Description of spatial pattern in seabird distributions along line transects using neighbour $K$ statistics

Richard L. O'Driscoll

Department of Marine Science, University of Otago, PO Box 56, Dunedin, New Zealand

ABSTRACT: Description of spatial pattern is an important step towards understanding the underlying pattern-generating processes, but few statistical techniques adequately describe both the grain and intensity of a spatial distribution. Counts of seabirds collected along line transects are particularly problematic because the data may contain a high proportion of zero counts, and the pattern at increased spatial scales is dependent on how the counts are grouped. Statistics based on Ripley's $K$ function seem well-suited to characterising scale-dependent spatial structure in such data. The spatial distributions of seabirds and fish schools off the Otago Peninsula, New Zealand, are described to illustrate the application of these statistics. Aggregation characteristics of sooty shearwaters $Puffinus griseus$, red- and black-billed gulls $Larus novaehollandiae$ and $L. bulleri$ and black-backed gulls $L. dominicanus$ were similar. Sitting seabirds were highly clustered, with patch lengths between 145 and 2850 m. The distributions of flying seabirds were less crowded, with patch lengths from 135 to 5290 m. Patches of krill-eating red- and black-billed gulls were longer along transects with high salinity structure. This may reflect a change in foraging strategy in response to visible surface fronts. There was no clear association between the distribution of sitting birds and the presence of schools of predatory, pelagic fish.

KEY WORDS: Spatial statistics · Patchiness · Seabird distribution · Otago · Fronts · Nyctiphanes australis · Thyrsites atun

INTRODUCTION

Counts of seabirds at sea typically exhibit spatial structure. Aggregations of birds occur over a wide range of scales, from tens of metres to thousands of kilometres (Hunt & Schneider 1987). The structure and scale of seabird aggregations may be related to a number of factors, including bird behaviour (e.g. Hoffman et al. 1981), fluid dynamics (e.g. Schneider 1991), the distribution of prey (e.g. Schneider & Piatt 1986), and the presence of other marine predators (e.g. Au & Pitman 1988). To understand the role of such processes in determining spatial distribution, it is first necessary to be able to measure and describe spatial pattern (Levin 1992). There are a number of difficulties encountered when attempting to quantify the spatial structure of seabird data.

First, a spatial pattern has 2 distinct characteristics, defined as the grain and the intensity (Pielou 1969). The grain of the pattern refers to the arrangement of areas of high density. Intensity describes the variation in density within the pattern. Tests of randomness based on the variance-to-mean ratio, for example Fisher's coefficient of dispersion (McClymont et al. 1989), measure the intensity of the distribution but reveal nothing about the grain (Andrew & Mapstone 1987).

Second, most indices are dependent on the scale at which the data were gathered. Because the patterns are often scale-dependent (Schneider 1992), the data must be examined over a range of scales. Seabird data are usually a series of time- or distance-integrated (binned) counts. The minimum spatial scale at which the distribution can be assessed is determined by the bin size. For analysis at longer scales data are grouped into larger bins by combining counts (Schneider &
Duffy 1985, Schneider & Piatt 1986). Tests may be sensitive to the manner in which counts are grouped. For example, the following is a series of seabird counts at a bin size of 1:

\[
3 5 2 3 6 6 0 0 7 9 3 1 0
\]

The data may be grouped into counts with bin size of 2 in at least 2 ways, resulting in very different values of the variance-to-mean ratio \(I'\):

\[
\begin{array}{cccccccc}
3 & 5 & 2 & 3 & 6 & 6 & 0 & 0 \\
8 & 25 & 62 & 0 & 16 & 4
\end{array}
I' = 27.1
\]

\[
\begin{array}{cccccccc}
3 & 5 & 2 & 3 & 6 & 6 & 0 & 0 \\
7 & 79 & 6 & 7 & 12 & 1
\end{array}
I' = 47.5
\]

Grouping data also reduces the sample size, decreasing the power of the test to detect significant aggregation at longer scales (Schneider 1990a).

Finally, data sets may be small (number of counts less than 30) and often contain a high proportion of zero counts. This precludes the use of time series methods such as spatial autocorrelation analysis (Schneider & Bajdik 1991) and spectral analysis (Schneider 1989), and geostatistical techniques such as variogram analysis (Begg & Reid 1997), which require large data sets with few zeros.

In this paper I describe statistics, based on the second-order moment (neighbour \(K\)) of the spatial distribution (Ripley 1981), which appear to be well-suited to the description of seabird and other count data. The index is based on the distance between neighbouring individuals along the transect, so is not influenced by the presence of zero counts. The test statistic may be computed at a wide range of scales without grouping the counts, and plots of the statistic as a function of spatial scale provide a description of both the intensity and the grain of the pattern.

To illustrate the utility of this approach, I use neighbour \(K\) statistics to investigate the spatial structure of seabird distributions off the coast of the Otago, New Zealand (Fig. 1). In summer and autumn many species of seabird breed on the Otago Peninsula and feed in the surrounding coastal waters. Coastal circulation off Otago plays an important role in the advection and retention of zooplankton (Murdock 1989), including the krill \textit{Nycitphanes australis}, which is an important food for many seabirds and pelagic fish (Mills 1969, Fenwick 1978, O'Brien 1988, McClatchie et al. 1989, O'Driscoll 1998). A surface band of lower salinity water, originating from riverine input to the south, is sometimes present within the neritic water mass over

![Fig. 1. Map of the study area and a simplified schematic of hydrography within the neritic water mass when a low-salinity band was present. The low-salinity band (hatched) flows northward and is deflected offshore by the Otago Peninsula. An eddy forms, and fronts (indicated by a row of dots) occur where the low-salinity band meets the more well-mixed water in the counter-current. The locations of the 3 main transects are indicated by the dashed lines. All 3 transects were run on 6 March 1995, 28 March 1995, 5 March 1996, 18 March 1996, and 11 April 1996. Transects 1 and 2 were run on 14 February 1994, 21 March 1995 and 26 February 1996. Transects 2 and 3 were run on 6 March 1996. On 19 March 1996 (twice) and 26 March 1996 (4 times) Transect 1 was run repeatedly. The 4 transects on 13 April 1995 were northeast-southwest, intersecting the marked transects at an angle of about 45°](image-url)
the inner and mid-shelf (Fig. 1; Hawke 1992). This low-salinity band usually occurs off Otago following periods of high river flows (O’Driscoll 1997), and visible surface fronts occur where the band meets more well-mixed coastal waters (Fig. 1). Krill appear to accumulate in the low-salinity band (O’Driscoll & McClatchie in press).

Previously, O’Driscoll et al. (1998) showed that seabird distributions off Otago were highly variable over spatial scales from hundreds of metres to kilometres. Here, I test 3 hypotheses about the processes influencing the spatial structure of seabird aggregations. First, I test whether aggregation characteristics were species specific. Second, I test whether the aggregation of krill-eating seabirds was related to the salinity gradients which influence the distribution of prey. Third, I test whether there was spatial association between seabirds and schools of barracouta Thysites atun, a predatory, pelagic fish.

MATERIALS AND METHODS

Survey design. Seabirds were counted from the 15 m RV ‘Munida’ during 12 surveys off the Otago Peninsula in the late summer and autumn of 1994 to 1996. The surveys were part of a side-scan sonar study of schooling fish off Otago (O’Driscoll 1997), and counts were made as the vessel steamed along the acoustic transects at a speed of ~6 knots. Transects were between 6 and 21 km long and most ran due east-west, across the predicted physical gradients, in a design of 3 evenly spaced transects (Fig. 1). Transect 1 was run fourteen times, Transect 2 nine times, and Transect 3 six times. On one day, 4 transects were run northeast-southwest, perpendicular to the bathymetric contours. This gave a total of 33 transects over the 12 days. Position data (within 100 m) were logged every 30 s from the global positioning system (GPS).

Seabird census. Seabirds were counted using binoculars by experienced observers following the method suggested by Tasker et al. (1984). Sitting seabirds were counted continuously in a 200 m swathe (100 m each side) just ahead of the vessel. Birds displaced by the vessel did not usually resettle within the swathe ahead of the vessel, so the probability of recounting birds was low. The counts were summed and recorded every 2 min. Flying seabirds were counted instantaneously at the start of each 2 min block in an area 200 m wide and about 400 m ahead of the vessel (equivalent to the area traversed in 2 min at a vessel speed of 6 knots). Instantaneous counts were used to avoid overestimating the abundance of flying birds (Tasker et al. 1984). Seabirds were usually identified to species level. It was impossible to reliably distinguish between red-billed gulls Larus novaehollandiae and black-billed gulls L. bulleri while steaming so these 2 similar species were grouped. Data were recorded manually by an assistant to the observer.

Collection of stomach contents. Five red-billed gulls and 5 sooty shearwaters Puffinus griseus were shot at sea on 20 March 1996 to collect stomach contents. These 2 species of seabird were selected because they were the most commonly observed in the study area. The sample size was small because these species are protected by law. Permits were obtained from the New Zealand Department of Conservation and the University of Otago Animal Ethics Committee (AEC Application No. 99-95). The stomachs of seabirds were removed at sea and preserved in 10% buffered formalin, and the contents sorted and weighed in the laboratory.

Krill sampling. Zooplankton were sampled from 2 m depth using a modified version of a small Hardy plankton sampler (Miller 1961) with a codend mesh size of 500 µm during 10 of the 12 surveys. Samples were preserved in 5% buffered formalin at sea, and adult and late developmental stages of krill Nyctiphanes australis were counted in the laboratory. Tow lengths were typically between 1 and 2 km, and the density of N. australis was expressed as the number per kilometre tow length.

Physical measurements. Temperature and salinity were measured using a Seabird SBE-19 conductivity temperature depth probe (CTD) during 11 of the 12 surveys. Vertical casts from the surface to the bottom were conducted at 1 nautical mile (1 nmi = 1850 m) intervals along 19 transects. On the remaining transects, seawater was passed continuously over the CTD in a simple flow-through system on the deck of the vessel. The water intake for the flow-through system was about 2 m below the sea surface, so I obtained a continuous trace of salinity and temperature at this depth. For comparison, I only extracted the salinity and temperature measurements at 2 m depth from the processed vertical casts. The vertical structure of the water column is described in detail by O’Driscoll & McClatchie (in press).

Acoustic detection of fish schools. Schools of fish were detected using a 130 kHz Klein 595 side-scan sonar. The towed body was deployed on the starboard side and towed ~10 m behind the vessel and ~3 m below the surface. The shallow depth of the towed body and the wide vertical beam angle of the transducers (40° between ~3 dB points, inclined 10° downwards) meant that all of the upper 15 m of the water column, where most schools occurred, was ensonified (O’Driscoll & McClatchie in press). Both surface and sub-surface schools could be detected out to 100 m either side of the vessel. The Klein 595 graphic recorder generates a high resolution black and white.
paper record. A 'school' was defined according to Misund (1993) as any grouping of fish on the paper record in which individual marks could not be resolved. The species composition of schools was determined visually, by fishing, or by characteristic features of the sonar mark (O’Driscoll & McClatchie in press). I binned the school data into 2-min bins corresponding to the seabird counts.

**Statistical analysis.** A spatial pattern may be described by its first- and second-order properties in the same way as the distribution of a random variable is described by its mean and variance. The first-order property of the pattern is its intensity function \( \lambda \), defined as the expected number of events within a unit area as the area tends towards zero. In practice, the mean density is used as an unbiased estimate of \( \lambda \) (Ripley 1981).

\[
\lambda = N/A
\]

(1)

where \( N \) is the total number of individuals within the study area, \( A \).

The second-order moment, \( K \), may be defined by the function:

\[
K(t) = \lambda^{-1} E[N(t)]
\]

(2)

where \( E[N(t)] \) is the expected number of further individuals within distance \( t \) of any individual in the distribution (Ripley 1981). When estimating \( K(t) \), this expectation is replaced by the empirical average (Cressie 1991). \( E[N(t)] \) is then estimated as the average number of individuals within distance \( t \) of any given individual in the distribution.

\[
E[N(t)] = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} I(t(u_{ij}))
\]

(3)

where \( u_{ij} \) is the distance from individual \( i \) to individual \( j \), and \( I(t(u_{ij})) \) is an indicator function, \( = 1 \) if \( u_{ij} < t \), 0 otherwise.

For a random arrangement of individuals in 2 dimensions \( K(t) = \pi t^2 \). If the individuals are regularly spaced \( K(t) \) will tend to be less than \( \pi t^2 \), whereas for a clustered distribution \( K(t) \) will tend to be greater than \( \pi t^2 \) (Cressie 1991). The significance of \( K(t) \), in terms of departure from a random arrangement, is usually assessed by Monte Carlo procedures (Ripley 1976, 1981; Lotwick & Silverman 1982, Cressie 1991; Andersen 1992).

**One-dimensional analysis.** The distributions of seabirds and fish along a transect were essentially 1-dimensional, because the swath width (200 m) was very small compared with the transect length (typically ~12 km). The range from the vessel to individual birds within the 200 m swath was not recorded. Although the side-scan sonar provided data in 2 dimensions, I did not consider the range dimension for 2 reasons. First, neighbour \( K \) analysis treats the distribution as a spatial point process where the size of the individual is negligible compared to the size of the study area (Ripley 1981, Getis & Franklin 1987). Some individual schools occupied a large proportion (over 50%) of the width of the transect, so the assumption that the size of the individuals is negligible did not hold in 2 dimensions. Second, the detection of schools was range dependent (O’Driscoll & McClatchie in press). Fewer schools than expected by a uniform probability were observed close to the tow-body and there was also a fall-off in the number of schools detected at longer range. I treated both seabird and fish data as 1-dimensional, making the assumption that all individuals were positioned exactly on the cruise track. This assumption introduces a maximum uncertainty in the position of each individual of ±100 m. This uncertainty is smaller than the smallest \( t \) at which the statistics were calculated (see below), so the effect of ignoring range is insignificant.

For a random arrangement of individuals in 1 dimension \( K(t) = 2t \). If the individuals are clustered \( K(t) > 2t \). I used \( E[N(t)] \) (Eq. 3) as the test statistic and assessed whether the distribution of individuals was significantly different from random by a Monte Carlo process. \( E[N(t)] \) retains information about the intensity of the distribution, and, as the average number of neighbours within \( t \) of an individual, is more easily interpreted in a biological sense.

**Correction for edge bias.** There will be a bias in the estimation of \( E[N(t)] \) from Eq. (3) because some individuals may lie less than distance \( t \) from the ends of the transect. The importance of this effect will be greater as \( t \) increases (Ripley 1981). There are at least 4 ways of correcting for edge effects described by Cressie (1991): a guard area correction, where \( E[N(t)] \) is calculated for a smaller area within the study area; a variable edge correction, where \( E[N(t)] \) is calculated only for those individuals that are greater than distance \( t \) from the nearest boundary; a toroidal correction, where the opposite ends of the study area are treated as adjoining; and a weighting correction (Ripley 1981), where each pairwise observation is weighted by the probability that a neighbour at that distance \( u_{ij} \) lies within the study area.

I tested 2 methods of correcting \( E[N(t)] \) for edge effects. The first applied a variable edge correction, and the second applied a weighting factor, \( w_{ij} \), so that:

\[
E[N(t)] = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}^{-1} I(t(u_{ij}))
\]

(4)

The weighting factor was equal to the proportion of the area within distance \( t \) of an individual \( i \) that lay
within the transect. For example, if 10 neighbours are
within \( t \) either side of \( i \) but only 80% of the block
(length 2\( t \)) is within the transect, \( w_i = 0.8 \) and individual
\( i \) is assumed to have \( 10/0.8 = 12.5 \) neighbours
within \( t \).

Effect of binning data. Previous applications of
neighbour \( K \) [e.g. Getis & Franklin 1987] have
required a mapped point process. Here, the data were
a series of binned 2-min counts. The exact position of
each individual along the transect was not recorded so
the bin size determined the increment of distance \( t \).
Bin size (\( B \)) in metres was calculated by dividing the
length of the transect (\( T \)) by the number of 2-min
counts. I assumed that, on average, the individual
around which the neighbours were being counted (the
focal individual) occurred in the centre of its bin, so the
minimum value of \( t \) is half the bin size (\( B/2 \)). That is, all
the neighbours in the same bin are considered to
within \( B/2 \) of a given individual. Distance then incre-
ments by \( B \) as the number of neighbours in adjacent
bins are added. For example, in the count series
0 1 3 2
1 0 1, an individual in the middle bin (in bold) has 1
neighbour in the same bin (\( t = B/2 \)), 5 neighbours
within \( t = 3B/2 \), 6 neighbours within \( t = 5B/2 \), and
7 neighbours within \( t = 7B/2 \). The assumption that the
focal individual always occurs in the centre of its bin
means that there will be an uncertainty associated with
\( t \). A neighbour classified as being within distance \( t \)
of the focal individual may be within distance \( t \pm B/2 \),
depending on the actual position of the focal individual
within its bin. This uncertainty is proportionally
greater at small \( t \).

The side-scan sonar provided a continuous record-
ing, so the exact position of each fish school along a
transect could be determined. To test the effect of bin
size I binned data from the transect with the most
schools of barracouta (19 March 1996, Transect 1) into
4 different size bins (\( B = 40, 80, 160 \) and 320 m) and
calculated the neighbour \( K \) statistics. At the smallest
bin size there was usually only 1 school per bin, so the
data may be regarded as a mapped point process.

Monte Carlo test for significance. I generated 999
further realisations of the data for each transect by ran-
domly re-allocating each individual to a bin along the
transect, and used a Matlab® routine to calculate \( E[N(t)] \) for the data and the 999 randomisations at dis-
tances, \( t = B/2, 3B/2, 5B/2 \) ... \( T \). The output of the rou-
tine was \( E[N(t)] \) for the data, and the minimum, maxi-
mum and mean values of \( E[N(t)] \) for the randomised
data, at each value of \( t \). The routine also calculated the
proportion of the 1000 realisations (data and 999 ran-
domisations) where the value of \( E[N(t)] \) was greater
than or equal to the value of \( E[N(t)] \) from the data. Indi-
viduals were considered to be clustered at distance
scale \( t \) when the proportion of realisations (\( p \)) in which
\( E[N(t)] \) exceeded the value from the data was less than
0.05. This is a one-sided test because regular arrange-
ments of individuals \( E[N(t)] \) from the data less than
\( E[N(t)] \) expected from a random arrangement were
seldom observed. The Matlab® code is available on
request from the author.

Display of results. The results of the analyses are
presented as plots of \( L(t) \) as a function of \( t \), where \( L(t) \) is
equal to \( E[N(t)] \) from the data minus the mean \( E[N(t)] \)
from the 999 randomisations [after Cressie (1991)]; note
that \( L(t) \) is equivalent to Cressie’s \( L(h) \) multiplied by the
intensity, \( \lambda = N/T \). \( L(t) \) is the average number of ‘extra’
neighbours within distance \( t \) of any given individual in
the distribution. The ‘extra’ neighbours are those
which were not expected if all the individuals were
arranged randomly along the transect. For a random
distribution \( L(t) = 0 \). Where individuals are clustered
\( L(t) > 0 \). I also plotted the upper confidence interval of
\( L(t) \), which is the 95th percentile of \( E[N(t)] \) from the
randomisations minus the mean \( E[N(t)] \), and repre-
sents the significance level of \( p = 0.05 \).

Bivariate analysis. I compared the distributions of
seabirds and fish using a bivariate \( K \) analysis (Lotwick
& Silverman 1982, Andersen 1992). The bivariate rou-
tine calculated the expected number of neighbours,
\( E[N(t)] \), from one group for each individual from a sec-
ond group, in this case the number of seabird ‘neigh-
bours’ within \( t \) of each school of fish. The counts of the
2 groups had to be along the same transect and binned
in the same manner. The Monte Carlo testing proce-
dure was carried out by randomly re-allocating all
the individuals in the first group only, and then re-cal-
culating the \( E[N(t)] \). This is not a true test of indepen-
dence because the second-order properties of one of
the distributions are changed during the randomisa-
tions (Lotwick & Silverman 1982). The hypothesis
being tested is whether there are significantly more
seabirds associated with a given individual fish school
than would be expected if the seabirds were distrib-
uted randomly. This hypothesis makes biological sense
if seabirds are aggregating in response to fish schools.
Results are again presented as plots of \( L(t) \) as a function
of \( t \).

RESULTS

Overall spatial distribution of seabirds

Sitting and flying seabirds were observed on all
transects. The most commonly observed species, in
decreasing order of abundance, were sooty shearwaters
(32.9% of 11071 total seabirds), red- and black-billed
gulls (31.3%), black-backed gulls \textit{Larus dominicanus}
(18.3%), white-fronted terns \textit{Sterna striata} (5.0%),
Distance along Transect (km)

Fig. 2. Distributions of (a) total flying and (b) total sitting seabirds along Transect 2 on 26 February 1996. The number of seabirds in each 2-min count is plotted as a function of distance along the transect.

Stewart Island shags *Leucocarbo chalconotus* (3.1%), Buller’s mollymawks *Diomedea bulleri* (3.1%), spotted shags *Sictocarbo punctatus* (2.2%), Cape pigeons *Daption capense* (1.6%), and shy mollymawks *Diomedea cauta* (1.5%). The remaining 1.0% of birds was made up of 10 species.

Birds were distributed throughout the study area. Over all days there were no obvious spatial trends in species composition or abundance (details in O’Driscoll et al. 1998). Multispecies groups were frequently observed.

Description of spatial pattern

The distribution of seabirds along a transect was typically heterogeneous (Fig. 2) and this was reflected in the plots of $L(t)$ as a function of $t$ (Fig. 3). Along Transect 2 on 26 February 1996, both total flying and sitting seabirds (all species combined) were significantly more clustered than would be expected in a random arrangement over a wide range of scales.

A description of the grain of the distribution pattern is obtained from the shape of the plot of $L(t)$ as a function of $t$. The distance to the first peak in $L(t)$ provides a description of the characteristic spatial scale of clustering, or patch length, independent of the intensity. Patch length was defined as the first distance $t$ at which there was a maximal significant difference between the observed number of neighbours and the expected number of neighbours from a random arrangement. On Transect 2 on 26 February 1996, $L(t)$ for flying seabirds peaks at a distance scale of 1350 m (Fig. 3a). The first peak in $L(t)$ for sitting seabirds occurs at 465 m (Fig. 3b). This indicates that sitting seabirds were aggregated in a smaller spatial group than flying seabirds along this transect (Fig. 2). As a comparison, maximal values of the variance-to-mean ratio ($\lambda$) [patch scale in Schneider & Piatt 1986] were at a spatial scale of 4200 m for both sitting and flying seabirds along this transect.

The height of the first peak in the plot of $L(t)$ may be used as a measure of the intensity of the distribution, or crowding. Crowding was defined as the difference between the observed and expected number of neighbours at the scale of the patch length. Crowding provides a measure of the number of individuals which are grouped together in a patch. The peak value of $L(t)$ for sitting seabirds (77.3) is higher than the value of crowding for flying seabirds (13.5) (Fig. 3), indicating that there were more individuals in a patch of sitting seabirds than in a patch of flying seabirds, even though the patch of sitting seabirds has smaller spatial dimensions.

The range of scales over which clustering appears significant is dependent on the intensity of the distribution. When many individuals occur in a group (high intensity) the observed number of neighbours within distance $t$ continues to significantly exceed the expected number of neighbours for values of $t$ much larger than the patch length (e.g. sitting seabirds in Fig. 3b).

The distance to subsequent peaks on plots of $L(t)$ versus $t$ are a function of the distance between high counts in the distribution. An example is the peak in Fig. 3b at $t = 9145$ m. This corresponds to the distance between the very high counts at the start of the transect and the much smaller counts that were recorded around 9 km along the transect.

In Fig. 3 $L(t)$ is plotted for values of $t$ up to the total length of the transect ($T$). At distances greater than $T/2$ ($t = 6620$ m in this example), the effect of edge bias is great and results must be interpreted with caution (Cressie 1991).

Values of $E[N(t)]$ and $L(t)$ as a function of $t$ for the 2 edge-bias corrections are plotted along with the uncorrected values for the distribution of flying seabirds along Transect 2 on 26 February 1996 in Fig. 4. The
Fig. 3. Plots of \( L(t) \) as a function of scale for distributions of: (a) total flying, and (b) total sitting seabirds, along Transect 2 on 26 February 1996 (data in Fig. 2). The solid line represents the observed number of neighbours per individual along the transect minus the expected number of neighbours if the individuals were distributed randomly. If the individuals are distributed randomly \( L(t) = 0 \) (heavy dashed line). The distribution is significantly clustered (\( p < 0.05 \)) at scale \( t \) when \( L(t) \) is above the 95% confidence level of \( L(t) \) from the 999 randomisations of the data, shown by the short-dashed line. The patch length was the value of \( t \) at the first peak in the \( L(t) \) plot. The height of the peak was used as a measure of crowding.

variable edge correction could only be calculated up to half the transect length because at \( t \) greater than \( T/2 \) none of the counts were further than distance \( t \) from an end of the transect. The corrected values of \( E[N(t)] \) increase linearly with \( t \) while the uncorrected values level off as \( t \) approaches the length of the transect. There were also differences in plots of \( L(t) \) although the distance to the first peak (patch length) was consistent for the uncorrected and weighted data.

The variable-width correction is useful because it does not make any assumptions about the distribution of individuals beyond the transect. However, the sample size from which \( E[N(t)] \) is estimated decreases non-linearly with increasing distance. This causes problems during Monte Carlo testing because sample sizes may differ widely between the observed distribution and the randomisations. At higher values of \( t \), the observed value of \( E[N(t)] \) was usually estimated from a smaller sample size than the values of \( E[N(t)] \) from the 999 randomisations, meaning \( L(t) \) was negative (Fig. 4b).

The weighting correction assumes that the areas beyond the ends of the transect have the same average density of individuals as were observed along the transect. This correction had the effect of increasing the patch length, the crowding, and the range of spatial scales at which the spatial pattern was considered significant (Fig. 4b). I do not believe that the assumption of equal density within and beyond the transect can be accepted without knowledge about the distribution of seabirds outside the study area.

I chose to use an unweighted analysis, in effect assuming no individuals occurred beyond the ends of the transect. Because the same assumption was made for the data and all randomisations, plots of \( L(t) \) described the distribution of the data well at distances up to half the length of the transect. The actual values of \( L(t) \) may be misleading at higher \( t \) if individuals also occur outside the study area.

There were no detectable differences between the shape of the \( L(t) \) plots calculated using 4 different bin sizes between 40 and 320 m over the scales of interest (Fig. 5). The major uncertainty associated with binning the data was the uncertainty in distance \( t \), equal to half the bin size. Typically, \( B/2 \) was about 150 m, which is of the same order as the uncertainty in the GPS positions.

Aggregation characteristics of different species

Flying and sitting seabirds were significantly aggregated along almost all transects (Table 1), but there was wide variation in the spatial structure of the aggregations. There were differences in aggregation characteristics between days, between transects, and even between repeated runs along the same transect on the same day. Because of this variation, I used paired-sample tests to compare aggregation characteristics of different groups along the same transects.

Sitting seabirds tended to occur in larger groups than flying seabirds (Table 1). Crowding of sitting seabirds was significantly greater than crowding of flying seabirds for total counts (Wilcoxon paired-sample test, \( Z = 2.84, p = 0.004 \)), red- and black-billed gulls (\( Z = 2.80, p = 0.005 \)) and black-backed gulls (\( Z = 2.75, p = 0.006 \)), but not for sooty shearwaters (\( Z = 0.89, p = 0.38 \)). In general, aggregations of sitting birds also occurred over a smaller spatial scale than groups of flying seabirds (Table 1), but the patch lengths were only significantly shorter for sooty shearwaters (Wilcoxon paired sample test, \( Z = -2.79 \)).
Aggregation characteristics along a transect did not appear to be species-specific. Friedman's paired-sample tests did not reveal significant species differences between patch length or crowding for either flying (patch length, $\chi^2 = 6.41$, $p = 0.08$; crowding, $\chi^2 = 3.17$, $p = 0.21$) or sitting counts (patch length, $\chi^2 = 5.16$, $p = 0.08$; crowding, $\chi^2 = 2.06$, $p = 0.27$). There was overlap between values of crowding and patch length for all 3 commonly observed species (Table 1).

Aggregation of krill-eating seabirds in relation to salinity gradient

A low salinity band of surface water was detected during 6 of the 11 CTD surveys (e.g. Fig. 6a). There was little variation in surface salinity and no fronts were observed on the other 5 days (e.g. Fig. 7a). The association between krill and hydrography is described in detail by O’Driscoll & McClatchie (in press). Briefly, on days when the low-salinity band was detected, mean catches of krill from 1 to 2 km net tows within this band (mean density = 503 krill km$^{-1}$, $n = 48$ tows) were over 5 times higher than catches in surrounding waters (94 krill km$^{-1}$, $n = 28$ tows) (Fig. 6b, Mann-Whitney U-test, $p = 0.0003$). Densities of krill in tows which crossed the front at the edge of the low-salinity band (710 krill km$^{-1}$, $n = 33$ tows) were not significantly different from those in the surrounding waters (503 krill km$^{-1}$, $n = 33$ tows) (Mann-Whitney U-test, $Z = -1.51$, $p = 0.13$), or black-backed gulls ($Z = -0.77$, $p = 0.44$).

There were no significant differences between patch lengths of flying and sitting total seabirds ($Z = -1.67$, $p = 0.09$), red- and black-billed gulls ($Z = -1.51$, $p = 0.13$), or black-backed gulls ($Z = -0.77$, $p = 0.44$).

### Table 1. Summary of aggregation characteristics of the total seabirds and of the 3 most abundant individual species (sooty shearwaters Puffinus griseus, red- and black-billed gulls Larus novaehollandiae and L. bulleri, and black-backed gulls L. dominicanus). Other species were too rare to compare their aggregation characteristics over a number of transects. Values of patch length and crowding were only calculated along transects where more than 5 individuals were observed. When fewer than 5 individuals were counted along a transect, it was impossible to detect significant aggregation because of the small sample size.

<table>
<thead>
<tr>
<th>Species</th>
<th>Count type</th>
<th>No. of transects with &gt;5 birds (max. = 33)</th>
<th>No. of transects with non-random aggregation</th>
<th>Patch length (m)</th>
<th>Crowding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total seabirds</td>
<td>Sitting</td>
<td>33</td>
<td>30</td>
<td>480</td>
<td>9.0-1030</td>
</tr>
<tr>
<td></td>
<td>Flying</td>
<td>33</td>
<td>31</td>
<td>800</td>
<td>1.4-111</td>
</tr>
<tr>
<td>Red- and black-billed gulls</td>
<td>Sitting</td>
<td>22</td>
<td>22</td>
<td>530</td>
<td>9.0-665</td>
</tr>
<tr>
<td></td>
<td>Flying</td>
<td>26</td>
<td>26</td>
<td>890</td>
<td>5.0-82</td>
</tr>
<tr>
<td>Black-backed gulls</td>
<td>Sitting</td>
<td>22</td>
<td>20</td>
<td>190</td>
<td>6.6-464</td>
</tr>
<tr>
<td></td>
<td>Flying</td>
<td>26</td>
<td>19</td>
<td>525</td>
<td>5.0-37</td>
</tr>
<tr>
<td>Sooty shearwaters</td>
<td>Sitting</td>
<td>20</td>
<td>19</td>
<td>480</td>
<td>5.8-122</td>
</tr>
<tr>
<td></td>
<td>Flying</td>
<td>33</td>
<td>27</td>
<td>800</td>
<td>1.0-122</td>
</tr>
</tbody>
</table>
48 tows with high densities of krill. Consequently, the correlation between the number of sitting red- and black-billed gulls and the abundance of krill was weak, but highly significant, at the 1 to 2 km scale of the plankton tows (Spearman’s rank correlation, n = 191 tows, \( \rho = 0.321, p < 0.0001 \)).

There was only weak evidence that sitting red- and black-billed gulls accumulated at salinity fronts. Values of salinity at 2 m depth at 1 nmi (1850 m) intervals along each transect were extracted from the CTD vertical casts or surface trace, and the change in salinity (dS) over each interval was calculated. Seabird counts were grouped into larger bins corresponding to the 1850 m intervals. High counts occurred at the frontal boundary on some transects, but the correlation between sitting red- and black-billed gulls and dS was very weak at this spatial scale (Spearman’s rank correlation, n = 214 counts, \( \rho = 0.157, p = 0.02 \)). Often, peak numbers of gulls were observed as the low-salinity side of the front, within the band of low-salinity water (Fig. 6).

Aggregation characteristics of sitting red- and black-billed gulls were correlated with the presence of fronts. I calculated the total salinity gradient along each transect by summing the absolute values of dS for each 1 nmi interval. The summed salinity gradient gave a measure of the physical structure along a transect. Patch lengths of sitting red- and black-billed gulls were longer (Spearman’s rank correlation, n = 19 transects, \( \rho = 0.64, p = 0.007 \)), and the crowding was greater (\( \rho = 0.57, p = 0.016 \)), along transects with high salinity structure. This indicates that, along transects when fronts were present, groups of sitting gulls were more spread out and contained more individuals than groups observed along transects when weaker salinity gradients were detected (e.g. Figs. 6 & 7). Aggregation characteristics of sitting black-backed gulls and sooty shearwaters were not significantly correlated with summed salinity gradient.

Association of seabirds with schooling barracouta

The side-scan sonar was not deployed on 14 February 1994. A total of 208 schools of fish were identified on the remaining 11 d. Most (191) were schools of barracouta. Schools of jack mackerel Trachurus murphyi (11) and slender tuna Allothunnus fallai (6) were also observed, but there were too few schools of these species to analyse their distribution or association with
Fig. 6. Observations on 6 March 1995, when a low-salinity band of surface water was present in the study area. Distribution of (a) salinity, (b) net catches of *Nyctiphanes australis*, and (c) sitting red- and black-billed gulls *Larus novaehollandiae* and *L. bulleri*.

Fig. 7. Observations on 18 March 1996, when there was little surface salinity structure and no fronts were detected. Distribution of (a) salinity, (b) net catches of *Nyctiphanes australis*, and (c) sitting red- and black-billed gulls *Larus novaehollandiae* and *L. bulleri*.
seabirds. Schools of barracouta were identified along 21 transects.

Schools of barracouta were significantly aggregated (p < 0.05) on 13 of the 14 transects when more than 5 schools of barracouta were detected. As for seabirds, aggregation characteristics were variable. Patch lengths were between 525 and 3230 m (median = 1510 m). Values of crowding revealed that, within a patch, each school had between 1.5 and 15.8 more neighbouring schools than would be expected if the schools were distributed randomly.

Only associations between sitting seabirds and barracouta were considered. A plot of bivariate L(t) from Transect 1 on 19 March 1996 is shown in Fig. 8 along with the distributions from which the statistics were calculated. This transect was chosen as an example because it had a large number of barracouta schools. More sitting seabirds were associated with every barracouta school than would be expected if the sitting birds were distributed randomly. The association was significant (p < 0.05) at distances between 160 and 2720 m (Fig. 8). The first peak in the bivariate L(t) plot is a measure of the distance between high counts of schools and sitting birds, and may be considered as the scale of maximum association. Sitting birds may be aggregating in response to schools of barracouta along transects where the scale of association is short (indicating that high counts of sitting birds occurred close to high numbers of schools). In Fig. 8, maximum association occurs at a distance of 800 m. The height of the peak in the bivariate L(t) plot is related to the intensity of the distribution of sitting birds.

There was no characteristic scale of association between total sitting seabirds and schools of barracouta. The first peaks in the L(t) plots occurred at distances between 165 and 10765 m (median = 1440 m). Distances greater than 1 km separated high counts of total sitting birds and schools of barracouta on 11 of the 18 transects in which the scale of association was significant. Sitting red- and black-billed gulls (median association scale = 1190 m) tended to be more closely associated with schools of barracouta than groups of sitting black-backed gulls (median association scale = 1770 m) or sitting sooty shearwaters (median association scale = 2730 m), but the difference was not significant (Friedman's paired-sample test, $\chi^2 = 1.60$, p = 0.45).

**DISCUSSION**

Neighbour K statistics can be used to identify and describe spatial pattern in count data collected along...
line transects. The analysis does not require the position of individuals to be mapped precisely, as long as the bin size at which the counts are grouped is small relative to the spatial scales of interest. This method of analysis was well-suited to description of spatial structure in seabird and fish data, and may also be useful for describing patterns in other 1-dimensional data sets.

Like other spatial statistics, neighbour K assumes stationarity (Ripley 1981). This requires all samples (counts) along a transect to be drawn from the same statistical distribution, regardless of time or location. Clearly, if seabirds are aggregating in response to fronts or patches of prey this will not be the case. In practice, the assumption of stationarity requires some large-scale homogeneity over the study area (Legendre & Fortin 1989). Most common species of seabird were encountered throughout the study area (O’Driscoll et al. 1998). Because of the high temporal and spatial variability in the physical environment and in the distribution of prey (O’Driscoll & McClatchie in press), groups of seabirds were likely to occur anywhere along a transect. I assumed the distribution of seabirds was weakly stationary (Ord 1979), where the spatial pattern depends only on the relative distance between individuals, and not their absolute positions in space and time (Simard et al. 1992).

Patch length and crowding provide measures of the grain and intensity of a spatial distribution along a transect. However, the results are 1-dimensional. Patch length cannot be used as a measure of overall patch dimensions because transects may intersect an aggregation of seabirds at any position or angle. Where seabirds aggregate in response to physical structure, their distribution is likely to be anisotropic (e.g. Begg & Reid 1997) and patch length will depend on transect orientation (Schneider & Duffy 1985). In this study, transects were perpendicular to the major physical gradients and probably crossed the short axis of seabird aggregations.

Red- and black-billed gulls, black-backed gulls and sooty shearwaters were all significantly clustered at scales from hundreds of metres to kilometres. Clustering at similar (<5 km) scales has been observed in other species, such as prions Pachyptila spp. (Goss et al. 1997), but in many studies seabird counts are grouped into bins which are too large to detect patchiness at small scales (e.g. Decker & Hunt 1996). If spatial pattern and process are coupled (Hunt & Schneider 1987), then it is likely that the processes which generate seabird patchiness off Otago operate over similar scales to the observed patterns in seabird distributions.

There were no characteristic, species-specific differences in the spatial structure of seabird aggregations off Otago. Patch length and crowding varied within species, and this intra-specific variation may have obscured differences between species. Variability in aggregation characteristics between transects may reflect different responses of seabirds to their environment, or spatial and temporal variability in the environment itself. An example in this study was the correlation between the spatial structure of aggregations of red- and black-billed gulls and salinity gradient.

The increased patch length and crowding of sitting red- and black-billed gulls along transects with high salinity structure may represent differences in the way gulls search for krill prey. During an aerial survey of schooling fish (O’Driscoll 1997) the front at the edge of the low-salinity band was clearly visible as a boundary between different coloured water masses. A foam line was also present. Seabirds may use the visible surface front at the edge of the low-salinity band as a cue to help locate patches of krill (Decker & Hunt 1996). The increased patch lengths of aggregations of sitting gulls along transects with high salinity gradient could result if birds initially aggregate at the frontal boundary, and then spread out, foraging into the low-salinity band where surface densities of krill were high. Alternatively, seabirds may be able to detect the low-salinity band directly, and restrict their foraging to this region, accumulating in large groups. When salinity structure was weaker, no surface fronts were visible, and it is unlikely birds were aggregating in response to the physical environment. Along these transects, smaller, shorter patch-length groups may result from attraction to other feeding birds (e.g. Hoffman et al. 1981).

As well as being more spatially predictable, patches of krill prey may also be more temporally persistent when salinity gradients are stronger. The implication from the differences in crowding between sitting and flying counts is that seabirds accumulate when sitting on the water to feed. Aggregation characteristics of sitting seabirds may change over time (Hoffman et al. 1981), with larger, longer groups forming as more individuals gather. If krill patches are more persistent when salinity gradients are stronger, aggregations of sitting red- and black-billed gulls will have longer to develop, potentially leading to increased patch length and crowding. It would be possible to test this hypothesis, monitoring the formation and dispersal of seabird and prey patches, by running repeated transects. Initial results, from repeated runs of Transect 1 on 26 March 1996, indicate that patch length and crowding of red- and black-billed gulls aggregations increased over ~2 h, providing support for the idea that seabird aggregation characteristics are temporally variable.

Much recent effort has focussed on how seabird abundance is related to fluid dynamics (reviews in Hunt & Schneider 1987, Schneider 1991), in particular the elevated densities of seabirds at fronts (e.g. Scnei-
O’Driscoll: Spatial pattern in seabird distributions


Begg GS, Reid JB (1997) Spatial variation in seabird density at tidal fronts surrounding the Pribilof Islands, Alaska USA in Marine Ecol Prog Ser 139:1–10


Hawke DJ (1996) Relatively infrequent seabird aggregation at nearshore fronts and tidal plumes at locations around Banks Peninsula, New Zealand. Notornis 43:66–70


