) and XC(\( t \)) are calculated from eqs. 1 and 4 as above. Bivariate XC(\( t \)) is a measure of the extra contact with type B individuals within distance \( t \) of a type A individual. The measure is easily interpreted. A value of zero at distance \( t \) indicates no association at that scale, values greater than zero indicate positive association or correlation, and values less than zero indicate negative association or segregation. The results are displayed as plots of XC(\( t \)) as a function of \( t \).

A randomisation procedure was again used to assess whether association was statistically significant, but it was more difficult to choose a null model. To be a valid test of association, the null (random) patterns of both type A and type B individuals should retain some of the spatial properties of the observed (data) patterns while randomising their relative positions (Roxburgh and Chesson 1998). This is achieved in part by only randomising the distribution of one group. In this paper, we compared four alternative null models based on different methods of randomisation: (i) random reallocation of density increments (individuals) as used in univariate analysis, (ii) random rearrangement of density values between sampling points, (iii) random toroidal shift of density values, and (iv) random repositioning of clump centres. The first two models change the second-order properties of the spatial distribution. However, both these methods have been used in the literature (Bult et al. 1998; O’Driscoll 1998). The toroidal shift method suggested by Lotwick and Silverman (1982) has been the most commonly used null model in analysis of bivariate point processes (e.g., Andersen 1992; Muotka and Penttinen 1994). Random repositioning of clump centres is a new method suggested by Coomes et al. (1999).

**Simulated data**

Continuous surface spatial patterns were generated in Matlab®. Univariate patterns were based on a random arrangement of groups of individuals (patches) within a \( 100 \times 100 \) km area. The size and number of groups were set and the routine positioned the patches at random within the area. Individual patches were usually circular (Fig. 1). In most cases, density varied across the patch in a sinusoidal fashion, with a peak density of 1 individual/km² at the centre (Fig. 1). Densities outside the patches were zero. All patterns were sampled using a regular \( 20 \times 20 \) grid of stations 5 km apart to obtain point estimates of density. We then investigated the influence of patch radius, number of patches, and density variation within a patch on contact XC(\( t \)) across a range of spatial scales, \( t = 5, 10, 15, \ldots, 140 \) km.

© 2000 NRC Canada
Fig. 1. Description of scale-dependent spatial pattern in simulated data. Plots on the left show a 20 × 20 grid of discrete samples from the simulated distribution. Plots on the right are of extra contact with neighbours as a function of scale (see text for details). Dotted lines are the 5 and 95% bounds for a random pattern generated from 100 randomisations of density increments. Values of $XC(t)$ above the upper confidence interval were considered statistically significant (nonrandom aggregation). Simulated patterns: (a) single circular patch of radius 30 km with sinusoidal density variation within the patch; (b) single circular patch of radius 30 km with uniform density within the patch; (c) two circular patches of radius 20 km; (d) five circular patches of radius 10 km; (e) 10 circular patches of radius 5 km; (f) 20 circular patches of radius 2 km; (g) two circular patches of radius 10 km and a single circular patch of radius 20 km; (h) anisotropic patch with dimensions 60 × 20 km.
Bivariate patterns were simulated by generating a second spatial pattern, where the centres of patches of type B individuals were separated by a known distance from the centres of patches of type A individuals. We used bivariate simulations to illustrate the effect of varying separation on the observed scale of association and also to investigate the choice of null model on the significance test.

**Bottom-trawl survey data**

Atlantic cod and capelin were caught during annual fall (October–November) bottom-trawl surveys off Newfoundland from 1985 to 1994. Trawl surveys were based on a stratified random design giving a total of between 357 and 543 trawl stations each year (Table 1). The survey design ensured relatively uniform spatial coverage with separation between adjacent stations of 1–60 km (mean separation 14.6 km, Table 1). Tows were made using an Engel-145 high-rise otter trawl with a mesh liner of 29 mm towed for 30 min at 3.5 km from either the 74-m *Gadus atlanticus* or the 50-m *Wilfred Templeman*. The constant wing spread value used for this trawl was 45 ft (13.5 m), meaning that the total bottom area covered in a 30-min tow was 0.044 km². Fishing sets covered the full 24-h period. Numbers of cod and mass (kilograms) of capelin in each catch were scaled by tow area and expressed per square kilometre with no correction for catchability. This gave point estimates of cod and capelin density at the position of the tow. We used these data to calculate the potential contact of cod with other cod and of cod with capelin at scales of t = 10, 20, 30, ..., 1260 km.

**Results**

**Simulation results: description of spatial pattern**

Plots of $XC(t)$ as a function of $t$ can be used to describe characteristics of a spatial pattern (Fig. 1). A single patch gives a single peak in $XC(t)$ (Fig. 1a). The value of $t$ at which this peak occurs is a measure of the patch extent. The patch extent is a function of both the patch dimensions and the arrangement of densities within the patch. Patterns in Figs. 1a and 1b have the same patch radius (30 km) but different surfaces. In Fig. 1a, peak $XC(t)$ occurs at $t = 35$ km, while in Fig. 1b, peak $XC(t)$ is at $t = 45$ km. The value of $XC(t)$ is a measure of the intensity of crowding within a patch. For example, the peak $XC(t = 35)$ in Fig. 1a indicates that each individual has 737 more neighbouring individuals within 35 km than would be expected in a random arrangement. The equivalent $XC(t = 35)$ in Fig. 1b is higher (1076) because density does not decline away from the centre of the patch.

When there are two or more patches in the pattern, the shape of the plot of $XC(t)$ as a function of $t$ reflects both the individual patch characteristics and the separation between patches (Figs. 1c–1g). The first peak in $XC(t)$ describes the patch characteristics. The first peak was defined as the first value of $t$ at which $XC(t)$ was higher than the values at the three succeeding values of $t$ (i.e., $XC(t) > XC(t + s)$, $XC(t + 2s)$, $XC(t + 3s)$, where $s$ is the increment between $t$ values). This criterion was used to avoid labelling small jumps in contact as peaks. In Figs. 1c–1f, the value of $t$ corresponding to the first peak in $XC(t)$ (patch extent) occurs at 25, 15, 5, and 5 km, respectively. $XC(t)$, corresponding to simulated patch radii of 20, 10, 5, and 2 km, respectively. Intensity of crowding (peak height) also decreases in Figs. 1c–1f because there are fewer individuals in the smaller patches. Second and higher peaks in $XC(t)$ correspond to the separation between patches. For example, in Fig. 1c, the second peak at $XC(t = 55)$ km corresponds to a distance separation between the centres of the two patches of 45 km. As more patches are added, interpretation becomes more difficult (Figs. 1d–f) due to constructive and destructive interference.

When patches of different sizes are present, patch extent provides a measure of the weighted average of patch size. In Fig 1g, two patches of radius 10 km and one patch of 20 km were sampled. Peak density in the larger patch was reduced, so total number of individuals obtained by integrating density over the patch surface was equivalent to the number of individuals in the two smaller patches. Patch extent, as determined by the first peak in $XC(t)$, was at $t = 20$ km (Fig. 1g), which is intermediate between patch extent determined for patches with radii of 10 km (extent = 15 km, Fig. 1d) and 20 km (extent = 25 km, Fig. 1c).

Our method assumes that patches are isotropic. When an anisotropic patch (e.g., Fig. 1h) was sampled, patch extent was a measure of the size of an equivalent isotropic patch. For example, the patch in Fig. 1h had overall dimensions of $60 \times 20$ km. The measured patch extent was 25 km, equivalent to the patch extent obtained from a circular patch with similar mean dimensions (radius 20 km, Fig. 1c).

In all these patterns, $XC(t)$ was significantly greater than zero for most values of $t$. The simulated patchy arrangement of individuals meant that contact with neighbours was significantly higher than for a random arrangement, over a wide range of scales.

**Description of spatial pattern in Atlantic cod distribution off Newfoundland**

Spatial distributions of cod from bottom-trawl catches were explored initially using expanding symbol plots (Fig. 2). Plots of $XC(t)$ as a function of $t$ (Fig. 2) provided a statistical description of the spatial pattern in cod data. We set the sampling scale at 10 km, which was slightly less than average separation between adjacent stations (Table 1). In eight of 10 years between 1985 and 1994, the measured patch extent of cod was between 100 and 250 km (Fig. 3a). This result suggests that cod were characteristically aggregated in patches at spatial scales of hundreds of kilometres at the time of the survey. Intensity of crowding (Fig. 3b) was related to patch extent (Fig. 3a), indicating that there were more extra neighbours in larger patches.

### Table 1. Arrangement of bottom-trawl stations during fall surveys, 1985–1994.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of stations</th>
<th>Station separation (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
</tr>
<tr>
<td>1985</td>
<td>543</td>
<td>2.74</td>
</tr>
<tr>
<td>1986</td>
<td>357</td>
<td>4.48</td>
</tr>
<tr>
<td>1987</td>
<td>453</td>
<td>3.34</td>
</tr>
<tr>
<td>1988</td>
<td>428</td>
<td>2.80</td>
</tr>
<tr>
<td>1989</td>
<td>450</td>
<td>2.18</td>
</tr>
<tr>
<td>1990</td>
<td>404</td>
<td>1.01</td>
</tr>
<tr>
<td>1991</td>
<td>532</td>
<td>1.61</td>
</tr>
<tr>
<td>1992</td>
<td>534</td>
<td>0.57</td>
</tr>
<tr>
<td>1993</td>
<td>416</td>
<td>1.90</td>
</tr>
<tr>
<td>1994</td>
<td>455</td>
<td>1.95</td>
</tr>
</tbody>
</table>
Fig. 2. Description of spatial pattern in Atlantic cod distributions from bottom-trawl data in (a) 1989, (b) 1992, and (c) 1994. Maps on left are expanding symbol plots of cod catch. Circle diameters are proportional to $\log_{10}$ number of cod per standard 30-min tow. A cross indicates that no cod were captured. Plots on the right are of extra contact as a function of scale. Dotted lines are the 5 and 95\% bounds for a random arrangement of cod generated from 100 randomisations of density increments. Horizontal bars indicate the first peaks that were used as a measure of patch extent.
As well as exploring patch characteristics using plots of \( XC(t) \) as a function of \( t \), it is useful ecologically to examine values of potential contact \( PC(t) \) directly. Because extra contact is the difference between observed and expected contact at scale \( t \) (eq. 4), \( XC(t) \) is a measure of patchiness of the distribution, independent of mean density. The decline in overall density of cod off Newfoundland in the 1990s has been well documented (e.g., Hutchings 1996) and is clearly seen in a plot of mean density of cod from bottom trawl catches from 1985 to 1994 (Fig. 4a). Potential contact \( PC(t) \) increases with increasing ambit \( t \) as a function of both patchiness and density. If we can estimate the ambit size of the organism being studied, we can test how changes in density (Fig. 4a) and in patchiness (Fig. 3) combine to influence contact with conspecifics. Little is known about the ambit of cod, but mean swimming speeds during the spring migration from 1990 to 1992 were between 6.6 and 24 km/day, with a peak of 59 km/day (Rose et al. 1995). Figure 4b compares potential contact between cod at ambits of \( t = 10, 20, \) and 60 km from 1985 to 1994. Absolute values of \( PC(t) \) need to be treated with caution because our calculation of contact (eq. 1) is based on the assumption that similar densities of fish were present outside the study area. This is clearly not the case because the survey area is bounded to the west by land and to the east by very deep water, which is not suitable habitat for cod (Fig. 2). As a result, \( PC(t) \) will tend to overestimate contact at larger values of \( t \). At small scales (Fig. 4b), almost all ambits are entirely within the surveyed area, so contact can be used as a measure of number of neighbours.

Changes in spatial distribution appeared to occur sequentially as number of cod declined. In the 1980s, there were several areas with consistently high densities of cod in fall, as previously described by Lilly (1994) and Hutchings (1996). These included patches in the northern (54–55°N, 54–56°W), central (51–52°N, 50–52°W), eastern (49–50°N, 49–51°W), and southern (46–47°N, 49–51°W) parts of the surveyed region (Fig. 2a). Patch extent was similar in all areas (hundreds of kilometres), and plots of \( XC(t) \) as a function of scale in the 1980s were complex, with multiple peaks at large scales reflecting separation between these areas of high density (Fig. 2a). Potential contact (Fig. 4b) and intensity of crowding (Fig. 3b) were highest in 1986, when catches of cod (Fig. 4a) were unusually large. The first noticeable change in spatial structure in the 1990s was a reduction in the number of areas with high densities of cod. In 1991, there was no patch in the north, and by 1992, only patches in the eastern and southern parts of the survey region remained (Fig. 2b). Plots of \( XC(t) \) became less complex, with fewer peaks and less structure at larger scales (Fig. 2b). Patch extent (Fig. 3) and small-scale potential contact (Fig. 4b) were similar because the remaining patches had spatial structure similar to that of patches in the 1980s. As numbers of cod continued to decline, patch size shrank, with a dramatic decrease in our measures of patch extent (Fig. 3a) and individuals within a patch (Fig. 3b) in 1993. Potential contact with neighbours at small scales also decreased in 1993 (Fig. 4b), but contact with neighbours at
ambits of 10 and 20 km was still similar to that observed in the 1980s (Fig. 4b). The final change in spatial structure was evident in 1994, when numbers of cod had declined to very low levels (Fig. 4a) and potential contact with neighbours fell by two orders of magnitude (Fig. 4b). There were no large patches of cod, and the plot of $XC(t)$ showed little major structure (Fig. 2c).

Simulation results: bivariate analysis

Contact statistics can be used to describe spatial association across a range of scales. For a simple example, consider a single patch of type A individuals separated by a known distance from a single patch of type B individuals (Fig. 5a). Association between the two groups is scale dependent. At scales less than the separation between groups, association is negative, indicating that there are fewer individuals of type B surrounding a type A individual than there would be if type B individuals were arranged randomly. At scales close to the separation between groups (30 km between patch centres in Fig. 5a), the association between groups is positive, indicating that there are more type B individuals than expected within distance $t$ of a type A individual. At larger scales, the association becomes negative again (Fig. 5a). The value of $t$ at which $XC(t)$ is maximal can be considered as the scale of maximum association. This is the scale at which each type A individual has the most extra type B individuals surrounding it. The value of $XC(t)$ is a measure of the intensity of the association. In Fig. 5a, the scale of maximum association is 40 km. Each type A individual had 62 more type B individuals within an ambit of 40 km than would be expected in a random arrangement. A more complex arrangement, where there are multiple patches of both type A and type B individuals, is illustrated in Fig. 5b. The resulting plot of $XC(t)$ again reveals scale-dependent association with maximum association at a scale of 40 km.

Effect of null model choice on significance of association

The statistical significance of the bivariate associations measured in Fig. 5 was highly dependent on the choice of null random model. Figure 6 shows plots of $XC(t)$ as a function of $t$ calculated from the distribution in Fig. 5b with four different null models. It is important to note that choice of null model does not influence the shape of the plot of $XC(t)$ or affect the scale of maximum association or its intensity (Fig. 6). However, maximum association ($t = 40$ km) was only considered significant using two of the four null models (random shift and random clump) that preserve the second-order spatial distribution of type B individuals indicated no significant association between type A and type B individuals at any scale (Figs. 6c and 6d).

Spatial association between cod and capelin prey off Newfoundland

Spatial association between cod and capelin off Newfoundland was scale dependent. Plots of extra contact between cod and capelin from bottom trawl data are shown in Fig. 7. The procedure to calculate PC($t_{\text{ran}}$) and confidence intervals was based on random rearrangement of cod density.

© 2000 NRC Canada
values (the second null model presented in the Methods section). Typically, there was no association between cod and capelin at scales less than 100 km, positive association at scales from 100 to 600 km, and negative association at larger scales (Fig. 7). From 1985 to 1994, scale of maximum association occurred at scales between 260 km (1990) and 920 km (1988). Cod and capelin were associated more closely from 1989 to 1994 than previously (Fig. 8a). Potential contact PC(t) with capelin for cod ambits of t = 10, 20, and 60 km was related to mean catch of capelin and showed no trend from 1985 to 1994 (Fig. 8b).

Discussion

Contact statistics provide a novel method of describing scale-dependent spatial pattern and assessing scale-dependent spatial pattern and assessing scale-dependent spatial pattern and assessing scale-dependent spatial pattern and assessing scale-dependent spatial pattern and assessing scale-dependent spatial pattern and assessing scale-dependent spatial pattern and assessing scale-dependent spatial pattern and assessing scale-dependent spatial pattern.
Fig. 7. Description of spatial association between Atlantic cod and capelin distributions from bottom-trawl data in (a) 1989, (b) 1992, and (c) 1994. Maps on left are expanding symbol plots of capelin catch. Circle diameters are proportional to log_{10} mass of capelin in kilograms per standard 30-min tow. A cross indicates that no capelin were captured. Cod distribution for the corresponding years is shown in Fig. 2. Plots on the right are of association as a function of scale. Dotted lines are the 5 and 95% bounds of spatial association for a random arrangement of cod generated from 100 randomisations of cod density values.
dependent spatial association. The results are easily interpreted ecologically because the output is a measure of the average number of type B individuals within a given ambit of a type A individual (Schneider 1994). Type B “individuals” could be conspecific neighbours, prey, competitors, or even abiotic features such as habitat characteristics (Bult et al. 1998). Contact may be used directly, as in this paper, to assess the potential frequency of social interactions, degree of competition, or encounters with prey. It could also be used as an input into dynamic models of social behaviour or foraging.

Our statistical methods share many similarities with second-order point pattern analysis techniques, such as Ripley’s $K$ function (Ripley 1981), but can be applied to sampled data from a continuous surface pattern as long as the sampling scale is smaller than the spatial scales of interest. Unlike alternative covariance-based measures such as geostatistics, contact statistics are robust against the presence of zero data values. Because contact at increasing spatial scale is calculated by incrementally increasing the ambit around each sampled point, problems with grouping data and decreased sample size at larger scales (e.g., Malatesta et al. 1992) are reduced. The major disadvantage of these methods is that analysis is computationally intensive, particularly when there are a large number of sample points. Analyses in this paper took from 4 to 24 h to run on a Pentium® II 450-MHz personal computer. Computing time may be reduced for exploratory analysis by decreasing the number of randomisations or by decreasing the number or range of spatial scales (values of $r$) over which contact is assessed.

Sampling design is an important consideration when calculating contact statistics, as it is for any statistical method that relies on sample data (e.g., Andrew and Mapstone 1987). The separation between adjacent sampling points determines the minimum spatial scale at which contact can be assessed and also the precision with which we can describe patch extent. Patch extents measured from Atlantic cod distribution data (20–340 km) were much greater than the distance between trawl stations and the number of samples was large, so the random sample design of bottom-trawl stations was sufficient to describe spatial pattern.

Our measurements of cod patch extent were similar to those determined by Warren (1997) based on variogram analysis of the trawl data from 1985 to 1992. Warren’s (1997) analysis divided the surveyed area into three subareas based on arbitrary Northwest Atlantic Fishery Organisation management divisions. To avoid problems with zeros, Warren (1997) first reduced trawl catches to an indicator variable (presence–absence) or to the nonzero component and then fitted variogram models. The range of the variogram may be considered as a measure of patch dimensions (Rossi et al. 1992). Observed ranges for variograms fitted to the presence–absence data in the three areas were between 30 and 330 km (Warren 1997), which were similar to the patch extents that we measured (20–340 km). However, there was no significant correlation at $p = 0.05$ between our measure of patch extent over the entire survey area and Warren’s (1997) estimates of variogram range in the northern ($r = 0.11, n = 8$), central ($r = 0.72, n = 6$), or southern ($r = -0.31, n = 8$) subareas between 1985 and 1992. Part of the reason for this was that cod patches tended to straddle the arbitrary boundaries between Northwest Atlantic Fishery Organisation divisions. We believe that our methods provide a better overall measure of spatial structure in the cod data than the analysis of Warren (1997) because we analysed the entire data set and did not need to reduce the data to avoid zero counts. Our results are also more easily interpreted from the standpoint of cod distribution ecology.

As expected, cod exhibited changes in patch structure and contact with neighbours during the period of changing spatial distribution and stock collapse in the 1990s. Our analysis indicates that changes in spatial structure occurred in three ways: (i) reduction in the number of patches from 1990, (ii) reduction in the size (spatial extent and number of individuals) of remaining patches from 1992, and (iii) reduction in local density and small-scale contact with neighbours from 1993. There is support for our interpretation from several sources. Lilly (1994) and Hutchings (1996) have both used density plots to show that patches of high cod density declined in the 1990s. The decline in the number of patches occurred from the north, resulting in a southward shift in the overall distribution of the cod, as indicated by density centroids (Rose et al. 1994) or ellipses (Atkinson et al. 1997). Warren (1997) showed that spatial structure measured from variograms had broken down (was not distinguishable from random) in the northern and central subareas by 1992 but...
It is widely acknowledged that overfishing was the major cause of the decline in the Newfoundland cod population, but we do not believe that the observed changes in cod spatial structure can be explained by overfishing alone, as suggested by Hutchings (1996). Under this model, as stock size declined, individual patches would be reduced at a rate proportional to the fishing effort. This does not fit with the observation that patches disappeared first in the north, where fishing pressure was lower (Rose et al. 2000), but persisted further south even up to the final collapse. Instead, we believe that other factors, perhaps indexed by anomalously low water temperatures (Rose et al. 1994) and changing distributions of important prey species (Carscadden and Nakashima 1997), affected cod habitat suitability in the early 1990s. Changing habitat suitability would alter the topography of the basin model (Fig. 9). In keeping with this interpretation, the loss of patches from the north was probably due to a decrease in habitat suitability related to adverse environmental conditions. Migrants from the north could then have joined and temporarily replenished patches to the south (Rose et al. 2000), which is equivalent to flow down the habitat preference gradient in the present model (Fig. 9b). This model accounts for the observed maintenance of patch structure and local density even after the population had declined to very low levels in 1992.

Observations of changes in spatial distribution of cod as the stock declined may be used as a guide to possible changes as the stock recovers. We speculate that cod spatial distribution during recovery will reverse the changes that occurred during the decline. Local density will increase first, followed by an increase in patch size and extent, with the final step being an increase in the number of patches. Whether the cod will return to their historical distribution pattern is dependent on whether hypothesised changes in habitat suitability are also reversible.

Contact statistics can be used to measure spatial association, but further work is required to determine an appropriate null model for assessing significance. The four alternative null models that we presented represent different null hypotheses about the process generating spatial structure. Random shift and random clump null models assume that spatial structure in species A is generated independently of species B. Power to detect significance using these models was low due to constraints imposed by the size of the sampled area. At large scales, some random arrangements would always align the patches of the two types of individuals more closely than the pattern observed in the data. Null models that randomised individual increments of density or density values were not constrained by maintaining spatial structure, and the possible set of randomisations was much greater. However, because these null models change the second-order properties of one of the distributions, they are not tests of independence (Roxburgh and Chesson 1998). As long as this violation is stated explicitly, the use of such random models may still provide useful ecological information. For example, the random values null model that we used to calculate confidence intervals for our trawl data was based on the hypothesis that sample estimates of cod density were independent and that second-order spatial properties were generated only by aggregation of cod in response to capelin. This hypothesis is valid if schools of cod are smaller than
the separation between trawl sample points and do not aggregate except in response to capelin. O’Driscoll (1998) randomised individual seabirds (random individuals null model) to test the hypothesis that more birds were associated with schools of fish than would be expected if the seabirds were distributed randomly. This hypothesis becomes biologically interpretable if seabirds are foraging individually and aggregating in response to fish schools.

Cod were only broadly associated with capelin prey in the bottom-trawl data from 1985 to 1994. This agrees with the interpretation of Lilly (1994) based on visual comparison of plots of cod and capelin distribution from 1985 to 1992. Scale of maximum association between cod and capelin decreased in 1989 and then remained relatively constant during the 1990s, reflecting common changes in the distribution patterns of the two species. Although distributions of both cod (Atkinson et al. 1997) and capelin (Carscadden and Nakashima 1997) contracted in the 1990s, the area occupied by the remaining fish was similar (Lilly 1994). Offshore concentrations of capelin shifted southward (Carscadden and Nakashima 1997) over the same time period that cod became concentrated in the southern part of their former range (Rose et al. 2000). The potential contact between cod and capelin within hypothetical cod daily home ranges of 10, 20, and 60 km did not decrease in the 1990s, suggesting that changes in the distribution of both cod and capelin did not negatively affect the likelihood of cod encountering capelin prey. This conclusion is supported by analysis of cod stomach contents. Lilly (1994) reported that feeding on capelin did not seem to decline in the early 1990s, except in the north of our survey area. The southward shift in cod distribution observed in the 1990s may have been driven in part by a need for cod to maintain broad-scale association with capelin.

There was no positive association between cod and capelin at small scales (10–100 km). There are several possible explanations for this lack of association. First, trawl surveys were conducted in the late fall, which is not regarded as the period of peak feeding on capelin by cod (Turuk 1968), although capelin were commonly found in stomachs of cod caught in trawls at this time (Lilly 1987, 1994). Second, the Engels bottom-trawl may not have adequately sampled capelin distribution. Capelin are a pelagic species and will only be caught in a bottom-trawl when concentrated near the bottom. There is also evidence that the Engels trawl is not an efficient gear for catching capelin. Estimates of trawlable biomass of capelin from fall surveys increased by an order of magnitude following a change of gear type from an Engels trawl to a Campelen trawl in 1998 (Lilly 1998). Third, Horne and Schneider (1994) have suggested, based on a physiological model, that cod do not have to be coupled spatially with capelin at small scales but can function as “sit and wait” predators. Even though cod were not closely associated with capelin, distributions were broadly similar and there was potential for contact with prey at scales similar to the scale of the hypothetical daily cod ambit.

Estimates of contact based on bottom-trawl data were relative because of uncertainties about the ability of the gear to adequately sample fish, particularly capelin. Concerns have also been expressed (e.g., Warren 1997) that changes in the timing of the trawl survey between years and also the time taken to complete the survey may influence observed spatial structure. Despite these limitations, our analysis provides a formal statistical assessment of changes in cod distribution and association with capelin over the period of the cod stock collapse. Our interpretation adds to the knowledge of the way that the Newfoundland cod declined and provides a predictive framework for the changes that we might expect to see as the stock recovers. With improved sampling methods, such as acoustic techniques that allow measurements of absolute fish densities at small scales, it may be possible to estimate both potential and realised contact at yet finer scales and apply these values to predictive dynamic models of social behaviour or foraging.

Acknowledgements

We thank Alec MacCall, Philip Dixon, Gareth Lawson, Gordie Swartzman, and three anonymous reviewers for comments on this manuscript. This work was funded by a New Zealand Foundation for Research Science and Technology postdoctoral Fellowship awarded to the first author and the Natural Sciences and Engineering Research Council of Canada Chair in Fisheries Conservation.

References


© 2000 NRC Canada


© 2000 NRC Canada