

The application of geostatistics in mapping and assessment of demersal resources. *Nephrops norvegicus* (L.) in the northwestern Mediterranean: a case study*

FRANCESC MAYNOU

Institut de Ciències del Mar (CSIC), P. Joan de Borbó, s/n, 08039 Barcelona, Spain.

SUMMARY: The application of geostatistics in mapping and assessing demersal resources is reviewed. The basic summary of the methodology is illustrated with a case study on data for *Nephrops norvegicus* off Barcelona (NE Spain) over commercial fishing grounds. The geostatistical analysis revealed that *N. norvegicus* is distributed over the continental slope in high-density patches of ~ 7 km, alternating with areas of low density. Mapping by disjunctive kriging helped visualize the spatial distribution of the resource and identified areas where the resource abundance was of commercial level.

Key words: Norway lobster, *Nephrops norvegicus*, geostatistics, variogram, kriging, resource assessment.

RESUMEN: APLICACIÓN DE MÉTODOS GEOESTADÍSTICOS EN LA CARTOGRAFÍA Y EVALUACIÓN DE RECURSOS DEMERSALES. ESTUDIO DE *NEPHROPS NORVEGICUS* EN EL MEDITERRÁNEO NOROCCIDENTAL. – Se presenta una revisión sobre la aplicación de métodos geoestadísticos para el cartografiado y evaluación de recursos demersales. El resumen básico de la metodología se ilustra con el estudio del recurso de cigala (*Nephrops norvegicus*) en caladores comerciales de la flota de Barcelona (NE España). El análisis geoestadístico permitió poner en evidencia que la cigala se distribuye en manchas de alta densidad de ~ 7 km, alternando con áreas de baja densidad. La cartografía por kriging disyuntivo permite visualizar la distribución espacial del recurso e identificar áreas en las que la abundancia del recurso alcanza niveles comerciales.

Palabras clave: Cigala, *Nephrops norvegicus*, geoestadística, variograma, kriging, evaluación de recursos.

INTRODUCTION

The classical methods used to obtain abundance estimates in demersal fish populations include bottom trawl surveys using standard trawling gears and specific sampling strategies, e.g. they are methods based on sampling theory. The sampling points (trawl hauls) can be randomly positioned over the sampling area, yielding an unbiased estimate of abundance and

its corresponding precision index (Cochran, 1977). When the random sampling (or variations thereof) is carried at the appropriate spatial scale, it effectively obliterates any underlying spatial structure. However, the scale of spatial distribution of the target species is usually unknown and this factor may introduce bias in the computation of abundance estimates. In experimental trawl surveys, when the average distance between hauls is smaller than the zone of spatial influence of the underlying spatial phenomenon, the abundance of organisms at one point can be partially predicted from neighbouring points.

*Received 12 October 1997. Accepted 20 February 1998.

The existence of spatial structure in the distribution of organisms (or any other ecological variable) originates spatial autocorrelation between pairs of samples. Spatial autocorrelation can be defined as the departure from randomness in the spatial distribution of organisms or ecological variables, and it is analogous to a covariance coefficient. The concept of spatial autocorrelation can also be seen as the non-null probability of predicting the value of a sample from neighbouring samples. The spatial autocorrelation present in a given data set can be analyzed, modeled mathematically and incorporated into spatially-explicit models which can be used to produce unbiased abundance estimates of demersal resources.

Spatial statistical methods were developed independently by two major lines of research: Human geography (and ecology) and mining geology. In the 1950's and 1960's autocorrelation models were first produced in human geography (Moran, 1950; Geary, 1954; Berry and Marble, 1968) and later adopted in ecology by the influential work of Cliff and Ord (1973). During the 1970's and 1980's, these models were widely applied in ecology and in the 1990's they have become mainstream statistical methods, although their use in marine ecology is still limited (Jumars *et al.*, 1977; Jumars, 1978; Mackas, 1984; Legendre and Troussellier, 1988). Excellent reviews of this methodology in ecology can be found in Sokal and Oden (1978), Legendre and Fortin (1989) and Legendre (1993).

In mining geology, spatial statistical methods were first developed in the 1960's in the context of ore-reserve estimation. These methods were termed *geostatistics*. The theoretical basis of geostatistics was laid out by Matheron (1962, 1963, 1969 and 1971). Later influential reviews and developments include the work of Matheron (1973, 1976), Journel and Huijbregts (1978), Isaaks and Srivastava (1989), and more recently, Cressie (1991). Geostatistics expanded from mining geology to other domains of the earth sciences in the 1980's (e.g. Burrough, 1983), and particularly to fisheries science (Conan, 1985; Conan *et al.*, 1988). Since the late 1980's and in the 1990's, geostatistics has been widely used for the estimation of abundance of demersal resources. To the pioneering applications of G.Y. Conan cited earlier, we can add: Conan *et al.* (1992); Conan *et al.* (1994); Freire *et al.* (1992); González Gurriarán *et al.* (1993), Maynou *et al.* (1996); Maynou *et al.* (1998); Pelletier and Parma (1994); Petitgas and

Poulard (1989); Simard *et al.* (1992); Sullivan (1991). They have also been applied to pelagic fish and acoustic survey data (Anonymous, 1990, 1991; Petitgas, 1993) and ecology (Rossi *et al.*, 1992).

In the geostatistical literature, an autocorrelated variable is termed a *regionalized variable*. The example of regionalized variable which will be used in this article is the density of *N. norvegicus* over the fishing grounds off Barcelona (NE Spain). The covariance function describing the spatial autocorrelation is called *structure function*. An example of structure function used in this article is the variogram. Other examples of structure functions are the autocorrelation functions used in spatial ecology, such as Moran's correlogram or Geary's correlogram.

In this review, the use of geostatistical tools for assessing the spatial structure, mapping and estimation of demersal resources is highlighted. These methods are illustrated with the case study of *N. norvegicus* and the underlying spatial phenomena that originate the spatial structure in this species are discussed.

METHODS

Sampling design

Two surveys (GEOESC-I and GEOESC-II) were specifically designed for the mapping and assessment of the harvestable resource of Norway lobster on the fishing grounds off Barcelona. For the purposes of this review, only the results for GEOESC-I will be presented. The survey site was chosen over muddy bottoms with gentle slope and depth contours parallel to the coast, limited by two submarine canyons (Fig. 1). A regular grid 1 by 2 nautical miles was set parallel to the coast and a start location for each tow was randomly selected within each cell. The total area covered by the sampling design was 790 km², comprising 115 cells, although due to logistical and practical constraints only 59 locations could be sampled during the first cruise. The depth of the sampling locations varied from 141 to 730 m, encompassing the whole depth range of Norway lobster in the area of study.

The experimental sampling gear was a specially designed otter trawl drawn by a single warp ('Maireta System Trawl', Sardà *et al.*, 1994). The codend stretched mesh was 12 mm in order to

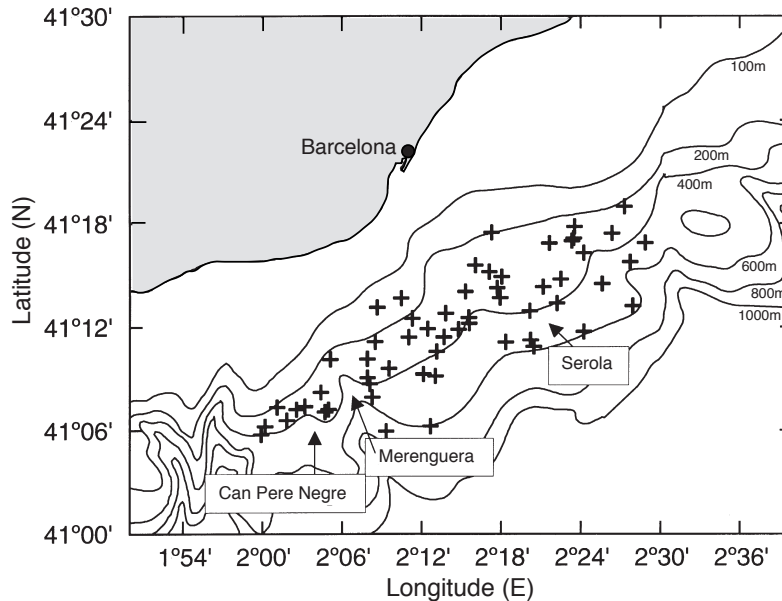


FIG. 1. – Sampling area of the experimental survey GEOESC-I. Crosses indicate mid position of trawl hauls. Main commercial fishing grounds for Norway lobster are shown.

retain small individuals not normally available to the commercial fishing gear. The actual opening of the trawl was measured using an acoustic system (SCANMAR) and stabilized at 14.0 m width by 2.0 m height. During the survey, tows were made parallel to the depth contours. The duration of each tow was set to exactly 15 min (time of effective trawling). The towing speed varied between 2.3 and 2.6 knots (mean 2.5 knots). Start and end locations for each tow were taken by GPS. The actual surface covered by each tow was computed from the GPS and the SCANMAR readings. Effective time of the survey was from 7:30 to 20:00 h each day. The survey was completed between 27 April and 5 May 1991.

The total catch of *N. norvegicus* was counted, weighed and measured. The density of Norway lobster was computed (number of individuals · km⁻²) from the total surface covered by each tow.

Statistical methods

The fundamentals of geostatistics, with emphasis on the methods employed here, are laid out in the Appendix in a rather mathematical way. What follows is an elementary summary of linear geostatistics and disjunctive kriging that should suffice in order to understand the methodological basis of our analysis and results from a fisheries perspective.

We consider the density of Norway lobster as the variable $Z(x_i)$: Number of individuals · km⁻² at location x_i , whose properties are relatively constant over the area at the spatial and temporal scales of study. For instance, it can be reasonably accepted that the abundance of Norway lobster does neither vary abruptly from one location to a neighbouring location nor shows a clear spatial trend, based on what is known from the biology of the species. These assumptions can be verified in practice by an exploratory data analysis previous to geostatistical analysis.

Under these conditions, the structure of spatial variability of Norway lobster abundance can be studied by computing an appropriate structure function, for instance a variogram. The computation of the variogram is summarized in Fig. 2, see also eq. (1) of the Appendix. A vector of distance classes or lags (h) is first specified, together with an angular tolerance. Then, from each sampling point x_i , the distance to each other location is computed and arranged according to the specified distance classes. The squared difference of density values for each pair of samples pertaining to a given lag is then computed. Finally the estimated variogram value for a given lag is obtained, dividing the sum of squared differences by the number of pairs of sampling points pertaining to this lag. This process is illustrated in Fig. 2 considering the first 6 sampling points in our data set.

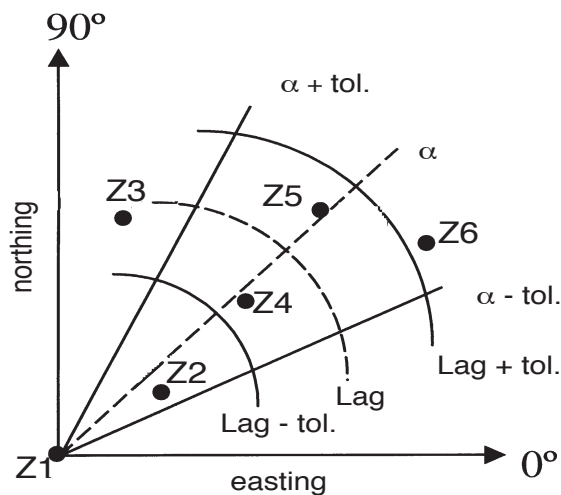


FIG. 2. – Computation of the experimental variogram for distance class Lag and direction α . Starting from the sampling point $Z1$ and specifying $\alpha \pm tol$, sampling points $Z3$, $Z4$ and $Z5$ would be included in the computation of the omnidirectional variogram. For a directional variogram specifying a direction α and an angle tolerance $\pm tol$, only the sampling points $Z4$ and $Z5$ would be taken into account. The computation of the experimental variogram proceeds in this manner using eq. (1) of the Appendix, starting from each sampling point at a time.

The shape of the variogram gives clues to the spatial structure of the variable under study. In ecology, the structure functions known as autocorrelograms (analogous to the variogram) are often the object of the study. They can also be tested statistically, and for this reason are often preferred, in ecological work, in order to make inferences on the spatial structure of ecological variables (Legendre and Fortin, 1989). In this review, we will assume that the variogram suffices to reveal the spatial structure of *N. norvegicus*. Then, by analogy with Sokal and Oden's (1978) simulations employing correlograms, we will attempt to establish the spatial characteristics of Norway lobster in the study area with the aid of the variogram.

Another aspect of geostatistics is the possibility of estimating the density of our variable at points not sampled or over areas where a global estimate is needed. This aspect is known as *kriging* and can be regarded as spatial prediction or estimation. In order to proceed with kriging, the experimental variogram has to be fitted to a theoretical variogram model (some models are given in the Appendix, section 2.2). When the experimental variogram has been computed correctly and there exists a clear underlying spatial structure, it is often easy to choose and fit a theoretical model. When there is no

obvious pattern in our variogram, it is always possible to contemplate other geostatistical models or assume that there is no spatial structure in our variable at the scale of study.

The behaviour of the variogram near the origin requires a detailed study. A variogram may show a discontinuity near the origin, called the nugget effect in geostatistics. The nugget effect can be attributed to measurement error, micro-scale variability or small-scale spatial structure. For instance, below ~ 1 km little can be said from our data set, as this is the average tow length in our study. A nugget component is usually included in most theoretical variogram models.

The variogram fitted can be used to predict the abundance of our variable at locations not sampled (point estimation by kriging) or over a user-defined region (block estimation by kriging). The estimate at a point or block is obtained from linear or non-linear weighting of the observed values. A structure function (for instance the variogram fitted) is used to optimally compute the kriging weights (linear estimation) or to estimate the kriging function (non-linear estimation). Kriging has the important property that for each estimate, an associated estimation variance can be computed.

The point estimation process is useful to produce an accurate map of the resource. For instance, by setting a grid of cells over the study area we can estimate the density of Norway lobster at each node (or cell) and colour-code the abundance values to obtain a map of the resource. In this sense, point kriging is an interpolation method with the advantage that it yields a precision index for each estimate on the map. The non-linear method of disjunctive kriging (Yates *et al.*, 1986; Appendix section 3.2) is employed here as a means to illustrate the use of kriging in mapping and in estimating the probability that Norway lobster density is above a certain profit threshold.

RESULTS

Exploratory Data Analysis

Before building the geostatistical model, it is advised to start with an exploratory data analysis. This preliminary analysis should help check for inconsistencies of the assumptions of the geostatistical model with the data or to suggest a specific geostatistical model to apply, i.e. linear or non-lin-

ear geostatistics, models with drift. A simple way to perform the exploratory data analysis is illustrated in Fig. 3, with the aid of scatterplots of Norway lobster density vs. some ancillary variables. More refined methods can be found in Cressie (1991, pp. 30-51).

The dispersion plot of Norway lobster density vs. depth did not evidence any depth-related trend (Fig. 3a). *Nephrops* density was related neither to time of the day (Fig. 3b) nor to a given geographical direction, such as easting or northing (Fig. 3c and 3d). Hence, a geostatistical model with no spatial drift can be applied to the variable density of Norway lobster. The spatial predictor to choose can be based on a linear predictor (Ordinary Kriging, Table A.1 of the Appendix) or a non-linear predictor (Disjunctive Kriging, Table A.1 of the Appendix).

Variogram computation

The variogram computation starts by choosing a vector h of distance classes, including a tolerance range for unequally spaced samples, as is our case. The choice of h is based on the maximum distance in the field of study divided by 10 or 20, in order to ensure a sufficient number of distance classes and

reveal a meaningful variogram. In the GEOESC-I survey the maximum distance in an E-W direction is ~ 43 km and ~ 25 km in a N-S direction. The maximum distance in the field of study is ~ 50 km, so our first choice of h is to set a distance class at each 2 km, from 0 to 48 km, yielding a total of 25 distance classes. The tolerance specified is ± 1 km. In Fig. 2 the variogram computation is illustrated, see also Appendix, eq. (1).

The variogram computed in this manner is shown in Fig. 4, bottom. The figure shows the estimate of $\gamma(h)$ at each distance class h , together with the number of pairs used to compute this estimate. The number of pairs used to compute each $\hat{\gamma}$ is useful in order to know the quality of the estimate. The weighted least squares method of fitting the variogram is an automated procedure for giving more weight to those distance classes with a higher number of pairs.

A way to obtain a well-defined experimental variogram, easier to fit, is to create a vector of distance classes with unequal spacing. For instance, $h = \{0.5 \pm 0.5, 1.5 \pm 0.5, 2.5 \pm 0.5, 3.5 \pm 0.5, 5.0 \pm 1.0, 7.0 \pm 1.0, 9.0 \pm 1.0, 12.5 \pm 2.5, \dots, 42.5 \pm 2.5\}$ yields a variogram (Fig. 4, top) which is easier to fit to the spherical model (Appendix, section 2.2.1.a, Fig. A.1) and with higher resolution at the small

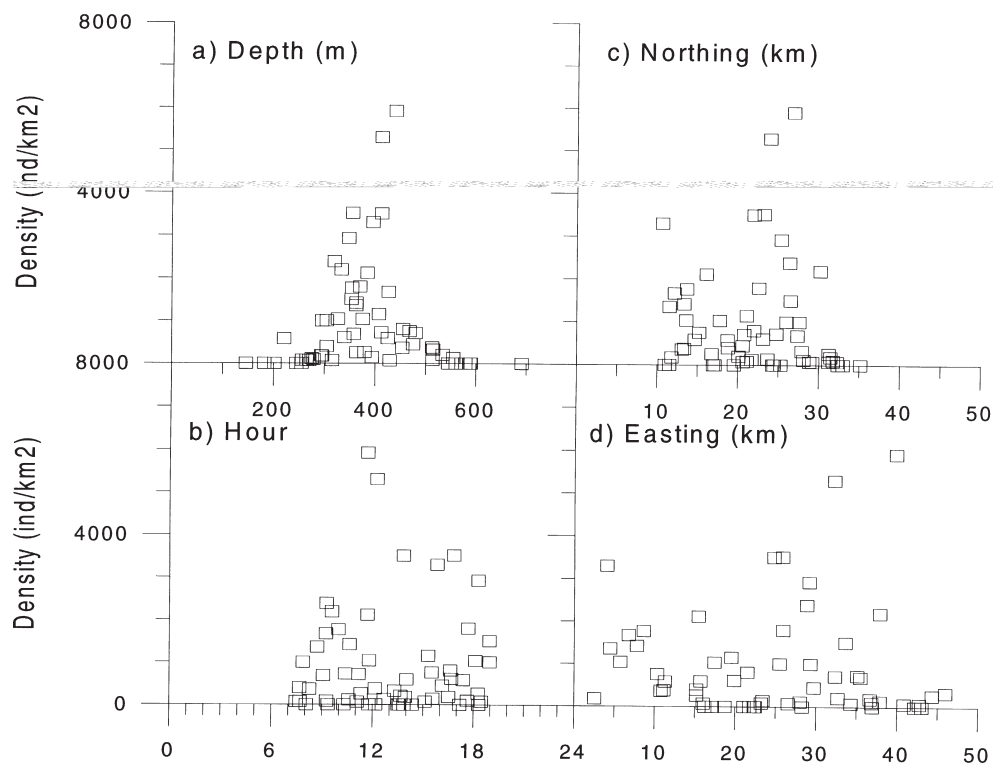


FIG. 3. – Preliminary data analysis of Norway lobster total density vs. some factors that could invalidate the stationary assumptions.

distance classes, thus making it easier to estimate the nugget parameter (c_0 in Appendix, section 2.2.1.a). The manual and weighted least squares (w.l.s.) fits to the variogram in Fig. 4 (top) are given in Fig. 5. Note that in variogram fitting, only those distance classes $|h| < h(max)/2$ were used, as justified in Journal and Huijbregts (1978, p. 194).

As the experimental variogram showed a clearly defined pattern, the two fitting procedures employed yielded essentially the same results, i.e. a spherical model with nugget 4.23×10^5 , sill 1.70×10^6 and range 6.46 km. The fraction of the variance not explained by our model is the ratio nugget to sill, which is 25% in the w.l.s. fit. This variance is attributed to the behaviour of the variable over short ranges, since ~ 2 km is the minimum resolvable distance in our data set. Although a model with zero nugget could also fit the experimental variogram, we consider that the small-scale variability of this highly territorial species (Chapman, 1980) justifies the nugget fitted.

The variogram model could be further refined by studying the existence of anisotropy.

Experimental variograms in an east (0°), north-east (45°), north (90°) and north-west (135°) directions were produced and they did not show any evidence of anisotropy (Maynou, 1995). The east and north-east variograms (along the field of study) showed essentially the same pattern as the omnidirectional variogram, and the north and north-west variograms (across the depth contours) contained few number of pairs and were not well defined.

The results of the variogram analysis for other categories of Norway lobster (juveniles, males, females) and for the second trawl survey (GEOESC-II, autumn 1991) showed essentially the same pattern illustrated here. The ranges of the spherical variograms fitted were between 6 and 10 km, with a mean of ~ 7 km. By analogy with Sokal and Oden's (1978) simulations with correlograms, the spatial structure of *Nephrops* populations in the north-western Mediterranean can be regarded as high-density patches of around 7 km, in the absence of anisotropy. A more detailed study of these data is given in Conan *et al.* (1992) and Maynou *et al.* (1998).

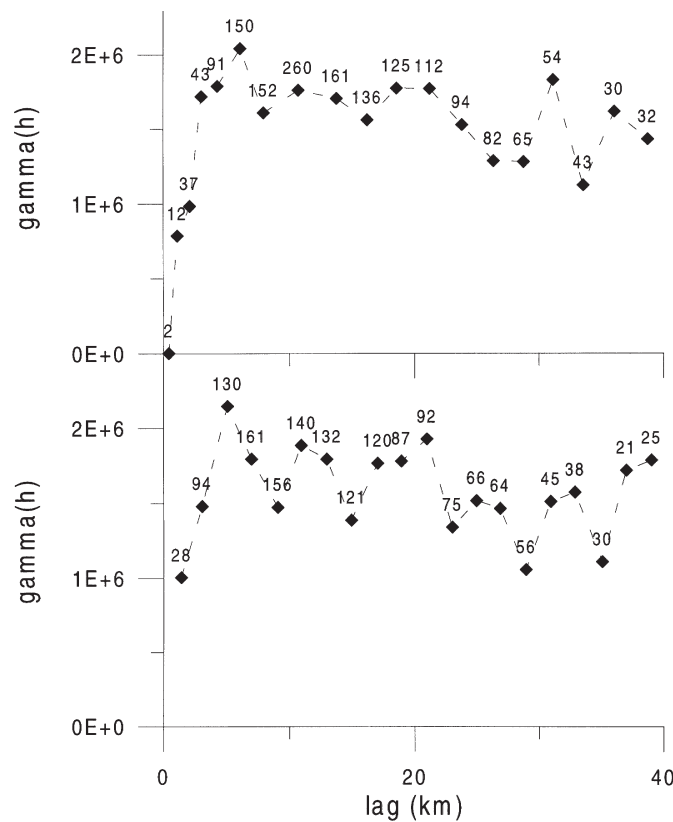


FIG. 4. – Experimental variogram of Norway lobster density using two different lag vectors. Bottom: an equally-spaced vector does not resolve accurately the small distance classes. Top: an unequally-spaced vector allows to better resolve the small distance classes and is easier to fit to a theoretical model. See text for vector specification.

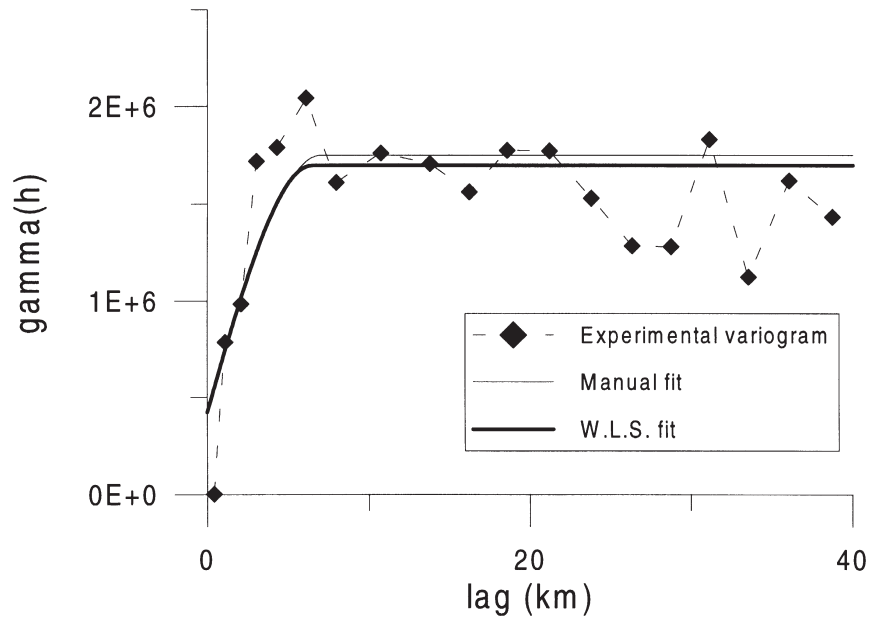


FIG. 5. – Comparison of a manual fit and an automated fit (weighted linear squares) to the experimental variogram in Fig. 4 (top). The variogram model is a spherical variogram (Appendix, section 2.2.1.a). The parameters of the manual fit are: Nugget= 4.50×10^5 , sill= 1.75×10^6 , range= 7.0 km. The parameters of the automated fit are: Nugget= 4.23×10^5 , sill= 1.70×10^6 , range= 6.46 km.

Kriging

To produce a precise map of the density of the resource in the study area, we set first a regular grid of 90×60 cells and estimated the density of *Nephrops* at each cell by disjunctive kriging (Appendix, section 3.2). Each cell had a surface of 0.25 km^2 . The cells falling outside an irregular polygon bounding the data set were not computed.

Disjunctive kriging requires approximating the non-linear predictor by a Hermite polynomial. We chose $K=10$ to truncate the Hermite polynomial. The map in Figure 6 (top left) was produced using the algorithm of Yates *et al.* (1986). The map produced by the disjunctive kriging technique does not differ from a map produced by ordinary kriging, which was presented in Maynou *et al.* (1998). Patches of high density of *Nephrops* can be observed over the fishing grounds known as Can Pere Negre and Serola, with densities over $3600 \text{ ind} \cdot \text{km}^2$ at some cells. The kriging variance obtained when solving the kriging system can be used as a precision index to evaluate the quality of the estimate at each cell. The kriging standard deviation for each cell is shown in Fig. 6 (top, right).

To illustrate the use of disjunctive kriging in computing densities above a user-specified threshold, we tried two levels of economic profit in the

variable *Nephrops* density. The first cut-off corresponded to the usual Norway lobster catch of Barcelona fishermen (for the purpose of this exercise, 15 kg/day). Using the appropriate conversion factors from our experimental trawl and fishing strategy and simulating the commercial fisherman's strategy we arrive at 'usual' catch of $1528 \text{ ind} \cdot \text{km}^2$ or $31 \text{ kg} \cdot \text{km}^2$ (Maynou, 1995). In the same way, we specified a 'high' catch cut-off value (25 kg/day), corresponding to $2,546 \text{ ind/km}^2$ or 51 kg/km^2 (Maynou, 1995). The first cut-off level (Fig. 6, bottom, left) revealed areas with probability $p > 0.6$ where fishing was economically profitable. These areas were the Serola, Can Pere Negre and Merenguera fishing grounds. The second cut-off level (Fig. 6, bottom, right) indicated that only in Serola could a 'high' catch be expected, with probability $p > 0.6$. This agrees with the fact that most (over 90%) Norway lobster landed at the Barcelona fish market is caught in the Serola fishing grounds (Sardà and Lleonart 1993).

DISCUSSION

In our geostatistical study of *N. norvegicus*, we uncovered a spatial pattern of high-density areas (patches), alternating with low-density areas. This

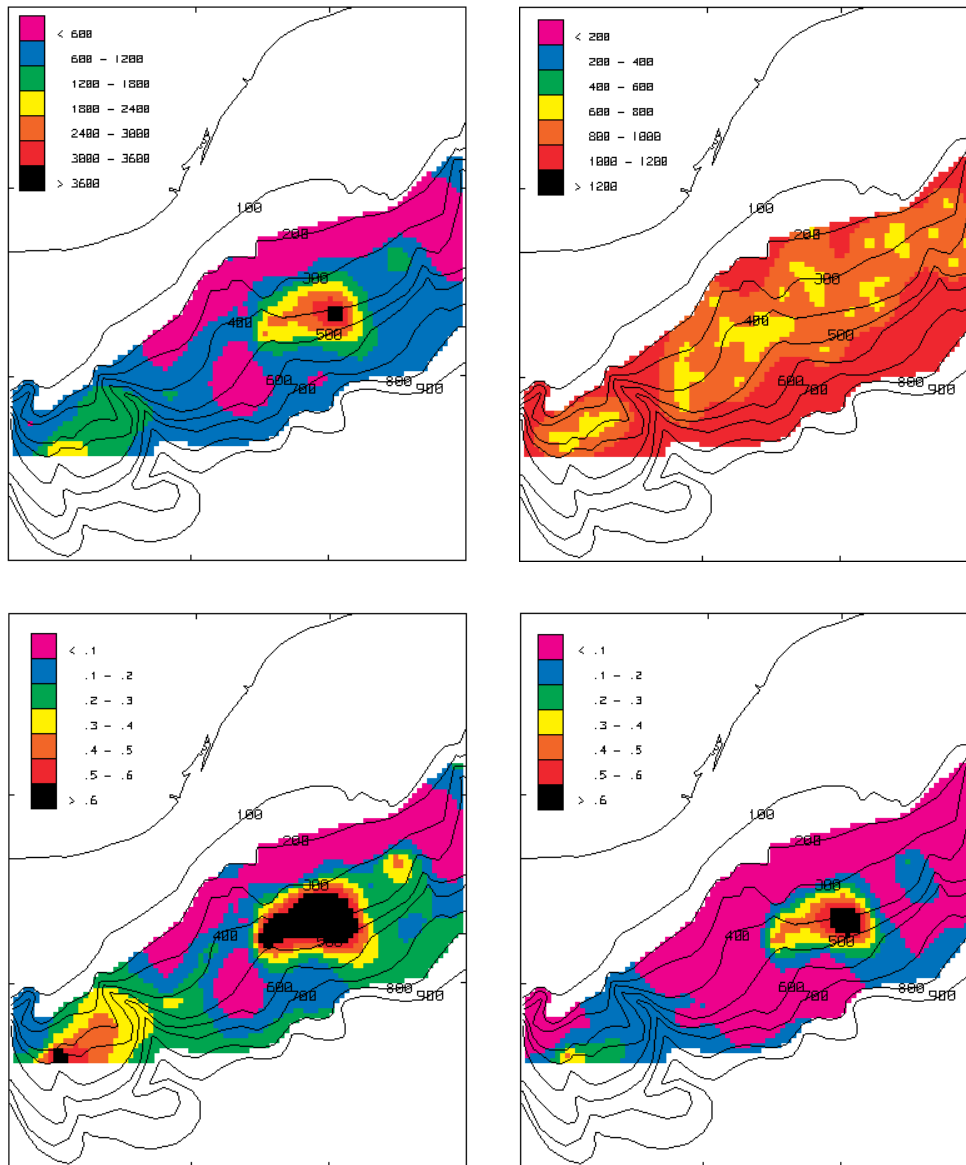


FIG. 6. – Results of disjunctive kriging. Top left: Map of *N. norvegicus* density in the study area, scale units are ind · km². Top right: Precision index (standard estimation of the point estimate) of the previous map. Bottom left: Probability levels for the first cut-off level: 1,528 ind · km² or 15 kg · day⁻¹. Bottom right: Probability levels for the first cut-off level: 2,546 ind · km² or 25 kg · day⁻¹.

patchy distribution is relatively stable throughout the year and across biological categories, as shown by Maynou *et al.* (1998) studying the variograms of juveniles, adults, males and females in two different seasons of the year (spring and autumn). The range of the variograms fitted varied from 6 to 10 km, with a mean of ~ 7 km. Our results can be contrasted with the results of Conan *et al.* (1988, 1996) in studies of the spatial distribution of the snow crab (*Chionoecetes opilio*). These authors showed that considerable sexual and size spatial segrega-

tion exists in snow crab due to its complex spatial behaviour.

Our results are not directly comparable to the results of Fariña *et al.* (1994) who studied the distribution of Norway lobster in Galician waters (NW Spain, Atlantic Ocean). They found a range of ~100 km in the spatial distribution of this species, but their minimum distance among stations was of the order of the range obtained in our study. Thus, their results showed rather a large-scale areal distribution of *Nephrops*.

High-resolution mapping by kriging has been used by fisheries biologists as a tool to forecast accurately the location and the spatial characteristics of an exploited resource. In this sense, several applications illustrate the potential of ordinary kriging to mapping and estimation of resource abundance, such as the contributions of Conan *et al.* (1988, 1992, 1994, 1996), Sullivan (1991), Simard *et al.* (1992) and others cited in the reference list. Disjunctive kriging has not previously been used in a fisheries context to our knowledge. It has the advantage of providing a probability level of commercial profit, allowing the identification of commercially viable areas in new or undeveloped resources.

Our knowledge on the behaviour of *N. norvegicus* indicates that this is a species of very low mobility at the spatio-temporal scales considered here (Chapman, 1980). Thus, due to its low mobility and benthic habits, *N. norvegicus* can be considered an ideal case for the application of geostatistics. Other sessile crustaceans and molluscs have been the object of geostatistical studies (Armstrong *et al.*, 1992; Freire *et al.*, 1992; González Gurriarán *et al.*, 1993; Maynou *et al.*, 1996; Simard *et al.*, 1992) from an ecological or fisheries perspective.

The underlying processes that generate the observed spatial structure have been discussed elsewhere by Maynou and Sardà (1997). In a statistical analysis of Norway lobster populations off the Ebro delta (located ~ 100 km south of the present study area), these authors found that substrate characteristics could be linked with certain population parameters of Norway lobster. The geostatistical analysis in that area revealed the same pattern of spatial structure shown here: High-density patches of ~ 7 km, no sex or size segregation (unpublished results). From the analysis of sediment variables, such as grain size and redox potential (indicative of organic matter content), the spatial structure of some environmental factors could be related to the patchy distribution of Norway lobster. Chapman and Howard (1988) and Tully and Hillis (1995) reported similar findings for Irish sea Norway lobster. However, other authors suggest that the patchy distribution of this species in the Irish Sea is attributable to larval dispersal determined by hydrological factors (White *et al.*, 1988; Hill and White, 1990).

We conclude that geostatistics can make a significant contribution to fisheries science in those aspects regarding the study of temporal and spatial distribution patterns of abundance in commercially exploited populations.

ACKNOWLEDGEMENTS

The research project NEMED, funded by the EC, DG. XIV(MED92/008) provided the opportunity to carry out the sampling cruises GEOESC-I and GEOESC-II. I thank all participants in the cruises for their assistance during the fieldwork, especially Dr. Joan Cartes, Dr. Joan B. Company and Dr. F. Sardà. My sincere appreciation to Dr. G. Conan for introducing me to geostatistics.

REFERENCES

- Anonymous. – 1990. Report of the study group on the applicability of spatial statistical techniques to acoustic survey data. *ICES*. CM 1990/D: 34.
- Anonymous. – 1991. Report of the Workshop on the Applicability of Spatial Statistical Techniques to Acoustic Survey Data. *ICES*. CM 1991/D: 40.
- Armstrong, M. and G. Matheron. – 1986a. Disjunctive kriging revisited: Part I. *J. Math. Geol.* 18: 711-728.
- Armstrong, M. and G. Matheron. – 1986b. Disjunctive kriging revisited: Part II. *J. Math. Geol.* 18: 729-742.
- Armstrong, M., D. Renard and P. Berthou. – 1992. Applying geostatistics to the estimation of a population of bivalves. *ICES*. CM 1989/K:37.
- Bárdossy, A., I. Bogardi and W.E. Kelly. – 1988. Imprecise (fuzzy) information in geostatistics. *J. Math. Geol.* 20(4): 287-311.
- Bárdossy, A., I. Bogardi and W.E. Kelly. – 1990. Kriging with imprecise (fuzzy) variograms. I: Theory. *J. Math. Geol.* 22(1): 63-94.
- Berry, B.J.L. and D.F. Marble. – 1968. *Spatial analysis. A reader in statistical geography*. Prentice-Hall: New Jersey.
- Besag, J.E. – 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society* B. 36: 192-225.
- Brooker, P.I. – 1986. A parametric study of robustness of kriging variance as a function of range and relative nugget effect for a spherical variogram. *J. Math. Geol.* 18(5): 477-488.
- Burrough, P.A. – 1983. Multiscale sources of spatial variation in soil. I. The application of fractal concepts to nested levels of soil variation. *J. Soil Sci.*, 34: 577-597.
- Chapman, C.J. – 1980. The ecology of juvenile and adults *Nephrops*. In: J.S. and B.F. Phillips (eds.). Vol. II, pp. 143-178. *The biology and management of lobsters*. Cobb, New York: Academic Press.
- Chapman, C.J. and F.G. Howard. – 1988. Environmental influences on Norway lobster (*Nephrops norvegicus*) populations and their implications for fishery management. In: Fincham, A.A. and P.S. Rainbow (eds.) *Symposia of the Zoological Society of London* No. 59. *Aspects of decapod crustacean biology*. Oxford: Clarendon Press.
- Cliff, A.D. and J.K. Ord. – 1973. *Spatial autocorrelation: Models and applications*. Pion. London.
- Cochran, W.G. – 1977. *Sampling techniques*. New York: Wiley.
- Conan, G.Y. – 1985. Assessment of shellfish stocks by geostatistical techniques. *ICES*. CM 1985/K: 30.
- Conan, G.Y., F. Maynou and F. Sardà. – 1992. Direct assessment of the harvestable biomass from a stock of *Nephrops norvegicus*, seasonal and spatial variations. *ICES*. CM 1992/K: 22.
- Conan, G.Y., F. Maynou, D. Stolyarenko and L. Mayer. – 1994. Mapping and assessment of fisheries resources with coastal and depth constraints, the case study of snow crab in the Bay of Islands fjord (Newfoundland). *ICES*. CM 1994/B:4.
- Conan, G.Y., M. Moriyasu, E. Wade and M. Comeau. – 1988. Assessment and spatial distribution surveys of snow crab stocks by geostatistics. *ICES*. CM 1988/K: 10.
- Conan, G.Y., M. Starr, M. Comeau, J.-C. Theriault, F. Maynou and G. Robichaud. – 1996. Life history strategies, recruitment fluctuations, and management of the Bonne Bay Fjord Atlantic

- Snow Crab (*Chionoecetes opilio*). In: *High latitude crabs: Biology, Management, and Economics*. pp. 60-97. Alaska Sea Grant College Program.
- Cressie, N. – 1991. *Statistics for spatial data*, Wiley-Interscience. New York.
- Cressie, N. and D.M. Hawkins. – 1980. Robust estimation of the variogram: I. *J. Math. Geol.*, 12(2): 115-125.
- David, M. 1977. *Geostatistical ore reserve estimation*, Elsevier. Amsterdam.
- Fariña, A.C., J. Freire and E. González-Gurriarán. – 1994. *Nephrops norvegicus* in the Galician continental shelf (NW Spain): abundance and distribution. *Fish. Res.*, 19: 333-347.
- Freire, J., E. González Gurriarán, I. Olaso. – 1992. Spatial distribution of *Munida intermedia* and *M. sarsi* (Crustacea: Anomura) on the Galician continental shelf (NW Spain): Application of geostatistical analysis. *Est. Coast. Shelf Sci.*, 35: 637-648.
- Geary, R.C. – 1954. The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 5: 115-141.
- González Gurriarán, E., J. Freire and L. Fernández. – 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.
- Hill, A.E. and R.G. White. – 1990. The dynamics of Norway lobster (*Nephrops norvegicus*, L.) populations on isolated mud patches. *J. Cons. Int. Explor. Mer.*, 46: 167-174.
- Isaaks, E.H. and R.M. Srivastava. – 1989. *An introduction to applied geostatistics*, Oxford University Press. New York.
- Journel, A.G. – 1980. The lognormal approach to predicting local distributions of selective mining unit grades. *J. Math. Geol.*, 12(4): 285-303.
- Journel, A.G. – 1983. Nonparametric estimation of spatial distributions. *J. Int. Assoc. Math. Geol.*, 15: 445-468.
- Journel, A.G. and C.J. Huijbregts. – 1978. *Mining Geostatistics*, Academic Press. New York.
- Jumars, P.A. – 1978. Spatial autocorrelation with RUM (Remote Underwater Manipulator): vertical and horizontal structure of a bathyal benthic community. *Deep-Sea Res.*, 25: 589-604.
- Jumars, P.A., D. Thistle and M.L. Jones. – 1977. Detecting two-dimensional spatial structure in biological data. *Oecologia*, 28: 109-123.
- Legendre, P. – 1993. Spatial autocorrelation: Trouble or new paradigm?. *Ecology*, 74(6): 1659-1673.
- Legendre, P. and M.-J. Fortin. – 1989. Spatial pattern and ecological analysis. *Vegetatio*, 80: 107-138.
- Legendre, P. and M. Troussellier. – 1988. Aquatic heterotrophic bacteria: Modeling in the presence of spatial autocorrelation. *Limnol. Oceanogr.*, 33(5): 1055-1067.
- Mackas, D.L. – 1984. Spatial autocorrelation of plankton community composition in a continental shelf ecosystem. *Limnol. Oceanogr.*, 29(3): 451-471.
- Margalef, R. – 1991. *Teoría de los sistemas ecológicos*. Publicacions de la Universitat de Barcelona. Barcelona.
- Matheron, G. – 1962. *Traité de Géostatistique appliquée, Tome I. Mémoires du Bureau de recherches géologiques et minières.*, No. 14. Editions Technip. Paris.
- Matheron, G. – 1963. *Traité de Géostatistique appliquée, Tome II: Le Krigeage. Mémoires du Bureau de recherches géologiques et minières.* No. 24. Editions du Bureau de recherches géologiques et minières. Paris.
- Matheron, G. – 1969. *Le Krigeage universel*, Cahiers du Centre de Morphologie Mathématique, No. 1. Fontainebleau, France.
- Matheron, G. – 1971. *The theory of regionalized variables and its applications*, Fontainebleau, France: Cahiers du Centre de Morphologie Mathématique, 5. (pàgines) llibre
- Matheron, G. – 1973. The intrinsic random functions and their applications. *Advances in Applied Probability*, 5: 439-468.
- Matheron, G. – 1976. A simple substitute for conditional expectation: The disjunctive kriging. In: M., M. David and C. Huijbregts (eds.). pp. 221-236. *Advanced Geostatistics in the Mining Industry*. Guarascio, Reidel. Dordrecht.
- Maynou, F. – 1995. *Modelització espacial de crustacis decàpodes. Aspectes ecològics i pesquers*. Ph.D. Thesis. Polytechnical University of Catalonia. Barcelona.
- Maynou, F., G.Y. Conan, J.E. Cartes, J.B. Company and F. Sardà. – 1996. Spatial structure and seasonality of decapod crustacean populations on the northwestern Mediterranean slope. *Limnol. Oceanogr.*, 41: 113-125.
- Maynou, F. and F. Sardà. – 1997. *Nephrops norvegicus* population and morphometrical characteristics in relation to substrate heterogeneity. *Fish. Res.*, 30: 139-149.
- Maynou, F., F. Sardà and G.Y. Conan. – 1998. Assessment of the spatial structure and biomass evaluation of *Nephrops norvegicus* (L.) populations in the northwestern Mediterranean by geostatistics. *ICES J. mar. Sci.*, 55: 102-120.
- Moran, P.A.P. – 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37: 17-23.
- Myers, D.E. – 1989. To be or not to be... stationary? That is the question. *J. Math. Geol.*, 21(3): 347-362.
- Pelletier, D. and A.M. Parma. – 1994. Spatial distribution of Pacific halibut (*Hippoglossus stenolepis*): An application of geostatistics to longline survey data. *Can. J. Fish. Aquat. Sci.*, 51: 1506-1518.
- Petitgas, P. – 1993. Geostatistics for fish stock assessments: review and an acoustic application. *ICES J. mar. Sci.*, 50: 285-298.
- Petitgas, P. and J.C. Poulard. – 1989. Applying stationary geostatistics to fisheries: a study on hake in the Bay of Biscay. *ICES. CM 1989/G: 62*.
- Rendu, J.M. – 1979. Normal and lognormal estimation. *J. Int. Assoc. Math. Geol.* 11: 407-422.
- Rendu, J.M. – 1980. Disjunctive kriging: A simplified theory. *J. Math. Geol.*, 12(4): 306-321.
- Rossi, R.E., D.J. Mulla, A.G. Journel and E.H. Franz. – 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecol. Monog.*, 62(2): 277-314.
- Sardà, F., and J. Leonart. – 1993. Evaluation of the Norway Lobster (*Nephrops norvegicus*, L.) resource of the "Serola" bank off Barcelona (western Mediterranean). *Sci. Mar.*, 57(2-3): 191-197.
- Sardà, F., J.E. Cartes and W. Norbis. – 1994. Spatio-temporal structure of the deep-water shrimp *Aristeus antennatus* (Decapoda: Aristeidae) population in the western Mediterranean. *Fish. Bull.*, 92: 599-607.
- Sheshinski, R. – 1979. Interpolation in the plane: The robustness to misspecified correlation models and different trend functions. In: Patil, G.P. and M. Rosenzweig (eds.) *Contemporary quantitative ecology and related econometrics*, pp. 399-420. International Co-operative Publishing House, Fairland, Maryland.
- Simard, Y., P. Legendre, G. Lavoie and D. Marcotte. – 1992. Mapping, estimating biomass, and optimizing sampling programs for spatially autocorrelated data: case study of the northern shrimp (*Pandalus borealis*). *Can. J. Fish. Aquat. Sci.*, 49: 32-45.
- Sokal, R.R. and Oden, N.L. – 1978. Spatial autocorrelation in biology. I. Methodology. *Biol. J. Linn. Soc.*, 10: 199-228.
- Sullivan, P.J. – 1991. Stock abundance estimation using depth-dependent trends and spatially correlated variation. *Can. J. Fish. Aquat. Sci.*, 48: 1691-1703.
- Tully, O. and J.P. Hillis. – 1995. Causes and spatial scales of variability in population structure of *Nephrops norvegicus* (L.) in the Irish Sea. *Fish. Res.*, 21: 329-347.
- White, R.G., A.E. Hill and D.A. Jones. – 1988. Distribution of *Nephrops norvegicus* (L.) larvae in the western Irish Sea: an example of advective control on recruitment. *J. Plankton Res.*, 10: 735-747.
- Yates, S.R., A.W. Warrick and D.E. Myers. – 1986. A disjunctive kriging program for two dimensions. *Computers and Geosciences*. 12(3): 281-313.

1. DEFINITIONS

1.1. Regionalized variables and random functions

The object of study of geostatistics (or surface pattern analysis) are continuously varying phenomena, i.e. phenomena which take a value at each point of space. These phenomena are termed regionalized variables in geostatistics.

A regionalized variable can be defined as a function taking a value $z(x)$ at each point of the space R^n , where $z(x)$ only depends on x (the geographical location). In this article, $n=2$; i.e., we consider only two dimensions of space. The density of *N. norvegicus* (Number of individuals/unit area) is then a regionalized variable $z(x)$. The bottom trawl sampling used to obtain the density of *N. norvegicus* over the study area (A) can be regarded as a realization of the regionalized variable, and the observed densities at the sampling points x_i are noted $Z(x_i)$.

The set of realizations of $z(x)$ is called a random function, $\{Z(x), x \in A\}$ (Journel and Huijbregts, 1978), which is defined by its distribution function:

$$F_{x_1 \dots x_n}(z_1 \dots z_n) = P \{Z(x_1) \leq z_1 \dots Z(x_n) \leq z_n\}.$$

It is not possible to make statistical inference with a single realization of the random function, because its distribution function cannot, in general, be known. Thus some assumptions on the statistical behaviour of the random function have to be made in order to proceed with the statistical analysis. Geostatistical models can become increasingly elaborate by increasing the number of assumptions, but it is preferable to keep the assumptions to a minimum in order to describe in a realistic manner the phenomenon being analyzed.

Most assumptions consider that the regionalized variable repeats itself in space and the sampling is representative of the regionalized variable, i.e. if we were to repeat the sampling over the same study area A we would obtain the same results (statistically). In linear geostatistics (the simpler and more widely used geostatistical methods in fisheries science), the assumptions taken are of stationarity and isotropy.

1.2. Stationarity

The stationarity assumptions are taken on the first and second order moments of the distribution function (Journel and Huijbregts, 1978; Myers, 1989).

The first order moment does not depend on location:

$$E [Z(x)] = m, \quad \forall x.$$

The second order moments (centered) are:

$$\begin{aligned} E [(Z(x) - m(x))^2] &= \text{Var}[Z(x)], && \text{variance;} \\ E [(Z(x_1) - m(x_1)) (Z(x_2) - m(x_2))] &= \\ E [Z(x_1) Z(x_2)] - m^2, && \forall x; \\ E [(Z(x_1) - Z(x_2))^2] - E^2 [Z(x_1) - Z(x_2)] &= \\ E [(Z(x_1) - Z(x_2))^2], && \forall x. \end{aligned}$$

Letting $x_1=x$ and $x_2=x+h$ (where h is a vector of distance) the last two expressions become:

$$\begin{aligned} E [Z(x) Z(x+h)] - m^2 &= C(h), \\ \text{covariogram function (E1);} \\ E [(Z(x) - Z(x+h))^2] / 2 &= \gamma (h), \\ \text{variogram function (E2).} \end{aligned}$$

Under the second-order stationarity hypothesis, the following identities are useful:

$$\begin{aligned} \text{Var} [Z(x)] &= C(0), \\ \gamma (h) &= C(0) - C(h). \end{aligned}$$

Both, $\gamma(h)$ and $C(h)$ can be used to characterize the structure of spatial variability of $Z(x)$ by means of their estimators (see below). Variations of these functions are also employed in the geostatistical literature, such as correlogram, relative variogram, madogram, etc. (Cressie, 1991; Maynou *et al.*, 1996).

$\gamma (h)$ is more general than $C(h)$ because it assumes only that the variance of the increments is finite. The concept of finite variance of the increments in the variogram is called the intrinsic hypothesis (Matheron, 1971). This hypothesis is weaker than the second-order stationarity hypothesis and includes it.

Second-order stationarity hypothesis and the intrinsic hypothesis require an unlimited domain, but in practice the analysis is limited to a finite field. Within the specific spatio-temporal scale of the problem to be addressed it is recommended to

check that our data do not depart to a great extent from these hypotheses. Preliminary exploratory data analysis can be used to verify the adequacy of the data to the assumptions being taken (Cressie, 1991; Maynou *et al.*, 1998).

1.3. Anisotropy

If the regionalized variable under study shows the same spatial structure on every direction (α) then the phenomenon is isotropic. When there exists anisotropy, then $\gamma(h)$ and $C(h)$ are functions of a also: $\gamma(h, \alpha)$ and $C(h, \alpha)$. There exist two types of anisotropy (David, 1977; Isaaks and Srivastava, 1989). In the case of geometric anisotropy the spatial continuity in one of the directions is different but of the same nature than in other directions. The problem can be reduced to isotropy by correcting the units of the anisotropic axis (affine transformation of the Euclidean space by an anisotropy coefficient, Cressie, 1991).

In the case of zonal anisotropy the nature of the spatial variability is different in one of the directions (Isaaks and Srivastava, 1989). This can be

illustrated by considering a water column (R^3): the spatial continuity over the plane x_1, x_2 is of the same nature, while in x_3 (vertical axis) it is fundamentally different and depends on the gravity (Margalef, 1991, pp. 136-139). In order to incorporate zonal anisotropy in the spatial modeling of $Z(x)$ it is necessary to obtain independent or external information.

2. STRUCTURE FUNCTIONS

2.1. Experimental variogram and covariogram:

The underlying autocorrelation function of a random function $Z(x)$ is often unknown but under the stationarity conditions detailed above, it can be estimated from the observations $Z(x_i)$. Experimental variograms $\hat{\gamma}(h)$, or covariograms $\hat{C}(h)$, are computed first and later fitted to a theoretical variogram function. The spatial autocorrelation descriptors are computed as follows (estimators obtained by the method of moments, Matheron, 1962):

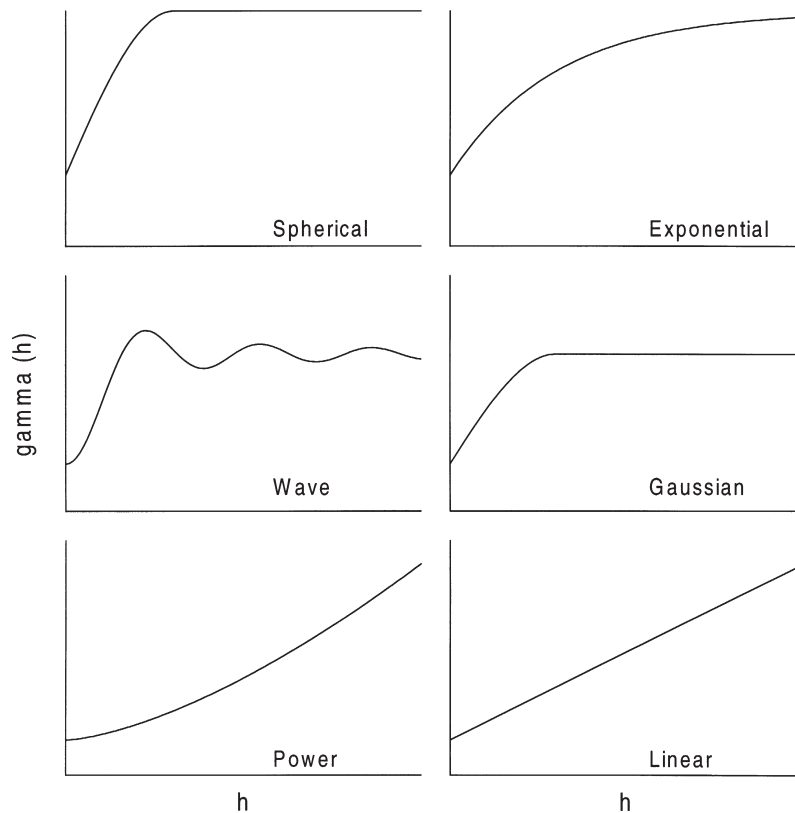


FIG. A – Some common variogram models.

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} (Z(x_i) - Z(x_j))^2, \quad (1)$$

where h is a vector of distance, including a user-specified tolerance in irregular samplings, and $N(h) = \{(x_i, x_j): x_i - x_j = h; i, j = 1 \dots n\}$ is the number of pairs used in computing the experimental variogram for each distance class. The experimental covariogram is (Matheron, 1962):

$$\hat{C}(h) = \frac{1}{N(h)} \sum_{N(h)} (Z(x_i) - \bar{Z})(Z(x_j) - \bar{Z}), \quad (2)$$

where: $\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z(x_i)$.

The vector h can be computed for specified directions α yielding directional variograms or covariograms.

2.2. Variogram models

Given that $\gamma(h)$ is of more general use than $C(h)$ or other related functions, the remaining of this review is illustrated with the variogram function and the experimental variogram. An advantage of using $\gamma(h)$ is that the variogram is defined for intrinsically stationary processes, which is a class including second order stationary processes. Also, $\hat{\gamma}(h)$ (Matheron, 1962) is an unbiased estimator of $\gamma(h)$, while $\hat{C}(h)$ is biased of order $O(1/n)$ when estimating $C(h)$ (Cressie, 1991, p.71); however, an unbiased estimator of $\hat{C}(h)$ has been proposed by Isaaks and Srivastava (1989).

To proceed with the geostatistical modeling the experimental variogram needs to be fitted to a theoretical variogram function. The latter has to comply with certain mathematical conditions (Matheron, 1971; Cressie, 1991). Some valid models defining variogram functions are shown in Fig. A (Journel and Huijbregts, 1978; Cressie, 1991). The models depicted include a constant c_0 , which is known as the nugget effect (see below). The formulae used to compute the variogram function are as follows:

2.2.1. Transition models: $\gamma(h)$ grows continuously up to a certain range α where it stabilizes around the sill $C(0) = c_0 + c'$. The range α is the distance beyond which there is no spatial correlation

among samples. $\gamma(h)$ reaches the sill only asymptotically in the exponential model and oscillates around the sill in the wave (or hole-effect) model.

a. Spherical model: $\gamma(h) = \begin{cases} c_0 + c' \left(\frac{3h}{2a} - \frac{1}{2} \frac{h^3}{a^3} \right) & h < a \\ c_0 + c' & h \geq a \end{cases}$,

when $h=a$, $\gamma(h) = C(0)$.

b. Exponential model: $\gamma(h) = c_0 + c' \left\{ 1 - e^{-\frac{h}{a}} \right\}$,

when $h=3a$, $\gamma(h) = 0.95 C(0)$.

c. Wave model: $\gamma(h) = c_0 + c' \left\{ 1 - \frac{a \sin\left(\frac{h}{a}\right)}{h} \right\}$,

reaches the sill periodically, for $h=\pi a$ and multiples.

d. Gaussian model: $\gamma(h) = c_0 + c' \left\{ 1 - e^{-\frac{h^2}{a^2}} \right\}$,

for $h = \sqrt{3a}$, $\gamma(h) = 0.95 C(0)$.

2.2.2. Non-transition models: Their general equation is $\gamma(h) = c_0 + p h^b$. When $b=1$, the equation yields the linear model; when $0 < b < 2$, the equation yields the power (or fractal) model. When the growth of the variogram is of order two or higher, the intrinsic hypothesis is violated. In such cases, there exists a drift or trend in the regionalized variable. The drift has to be separated from $z(x)$ and modeled independently (Sullivan, 1991).

A combination of the models given in 2.2.1 and 2.2.2 can be used to model nested structures. Nested structures are useful to characterise spatial phenomena that present different variogram models at different distance ranges.

There exist a number of methods for fitting the experimental variogram to a theoretical variogram model. The most common method consists in graphically fitting the two or three parameters that define the theoretical model. This simple method is very useful for well-defined experimental variograms, which easily suggest the variogram function to choose. Some mathematical methods for variogram fitting are reviewed by Cressie (1991,

pp. 90-104) and include maximum likelihood methods, least squares methods or cross-validation. However, a very accurate fitting of the variogram model is not essential when using the fitted model in spatial interpolation, as our own (Maynou, 1995) or Sheshinski's (1979) simulation results show.

2.2.3 *Some problems encountered when using Matheron's estimator:* 1) This estimator is not robust to outliers in the data set. A robust estimator has been proposed by Cressie and Hawkins (1980). 2) The fitting of the experimental variogram reduces the spatial variability of the original data set. This implies that the estimation variance computed by kriging (see below) underestimates the real variance. An approach to this problem was investigated by Bárdossy *et al.* (1988, 1990) using fuzzy set theory.

2.3. Nugget effect

Although strictly $\gamma(0)=0$, the function $\gamma(h)$ often shows a discontinuity at the origin due to small-scale spatial phenomena. It is termed the nugget effect and is simbolized as c_0 in the equations. It originates from the fact that at a scale smaller than the minimum distance between samples, $\gamma(h)$ cannot be studied from the experimental data set. It can be interpreted in a number of ways, depending on the phenomenon under study: as white noise (Matheron, 1962), measurement error, or existence of spatial structure at distances smaller than $\min \|x_i - x_j\|$ (Cressie, 1991). When the nugget effect is observed over the entire range of h (flat variogram), the absence of spatial structure at the scale of study can be assumed.

To ascertain the contribution of a measurement-error nugget effect to the overall variance of

TABLE A.1. Types of kriging classified according to assumptions on the data and on the predictor. $m(x)=E[Z(x)]$ is the mean of the regionalized variable, which can be constant (m) or not ($m(x)$) over the field of study. $Z(x)$ and $Y(x)$ are random functions. See review in Cressie (1991).

Assumptions	Linear predictor	Non-linear predictor
m and $\gamma(h)$ known	Simple Kriging (Matheron, 1962)	
m stationary, $\gamma(h)$ estimated from $Z(x)$	Ordinary Kriging (Matheron, 1962)	Disjunctive Kriging (Matheron, 1976) Indicator Kriging (Journel, 1983) Isofactorial Models (Armstrong and Matheron, 1986a,b)
$Z(x)=m(x)+Y(x)$, $m(x)$ deterministic, $Y(x)$ stationary and $\gamma(h)$ unknown	Universal Kriging (Matheron, 1969)	
$Z(x)=m(x)+Y(x)$, filtering of $m(x)$, general covariance of $Y(x)$ computed from $Y(x)$.	Intrinsic Random Functions of order k (IRF- k) (Matheron, 1973)	
$Z(x)=m(x)+Y(x)$, $m(x)$ known, distribution of $Y(x)$ multinormal, covariance function known.	Markov Random Fields (MRF) (Besag, 1974)	
Normalization of data: $Z(x)=\phi(G(x))$, where $G(x)$ is multinormal	Trans-Gaussian Kriging (Cressie, 1991)	
Logarithmic transformation of data	Lognormal Kriging (Rendu, 1979; Journel, 1980)	

the phenomenon it is necessary to sample with replicates. When the nugget effect is thought to originate from a small-scale spatial phenomenon, it is advised to sample with different spatial intensities, with areas of increased sampling effort at small distances. The contribution of c_0 to the overall estimation variance is important and special attention has to be given to correctly model this parameter (Brooker, 1986).

3. SPATIAL PREDICTION

One of the most important aspects of spatial modeling regards the possibility of making statistical inference of a spatially autocorrelated variable. Here we will focus on the aspects of spatial prediction (estimation) of the value of $Z(x)$ at points not sampled (interpolation) or over areas where a mean or global estimate has to be obtained. The process of estimating the value of $Z(x)$ at unknown locations is termed kriging (Matheron, 1963). Kriging over a grid of points is very useful to obtain an accurate cartography of a resource and in this regard can be seen as an interpolation or mapping method (point kriging). Estimating the value of $Z(x)$ over a bounded area is known as block kriging and can be used as a direct method of biomass assessment in fisheries.

The predictor (estimator) of $Z(x)$ at x_0 is noted $Z^*(x_0)$ and its estimation error $Z(x_0) - Z^*(x_0)$. The theoretical estimation variance is: $E(Z(x_0) - Z^*(x_0))^2$. Theoretical considerations (Cressie, 1991, p. 109) indicate that it is possible to obtain two general classes of unbiased predictors: linear predictors and non-linear predictors. Combining the several assumptions and the two classes of predictors, different types of kriging are obtained. They are summarized in Table A.1, with the most significant reference citations added. Here we will illustrate in detail the linear geostatistical method of ordinary kriging and the non-linear method of disjunctive kriging. The former is the general method employed in most geostatistical fisheries papers (Conan *et al.*, 1988; Maynou *et al.*, 1996, 1998; Simard *et al.*, 1992). The latter has a specific interest in fisheries as it allows estimating the probability that $Z^*(x_0)$ be above a cut-off value. It can be seen that the simpler kriging methods require less assumptions and less parameters to be estimated and they are probably more realistic.

3.1. Ordinary kriging

3.1.1. Point kriging: If $Z(x)$ is a stationary random function, the difference $R(x) = Z(x_0) - Z^*(x_0)$ is also a random function, where $Z(x_0)$ is the real value of $Z(x)$ at x_0 and $Z^*(x_0)$ is its estimator. As $Z(x_0)$ is unknown, $R(x)$ is also unknown but under the stationary hypotheses specified in section 1.2 of the Appendix, the computation of its two first moments becomes possible:

$$m_E = E [R(x)] \text{ and } \sigma_E^2 = \text{Var} [R(x)], \text{ called the estimation variance}$$

With these two moments, we can specify the quality of the estimation:

$$m_E = 0, \text{ unbiased and } \sigma_E^2 \text{ small.}$$

Matheron (1963) proposed a linear unbiased estimator $Z^*(x_0)$ of $Z(x_0)$ from the observed values $Z(x_i)$. This estimator and its variance can be computed analytically and minimized under certain conditions (Matheron, 1971; Journel and Huijbregts, 1978). The linear estimator of $Z(x_0)$ is:

$$Z^*(x_0) = \sum_i w_i Z(x_i), \quad (3)$$

where w_i are the kriging weights attributed to each $Z(x_i)$, subject to $\sum_i w_i = 1$ in order to guarantee unbiasedness. The vector of weights to be given to each observed value is obtained by solving a linear kriging system of equations. Kriging is then an optimized weighting of the samples using their autocorrelation structure (Matheron, 1971). The system of equations is (Matheron, 1971):

$$\begin{cases} -\sum_{j=1}^n w_j \gamma(x_0, x_j) - \mu + \gamma(x_0, x_0) = 0, & i = 1 \dots n \\ \sum_{j=1}^n w_j = 1 \end{cases} \quad (4)$$

where μ is a Lagrange parameter and $\gamma(a,b)$ is the value of the fitted variogram function from point a to b . The kriging (or estimation) variance is (Matheron, 1971):

$$\sigma_k^2(x_0) = \sum_{j=1}^n w_j \gamma(x_0, x_j) + \mu \quad (5)$$

The kriging variance can be used to give a precision index to the point estimates $Z^*(x_0)$ or to construct confidence intervals for the estimate, e.g. a confidence interval for $Z^*(x_0)$ at the 0.05 p-level is $Z^*(x_0) \pm 1.96 \sigma_k(x_0)$, under the assumption of normality of the data $Z(x_i)$ (Cressie, 1991).

The kriging variance is affected by several factors, such as the type of kriging or the sampling design. In order to minimize the variance and to obtain precise estimates it is useful to bear in mind that:

- The difference between the real underlying variogram function and the fitted model from the experimental variogram may be important. When solving the kriging system of equations (4), the fitted model is considered correct and the kriging variance does not include the difference between the underlying variogram function and the fitted model. An approach to this problem has been investigated by Bárdossy *et al.* (1988, 1990) using fuzzy set theory.

- The actual sampling design (equally or unequally spaced samples, preferential sampling) influences the kriging variance.

- The data itself $Z(x_i)$ can have high or low variance. Specially, the ratio c_0 to c' is very important. High c_0/c' ratios contribute to a high estimation variance. It is essential to give special attention to the modeling of the first distance classes of the experimental variogram to accurately assess the nugget value (Brooker, 1986; Cressie, 1991).

3.1.2. Global kriging and change of support:

The mean density of the resource, $Z(B)$, over the estimation area B is (Matheron, 1971):

$$Z(B) = \frac{1}{|B|} \int_B Z(x) dx, \quad (6)$$

where $|B|$ is the area (or volume): $|B| = \int_B dx$. The global abundance of the resource is $Z_G(B) = Z(B) \cdot |B|$.

The estimation of $Z(B)$ from the data $Z(x_i)$ is complicated by the problem of change of support (Matheron, 1971; Journel and Huijbregts, 1978). The samples are defined over a support b , while $Z(B)$ is defined over a block $|B|$. Additional variance terms (Journel and Huijbregts, 1978, pp. 77-123, including computer programs) are added to the kriging equations given above to include the error made when estimating $Z(B)$ from the $Z(x_i)$.

These variance terms can only be computed analytically when working over simple geometric shapes, and they have to be solved by numerical approximations in cases where global estimates have to be produced over irregular contours or in places of complex topography (Conan *et al.*, 1994).

The problem of change of support is encountered when trying to estimate the global fish abundance over a large area (in our case study $\sim 100 \text{ km}^2$) from trawl hauls which can be considered point samples ($\sim 0.01 \text{ km}^2$).

Other types of kriging, e.g. universal kriging or kriging with intrinsic random functions of order k can be more suitable to specific applications, especially when the data have a clear spatial trend, which can be modeled and included in these more advanced types of kriging (Matheron, 1969; 1973; Besag, 1974; Rendu, 1979; Journel, 1980; Cressie, 1991). For a fisheries application, see Sullivan (1991), who modeled the spatial trend stemming from the dependence of fish abundance with depth.

3.2. Non-linear geostatistics and disjunctive kriging

Non-linear geostatistics are of interest in fisheries science as it allows computing the probability that an estimated level of abundance is higher than a user-defined cut-off value. This cut-off value can be chosen, for instance, as a value beyond which the resource becomes economically exploitable. In mathematical terms, the problem is to find an approximation to the conditional expectation $P\{Z(x_0) \geq z_0 \mid Z(x)\}$, from the data values only. One such approximation is illustrated by disjunctive kriging (Matheron, 1976). Other non-linear methods are indicator kriging (Journel, 1983) or isofactorial kriging models (Armstrong and Matheron, 1986a; 1986b).

An essential step in disjunctive kriging is to find a transformation ϕ of the original data, $Z(x) = \phi(Y(x))$ such as to guarantee marginal normality in the bivariate distribution $F_{1,2}(z_1, z_2)$. Then $F_{1,2}$ is expressed in terms of F_1 and F_2 (the marginal distribution functions). The non-linear estimator is:

$$Z^*(x_0) = \sum_{i=1}^n f_i(Z(x_i)),$$

which cannot generally be solved analytically.

The f_i are approximated by expansions of Hermite polynomials of order K , as follows:

$$h_i(y) = f_i(\phi(y)) = \sum_{k=0}^K a_{ik} \eta_k(y), \quad i = 1 \dots n,$$

$$\phi(y) = \sum_{k=0}^K b_k \eta_k(y),$$

where η_k are Hermite polynomials of order K and the coefficients a_{ik} are $n \times K$ unknowns. The $k=1 \dots K$ system of equations is:

$$\sum_{i=1}^n a_{ik} \rho_{i,j}^k = b_k \rho_{0,j}^k, \quad j = 1 \dots n,$$

where $\rho_{i,j} = 1 - \gamma_Y(h)/C(0)$, Rendu (1980). These a_{ik} are employed in the computation of approximate estimator:

$$Z^*(x_0) \cong \sum_{i=1}^n \sum_{k=0}^K a_{ik} \eta_k(\phi^{-1}(Z(x_i))).$$

Its estimation variance is given by:

$$\sigma_{dk}^2(x_0) \cong \sum_{k=0}^K b_k^2 - \sum_{k=0}^K b_k \sum_{i=1}^n a_{ik} \rho_{0,i}^k.$$

The value of K chosen to truncate the Hermite polynomial is an important and subjective factor that may influence the quality of the estimation. Values of $K=10$ or higher can be recommended. An additional practical problem is the need to invert K matrices of $n \times n$, which can slow down the computing process considerably. The estimation of $Z(B)$ is analogous to $Z(x_0)$, taking into consideration the point-to-block correction as discussed in 3.1.2 of the Appendix. Disjunctive kriging algorithms can be found in Yates *et al.* (1986).