Plant and Soil 120, 103–110 (1989).

PLSO 8119

Spatial variability of soil nitrate nitrogen after potatoes and its change during winter

M. VAN MEIRVENNE and G. HOFMAN
Laboratory of Agricultural Soil Science, Faculty of Agriculture, State University of Ghent, Coupure 653, B-9000 Gent, Belgium

Received 30 January 1989. Revised June 1989

Key words: (co)kriging, frequency distribution of NO_3-N, sampling strategy, spatial variability of NO_3-N

Abstract

Knowledge of the frequency distribution and spatial structure of the soil NO_3-N is required to develop an efficient sampling strategy. A 1 ha polder field was sampled after the harvest of potatoes in October 1987, and in February and April 1988, without being fertilized since March 1987. These data sets were examined by a classical statistical as well as a spatial structure analysis. The October and February data sets were found to be lognormally distributed, the April data showed a normal frequency distribution. All three data sets had a spatial structure, although the October data were anisotropic and needed removal of a trend. The spatial variability of soil NO_3-N decreased, became isotropic and evolved towards a larger range of spatial dependence during the winter. Knowledge of this structure permitted to krige or cokrige the data. The number of samples required to estimate the mean NO_3-N content with an acceptable precision was found to be 39, 43 and 17 in October, February and April respectively.

Introduction

An accurate determination of the soil NO_3-N content at the end of the winter is essential for fertilization recommendations (Boon, 1981; Hofman, 1983; Neeteson and Smilde, 1983; Wehrmann and Scharpf, 1979). However, specification of an efficient sampling programme requires knowledge of the frequency distribution and spatial structure of the NO_3-N concentration. Furthermore, it is believed that the simulation of soil processes could be improved by including the spatial variability of their input variables (Dahiya *et al.*, 1984; Warrick and Nielsen, 1980).

Soil NO_3-N frequency distributions have been described as normal (Dahiya *et al.*, 1984), or more commonly as lognormal (Cameron and Wild, 1984; Tabor *et al.*, 1985). In a several years' investment of soil NO_3-N under grassland, White *et al.* (1987) found predominantly a lognormal frequency distribution, irrespective of time of year or depth.

Only little information is available about the spatial structure of soil NO_3-N. Dahiya *et al.* (1985) found no spatial dependence for distances larger than 30 m, which was their closest sampling spacing. White *et al.* (1987) found only scarce evidence for strong spatial dependence of soil NO_3-N over distances > 1 m. However, Tabor *et al.* (1985) describe such a dependence over 150 m.

This paper focusses on the soil NO_3-N concentration in an agricultural field after the cultivation of potatoes, which were chosen because they often leave large NO_3-N residues in the soil profile (Hofman *et al.*, 1984). It has three objectives: (i) to describe the soil NO_3-N frequency distribution and to analyse its spatial structure, (ii) to investigate the sampling strategy for soil NO_3-N and (iii) to discuss the change of these findings over winter.

Materials and methods

The experimental field was a 1 ha subarea of a 15 ha field, situated in the polder area of north East-Flanders (Belgium). The soil was classified as Aquic Udifluvent. The top 110 cm are lime rich Holocene deposits with a lithologic discontinuity, generally at a depth of 40 to 45 cm, between loamy and sandy loam material (Van Meirvenne and Hofman, 1989). Deeper, loamy sand of Pleistocene age is found.

The entire field was planted with potatoes in April 1987. Prior to planting, the field was fertilized in January with 300 kg K_2O ha^{-1} (KCl 60%) and in March with 240 kg N ha^{-1} together with 130 kg P_2O_5 ha^{-1} (compound fertilizer 26 N + 14 P_2O_5). The crop of the entire field was harvested between September 25 and October 10; on the experimental site, potatoes were harvested October 6 and 7. The yield amounted to 58600 kg potatoes ha^{-1} (all data supplied by the farmer).

Between 12 and 15 October, the experimental field was sampled at 247 locations (Fig. 1). At each location, 3 borings were done within 1 m^2 with a gouge auger (ϕ 19 mm) and all the soil up to 1 m depth was collected and mixed. Reference marks were placed to enable the exact reconstruction of the sampling grid after the soil was plowed and sowed with winter wheat in November.

February 17, 1988, the same 247 locations were identically sampled as in October 1987.

Due to the very wet climatic conditions, the farmer could not fertilize the crop before April, instead of the usual February. So, it was possible to sample the soil a third time (on April 1), since it was last fertilized in March 1987. However, the amount of sampling points was reduced to 50 to minimize crop damage. These 50 points were selected by a random number generator and their positions are indicated in Fig. 1. Sampling was done the same way as before.

All soil samples were analysed within two days. To prevent mineralization, samples were kept in a refrigerator or sometimes even frozen. The field moist samples were mixed thoroughly and shaken in 1% $KAl(SO_4)_2$ at a 1:2 soil : solution ratio for 1 h. Filtered extracts were analysed for NO_3-N with a nitrate specific electrode (Cottenie and Velghe, 1973). All analyses were carried out twice. The average coefficient of variation of this procedure was earlier found to be 2% (Hofman *et al.*, 1982).

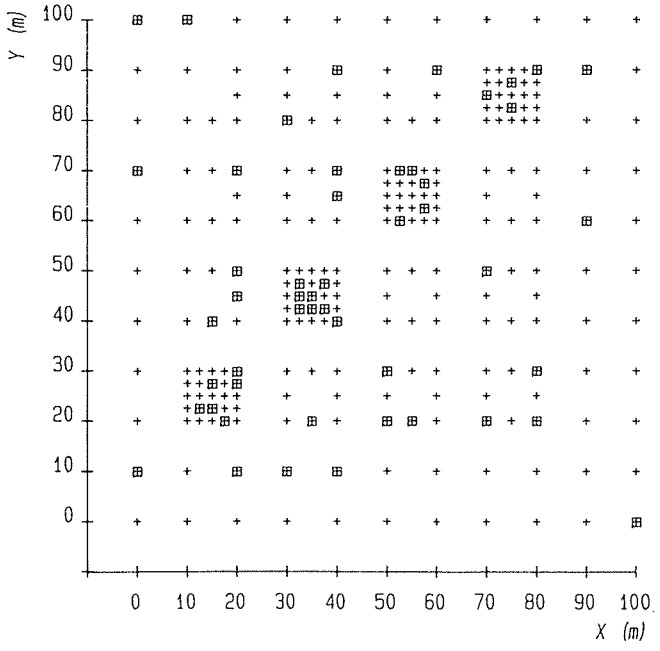

Fig. 1. Sampling locations (crosses were sampled in October and February, squares in April).

Results and discussion

Classical statistical analysis

In this analysis, it was assumed that all observations of each data set were spatially independent.

Sample mean, variance, coefficient of variation, extreme values, skewness (g_1) and kurtosis (g_2) of the three data sets are given in Table 1. In order to verify whether or not the NO_3-N contents were normally distributed, the significance of g_1 and g_2-3 was tested according to Snedecor (1959).

In October and February, the NO_3-N of the top 1 m of the soil was found to be positively skewed ($g_1 > 0$ or extreme values tail off to the right) but normally peaked (g_2 statistically not different from 3). Only the April data could be considered to be normally distributed. To investigate both skewed distributions for lognormality, they were ln transformed. Both g_1's were not significant at $\alpha = 0.05$. So the October and February data sets could be fitted by a lognormal frequency distribution.

Since the values of the statistics describing a population depend on its frequency distribution, recalculation of them was necessary for the October and February data (Vauclin, 1983) (Table 1). By assuming a normal distribution for these two data sets, only a limited error was made by calculating the mean, but an error of 8.1% and 13.6% occurred in estimating s^2 of the October and February observations respectively.

Having characterized the frequency distributions, it is possible to determine the number of observations needed to obtain an estimation of the mean value with a given degree of precision and significance level. This is calculated by (Dahiya *et al.*, 1984; Vauclin, 1983):

$$N = \{[t(\alpha)/\varepsilon]CV\}^2 \quad (1)$$

with N the number of required samples, $t(\alpha)$ the Student's t value for probability $(1 - \alpha)$ and ε the degree of precision requested to estimate the mean value. Table 2 gives some results from $\alpha = 0.05$ for three different degrees of precision.

With an acceptable degree of precision of 10%, 39, 43 and 17 samples would have been necessary to estimate the mean NO_3-N concentration with 95% probability in October, February and April respectively. When a higher precision is desired, the number of samples increases substantially.

Spatial structure analysis

The theory of regionalized variables, first developed by Matheron (1963, 1971), enables the investigation of the structure of the spatial variation of a soil property. The nature of spatial dependence can be measured by calculating the semivariance by (Journel and Huijbregts, 1978):

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

in which $\gamma(h)$ is the semivariance for lag distance h and N(h) the number of pairs of observations $\{Z(i), Z(i + h)\}$ separated by a distance h. Under the condition of stationarity of second order (de Marsily, 1986), the semivariogram (which is a plot of $\gamma(h)$ *versus* h) is a function increasing from zero to a constant value called the sill. The distance over

Table 1. Mean values (m), variances (s^2), coefficient of variation (CV), extreme values, skewness (g_1) and kurtosis (g_2) of the original and ln transformed data sets (Oct, Feb: n = 247, Apr: n = 50)

	m ($mg\,kg^{-1}$)	s^2 ($(mg\,kg^{-1})^2$)	CV (%)	Extremes ($mg\,kg^{-1}$)	g_1	g_2
Original data						
Oct	8.85	7.40	30.7	3.13–16.71	0.58***	2.93
Feb	3.22	1.03	31.5	0.97–6.15	0.47**	2.89
Apr	3.67	0.61	21.2	2.16–5.78	0.33	2.81
Ln transformed data						
Oct	8.87	8.00	31.9		−0.19	2.85
Feb	3.23	1.17	33.6		−0.26	3.00

, * significant at $\alpha = 0.01$ or 0.001 respectively.

Table 2. Number of samples needed to estimate the mean value of the three data sets with a probability of 95% for three different degrees of precision

	ε = 5%	ε = 10%	ε = 20%
Oct	156	39	10
Feb	173	43	11
Apr	69	17	4

which $\gamma(h)$ reaches the sill is called the range, representing the scale over which observations are spatially dependent. Although, by definition $\gamma(0) = 0$, it is often found that as h approaches 0, $\gamma(h)$ remains a positive finite value, named nugget effect.

To investigate whether the data vary isotropically, the semivariograms were calculated in the X (perpendicular to the tillage) and Y (parallel to the tillage) directions. For the October data, a strong anisotropic behaviour was found (Fig. 2). Moreover, the X semivariogram decreased sharply beyond a lag distance of 30 m, indicating a periodic fluctuation of the data in this direction.

To remove the anisotropy of the October data, a trend analysis was carried out. For every X position, the mean NO_3-N concentration in the Y direction was calculated, and a fifth degree polynomial regression was found to fit these data best (higher order terms were no longer significant). The result is plotted in Fig. 3. This trend was subtracted from all 247 data and the X directional semivariogram was recalculated. The residuals could be considered to be isotropic (Fig. 2).

Directional semivariograms of the February data displayed a small anisotropy for large lag distances (> 20 m), but both semivariograms differed little for small lags. Therefore, isotropy was assumed. The April data were isotropic.

Isotropic semivariograms were calculated for the residuals of the October data and for the February[1] and April data and grouped per 5 m class. To these values a theoretical model was fitted using an iteration technique minimizing the mean square differences with the number of data pairs as weight. In practice, only a few theoretical models are authorized to fit experimental semivariograms (McBratney and Webster, 1986). The models used here are (in the following, C_0 is the nugget effect, C_1 the sill and a the range):

– Spherical

$$\gamma(h) = C_0 + (C_1 - C_0)[(3h/2a) - \tfrac{1}{2}(h/a)^3]$$

for $0 \leqslant h \leqslant a$

$$\gamma(h) = C_1 \text{ for } h > a \qquad (3)$$

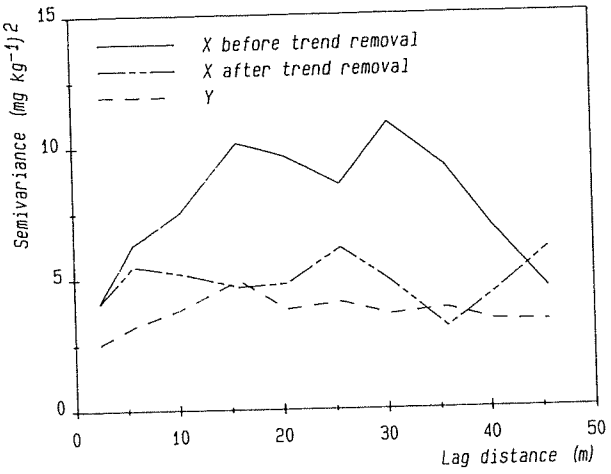

Fig. 2. Bi-directional (X and Y) experimental semivariograms of the October data before and after trend removal.

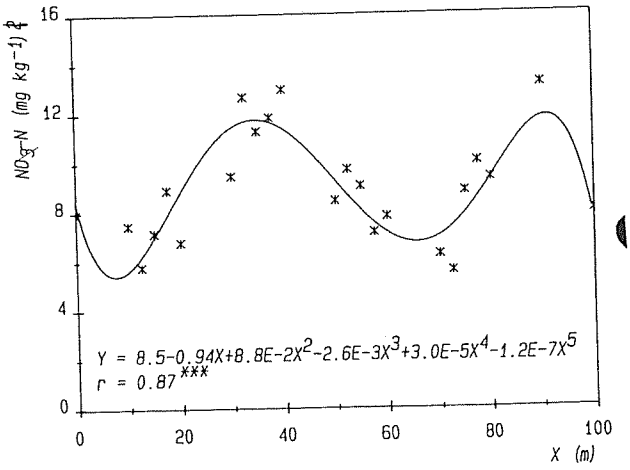

Fig. 3. Mean NO_3-N content for every X position and fitted fifth degree polynomial regression (October data).

[1] Lognormal distributions are often ln transformed to stabilize the semivariance. Since the residuals were normally distributed, this was tested for the February data only. No difference in spatial structure between the ln transformed and original data was found, so the latter were preferred for comparative reasons.

– Exponential

$$\gamma(h) = C_0 + (C_1 - C_0)[1 - \exp(-h/b)] \quad \text{with } a = 3b. \tag{4}$$

Figure 4 shows the results, and Table 3 gives the values of the fitted theoretical semivariograms.

All data sets showed a spatial structure, even the detrended residuals of the October data. The smallest limit of spatial dependence of the soil NO_3-N was found for the residuals of the October data (9.5 m), the largest for the April data (34 m). This indicates an increasing uniformity of the NO_3-N concentration between October and April. Important nugget effects indicate that large parts of the variability were present at distances closer than the smallest sampling distance (2.5 m), or were due to measurement errors.

The spatial structure of the observations allows the use of kriging to interpolate the NO_3-N content at unsampled locations. In punctual-kriging, the estimated value Z* at an unsampled location x_0 is taken to be a linear combination of N measured neighbours Z at locations x_i:

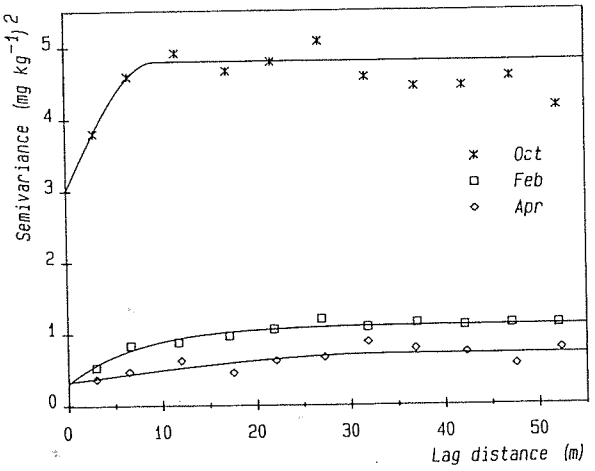

Fig. 4. Experimental (points) and theoretical (lines) semivariograms of the three NO_3-N data sets.

Table 3. Theoretical models fitted to the experimental semivariograms

	Eq. no.	C_0 ((mg kg^{-1})^2)	C_1 ((mg kg^{-1})^2)	a (m)
Oct	3	3.00	4.80	9.5
Feb	4	0.31	1.11	23
Apr	3	0.33	0.71	34

$$Z^*(x_0) = \sum_{i=1}^{N} \lambda_i Z(x_i) \tag{5}$$

λ_i are the weights attributed to the neighbours which are subject to the restriction that:

$$\sum_{i=1}^{N} \lambda_i = 1. \tag{6}$$

The weights can be found by solving the kriging equation system as described by Journel and Huijbregts (1978).

Kriging is done under the conditions of unbiasedness and minimum variance in which sense it is an optimal interpolation technique (Burgess and Webster, 1980). Kriging also provides an estimation variance, s_k^2, of each interpolation. This kriging variance can be used to construct a confidence interval about the estimate (de Marsily, 1986).

A moving neighbourhood of 10 m and 20 m was used to punctual-krige the residuals of the October data and the February data every 2.5 m. Finally, the fifth order polynomial regression was added to the kriged residuals. The results are shown in Figs. 5 and 6.

In April only 50 soil samples were taken. Kriging of these values is possible, but a gain in accuracy is obtained by using additional information supplied by the denser sampled February measurements and their mutual correlation. Cokriging allows the use

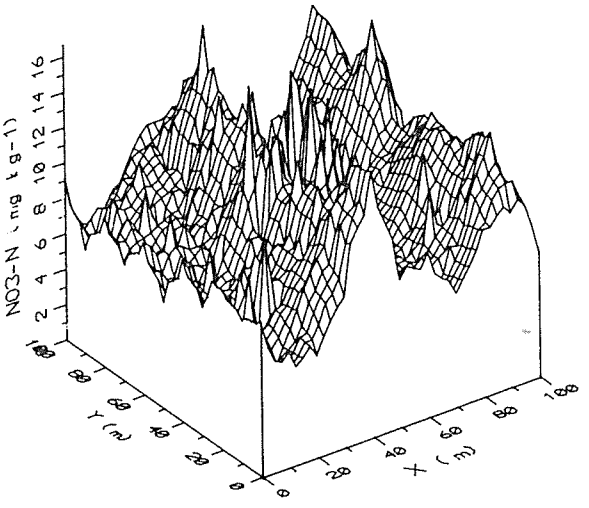

Fig. 5. Punctual-kriged nitrate nitrogen concentration in October.

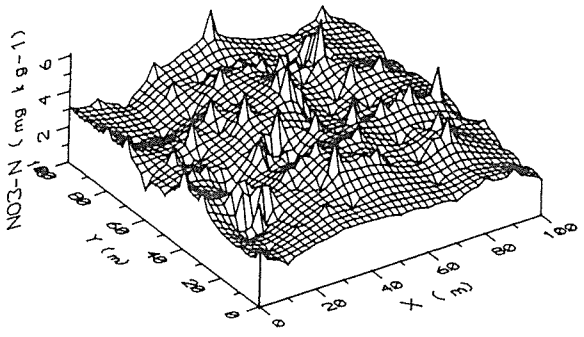

Fig. 6. Punctual-kriged nitrate nitrogen concentration in February.

of this spatial cross correlation (McBratney and Webster, 1983; Vauclin *et al.*, 1983). Therefore, the cross semivariance between both data sets was calculated by:

$$\gamma_{12}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z_1(i) - Z_1(i+h)] \times [Z_2(i) - Z_2(i+h)] \quad (7)$$

with $\gamma_{12}(h)$ the cross semivariance between data set 1 (April data) and 2 (February data). A theoretical model was fitted to the experimental isotropic cross semivariogram in the same way as described before (Fig. 7). Similar to punctual-kriging, in punctual-cokriging the estimator has the form:

$$Z_1^*(x_0) = \sum_{i=1}^{N_1} \lambda_{1i} Z_1(x_{1i}) + \sum_{j=1}^{N_2} \lambda_{2j} Z_2(x_{2j}). \quad (8)$$

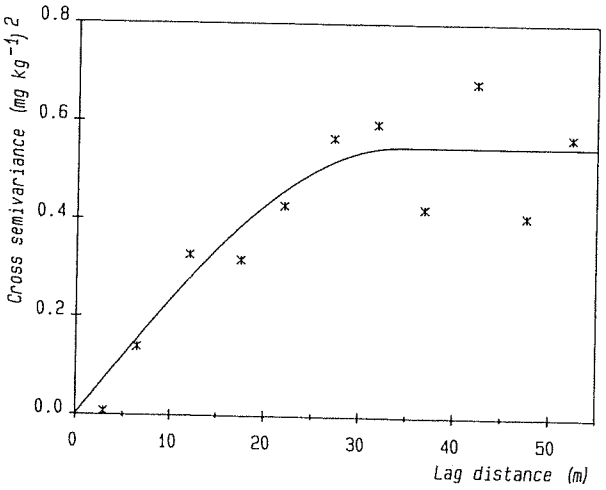

Fig. 7. Experimental (points) and theoretical (lines) spherical cross semivariogram (Eq. 3) between the February and April data ($C_0 = 0$ and $C_1 = 0.55$ (mg kg^{-1})^2 and a = 34 m).

The weights are restricted by the following conditions:

$$\sum_{i=1}^{N_1} \lambda_{1i} = 1 \quad \text{and} \quad \sum_{j=1}^{N_2} \lambda_{2j} = 0. \quad (9)$$

Cokriging also provides an estimation variance for every interpolation.

The cross semivariogram and the semivariograms of the February and April data were used to punctual-cokrige the April data every 2.5 m, using a moving neighbourhood of 20 m (Fig. 8).

The accuracy of the kriging or cokriging estimates can be tested by interpolating sampled locations pretending not to know them (Vieira *et al.*, 1981). In this way, two data sets with the same X, Y coordinates become available, a measured and an interpolated one. This allows the calculation of the mean estimation error (MEE), which should be close to zero to indicate no systematic bias, and the reduced variance (RV), which should be close to unity to verify whether the (co)kriging errors are consistent with the predicted variances (de Marsily, 1986; Vauclin *et al.*, 1983). The results of these calculations are grouped in Table 4, together with the t values to test whether the MEE's are significantly different from zero.

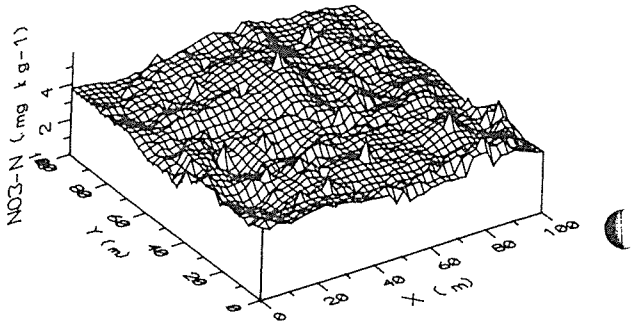

Fig. 8. Punctual-cokriged nitrate nitrogen concentration in April.

Table 4. Mean estimation errors (MEE), its associated t values and reduced variances (RV) used to test the (co)kriging of the different data sets

	MEE	t value	RV
Oct (kriging)	4.64E-03	7.26E-03	0.87
Feb (kriging)	7.30E-03	1.14E-02	0.75
Apr (cokriging)	−3.59E-02	−2.43E-02	0.97

All MEE's were not significantly different from zero at $\alpha = 0.05$, and the RV's were close to one, so it was concluded that the options chosen and the assumptions made were reasonable.

Since all three data sets show a spatial structure, the assumption, made in the preceding chapter, that all observations are independent is not correct. Only data separated by a distance larger than the range of their semivariogram fulfil this assumption. Therefore, the results obtained in the classical statistical analysis should be treated with care. To investigate the impact of the spatial structure of the October data set on this analysis, 121 points separated 10 m apart were used to recalculate the frequency distribution and associated statistics. The spatially independent data were lognormally distributed ($g_1 = 0.73^{***}$ before ln transformation and -0.06 after), the statistics were: $m = 8.82\,mg\,kg^{-1}$, $s^2 = 7.09$ $(mg\,kg^{-1})^2$ and CV = 30.2%. Compared to Table 1, the variance differed most (11.4% smaller). Recalculation of the number of samples necessary to estimate the mean with a given precision and probability (Eq. 1), also resulted in a small reduction only. For $\alpha = 0.05$ and $\varepsilon = 5$, 10 and 20%, N was 140, 35 and 9 respectively. This indicates that the error made by considering these data to be spatially independent was relatively small. Since the range of the semivariogram of the other two data sets was much larger, too few spatially independent points remained to repeat this calculation.

Conclusions

The frequency distributions were found to be highly skewed in October, less skewed, but still lognormal in February, and normally distributed in April. This suggests that, due to a homogenization of the soil NO_3-N concentration, as indicated by the increasing range of spatial dependence, a shift to a normal frequency distribution took place towards the end of the winter. An uneven spreading of N fertilization is most likely responsible for the skewness of the soil NO_3-N content (White *et al.*, 1987).

All data sets showed a spatial structure. The spatial variability of the October NO_3-N data was anisotropic and needed removal of a trend in the X direction. The kriged October data (Fig. 5) showed two ridges of high NO_3-N contents running parallel with the direction of tillage. At the top of these ridges the NO_3-N concentrations were about twice as high as in the intermediate valley (top $\approx$ 13, valley $\approx 6\,mg\,kg^{-1}$). Only one mineral N fertilization was given by the farmer, so it is very likely that uneven spreading of the fertilizer was responsible for these large spatial differences in the soil NO_3-N residues after the cultivation of potatoes. Informed about these results, the farmer was surprised and unaware of such a large variability of his fertilizer application.

In February, the pattern of the ridges and valleys was still detectable, although much less pronounced. Therefore, isotropic conditions could be assumed. This process of homogenization of the nitrate concentration in the soil continued, and in April spatial variability was the smallest and the range of spatial dependence the largest.

Under the assumption of independent observations, 39, 43 and 17 samples would have been necessary to estimate the mean NO_3-N concentration in the soil of this field with 95% probability and with an acceptable degree of precision of 10% in October 1987, February and April 1988 respectively. Due to the spatial structure of these data sets, samples should have been taken at least 9.5, 23 and 34 m apart respectively.

References

Boon R 1981 Stikstofadvies op basis van profielanalyse voor wintergraan en suikerbieten op diepe leem- en zandleemgronden. Pedolgie 31, 347–363.

Burgess T M and Webster R 1980 Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. J. Soil Science 31, 315–331.

Cameron K C and Wild A 1984 Potential aquifer pollution from nitrate leaching following the plowing of temporary grassland. J. Environ. Qual. 13, 274–277.

Cottenie A and Velghe G 1973 Het gebruik van de specifieke nitraat electrode voor de bepaling van nitraten in gronden en planten. Mededelingen Faculteit Landbouwwetenschappen, Rijksuniversiteit Gent 38, 560–568.

Dahiya I S, Kersebaum K C and Richter J 1984 Spatial variability of some nutrient constituents of an Alfisol from loess. I. Classical statistical analysis. Z. Planzenernaehr. Bodenkd. 147, 695–703.

Dahiya I S, Anlauf R, Kersebaum K C and Richter J 1985 Spatial variability of some nutrient constituents of an Alfisol from loess. II. Geostatistical analysis. Z. Planzenernaehr. Bodenkd. 148, 268–277.

Hofman G, Ossemerct C and Ide G 1982 Variability of the mineral nitrogen content in sandy loam soils. Mededelingen Faculteit Landbouwwetenschappen, Rijksuniversiteit Gent 47, 1225–1231.

Hofman G 1983 Minerale stikstofevolutie in zandleemprofielen. PhD Thesis, Univ. of Ghent, Belgium, 183 p.

Hofman G, Ossemerct C, Ide G and Van Ruymbeke M 1984 Significance of the latent mineral N-residue in the soil profile in nitrogen fertilization advices. *In* Fight Against Hunger through improved Plant Nutrition. Proceedings of the 9th World Fertilizer Congress, Budapest, Hungary, 11–16 June. Eds. E Welte and I Szabolcs pp. 225–229.

Journel A G and Huijbregts C J 1978 Mining Geostatistics. Academic Press Inc., London, U.K., 600 p.

Marsily G de 1986 Quantitative Hydrology. Academic Press Inc., London, U.K., 434 p.

Matheron G 1963 Principles of geostatistics. Econ. Geol. 58, 1246–1266.

Matheron G 1971 The Theory of Regionalized Variables and its Applications. Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau no. 5, Ecole Nationale Superieure des Mines de Paris, France, 208 p.

McBratney A B and Webster R 1983 Optimal interpolation and isarithmic mapping of soil properties. V. Coregionalization and multiple sampling strategy. J. Soil Sci. 34, 137–162.

McBratney A B and Webster R 1986 Choosing functions for semi-variogram of soil properties and fitting them to sampling estimates. J. Soil Sci. 37, 617–639.

Neeteson J J and Smilde K W 1983 Correlative methods of estimating the optimum nitrogen fertilizer rate for sugar beet as based on soil mineral nitrogen at the end of the winter period. *In* Proceedings Symposium 'Nitrogen and Sugar Beet'. Ed. IIRB, Brussels, Belgium, pp. 409–421.

Snedecor G W 1959 Statistical Methods. The Iowa State College Press, Ames, Iowa, U.S.A. 534 p.

Tabor J A, Warrick A W, Myers D E and Pennington D A 1985 Spatial variability of nitrate in irrigated cotton. II. Soil nitrate and correlated variables. Soil Sci. Soc. Am. J. 49, 390–394.

Van Meirvenne M and Hofman G 1989 Spatial variability of soil texture in a polder area. I. Kriging. Pedologie 34, 69–87.

Vauclin M 1983 Méthodes d'étude de la variabilité spatiale des propriétés d'un sol. *In* Variabilité Spatiale des Processus de Transfert dans les Sols, Avignon, 24–25 juin 1982. pp 9–43. Ed. INRA Publ. (Les Colloques de l'INRA no. 15).

Vauclin M, Vieira S R, Vachaud G and Nielsen D R 1983 The use of cokriging with limited field soil observations. Soil Sci. Soc. Am. J. 47, 175–184.

Vieira S R, Nielsen D R and Biggar J W 1981 Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J. 45, 1040–1048.

Warrick W R and Nielsen D R 1980 Spatial variability of soil physical properties in the field. *In* Applications of Soil Physics. Ed. D Hillel. pp. 319–344. Academic Press Inc., New York, U.S.A.

Wehrmann J and Scharpf H 1979 Der Mineralstickstoffgehalt des Bodens als Massstab für den Stickstoffdüngerbedarf (N_{min}-Methode). Plant and Soil 37, 109–126.

White R E, Haigh R A and Macduff J H 1987 Frequency distributions and spatially dependent variability of ammonium and nitrate concentrations in soil under grazed and ungrazed grassland. Fert. Res. 11, 193–208.