

Spatial structure of planktonic ciliate patches in a tropical coastal lagoon: an application of geostatistical methods

Celia Bulit^{1,*}, Carlos Díaz-Avalos², Martha Signoret¹, David J. S. Montagnes³

¹Departamento El Hombre y su Ambiente, Universidad Autónoma Metropolitana-Xochimilco, Calzada del Hueso 1100, 04960 México DF, México

²Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas, Universidad Nacional Autónoma de México, Apartado Postal 20-726, 01000 México DF, México

³Port Erin Marine Laboratory, School of Biological Sciences, University of Liverpool, Port Erin, Isle of Man, IM9 6JA, British Isles

ABSTRACT: The distribution of ciliates in a Mexican coastal lagoon was studied. The 4 goals were to: examine small-scale (<100 m) patches; indicate how geostatistical techniques can be used to examine these patches; make inferences concerning ciliate distribution and behaviour in the lagoon using geostatistical techniques; and assess geostatistics as a method for modelling ciliate distributions. Underlying these goals we attempt to make geostatistical techniques accessible to the non-expert. We provide an overview of the methodology, references to the geostatistical literature, and use our system as an example. Ciliates were sampled in a 40 × 40 m grid, divided at 10 m intervals; the grid was further divided into subsets, to determine 1 to 10 m scale variation. Between 30 and 35 points were sampled on 2 occasions (January and October). Ciliates were preserved with Lugol's iodine; abundance and species composition were determined by standard inverted microscopy. The work focused on 4 abundant ciliate species. We indicate, using the variographic analysis, that the abundance of 3 of the 4 ciliates is neither randomly nor homogeneously distributed, but exhibits a structured small-scale patchy distribution. We indicate that species-specific patterns of patchiness exist in stratified and in mixed waters, supporting the notion of behavioural niche-separation of planktonic ciliates. Patches of <13, <18, and <77 m were formed by *Lohmaniella oviformis*, *Tintinnopsis* sp. and *Strombidium* sp., respectively. In contrast, *Pleuronema* sp. formed patches below the detection limits of the analysis (<1 m). Using geostatistical techniques, we established variograms and used them to model ciliate distribution and predict ciliate behaviour. Distribution maps were then generated that depicted the shape, distinctness, and gradient of the different patches. After analysing the data, we proposed a working definition of a 'ciliate patch': regions with abundance above the cut-off of the upper quartile from the kriging prediction model. Finally, error-maps were developed, indicating the coefficient of variation of the predicted distributions. We conclude that geostatistical analysis is a powerful tool to examine microzooplankton at small-scales, and we support its further application in the field.

KEY WORDS: Microzooplankton patchiness · Variographic analysis · Ordinary kriging · *Tintinnopsis* · *Lohmaniella* · *Strombidium* · *Pleuronema* · Mexico

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Plankton are distributed in patches (Cushing 1953, Steele 1978) which can be conceptualised as rare dis-

continuous local regions of high density, sparsely distributed over a larger background of low density. However, the extent and abundance of such patches are difficult to quantify. Geostatistics is a powerful set of tools that can assess patchiness (Legendre & Fortin 1989, Legendre & Legendre 1998). Underlying this

*Email: cbulit@cueyat1.uam.mx

approach is the expectation that, on average, samples close together have values (e.g. of abundance) more similar than those further apart. This methodology has been applied to examine the spatial distribution of a number of aquatic populations in various systems (e.g. Mackas 1984, Legendre & Troussellier 1988, Freire et al. 1992, González-Gurriarán et al. 1993, Pelletier & Parma 1994, Maravelias et al. 1996, Pinel-Alloul et al. 1999, Pinca & Huntley 2000, Roa & Tapia 2000, Passy 2001, Rueda 2001, Defeo & Rueda 2002), but to our knowledge, it has not been applied to small-scale distributions of microzooplankton.

Many works that use geostatistics either provide detailed methodology or assume expert knowledge on the part of the reader. While we recognise the existence of an entire literature devoted strictly to these details (e.g. Isaaks & Srivastava 1989, Cressie 1993, Goovaerts 1997, Armstrong 1998, Chilès & Delfiner 1999 and references therein), with some notable exceptions (Legendre & Fortin 1989, Liebhold et al. 1993), there is a lack of simple 'how-to' instructions on the processes of geostatistics, using case-specific examples. Thus, as a basis for our own work, and to encourage other aquatic microbial ecologists to apply these techniques, we provide an overview of the methodology, references to the appropriate literature for geostatistical details, and finally use our system as an example. Specifically, we refrained from complicating this description with jargon and detailed formulae. In taking this approach, we deviate somewhat from a standard presentation; for clarity some specific methods are incorporated into the 'Results and discussion' section.

This study has 4 main goals: (1) to examine the small-scale (<100 m) distribution patterns of ciliates in a coastal lagoon; (2) to assess how geostatistical techniques can be applied to examine and quantify small-scale patches of plankton; (3) to use these geostatistical techniques to make inferences concerning ciliate distribution and behaviour in the lagoon; and (4) to assess the ability of these techniques to model ciliate distributions in the lagoon.

MATERIALS AND METHODS

General overview of geostatistics. It is simple to create contour/density maps from grid data (e.g. Fig. 1a,b) using many existing software packages (e.g. Surfer, Golden Software; S-Plus, MathSoft). Such maps provide some information regarding microplankton distribution (e.g. Montagnes et al. 1999) but may be inadequate or misleading as they: (1) use simplistic methods to create smoothed contours; (2) typically, do not provide a method to interpolate between contours or extrapolate beyond the data set; (3) do not provide

detailed, quantifiable data regarding patch composition; and (4) do not provide an estimate of the error associated with predicted values (e.g. of abundance). Geostatistical analysis provides a method of solving these problems, by assuming an underlying probabilistic model that accounts for the pattern of spatial variability (Cressie 1993, Chilès & Delfiner 1999). Geostatistics is based on the concept of random spatial distributions; in our case, ciliate assemblages are viewed as random patches in space, and the data collected are a discrete observation of that distribution.

Here, we offer a simple step-by-step description of the basic principles underlying this analysis, and we assume that the variogram is valid over the area covered by the population (i.e. we assume a second order stationarity hypothesis; see Roa & Tapia 2000).

First, data are collected over a region, ideally, but not necessarily, forming a regular grid (Fig. 1a). Next, comparisons are made among data (in our case, ciliate abundance) at discrete distances; e.g. over all 1 m intervals (short arrow, Fig. 1a), all $\sqrt{2}$ m intervals (medium arrow, Fig. 1a), all 2 m intervals (long arrow, Fig. 1a), etc. For each discrete distance (referred to as a lag, h), data-pairs can be plotted on a scattergram to examine the strength of their association (Fig. 1c). If 2 estimates from grid-points have an identical value, then that point will fall on the 45° slope of the scattergram, and if they differ substantially, they will fall far from this line. An estimate of the variance of these data is then computed for the scattergram at each lag following:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \quad (1)$$

where $\hat{\gamma}(h)$ is the estimated variance for any lag (h), $z(x_i)$ is a datum-value at any one point, $z(x_i + h)$ is a datum value at lag (distance) h from $z(x_i)$, and $N(h)$ is the number of pairs of points separated by h . Ideally, to reduce biases, only lags less than half the maximum lag are used, and all points of the variograms include >30 data pairs (Isaaks & Srivastava 1989).

Each variance value, $\hat{\gamma}(h)$, is then plotted against the associated lag (h) to produce an 'empirical variogram' (Fig. 1d). Patchy systems typically follow the pattern illustrated in Fig. 1d: as distances increase between samples, the variance increases until variance remains constant regardless of the lag, i.e. 2 samples very close together are probably similar; as distance between samples increases, they become less similar. Ultimately, regardless of distance, there is a constant dissimilarity between samples.

To assess the shape of the empirical variogram, functions are fit to the data to provide a 'model variogram'. There are a number of potential functions (Isaaks & Srivastava 1989, Goovaerts 1997, Armstrong 1998), but

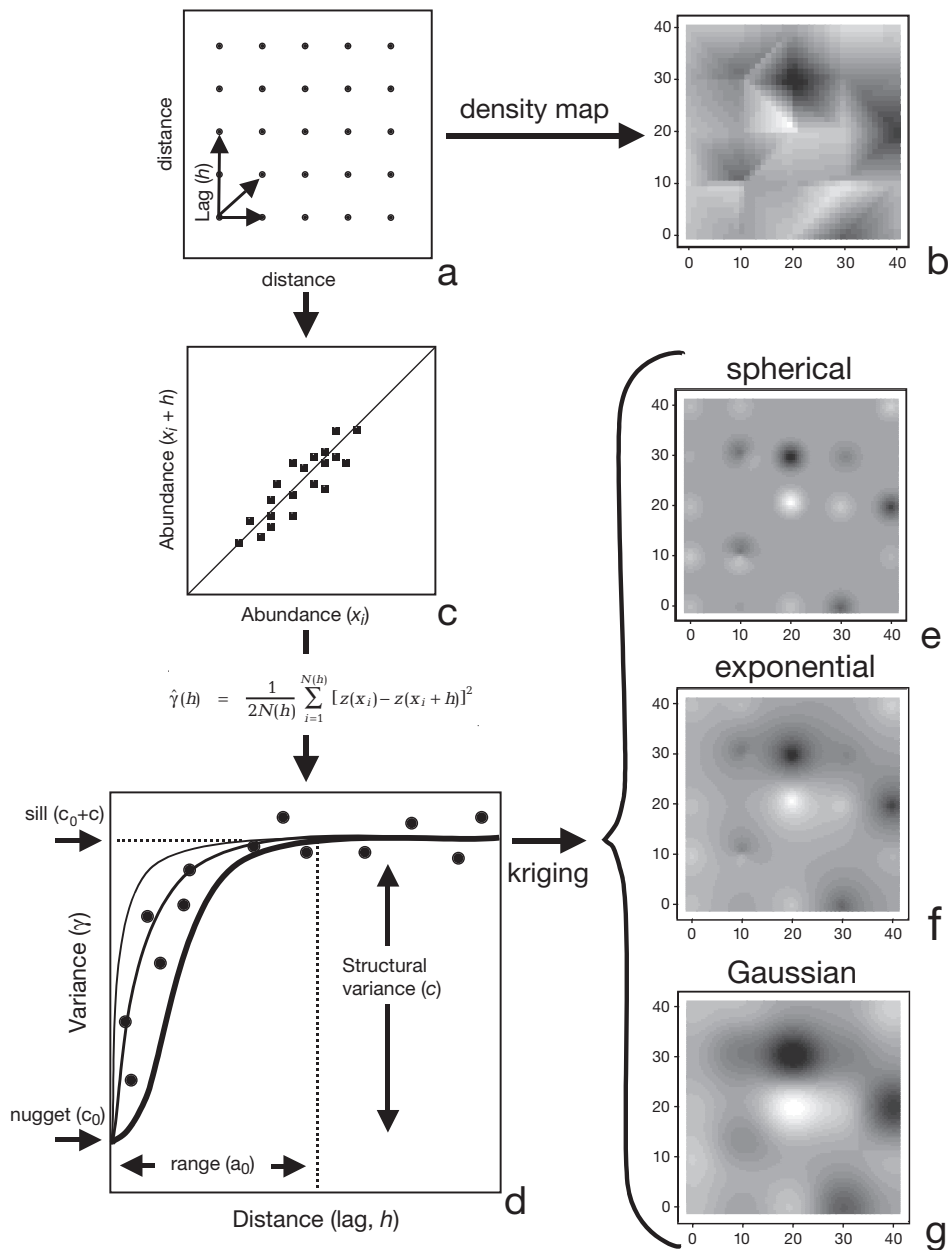


Fig. 1. Schematic description of the geostatistical method, using ciliate distribution as an example. Ciliates are collected from points, generally forming a grid (a). To assess for patchiness, data can be processed by a simple contour-mapping function, to produce contour (or density) plots (b). If patchiness appears to exist, geostatistical analysis is then applied. First, ciliate abundance at points separated by a common distance (lag, h) are calculated; data pairs can then be plotted on a scattergram (c). Next, this illustration of similarity is quantified: for a single lag, differences are summed, squared, and divided by twice the number of pairs to yield a variance (γ , equation). This process is repeated for each lag; e.g. the short, medium, and long arrows in (a) are examples of the first 3 lags for which γ is calculated. Each of these estimates of γ is then plotted against the lag to produce an empirical variogram (points in d). Then, a model is fit to the variogram data (lines in d); these are used to predict abundance at unsampled points and to assess the behaviour of the ciliate (see 'Results and discussion' on species). Commonly, exponential, spherical, or Gaussian models are fit to the data (thin, medium, thick lines, respectively, in d). There are 3 main components of the variogram: the nugget (c_0); the range (a_0); and the sill ($c_0 + c$), composed of the nugget variance (c_0) and the structural variance (c) (see 'General overview of geostatistics'). Once models are fit to the variogram, they are used to map ciliate abundance by kriging (see 'Cross-validation and kriging'). Note that each model produces a different predicted distribution (e–g, see 'General overview of geostatistics')

the 3 most commonly used are spherical, exponential, and Gaussian functions (Fig. 1d). The spherical function indicates a system where patches are more structured (Fig. 1e); the exponential function indicates that patches have fuzzy edges (Fig. 1f); the Gaussian function indicates an extremely continuous distribution with patches that fade off smoothly (Fig. 1g). Thus, the shape of the function (or model) fit to the variogram also provides useful information concerning the structure of the system and the behaviour of the organism (ciliates in our case).

The relationship illustrated in Fig. 1d indicates features of the patchiness: the distance between the smallest scale sampled and where the data become asymptotic (Fig. 1d: range, a_0) indicates the range over which patches will be observed; the asymptote (Fig. 1d: sill, $c_0 + c$) provides an indication of the variability of the data in the sampling region; and the y -intercept (Fig. 1d). Note, the sill is composed of the structural variance (c) and the nugget (c_0). The nugget is an unfortunate term stemming from the geological roots of the methodology rather than being named after a general phenomenon; it indicates the difference between 2 samples at an infinitely small distance from each other, either due to natural variation, experimental error, or both. These 3 aspects of the data tend to exist for most spatial distributions, but the shape of the variogram varies depending on how data are spatially distributed.

Furthermore, using a process called kriging (named after D. G. Krige, who devised the method), the model can then be used to interpolate between sampling points on the grid (Fig. 1a) to provide a better prediction of distribution (cf. Fig. 1b vs Fig. 1e-g). Finally, the model can be incorporated into ecosystem models to extrapolate patchiness to larger scales (e.g. rather than initiating dynamic models with even or randomly distributed plankton, species-specific patterns of patchiness could be imposed). Thus, geostatistical analysis provides considerable information about a system.

The above description is simple, and the reader is directed to Isaaks & Srivastava (1989), Cressie (1993), Goovaerts (1997), Armstrong (1998), Chilès & Delfiner (1999) and to the reviews of Legendre & Fortin (1989), Rossi et al. (1992), and Burrough (1995). Below (see 'Results and discussion') we illustrate the usefulness of this method, as applied to the distribution of 4 abundant ciliate species in a coastal lagoon. We use these data to assess the behaviour of these ciliates and to indicate some of the subtle inferences that can be made from geostatistical analysis concerning the biology of ciliates in general.

Sampling design, collection, and enumeration. A 40×40 m grid, divided at 10 m intervals, was designed following the requirements of the geostatistical analysis (Yfantis et al. 1987, Burrough 1995). Ciliate abun-

dance was determined on 2 occasions (25 October 2000 and 23 January 2001) at a sampling point located in the node of each cell. Five additional (October), and 10 additional (January) sampling points were added at a distance of 1 m from some of the above-mentioned points, to provide information for a better estimation of the small-scale variability (Fig. 2).

Sampling took place near the centre of Chautengo lagoon, Mexico ($16^\circ 37.4' N$, $99^\circ 06.6' W$), at the 30 to 35 points of the grid described above (Fig. 2). On each date, all samples were taken within 2 h, before midday when thermally induced winds occur. Water column temperature and salinity were recorded with a thermometer and an American Optical refractometer. Transparency was measured with a Secchi disk. Sampling was conducted at 0.4 m by deploying 400 ml Niskin-type bottles, built for use in shallow waters. Individual 400 ml water samples were homogenised, and for each, 100 ml subsamples were preserved with acid Lugol's iodine (2% final concentration) (Thronsdon 1978).

Ciliates were enumerated by settling and examining 5 ml, following the Utermöhl method (Hasle 1978). Samples were viewed using an inverted microscope ($\times 200$ and $\times 400$), equipped with phase-contrast optics. Although all ciliate taxa were quantified, in this study only 4 case-species were examined: a tintinnid and a scuticociliate from the October 2000 survey, and a strobilidiid and a strombidiid from the January 2001 survey.

Data analysis. Simple contour maps of ciliate abundance were made for exploratory analyses (see below) using the triangulation method of Surfer (Golden Software). Spatial analysis was conducted using the variogram as the basic tool for the characterisation of spatial structure and mapping of ciliate abundance (see above, Rossi et al. 1992, Goovaerts 1997). Model variograms were fit to the empirical variogram data for each ciliate; for each case (each ciliate), 3 different

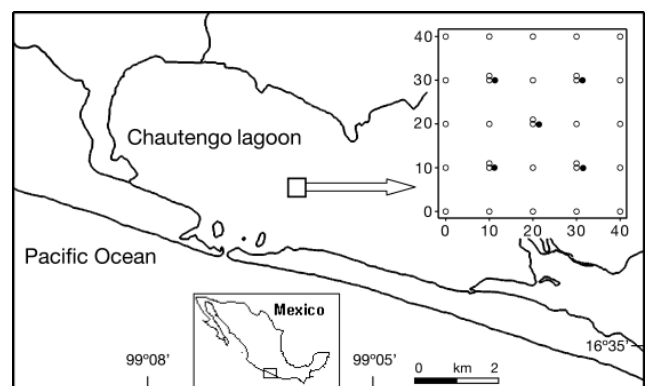


Fig. 2. Chautengo lagoon, Mexico, indicating the location of the 40×40 m sampling grid of 30 points (empty circles, October 2000), and 35 points (empty and solid circles, January 2001)

models were fit to the variogram according to an approximate weighted least squares (WLS) procedure, as described in Cressie (1993) (the best fit models are indicated in Fig. 4). Mapping was based on ordinary kriging, an interpolation technique that predicts values at unsampled points, using the model variogram (Isaaks & Srivastava 1989). To provide an indication of the precision of estimates, the coefficient of variation associated with each estimate was calculated as the ratio of the standard deviation to the mean, both of which are determined by kriging (Isaaks & Srivastava 1989). Some further details of the methodology associated with specific conditions are presented below.

RESULTS AND DISCUSSION

Environmental and biological variables

In October, the 1 m water column (i.e. from the surface to the sediment) was stratified at ~0.8 m: salinity was 16 psu at the surface and 23 psu near the bottom, while temperature was uniform, at 30°C throughout the column. The Secchi depth in October was 0.6 m, therefore, the entire water column was well illuminated, assuming the euphotic zone is 2 to 2.7 times the Secchi depth (Parsons et al. 1977). In January, the 0.8 m water column was mixed: salinity was 30 psu and temperature was 29°C; the Secchi depth was 0.3 m. Thus, for each individual month (October stratified, January mixed water), samples taken at 0.4 m across the grid were from waters with similar physical conditions.

Taxa for geostatistical analysis

In October, the ciliate community was dominated by a tintinnid (*Tintinnopsis* sp.) and a scuticociliate (*Pleuronema* sp.), which constituted up to ~70 and 10% of the total ciliate numbers, respectively. *Tintinnopsis* sp. abundance varied between 21 and 37 cells ml⁻¹, whereas *Pleuronema* sp. varied between 2 and 7 cells ml⁻¹. In January, *Lohmaniella oviformis* (5 to 23 cells ml⁻¹) and a medium sized (30 to 50 µm) *Strombidium* sp. (2 to 14 cells ml⁻¹) were the dominant ciliates, composing up to ~50 and 30% of the numbers, respectively. The following sections are a step-by-step presentation of the analysis conducted to uncover the underlying spatial structure of the abundance of these 4 ciliates.

Exploratory data analysis

Before starting the variographic analysis, an exploratory data analysis was performed to assess the

main features of the data and examine for outliers that may influence the results of the analysis (Tukey 1977, Cressie 1993). We used several of the exploratory techniques but here we present only the contour maps of the data.

These maps indicated the overall trends in the abundance data (Fig. 3). Local maxima of *Tintinnopsis* sp. were apparent: 2 distinct peaks of abundance occurred over the background surface (Fig. 3a). The contour map for *Lohmaniella oviformis* (Fig. 3b) indicated several areas of high abundance. The spatial distribution for *Strombidium* sp. (Fig. 3c) suggested 3 peaks of abundance. The contour map for *Pleuronema* sp. indicated higher values in the upper region and a small peak near the centre of the grid (Fig. 3d). The contour plots thus suggest that patchiness existed for all 4 species. However, as stated above, contour maps have limitations.

Structural analysis

Following the procedure outlined above (see 'Materials and methods' and Fig. 1), the next step was to calculate variograms for the 4 species (Fig. 4). However a preliminary requirement, not outlined above, was to determine if the patch structure has a directional component. For example, Langmuir currents may induce elongated planktonic patches (Parsons et al. 1977), in which case, the spatial variability will be higher when perpendicular to the wind direction and smaller when parallel to it.

A phenomenon with spatial variability, the same in any direction, is known as isotropy, whereas directionality results in anisotropy. Empirical variograms (i.e. variograms presenting lag-data only; e.g. points on Fig. 4 a–d) can be used to test for anisotropy (for details see Rossi et al. 1992). To assess for anisotropic distribution, empirical variograms for all 4 ciliates were computed for vertical and horizontal directions—given our experimental design these were the only directions that could be assessed. There was no marked anisotropic effect, so the spatial distribution was assumed to be isotropic. Furthermore, stable conditions of weather and water column led us to adopt an omnidirectional variogram. It was thus appropriate to use empirical omnidirectional variograms in the analysis. To perform variographic analysis we used 60% (~30 m) of the maximum lag as the active lag distance, as variograms decompose at intervals close to the maximum (Isaaks & Srivastava 1989). All point of the variograms, except the first one, included >30 data pairs (see Fig. 4).

Variograms are sensitive to extreme data values (potential outliers), as differences are squared (Eq. 1). To

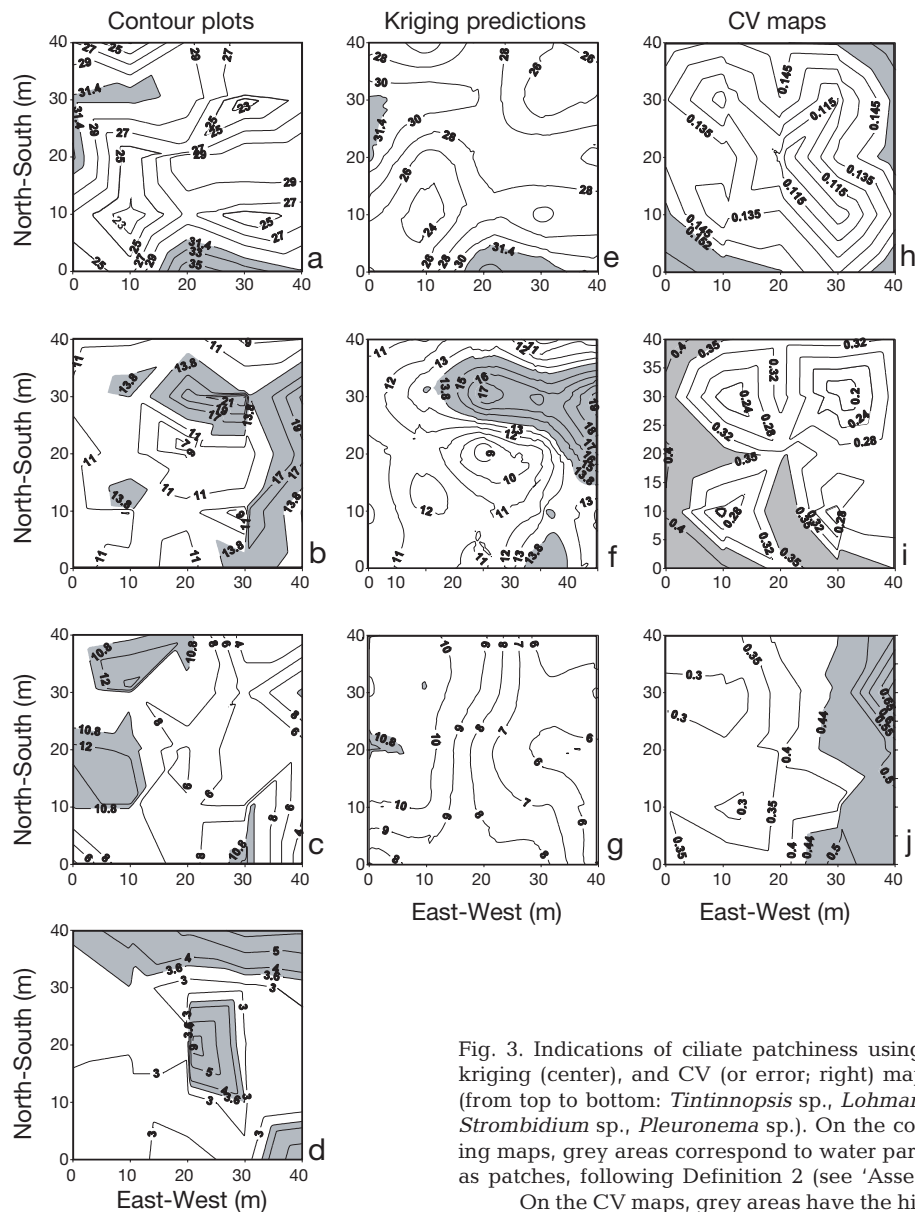


Fig. 3. Indications of ciliate patchiness using contour (left), kriging (center), and CV (or error; right) maps for 4 ciliates (from top to bottom: *Tintinnopsis* sp., *Lohmaniella oviformis*, *Strombidium* sp., *Pleuronema* sp.). On the contour and kriging maps, grey areas correspond to water parcels considered as patches, following Definition 2 (see 'Assessing patches'). On the CV maps, grey areas have the highest CV

assess for the influence of such values, a further test was conducted; a 'robust' variogram was computed which can reduce the effect of outliers, without removing data (Cressie & Hawkins 1980, Maravelias et al. 1996). The comparison between the 'classical' and the 'robust' variogram indicated no remarkable differences in terms of the underlying spatial structure. Thus, the presence of high values did not affect the detection of the patches, and these high values may actually represent the core of patches. Consequently, in subsequent analyses, classical variograms were used.

The next step for data analysis was to fit models to the empirical variogram data (i.e. lines on Fig. 4). Using such models we can characterise the spatial

structure of ciliate abundance and estimate abundance at unsampled points. The fitting of models was conducted following the approximate WLS procedure (Cressie 1993). This method gives more weight to the lags with more observations, and with lower empirical variogram values, allowing for a better fit of the variogram near the origin, which can be the most important part of the model. Note: in practice most geostatistical packages determine average lags from the data, and thus variograms generally do not present lags identical to those measured (cf. Fig. 2 with Fig. 4).

In fitting the omnidirectional variograms, the 3 models illustrated in Fig. 1 were considered, but only 2 adequately fit the data (based on the least weighted

squares, Cressie 1993): a 'spherical' model with a nugget effect and an 'exponential' model, also with a nugget effect. These models are given by:

$$\hat{\gamma}(|h|) = \begin{cases} 0 & \text{for } |h| = 0 \\ c_0 + c \left[\frac{3}{2} \left(\frac{|h|}{a_0} \right) - \frac{1}{2} \left(\frac{|h|}{a_0} \right)^3 \right] & \text{for } 0 < |h| \leq a_0 \\ c_0 + c & \text{for } |h| \geq a_0 \end{cases} \quad (2)$$

for the spherical model, and:

$$\hat{\gamma}(|h|) = \begin{cases} 0 & \text{for } |h| = 0 \\ c_0 + c \left[1 - \exp\left(-\frac{3|h|^2}{a_0^2}\right) \right] & \text{for } |h| \neq 0 \end{cases} \quad (3)$$

for the exponential model, where c_0 is the nugget, $c_0 + c$ is the sill and a_0 is the range of the variogram (Cressie 1993). These models were fit to the empirical variograms using the interactive feature of S-Plus 2000 (MathSoft).

The variograms for *Tintinnopsis* sp. and *Lohmaniella oviformis* both illustrated typical cases, similar to that shown in Fig. 1d, but there were subtle differences between them. In contrast, the variogram for *Strombidium* sp. failed to reach a sill over the sampled distance, while that of *Pleuronema* sp. appeared to have already reached the sill. These 4 cases are discussed in detail below.

Tintinnopsis sp. (Fig. 4a)

Tintinnopsis sp. was the dominant ciliate in October, when the shallow water column was stratified. A spherical function best fit the data of the empirical variogram for *Tintinnopsis* sp., suggesting clearly defined structured ciliate patches within the sampling grid (see 'General overview of geostatistics' above). The range (Fig. 1d) for this species is ~18 m (Table 1); this is the distance below which discrete patches might be observed, e.g. patches could be 1, 10 or 15 m, but unlikely 20 m in size. Patches on the order of 1 to 100 m have been suggested for other tintinnids (e.g. Stoecker et al. 1984), but size limitations of such patches have generally not been proposed. Using geostatistical techniques, we have now provided observational limits of these patches.

The spatially structured component ($c/(c_0 + c) \times 100$, Fig. 1d) was 74 % of the total spatial variation (Table 1). This ratio can be interpreted as the portion of the total variance explained by a spatial model. The nugget (Fig. 1d) accounts for the microscale (<1 m) variations and the measurement error (the difference between repeated ciliate counts made with the same water sample), and it represented the remaining 26 % of the vari-

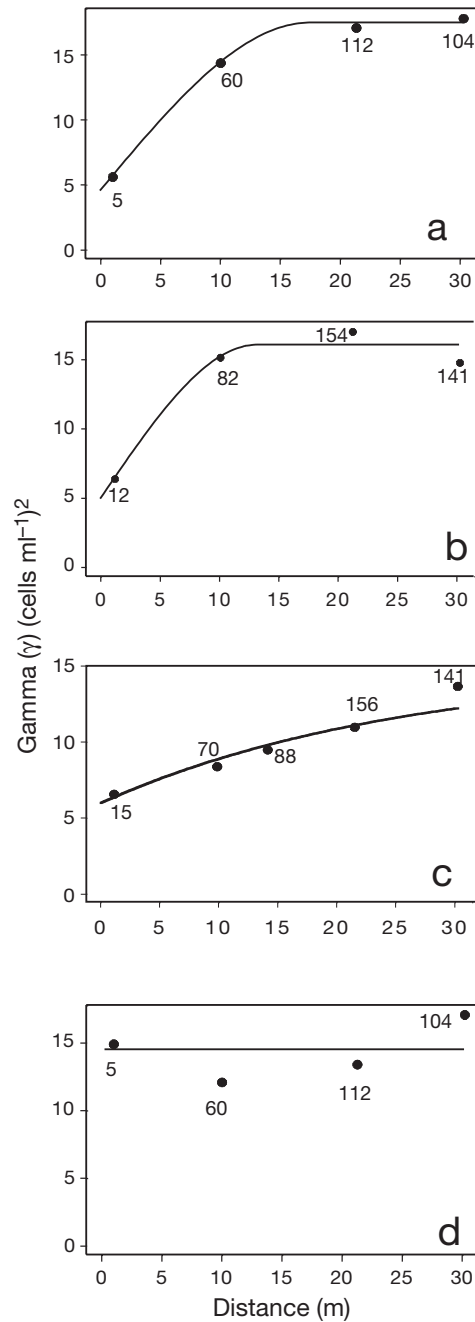


Fig. 4. Empirical variograms (points) and fitted models (lines). Only 60 % of the total distance (~30 m), measured in any direction over the sampling space (see Fig. 1a), was represented in the variograms (see 'General overview of geostatistics'). Spherical models were fit to the data for (a) *Tintinnopsis* sp. and (b) *Lohmaniella oviformis*. An exponential model was fit to data for (c) *Strombidium* sp.. The spatial distribution of (d) *Pleuronema* sp. was random at the measured scale; therefore a spatial model cannot explain this variability, and a random model explains it. See Table 1 for the parameters associated with these curves. The number of pairs of points for each lag distance is indicated. Note: in practice most geostatistical packages determine average lags from the data, and thus variograms generally do not present lags identical to those measured (cf. Fig. 2)

Table 1. Abundance statistics (\bar{x} : mean; SD: standard deviation, measured as cells ml⁻¹) and parameters of the different variogram models for ciliate abundance, determined through cross validation (see 'Cross-validation and kriging' for details). The best fit, the lowest according to Cressie (1993), for each ciliate is presented graphically in Fig. 4. See Fig. 1 for a definition of symbols (e.g. c_0 , nugget effect; $c_0 + c$, sill; a_0 , range). Cross-validation results of models and parameters selected to map the predictions of ciliate abundance by ordinary kriging (see 'Cross-validation and kriging' for details): Bias, the mean prediction error (estimated – observed); SD²: the variance of the error; MSE (mean squared error) = SD² + bias². See 'Results and discussion' for discussion of the models. *Lowest MSE

Ciliate	Abundance		Model	Parameters			Spatially structured component (%) $c/(c_0 + c) \times 100$	Model statistics		
	\bar{x}	SD		c_0 (γ , cells ml ⁻¹) ²	$c_0 + c$	a_0 (m)		Bias	SD ²	MSE
<i>Tintinnopsis</i> sp.	27.8	4.0	Spherical	4.5	17.5	17.5	74	-0.22	12.25	12.29*
			Exponential	6.6	16.8	24.5	61	-0.20	13.10	13.14
<i>Lohmaniella oviformis</i>	12.4	4	Spherical	5	16	13	69	0.13	18.14	18.16*
			Exponential	2.5	16.5	10	87	0.12	18.31	18.33
<i>Strombidium</i> sp.	8	3.4	Exponential	6	15	25.6	60	0.086	11.08	11.09*
			Spherical	6	18	70	66	0.081	11.15	11.16
<i>Pleuronema</i> sp.	3.4	1.3	Nugget	14.5	14.5	<1	0			

ability. The nugget effect thus reflects, in plankton distribution, the existence of ciliate aggregations that cannot be further resolved by our sampling. Buskey & Stoecker (1988, 1989) have indicated that swimming of the tintinnid *Favella* sp. would allow the ciliate to remain in small-scale patches, and Kils (1993) has demonstrated the existence of dense patches on the order of centimetres. Such patches may produce the microscale variance of the nugget that we estimated for *Tintinnopsis* sp. The nugget may also represent measurement error (e.g. variation between subsamples); we are presently examining this error-term in a separate study (Bulit et al. unpubl.). The nugget also reflects how tight the patches are (Dalthorp et al. 2000); a small nugget indicates a tight patch. Therefore, on the basis of the properties derived from the nugget and range, patches of *Tintinnopsis* sp. are of medium size relative to the sampled grid and tighter than those of *Lohmaniella oviformis* and *Strombidium* sp. patches (see below).

Lohmaniella oviformis (Fig. 4b)

This small strobilidiid, close in size to the nanoplankton, was the dominant ciliate in January, when the water column was mixed. Like *Tintinnopsis* sp., a spherical function best fit the data of the empirical variogram for *Lohmaniella oviformis*, although the exponential function provided almost as good a fit (Table 1). The range for this species was 13 m, suggesting that patches of the strobilidiid were slightly smaller than those of *Tintinnopsis* sp. The sill of the *L. oviformis* model was also lower than that of *Tintinnopsis* sp. (Table 1), indicating less variability in its distribution. These small patches could result from repeated cell divisions, as small cili-

ates can have generation times on the order of hours in warm waters (Montagnes 1996), but patches might also be due to swimming behaviour (see below).

For *Lohmaniella oviformis*, the spatially structured component was 69% of the total spatial variation, with the nugget representing 31% (Table 1). These results suggest a similar structure in patches to that of *Tintinnopsis* sp.; possibly, the slightly larger nugget of *L. oviformis* was due to its lower density (Fig. 3), and thus potentially a larger measurement error. To date, there are few studies on *L. oviformis*. Jonsson (1989) examined the vertical distribution of *L. oviformis*, indicating a swimming velocity of 340 $\mu\text{m s}^{-1}$. In contrast, Jonsson (1989) found *T. campanula* to have a vertical swimming velocity of 260 $\mu\text{m s}^{-1}$. Furthermore, small, naked strobilidiids tend to have a higher tumbling-reorientation rate than larger tintinnids (D.J.S.M. pers. obs.). Thus, although it is small, the ability of *L. oviformis* to accumulate in patches by rapid swimming and reorientation might be considerable, and we speculate that this may explain the patches observed in the lagoon.

Strombidium sp. (Fig. 4c)

Strombidium sp. also dominated in January although its mean abundance (8.0 ± 3.4 cells ml⁻¹) was lower than that of *Lohmaniella oviformis*. The variogram for this strobilidiid shows that the variance continued to increase without reaching a sill. This suggests that patches tend to be larger than the 40 m sampling grid. Some geostatistical packages provide an option to fit a linear variogram, but this does not allow prediction at unsampled points (e.g. Fig. 1d). Thus, one option is to fit an exponential model to the data, approximating the linear shape with a range beyond the survey's maxi-

mum allowable distance (i.e. 30 m for this study); we followed this second option (see Goovaerts 1997). As the exponential model approaches its sill asymptotically, the practical range is calculated as $3 \times a_0$, as the distance at which the model reaches 95% of its sill (Goovaerts 1997, Armstrong 1998). In the case of *Strombidium* sp., a_0 was 25.6 m; therefore the practical range was ~77 m. Thus, if the grid were larger, discrete patches of this ciliate are likely to be have been observed. The spatially structured component was 60% of the total spatial variation (Table 1).

It is unclear why this ciliate would form large diffuse patches. However, there are many mixotrophic species of *Strombidium* (Stoecker et al. 1989, Bernard & Rasoulzadegan 1994, Dolan & Pérez 2000); mixotrophs might form patches larger than exclusively heterotrophic ciliates, as autotrophy would allow these ciliates to occupy areas where there are no prey, assuming that prey are also distributed in patches. Dale & Dahl (1987) suggested that swimming behaviour of strombidiid ciliates might cause concentration of organisms. Jonsson (1989) indicated that some species of *Strombidium* have swimming velocities of 450 to 680 $\mu\text{m s}^{-1}$. Thus, under very calm weather conditions and water stability like those we found, this speed can be of significance, and patches of *Strombidium* might be large. However, the Strombidiidae comprises a diverse group, and it may be imprudent to generalise based on the few well-studied taxa.

***Pleuronema* sp. (Fig. 4d)**

This species occurred in October at a lower abundance (3.4 ± 1.3 cells ml^{-1}) than *Tintinnopsis* sp. and it differs in distribution from those species discussed above. It appears from the variogram that a sill was reached below the minimum sampling distance (1 m). Thus, a pure nugget model was fit to the data (Table 1), which indicates a lack of spatial resolution; i.e. patches were not detected. Ideally, we would return to the field and sample for this ciliate at a smaller scale. Note that this analysis contradicts the observations predicted by a simple contour map (i.e. Fig. 3d). Thus, the application of geostatistical analysis has allowed a different, and likely more accurate, evaluation of the distribution of this ciliate.

We might predict that the bacterivorous nature of this scuticociliate (Fenchel 1987, Dolan & Coats 1991, Ederington et al. 1995) means that it forms patches around detrital clumps. Furthermore scuticociliates appear to be the most abundant ciliates in aggregates of suspended matter (Artolozaga et al. 2000), and they have been associated with detritus (Silver et al. 1984, Sherr et al. 1986). Such clumps would undoubtedly be

microns to millimetres in size, and thus the patches would be $\ll 1$ m. The samples of *Pleuronema* sp. that have high values (13% of the total samples) may, therefore, be chance collection of detrital material with associated ciliates.

We have shown, using variographic analysis, that for 3 of the 4 evaluated ciliates, the abundance is neither randomly nor homogeneously distributed, but rather exhibits a structured small-scale patchy distribution. We have also indicated that simply applying standard contour fits to the data, using existing packages (see 'General overview of geostatistics' above) may provide erroneous conclusions regarding patchiness (e.g. *Pleuronema* sp. did not form patches when geostatistically analysed). Furthermore, we have indicated that spherical and exponential functions, and not the Gaussian function, best model ciliate patches; this suggests tight, rather than diffuse, patches (see Fig. 1). These are potentially useful observations of patches, and we thus support the application of these variographic techniques to analyze microzooplankton distributions.

Cross-validation and kriging

In the next step of the analysis, we indicate how prediction maps are generated and associated with the coefficients of variation. In this phase, the variogram models are used to predict abundance values at unsampled localities, using a linear prediction method called ordinary kriging (Isaaks & Srivastava 1989, Goovaerts 1997, Armstrong 1998).

In practice, the prediction is computed using only those observations inside a selected 'search radius'. The performance of the best model, search radius, and number of points to be used in the kriging system were determined through a cross validation procedure. This procedure consists of incrementally removing observations (abundance values) from the data set, one by one, and re-estimating the value of the removed observation using the parameters being validated (Goovaerts 1997). Thus, a prediction error (predicted abundance – measured abundance) was generated at each sampling location.

The goal of this validation is to select the model and the search-strategy parameters that give the minimum mean squared error (MSE), by making a trade-off between bias and variance (see Table 1). The MSE is the sum of the squared bias (the mean of the error) and the variance (the spread of the error) of the residuals (Isaaks & Srivastava 1989). The X-valid procedure (Geo-Eas 1.2.1; Englund & Sparks 1991) was used to generate the residual vectors. Table 1 summarises the results of crossed validated models and parameters from the 2 best models assessed (out of 8 to 10 possible

models). The models with the lowest MSE (values with asterisks, Table 1) were then used to map the distributions of *Tintinnopsis* sp., *Lohmaniella oviformis*, and *Strombidium* sp. (Fig. 3e–g). These distributions can then be used to assess patchiness.

Assessing patches

Although many studies describe patches, a 'patch' is rarely quantifiably defined; we offer 2 definitions and then assess if either is appropriate. Following our initial argument that patches are rare, relative to a background of lower abundance, a patch might then be defined as a region where the ciliate abundance exceeds a cut-off value of the kriging predictions map, such as (1) the median value or (2) the upper quartile value. Below, we use our data to assess these 2 definitions.

The search strategy with the best performance for predicting tintinnid abundance used a circular search neighbourhood of 12 m and the nearest 10 observations (Table 1). The kriging-abundance predictions for tintinnids (Fig. 3e) were mapped using the spherical model and these parameters. The resulting distribution differed from that obtained from the contour plot (Fig. 3a). The modelled distribution produced one extended patch (70% of the total area), following Definition 1. In contrast, when Definition 2 was applied, 2 small patches existed and represented the 4.4% of the total area (Fig. 3e). Following the initial premise that patches are rare, it appears that Definition 2 is more appropriate for the tintinnid. Similarly, the prediction map for *Lohmaniella oviformis* was made using a spherical model, a circular search neighbourhood of 12 m in radius, and considered the 10 nearest points (Table 1). This modelled distribution (Fig. 3f) differed from the contour map (Fig. 3b), where several areas were presumptively considered as patches, and allowed quantification of the distribution: by Definition 1, patches occupied 52% of the surface, while by Definition 2, patches covered 23% of the total area. Again, assuming patches are rare, the second definition of a patch seems more appropriate. For *Strombidium* sp., an exponential model was used, a circular search neighbourhood of 20 m in radius, and the nearest 20 points were considered to make the kriging map (Table 1). The modelled distribution (Fig. 3g) was distinctly different from that of the contour map (Fig. 3c). Following Definition 1, the modelled distribution indicated that a portion of a patch occupied 50% of the area, but by Definition 2, there was a tiny part of a patch, located at coordinates 0, 20, covering only 1% of the area, yet again supporting Definition 2. In contrast to the 3 taxa examined above, fitting of a pure nugget model to the empirical variogram for *Pleuronema* sp. precluded the use of kriging (Goovaerts 1997). Therefore, at this

scale, *Pleuronema* sp. is considered to be randomly distributed, with no distinct patches.

We have thus assessed that in the lagoonal environment, structured ciliate patches exist at a small-scale in stratified as well as in mixed waters. Following the premise that patches are rare (covering <<50% of the sampled area), and using quantitative criteria provided by geostatistics, we have put some limits on the definition of these ciliate patches; they are nearer to being associated with the abundance above the density cut-off of the upper quartile range from kriging maps, rather than above the median value. We have also indicated a difference in predicted distribution when kriging is applied to the data, relative to using traditional interpolation methods. This is because the contour levels obtained by kriging are dictated by the variogram model and are not simply based on isolated points measured over the grid. Furthermore, the models illustrated in Fig. 4 can be used to predict patchiness on a larger scale and thus could be used for ecosystem models. However, like any estimates there is error associated with these predictions; the next section examines this error.

Estimating error of the modelled distribution

The maps for the coefficient of variation (CV) indicate the precision of the estimated distribution (Fig. 3h–j). This error term can then be used to assess the predictable nature of the kriging map, taking into account the variability of the predictions. The CV for the predicted distribution of *Tintinnopsis* sp. varied between 9 and 17% (Fig. 3h); the CV for *Lohmaniella oviformis* ranged from 16 to 43% (Fig. 3i); the CV for *Strombidium* sp., ranged from 25 to 74% (Fig. 3j). As the CV was <100% in all cases, we can conclude that outlying, or erratic, values that affect the estimated distribution are rare or non-existent (Isaaks & Srivastava 1989). By examining the kriging and the CV maps for all the ciliates, it can be seen that at higher ciliate abundance, the CV forms a homogeneous spatial pattern, increasing towards the edges, where fewer points were sampled; this is especially clear for the case of *Tintinnopsis* sp. (Fig. 3h). However, the CV may also vary inversely with the predicted abundance, as depicted for *L. oviformis* (Fig. 3i) and for *Strombidium* sp. (Fig. 3j). This suggests that the uncertainty on the predicted values increases at lower abundance and could be reduced by adding more sampling points in those areas. These estimates of error illustrate the precision of modelling distributions and may thus ultimately be used to help assess the predictability of larger scale food-web models.

To our knowledge, this study constitutes the first application of geostatistical techniques to model the small-scale spatial structure of microplanktonic populations. Our results indicate different species-specific patterns of patchiness at a small-scale and in different hydrodynamic conditions, supporting the notion of behavioural niche-separation of planktonic ciliates. The size, shape, and distinctness of the patches were characterised by the spatial dependence, which we indicate can be summarised by 3 parameters: the range, sill, and nugget. And finally, a working definition of patch was proposed and was used to characterise the generated kriging predictions. Thus, geostatistical analysis appears to be a powerful tool to examine microzooplankton at small-scales, and we support its further application in the field.

Acknowledgements. This paper forms part of the PhD thesis of C.B. at the University of Liverpool and Universidad Autónoma Metropolitana-Xochimilco. We thank Dr. Andrés Boltovskoy, Eusebio Cueva, and Héctor Chagoyan for their assistance during fieldwork. We also thank PROMEP-SEP for support in purchasing a microscope. Rubén Roa and Fabián Tapia kindly provided the templates for cross validation procedure. Finally, we thank 2 anonymous reviewers for their constructive comments.

LITERATURE CITED

- Armstrong M (1998) Basic linear geostatistics. Springer-Verlag, Berlin
- Artolozaga I, Ayo B, Latatu A, Azúa I, Unanue M, Iriberry J (2000) Spatial distribution of protists in the presence of macroaggregates in a marine system. *FEMS Microbiol Ecol* 33:191–196
- Bernard C, Rassoulzadegan F (1994) Seasonal variations of mixotrophic ciliates in the northwest Mediterranean Sea. *Mar Ecol Prog Ser* 108(3):295–301
- Burrough P (1995) Spatial aspects of ecological data. In: Jongman R, Ter Braak C, Van Tongeren O (eds) *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge, p 213–251
- Buskey EJ, Stoecker DK (1988) Locomotory patterns of the planktonic ciliate *Favella* sp.: adaptations for remaining within food patches. *Bull Mar Sci* 43:783–796
- Buskey EJ, Stoecker DK (1989) Behavioral responses of the marine tintinnid *Favella* sp. to phytoplankton: influence of chemical, mechanical and photic stimuli. *J Exp Mar Biol Ecol* 132:1–16
- Chilès JP, Delfiner P (1999) *Geostatistics: modeling spatial uncertainty*. Wiley, New York
- Cressie NAC (1993) *Statistics for spatial data*. Wiley, New York
- Cressie NAC, Hawkins DM (1980) Robust estimation of the variogram. *I. Math Geol* 12:115–125
- Cushing DH (1953) Studies on plankton populations. *J Cons Perm Int Explor Mer* 19:1–22
- Dale T, Dahl E (1987) Mass occurrence of planktonic oligotrichous ciliates in a bay in southern Norway. *J Plankton Res* 9:871–879
- Dalthorp D, Nyrop J, Villani MG (2000) Foundations of spatial ecology: the reification of patches through quantitative description of patterns and pattern repetition. *Entomol Exp Appl* 96:119–127
- Defeo O, Rueda M (2002) Spatial structure, sampling design and abundance estimates in sandy beach macroinfauna: some warnings and new perspectives. *Mar Biol* 140: 1215–1225
- Dolan JR, Coats WD (1991) A study of feeding in predacious ciliates using prey ciliates labeled with fluorescent microspheres. *J Plankton Res* 13:609–627
- Dolan JR, Pérez MT (2000) Costs, benefits and characteristics of mixotrophy in marine oligotrichs. *Freshw Biol* 45(2): 227–238
- Ederington MC, McManus GB, Harvey HR (1995) Trophic transfer of fatty acids, sterols, and a triterpenoid alcohol between bacteria, a ciliate, and the copepod *Acartia tonsa*. *Limnol Oceanogr* 40:860–867
- Englund E, Sparks A (1991) *GEO-EAS 1.2.1 User's guide*. Environmental Protection Agency, Las Vegas, NV
- Fenchel T (1987) *Ecology of Protozoa: the biology of free-living phagotrophic protists*. Brock/Springer Series in Contemporary Bioscience. Science Tech Publishers, Madison, WI
- Freire J, González-Gurriarán E, Olaso I (1992) Spatial distribution of *Munida intermedia* and *M. sarsi* (Crustacea, Anomura) on the Galician continental-shelf (NW Spain): application of geostatistical analysis. *Estuar Coast Shelf Sci* 35: 637–648
- González-Gurriarán E, Freire J, Fernández L (1993) Geostatistical analysis of spatial-distribution of *Liocarcinus depurator*, *Macropipus tuberculatus* and *Polybius henslowii* (Crustacea, Brachyura) over the Galician continental-shelf (NW Spain). *Mar Biol* 115:453–461
- Goovaerts P (1997) *Geostatistics for natural resources evaluation*. Oxford University Press, New York
- Hasle G (1978) The inverted microscope method. In: Sournia A (ed) *Phytoplankton manual*. UNESCO, Paris, p 88–96
- Isaaks EH, Srivastava RM (1989) *An introduction to applied geostatistics*. Oxford University Press, New York
- Jonsson PR (1989) Vertical distribution of planktonic ciliates—an experimental analysis of swimming behaviour. *Mar Ecol Prog Ser* 52:39–53
- Kils U (1993) Formation of micropatches by zooplankton-driven microturbulences. *B Mar Sci* 53(1):160–169
- Legendre P, Fortin MJ (1989) Spatial pattern and ecological analysis. *Vegetatio* 80:107–138
- Legendre P, Legendre L (1998) *Numerical ecology*. Elsevier Science, Amsterdam
- Legendre P, Troussellier M (1988) Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. *Limnol Oceanogr* 33:1055–1067
- Liebhöf A, Rossi R, Kemp W (1993) Geostatistics and geographic information systems in applied insect ecology. *Annu Rev Entomol* 38:303–327
- Mackas DL (1984) Spatial autocorrelation of plankton community composition in a continental shelf ecosystem. *Limnol Oceanogr* 29:451–471
- Maravelias CD, Reid DG, Simmonds EJ, Haralabous J (1996) Spatial analysis and mapping of acoustic survey data in the presence of high local variability: geostatistical application to North Sea herring (*Clupea harengus*). *Can J Fish Aquat Sci* 53:1497–1505
- Montagnes DJS (1996) Growth responses of planktonic ciliates in the genera *Strobilidium* and *Strombidium*. *Mar Ecol Prog Ser* 130:241–254
- Montagnes DJS, Poulton AJ, Shammon TM (1999) Mesoscale,

- finescale and microscale distribution of micro- and nanoplankton in the Irish Sea, with emphasis on ciliates and their prey. *Mar Biol* 134:167–179
- Parsons T, Takahashi M, Hargrave B (1977) Biological oceanographic processes. Pergamon Press, Oxford
- Passy SI (2001) Spatial paradigms of lotic diatom distribution: a landscape ecology perspective. *J Phycol* 37:370–378
- Pelletier D, Parma AM (1994) Spatial distribution of Pacific halibut (*Hippoglossus stenolepis*): an application of geostatistics to longline survey data. *Can J Fish Aquat Sci* 51:1506–1518
- Pinca S, Huntley ME (2000) Spatial organization of particle size composition in an eddy-jet system off California. *Deep-Sea Res* 47:973–996
- Pinel-Alloul B, Guay C, Angeli N, Legendre P, Dutilleul P, Balvay G, Gerdeaux D, Guillard J (1999) Large-scale spatial heterogeneity of macrozooplankton in Lake of Geneva. *Can J Fish Aquat Sci* 56:1437–1451
- Roa R, Tapia F (2000) Cohorts in space: geostatistical mapping of the age structure of the squat lobster *Pleuroncodes monodon* population off central Chile. *Mar Ecol Prog Ser* 196:239–251
- Rossi RE, Mulla DJ, Journel AG, Franz EH (1992) Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecol Monogr* 62:277–314
- Rueda M (2001) Spatial distribution of fish species in a tropical estuarine lagoon: a geostatistical appraisal. *Mar Ecol Prog Ser* 222:217–226
- Sherr EB, Sherr BF, Fallon RD, Newell SY (1986) Small, aloricate ciliates as a major component of the marine heterotrophic nanoplankton. *Limnol Oceanogr* 31:177–183
- Silver MW, Gowing MM, Brownlee DC, Corliss JO (1984) Ciliated protozoa associated with oceanic sinking detritus. *Nature* 309:246–248
- Steele JH (1978) Spatial pattern in plankton communities. Plenum Press, New York
- Stoecker DK, Davis LH, Anderson DM (1984) Fine scale spatial correlations between planktonic ciliates and dinoflagellates. *J Plankton Res* 6:829–842
- Stoecker DK, Taniguchi A, Michaels AE (1989) Abundance of autotrophic, mixotrophic and heterotrophic planktonic ciliates in shelf and slope waters. *Mar Ecol Prog Ser* 50:241–254
- Thronsen J (1978) Preservation and storage. In: Sournia A (ed) *Phytoplankton manual*. UNESCO, Paris, p 69–74
- Tukey JW (1977) *Exploratory data analysis*. Addison-Wesley, Reading, MA
- Yfantis E, Flatman GT, Behar JV (1987) Efficiency of kriging estimation for square, triangular and hexagonal grids. *Math Geol* 19:183–205

*Editorial responsibility: John Dolan,
Villefranche-sur-Mer, France*

*Submitted: March 22, 2002; Accepted: September 17, 2002
Proofs received from author(s): December 16, 2002*