

How many sampling points are needed to estimate the mean nitrate-N content of agricultural fields? A geostatistical simulation approach with uncertain variograms

Siegfried C.K. Hofman ^a, D.J. Brus ^{b,*}

^a VITO, Boeretang 20, 2400 Mol, Belgium

^b Biometris, Wageningen University and Research, PO Box 16, 6700 AA Wageningen, Netherlands

ARTICLE INFO

Handling Editor: Morgan Cristine L.S.

Keywords:

Pedometrics
Stratified random sampling
Bayes theorem
Markov chain Monte Carlo (MCMC)

ABSTRACT

Knowledge of how many sampling points are needed to estimate the mean content of soil nutrients in agricultural fields, given a precision requirement on the estimated mean, is limited. This paper describes a versatile geostatistical simulation approach for predicting the variance of the mean nitrate-N (NO₃-N) content within an agricultural field estimated by random sampling. In fall of 2016 sixteen agricultural fields were sampled on a square grid to model the spatial variation of NO₃-N. On twelve out of sixteen fields NO₃-N showed a lognormal distribution rather than a normal distribution. Variograms for (log-transformed) NO₃-N are estimated using a Bayesian approach, resulting in 100 vectors with possible variogram parameters per field, obtained by MCMC sampling from the posterior distribution. Each of these variograms is used to simulate 100 maps of NO₃-N, resulting in 100 × 100 maps of NO₃-N per field. Each map is used to compute the variance of the estimated mean with stratified simple random sampling of 5, 10, ..., 50 points, with one point per compact geographical stratum. For each sample size (number of sampling points) the mean, median and P90 of the uncertainty distribution are computed. Based on the medians, the sample size required for a maximum expanded measurement uncertainty of 50% varies from < 5 to > 50. This large variation in required sample size shows the large variation among the sixteen fields in variance of NO₃-N within a field.

1. Introduction

In a world where the demand for agricultural produce is predicted to increase with 50% by 2050 (United Nations, 2017), the need for substantial improvements in both crop production and sustainable use of agricultural resources is of utmost importance. Within any programme to develop site specific management in order to increase crop yield, basic knowledge of the fertility status on the field scale plays an important role. Although analytical procedures for the determination of common physical properties and chemical composition of agricultural soils have since long been optimized and even faster and cheaper techniques e.g. by use of NIRS-technology are up-and-coming (Nie et al., 2017), this is not always true for the preliminary sampling of the soil. Despite numerous studies that show the importance of sampling procedures inspired by the specific characteristics of the measurands, improved procedures tend to be neglected in favour of long-time practice or ease-of-use. This results in sub optimal laboratory samples and subsequent

analytical results (Starr et al., 1992; Clay et al., 1997; Hennart et al., 2004).

Knowledge of NO₃-N availability is paramount when designing any plan to fulfill the crops NO₃-N requirement. Alternatives to sampling and analysis such as the use of sensor networks for use in precision agriculture might prove their worth in future but are still far away from becoming daily practice and come with a hefty investment (Shaw et al., 2016). Therefore there is a need for sampling procedures that take into account the spatial variability of nutrients as to ensure that the collected sample is a good enough representation of the plot's fertility status. Not only would faulty or inaccurate assessment of the soils nitrogen level have a direct derogatory effect on the crops quality and/or quantity, it also tends to lead to overly use of fertiliser which then results in high post-harvest NO₃-N residues. NO₃-N then leaches into ground- and surface water causing eutrophication and poses a direct health risk in those areas where the population is dependent on natural water for use as drinking water, e.g. overly use of NO₃-N fertilisers has been proposed

* Corresponding author.

E-mail address: dick.brus@wur.nl (D.J. Brus).

as one of the governing factors in the occurrence of infant methemoglobinemia (blue baby syndrome) (Kathik et al., 2011).

Sampling for NO₃-N determination poses to be more challenging when compared to other nutrients as a result of its transient nature in soil. Not only is NO₃-N highly mobile making its availability dependent on local weather conditions and irrigation, it is also both formed and removed by nitrification and denitrification. The speeds of these processes is governed by local factors such as water and the presence of organic material that promote bacterial growth (James et al., 1971). In addition, the amount of nitrogen extracted by the crop from the soil in a single growing season can be high: e.g. typical fertilisation advice for potatoes is in the order of 210 kg N/ha (UGent, 2000), but can, in extremis, reach up to 860 kg N/ha in banana plantations (Kathik et al., 2011), whereas an ecologically safe post-harvest soil NO₃-N residue varies somewhere between 20 and 50 kg N/ha (UGent, 2000). As a result of these fast temporal variations and the many processes by which it is formed and removed also the spatial variability of NO₃-N concentrations in soil is high in comparison with other nutrients since natural or man-made factors (e.g. tillage) that tend to homogenise have little effect on the short-lived NO₃-N.

Following the European Council Directive 91/676/EEC a.k.a. the EU Nitrates directive (EU, 1991), member states must, among others, identify and monitor NO₃-N concentrations in waters that are susceptible to NO₃-N pollution. Furthermore, an action programme to reduce water pollution by NO₃-N must be established, but member states are free when it comes to the design of such a programme. In Flanders one of the measures is the monitoring of residual NO₃-N of fields in the fall soon after the crop is harvested, following the rationale that the risk of NO₃-N leaching is independent of the amount of fertiliser as long as it is used by the crops before winter. This allows for more crop-friendly fertilisation practices while still preventing NO₃-N contamination. With this approach accurate monitoring of residual NO₃-N in the soil of the agricultural fields is a prerequisite.

Numerous studies have been performed on the variation of soil NO₃-N under specific soil conditions, crops or fertilisation techniques and its implications for sampling design (White et al., 1987; Starr et al., 1992; Franzen et al., 2011; Kathik et al., 2011; Parmodh et al., 2011). More detailed studies into the spatial variation of soil NO₃-N report both a low proportion of spatially structured variance with high variance at small distances, and high temporal instability (Bogaert et al., 2000; Stenger et al., 2002). Although these studies give important insights in the problems faced when sampling for NO₃-N, they seldom propose a workable sampling method for use on fields for which there is only limited or no prior knowledge on the amount and variability of available NO₃-N. Current sampling practices often use a workaround to compensate for the high spatial variability of NO₃-N in soil and the bad reproducibility in results it induces. Mostly these workarounds consist of limiting the sampling to part of the field instead of sampling the whole surface of the field under study. Examples are cross- or zigzag-shaped sampling patterns used in Flanders (VITO, 2010), sampling around a 30 m diameter circle on a homogeneous presumed part of the field in the Belgian Walloon region (ISSEP, 2014), or limiting the sampling to a “representative” 100 m by 100 m square on a 10 ha field in Germany (Sächsische, 2012). The advantage and often the origins of these techniques lie in their simplicity, ease of use and general applicability on the field. The quality of the results from these procedures however are questionable: an unknown systematic error in the results can be there as only a part of the field is sampled, sampling locations are chosen subjectively and the variance of the random error in the results cannot be quantified.

Identifying sampling locations by their GPS-coordinates, as proposed hereafter, might seem to complicate the process of sampling both in the office and on the field significantly and result in a price increase. One must however bear in mind that contemporary techniques originate from the pre-GPS area. Nowadays the use of dedicated handheld GPS-devices to find specific (predefined) sampling locations is

commonplace and greatly simplifies the task.

We propose to sample the entire field by probability sampling, so that an unbiased estimate of the mean NO₃-N concentration of the field can be obtained (de Gruijter et al., 2006). Following Brus et al. (1999) we propose to sample fields by stratified simple random sampling, using compact geographical strata (geostrata) of equal area. From each geostratum one location is randomly selected, so that the number of geostrata equals the number of soil cores (sampling points). This sampling design is prescribed by the national government of the Netherlands in case a farmer wants to be set free from EU regulations on the application of phosphate on agricultural fields with low phosphate levels. The overall precision of a sampling method is increased by spreading sampling points as evenly over the field as possible. Preliminary splitting up the field in compact geostrata of equal area and randomly selecting one sampling point from these strata largely prevents the sampling locations from spatial clustering while at the same time ensuring that each possible sampling location has an equal probability of ending up in the bulked sample and thus guaranteeing a bias-free composite sample. To save laboratory analysis costs, in the proposed sampling scheme the soil cores collected at the sampling locations are bulked into one composite sample and subsampled for analysis in the lab after thorough homogenisation. Core diameter and sampling depth are kept identical to current schemes: the soil is sampled in 13 mm cores up to a depth of 90 cm.

The question then is how many sampling points do we need given some requirement on the accuracy of the estimated mean. In sampling literature the number of selected population units (in our case sampling points) is referred to as the sample size. In principle this question can be answered by selecting multiple random samples with a given type sampling design (for instance simple random sampling or stratified simple random sampling) and a range of sample sizes from a field, to plot the variance of the estimated mean against the sample size, and use this plot to derive the required sample size given a maximum variance. However, this would be too costly if we want to repeat this for a large number of sample sizes and fields. The alternative is to collect data from a field for modelling the spatial variation of NO₃-N in that field, and to predict the variance of the estimated mean from this model (Domburg et al., 1994). In practice we are always uncertain about the model of spatial variation. In this paper a method is worked out that accounts for uncertainty about the variogram to predict the variance of the estimated mean with stratified simple random sampling and a series of sample sizes. In this method a Bayesian approach is followed to assess the uncertainty about the variogram parameters.

The aim of this research was to collect data on the spatial variation of NO₃-N in agricultural fields, to use these data to build a model of the spatial variation for each field, and to use these models to predict the sampling variance of the estimated mean for the proposed random sampling design for a range of sample sizes.

2. Theory

2.1. Predicting the sampling variance from a variogram

The sampling variance of the estimated population mean with stratified simple random sampling is (de Gruijter et al., 2006)

$$V(\hat{z}) = \sum_{h=1}^L w_h^2 \frac{S_h^2}{n_h}, \quad (1)$$

with L the total number of strata, \hat{z} the estimated population mean, w_h the weight of stratum h which is equal to the relative size of stratum h : $w_h = A_h/A$ (A_h is the area of stratum h , A is the area of the field), S_h^2 the variance of the variable of interest (NO₃-N) within stratum h , and n_h the number of sampling points in stratum h . With strata of equal size so that $w_h = 1/L$, and one point per stratum ($n_h = 1$), the sampling variance reduces to

$$V(\bar{z}) = \frac{1}{L^2} \sum_{h=1}^L S_h^2. \quad (2)$$

Note that with the proposed sampling design described in Section 1 the sampling variance cannot be estimated from the data, because the soil aliquots are bulked into a composite sample. The result is only one number, the measured NO₃-N content of the composite sample, used as an estimate of the mean of the field. Even when the soil aliquots would have been analyzed separately, an unbiased estimator of the sampling variance is not available, because we have one point per stratum only ($n_h = 1$) so that we do not have estimates of the stratum variances S_h^2 . However, the sampling variance of the estimated mean can still be predicted from a variogram, as explained hereafter.

Given a variogram, under the assumption of a constant mean within a stratum, the stratum variance can be predicted by the mean semivariance of all pairs of points within that stratum

$$E_\xi[S_h^2] = \bar{\gamma}_h, \quad (3)$$

with $E_\xi[\cdot]$ the model expectation, and $\bar{\gamma}_h$ the mean semivariance of stratum h . By plugging these model-based predictions of the stratum variances in Eq. 2, a model-based prediction of the sampling variance is obtained:

$$E_\xi[V(\bar{z})] = \frac{1}{L^2} \sum_{h=1}^L \bar{\gamma}_h. \quad (4)$$

In practice $\bar{\gamma}_h$ is approximated by discretizing a stratum by a fine square grid, computing a matrix with distances between all pairs of points that can be formed with these grid nodes, transforming this matrix into a semivariance matrix, and averaging.

An alternative approach is to simulate with the variogram a large number of maps with possible NO₃-N concentrations, followed by computing for each simulated map the variances of the simulated values within the strata (S_h^2 in Eq. 2), and computing the sampling variance of the estimated mean using Eq. 2. This results in as many sampling variances as we have simulated maps. In the case study hereafter we applied this alternative procedure so that we obtain information about the uncertainty of the sampling variance of the estimated mean, given a variogram used in geostatistical simulation.

2.2. Bayesian approach for estimating the variogram

It is evident that we do not have perfect knowledge of the variogram of a given sampled agricultural field. The uncertainty about the variogram parameters will propagate to our uncertainty about the sampling variance of the estimated mean. For that reason, we repeated the simulation procedure described above, for a large number of variograms estimated from the sample data. These variograms are obtained by a Bayesian approach, resulting in a large sample of variogram parameters, sampled from the posterior multivariate distribution of the variogram parameters, rather than in a single variogram. Bayes Rule can be applied to the problem at hand as (Gelman et al., 2013, p. 7):

$$f(\Theta|Z, M) = \frac{f(Z|\Theta, M)f(\Theta|M)}{f(Z)}, \quad (5)$$

with Θ the vector with variogram parameters, Z the available sample data for estimating the variogram, M the variogram model type (spherical, exponential et cetera), $f(\Theta|M)$ our prior belief in the parameters, given a variogram model type, specified by a probability density function, $f(Z|\Theta, M)$ the likelihood function, $f(Z) = \int_\Theta f(Z|\Theta)d\Theta$ the probability density of the data, and $f(\Theta|Z, M)$ the posterior distribution function, i.e. the multivariate probability density function of the variogram parameters given the sample data and the variogram model type.

We assumed (after a log-transform of the NO₃-N data of most fields,

see hereafter) a multivariate normal distribution for the data. For this distribution the likelihood function is (Gelman et al., 2013, p. 578)

$$f(Z|\Theta, M) = 2\pi^{-n/2}|\mathbf{C}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{Z} - \boldsymbol{\mu})^T \mathbf{C}^{-1}(\mathbf{Z} - \boldsymbol{\mu})\right\}, \quad (6)$$

with n the number of sampling locations used for estimating the variogram, $\boldsymbol{\mu}$ the vector with means, \mathbf{C} the $n \times n$ matrix with variances and covariances of the data, which is a function of the semivariogram parameters Θ , and $|\mathbf{C}|$ the determinant of the covariance matrix.

In the proposed sampling scheme the soil cores collected at the points of the stratified random sample are bulked into a composite sample. Not all soil cores are analyzed separately, only the single composite sample is analyzed. This implies that the contribution of the laboratory measurement error to the variance of the estimated mean is larger than when all soil aliquots would have been analyzed separately. To quantify the total variance of the mean estimated by the concentration of a composite sample, we need to quantify the contribution of the sampling error and the measurement error separately. To predict the variance of the sampling error we need to estimate the variogram parameters of errorless measurements of NO₃-N. The variance of the error due to analysis of the composite sample is then added to this pure sampling variance to obtain the total variance.

The variogram parameters of errorless measurements of NO₃-N are obtained by adding the measurement error variances to the diagonal of the variance-covariance matrix of errorless measurements:

$$\mathbf{C} = \mathbf{C}^* + \mathbf{E}, \quad (7)$$

with \mathbf{C}^* the variance-covariance matrix computed with the variogram of errorless data, and \mathbf{E} a diagonal matrix with the measurement error variances on the diagonal.

2.3. Markov chain Monte Carlo (MCMC) sampling

The posterior distribution function $f(\Theta|Z, M)$ is only fully determined in closed-form for specific combinations of the prior probability density function $f(\Theta|M)$ and the likelihood function $f(Z|\Theta, M)$. For these combinations the posterior and prior distributions are of the same family: the two distributions are called conjugate distributions, and the prior is called a conjugate prior for the likelihood function.

For priors that are not conjugate with the likelihood, the posterior distribution function can be approximated up to proportionality by Markov chain Monte Carlo (MCMC) sampling from the posterior distribution. Various algorithms are available for this, that differ in the way a “walk” is generated through the multivariate space spanned by the variogram parameters. Starting from a randomly chosen vector with initial variogram parameter values, we jump to a new “location”. The proposed vector with new variogram parameter values is accepted with a probability equal to:

$$Pr = \frac{f(Z|\Theta, M)f(\Theta|M)}{f(Z|\Theta_0, M)f(\Theta_0|M)}, \quad (8)$$

with Θ and Θ_0 the proposed and current vector with variogram parameters, respectively.

2.4. Decomposing the variance of the sampling variance

MCMC sampling from the posterior distribution of the variogram parameters results in a large number of vectors with possible variogram parameters. Each vector is used to simulate a large number of maps with possible NO₃-N values, see Section 3.3. So, there are two sources of randomness that cause variation in the sampling variance of the estimated mean:

- MCMC sampling from posterior distribution of variogram parameters
- Geostatistical simulation of maps with a given variogram

To get insight in the contribution of these two sources of randomness to our uncertainty about the sampling variance, the variance of the sampling variance can be decomposed as follows

$$V_{\text{total}}(V(\bar{z})) = V_{\text{MCMC}} \left[E_{\xi}(V(\bar{z})) \right] + E_{\text{MCMC}} \left[V_{\xi}(V(\bar{z})) \right]. \quad (9)$$

The first variance component is the contribution due to uncertainty about the variogram, the second variance component is the contribution due to uncertainty about the variances of z (NO3-N) within the geostrata given a variogram.

2.5. Expanded measurement uncertainty

The sampling variances of the estimated mean are calculated with a variogram for errorless measurements of NO3-N. The soil cores are bulked into a composite sample which is analyzed in a laboratory. The error variance in the measured NO3-N concentration of this composite sample is added to the sampling variance. We computed the expanded measurement uncertainty U defined as (ISO, 1995):

$$U = 100 \frac{2\sqrt{V_{\text{fld}} + V_{\text{lab}}}}{\bar{z}}. \quad (10)$$

with V_{fld} the sampling variance of the estimated mean for the proposed stratified random design, using errorless measurements of NO3-N, V_{lab} the variance of the laboratory measurement error, and \bar{z} the mean of the field. In ISO Guide 98 (ISO, 1995) the expanded measurement uncertainty is defined as “a quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand”. Used with the multiplier 2 (the coverage factor) the expanded measurement uncertainty encompasses approximately the 95% interval around the measured value in which the true values can be expected to lie.

3. Materials and methods

3.1. Data on spatial variation of NO3-N within a field

This study was limited to fields with crops not sown in rows such as grasses, or to those crops where the inter row distance is small such as with cereals. We expect that when crops are sown in rows and the inter row distance is large compared to the size of a single soil sample (here a 13 mm auger), there is an extra short-distance variance component as a result of both the fertiliser techniques and the extraction of NO3-N from the soil by the plants. This extra source of variance then should be accounted for in the sampling scheme, which will be object of further study.

Based on NO3-N residues after earlier crops and the use of fertiliser over the last year, we selected a set of fields that covers a large range of suspected NO3-N. In total sixteen fields were sampled post harvest in the fall of 2016 (Table 1). All fields were situated in Flanders in the region between the city of Ghent and the German border. A summary of some properties of the fields is given in Table 1. On each field a rectangular part with a surface of approximately one hectare was sampled using an approximately orthogonal grid consisting of 30 sampling points resulting in a distance of approximately 15 to 20 meter between adjacent points; For fields 30 and 35 the short-distance variance was assessed by sampling a subplot of 1 m² on a very fine grid, with cores spaced 10 cm apart, resulting in 100 points. Samples were collected with a hand auger up to a depth of 90 cm, GPS-coordinates were recorded for each sample location and each core was analysed separately.

Table 1
Properties of the sixteen sampled fields.

Field	Location	Soil texture	Crop
1	OLV Waver	Sand	Pasture
2	Leest	Sandy loam	Pasture
3	Putte	Sand	Pasture
4	Hingene	Sandy loam	Ryegrass
5	Leest	Sandy loam	Cereal
6	Oksdonk	Silt loam	Cereal
13	Hombeek	Silt loam	Pasture
14	Gontrode	Silt loam	Pasture
15	Melle	Sand	Pasture
16	Melle	Sand	Ryegrass
17	Melle	Sand	Ryegrass
18	Vilvoorde	Loam	Cereal
19	Millen	Loam	Cereal
20	Kampenhout	Silt loam	Cereal
30	Leest	Silt loam	Cereal
35	Retie	Sand	Ryegrass

3.2. Chemical analysis and analytical measurement error

Directly after sampling the samples were sealed in plastic bags to avoid overmuch contact with atmospheric oxygen and stored at 5 degrees Celsius pending analysis. NO3-N was determined in the field moist samples after manual subsampling and extraction with 1 M KCl following ISO 14256, the official method for the determination of NO3-N in Flanders. Results were calculated as kg NO3-N/ha using a standard soil density of 1500 kg/m³. In those few cases (five samples of field 35) where the NO3-N concentration was lower than the analytical level of quantification (LOQ) a higher bound approach was used, the concentrations were set equal to the LOQ of 0.5 kg N/ha.

Besides sampling variance we have variance due to laboratory measurement error. We used results from the laboratories quality control to estimate this variance component. Following the guidelines of ISO 17025 and the requirements for laboratories recognised by the Flemish government, laboratories have to analyse a reference sample with each run. By pooling the variances of in total 184 results for the same reference sample analysed by seven different labs we calculated the reproducibility standard deviation to be 6,4% of the measured value. Validation of the test by these labs show that this standard deviation is constant over the range of the reported results.

Since reference materials have controlled homogeneity, this variance component does not include a variance component induced by subsampling the bulked sample. However regular quality control performed by repeating the subsampling showed this variance to be negligible.

3.3. Methods

As a first step we checked the assumption that the data come from a normal distribution, or whether a lognormal distribution is more realistic. This is done by making Q-Q plots, and the Shapiro-Wilk test.

For the agricultural field with an assumed normal distribution of the NO3-N data (see Section 3.3.1) the following procedure is implemented:

1. discretize the field by a fine square grid of (about) 2000 nodes
2. stratify the field into $L = 5, 10, \dots, 50$ compact geostrata of equal size
3. sample 100 variograms from the multivariate posterior distribution of the variogram parameters by MCMC
4. simulate with the first sampled variogram 100 maps
5. compute for the first simulated map the variance within the five geostrata, and compute the sampling variance of the estimated mean (Eq. 2). Repeat this for $L = 10, \dots, 50$. This results in ten sampling variances (one per sample size)

6. repeat step 5 for the other 99 maps. After this step we have 100 sampling variances for sample size $n = 5$, 100 sampling variances for $n = 10$, et cetera
7. repeat steps 4–6 for the other sampled variograms. After this step we have 100×100 sampling variances for $n = 5$, 100 sampling variances for $n = 10$, et cetera
8. compute for each sample size the mean, P50 and P90 of the 100×100 sampling variances
9. compute the variance over the 100 sampled variograms of the average sampling variances (averaged over the 100 maps) for a given variogram. This is an estimate of the first variance component $V_{\text{MCMC}} [E_{\xi}(V(\hat{\mathbf{z}}))]$ (Eq. 2).
10. compute the average over the 100 sampled variograms of the variance of the sampling variance for a given variogram. This is an estimate of the second variance component $E_{\text{MCMC}} [V_{\xi}(V(\hat{\mathbf{z}}))]$ (Eq. 2).

For the fields with an assumed lognormal distribution, we need some extra steps. Before step 3 the natural logarithms of the NO3-N data are computed, so that the sampled variograms of step 3 are on the log scale, as well as the simulated maps of step 4. Before step 5 the estimated model mean (μ of Eq. 6, which is assumed constant throughout the field) is added to the simulated values. The sums are backtransformed to the original scale, before proceeding with the next steps.

We used R package spcosa (Walvoort et al., 2010) to compute the compact geostrata of equal size. Simulation of the maps with a given sampled variogram was done by Cholesky decomposition of the 2000×2000 matrix with variances and covariances computed with a given sampled variogram (Gelman et al., 2013, p. 582).

3.3.1. MCMC sampling of variogram parameters

Further, we assumed a constant mean, and an exponential variogram with nugget (Webster and Oliver, 2007), so that we have three variogram parameters: the nugget variance, the partial sill and the distance parameter. We used a bounded uniform prior for the inverse of the sill (nugget + partial sill), with a lower and upper bound of 10^{-6} and 1, respectively. A uniform prior for the inverse of the sill instead of for the sill itself is commonly used (Gelman et al., 2013, p. 52). For the nugget-to-sill ratio we used a bounded uniform prior with lower and upper bound of 0 and 1, respectively. Finally, for the distance parameter we used a bounded uniform prior with lower and upper bound of 10^{-6} and three times the maximum distance in the data set, respectively. We assumed that the variogram parameters are independent, so that the multivariate prior density can be computed as the product of the univariate prior densities.

The posterior distribution was sampled with a Differential-Evolution sampler (ter Braak and Vrugt, 2008). We used R package Bayesian Tools for this (Hartig, 2018).

4. Results

4.1. Analytical results

The overall analytical results for all fields are given in Table 2. The average NO3-N concentrations in the fields calculated as the arithmetic mean of all measurements on a field, varied between 2.56 kg N/ha, which is close to the reporting limit of the analytical method, and 314 kg N/ha. The latter value is about six times the amount that is considered ecologically safe vis à vis the danger of trespassing the maximum nitrate concentration limit of 50 mg NO3-N/l. To get an idea of the extent of the variability of NO3-N within a field, the variability was calculated as the coefficient of variation (CV) of all measurements made on a field. The standard deviation ranged between 26% and 256% of the sample averages. No significant correlation between the sample average and CV

Table 2

Average, coefficient of variation (CV) and skewness of sample data for the sixteen fields. Fields are sorted on the coefficient of variation.

Field	Sample average	CV	Skewness
6	252	0.26	0.48
5	64.6	0.42	0.72
1	10.7	0.49	0.27
20	314	0.50	0.90
17	11.1	0.54	1.20
4	4.29	0.71	1.34
13	8.26	0.80	0.93
30	41.9	0.83	2.33
15	102	0.94	1.57
35	2.56	0.97	3.73
16	13.2	1.08	2.94
18	31.9	1.11	3.00
19	13.9	1.20	4.06
2	35.0	1.23	2.40
14	200	2.03	4.43
3	29.2	2.56	4.06

was found with a Spearman's rank correlation test at the 5% significance level: a high average NO3-N concentration does not go hand in hand with a high CV.

4.2. Checking assumption of normal distribution

Based on the Q-Q plots and the Shapiro–Wilk test we assumed a lognormal distribution for twelve out of the sixteen fields (Table 3). For these fields the natural logarithm of the NO3-N data was used to estimate the variogram on the log-scale. Only for fields 1, 5, 6 and 20 we did not have enough evidence against a normal distribution; for these fields we used the untransformed NO3-N data. Note that these four fields are the fields with the smallest CV (Table 2). Fig. 1 shows the Q-Q plots for field 17 and 6, which are used in this paper to illustrate the results for the fields with an assumed lognormal and normal distribution, respectively.

4.3. Geographical stratification

Fig. 3 shows the compact geographical stratification for field 17 for $L = 10$ and 25.

4.4. MCMC sample of variogram parameters

The first twenty variograms sampled by MCMC for fields 17 and 6 are shown in Fig. 2, 3. Table 3 shows the means and standard deviations of the sampled marginal posterior distributions of the variogram

Table 3

Means and, between brackets, standard deviations of MCMC samples of parameters of an exponential variogram of (natural logs of) soil NO3-N for the sixteen fields.

Field	Distribution	Relative nugget	Sill	Distance parameter
6	Normal	0.67 (0.26)	4371 (2263)	164 (151)
5	Normal	0.48 (0.27)	804 (358)	90 (89)
1	Normal	0.62 (0.30)	33.0 (16.3)	108 (98)
20	Normal	0.84 (0.15)	26378 (11869)	205 (121)
17	Lognormal	0.147 (0.147)	0.361 (0.215)	22.3 (17.3)
4	Lognormal	0.426 (0.254)	0.722 (0.505)	114 (110)
13	Lognormal	0.584 (0.227)	1.13 (0.738)	160 (124)
30	Lognormal	0.081 (0.072)	0.672 (0.182)	0.515 (0.480)
15	Lognormal	0.777 (0.0755)	1.07 (0.427)	147 (103)
35	Lognormal	0.245 (0.039)	1.04 (0.269)	4.81 (3.30)
16	Lognormal	0.520 (0.235)	0.974 (0.755)	91.0 (79.1)
18	Lognormal	0.810 (0.168)	0.819 (0.284)	157 (109)
19	Lognormal	0.366 (0.212)	0.631 (0.343)	162 (107)
2	Lognormal	0.700 (0.235)	0.938 (0.448)	162 (123)
14	Lognormal	0.584 (0.253)	1.58 (0.858)	160 (106)
3	Lognormal	0.497 (0.225)	2.24 (1.16)	100 (75.7)

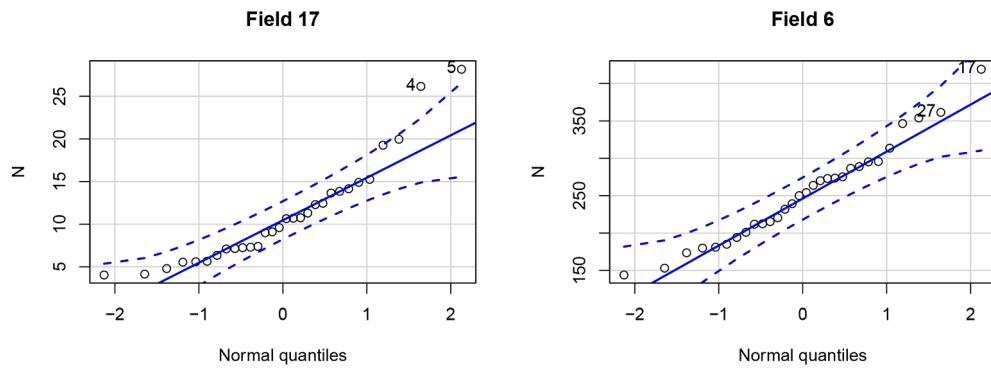


Fig. 1. Q-Q plot of NO₃-N in fields 17 and 6. The p-value of the Shapiro-Wilk test was 0.0028 for field 17 and 0.55 for field 6.

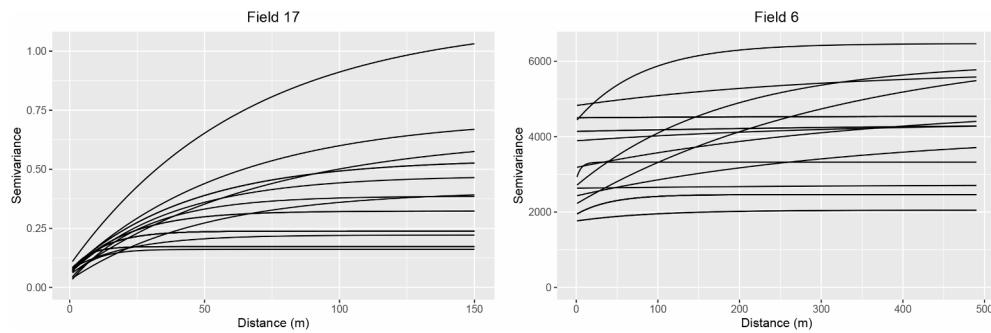


Fig. 2. The first twenty sampled variograms of fields 17 and 6.

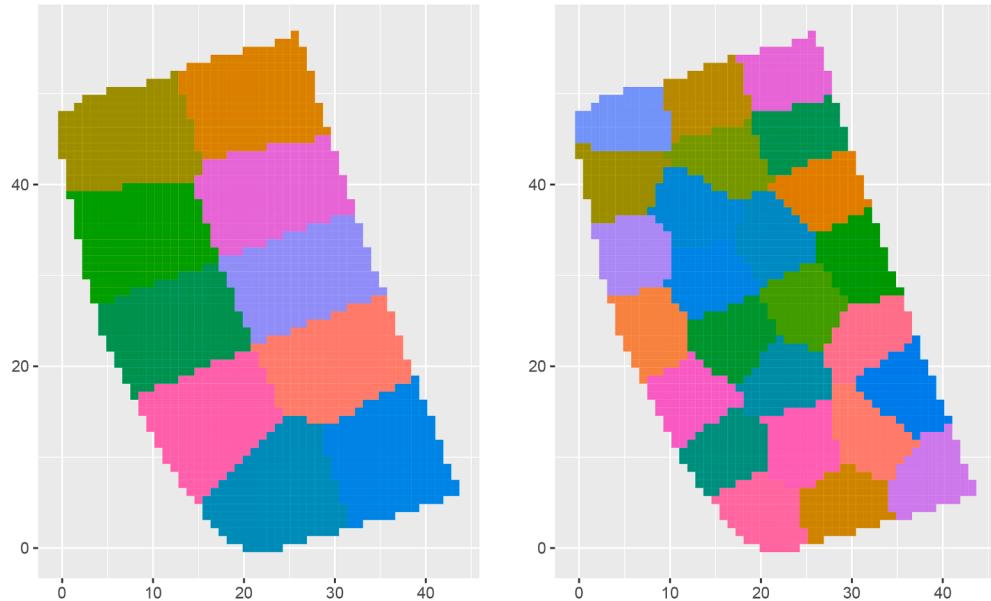


Fig. 3. Ten and twenty-five compact geographical strata for field 17. Sampling variances are predicted for stratified simple random sampling with one point per stratum.

parameters for all sixteen fields. The distance parameters of the variograms with the additional data collected in a subplot of 1 m² (fields 30 and 35) were considerably smaller than for the other fields. For all other fields the average distance parameter value ranged from 22 m (field 17) to 205 m (field 20). The relative nuggets of the variograms for field 30 were much smaller than of the other fields. For all other fields the mean values of the relative nugget were quite large, and ranged from 0.15 (field 17) to 0.84 (field 20). The mean values of the sill of the four variograms on the original scale ranged from 33 to 26,378, and of the

remaining variograms on the log-scale from 0.361 to 19.3. For all variogram parameters the standard deviations were large, showing that we are uncertain about all three variogram parameters for all fields, including fields 30 and 35 with the additional fine grid data.

4.5. Predicted sampling variance of estimated mean

Sampling variances are computed for stratified simple random sampling for sample sizes 5, 10, ..., 50, with one point per stratum.

As an illustration Fig. 4 shows twenty simulated maps of NO₃-N for field 17, simulated with the first four sampled variograms (rows 1 through 4), and using each variogram to simulate five maps (columns a through e). The variance of the simulated values within the $L = 5, 10, \dots, 50$ strata is computed, and these stratum variances are used to compute the ten sampling variances of the estimated mean for that simulated map. For each value of L , the rowwise averages of the sampling variances are computed (in the experiment not the average of five variances, but of 1000 variances), as well as the rowwise variances. The variance of the rowwise averages is an estimate of the first variance component of Eq. 9 quantifying the uncertainty in the sampling variance of the estimated mean due to uncertainty about the variogram parameters, and the average of the rowwise variances is an estimate of the second variance component, quantifying the uncertainty in the sampling variance of the estimated mean due to uncertainty in the within-stratum variances of NO₃-N values simulated with a given variogram.

Fig. 5 shows for fields 17 and 6 the mean, P50 and P90 of the uncertainty distribution of the sampling variance as a function of the sample size. As expected, the sampling variance decreases rapidly with the sample size. For the fields with an assumed lognormal distribution the mean was (much) larger than the median, showing the strong positive skew of the uncertainty distribution of the sampling variance. For five fields the mean was even larger than the P90 of the uncertainty distribution (fields 2, 3, 13, 14 and 15). For the four fields with an assumed normal distribution the mean and median were about equal. we

believe that with a strong positive skew of the uncertainty distribution the median sampling variance is a more sensible and practical prediction of the sampling variance of the estimated mean than the mean of this uncertainty distribution. Therefore, hereafter the median is used to derive the required sample size, as well as the P90 of the uncertainty distribution. Especially for the fields with an assumed lognormal distribution the P90 was substantially larger than the median.

4.6. Variance contributions

Fig. 6, shows for fields 17 and 6 the contribution of uncertainty about the variogram parameters and that of uncertainty about the variance of NO₃-N within strata given a variogram, to the total uncertainty about the sampling variance of the estimated mean. For field 6 with an assumed normal distribution the contribution of variogram parameter uncertainty dominates, whereas reversely for field 17 with an assumed lognormal distribution the contribution of the uncertainty about the stratum variances is dominant.

4.7. Expanded measurement uncertainty

Fig. 7 shows the P50 and P90 of the expanded measurement uncertainty U as a function of the sample size for fields 17 and 6. U decreases with the sample size. For a given sample size U largely differs among the fields. Table 4 shows U for a sample size of 15, which is the current

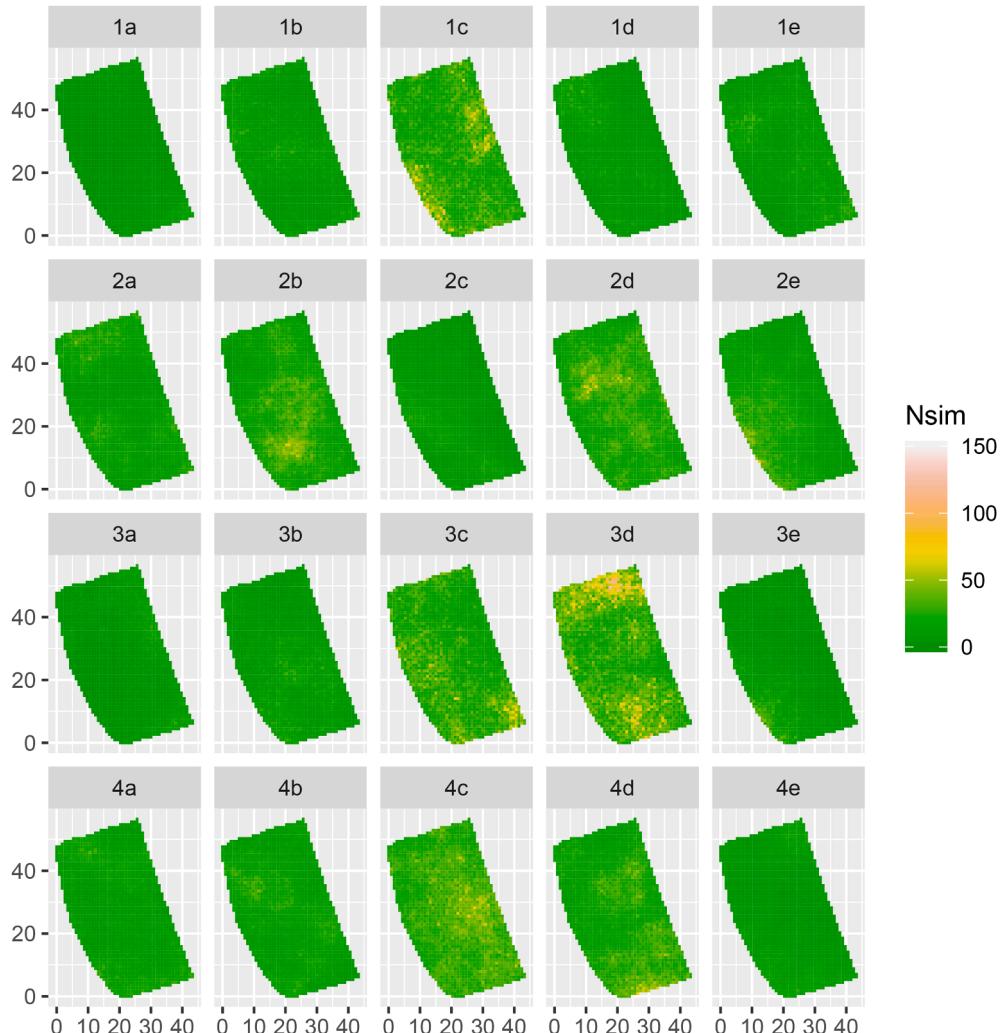


Fig. 4. Twenty simulated maps of NO₃-N values (Nsim) for agricultural field 17, using first four sampled variograms (rows 1 through 4). With each variogram five simulated maps are shown (columns a through e).

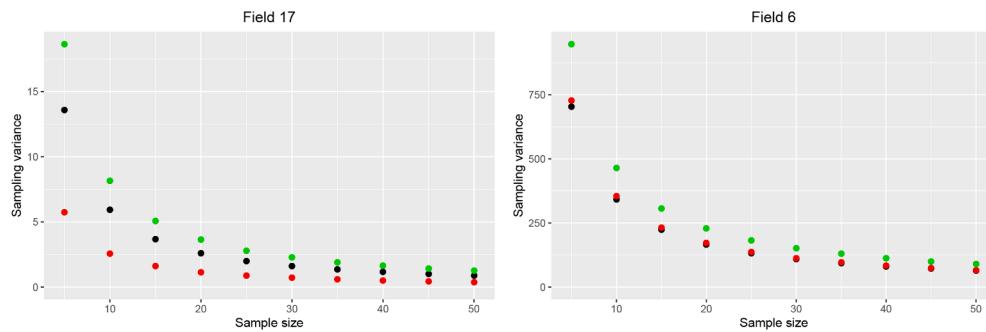


Fig. 5. Mean (black dots), P50 (red dots) and P90 (green dots) of the uncertainty distribution of the sampling variance of estimated mean for fields 17 and 6. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

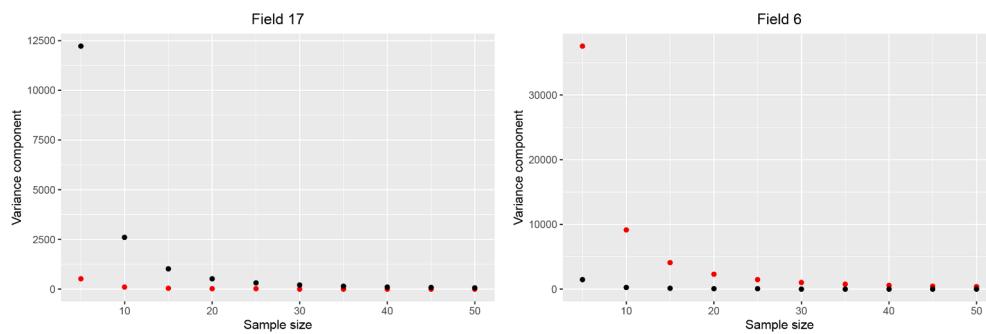


Fig. 6. Contribution of uncertainty about the variogram parameters (red dots) and of uncertainty about the variance of NO3-N within strata given a variogram (black dots), to the total uncertainty about the sampling variance of the estimated mean, for fields 17 and 6. For field 17 a lognormal distribution is assumed, for field 6 a normal distribution is assumed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

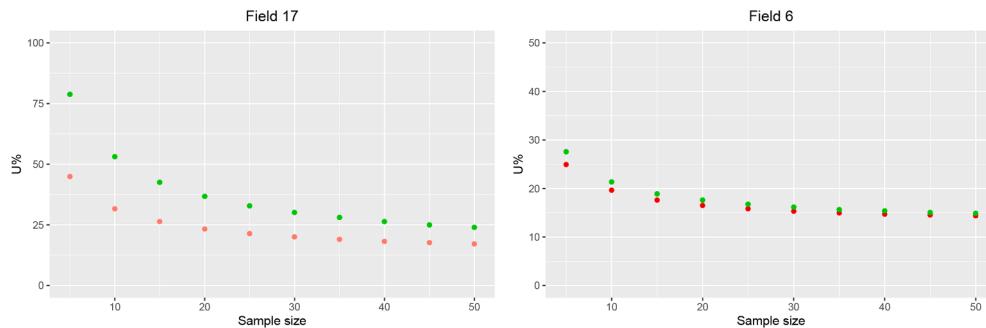


Fig. 7. P50 (red dots) and P90 (green dots) of expanded measurement uncertainty. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sample size in Flanders. With this sample size U is larger than 50% for six fields at P50 and for 11 fields at P90.

4.8. Required sample size

Table 5 shows the sample sizes based on the P50 and P90 of the uncertainty distribution of the sampling variance required for a maximum of the expanded measurement uncertainty of 50%. Based on the median (P50) of the empirical uncertainty distribution of U , for five agricultural fields less than five sampling points are needed to achieve an expanded measurement uncertainty of 50%, and for one field we need more than 50 points. If we use P90, somewhat larger sample sizes are required for a maximum U of 50%.

4.9. Stratification effect

We also computed the sampling variance of the estimated mean for

simple random sampling, i.e. no stratification. This is straightforward: in step 4 of the procedure described in Section 3.3 the variance of the simulated values within the entire field is computed, instead of the variances within the strata. Eq. 2 still can be used to compute the sampling variance, by setting L equal to 1. By comparing the sampling variance of the estimated means with stratified simple random sampling and simple random sampling at an equal number of sampling points we get insight in the stratification effect. For most fields the stratification effect, quantified by the ratio of the variance with simple random sampling to the variance with stratified simple random sampling at an equal sample size, was small but not irrelevant (Fig. 8). This was expected, given the limited spatial structure of NO3-N (large relative nugget, small range) for most fields. It increases with the sample size. For fields 17, 19 and 4, the top three curves in Fig. 8, the stratification effect was considerable. For these fields the gain in precision increased to 1.79, 1.48 and 1.37, respectively. NO3-N on these fields showed relatively strong spatial structure, a relatively small relative nugget in

Table 4

Expanded measurement uncertainty for stratified simple random sampling of 15 points (1 point per geostratum).

Field	P50	P90
6	17.6	18.9
5	21.5	24.6
1	26.4	30.2
20	26.7	34.4
17	26.3	42.6
4	41.3	78.2
13	60.1	139
30	75.4	131
15	62.4	135
16	41.0	66.1
18	47.8	83.3
35	118	175
19	29.3	54.3
2	42.9	91.8
14	62.8	133
3	53.7	202

Table 5

Sample sizes required for a maximum expanded measurement uncertainty of 50%, based on the P50 and P90 of the uncertainty distribution of U .

Field	P50	P90
6	< 5	< 5
5	< 5	< 5
1	< 5	6
20	< 5	8
17	< 5	12
4	11	36
13	23	> 50
30	36	> 50
15	24	> 50
16	10	27
18	14	44
35	> 50	> 50
19	6	18
2	12	> 50
14	25	> 50
3	18	> 50

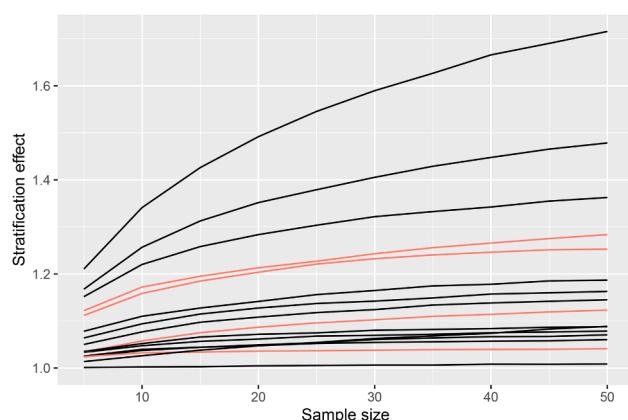


Fig. 8. Stratification effect. Red lines are for the fields with assumed normal distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

combination with a large effective range (Table 3). At a sample size of 25, for nine of the fields the stratification effect was 1.10 or larger, implying that at least 10% less samples are needed with stratified sampling to obtain the same precision as with simple random sampling.

5. Discussion

Both Fig. 2 and the standard deviations of the MCMC samples in Table 3 clearly show that we were quite uncertain about the variogram parameters. For the fields with a normal distribution, uncertainty about the variogram parameters contributed stronger to the total uncertainty about the variance of the estimated mean, than uncertainty about the spatial variance within strata given a variogram. So for these fields narrowing down our uncertainty about the variogram by collecting more data with a well-designed sample, would help to narrow down our uncertainty about the sampling variance of the estimated mean. For reliable estimation of the variogram using the method-of-moments 150–225 point observations are needed (Webster and Oliver, 1992); with maximum likelihood estimation of the variogram fewer observations are required (Lark, 2000). However, for the fields with a lognormal distribution the contribution of the uncertainty about the variance within strata given a variogram was much stronger. So even when we would have perfect knowledge of the variogram, for fields with a lognormal distribution considerable uncertainty remains about the sampling variance of the estimated mean.

Various modelling assumptions are made to compute the required sample sizes. We assumed a constant mean, μ in Eq. 6. In cases where NO3-N shows a spatial trend or is related to a covariate of which a map is available, a model with a spatially varying mean can be more realistic. We expect that the predicted sampling variance using a non-stationary model will not differ much from the variance obtained with a model with a constant mean: the variance of the simulated NO3-N values within the strata will not differ much.

Secondly, we either assumed a normal or a lognormal distribution for the data. For field 20 the Q-Q plot showed some deviation of the normal distribution, but the p-value of the Shapiro–Wilk test was only marginally smaller than 0.05, so we decided to assume a normal distribution for this field. We also computed the required sample size with a lognormal distribution. The required sample sizes were equal to those with the normal distribution. However, for other fields there were large differences in required sample sizes for the two distributions. For instance, for field 4 the required sample sizes computed with a normal distribution were 6 and 9 based on the P50 and P90 of the uncertainty distribution, respectively, which were both substantially smaller than with the lognormal distribution, being 11 and 36, respectively. This shows how important it is to check carefully the type of distribution.

The proposed simulation approach for predicting the variance of the estimated mean is versatile. It can also be used for other random sampling designs, such as systematic random sampling, cluster random sampling, two-stage random sampling et cetera. For stratified simple random sampling the sampling variance of the estimated mean for a given simulated map can be computed from the variances of the simulated values within the strata (Eq. 1). However, for some other sampling designs the sampling variance must be approximated by simulating a large number of samples with the sampling design under study, and using each sample to estimate the mean. The simulation approach is also versatile with respect to the distribution. Other transformations can be tried, for instance the square root transformation, or the normal score transform (Goovaerts, 1997).

To predict the sampling variance for the proposed sampling design prior knowledge of the stratum variances is required. As shown in this paper these stratum variances can be derived from an a priori variogram. In practice, we often do not have data from which we can estimate the variogram of the field under study. For the sixteen fields sampled in this study the variograms were largely different, so using an average variogram to predict the sampling variance of the estimated mean is not a

good option. For most fields, most MCMC variograms showed large relative nuggets and/or small ranges, indicating weak spatial structure. The gain in precision due to the stratification was for most fields rather small (Fig. 8). Therefore, to predict the variance for unsampled fields we propose to assume that the NO₃-N concentrations shows no spatial structure, i.e. to postulate a pure nugget variogram. We then require a prior estimate of the variance within the entire field (nugget variance) only. The sampling variance of the estimated mean can then be predicted by dividing this prior estimate of the within-field variance through the sample size. If a lognormal distribution is assumed, a prior estimate of the variance on the log scale as well as of the mean is required. The variance on the original scale can then be approximated by the variance on the log scale multiplied by the square of the mean. This results in a conservative (safe) estimate of the required sample size: if in reality there is spatial structure, the required sample size will be smaller.

The question is how to obtain this prior estimate of the within-field variance. In this study we found that the coefficient of variation largely differed among the sixteen fields (Table 2) so prior knowledge of the mean NO₃-N concentration of a field is of little use for guessing the standard deviation within that field. How to obtain an estimate of the standard deviation (variance) of NO₃-N within unsampled fields will be subject of further study. Even in case no usable correlation can be found between the standard deviation within a field and field properties such as cropping history, soil texture or fertilisation practice, knowledge of the variation among agricultural fields of the within-field standard deviation of NO₃-N in fall can be used to derive a worst-case sampling design.

6. Conclusions

The proposed methodology for predicting the sampling variance of the mean estimated by stratified simple random sample, using an estimated variogram is a versatile approach, and takes into account uncertainty about the variogram parameters, as well as uncertainty about the spatial variance of NO₃-N within strata for a given variogram. For most fields a lognormal distribution of NO₃-N was more realistic than a normal distribution. For the fields with an assumed lognormal distribution, the empirical uncertainty distribution of the sampling variance showed very strong positive skew. For these fields the median sampling variance is a more sensible and practical prediction of the sampling variance of the estimated mean than the mean of this uncertainty distribution.

Analytical results for the sixteen fields showed very large differences in the mean NO₃-N concentration, their within-field variance, and in the coefficient of variation. As a result there are large differences in the sample sizes needed to be assured of a maximum allowable expanded measurement uncertainty. If a 50% expanded measurement uncertainty is acceptable, based on the median of the uncertainty distribution, the required sample size varied from less than five to more than 50. For six fields using a sample size of fifteen borings per field, which is common practice in Flanders, results in an expanded measurement uncertainty of more than 50%.

The gain in precision due to the geographical stratification was rather limited for most fields due to the generally weak spatial structure.

Supplementary material

The R script and the data can be downloaded from https://github.com/DickBrus/GeodermaPaper2021_HofmanBrus.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Acknowledgements

This research programme was made possible through funding by the Vlaamse Landmaatschappij (VLM)

References

Sächsische Landesanstalt für Landwirtschaft, 2012. Umsetzung der düngeverordnung - hinweise und richtwerte für die praxis, URL:<https://publikationen.sachsen.de/bdb/artikel/15242/documents/18421>.

Bogaert, N., Salomez, J., Vermoesen, A., Hofman, G., Cleemput, O.V., Meirvenne, M.V., 2000. Within-field variability of mineral nitrogen in grassland. *Biol. Fertil. Soils* 32, 186–193.

ter Braak, C.J., Vrugt, A.J., 2008. Differential evolution markov chain with snooker updaters and fewer chains. *Stat. Comput.* 18, 435–446.

Brus, D.J., Späthens, L.E., de Gruijter, J.J., 1999. A sampling scheme for estimating the mean extractable phosphorus concentration of fields for environmental regulation. *Geoderma* 89, 129–148.

Clay, D., Carlson, C., Brix-Davis, K., Berg, B., J.O., 1997. Soil sampling strategies for estimating residual nitrogen. *J. Prod. Agric.* 10 no. 3, 446–452.

Domburg, P., de Gruijter, J.J., Brus, D.J., 1994. A structured approach to designing soil survey schemes with prediction of sampling error from variograms. *Geoderma* 62, 151–164.

EU, 1991. Council directive of 12 december 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources.

Franzen, D., Long, D., Sims, A., Lamb, J., Casey, F., Staricka, J., Hofman, V., 2011. Evaluation of methods to determine residual soil nitrate zones across the northern great plains of the USA. *Precis. Agric.* 12, 594–606.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. *Bayesian Data Analysis*, third ed. CRC Press.

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

de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Knotter, M., 2006. *Sampling for Natural Resource Monitoring*. Springer.

Hartig, F., 2018. Bayesian Tools. URL:<https://cran.r-project.org/web/packages/BayesianTools/vignettes/BayesianTools.html>.

Hennart, S., Stilmant, D., Vandenberghe, C., 2004. Accuracy of the nitrate sampling scheme, proposed in wallon area, for grazed grasslands. Proceedings of the 14th N Workshop. Maastricht, the Netherlands, pp. 300–302.

ISO/IEC, 1995. Iso/iec guide 98-1:2009 – uncertainty of measurement – part 1: Introduction to the expression of uncertainty in measurement.

ISSEP, 2014. Echantillonage des sols agricoles. URL:<https://www.issep.be/wp-content/uploads/11-P-11-Echantillonage-sols-agricoles.pdf>.

James, D., Richards, A.W., Weaver, W.H., Reeder, R.L., 1971. Residual soil nitrate measurement as a basis for managing nitrogen fertilizer practices for sugarbeets. *J. A.S.S.B.T.* 16 (4), 313–322.

Kathik, V.A., Sarode, D.B., Jadhav, R.N., Ingle, S.T., Attarde, S.B., 2011. Assessment on residual soil nitrate of intensively fertilized banana farms of Jalgaon Region. *Malay. J. Soil Sci.* 15, 87–99.

Lark, R., 2000. Estimating variograms of soil properties by the method-of-moments and maximum likelihood. *Eur. J. Soil Sci.* 51, 717–728.

Nie, P., Tao, D., Yong, H., Fangfang, Q., 2017. Detection of soil nitrogen using near infrared sensors based on soil pretreatment and algorithms. *Sensors* 17 (1102). URL: <https://www.mdpi.com/1424-8220/17/5/1102>.

Parmodh, S., Manoj, K.S., Mexal, J.G., 2011. Spatial variability of soil properties in agricultural fields of southern new Mexico. *Soil Sci.* 176 (2), 288–302.

Shaw, R., Lark, R., Williams, A., Jones, D.L., D.C., 2016. Characterising the within-field scale spatial variation of nitrogen in a grassland soil to inform the efficient design of in-situ nitrogen sensor networks for precision agriculture. *Agric., Ecosyst. Environ.* 230, 294–306.

Starr, J.L., Parkin, T.B., Meisinger, J.J., 1992. Sample size consideration in the determination of soil nitrate. *Soil Sci. Soc. Am. J.* 56 (6), 1824–1830.

Stenger, R., Priesack, E., Beese, F., 2002. Spatial variation of nitrate-nitrogen and related soil properties at the plot scale. *Geoderma* 105, 259–275.

UGent, KUL, 2000. N-(eco)2: Bepaling van de hoeveelheid minerale stikstof in de bodem als beleidsinstrument.

United Nations, 2017. *The Future of Food and Agriculture – Trends and Challenges*. FAO, Rome.

VITO, 2010. Bam deel 1.01. URL:<https://emis.vito.be/nl/erkende-laboratoria/mestbodem-en-veevoeder-vlm/compendium-bam>.

Walvoort, D.J.J., Brus, D.J., de Gruijter, J.J., 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. *Comput. Geosci.* 36, 1261–1267. <https://doi.org/10.1016/j.cageo.2010.04.005>.

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

Webster, R., Oliver, M.A., 2007. *Geostatistics for Environmental Scientists*, second ed. Wiley, Chichester.

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