Available online at www.sciencedirect.com

Applied Soil Ecology 27 (2004) 189–196

Applied Soil Ecology

www.elsevier.com/locate/apsoil

The effect of sampling unit size on the perception of the spatial pattern of earthworm (*Lumbricus terrestris* L.) middens

J.-P. Rossi [a,*], V. Nuutinen [b]

[a] *UMR 137 Biodiversité et Fonctionnement des Sols IRD/Université Pierre et Marie Curie, 32 Avenue Varagnat, F-93143 Bondy, France*
[b] *MTT Agrifood Research Finland, Soils and Environment, FIN-31600 Jokioinen, Finland*

Received 23 June 2003; received in revised form 27 February 2004; accepted 2 March 2004

Abstract

Sampling strategy, in terms of physical size and positioning of the sampling units, may affect strongly the results of spatial surveys. The aim of this study was to analyse the effect of sampling unit size on the perception of spatial patterns of earthworm (*Lumbricus terrestris* L.) middens. The implications for optimal sampling strategy for spatial interpolation were assessed. Spatial variation of midden density was investigated in Vaisakko forest, south-western Finland, using 225 sample points distributed on a square grid with a minimum distance of 25 m between samples. At each point, middens were counted within samples of sequentially increasing size (sample surface 0.125, 0.25, 1 m^2) and analysed by means of geostatistics. The results showed significant spatial continuity of midden distribution in all cases. Whereas, neither the estimate of mean middens density nor the global distribution pattern were markedly affected by sample unit size, the total variance increased considerably with decreasing sample unit area. Isotropic variograms for different sample unit sizes were all spherical but large discrepancies in the model parameters were observed. The nugget variance tended to decrease with increasing sample unit size while the spatial variance increased slightly. Since changing sample unit size affected the variogram we also investigated the consequences in terms of optimal sampling strategy for spatial interpolation by punctual kriging. Increasing the quadrat size from 0.25 to 1 m^2 and simultaneously increasing the sample spacing from 25 to 50 m, so that the sampling effort was constant in terms of total surface investigated, did not affect the kriging standard deviation. The positive effect of increasing quadrat size was thus enough to compensate the negative effect of the correspondingly sparser sampling grid. The results showed that while the sampling unit size did not have strong effect on the perception of general midden distribution in the forest, it did have marked consequences in the spatial modelling of the phenomenon. The issue of sampling unit size is clearly worthy of careful consideration in the planning of field studies and geostatistical tools can be put to good use in evaluating the pros and cons of different sampling strategies.

Keywords: Spatial distribution; Sampling unit size; Geostatistics; *Lumbricus terrestris*; Earthworm middens; Kriging

1. Introduction

The importance of sampling unit size for the perception of ecological phenomena has long been acknowledged (see Levin, 1992 and references therein). Changing the sample unit size has important conse-

---

[*] Corresponding author. Present address: I.N.R.A., Centre de Bordeaux – UMR BIOGECO Biodiversité, Gènes et Ecosystèmes, Domaine de l'Hermitage, 69, route d'Arcachon, F-33612 Cestas cedex, France. Tel.: +33-5-5712-2859; fax: +33-5-5668-0546.
E-mail address: rossi@pierroton.inra.fr (J.-P. Rossi).

0929-1393/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.apsoil.2004.03.001

quences for the sample variance and can hence affect the results of spatial analyses. In soil ecology standard sample sizes, proportional to the size of the study organism, are typically used (e.g. Stein and Ettema, 2003). There is, however, startlingly little information available on the importance of sample unit size on the perception of spatial population patterns in soil organisms.

The present study explores the effect of sample unit size upon the estimation of earthworm *Lumbricus terrestris* L. midden distribution. Mapping of *L. terrestris* midden density was carried out in a forest habitat using different quadrat sizes. Middens, which are formed on the soil surface at the openings of the permanent home burrows of *L. terrestris*, consist of collected surface litter and of castings. The diameter of middens varies widely, but it is often 7–10 cm. *L. terrestris* is responsible for high bioturbation in soil and hence contributes to soil structure dynamics as well as nutrient turn-over (Brown et al., 2000). For instance, in two studies carried out in temperate deciduous forests it was estimated that *L. terrestris* population alone could bury the entire litter fall during the course of the year (Nielsen and Hole, 1964; Satchell, 1967). Earthworm populations are often spatially structured at the field scale (Phillipson et al., 1976; Poier and Richter, 1992; Stein et al., 1992; Nuutinen et al., 1998). The spatial patterns result from various controlling factors, such as soil physico-chemical characteristics and intra- and interspecific population interactions, while their relative importance often remains poorly understood.

The first aim of the work was to examine the magnitude of quadrat size effect on the sample mean and variance. The main motivation for this elementary investigation was that in earthworm ecology explicit justification of sample unit size is often lacking. Evaluation of variance as a function of sample unit size is a natural first step in making the planning of sampling more transparent and objective.

The second objective was to explore with variography the nature of spatial continuity and assess the effect of quadrat size on the variogram parameters. Theoretically, the main effect of decreasing the sample unit size is an increase in the variogram magnitude, i.e. the sill variance (Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989). Apart from that we were particularly interested in the impact of quadrat size upon the nugget variance. One should expect a linear decrease in this spatially independent term with increasing quadrat size (Bellehumeur and Legendre, 1997) and we wanted to determine to what extent this prediction holds for midden density. We suspected that the prediction would not be verified with our data because of the short-range spatial autocorrelation in earthworm spatial pattern (Rossi, 2003). In this situation the nugget variance would encompass some nested spatial variance and therefore would not be a truly spatially independent term.

Thirdly, we assessed the impact of quadrat size upon the maximum punctual kriging variance. Kriging is an interpolation method commonly used in soil ecology to perform spatial estimation and construct contour maps (Ettema and Wardle, 2002). Each estimate is accompanied by its standard deviation, which allows the evaluation of the estimation's accuracy. An important question in the study of spatial patterning is how different combinations of sampling unit sizes and spacing compare in their performance regarding kriging variance. To tackle this question we compared the spatial interpolation variances associated with different combinations of sample unit sizes and spacing.

2. Materials and methods

2.1. Study site

Field data were collected in the summer of 1993 at the eastern part of the Vaisakko nature sanctuary, which is situated at the south-western coast of Finland (60°21′N, 23°02′) (Anonymous, 1990). The size of study area was approximately 12 ha. In its western margin the area covered a strip of an old growth forest with lime (*Tilia cordata*), spruce (*Picea abies*), birch (*Betula pendula*) and oak (*Quercus robur*). Towards the east, the area extended across an ancient strait, now a few meters above sea level due to land uplift. This area was previously used as a meadow and planted with birch in 1971. The central/eastern parts of the study area were characterised by notable topographical variation; a few bedrock highs with shallow, rocky soils and even exposed rock, with pine (*Pinus sylvestris*) and spruce and to a lesser extent aspen (*Populus tremula*) and oak. The lower elevation sites surrounding the highs had mostly lush, deciduous vegetation with maple (*Acer platanoides*), hazel

(*Coryllus avellana*), lime, oak and birch. In its easternmost corner the study area was delimited by the seashore. No extensive soil survey has been carried out at the site, but according to limited soil sampling done in spring 1997 the textural class varied widely from clay at the birch plantation to gravelley sand near the seashore. In addition to *L. terrestris* L. 4 earthworm species were encountered during the study: *Lumbricus castaneus* (Sav.), *Aporrectodea caliginosa* (Sav.), *Aporrectodea rosea* (Sav.) and *Dendrobaena octaedra* (Sav.). None of the four species forms structures that could be mistaken for *L. terrestris* middens.

2.2. Sampling strategy

L. terrestris middens were counted in 225 sampling sites distributed on the nodes of a traversed, square grid at intervals of 25 m (Fig. 1A). Distances between pairs of sample locations ranged from 25 to 550 m. Three different sampling unit sizes were used: 0.125, 0.25 and 1 m^2. Due to the experimental design two sets of 0.125 m^2-samples and three sets of 0.25 m^2-samples were studied (Fig. 1B). The total surface investigated at each location was 1 m^2. In each sub-area, the number of earthworm middens was determined. Estimate for per square meter midden density were calculated based on each sampling unit size (see below).

2.3. Statistical analyses

The resulting data sets were used to compute experimental variograms (Goovaerts, 1997). Data pairs were grouped into 13 distance classes between 0 and 200 m. The spatial lag was 15 m and the sampling locations in the first distance class were separated by a distance of 25 m. The number of pairs in semi-variance estimation ranged from 391 to 1873. For variogram models, the standard functions were fitted to observed values using weighted least squares fit (Isaaks and Srivastava, 1989). Computations were performed using GSTAT (Pebesma and Wesseling, 1998).

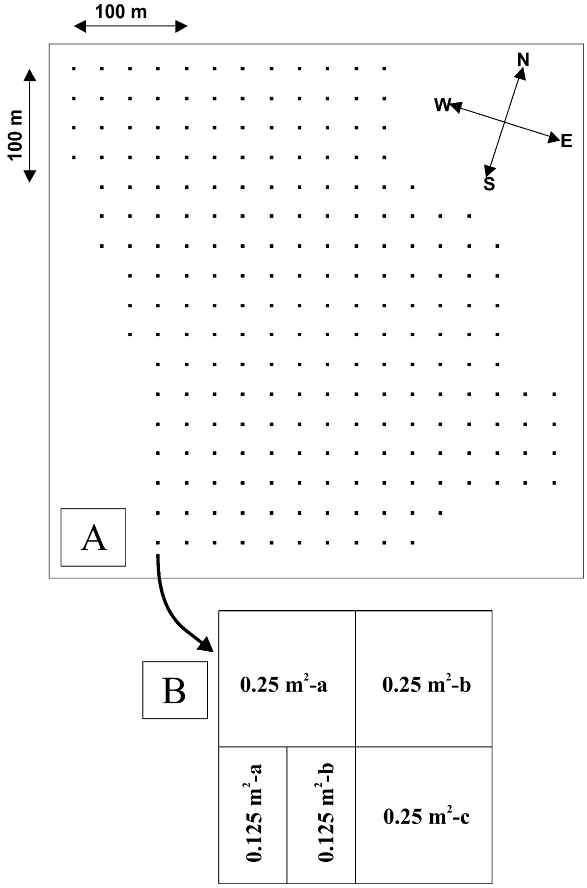

Fig. 1. Sampling protocol used in counting *L. terrestris* middens. (A) Spatial distribution of sampling units ($n = 225$). (B) Structure of each sampling unit consisting of five subunits.

Classical statistical theory assumes a linear relationship between sample unit size and the sample variance (e.g. Bellehumeur and Legendre, 1997). When small sampling units are aggregated to form larger samples, the variance tends to decrease as a linear function of the sample size according to:

$$\mathrm{Var}(V|A) = \frac{\mathrm{Var}(v|A)}{N} \tag{1}$$

where Var($V \mid A$) is the variance of large sample units V in the area A; Var($v \mid A$) is the variance of the small sampling units v in the area A and N is the number of units v aggregated to form the large sample V. The latter relationship is valid only in the case of independent values, that is, in the absence of spatial autocorrelation. Therefore, it can properly predict the values of the nugget variance (a spatially independent term) but cannot be used to predict the sample variance in the presence of autocorrelation. In that case certain geostatistical procedures can be used to predict the change of variance associated with changes in sampling unit size (Journel and Huijbregts, 1978).

The maximum kriging variances associated with two different sampling schemes (square grids) with equal total sample surface area were estimated using the method proposed by Burgess et al. (1981) and the software OSSFIM (McBratney and Webster, 1981). The two sampling schemes compared were: a sampling unit size of 1 m^2 with sampling points separated by 50 m and a unit size of 0.25 m^2 with samples separated by 25 m.

3. Results

The mean value of midden density varied only slightly according to the sampling unit size but sample variance markedly increased with decreasing sample size (Table 1). The skewness tended to decrease somewhat, but not consistently, as the quadrat size increased. The distribution of *L. terrestris* was very heterogeneous and the maps of midden density based on different sampling unit sizes all exhibited the same patchy distribution pattern (Fig. 2). Midden densities were low in the birch plantation running N–S near the western margin of the study area. The other areas of low densities at the more central and eastern parts of the area relate to the nearness of bedrock highs.

Isotropic variograms indicated the presence of positive autocorrelation for all sampling unit sizes and spherical functions fitted satisfactorily to the observed semi-variance values (Table 2). Irrespective of the quadrat size, the shape of experimental variograms remained unchanged (Fig. 3). By contrast, the model parameters changed markedly with support size (Table 2). The sill variance tended to increase as quadrat size decreased and the same trend appeared also for the nugget and the structural variances (Table 2). However, the nugget variance remained relatively high whatever the sample size. Among the resulting variograms, one exhibited a very large range compared to the others (0.25 m^2-c; see Fig. 3 and Table 2).

The sample variance predicted for a quadrat of 1 m^2 using Eq. (1) and 0.25 m^2 quadrats *a–c* were 15.30, 12.67 and 15.25, respectively. These values are much smaller than the observed variance for 1 m^2 estimates (Table 1). The observed nugget variance for quadrats of 1 m^2 was 19.45. Its associated predictions (Eq. (1))

Table 1
Summary statistics for *L. terrestris* midden density (number of items per square metre) for different sampling unit sizes (cf. Fig. 1; $n = 225$ sample locations)

Sample unit size	Mean	Variance	Minimum	Maximum	Skewness
0.125 m^2-a	5.87	93.71	0	48	1.72
0.125 m^2-b	5.48	75.03	0	40	1.76
0.25 m^2-a	5.3	61.24	0	40	1.64
0.25 m^2-b	4.48	50.7	0	40	1.93
0.25 m^2-c	5.24	61.02	0	40	1.66
1 m^2	5.17	49.99	0	34	1.57

Table 2
Variogram model parameters for *L. terrestris* midden density

Sample unit size	C_0	C	a	$C_0 + C$	$C/(C_0 + C)$
0.125 m^2-a	53.64	29.09	79.72	82.73	0.35
0.125 m^2-b	37.53	27.46	84.67	64.99	0.42
0.25 m^2-a	32.0	20.23	98.07	52.23	0.39
0.25 m^2-b	24.26	20.41	81.78	44.67	0.46
0.25 m^2-c	35.95	18.97	153.86	54.92	0.35
1 m^2	19.45	23.24	93.82	42.68	0.54

C_0: nugget variance, a: range; the distance (m) over which autocorrelation is expressed, $C_0 + C$ = sill variance or the variogram plateau, $C/(C_0 + C)$ = relative structural variance. In all cases the models are spherical functions, i.e. for $h \leq a$, $\gamma(h) = C_0 + C(1.5(h/a) - 0.5(h/a)^3)$; for $h > a$, $\gamma(h) = C_0 + C$.

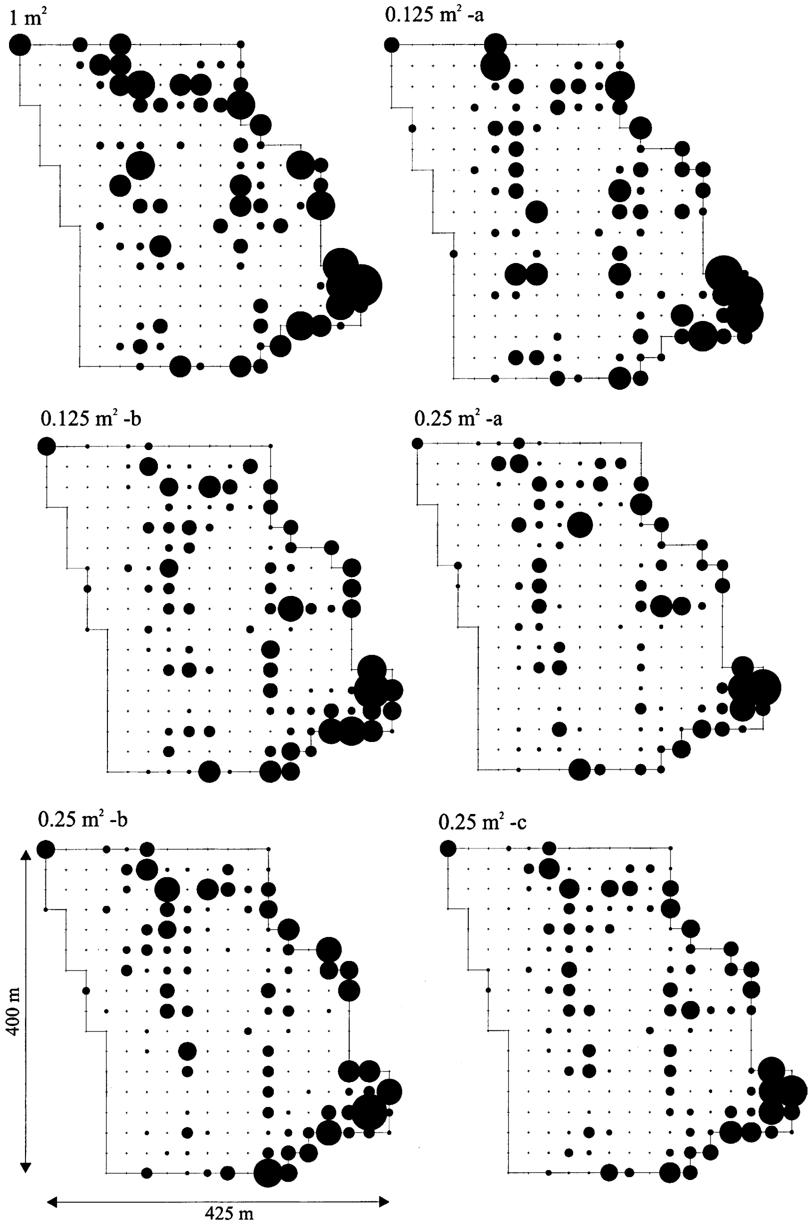

Fig. 2. The spatial pattern of *L. terrestris* midden density (number of items per square meter) for different sampling unit sizes at the eastern part of the Vaisakko nature sanctuary. The size of the symbols is linearly related to the observed density.

from 0.25 m^2 quadrats were 8.00, 6.07 and 8.99 showing a notable underestimation of this term.

In the sampling scheme comparison, using 1 m^2 quadrats with 50 m sample spacing, the maximum punctual kriging variance was 34.8 while reducing the quadrat size to 0.25 m^2 and using a grid spacing of 25 m resulted in a maximum punctual kriging variance of 41.0, 33.8 and 42.3 depending on the variogram used (models established for 0.25 m^2-a–c, respectively).

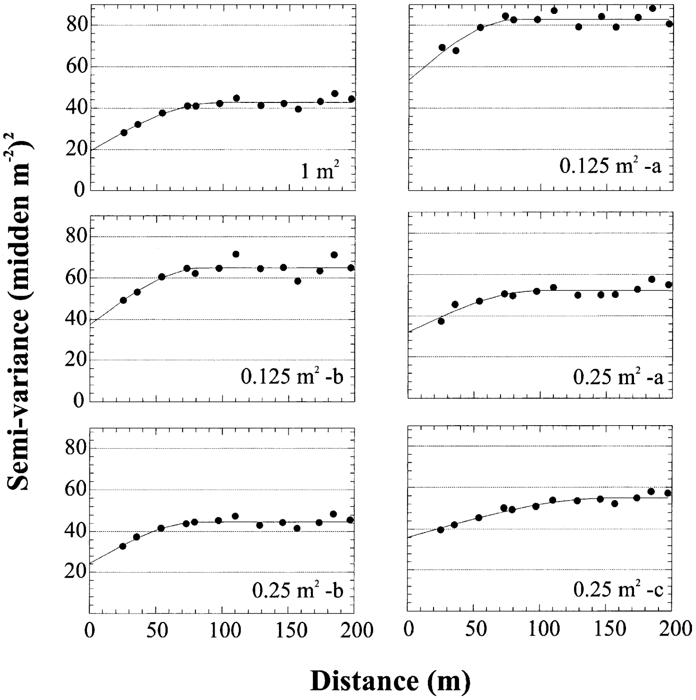

Fig. 3. Experimental and theoretical variograms for *L. terrestris* midden density for different sampling unit sizes ($n = 225$).

4. Discussion

The increase in the variance of *L. terrestris* midden density with decreasing sampling unit size, while the sample mean remained essentially unchanged, corresponds to the dilution of the extreme values when larger quadrats are used (Isaaks and Srivastava, 1989). Increasing the sampling unit size has the effect of reducing the spread of data values and the skewness.

The variation in *L. terrestris* midden density was spatially dependent irrespective of the sampling unit size. Although the overall distribution pattern was similarly captured by different sample sizes, some notable sample unit size effects were revealed by the variogram analysis. While all variograms showed a clear increase in semi-variance with increasing spatial lag until levelling off to the sill variance, the ranges tended to decrease with decreasing quadrat size except in one case.

The main effects of increasing the sampling unit size on the variogram were obvious: the larger the quadrat size, the lower the sill and the smaller the nugget variance. The part of the variability ascribed to random-like small scale variation thus tended to decrease when sample size increased. Conversely, the structural variance increased with quadrat size. Such a trend was also reported by Bellehumeur et al. (1997). The sill variance increased with decreasing quadrat size due to simultaneous increase of nugget variance, simply reflecting the fact that smaller quadrats are more likely to contain extreme values leading to larger variances.

The observed nugget variance for 1 m^2 quadrats was 2–3 times larger than predictions based on the linear relationship. We believe that this discrepancy between observed and predicted values is related to the nature of small-scale spatial patterns in midden distribution. When autocorrelated sampling units are aggregated, the decrease in the variance with increasing sample unit size is less than predicted for independent units. This situation may occur in the case of earthworm middens because earthworms often display strongly aggregated patterns (Boag et al., 1994; Jiménez et al., 2001). Not much is known about the small-scale spatial patterns in earthworm populations but, for example, in some tropical species, patches of individuals may be a few metres in diameter (Rossi, 2003). The present failure to predict the nugget variance most likely points to the presence of very local spatial structures in *L. terrestris* density. It can be hypothesised that the short-range spatial variability is mainly caused by local habitat features such as root architecture or small scale spatial patterns in the plant community (e.g. the "single-tree effect"; Wardle and Lavelle, 1997). Patterns at larger scales are likely related to the interrelated variation of organic matter quantity and quality, vegetation structure and soil type and depth.

Our simulation with two different sampling schemes showed that reducing quadrat size to one quarter and halving the minimum distance between sampling locations slightly diminished the kriging variance in only one case (quadrat 0.25 m^2-b). This underlines the importance of the fact that larger quadrats result in lower nugget variances. In this instance the positive effect of increasing quadrat size was enough to override the negative effect of the correspondingly sparser sampling grid. It therefore would seem of questionable value to decrease the size of the quadrat to increase their number and density, provided that the number of quadrats is large enough to allow satisfactory estimation of the variogram (see Webster and Oliver, 1992).

In choosing the sampling unit size it is important to consider the relationship between the size of object (midden) and the sampling unit. The sampling variance is proportional to the linear size of the object and inversely proportional to the sampling unit size (Bellehumeur et al., 1997). Thus, larger sampling units lead to lower variance. In practice a compromise between the number of samples and the size of each unit needs to be found. Small units with regard to the study objects would yield extremely high variances. On the other hand, increasing the size of the sampling units may lead to a total number of sampling units too small to allow a good estimation of the variogram.

In the present study the time spent in making the measurement in an individual sampling station was small even for the largest sample unit size compared with the time needed for locating and accessing the stations. It would therefore seem advisable to use in this instance the 1 m^2 sample unit size to obtain the minimum level of variance. In most soil ecological studies with time consuming soil sampling and lengthy laboratory treatment of the material very different reasoning would obviously apply. For instance, carrying out an earthworm survey by soil hand-sorting in the grid of the present survey and by using 1 m^2 sampling unit size would be a dubious effort because of the enormous workload and excessive soil disturbance. In such situation sampling unit size must be decreased to the predetermined level of maximum acceptable variance. In many applications, such as in the measurement of certain soil physico-chemical attributes, a possible solution to improve the local representativeness (at a given sampling location) would be to perform composite sampling (Rohde, 1976).

Acknowledgements

We thank the Finnish Forest Research Institute for the permission to work at the Vaisakko forest. We are grateful to Dr. Monika Joschko (ZALF, Müncheberg, Germany), Dr. John Measey (IRD, Bondy, France) and the referees for very helpful comments on the manuscript. VN acknowledges financial support of Maili Autio Foundation.

References

Anonymous, 1990. Vaisakon luonnonsuojelualueen hoito- ja käyttösuunnitelma. Helsinki, 65 pp.

Bellehumeur, C., Legendre, P., 1997. Aggregation of sampling units: an analytical solution to predict variance. Geogr. Anal. 29, 258–266.

Bellehumeur, C., Legendre, P., Marcotte, D., 1997. Variance and spatial scales in a tropical rain forest: changing the size of sampling units. Plant Ecol. 130, 89–98.

Boag, B., Legg, R.K., Neilson, R., Palmer, L.F., Hackett, C.A., 1994. The use of Taylor's Power Law to describe the aggregated distribution of earthworms in permanent pasture and arable soil in Scotland. Pedobiologia 38, 303–306.

Brown, G.G., Barois, I., Lavelle, P., 2000. Regulation of soil organic matter dynamics and microbial activity in the drilosphere and the role of interactions with other edaphic functional domains. Eur. J. Soil Biol. 36, 177–198.

Burgess, T.M., Webster, R., McBratney, A.B., 1981. Optimal interpolation and isarithmic mapping of soil properties. 4. Sampling strategy. J. Soil Sci. 32, 643–659.

Ettema, C.H., Wardle, D.A., 2002. Spatial soil ecology. Trends Ecol. E 17, 177–183.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, Oxford, 512 pp.

Isaaks, E.H., Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, Oxford, 561 pp.

Jiménez, J.J., Rossi, J.P., Lavelle, P., 2001. Spatial distribution of earthworms in acid-soil savannas of the eastern plains of Colombia. Appl. Soil Ecol. 17, 267–278.

Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London, 600 pp.

Levin, S.A., 1992. The problem of pattern and scale in ecology. Ecology 73, 1943–1967.

McBratney, A.B., Webster, R., 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. 2. Program and examples. Comp. Geosci. 7, 335–365.

Nielsen, G.A., Hole, F.E., 1964. Earthworms and the development of coprogenous A-horizons in forest soils of Wisconsin. Soil Sci. Soc. Am. J. 28, 426–430.

Nuutinen, V., Pitkänen, J., Kuusela, E., Wildbom, T., Lohilahti, H., 1998. Spatial variation of earthworm community in relation to soil properties and yield on a grass-clover field. Appl. Soil Ecol. 8, 85–94.

Pebesma, E.J., Wesseling, C.G., 1998. Gstat, a program for geostatistical modelling, prediction and simulation. Comp. Geosci. 24, 17–31.

Phillipson, J., Abel, R., Steel, J., Woodell, S.R.J., 1976. Earthworms and the factors governing their distribution in an English beechwood. Pedobiologia 16, 258–285.

Poier, K.R., Richter, J., 1992. Spatial distribution of earthworms and soil properties in an arable loess soil. Soil Biol. Biochem. 24, 1601–1608.

Rohde, C.A., 1976. Composite sampling. Biometrics 32, 272–282.

Rossi, J.P., 2003. Short-range structures in earthworm spatial distribution. Pedobiologia 47, 582–587.

Satchell, J.E., 1967. Lumbricidae. In: Burges, A., Raw, F. (Eds.), Soil Biology. Academic Press, London, pp. 259–322.

Stein, A., Bekker, R.M., Blom, J.H.C., Rogaar, H., 1992. Spatial variability of earthworm populations in a permanent polder grassland. Biol. Fert. Soils 14, 260–266.

Stein, A., Ettema, C., 2003. An overview of spatial sampling procedures and experimental design of spatial studies for ecosystem comparisons. Agr. Ecosyst. Environ. 94, 31–47.

Wardle, D.A., Lavelle, P., 1997. Linkages between soil biota, plant litter quality and decomposition. In: Cadisch, G., Giller, K. (Eds.), Driven by Nature: Plant Litter Quality and Decomposition. CAB-International, Wallingford, pp. 107–124.

Webster, R., Oliver, M.A., 1992. Sample adequately to estimate variograms of soil properties. J. Soil Sci. 43, 177–192.