April 2012

Cost £8.75

Project Report No. 485

Cost-effective sampling strategies for soil management

by

B. P. Marchant, A. G. Dailey and R. M. Lark

Rothamsted Research, Harpenden, AL5 2JQ

This is the final report of a 24 month project (RD-2005-3189) which started in April 2009. The work was funded by a contract for £113,744 from HGCA.

HGCA is the cereals and oilseeds division of the Agriculture and Horticulture Development Board.

CONTENTS

1. ABSTRACT

Efficient fertiliser management requires information about the nutrient status of each management area or field. This information can be gathered by observing soil nutrients at a number of sites in the field. The quality of this information is dependent on the sampling strategy that is employed. The sampling strategies suggested in fertiliser recommendations such as RB209 are generally based on anecdotal evidence regarding the number of soil cores required or are designed to ensure that the errors in estimating soil nutrient concentrations are less than an arbitrarily defined threshold. Such strategies do not directly link the sampling effort to the consequences of erroneous soil nutrient information, which may include reduced profitability or the long term development of nutrient excess or deficiency. We develop a quantitative framework to study the effectiveness of different sampling designs so that rational sampling recommendations for phosphorus (P), potassium (K) and nitrogen (N) can be developed.

For all nutrients, current recommendations suggest that measurements should be regularly spaced on a 'W' design which covers the field. Four alternatives to the 'W' are tested: an optimized sample configuration, stratified random sampling, rank set sampling and a clustered or bad practice design. We quantify the errors associated with each design, determine the management decisions that will be made by the farmer based on this erroneous information and then model the effects of these decisions. Thus we are able to relate the resources devoted to sampling to the expected profitability or long-term nutrient status of the field.

Our study shows, that for a particular sampling effort, sample designs can be optimized to give smaller errors than the 'W' design. However we also find that the errors from estimating soil-nutrient status with a 'W' are not large enough to substantially affect the quality of soil nutrient management. This is because once a certain accuracy in estimating soil-nutrient concentrations has been achieved, the quality of the management recommendations are limited by other sources of uncertainty in predicting the amount of nutrients the crops will access from the soil. Therefore the benefits of using optimized designs do not outweigh the extra complexity which they entail. If in the future fertiliser recommendations are more sensitive to soil information, say for example if nitrous oxide emissions had to be carefully controlled, then the use of optimized sample designs should be re-explored.

We find that in the scenarios explored in this project, decisions regarding K require less accurate information than P. A bulked sample every four years of 10 soil cores is sufficient to maintain both soil P and K stocks within a target range. This is less than half of the number of cores which is currently recommended. For N, rational sampling effort varies according to the expected SNS in the field, and the field size. Bulked samples of 10-15 soil cores are adequate for most fields. Including more than 10 cores in the bulked sample is warranted when fields are larger than 20 ha or if SNS is expected to be high (>160 kg/ha). The largest financial benefit from sampling occurs when soil nitrogen supply is around 175 kg/ha since at these concentrations the yield is most sensitive to sampling errors and erroneous decisions. There is a smaller benefit when the expected SMN is much larger or much smaller since in these circumstances it is clear that either a small or large amount of N fertiliser should be added. We determine the circumstances in which SMN measurement and the use of barometer fields are cost effective in comparison to prior knowledge of SMN. We do not consider alternative methods of estimating soil nitrogen supply such as the Field Assessment Method.

The framework developed in this project is not only suitable to assess the cost-effectiveness of different sampling designs under current fertiliser recommendations but also to develop and assess the cost-effectiveness of modifications to these recommendations.

2. SUMMARY

2.1. Introduction

Efficient fertiliser management requires information about the soil-nutrient status within each management area or field prior to fertiliser additions. The soil-nutrient status can be estimated by bulking a number of soil cores from different sites within a field. The accuracy of such an estimate increases with the number of cores extracted but so do the time taken and the costs of sampling. Current recommendations for soil sampling are based primarily on anecdotal evidence of what sampling is sufficient. They do not relate the sampling effort to the consequences of erroneous soil nutrient information. These consequences might include reduced profitability or the long term development of nutrient excess or deficiency.

RB209 suggests that for phosphorus (P) and potassium (K) a bulked sample of 25 cores will be adequate for a uniform area and that a 'W' design will ensure an even distribution over the whole field. For soil mineral nitrogen (SMN), RB209 suggests a minimum of 10 individual sub-samples should be taken from the sampled area and more if practically feasible. The HGCA *Nitrogen for winter wheat – management guidelines* suggest at least 15 cores in each field where SMN analysis is undertaken and 20 in more variable fields. Both RB209 and the HGCA guidelines emphasize that costs of SMN analysis will prohibit sampling in every field and that sampling is most worthwhile where large and uncertain soil nitrogen residues are expected.

We showed in a previous HGCA project that it is possible to optimize the sample designs to perform better than the 'W' in terms of sampling errors per core taken. However such schemes might not be as simple to implement in the field. These factors suggest there is a need to assess the sampling requirements for soil nutrient management so that the best sampling design and the rational sampling effort (i.e. number of cores) can be determined.

2.2. A framework for assessing sampling requirements for nutrient management

Field experiments to compare the relative merits of different sample designs would require an impractical number of soil cores to be extracted and analysed. Therefore a computer-based framework to compare sample designs is developed in this study. Typical patterns of the variation of soil nutrients are simulated based on existing datasets. These simulations are sampled and the mean soil nutrient requirement is estimated according to the various proposed sample designs. The implications of basing fertiliser management on these estimated values are explored and compared using mathematical models of soil nutrient management decisions, soil nutrient dynamics and crop responses to nutrients.

2.3. Comparison of different sample designs for estimating field-mean soil nutrient status.

The simulation tests confirmed that optimized sample designs do estimate the field-mean P, K and SMN concentration more efficiently than the 'W' design. However the improvement in efficiency is only substantial when more than 15 cores are bulked. It was found that the accuracy of the 'W' design is sufficient for current fertiliser recommendations. However if future fertiliser recommendations require more precise fertiliser management, such as if nutrient leaching or nitrous oxide emissions have to be more carefully controlled, then the use of optimized sample designs should be reconsidered.

2.4. Sampling recommendations for soil nutrient status

The sampling tests suggest that 10 bulked cores selected from a 'W' design every four years are sufficient to maintain soil P and K concentrations within the ideal RB209 index. The number of cores required is largely independent of field size and is less than half of the number of cores currently recommended.

For SMN, 10-15 cores are adequate for most fields. More than 10 cores are warranted for large fields (>20 ha) where SNS is expected to be high (>160 kg/ha). Sampling requirements increase with field size and 23 cores would be most cost-effective for a highly variable 60 ha field. We do that envisage that such fields are common in the UK. The study suggests that for some small fields where SNS is likely to be small or moderate (below 120 kg/ha) a bulked sample consisting of fewer than 10 cores is most cost-effective. However in such circumstances it is likely that the field assessment method would be used in preference to direct measurements. Therefore where SMN is measured we recommend that a bulked sample should be formed from 10-15 cores which are located on a 'W'. The use of barometer fields is seen to be efficient for small target areas with low variability such as clusters of small fields where N requirements are not expected to differ. More widespread sampling should be considered in other circumstances.

The scenarios and models for P and K in this project were relatively robust. However the scenarios for SMN required assumptions about the price of wheat and fertiliser and the quality of the practitioner's prior knowledge of nutrient stocks. If these assumed values change or if more information of the complex dynamics of SMN become available then the sampling recommendations should be recalculated. Alternative methods of estimating soil nitrogen supply such as the Field Assessment Method are not considered here but are addressed in *HGCA Project 3425 Establishing best practice for SNS estimation*.

2.5. Delineation of regions of soil-nutrient excess or deficiency within fields

Previous studies have demonstrated the potential for regions of soil-nutrient excess or deficiency to appear in fields under uniform management. These might arise because of variation in yield and hence nutrient off take. There is a need to identify these regions so that limitations on yield and excessive emissions to the environment can be avoided. In surveys of soil pollution, sequential sweep-out methods are used to efficiently identify pollution hotspots. These methods adopt an efficient bulking strategy to reduce the laboratory costs of conducting the survey. However we found them to be ineffective for soil-nutrient surveys because the variation of the nutrients is much less extreme than the variation of contaminants observed close to sources of pollution.

2.6. Future application of the computer-based framework

The framework used to assess sample designs in this framework could be extended to compare and evaluate future fertiliser management recommendations. This would require extensive analyses of existing soil-nutrient data and perhaps further field experiments to determine the relationships between soil nutrient status, fertiliser additions and crop yields under various soil and climate conditions. Such issues have been addressed in HGCA Project 3425 *Establishing best practice for SNS estimation*. Critically the uncertainty associated with these relationships should be quantified. It is also important to understand how the relationships can be scaled from experimental plots up to the field or farm scale. The framework developed in this project would then be suitable to process this information and rationally determine the most cost-effective fertiliser management procedures. It would also be possible to devise recommendations which efficiently combine SMN observations with the Field Assessment Method or to develop decision support systems for individual farms which, over a number of seasons, adapt recommendations to the specific needs of the farm.

3. TECHNICAL DETAIL

3.1. Introduction

Efficient fertiliser management requires accurate knowledge of the soil nutrient status and the potential to supply nutrient to the crop. At the field-scale this knowledge is commonly acquired by bulking a number of soil cores extracted from representative sites in the management area or field and then using laboratory analyses to estimate the mean nutrient content of the bulked sample and hence the field. The degree to which the bulked sample is representative of the field as a whole depends on the sample design used to select the position of the cores and the number of cores extracted. In general the more cores that are included in the bulked sample the more accurate is the estimate of the nutrient status within the field but this increased accuracy must be balanced against the labour costs associated with extracting extra cores. Recommendations for UK farmers suggest that cores should be extracted from a 'W' design since this leads to a fairly even spatial coverage of the field and hence a reasonably representative bulked sample. However it is known that more efficient sampling designs are available (Marchant *et al*., 2005).

The recommendations for soil nutrient sampling in RB209 (DEFRA, 2010) and other guidelines are largely based on anecdotal evidence of what sampling is sufficient. RB209 suggests that for P and K a sample of 25 individual sub-samples will be adequate for a uniform area and that a 'W' design will ensure an even distribution over the whole field. For soil mineral nitrogen (SMN), RB209 suggests a minimum of 10 individual sub-samples should be taken from the sampled area and more if practically feasible. The HGCA *Nitrogen for winter wheat – management guidelines* suggest at least 15 cores in each field where SMN analysis is undertaken and 20 in more variable fields. If it is thought that the farm can be divided into zones of relatively uniform SMN then these zones should be analysed separately. Both RB209 and the HGCA guidelines emphasize that costs will prohibit SMN analyses in every field and that they are most worthwhile where large and uncertain soil nitrogen residues are expected.

Few studies have attempted to rationally determine the effort which should be devoted to soil nutrient sampling and the design which should be used. One exception was a study by Oliver *et al*. (1997) who considered the bulking required for within-field mapping of soil phosphorus (P) and potassium (K) rather than the field averages considered here. They suggested that such surveys should be conducted on a grid and each observation should consist of 16 bulked cores extracted from within an area of 5 m^2 about the grid node. This recommendation was based on two surveys of around a hundred P and K observations. Sixteen cores were suggested because on the two fields they ensured that the sampling errors in estimating the mean nutrient content of the 5 m^2 were less than 1 and 7 mg/l for P and K respectively. This arbitrary criterion does not relate the sampling effort to the implications of sampling errors such as loss of soil nutrient status and decrease in profits.

In this study we form a framework to determine the rational sampling effort that directly relates this effort to the maintenance of soil nutrient stocks and profitability. Using this framework we can compare the effectiveness of different sampling designs and, if suitable information were available, we could extend it to test the cost-effectiveness of the actual nutrient management recommendations. The framework uses realistic simulations of the variation of soil nutrients in fields which are based upon observed soil nutrient data. It is possible to generate thousands of these simulations and hence to form reliable averages of the results.

The situation we consider is where a practitioner makes uniform fertiliser applications across a target area (which we nominally refer to as a field) and in which they determine the amount of fertiliser to apply from measurements of soil nutrient status and the RB209 guidelines. In our framework the only source of uncertainty is from sampling errors. This means that we assume that if perfect soil nutrient information is available the RB209 recommendations will ensure that the soil nutrient status and crop yields are exactly as required. We estimate the expected evolution of soil nutrient status, crop yields and profit when the management strategies are based on imperfect estimates from bulked samples. In our framework, soil nutrient status is simulated across each field, different sample designs are used to estimate the average status and the errors associated with these estimates are calculated. Mathematical models are used to calculate the effects of these errors. Through conducting an ensemble of such simulation tests in different circumstances we generate rational recommendations for the soil sampling which should be conducted for soil nutrient management.

Throughout the project, guidance was provided by the steering group consisting of Simon Griffin (SOYL), James Holmes (HGCA), Stuart Knight (NIAB TAG) and Peter Taylor (AICC Independent Agronomist, Peter Taylor Agronomy Ltd).

3.2. Materials and methods

3.2.1. Overview

The computer-based framework developed in this project was designed to test the effectiveness of different sampling approaches for soil nutrients. The main components of this framework were (i) the simulation of realistic soil nutrient variation within exemplar fields (ii) the sampling of these simulations by different approaches and quantification of the expected errors (iii) the use of mathematical models to determine the implications of these errors in terms of the evolution of soil nutrient stocks and of fertiliser cost and yield, and hence profitability. Once this framework was constructed a number of simulation tests were conducted to compare the performance of the different sampling approaches in different circumstances.

3.2.2. Simulation of soil nutrient variation within fields

The comparisons of different sample schemes within this study required thousands of realistic simulations of soil nutrient variation within fields. These were generated by geostatistical methods and based upon spatial models fitted to existing datasets. The steps in this process were:

1. Compilation of a database of within-field variation of soil nutrients
2. Geostatistical estimation of models describing the variation of the soil nutrients
3. Selection of exemplar fields
4. Stochastic simulation of soil nutrients concentrations at all grid points within exemplar fields

Compilation of soil nutrient datasets

The simulations used in this study were based upon statistical models fitted to existing datasets. This required geo-referenced within-field data rather than the field scale averages which would normally be used to manage field-scale fertiliser applications. The simulations should represent variation at the scale of a soil core and therefore ideally each observation within the datasets would have been from individual non-bulked soil cores. This is generally the case for observations of SMN for precision agriculture research and applications because of the effort involved in collecting multiple 90 cm cores to be bulked. For P and K soil cores are generally bulked prior to laboratory analysis because management decisions are made at scales much larger than a single core and the variation between different cores collected over a small area is of little relevance for management. Therefore our models of P and K variation were initially fitted to measurements from bulked cores. Then this variation was downscaled to the core scale using the results of a study by Oliver *et al*. (1997). Within this study single cores were collected from two fields in order to determine the amount of bulking required to construct within-field maps of nutrient variation.

The datasets for P and K were provided by SOYL (SOYL, a division of Frontier Agriculture Ltd) and were all collected since 2002. SOYL generally extract samples on a regular square grid of length 100m but their data include occasional additional observations at closer intervals. Each sample consists of 16 bulked cores extracted from an area of 5 m^2. As recommended in RB209, P content (mg/l) is measured as Olsen's P and K content (mg/l) by ammonium nitrate extraction following the procedures described in *Specification for Topsoil (British Standard 3882)* or *The Analysis of Agricultural Materials (MAFF RB427)*. We consider data from four farms: Hamilton (1998 observations over two years), Kemble (1467 observations over three years), Welford (1151 observations over six years) and Roxhill (710 observations over 3 years). SOYL also provided geo-reference yield data for each field on each farm and details of the crop grown.

SMN observations were taken from seven previous studies of within-field N variation and two sets of measurements on transects. Full details of the sample schemes and analysis techniques used are contained in the works cited below. The two transects, one consisting of 256 observations each separated by 4 m (Lark *et al*., 2004) and one consisting of 256 observations each separated by 29.44 m (Haskard *et al*., 2010) were both located in Silsoe, Bedfordshire. The longer transect included observations from non-arable land which were removed. Observations from four fields in Bedfordshire were provided by the University of Reading (Baxter, 2002). These surveys all consisted of more than 100 observations and two of them included close pairs of observations so that short-scale variation could be explored. A survey of 100 observations on a regular 5 m grid at Rothamsted Research Station (Cordova, 2011) and 100 observations on a 50 m grid with additional short-scale comparisons from Silsoe (Lark *et al*., 1998) were also included. The RB209 recommendations are based on SMN measurements to 90 cm depth. Some of the observations collected in these surveys were only measured to 30 or 60 cm but these were scaled to 90 cm using factors determined where all three depths (0-30, 30-60 and 60-90 cm) had been recorded.

Models of P and K variation

The P and K data from each farm were filtered prior to model fitting to only include fields where wheat had been grown in the previous season. The procedure described below was used to fit a model of variation for each farm. Observations collected in different years were assumed to be independent so that the fitted model represented the spatial correlation between observations made in a single year.

Geostatistical models express the spatial dependence of a property in terms of the variogram (Webster and Oliver, 2007) which describes how the expected squared difference or semi-variance between observations of a nutrient increases with separation distance. Geostatistical models are generally based on the assumption that the variable has a Normal distribution. However within-field observations of soil nutrients can include extreme values or hotspots which are not consistent with this assumption. It is important that our spatial models include the potential for hotspots since they will have a large effect on the errors in field-scale estimates based on bulked cores.

The models of P and K variation assumed a Generalized Extreme Value (GEV) distribution, rather than a Normal distribution. Such a model has previously been used to map soil contamination by heavy metals (Marchant *et al*., 2011). The parameters which describe the GEV include a shape parameter which controls the likelihood of hotspots being present. The model has two components – the fitted GEV probability density function (pdf) which describes the marginal distribution or histogram of the observations and the variogram which describes the spatial correlation between observations.

The full procedure for simultaneously fitting the two components by maximum likelihood is described by Marchant *et al*. (2011). In the current study an approximate procedure was used because of the size and number of datasets being considered. Initially the spatial correlation was ignored and a GEV distribution was fitted to data from each farm. The observations from each farm were then transformed twice: first to their quantile values using the fitted GEV distribution and second to an approximate Normal distribution of mean zero and variance one, using the inverse Normal distribution function. An exponential variogram was fitted to the approximate Normal distribution by maximum likelihood (ML; Webster & Oliver, 2007).

One of the sampling methods tested was based on the yield map from the previous season. To generate this sample design we therefore required a simulation of the yield from the previous season which corresponded to a simulation of P or K. These correlated simulations were generated from a model known as the linear model of co-regionalization (LMCR) which is often used to represent the relationship between spatially correlated soil properties. It describes the distances over which the soil properties are related to each other and the strength of this relationship. The LMCR was fitted to P or K observations and corresponding yield observations by the ML procedure described by Marchant & Lark (2007a). All of the fitted models suggested that the relationship between yield from the previous season and observed nutrient content was fairly weak.

Models of SMN variation

RB209 underlines the importance of distinguishing between SMN and SNS. The SMN is the most important component of SNS, but SNS also includes the total crop nitrogen content and the mineralisable nitrogen in the soil. We describe in Section 3.2.4 how in this project we assume that the percentage errors in estimating SNS are identical to the percentage errors in estimating SMN. Therefore for the purposes of our calculations the two terms become synonymous. However in a more general framework where the errors in estimating the total crop nitrogen content and the mineralisable nitrogen were better understood, the distinction would be important. We refer to the quantity that is measured from the bulked sample as the SMN and the total of SMN, mineralisable and crop N as the SNS. Therefore in the discussion below it is the expected SNS which controls the variation of SMN.

The within-field variation of SMN is known to increase with the mean SNS within a field. For this reason the HGCA Nitrogen for winter wheat management guidelines recommend that more sampling is conducted on fields where the SNS is expected to be large. To conduct tests that were consistent with these recommendations we required a model that explicitly related SMN variability to expected SNS.

The model for SMN assumed that the standard deviation of SMN was linearly related to the expected SNS. A single model of this form was fitted to the data from all nine of the SMN datasets used in this study. The observations from each dataset were divided by the mean SMN concentration for that dataset. Then an exponential variogram was fitted to the combined scaled datasets by ML, assuming that spatial correlation only occurred within datasets and not between observations from different datasets. This variogram can be thought of as the spatial model when the expected SNS is 1 kg/ha. The model for larger expected SMN concentrations can be determined by scaling this model.

Selection of exemplar fields

The yield maps provided by SOYL were separated on a field by field basis. Therefore the outlines of exemplar fields could be extracted from these maps. Six fields of sizes 5, 10, 20, 30 and 60 ha (Figure 1) were selected as representative of a range of typical UK arable fields.

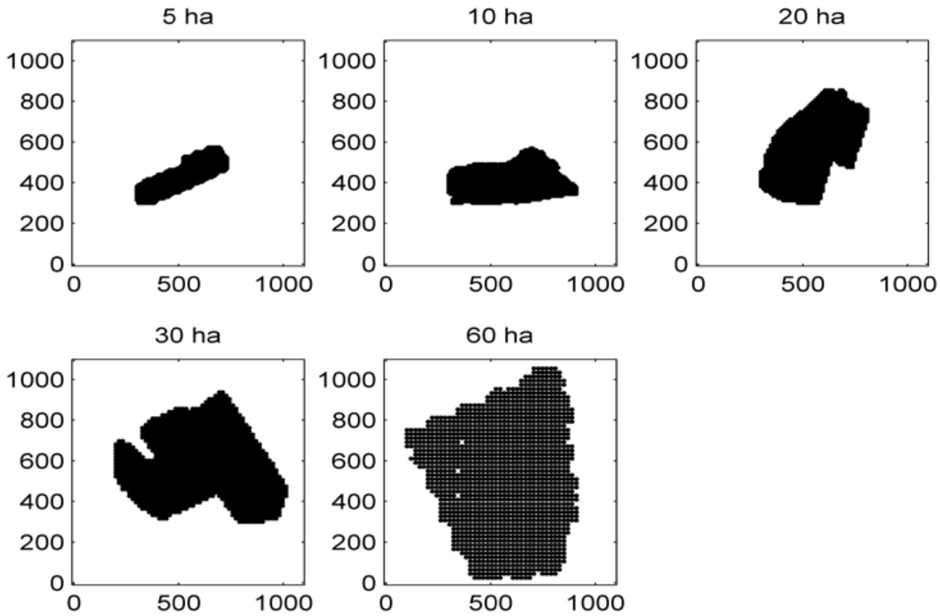

Figure 1. Outlines of exemplar fields, coordinates on axes are measured in m.

Simulation of soil nutrient concentrations within exemplar fields

For P, K and SMN, multiple simulations of the nutrients across each exemplar field were generated by an LU simulation algorithm (Webster and Oliver, 2007). This algorithm produces simulations of a spatial property which are consistent with its model of variation. For SMN the model fitted in the manner described above leads to a realization of mean 1 kg/ha. This is then scaled by the expected SMN concentration for the scenario being considered.

For P and K we used the models fitted to the SOYL data to simulate variation at the 5 m^2 scale. An additional component of uncertainty is required to downscale this variation to the core scale. The size of this component was based on the results of the Oliver *et al*. (1997) study of bulking strategies for P and K. Their results suggested that when the mean P concentration was 21.5 mg/l, the additional variance upon down-scaling to a single core from 16 cores over 5 m^2 was approximately 8.5 $(mg/l)^2$. When the mean K concentration was 294 mg/l, the additional variance was 800 $(mg/l)^2$. This additional variation is likely to be larger for larger P and K concentrations, as observed in the Netherlands by Brus *et al*. (1999). We therefore assume that the standard deviation of the additional variation is proportional to the mean variation at the 5 m^2 scale. Based on the observations of Oliver *et al*. (1997) we assumed that this variation had a coefficient of variation of 0.15 for P and 0.1 for K. A realization of a white noise (i.e. Normal) process with zero mean and the appropriate coefficient of variation was added to each simulated value at the 5 m^2 scale.

3.2.3. Quantification of sample errors by different designs

Proposed sample designs

The performance of five different types of sample scheme was compared within the project. Examples of each of these designs for a bulked sample consisting of 20 cores are shown in Figure 2.

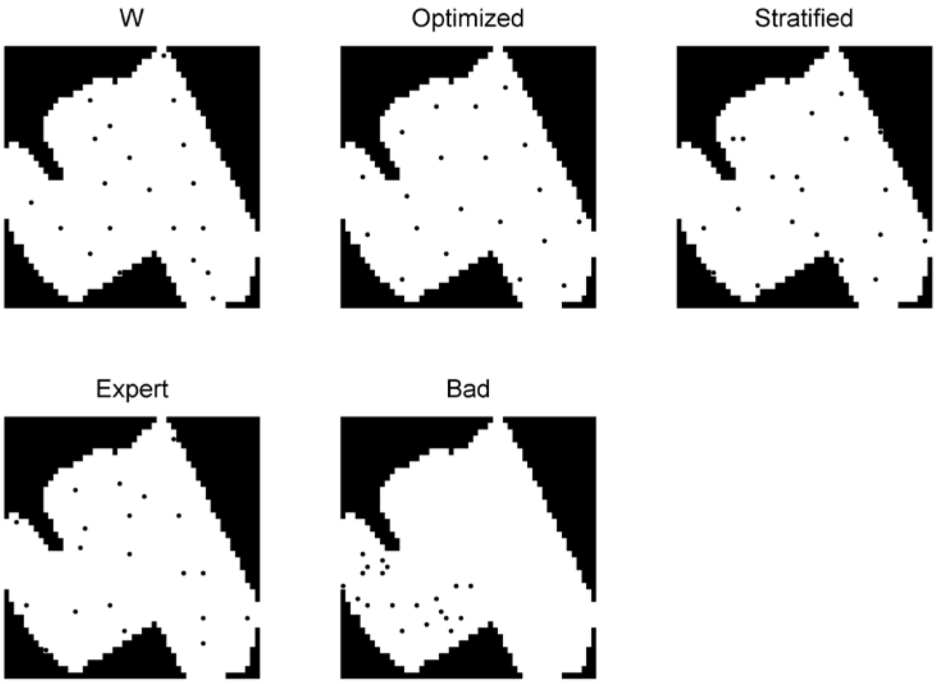

Figure 2. Twenty point examples of five sample designs tested in this project.

The 'W' design

The 'W' design is commonly used by agronomists to determine the field mean nutrient content and is the design recommended in RB209. It requires the practitioner to walk in a 'W' pattern across the field and extract soil cores at regular distance. The 'W' should cover as much of the field as is possible. The design is favoured because of its simplicity. There is no need to use statistical algorithms or to exactly find sampling sites with a GPS. It does disperse points within the field. However there is potential for inefficiency at each apex of the 'W' since two cores might be extracted close together and as the number of cores increases there is a limit to how accurate the estimate of the field mean becomes. This is because sites not on the 'W' are never sampled.

The optimized design

The optimized design uses a computational algorithm known as spatial simulated annealing (van Groenigen, 1999) to select the sites of soil cores which will lead to the most accurate estimate of the field mean. The algorithm uses a formula suggested by Burgess and Webster (1984) to calculate the expected error associated with each potential design. This formula is written in terms of the spatial model that describes the data. In reality the spatial model for a particular soil nutrient will not be known prior to sampling but the sample design that results is very insensitive to the assumed model. Throughout this study we consider a simple variogram model where the semi-variance is always increasing. This leads to designs such as Figure 2, 'optimized', where the points are evenly dispersed across the study region. In a previous HGCA project (Marchant *et al*., 2005) we demonstrated that an optimized scheme can lead to smaller errors than the 'W'.

Spatially stratified random sampling

Stratified random sampling is another method used to ensure that soil cores are fairly evenly dispersed throughout the field. A computer algorithm (Walvoort *et al*., 2010) is used to divide the field into a number of continuous sub-regions or strata of equal area. The algorithm ensures that these sub-regions are as compact as possible. This means they resemble circles or squares rather than long thin shapes. An example of the stratification is shown in Figure 3. A location is selected at random within each strata leading to the design shown in Figure 2 'stratified'.

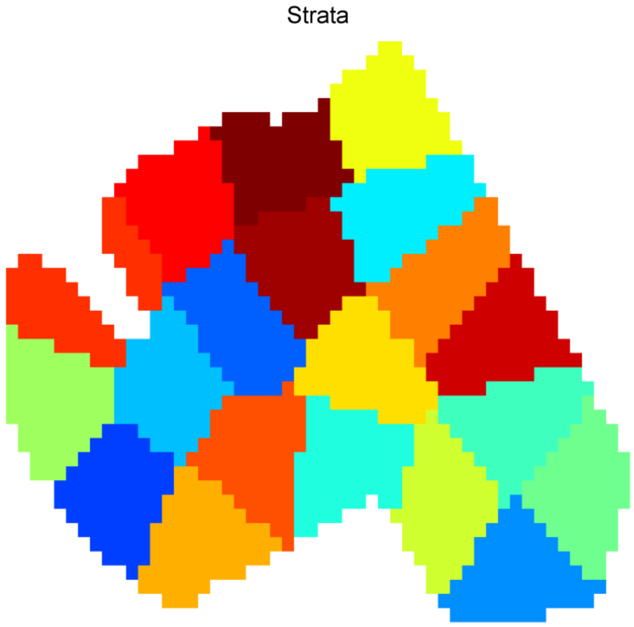

Figure 3. The 30 ha field divided into 20 spatial strata.

Using expert knowledge

Where a practitioner has detailed knowledge of the causes of variation in a field or of yield maps from previous years, it is possible to incorporate this knowledge in the sample design to make the bulked sample more representative of the field as a whole. This is achieved through an iterative method known as rank set sampling (Chen *et al*., 2003). If the bulked sample is to consist of n cores, then in the first iteration n sites in the field are selected at random and the expert decides which of these sites they expect will have the largest concentration. The first core of the bulked sample is taken from this site. The second site is chosen by selecting a different n random sites and requesting that the expert selects the site he expects to have the second largest concentration. The algorithm continues until the n^{th} site is selected. To some degree, the success of this strategy depends on the quality of this expert knowledge, but the theory shows that it is an unbiased sampling strategy and never less efficient than simple random sampling. For the purposes of this study it is difficult to objectively determine this quality so we replace the expert knowledge by the yield map from the previous season, so that the first core is the member of the set of randomly located sites with the largest yield.

Bad practice

The bad practice or clustered design is included to quantify the costs associated with non-representative sample designs. The first point of this design is chosen at random and then all subsequent points are within 30 m of the preceding one leading to the clustering of points seen in Figure 2 ‘Bad’.

Simulation tests

The criterion used in the initial simulation tests is the mean absolute error (MAE) upon estimation of the field mean nutrient content by the proposed sample designs described above. One thousand simulations of P and K were generated for each combination of the five exemplar fields and the fitted models from the four farms. Each simulation was then sampled using the five designs. Four of the five designs (‘W’, stratified random sampling, expert knowledge and bad practice) included a random component. In the ‘W’ design this arose by varying the distance along the ‘W’ where the first core was extracted. Therefore it is probable that chance could cause some realizations of these designs to be better than others. This effect was compensated for by generating a unique realization of the sample scheme for each nutrient simulation. This was not necessary for the optimized design which was generated by a deterministic rather than random method. The actual mean nutrient content of each simulated field was recorded along with the estimated means according to the different sample designs.

A similar procedure was followed for SMN sampling with various expected SMN concentrations for the field. Again the actual mean SMN for each simulated field was compared with the estimates from the different sample designs.

3.2.4. Determination of the implications of errors for P and K

The RB209 fertiliser recommendations for P and K aim to maintain these nutrients within Index 2 of the classification. For Olsen's P this corresponds to levels between 16 and 25 mg/l and for extractable K between 121 and 240 mg/l. In conducting this study we assume that the RB209 recommendations are based on an accurate understanding of the dynamics of soil nutrients and that if the soil nutrient content of the soil is known exactly and the recommendations followed then the optimal indices will be maintained. However, the errors associated with estimating soil nutrient concentrations from a bulked sample will lead to sub-optimal management. The practical effects of these errors and the sub-optimal soil management decisions that result were explored via mathematical models of the evolution of soil nutrient concentrations.

Mathematical models of long-term temporal variation of field mean soil P and K

The models of long-term variation of field mean soil P and K accounted for four key processes:

1. The decision about the amount of fertiliser that is added
2. The increase in soil nutrients because of this fertiliser
3. The effect of the soil nutrient concentrations on yield
4. The loss of nutrients from the soil to the crop

We note that the Olsen P and ammonium nitrate extract test for K only measure a portion of the total P and K in the soil. Our models are expressed in terms of this portion rather than the total concentrations to maintain consistency with the measured values. We denote the actual concentration of the test portions as P_{Test} and K_{Test} and the estimated concentrations by P^{*}_{Test} and $K^{*}_{\mathrm{Test.}}$

Fertiliser Decisions

The RB209 fertiliser recommendations for P and K on cereals are on an index-by-index basis. For our mathematical representation of the crop system it was more convenient to assume that the additions varied continuously with the P^{*}_{Test} and K^{*}_{Test} values rather than jumping at the boundary between indices. This was consistent with a practitioner adjusting fertiliser additions for fields at the top or bottom of an index. The RB209 recommendations also account for the nutrients that are removed by the crop if the desired yield is achieved. RB209 assumes that 7.8 kg of phosphate and 5.6 kg of potash is lost from the field per tonne of cereal grain yield. The target yield is denoted by Y_T (t/ha). Based on the RB209 recommendations we assumed that when $P^{*}_{\mathrm{Test}} < 20 + 7.8\,Y_T$:

$$F = 80 - 4\,P^{*}_{\mathrm{Test}} + 7.8 Y_T, \qquad (1a)$$

and when $K^{*}_{\mathrm{Test}} < 150 + 5.6\,Y_T$:

$$F = 75 - 0.5\,K^{*}_{\mathrm{Test}} + 5.6\,Y_T, \qquad (1b)$$

where F is the amount (kg/ha) of added phosphate for P or potash for K. For larger observed nutrient values, no fertiliser was added.

Effect of fertiliser additions on soil test values

Only a proportion of the added fertiliser is partitioned into soil test P and K. For P, this proportion was determined by linear regression on unpublished data of P depletion and build-up provided by AE Johnston of Rothamsted Research. Observations for K were available from Johnston and Goulding (1990). The proportions which resulted were 0.18 for P and 0.34 for K. If we assume that the added nutrient is uniformly mixed in the top 30 cm of soil then the increase in soil test values must be further scaled by 0.33 to convert from kg/ha to mg/kg.

Yield response to soil-test values

The relationship between crop yield and soil-test P and K is generally overshadowed by variation in SMN and therefore limited suitable data are available to fit a model of this relationship. For P a model was fitted to three previously published response curves (Syers *et al*., 2008; Johnston, 2005). For K a model was fitted to data from seven yield response experiments (Milford and Johnston, 2007, Johnston and Goulding, 1988). In both cases the data corresponded to a range of crops so each yield response was normalized such that the yield at the target nutrient level was 8.8t. The fitted responses were of the form:

$$Y = Y_T\,(1 + A \times B^N), \qquad (2)$$

where Y was the realized yield, N the soil test nutrient value (P_{Test} or K_{Test}), and A and B were fitted parameters. The fitted values were A=-1.33 and B= 0.68 for P and A=-2.01 and B = 0.96 for K. The yield curves which resulted are shown in Figure 4.

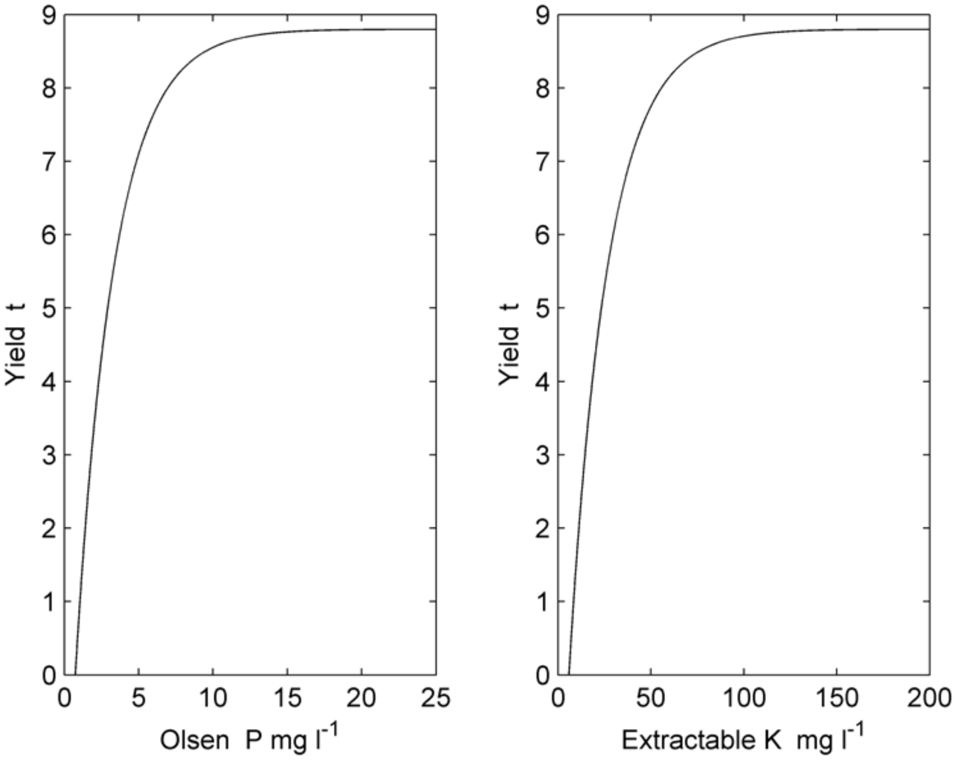

Figure 4. Fitted yield responses for P and K.

Soil-test nutrient removed by the crop

Following the RB209 assumptions the total phosphate and potash removed by the crop are 7.8 kg/t and 5.6 kg/t respectively. However only a proportion of the P and K will come from the soil-test partitions and experiments by Johnston and Poulton (1992) and Johnston (1986) suggest that this proportion increases as the soil-test concentrations increase. For P, this proportion was approximated by $0.062 + 0.0031\ P_{\text{Test}} + 0.00017\ P_{\text{Test}}^{2}$ and for K it was approximated by $-0.21 + 0.0037\ K_{\text{Test}}$. These relationships were fitted to data where the proportion of soil-test P lost varied between 0.08 and 0.47 and the proportion of soil-test K lost varied between 0.05 and 0.85. We constrained the proportions within these bounds.

Simulation test methodology

The components described above provide sufficient information to model the long term dynamics of P and K in the soil. Our simulation tests considered the long-term effects of different sampling strategies on these dynamics. We assumed that at the start of the experiments the soil nutrient status was optimal (i.e. $P_{\text{Test}} = 21$ mg/l and $K_{\text{Test}} = 180$ mg/l). The soil was sampled according to the specified strategy every four years. The error in estimating the mean field nutrient concentration was extracted from each of the 1000 experiments for each nutrient/farm/field combination. The models were used to determine the responses when fertiliser management decisions were based on the actual nutrient level plus this error. This assumes that changes in the nutrient content are uniform throughout the field and that the sampling errors are unaffected by them.

The annual phosphate additions until the next phase of sampling were determined by substituting the observed P^{*}_{Test} value into Equation 1a. The changes in soil test P which resulted were modelled based on the actual concentration P_{Test}. The model system was run until 1000 sampling phases had been conducted. The modelling framework ensured that if the observed soil test P was equal to the actual soil test P throughout the run then the soil test P would remain at the optimum. The errors from sampling P cause the soil test P to fluctuate around the optimum value as shown in Figure 5 and the extent of these fluctuations were used to assess the effectiveness of the sample scheme. The same procedure was followed for K.

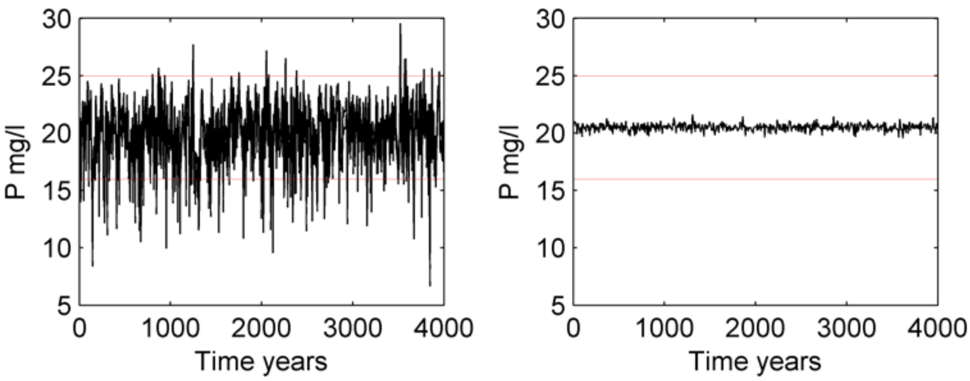

Figure 5. Long term evolution of soil P when mean absolute sampling errors are 5 mg/l (left) and 1 mg/l (right).

3.2.5. Determination of the implications of sampling errors for SMN

Since nitrogen is often the limiting factor on crop yield and the effect of nitrogen additions can be observed on an annual basis, the cost effectiveness of sampling for N is assessed in terms of the effects on these yields and the resultant effects on profit. The relationship between soil nitrogen supply (SNS) and observed SMN is complicated and the subject of HGCA Project 3425 *Establishing best practice for SNS estimation*. In that project the effect of various factors such as the time of SMN measurement, soil texture, rainfall, and sample storage are being considered. However for the purposes of this project we assume that the percentage errors in estimating SNS are the same as those from estimating SMN. If the relationship between SMN and SNS were to be more fully understood then this information could be incorporated into our simulation tests.

Mathematical models of the relationship between profitability and N management

A previous HGCA project (Sylvester-Bradley *et al*. 2008) considered the net reduction in profit from sub-optimal N management. This project showed that errors of less than 20 kg/ha in determining N applications had little effect on the total profit. However the profit decreased rapidly with N additions applications that were more than 100 kg/ha from the optimum.

Our N modelling framework was based on a yield response curve fitted to the data from (Sylvester-Bradley *et al*. 2008). The data included 83 trials of the yield response to N for contemporary winter wheat cultivars. The fitted linear exponential model was of the form

$$Y = a + b \times r^{N} - c\, N \quad (3)$$

where N is the SNS kg/ha, Y the yield in t/ha. The fitted parameters were a=37.5, b=-37.8, c=-0.0362, and r= 0.9976. The yield response curve has a maximum yield of 8.58 t/ha for 380 kg/ha SNS.

When the cost of fertiliser is accounted for, the maximum profit occurs when the derivative of Equation (3) is equal to the ratio of the price of fertiliser N (£/kg) to the price of the crop (£/t). This ratio is referred to as the breakeven ratio (BER). A formula for the optimal SNS, denoted N_{opt}, as a function of the BER can be determined by calculation of the derivative of (3). The nutrient management strategy requires the practitioner to add sufficient fertiliser to match the deficit between the SNS prior to fertiliser application and N_{opt}. In this project we followed the HGCA *Nitrogen for winter wheat – management guidelines* and assumed a BER of 5 and additionally assumed a wheat price of £100 / t. Figure 6 shows the losses which occur because of erroneous SNS estimation and hence sub-optimal N additions under these assumptions.

If we denote the error in estimating SMN (and hence SNS) by E (which can be positive or negative) and if the practitioner follows the recommendations but bases his decision on the erroneous SMN estimate, then the total nitrogen supplied to the crop will equal $N_{opt} + E$. The yield which results from this application can be calculated from Equation (3). This yield is multiplied by the price of wheat to determine the income from the field. The cost of fertiliser and sampling costs are subtracted to determine the profit. The sampling costs were based on advice from the project steering group and were assumed to be £5.33 per extracted core plus £86 to conduct the laboratory analysis of the bulked sample over three depths (0-30, 30-60, 60-90 cm).

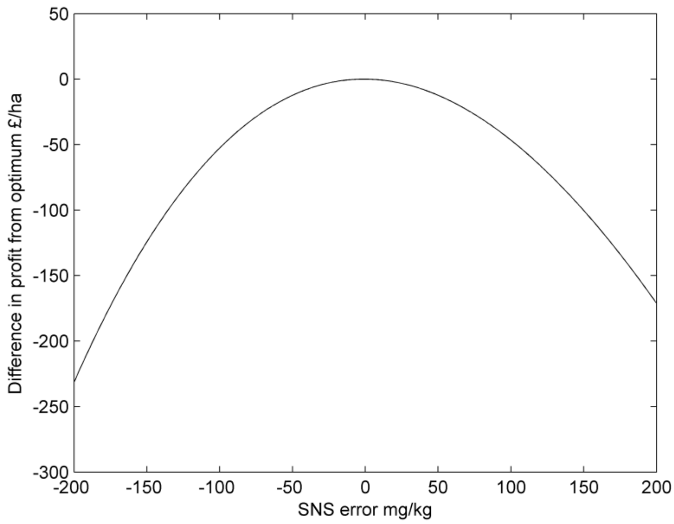

Figure 6. Reduction in profit because of estimation of SMN and hence sub-optimal N application for BER of 5 and wheat price £100/t.

Simulation test scenarios

RB209 and the HGCA Nitrogen for winter wheat – management guidelines both state that the variability of SMN increases with expected SNS. Therefore the sampling requirements for N are likely to change according to the expected SNS and our simulation tests must account for this.

We represented the practitioner's uncertain knowledge of SNS levels prior to sampling by a probability density function (pdf). This knowledge would be based on factors such as climate, soil type and previous management decisions. The simulation tests quantify the benefits of sampling beyond basing N fertiliser management decisions on this prior knowledge. Examples of the pdfs used in this study (with modal values of 50, 150 and 250 kg/ha) are shown in Figure 7. The level of uncertainty is such that when, for example, the practitioner expects that the SNS is around 50 kg/ha there is a 2% chance that the actual content is greater than 200 kg/ha. Such an error might arise if previous applications have been poorly documented.

The simulation tests were repeated for circumstances where the modal value of the practitioner's prior SNS pdf were 50, 100, 150, 200, 250 kg/ha. The baseline was the profit which would have resulted from determining fertiliser additions from the modal value. The actual expected SNS value was randomly selected from the pdf and a simulation generated based on this value. This simulation was sampled according to the procedures described in Section 3.2.4 and the field estimate of SNS was determined. The profits that result from using this estimated value are calculated and the baseline profit is subtracted so that the remaining profit is the benefit from sampling. For each sampling configuration being tested this process is repeated 1000 times and the average profit above the baseline is recorded. The number of soil cores that leads to the largest average profit is the rational sampling effort for the expected SNS.

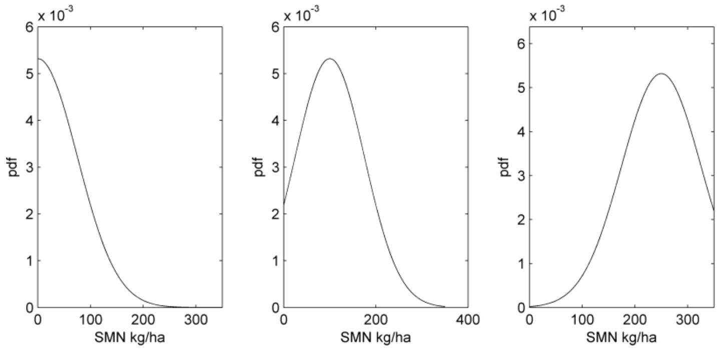

Figure 7. Probability density functions of the practitioner's prior knowledge of SNS.

When factors such as soil type, previous cropping and fertiliser additions are considered to be uniform over a large area the use of barometer fields might be recommended to reduce sampling effort. The barometer field is a portion of the total target area. We tested the implications of only sampling in 10 ha barometer fields when the target area was 20, 30 or 60 ha. These tests followed the same procedure described above. In each simulation we assumed that the target area was relatively uniform and that the same expected SMN concentration applied throughout. The barometer field was located at the edge of the target area. Again the rational sampling effort was determined and the profit which resulted from this rational effort was compared with the profit from rationally sampling across the entire target area.

3.2.6. Efficient detection of regions of nutrient excess or deficiency

In a study of the change in available soil K over a single season, Bishop & Lark (2007) demonstrated how a relatively stable nutrient is likely to develop regions of deficiency, 'coldspots' and excess 'hotspots' under uniform application. There is therefore a need to identify coldspots and hotspots to avoid limitations on production and emissions to the environment. At present, information on spatial variation of properties within fields, where it is collected at all, is obtained on square grids, typically with a 100-m interval. However, such practice is unusual, somewhat expensive and it might not be sufficient to adequately resolve spatial variation (Oliver and Carroll, 2004). The resolution of coldspots and hotspots could be improved with more intense sampling but the additional labour and laboratory analyses mean that this is unlikely to be cost effective.

A similar problem exists in surveys of soil pollution which require cost effective methods to determine where remediation is required (de Gruijter *et al.* 2006). Practitioners have noted that the laboratory costs are a substantial proportion of the costs of a soil pollution study. They have therefore developed sophisticated bulking strategies which reduce the number of laboratory analyses that are required without sacrificing the resolution of the maps which result. One such approach is the sequential sweep-out method of Gore *et al.* (1996). We test whether the sequential sweep-out method can lead to more efficient identification of soil-nutrient hotspots.

We consider a situation where soil cores have been collected upon a regular grid within a field. We seek to identify the cores for which the soil nutrient concentration is greater than a critical threshold with as few laboratory analyses as possible through an efficient bulking strategy. Initially bulked samples are formed for each row and column of the sampling grid and the nutrient content of each bulked sample is determined. Some soil is held back from each site as this might be required for subsequent analyses.

The maximum possible concentration is then determined for each site in turn based upon the measured concentrations of the bulked samples. This maximum corresponds to the situation where the site in question is the only one contributing nutrient to the bulked row or column core. If all of these maximums are less than the critical threshold then the survey is complete. Otherwise the core from the site with the largest attainable value is analysed so that the nutrient concentration at this site is known. The concentrations of the two bulked samples which include this site are then adjusted to account only for the sites where the concentration is unknown and the maximum attainable value at each site is re-calculated. This procedure continues until the maximum attainable concentration at each unmeasured site is less than the critical threshold. At this stage an indicator kriging method is used to interpolate a map of where the threshold is exceeded based upon the knowledge of which cores exceed the threshold.

We applied this approach for both P and K on a grid of 9 columns and 10 rows each separated by 50 m. The nutrient values at each site were simulated at each site based in the models of variation determined in Section 3.2.2. The sequential sweep-out method was then conducted based on these simulated values and the number of analyses required to delineate between the cores above and below a specified threshold was determined.

3.3. Results

3.3.1. Models of spatial variation of soil nutrients

The fitted models of within-field P (Figure 8) and K (Figure 9) variation all showed evidence of non-Normal variation. The P pdfs on all four farms had a slowly decaying tail to the right which indicated that hotspots of P were present on the farms. The pdfs for K are more symmetric than those of P but there is still some evidence of hotspots. There was spatial correlation on all farms for both nutrients although the range of this correlation varied between 1 and 3 km. Differences between the models of variation on different farms are evident. For P, the right hand tail of the pdf decays much more slowly on Kemble and Welford and hence P concentrations on these farms are more variable. For K, Kemble has the slowest decaying pdf and Hamilton has larger concentrations than the other farms.

The model of SMN (Figure 10) suggested that SMN is spatially correlated up to around 200 m. This is a substantially shorter range than for P and K. The semi-variance between distant observations of SMN is 0.29 $(kg/ha)^2$ which corresponds to a coefficient of variation of 0.54. This is comparable with coefficients of variation observed in previous SMN surveys, such as those reviewed by McBratney and Pringle (1999).

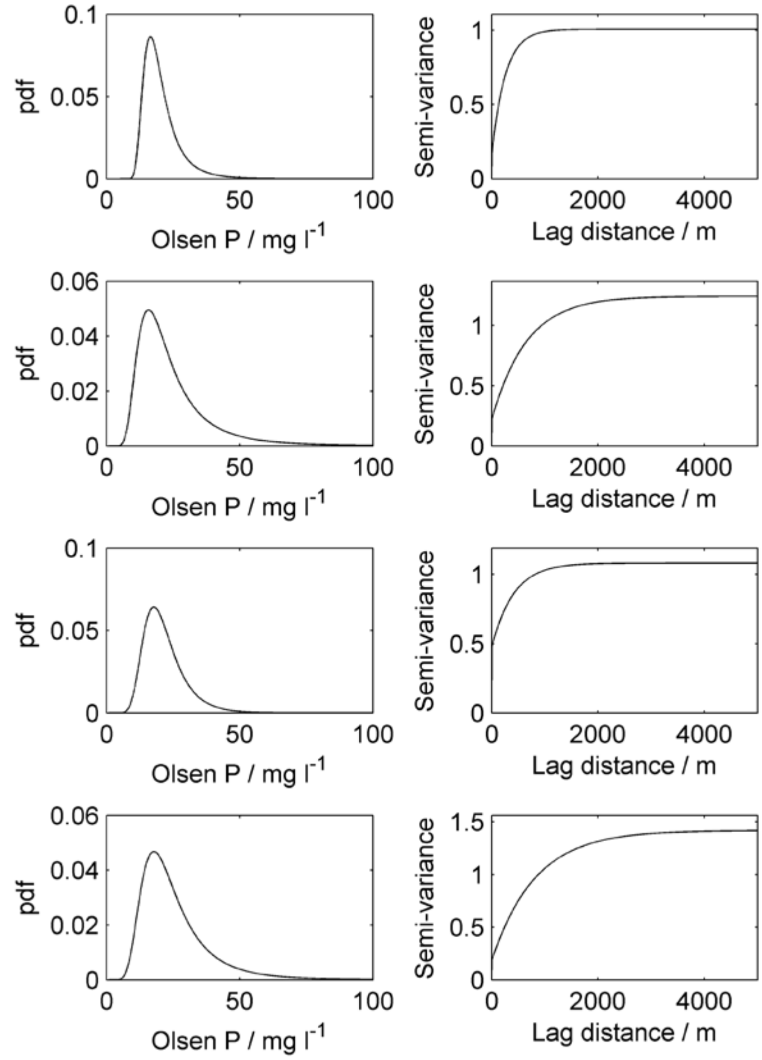

Figure 8. Fitted pdfs and variograms for P at (top-bottom) Hamilton, Kemble, Roxhill and Welford.

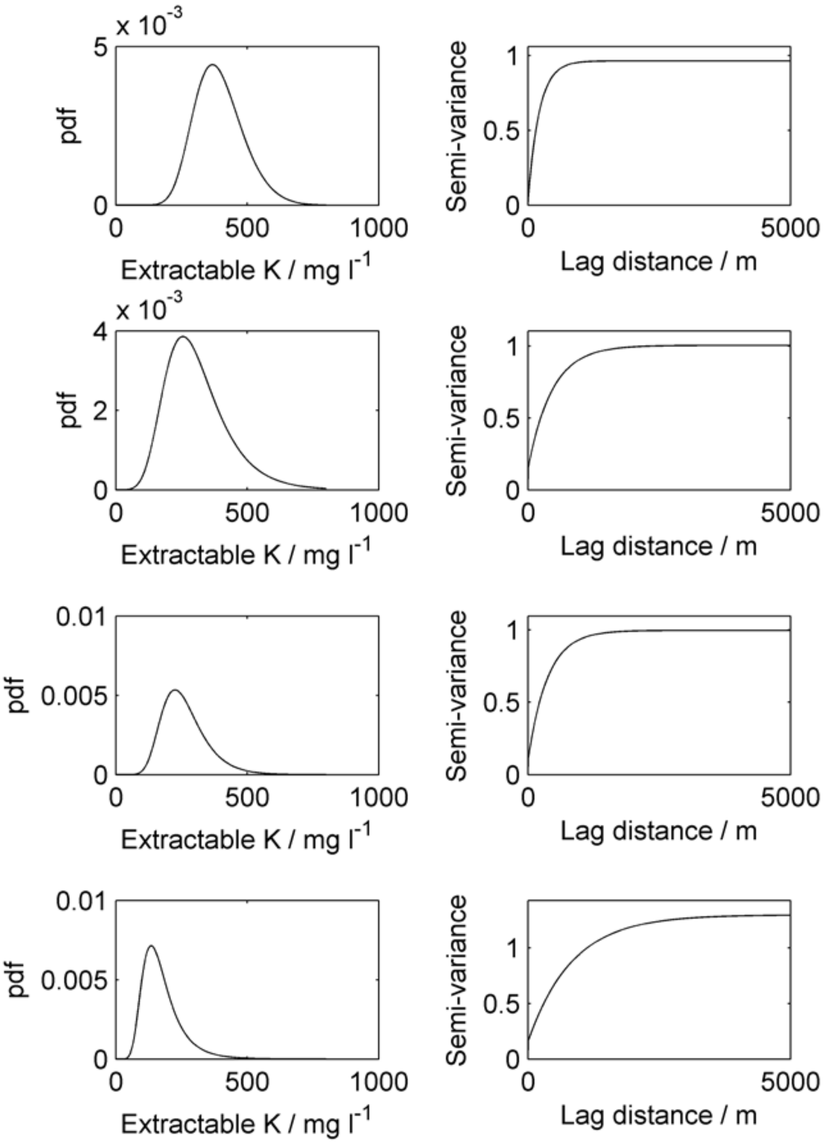

Figure 9. Fitted pdfs and variograms for K at (top-bottom) Hamilton, Kemble, Roxhill and Welford.

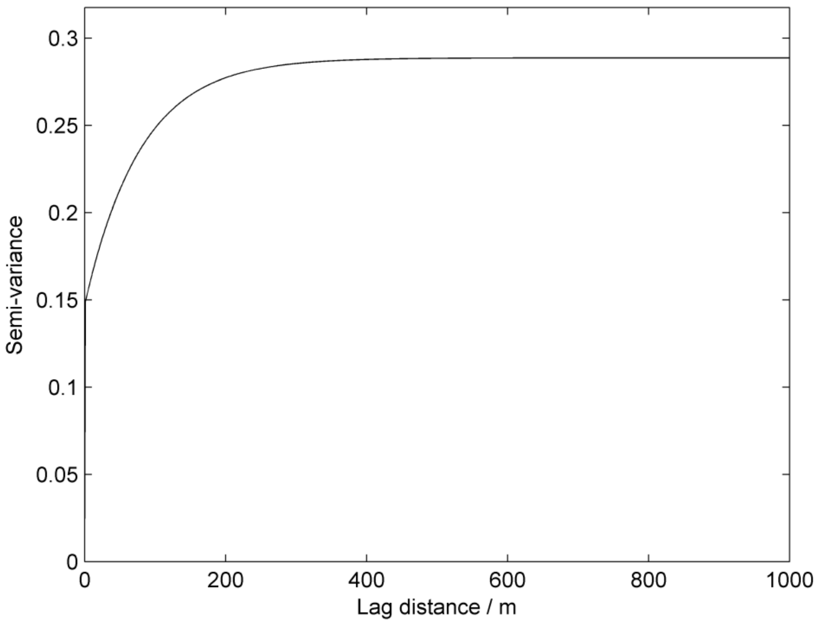

Figure 10. Fitted variogram model for SMN after scaling by expected SMN.

3.3.2. Simulations of soil nutrient variation within fields

Spatial correlation and hotspots are evident in all of the simulations of soil nutrients across exemplar fields (Figures 11-13). In Figure 12 the simulated yield from the previous season is shown alongside the corresponding simulation of K. There is some correlation between the yield and the soil-test K but this relationship is fairly weak. The three SMN simulations in Figure 13 are scaled by different expected SMN concentrations for the field. As the expected SMN increases from left to right, so does the amount of variation within the simulations.

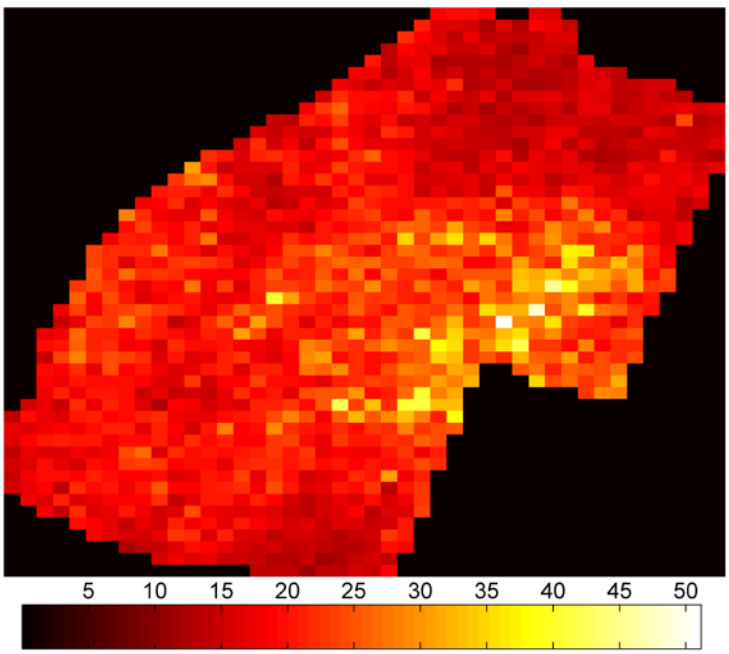

Figure 11. Simulated variation of P (mg/l) across the 20 ha field using the model fitted to Hamilton data.

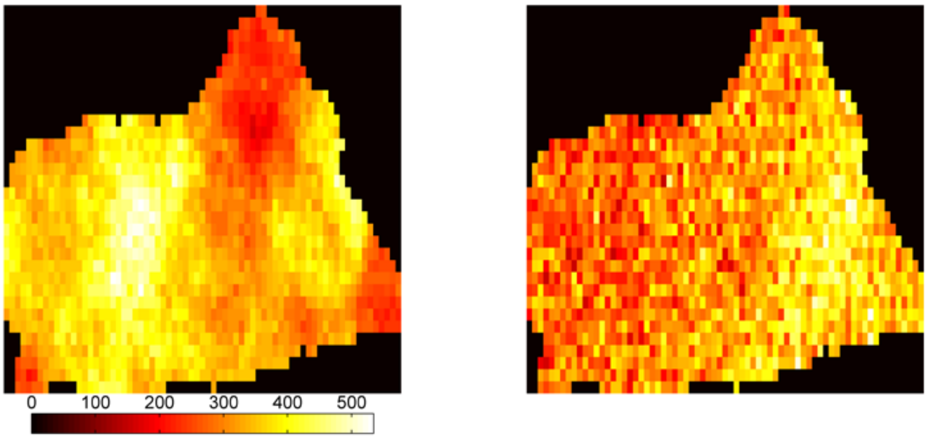

Figure 12. Simulated variation of K (mg/l) within 10 ha field using model fitted to Hamilton data (left) and corresponding normalized yield map (right).

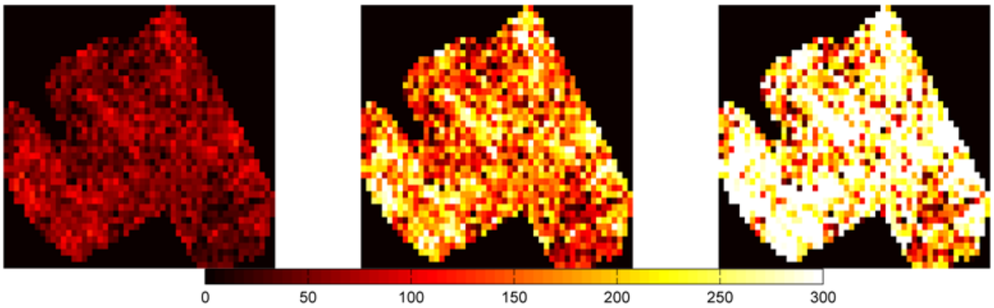

Figure 13. Simulated variation of SMN (kg/ha) for 30 ha field and expected SNS of 50 mg/ha (left), 150 mg/ha (centre) and 250 mg/ha (right).

3.3.3. Sampling errors for different designs

The sampling experiments demonstrated that for all three nutrients both the optimized and stratified sample designs led to smaller errors than the 'W' design (Figures 14-16). With the exception of the bad practice design, the errors for all of the designs reduced rapidly as the number of cores increased from one to ten. The rate of decrease in errors then slowed. For fewer than ten cores the difference between the optimized and 'W' designs is small. However for more than 20 cores a more obvious difference is evident. The expert design performs slightly worse than the 'W', optimized and stratified designs. This is probably because of the weak relationship between the nutrient concentrations and the expert knowledge (i.e. the yield maps). The errors would be reduced by better expert knowledge but the statistical assumptions made about the variation of the soil nutrients mean that the expert design will not out-perform the optimized design. The errors decay slowly for the bad practice design and illustrate the importance of ensuring that the sampled cores are representative of the target area as a whole.

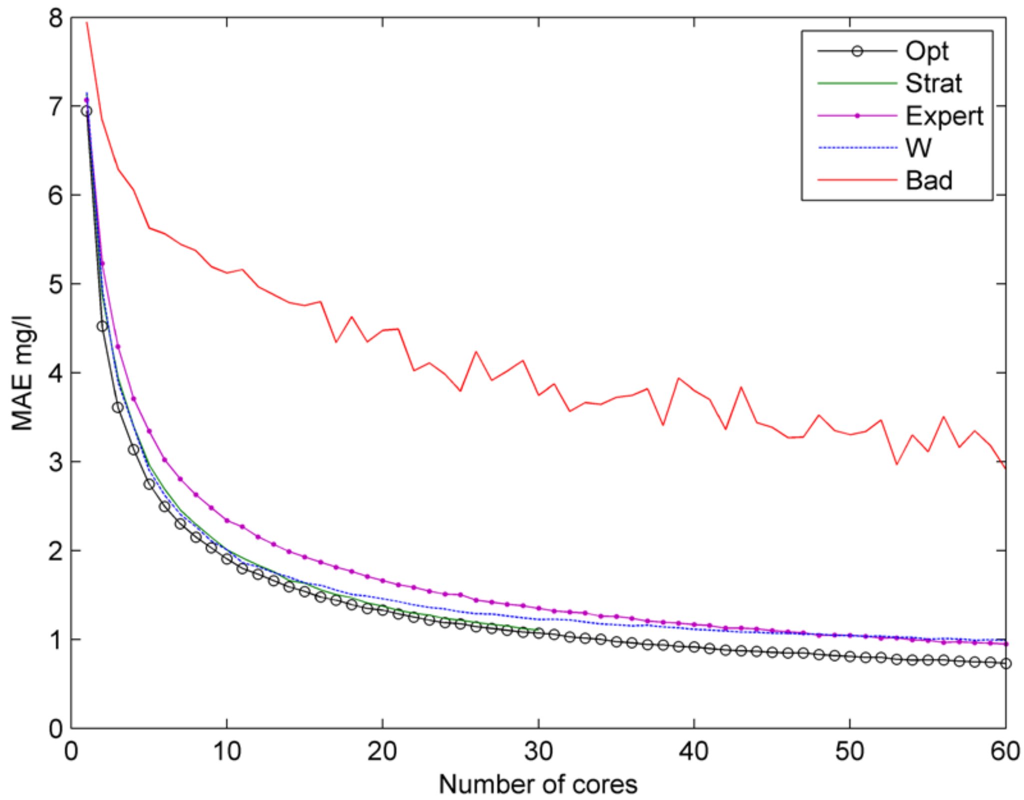

Figure 14. Sampling errors for P by different designs averaged over all farms and fields.

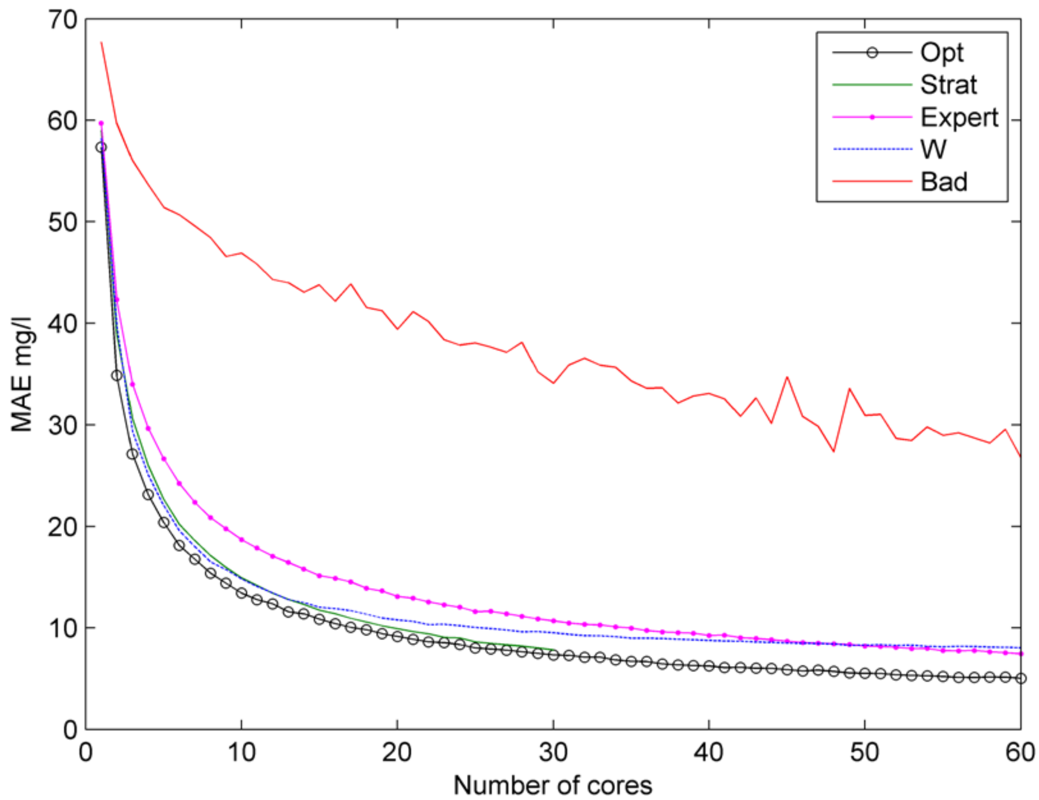

Figure 15. Sampling errors for K by different designs averaged over all farms and fields.

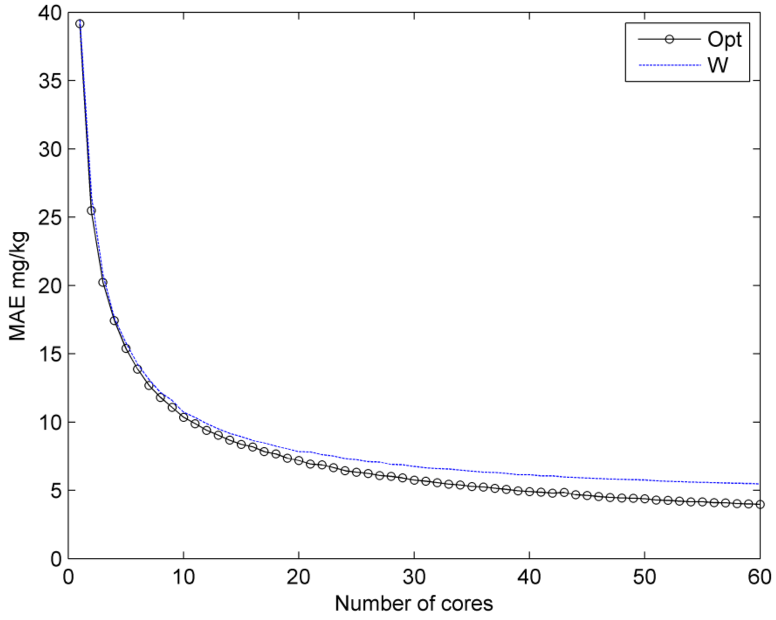

Figure 16. Sampling errors from 'W' and optimized designs for SMN on 10 ha field with expected concentration of 100 kg/ha.

3.3.4. Implications of sampling errors for soil P and K stocks

The effectiveness of the sample designs for P and K management are assessed in terms of the probability of sampling errors leading to the soil nutrient concentrations being outside Index 2 (16-25 mg/l for P and 121-240) of the RB209 recommendations. Figure 17 (left) shows how these probabilities for P vary with the number of bulked cores for the Kemble farm. For bulked samples consisting of fewer than 6 cores there is a substantial probability of the soil concentrations leaving Index 2. This quickly decreases with the number of bulked cores and the probability is negligible for more than 20 cores. There is very little difference between the W and optimized schemes in terms of the number of cores required to ensure that the P concentration leaves Index 2 in less than 2.5 % of years. However if the target range for P is halved to be between 18.25 and 22.75 (Figure 17; right) then more samples are required to maintain this target and the optimal design performs substantially better than the W. This reflects the observations in Section 3.3.3 that for more than 20 cores the errors on using the optimized scheme are substantially less than those for the W (Figures 12-13).

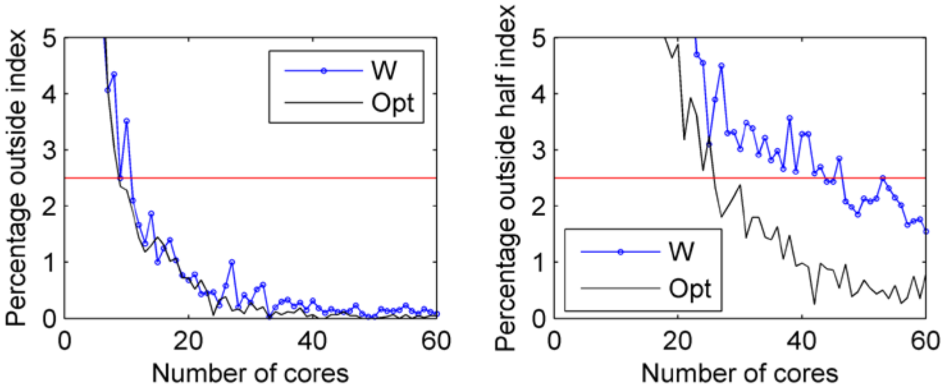

Figure 17. Comparison of ‘W’ and optimized designs in terms of percentage of years for which soil P concentration is outside RB209 Index 2 on 20 ha Kemble field (left) and percentage of years for which soil P concentration is outside half of RB209 Index 2 (i.e. 18.25-22.75) on 20 ha Kemble field (right).

3.3.5. Implications of sampling errors for N

Figure 18 shows the relationship between the expected profit after measuring SMN and the number of bulked cores on a 10 ha field with an expected SNS of 100 kg/ha. For this situation the largest expected profit of £750.80 per ha results from bulking 8 cores. The difference between the profit from using the optimized and ‘W’ schemes is a small proportion of the benefits of sampling. A small benefit from using the optimized scheme was seen across the simulation tests for different expected SNS levels and field sizes. However it never exceeded £0.20/ha of the benefits of sampling and therefore all further results of simulation tests use the W design only.

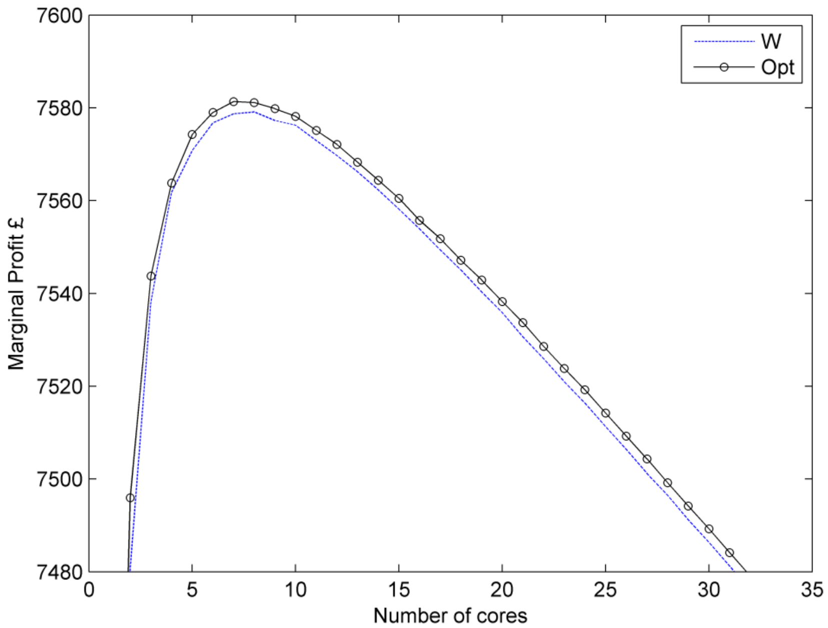

Figure 18. Comparison of ‘W’ and optimized sample designs in terms of total expected profit on a 10 ha field with expected concentration of 100 kg/ha.

3.3.6. Sampling requirements for P and K

According to our sampling experiments 10 bulked cores taken from a 'W' sample design every four years will be sufficient to ensure that both P and K concentrations remain in Index 2 of the RB209 guidelines for more than 97.5% of years (Tables 1-2). Substantially less sampling is required for K than P. There is variation in the sampling requirements on the different farms with only 3 and 4 cores being required on the less variable farms. In practice if fewer than 4 cores are required then the 'W' design would not be used but the practitioner would ensure that the cores were extracted a large distance apart.

Table 1. Number of cores taken from a 'W' every four years required to ensure that field-mean P concentration remains in Index 2 for more than 97.5 % of years.

Farm	5 ha	10 ha	20 ha	30 ha	60 ha
Hamilton	3	3	3	3	3
Kemble	7	7	9	9	9
Roxhill	4	3	3	3	4
Welford	7	9	8	9	9

Table 2. Number of cores taken from a 'W' every four years required to ensure that field-mean K concentration remains in Index 2 for more than 97.5 % of years.

Farm	5 ha	10 ha	20 ha	30 ha	60 ha
Hamilton	2	2	2	2	2
Kemble	2	2	2	2	3
Roxhill	1	1	1	2	2
Welford	1	1	1	1	2

3.3.7. Cost-effectiveness of SMN measurements and sampling requirements

The results of the sampling experiments for SMN are shown in Tables 3-8. The optimal intensity of sampling increases with both field size and expected SNS (Table 5). More sampling is cost-effective on larger fields because of the potential for larger total yield and profit. More sampling is required when the expected SNS is large because this leads to large within field variability of SMN. On the 60 ha fields with expected SNS of 275 kg/ha, 23 cores lead to the largest profit. This is slightly greater than the recommendation of 20 cores on variable fields from the HGCA *Nitrogen for winter wheat – management guidelines.* These guidelines suggest that where sampling is conducted at least 15 cores should be extracted. However the sampling experiments suggest that a smaller number of soil cores might be beneficial on small fields with expected SNS of less than 50 kg/ha.

When there is no prior knowledge of the SNS for the target area the optimal number of cores increases from 7 cores for a 10 ha field up to 18 cores for a 60 ha field. The extra profits from sampling are much greater than when there is prior knowledge of SNS (Table 6). In these 'no prior knowledge' tests a different SNS value was used for each simulated realization. These values were sampled from a pdf which reflected the variability of field mean SNS values within the datasets used in this project.

In the tests, the measurement of SMN leads to increased profit for all field sizes and expected levels of SNS in comparison to the profit which would have resulted from assuming that the SNS was equal to the mode of the prior knowledge pdf. Using the modal value of an assumed pdf is equivalent to trusting uncertain prior information regarding the soil-nutrient status from sources such as the Field Assessment Method. We emphasise that the models used in the tests made several assumptions about the relationship between SNS and optimal fertiliser N and the relative size of SNS and SMN errors. These assumptions were detailed in Section 3.2.5. If future studies provide more detailed information about these relationships then the models may be adjusted and the simulation tests repeated. Also, increased profitability from SMN measurement might not arise if the prior knowledge of SNS was more certain. In terms of increased profit the largest benefits of sampling occur for fields with expected SNS of around 175 kg/ha (Tables 3-4). This is because at this SNS level both under- and over-estimates of the SMN lead to substantial inefficiencies and loss of profit. In contrast, for SMN of greater than 250 kg/ha only a small amount of N should be added and therefore if the SNS is overestimated this will have little effect on the additions and hence the profit. Similarly for small SNS, close to the maximum additions will be recommended and therefore an underestimate of the SMN will be of little consequence.

The use of a 10-ha barometer field leads to a decrease in the expected profit (Table 8). However for target areas of 30 ha or less and/or expected SNS of 100 kg/ha or less this expected loss is less than £50 and possibly not large enough to justify visiting the entire target area. The losses are more substantial in other circumstances when the practitioner should consider either using a larger barometer field or sampling throughout the target area.

Table 3. Additional profit per hectare from measuring SMN on ‘W’ across target area (£/ha).

Target	**Expected SNS**					
	25 kg/ha	**75 kg/ha**	**125 kg/ha**	**175 kg/ha**	**225 kg/ha**	**275 kg/ha**
5 ha	0.68	4.74	9.66	11.79	9.61	4.66
10 ha	5.58	10.77	14.97	17.04	15.16	10.20
20 ha	8.49	13.14	18.28	20.41	17.91	13.46
30 ha	9.37	13.98	19.01	20.93	19.26	13.93
60 ha	10.53	15.39	20.25	22.07	19.67	15.09

Table 4. Total additional profit from whole target area upon measuring SMN on ‘W’ across target area (£).

Target	**Expected SNS**					
	25 kg/ha	**75 kg/ha**	**125 kg/ha**	**175 kg/ha**	**225 kg/ha**	**275 kg/ha**
5 ha	3.40	23.71	48.32	58.93	48.05	23.28
10 ha	55.83	107.65	149.67	170.38	151.57	101.99
20 ha	169.74	262.80	365.66	408.23	358.28	269.17
30 ha	281.07	419.43	570.15	627.88	577.71	418.02
60 ha	631.71	923.10	1214.76	1324.20	1180.41	905.37

Table 5. Optimal number of cores on 'W' when sampling SMN throughout the target area.

Target	Expected SNS					
	25 kg/ha	**75 kg/ha**	**125 kg/ha**	**175 kg/ha**	**225 kg/ha**	**275 kg/ha**
5 ha	3	4	4	5	6	6
10 ha	4	6	6	8	8	9
20 ha	5	8	8	10	10	13
30 ha	7	10	12	12	14	18
60 ha	10	14	15	18	23	23

Table 6. Optimal number of cores and additional profit from sampling when measuring SMN on 'W' throughout the target area without prior knowledge of SNS.

Target	Number of cores	Extra profit £/ha	Extra profit £
10 ha	7	43.89	438.93
20 ha	10	47.46	949.11
30 ha	14	48.40	1452.01
60 ha	18	49.96	2997.55

Table 7. Loss resulting from sampling SMN on 10 ha barometer field rather than throughout target area (£/ha).

Target	Expected SNS					
	25 kg/ha	**75 kg/ha**	**125 kg/ha**	**175 kg/ha**	**225 kg/ha**	**275 kg/ha**
5 ha	na	na	na	na	na	na
10 ha	0.00	0.00	0.00	0.00	0.00	0.00
20 ha	0.46	0.63	1.36	1.88	1.87	2.69
30 ha	0.50	0.56	1.26	1.90	2.45	2.45
60 ha	0.75	0.59	1.26	1.71	1.87	3.65

Table 8. Lost profit resulting from sampling SMN on 10 ha barometer field rather than throughout target area (£).

Target	Expected SNS					
	25 kg/ha	**75 kg/ha**	**125 kg/ha**	**175 kg/ha**	**225 kg/ha**	**275 kg/ha**
5 ha	na	na	na	na	na	na
10 ha	0.00	0.00	0.00	0.00	0.00	0.00
20 ha	9.25	12.59	27.21	37.51	37.50	53.84
30 ha	14.88	16.94	37.86	57.08	73.46	73.42
60 ha	44.89	35.44	75.81	102.37	112.44	219.21

3.3.8. Delineation of hotspots and coldspots by sequential sweep-out methods

Some exemplar results of the sweep-out experiments are shown in Tables 9 and 10. Each experiment consists of 90 cores. The number of analyses required to delineate between the above- and below-threshold cores decreases as the threshold increases relative to the underlying distribution of the nutrient. For example, if the threshold is the 99.9% confidence limit of the distribution of P at Welford Farm then 55 analyses are required to complete the sweep-out method. This is a saving of 35 analyses upon analysing each individual core. Smaller savings are seen at other farms. The farms where the sweep-out methods leads to large savings are those where the distributions of the nutrients are most skewed (see Figures 8 and 9). The savings generally disappear if the threshold is less than the 95 % confidence limit of the distribution. Hence the method is not cost-effective for finding coldspots and the saving for finding hotspots does not compensate for the extra complexity of the method and the multiple phases of analysis that are required.

Table 9. Number of laboratory analyses required to complete sweep-out method for phosphorus observed at 90 sites for different critical thresholds.

Farm	Threshold Quantile				
	0.5	**0.9**	**0.95**	**0.99**	**0.999**
Hamilton	102.7	98.6	95.9	90.0	77.1
Kemble	102.0	92.6	87.8	72.4	70.3
Roxhill	102.7	97.6	95.3	90.1	82.3
Welford	101.9	93.7	90.0	79.1	55.9

Table 10. Number of laboratory analyses required to complete sweep-out method for potassium observed at 90 sites for different critical thresholds.

Farm	Threshold Quantile				
	0.5	**0.9**	**0.95**	**0.99**	**0.999**
Hamilton	103.0	99.2	98.6	97.2	94.6
Kemble	102.6	96.6	94.3	89.5	85.2
Roxhill	102.9	98.0	96.3	92.2	87.2
Welford	102.3	95.8	93.0	86.1	74.5

3.4. Discussion

The simulation tests conducted within this project demonstrated that optimized sample designs can lead to more accurate estimates of field mean nutrient concentrations than the recommended ‘W’ design. For fewer than 10 cores the sampling errors on using the optimized design were only slightly smaller than those from the ‘W’. The improvements from the optimized designs were more evident when the number of cores to be bulked exceeded 15. Once this sampling intensity had been reached negligible further reductions in sampling errors were achievable with the ‘W’ design. More significant reductions would have required cores to be extracted away from the ‘W’.

The soil nutrients observations within the datasets used in this study exhibited non-Normal variation caused by local hotspots. The presence of these hotspots increases the expected sampling errors and it is therefore important to account for them when making an assessment of sampling requirements. For P and K the potential for hotspots was included in the model of variation through the assumption of a GEV distribution. The variability of SMN was assumed to increase with the expected value.

When the effectiveness of the sample designs was considered in terms of the long term management of P and K stocks or the cost-effectiveness of measuring SMN the benefit from using the optimized design instead of the ‘W’ was small. This was because the errors associated with a ‘W’ design with the appropriate number of cores were small enough that other uncertainties in the estimation of nutrient uptake by the crop limited the quality of the fertiliser management recommendations. These results suggest that the ‘W’ should continue to be recommended because of the simplicity of implementing the design. However if future fertiliser management recommendations require more accurate estimates of soil nutrient status then the use of the optimized designs should be reassessed.

The performance of the stratified and expert knowledge sample designs was comparable to the ‘W’ however these designs were more complicated to implement. Better quality expert knowledge could lead to an improvement in these designs but under the type of nutrient variation assumed in this project they would not outperform the optimized scheme. The clustered design performed very poorly which is a clear warning of the problems that can arise from haphazard sampling schemes.

The tests on P and K sampling suggested that bulking 10 cores every four years would be sufficient to ensure that soil nutrient levels remained in RB209 Index 2 for more than 97.5 % of years. This is substantially fewer cores than the currently recommended 25. The current recommendations are designed to ensure that the sampling errors for these nutrients do not exceed arbitrary thresholds rather than considering the implications of the errors. The sampling requirements for P and K are insensitive to field size. This is because there were two competing effects in larger fields which cancelled each other out. The cores were taken from more widely dispersed locations and were therefore on average less correlated, but more variation occurred within the larger field.

For SMN, 10-15 cores on a ‘W’ were adequate for most fields. The sampling requirements increased with field size and expected SNS. More than 10 cores were warranted for fields of greater than 20 ha where the expected SNS was high (defined as >160 kg/ha in RB209). The most variable 60 ha fields considered required 23 cores but such fields rarely occur in the UK. Fewer than ten cores were cost-effective on some small fields where SNS was likely to be small or moderate. However we envisage that the FAM would be used in such circumstances in preference to direct measurements. Therefore we generally recommend that bulked samples should be formed from 10-15 cores. More than 15 cores would only be required for exceptionally large or variable fields. The measurement of SMN was most beneficial when the expected SNS was close to 175 kg/ha because at such concentrations both under- and over-estimates of N status have substantial implications for profitability.

The use of 10-ha barometer fields was reasonably effective for management of neighbouring fields smaller than 60 ha with expected SMN of less than 100 kg/ha. However for larger or more variable fields the use of larger barometer fields or sampling of the entire target area should be considered. It should be noted that the barometer field experiments were based on a fixed expected SNS value for the whole target area. We have not addressed the problem of how such zones of relatively uniform SNS should be delineated. More detailed assessments of the profitability of SMN measurement in various circumstances and of the relative effectiveness of the Field Assessment Method have been conducted by Kindred et al. (2011).

In contrast to the P and K scenarios, the SMN experiments required a number of assumptions about the price of wheat and fertiliser and the practitioner's prior knowledge of the SNS. The sampling recommendations should be recalculated if any of the parameters change drastically or if the relationship between SMN and SNS becomes better understood. In would be possible to replace the pdf of the practitioner's prior knowledge of SNS with one that reflects the expectation and uncertainty of SNS according to the Farm Assessment Method. This would require a comprehensive dataset of the type collected by Kindred et al. (2011) which observes SNS under different conditions.

The framework developed in this project was the first attempt to relate sampling requirements to the implications of fertiliser management decisions. Such a framework can lead to improved recommendations in which practitioners can have more confidence. The framework assumes that current fertiliser recommendations are appropriate. It could be extended to test the effectiveness of the recommendations but this would require further field experiments to be conducted to determine the relationships between nutrient levels prior to application, fertiliser additions and crop yields. The uncertainty associated with these relationships must be quantified if a thorough cost-effectiveness study is to be conducted. The experiments should be designed to explain how relationships at the scale of experiment plots can be up-scaled to the field scale. A similar framework could be used to assess the cost-effectiveness of precision agriculture techniques that vary fertiliser additions within the field.

The sequential sweep-out method proved to be an inefficient way to delineate nutrient hotspots and coldspots within a field. This method has previously been effective at determining areas which require remediation in studies of soil pollution. In these studies the critical thresholds are often an order of magnitude greater than the background concentrations of the pollutant. Such extreme behaviour is not seen in our models of soil nutrient variation.

This project has demonstrated that a mathematical framework can be used to make rational recommendations about sampling requirements for soil nutrient management. The framework could be expanded to improve the soil nutrient management recommendations themselves by integrating new experimental evidence or to develop decision support systems which could operate on specific farms and adapt, over multiple seasons, to the requirements of the farm.

There are many complexities and sources of uncertainty in the soil nutrient management system. In the case of N management these include the relationships between observed SMN and realized SNS on different soil types and under different climatic conditions. Such issues have been addressed in an HGCA project (HGCA Project 3425 *Establishing best practice for SNS estimation*) which investigated these relationships at the plot-scale. What is less well understood is how this knowledge can be integrated to the field or management scale to form rational fertiliser recommendations. The discrepancy between the scales at which soil processes are understood and the scales at which management measures can be implemented is a familiar issue in other environmental systems such as the emissions of ammonia from the soil (Corstanje et al., 2008). It is not currently clear whether the uncertainty observed in the plot-scale SNS experiments is because of random within-field variation or because of the variation in the nutrient requirements of different fields or farms. If it is the former then the effects of this variation will disappear upon up-scaling. However if it is the later then the variation of nutrient requirements must be understood and accounted for before cost-effective recommendations can be devised.

A number of specific issues relevant to soil N management could be addressed through expansions of the mathematical framework presented in this project. These include

1. How knowledge of soil processes at the plot-scale can be integrated to management recommendations at larger scales.
2. How knowledge of soil type, climate, previous crops can be combined with SMN measurements to efficiently determine the N requirements of a particular management area.
3. When management areas are sufficiently uniform for the use of barometer fields to be efficient.
4. When precision agriculture techniques and within-field variation of nutrient inputs is cost effective.

These problems can only be addressed if additional data are collected which explore how the variation in the soil N system is divided between the regional, farm and within-field scales. Our current understanding can inform how new observations should be divided between these scales and how they should be stratified over different soil types and climate regions.

The question of integrating knowledge from different sources could be addressed by an approach used in this project. We formed plausible probability distributions of the knowledge of SMN content prior to sampling using expert knowledge. The data from HGCA Project 3425 could be used to determine the uncertainty of the Field Assessment Method (DEFRA, 2010) and to describe it by a probability distribution. Then this probability distribution could be updated after SMN measurements have been made by Bayesian statistical techniques (Cressie & Wikle, 2011). If such techniques were used in a decision support system it would also be possible to include information from previous seasons.

The results from this project suggested that barometer fields could be cost-effective but the issue is to understand when the management zones are sufficiently homogenous. This could be achieved through a Bayesian decision support system which utilizes observations from previous systems and suggests locations for new SMN measurements.

Precision agriculture systems for within-field management will require a detailed understanding of the variation of the key relationships in the within-field soil N system. Sweep-out methods proved to be inefficient with within field soil nutrient mapping. Therefore we suggest that the use of efficient sampling and interpolation techniques (e.g. Marchant & Lark, 2007b) to map the SMN variation should be explored.

3.5. Conclusions

- Statistical tools such as the framework developed in this project are required to rationally determine the requirements for soil nutrient sampling and to determine when it is cost effective to measure nutrients.

- Optimized sample designs perform better than the ‘W’ design for large sample sizes and their use should be considered if future management recommendations require more accurate soil nutrient information.

- The ‘W’ design is sufficient for current fertiliser management recommendations.

- P and K sampling recommendations can be reduced to 10 cores per field every four years from the current 25 cores.

- For SMN 10-15 cores are adequate for most fields. More than 10 cores should be used for large fields (>20 ha) or if SNS is expected to be high (>160 kg/ha). More than 15 cores are only appropriate for very large fields (>30 ha) where SNS is expected to be high.

- Ten-hectare barometer fields can be effective to infer SNS over similar neighbouring target areas of less than 60 ha when the expected SNS is less than 100 kg/ha. In other circumstances more widespread sampling should be considered.

- The recommendations for N are based on assumed costs and simplifications of the relationship between SMN and SNS. These recommendations should be recalculated if costs change drastically or if the relationships become better understood.

- Alternative methods of estimating SNS, such as the Field Assessment Method should also be considered.

- Sweep-out methods are inefficient at delineating soil-nutrient hotspots and coldspots and should not be adopted.

3.6. Acknowledgements

The authors are grateful to the members of the project steering group namely Simon Griffin (SOYL), James Holmes (HGCA), Stuart Knight (NIAB TAG) and Peter Taylor (AICC Independent Agronomist, Peter Taylor Agronomy Ltd.) who provided helpful advice and feedback throughout this study. The authors would also like to acknowledge the various individuals and organisations that made data available for this project: Simon Griffin (SOYL), Samantha Baxter, Zoe Frogbrook and Margaret Oliver (The University of Reading), Johnny Johnson (Rothamsted Research) and Roger Sylvester-Bradley (ADAS).

3.7. References

Baxter, S.J. 2002. The spatial variation of plant available nitrogen within arable fields. PhD Thesis, University of Reading.

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

Burgess, T.M., Webster, R. 1984. Sampling and bulking strategies for estimating soil properties in small regions. *Journal of Soil Science*, **35**, 127-140.

Chen, Z., Bai, Z., Sinha, B.K. 2003. *Ranked Set Sampling: Theory and Applications*. Springer-Verlag.

Cordova, S.C. 2011. Spatial variability of soil organic matter fractions in arable and grassland soils – implications for soil N supply. PhD Thesis, University of Reading.

Corstanje, R., Kirk, G.J.D., Pawlett, M., Read, R., Lark, R.M. 2008. Spatial variation of ammonia volatilization from soil and its scale-dependent correlation with soil properties. *European Journal of Soil Science*, **59**, 1260-1270.

Cressie, N., Wikle, C.K. 2011. *Statistics for spatio-temporal data*. Wiley, New Jersey.

DEFRA, 2010. *The Fertiliser Manual (RB209)*. 8th ed. The Stationery Office, London.

Haskard, K.A., Welham, S.J., Lark, R.M. 2010. Spatial tempering to model non-stationary covariances of nitrous oxide emissions from soil using continuous or categorical explanatory variables at a landscape scale. *Geoderma*, **159**, 358-370.

Johnston, A.E., Goulding, K.W.T. 1988. Rational potassium manuring for arable cropping systems. *Journal of the Science of Food and Agriculture*, **46**, 1-11.

Johnston, A.E., Goulding, K.W.T. 1990. *The Use of Plant and Soil Analyses to Predict the Potassium Supplying Capacity of Soil*. In: Proceedings of 22nd Colloquium of the International Potash Institute. Bern, Switzerland. p 177-204.

Johnston, A.E, Poulton, P.R. 1992. *The role of phosphorus in crop production and soil fertility: 150 years of field experiments at Rothamsted, United kingdom*. In: Schultz, J.J. (Ed). Proceedings of International Workshop on Phosphate Fertiliser and the Environment. International Fertiliser Development Centre, Tampa, Florida, USA, pp. 45-63.

Johnston, A.E. 1986. *Potassium fertilization to maintain a K-balance under various farming systems.* p. 199-226. In Nutrient balances and the need for potassium. Proc. 13th IPI-Congress. Int. Potash Inst., Worblaufen-Bern, Switzerland

Johnston, A.E. 2005. *Phosphorus nutrition of arable crops*. In: Sims, J.T. and Sharpley, A.N. (eds.) Phosphorus: Agriculture and the Environment. Agronomy Monograph No. 46, ASA-CSSA-SSSA, 495-519.

Kindred, D., Knight, S., Berry, P., Sylvester-Bradley, R., Hatley, D., Morris, N., Hoad, S., White, C. 2011. *Establishing Best Practice for Estimation of Soil N Supply*. HGCA Report for Project 3425. Stoneleigh: HGCA.

Lark, R.M., Milne, A.E., Addiscott, T.M., Goulding, K.W.T., Webster, C.P., O'Flaherty, S. 2004. Scale- and location-dependent correlation of nitrous oxide emissions with soil properties: an analysis using wavelets. *European Journal of Soil Science*, **55**, 611-627.

Lark, R.M., Catt, J.A. & Stafford, J.V. 1998. Towards the explanation of within-field variability of yield of winter barley: soil series differences. *Journal of Agricultural Science*, **131**, 409–416.

Marchant, B.P., Lark, R.M. 2007a. Estimation of linear models of coregionalization by residual maximum likelihood. *European Journal of Soil Science* **58**, 1506-1513.

Marchant B.P., Lark R.M., 2007b. Optimized sample schemes for geostatistical surveys. *Mathematical Geology*, **39**, 113-134.

Marchant, B.P., Lark, R.M., Wheeler, H.C. 2005. *Developing methods to improve sampling efficiency for automated soil mapping*. London, HGCA Project Report No. 364.

Marchant B.P., Saby, N.P.A., Jolivet, C.C., Arrouays, D., Lark, R.M. 2011. Spatial prediction of soil properties with copulas. *Geoderma*, **162**, 327-324.

McBratney, A.B., Pringle, M.J. 1999. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. *Precision Agriculture*, **1**, 125-152.

Milford G F J and Johnston A E 2007 *Potassium and nitrogen interactions in crop production*. Proceedings No 615 - International Fertiliser Society, 1-22.

Oliver, M.A., Carroll, Z.L., 2004. *Description of spatial variation in soil to optimize cereal management*. London, HGCA Project Report No. 330.;

Oliver, M.A., Frogbrook, Z.L., Webster, R., Dawson, C.J. 1997. *A rational strategy for determining the number of cores for bulked sampling in soil*. In J.V. Stafford (Ed.), Precision agriculture '97. Volume I, spatial variability in soil and crop (pp. 155-162). Oxford: BIOS Scientific Publications,

Syers, J. K., Johnston, A. E., Curtin, D., 2008. Efficiency of soil and fertilizer phosphorus use: reconciling changing concepts of soil phosphorus behaviour with agronomic information. Food and Agriculture Organization of the United Nations, Rome, Bulletin 18. ftp://ftp.fao.org/docrep/fao/010/a1595e/a1595e00.pdf

Sylvester-Bradley R., Kindred D.R., Blake J., Dyer C.J., Sinclair A.H. 2008. *Optimising fertilizer nitrogen for modern wheat and barley crops*. London: HGCA. Project report No. 438.

Van Groenigen, J.W., Siderus, W, Stein, A. 1999. Constrained optimisation of soil sampling for minimisation of the kriging variance. *Geoderma*, **55**, 239-255.

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. *Computers and Geosciences*, **36**, 1261-1267.

Webster, R., Oliver, M.A. 2007. *Geostatistics for Environmental Scientists*. 2nd ed. John Wiley & Sons Ltd, Chichester.