

Original papers

Assessing sampling designs for determining fertilizer practice from yield data

S.E. Muhammed^a, B.P. Marchant^b, R. Webster^a, A.P. Whitmore^a, G. Dailey^a, A.E. Milne^{a,*}^a Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK^b British Geological Survey, Keyworth, Nottingham NG12 5GG, UK

ARTICLE INFO

Article history:

Received 18 December 2015

Received in revised form 2 February 2017

Accepted 5 February 2017

Available online 17 February 2017

Keywords:

Yield maps

Phosphorus

Sampling

Plant nutrients

Variograms

ABSTRACT

Many farmers sample their soil to measure the concentrations of plant nutrients, so as to decide how much fertilizer to apply. Now that fertilizer can be applied at variable rates farmers want to know whether maps of nutrient concentration made from grid samples or of field subdivisions (zones within their fields) are merited: do such maps lead to greater profit than would a single measurement on a bulked sample for each field when all costs are taken into account? We have examined the merits of grid-based and zone-based sampling strategies over single field-based averages using continuous spatial data on wheat yields at harvest in six fields in southern England and simulated concentrations of phosphorus (P) in the soil. We have taken into account current prices of wheat, P fertilizer and sampling and laboratory analysis. Variograms of yield provide guides for sampling. We show that where variograms have large variances and long effective ranges grid-sampling and mapping are feasible and have large probabilities of being cost-effective. Where effective ranges are short, sampling must be dense to reveal the spatial variation and be expensive, and variable-rate application of fertilizer is likely to be impracticable and almost certainly not cost-effective. We found zone-based sampling was less likely to be cost effective in a similar situation when the management zones were poorly correlated to P concentrations. Crown Copyright © 2017 Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

We have known for more than 150 years that shortages of phosphorus and potassium in the soil limit crop growth. Thousands of experiments have been done to estimate the responses of crops to additions of these nutrients and to calculate the needs for fertilizers. Farmers now want to use these results to vary their applications within fields. Perhaps surprisingly, there are few reports linking variation in the concentrations of these elements in the soil to yields within individual fields on commercial farms. There are examples, however, where positive correlations were found for cereals (Frogbrook et al., 2002) and pastures (McCormick et al., 2009; Serrano et al., 2011). Many farmers in the United Kingdom sample their soil every four years to measure the nutrients, in particular phosphorus (P) and potassium (K), in the soil so as to decide how much fertilizer to apply to their crops. Sampling is often done at points on a ‘W’ shape across each field, and then individual samples are bulked before analysis in the laboratory (PDA, 2011). Even though this sampling configuration does not follow the principles of design-based statistics, it has been widely adopted by farmers

as the results are generally no less accurate than those obtained from stratified random sampling (Marchant et al., 2012). By bulk-ing the sample, however, all information on the variation of the nutrients across the field is lost, and so any local deficiency or excess is obscured. If a farmer wants to map the variation in nutrients across a field, so that fertilizer could be adjusted spatially, for example, then he or she should ideally sample the soil on a grid (with perhaps some additional points at closer spacings) and measure the nutrient content in each sample of soil separately (Mallarino and Wittry, 2004; Sawchik and Mallarino, 2007; Fu et al., 2013). Kriging, which makes best use of such data, could then be used to map the variation in nutrients (Kravchenko, 2003; Webster and Oliver, 2007). To kriging, however, one needs an accurate estimate of the variogram or covariance function for the variable of interest, and for that at least 100 measurements are needed (Oliver and Webster, 2014). This creates a problem because measuring the concentration of P or K of each sample in the laboratory is costly.

Often the reason for the variation in yield across a field will be obvious to the farmer. For example, the farmer might know that a particular part of the field is prone to drought and that this is a major cause of the variation in yield. If the farmer suspects a local nutrient deficiency then it would make sense to divide the land

* Corresponding author.

E-mail address: alice.milne@rothamsted.ac.uk (A.E. Milne).

Nomenclature

$\mathbf{x} \equiv \{x_1, x_2\}$	spatial coordinates in two dimensions	<i>variogram parameters</i>	
y	yield of crop	c_0	nugget variance
\bar{y}	standardized yield	c_1	variance of spatially correlated structure
y_r	realized yield	a	distance parameter
y_0	target yield		
z	quantity of phosphorus, P	<i>Costs</i>	
z^*	realization of z	G_{wheat}	price of grain, assumed to be £150 t ⁻¹
Z_{fert}	quantity of fertilizer P	G_{fert}	price of P fertilizer, assumed to be £0.31 kg ⁻¹
Z_{soil}	initial quantity of P in the soil	G_{sample}	cost of soil analysis, assumed to be £5 sample ⁻¹
Z_{total}	$Z_{\text{fert}} + Z_{\text{soil}}$		
k	the number of classes in the k -means classification		

into management zones and estimate the nutrient status for each zone separately.

So, which of these three sampling approaches, commonly used by farmers and advisors, should be adopted in any particular situation to apply P and K fertilizer spatially? It is a question taxing agricultural advisers who want to advise farmers on best practice for within-field sampling for plant nutrients—see, for example, Oliver and Kerry (2013), Mylavarapu and Wonsok (2014) and Hawkins et al. (2016). The grid-based approach should result in the most accurate prediction of fertilizer requirement (Mallarino and Wittry, 2004; Sawchik and Mallarino, 2007). But the money saved by varying the application of the fertilizer locally within fields to match the requirement of the crop might be less than the cost of sampling and measurement (Fleming et al., 2000; Mallarino and Wittry, 2004). The balance of the two, and therefore the merit of the approach, depend on the magnitude and variation of the nutrient content of the soil (Sawchik and Mallarino, 2007). These two variables can be determined accurately only from measurements made on samples. The variation can, however, be assessed indirectly from crop yields. Many farmers are already monitoring yields as they harvest their crops, and variation in the data they record is in many instances a reflection of the variation in the availability of the nutrients in the soil (Stafford et al., 1999; Diker et al., 2004; Flowers et al., 2005). If the variation in yield is small then it is unlikely that the nutrients vary substantially. We know that factors other than nutrient supply can cause large variations in yield. Nevertheless, nutrient supply does dominate yield variation in many cases, and for present purpose we proceed on that assumption. In such circumstances Lark et al. (2003) proposed metrics based on the variogram of yield data (which captures the magnitude of variation of the yield) to assess the scope for variable rate management. Their approach was to use the metrics as factors in a decision tree designed to determine the potential for variable rate management.

In this context we aimed (i) to compare the merits of measuring plant nutrients by three sampling schemes (whole fields, zones within fields and grids) and (ii) to assess the extent to which yield maps might be used to determine the most cost-effective sampling. Which of the above sampling approaches is suitable for a given situation depends on a farmer's profit margin over the cost of fertilizer and soil sampling. It is not possible to do such a comparison in the field because the test requires perfect knowledge of how the nutrients vary across the field, we therefore resorted to simulation. In the approach presented here we simulated the variation in nutrients across fields from geostatistical models of the nutrients and used these to test the cost effectiveness of each sampling scheme. We modelled the associated yield variation for each realization and explored the use of the metrics of the yield variogram to decide which sampling strategy was likely to be most cost-effective.

2. Method

2.1. Data

We collated yield data, denoted y , from monitors on board combine harvesters and measurements of extractable P, denoted z , in the soil for six fields on a farm near Newbury, England (Table 1). Soils were medium to heavy textured with slight to moderate stoniness. The yields were of winter wheat from the seasons 2001 to 2011 recorded at approximately 20-m intervals by the monitor. The measurements of Olsen extractable P (sodium bicarbonate extract at pH 8.2) were made on a 100-m grid at 24–36 locations across each field. Fig. 1 displays yield maps for the six fields for a single year.

2.2. Determining management zones

Several methods have been devised for creating management zones (Oliver and Webster, 1989; Fleming et al., 2000; Diker et al., 2004; Flowers et al., 2005; Zhang et al., 2009). The multivariate technique of Dray et al. (2006) has most recently been applied successfully by Peralta et al. (2015) for wheat farming and by Córdoba et al. (2016) for grain cropping more generally.

We created management zones from the yield data using a spatially smoothed version of a fuzzy k -means classification devised by Lark (1998) (see also Milne et al., 2012). The data for the classification consisted of yields of wheat for p years at the n nodes of a square grid at intervals of 10 m. We denote the grid coordinates as $\mathbf{x} \equiv \{x_1, x_2\}$ and the yields in the p years as $y_1(\mathbf{x}), y_2(\mathbf{x}), \dots, y_p(\mathbf{x})$.

From these data we created a classification by a 'hard' k -means algorithm. We standardized each of the $y_j, j = 1, 2, \dots, y_p$ to zero mean and variance of 1, denoted \bar{y}_j . For this method we choose k , the number of classes. We divided the whole set of the standardized data into that number of classes in such a way as to minimize the trace or determinant of the within-classes variance-covariance matrix. Each grid node then belonged to one and only one of the k classes, and in general it resembled other members of its class more than the members of the other classes.

For fuzzy k -means classification we first define a measure of dissimilarity, d , between an individual node i and a class q . A convenient measure is the Euclidean distance in the vector space:

$$d_{iq} = \sqrt{\sum_{j=1}^p (\bar{y}_{ij} - \bar{y}_{jq})^2}, \quad (1)$$

where \bar{y}_{ij} is the standardized yield at node i in the j th year, and \bar{y}_{jq} is the mean of \bar{y} in class q in that year.

Table 1
Soil type and summary statistics for measured P in the six study fields.

Field name	P (mg l ⁻¹)			Size of sample
	Range	Mean	Standard deviation	
Bassdown	13–31	19.27	3.99	28
Easton Right	16–25	21.7	4.14	24
Gravelpit	37–165	95.04	36.67	23
Limes	19–52	28.04	6.79	26
Mantclose	12–35	16.9	6.07	24
Roodhill	21–163	83.19	40.41	20

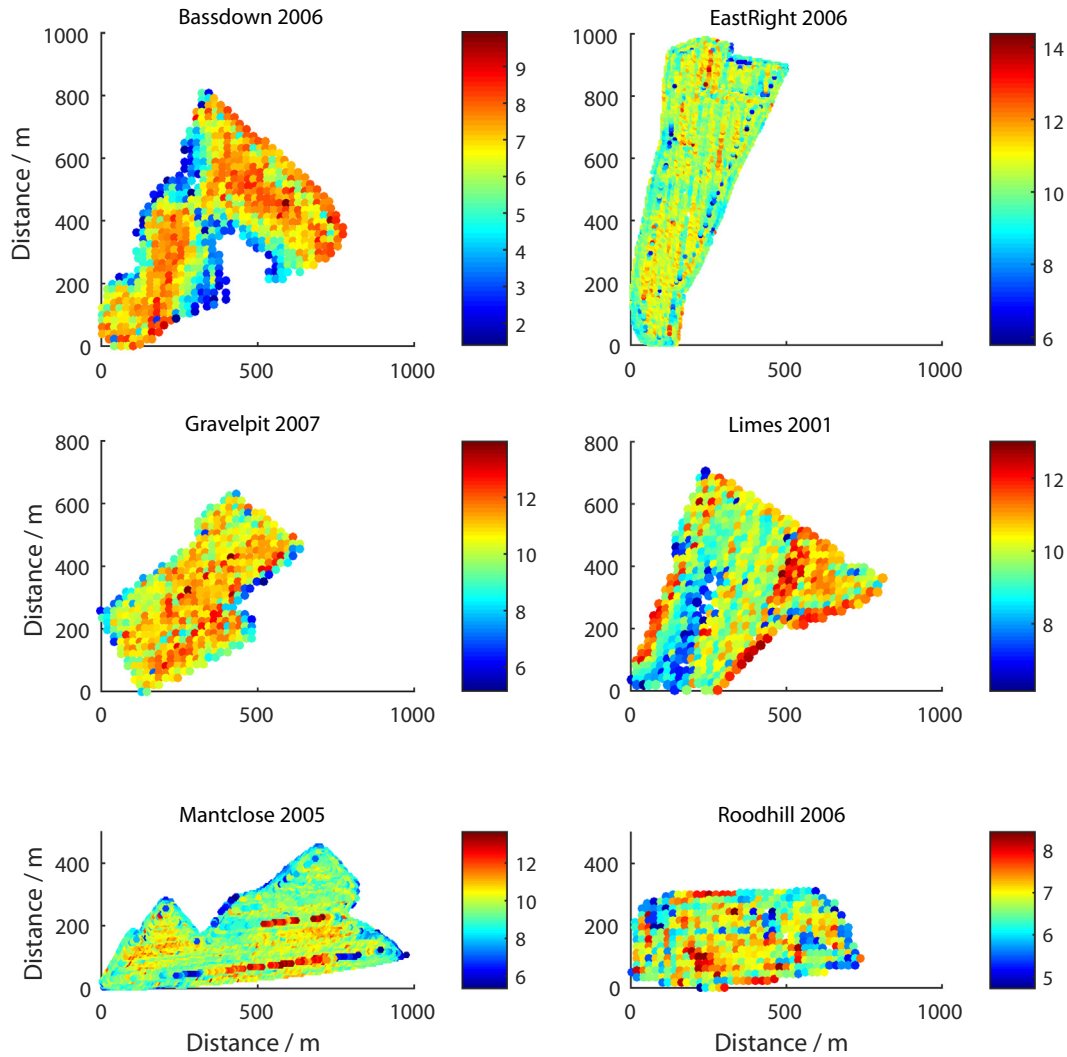


Fig. 1. Maps of winter wheat yield from combine harvester yield monitors (t ha⁻¹).

For this method we assume that each node belongs to some degree to every class, and we create a classification by minimizing a pooled ‘belongingness’, say b :

$$b = \sum_{q=1}^k \sum_{i=1}^n d_{iq}^2 u_{iq}^\omega, \tag{2}$$

in which u_{iq} is the degree of membership of node i to class q , and ω is the fuzziness parameter.

The membership across all classes must sum to 1:

$$\sum_{q=1}^k u_{iq} = 1. \tag{3}$$

The parameter ω must lie between 1 (in which case we obtain a hard classification) and 2. We set $\omega = 1.25$ to create our classifications.

As above, we needed to choose k . We did this by experimenting with several values between 2 and 5 (the most that a farmer is likely to distinguish). For each class we then computed the normalized classification entropy $\xi(k)$, proposed by Dunn (1977):

$$\xi(k) = -\frac{1}{\ln k} \sum_{q=1}^k \sum_{i=1}^n \frac{1}{n} u_{iq} \ln u_{iq}. \tag{4}$$

We then plotted $\xi(k)$ against k and identified the value of k at which $\xi(k)$ falls below the overall trend. This was the value we chose. The

procedure is analogous to that in hard k -means classification in which the trace of the within-classes variance-covariance matrix or its determinant is plotted against k .

The next step in the zonation is to smooth the classes. It turned out that the distributions of the memberships of the nodes were strongly bimodal, and so, following Lark (1998), we transformed them with the symmetric log-ratio to unimodal distributions. We smoothed the transformed membership (denoted \tilde{u}_{iq}) using a weighted average of the transformed memberships in circular neighbourhoods, R , of radius r :

$$\tilde{u}_{iq}^* = \sum_{j \in R} w(i,j) \tilde{u}_{iq}. \tag{5}$$

Like the original memberships, the transformed memberships must lie in the range 0–1 and must sum to 1. This means that the weights in R must sum to 1.

The weights were derived from the variogram, Eq. (6). We can write a simple bounded model in general as

$$\gamma(\mathbf{h}) = c_0 + c_1 f(\mathbf{h}). \tag{6}$$

In this equation c_0 is a spatially uncorrelated variance, the ‘nugget variance’, corresponding to white noise and c_1 is the spatially correlated component of variance; their sum, $c_0 + c_1$, constitute the ‘sill’. The function $f(\mathbf{h})$ defines the form of the variogram and contains a distance parameter. The weights are obtained as

$$w(i,j) = \frac{1 - f(\mathbf{h}_{ij})}{\sum_{j \in R} \{1 - f(\mathbf{h}_{ij})\}} \text{ for all } j \in R, \tag{7}$$

where \mathbf{h}_{ij} is the separation in distance and direction, the lag, between nodes i and j . Note that only the functional form of the variogram and its distance parameter affect the weights; neither c_0 nor c_1 do so. The neighbourhood R defines the region over which the membership values are smoothed, unless the variogram reaches its sill within it. In the latter case the effective range of the variogram defines the smoothing region.

The farmer, of course, must have a hard classification; for practical management he must have each position in the field belonging to one class and one class only. So the final stage in the zonation is therefore to assign each node to the class for which its smoothed membership is greatest.

Note that the size of R affects the results. The larger it is the greater is the smoothing. If R is small then the classification is likely to be too fragmented; if it is too large then the memberships will be smoothed too much and the final classes not sufficiently homogeneous. Lark (1998) proposed a coherency index to identify an appropriate radius for R defined as

$$H = \frac{\eta_\alpha}{\sum_{q=1}^k \psi_q^2}, \tag{8}$$

where η_α is the proportion of pairs of nodes within a distance α that belong to the same class, and ψ_q is the proportion of nodes that belong to class q . The larger is the value of H , the more spatially coherent are the classes. We chose $\alpha = 10\sqrt{2}$ m (the distance between two points on the diagonal of our 10-m grid) so that we were effectively comparing each node with its neighbours along the rows, columns and diagonals on the grid.

2.3. Modelling phosphorus in soil

For each of the six fields, we used the measurements, z , of P to create realizations of P concentrations with plausible means and spatial variations that could differ from one management zone to another. Our data came from well managed fields, but we wanted our simulated values of P to limit the yield. Therefore we scaled the

measured data by 0.5 before fitting the models of spatial variation for all fields except Mantclose. We used a similar approach to that described by Marchant et al. (2012). First we standardized the measurements, $z_i, i = 1, 2, \dots, n$, to have a variance = 1 by dividing by the standard deviation, s , of the data to give values $\tilde{z}_i, i = 1, 2, \dots, n$. Then we characterized the mean and spatial variation within each of the k zones of the field by fitting a linear mixed model to the transforms, $\tilde{\mathbf{z}}$:

$$\tilde{\mathbf{z}} = \mathbf{M}\boldsymbol{\beta} + \boldsymbol{\eta}, \tag{9}$$

where \mathbf{M} is an $n \times k$ fixed effects design matrix which permits the mean concentrations to differ between zones. If the i th value of $\tilde{\mathbf{z}}$ is in zone j then the element in i th row of the j th column of \mathbf{M} is 1, otherwise it is 0. The vector $\boldsymbol{\beta}$ is of length k and contains the coefficients of the fixed effects (i.e. the mean concentration within each zone). The component $\boldsymbol{\eta} \sim \mathbf{UN}(0, \mathbf{V})\mathbf{U}$ where $\mathbf{N}(0, \mathbf{V})$ is a vector of spatially correlated random residuals with a normal distribution with zero mean and covariance matrix \mathbf{V} , and \mathbf{U} is a diagonal matrix where element $\mathbf{U}(i, i) = \sigma_j$ when the i th datum of \mathbf{z} is in zone j (i.e. this is a zone-dependent scaling factor of the variance). If we assume second-order stationarity, so that the covariance function exists and can be obtained from the variogram parameters, then these and the fixed effects coefficients $\boldsymbol{\beta}$ can be estimated simultaneously by residual maximum likelihood, REML, (Patterson and Thompson, 1971). We assumed that the spatial variation is represented by an isotropic exponential variogram model:

$$\gamma(h) = c_0 + c_1 \left\{ 1 - \exp\left(-\frac{h}{a}\right) \right\}, \tag{10}$$

in which c_0 and c_1 are the nugget and spatially correlated components of the variance, as mentioned above, and a is a distance parameter. The function approaches its maximum, $c_0 + c_1$, asymptotically, and the distance $3a$ is often taken to be the effective range of the spatial correlation (see Webster and Oliver, 2007). In the discussion below we shall often refer to the ‘effective range’ with this meaning.

We simulated values for P on a 10 m \times 10 m grid across each field using the Cholesky decomposition technique, also known as lower-upper or LU technique (Webster and Oliver, 2007). We used the variogram model that we fitted to the transformed data to create an $t \times t$ covariance matrix \mathbf{C} and scaled this for each zone independently as $\hat{\mathbf{U}}\hat{\mathbf{C}}\hat{\mathbf{U}}$, where t is the number of simulated points on the 10 m \times 10 m grid and $\hat{\mathbf{U}}$ is a diagonal matrix where element $\hat{\mathbf{U}}(i, i) = \sigma_j$ when the i th simulated value is in zone j . This was then decomposed into its lower and upper triangular form where

$$\hat{\mathbf{U}}\hat{\mathbf{C}}\hat{\mathbf{U}} = \mathbf{L}\mathbf{L}^T. \tag{11}$$

The simulated values, \mathbf{z}^* , are then given by

$$\mathbf{z}^* = s(\mathbf{L}\mathbf{g} + \mathbf{M}_{\text{sim}}\boldsymbol{\beta}), \tag{12}$$

where \mathbf{g} is a $t \times 1$ vector of random numbers drawn from a standard normal distribution, and \mathbf{M}_{sim} is a $t \times k$ design matrix. If the i th datum is in zone j then the element in the i th row of the j th column of \mathbf{M}_{sim} is 1, otherwise it is 0. Fig. 2 shows examples of the simulated values of P. For ease of calculating the yield response (see Section 2.5) we converted our simulated P values to mg kg⁻¹ by assuming the soil has a bulk density of 1.1 g cm⁻³.

To simulate field measurements with larger variances than in the observed data we also scaled the values of σ_j and simulated values of P. In each additional zone, concentrations of P were modified at four scales (0.5, 1, 2, 3) resulting in four sets of runs for fields with two zones and 16 sets of runs for fields with three zones.

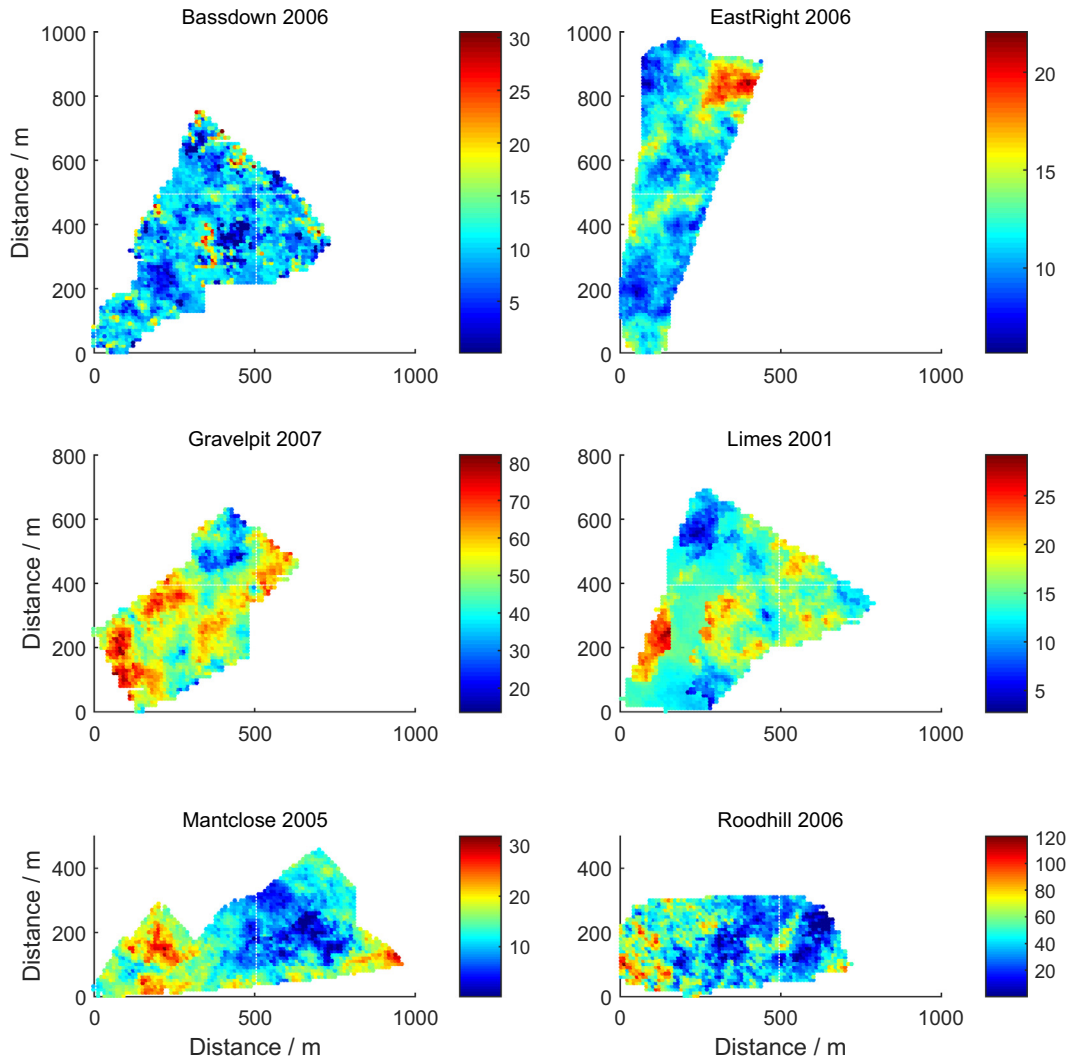


Fig. 2. Realizations of the simulated phosphorus concentrations (mg kg^{-1}).

2.4. Modelling yield and quantifying its spatial variation

For each realization of simulated phosphorus, \mathbf{z}^* , we simulated the associated yields (\mathbf{y}) using yield response models for P (see Section 2.5) and computed experimental variograms from simulated yield values by the method of moments:

$$\hat{\gamma}(h) = \frac{1}{2m(\mathbf{h})} \sum_{j=1}^{m(\mathbf{h})} \{y(\mathbf{x}_j) - y(\mathbf{x}_j + \mathbf{h})\}^2, \quad (13)$$

where $y(\mathbf{x}_j)$ and $y(\mathbf{x}_j + \mathbf{h})$ are the simulated values at positions \mathbf{x}_j and $\mathbf{x}_j + \mathbf{h}$ separated by the lag \mathbf{h} , which in the isotropic case is the scalar distance h , and $m(\mathbf{h})$ is the number of comparisons at that lag. By changing \mathbf{h} we obtained the ordered set to which we fitted exponential models, Eq. (10), by weighted least-squares approximation.

2.5. Yield response model

The yield response model for P was derived by Marchant et al. (2012) from published data (Syers et al., 2008; Johnston, 2005; Milford and Johnston, 2007; Johnston and Goulding, 1988). For every 1 kg of P added in fertilizer, we assumed that 0.18 kg is available to the crop. We also assumed that this addition is contained in the top 30 cm of soil and the soil has a bulk density of 1.1 g cm^{-3} .

This means that an addition of 1 kg P ha^{-1} leads to an increase in the concentration of this layer of $\kappa = 0.054 \text{ mg kg}^{-1}$. Thus the total nutrient available after addition of a quantity of fertilizer z_{fert} was

$$z_{\text{total}} = \kappa z_{\text{fert}} + z_{\text{soil}}, \quad (14)$$

where z_{soil} is the initial concentration of the nutrient in the soil.

The yield response to added nutrients was modelled by

$$y_r = y_0(1 - AB^{z_{\text{total}}}), \quad (15)$$

where y_r is the realized yield, y_0 is target yields and A and B are parameters. We set our target yield at 8.8 t ha^{-1} . The model parameters were $A = 1.33$ and $B = 0.68$ (see Fig. 3).

2.6. Sampling strategies

In each of the simulated fields, we compared three sampling schemes to see which would result in a treatment map that gave the greatest profit. The sampling schemes we considered were

- (i) a W-shaped design across the whole field,
- (ii) a zone-based scheme with W-shaped designs within each zone and
- (iii) a grid-based scheme with samples taken on a $100 \text{ m} \times 100 \text{ m}$ grid.

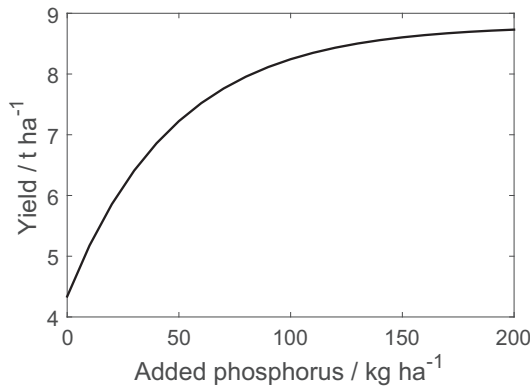


Fig. 3. Response of yield to added P to a soil with an Olsen extractable P of 2.5 mg kg^{-1} .

Each W-shaped design comprised 10 sampling points. In practice soil samples taken according to schemes (i) and (ii) are bulked before analysis resulting in either a single value for each field or each zone within a field. To simulate these we calculated the average of the nutrient values from the sample points on each W-shaped design. In grid-based designs the aim is to map the variation in a nutrient so that fertilizer rates can be adjusted accordingly. Therefore these samples are not bulked before laboratory analysis; rather a measurement is made on each sample. The number of sampling points depends on the size of the field, but typically there are too few for prediction by kriging (the fields we considered here vary from 20 to 30 ha in size). The reason is that a minimum of about 100 points is needed for an accurate estimate of the variogram. Therefore inverse distance weighting is often used for the predictions instead of kriging, but these predictions are likely to be sub-optimal because practitioners do not know the spatial scale(s) of the variation.

2.7. Evaluation of sample designs

The profit from applying any of the strategies we have set out can be assessed simply as the difference between the cost of the measurements plus that of the fertilizer and price of the crop that is produced, in this case wheat. This is the net gain, which we denote Δ for unit area, thus:

$$\Delta = yG_{\text{wheat}} - z_{\text{fert}}G_{\text{fert}} - nG_{\text{sample}}, \quad (16)$$

where G_{wheat} is the price of the grain, which we assumed to be $\text{£}150 \text{ t}^{-1}$, G_{fert} is the price of fertilizer P, assumed to be $\text{£}0.31 \text{ kg}^{-1}$, n is the number of individual soil samples analyzed in the laboratory, each costing $G_{\text{sample}} = \text{£}5$, and y and z_{fert} are the yield and quantity of fertilizer added, respectively.

Table 2
The number of management zones derived for each field and the parameters of the linear mixed model (Eq. (9)) fitted to the measured values of P. In each case $\sigma_1 = 1$.

Field name	Number of zones	Fitted means			St. dev. (s)	Mean times St. dev.			Variogram structure parameters				
		$\beta(1)$	$\beta(2)$	$\beta(3)$		$s\beta(1)$	$s\beta(2)$	$s\beta(3)$	c_0	c_1	$3a^a$	Zone 2 scaling parameter (σ_2)	Zone 3 scaling parameter (σ_3)
Bassdown	3	4.92	4.99	4.83	1.99	9.79	9.93	9.61	0.00	0.39	63.0	2.22	5.72
Easton	3	4.79	5.53	5.12	2.07	9.92	11.45	10.60	0.00	0.74	241.5	1.24	1.43
Right													
Gravelpit	3	2.27	2.78	1.92	18.33	41.61	50.96	35.19	0.26	0.93	471.9	1.15	0.61
Limes	2	3.81	4.21	–	3.40	12.95	14.31	–	0.01	0.08	480.0	4.46	–
Mantclose	2	2.81	2.92	–	6.07	17.06	17.72.31	–	0.00	0.88	522.6	1.68	–
Roodhill	3	2.06	1.40	2.70	20.21	41.63	28.29	54.57	0.00	0.97	84.9	0.58	1.16

^a Strictly the distance parameter of the exponential variogram model is a . We have inserted the quantity $3a$ in the table so that readers see more readily the effective range of the models.

We estimated the nutrient concentration in the soil at each location on the $10 \text{ m} \times 10 \text{ m}$ grid using each of the three sampling schemes. This resulted in a single estimate for the whole-field sampling scheme, an estimate for each zone for the zone-based scheme and a spatially varying estimate for the grid-based scheme. We denote these estimates of z_{soil} by \hat{z}_{field} , \hat{z}_{zone} and \hat{z}_{grid} . Using these estimates with Eqs. (14)–(16) we calculated the amount of fertilizer that should be added (\hat{z}_{fert}), noting that this value is not the true optimum as it is based on the estimated nutrient supply not the true nutrient supply, z_{soil} .

For each realization, i , of the fields we calculated the profit margin under each sampling scheme. We computed the difference in profit margin given by the zone-based, $\Delta_{\text{zone}}(i)$, and grid based, $\Delta_{\text{grid}}(i)$, schemes compared with that of the field-based $\Delta_{\text{field}}(i)$ scheme. We also computed the excess fertilizer applied under the zone- and grid-based schemes and compared this with the field-based scheme.

2.8. Using metrics of variation to guide sampling strategies

For each set of simulations, we used multiple linear regression to see how much of the variation in $\Delta_{\text{grid}} - \Delta_{\text{field}}$ and $\Delta_{\text{zone}} - \Delta_{\text{field}}$ could be explained by the distance parameter of the yield variogram a and the variance parameter c_1 . More importantly we wanted to compute the probability that the grid- or zone-based sampling strategies were more profitable than the field-based strategy for given parameters of a and c_1 . For each field we fitted the model

$$\Delta_{\text{scheme}} - \Delta_{\text{field}} = b_0 + b_1a + b_2c_1 + b_3ac_1, \quad (17)$$

to the data, where ‘scheme’ is ‘zone’ or ‘grid’. The assumption underlying the model is that the residuals are normally distributed about the mean prediction and that the standard error, s_{obs} for predicting a single observation is given by

$$s_{\text{obs}}^2 = s_{\text{mse}}^2 + \mathbf{b}^T \mathbf{V}(\mathbf{b}) \mathbf{b} \quad (18)$$

where $\mathbf{V}(\mathbf{b})$ is the covariance function for the parameter estimates $\mathbf{b} \equiv \{b_0, b_1, b_2, b_3\}$ and s_{mse}^2 is the mean square error. From this we calculated the probability that $\Delta_j - \Delta_{\text{field}} > 0$ for any given combination of a and c_1 . We did a similar analysis for excess fertilizer.

3. Results

Based on the normalized classification entropy, two or three management zones were identified for each field (Table 2). Table 2 lists the parameters of the models fitted to the nutrient measurements, Eq. (9), and the fitted variograms are shown in Fig. 4. The sill variances, $c_0 + c_1$, vary from 0.09 to 2.23, and the effective range from 63.0 m to 522.6 m showing substantial differences in the spatial structure of P between the fields.

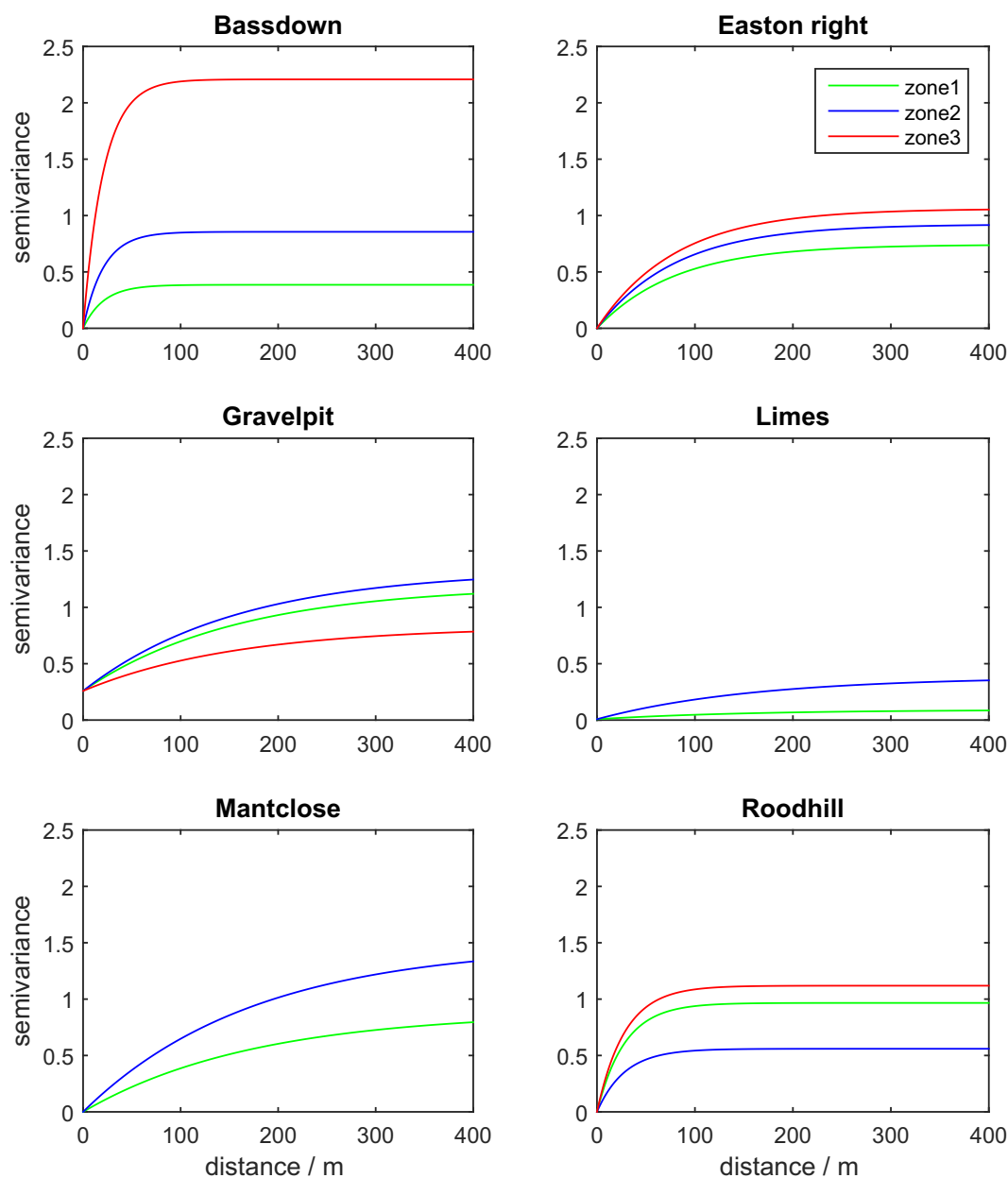


Fig. 4. The spatial structure fitted to the phosphorus data in the six fields.

Table 3

Simulated field averages of soil P concentrations with the average difference in net profit for zone- and grid-based sampling compared with the field-based sampling across all simulations.

Field name	Range of simulated averages of P (mg kg ⁻¹)	Mean net profit (£ ha ⁻¹)				Mean excess nutrient (kg ha ⁻¹)			
		Zone-Field		Grid-Field		Zone-Field		Grid-Field	
		Profit	SE ^a	Profit	SE	Profit	SE	Profit	SE
Bassdown	19–21	-13	0.81	12.1	0.75	-2.49	0.17	1.07	0.17
Easton Right	17–25	0.93	0.23	14.8	0.21	0.14	0.11	-1.9	0.10
Gravelpit	48–131	2.6	0.36	10.3	0.46	0.14	0.02	0.64	0.02
Limes	22–30	12.9	1.2	57.7	1.4	-2.7	0.22	-0.35	0.22
Mantclose	7–34	16.7	1.05	46.9	1.11	-2.06	0.16	-0.10	0.16
Roodhill	81–90	-0.18	0.05	2.4	0.08	0.07	0.01	0.98	0.01

^a Standard error.

3.1. Comparison of P estimated by the sampling schemes

The range of the simulated field averages for P (Table 3) were consistent with those reported by analytical laboratories for arable

fields in England and Wales (NRM Laboratories, 2017) and within the measured range (Table 1). For each realization we computed the difference in profit margin given by the zone-based (Δ_{zone}) and grid based (Δ_{grid}) schemes compared with that of the

field-based (Δ_{field}) scheme. Thus, the differences are ($\Delta_{\text{grid}} - \Delta_{\text{field}}$) and ($\Delta_{\text{zone}} - \Delta_{\text{field}}$) for the grid-based and zone-based values, respectively. Table 3 reports the means and standard errors for these differences, which are also shown in Fig. 5. In all cases the grid-based estimates gave larger profits than did the zone-based estimates. This was largely because the cluster classes were not significant factors in explaining the variation in the nutrients

(see supplementary information). The smallest differences between the zone-based estimate and the grid-based one were for the fields with the largest differences in mean concentrations of P between zones (Gravelpit and Roodhill). Bassdown had the smallest mean of $\Delta_{\text{zone}} - \Delta_{\text{field}}$, and this is likely to result from both the small difference in mean values of P between the zones, $s\beta(1), s\beta(2)$ and $s\beta(3)$, and the short-range variation (effective

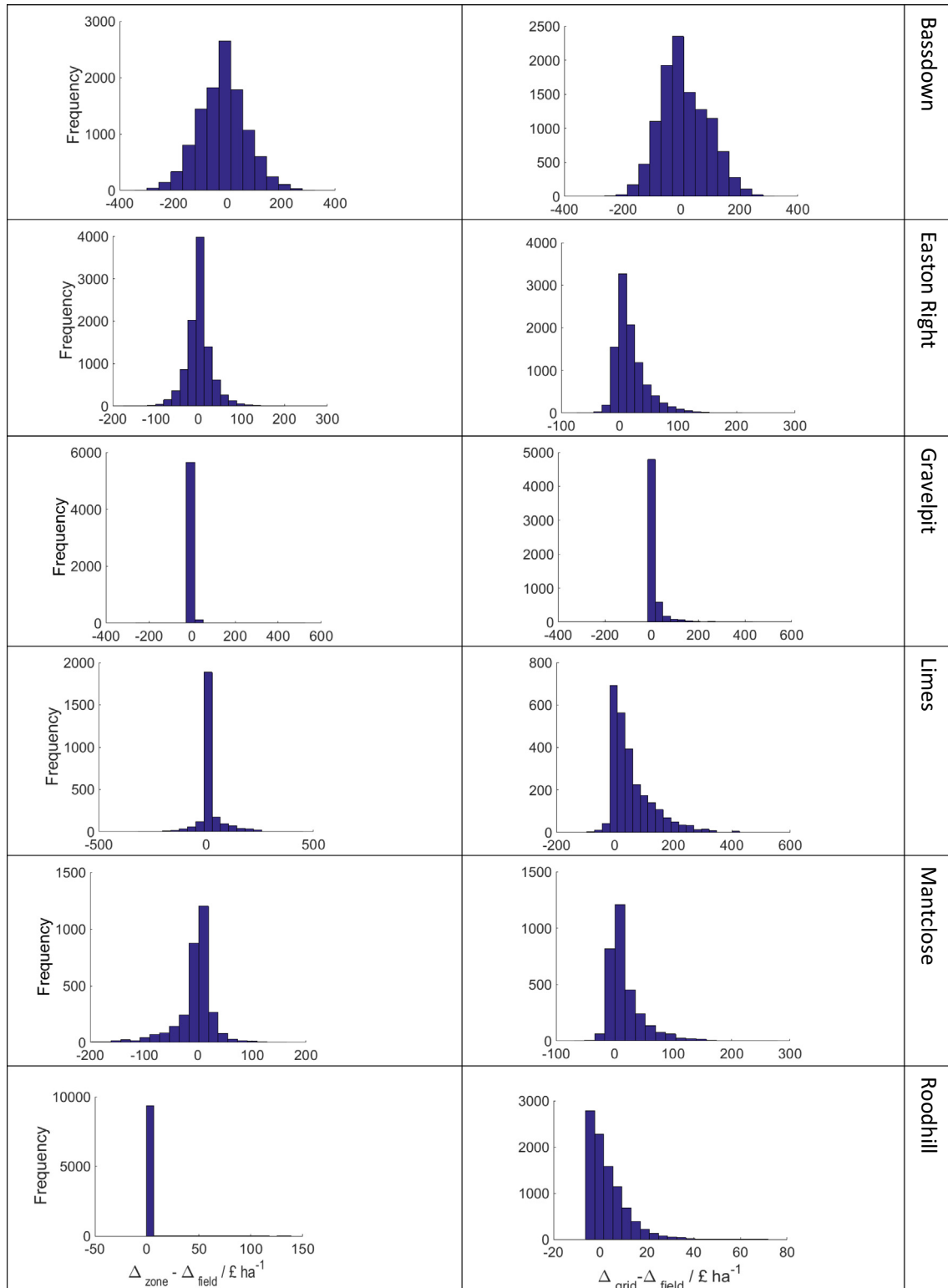


Fig. 5. Histograms of $\Delta_{\text{zone}} - \Delta_{\text{field}}$ and $\Delta_{\text{grid}} - \Delta_{\text{field}}$ data.

Table 4

The percentage variance accounted for and model parameters for the multiple linear regression models (Eq. (17)) fitted to the $\Delta_{\text{zone}} - \Delta_{\text{field}}$ and $\Delta_{\text{grid}} - \Delta_{\text{field}}$ data.

Field name	Model parameters				Percentage variance accounted for
	b_0	b_1	b_2	b_3	
<i>Zone sampling</i>					
Bassdown	-7.61	-0.1907	-7.82	0.1552	0.1
Easton Right	-2.07	-0.0082	4.275	-0.00428	2.1
Gravelpit ^a	-7.2	0.018	12.53	-0.0075	16.0
Limes	-4.03	0.0008	3.882	0.00084	6.7
Mantclose	-17.4	0.253	11.61	-0.0378	3.8
Roodhill ^a	189.6	-1.55	-25.98	0.25	10.2
<i>Grid sampling</i>					
Bassdown	6.01	-0.013	-10.03	0.2322	0.5
Easton Right	5.762	0.00653	6.629	0.01229	18.7
Gravelpit	-8.065	-0.00326	5.3	0.02439	50.2
Limes	-9.03	0.1492	10.046	-0.00218	33.9
Mantclose	4.654	-0.0205	-1.065	0.0049	45.0
Roodhill	-7.117	0.0955	1.341	0.0145	21.9

^a Model fitted to realizations only where estimated P was limiting.

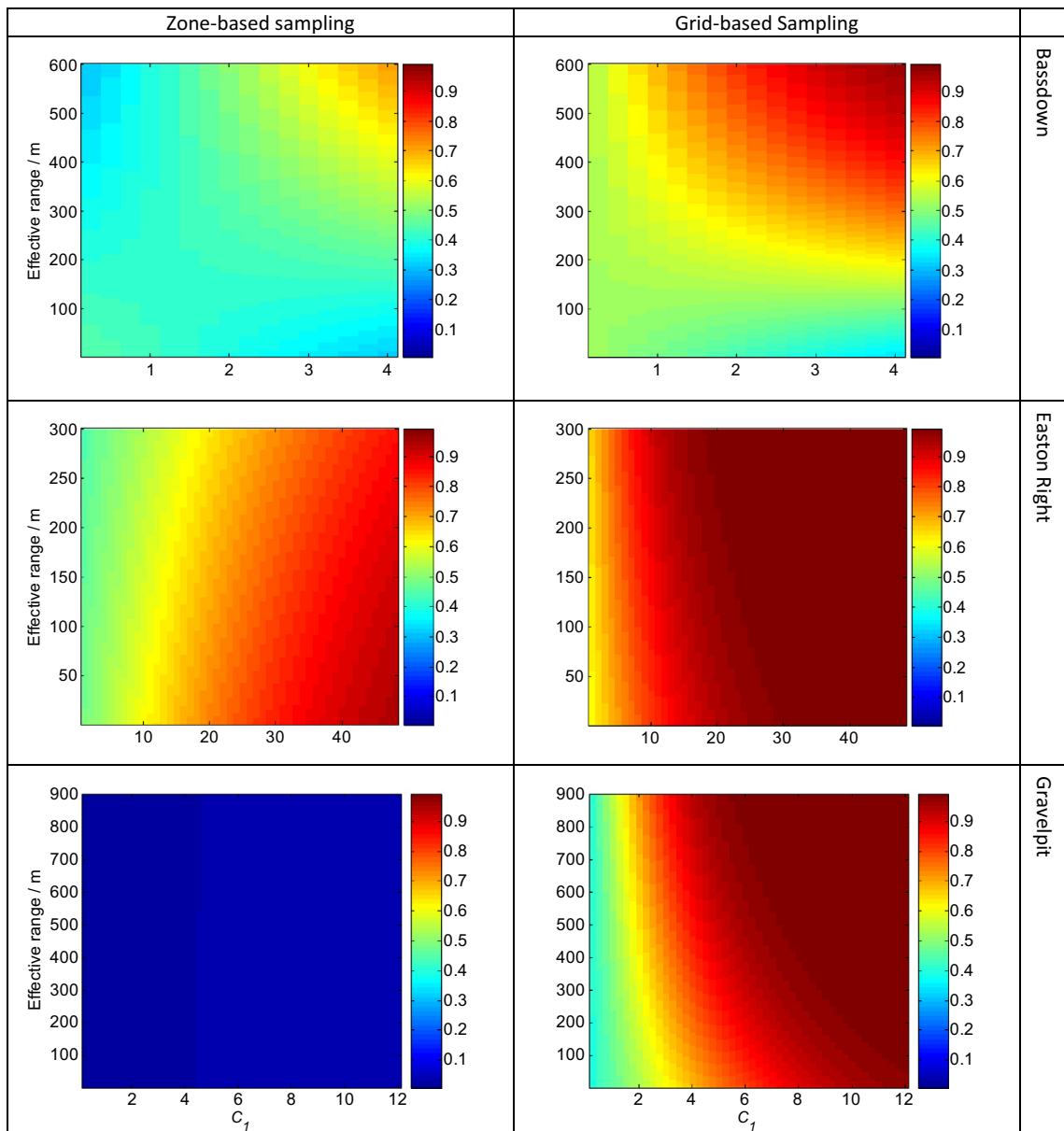


Fig. 6a. Maps of Bassdown, Easton Right and Gravelpit showing the probability that zone-based and grid-based sampling are more profitable than field-based sampling.

range 63.0 m) which is smaller than the distance across feasible management zones. The largest mean profits were for Limes and Mantclose which have a long spatial structure (Fig. 4) with values of P in a treatable range (Table 3 and Fig. 3).

The distributions of $\Delta_{zone} - \Delta_{field}$ are more symmetrically distributed than for $\Delta_{grid} - \Delta_{field}$ for all fields except for Roodhill, (as illustrated in Fig. 5). The distributions of $\Delta_{grid} - \Delta_{field}$ are positively skewed. For Roodhill and Gravelpit a large number of realizations had values of \hat{z}_{field} and \hat{z}_{zone} that were not limiting (99% and 66% respectively). This resulted in recommendations of no fertilizer application, and so $\Delta_{zone} - \Delta_{field}$ is simply the difference in sampling costs. These correspond to the large peaks in the distributions of $\Delta_{grid} - \Delta_{field}$.

We also computed the difference in excess fertilizer applied when estimates were based on the zone-based (\hat{z}_{zone}) and grid based (\hat{z}_{grid}) sampling schemes compared with the field-based (\hat{z}_{field}) scheme. We define excess fertilizer as the amount applied

over and above that which would have been applied if we had perfect knowledge of the true variation in P across the field. Table 3 reports the means and standard error for these differences ($\hat{z}_{zone} - \hat{z}_{field}$ and $\hat{z}_{grid} - \hat{z}_{field}$).

There was no consistent pattern to the mean responses. In some fields (Easton Right, Gravelpit and Roodhill) the field-based sampling resulted in less of an excess than the zone-based sampling with positive differences and in some fields (Bassdown, Gravelpit and Roodhill) the field-based sampling resulted in less of an excess than the grid-based sampling (Table 3). Similarly there was no consistent pattern between grid- and zone-based sampling.

3.2. Assessing the extent to which yield maps can be used to predict the most appropriate sampling scheme

The parameters for the models fitted by multiple linear regression are listed in Table 4 along with percentage variance accounted

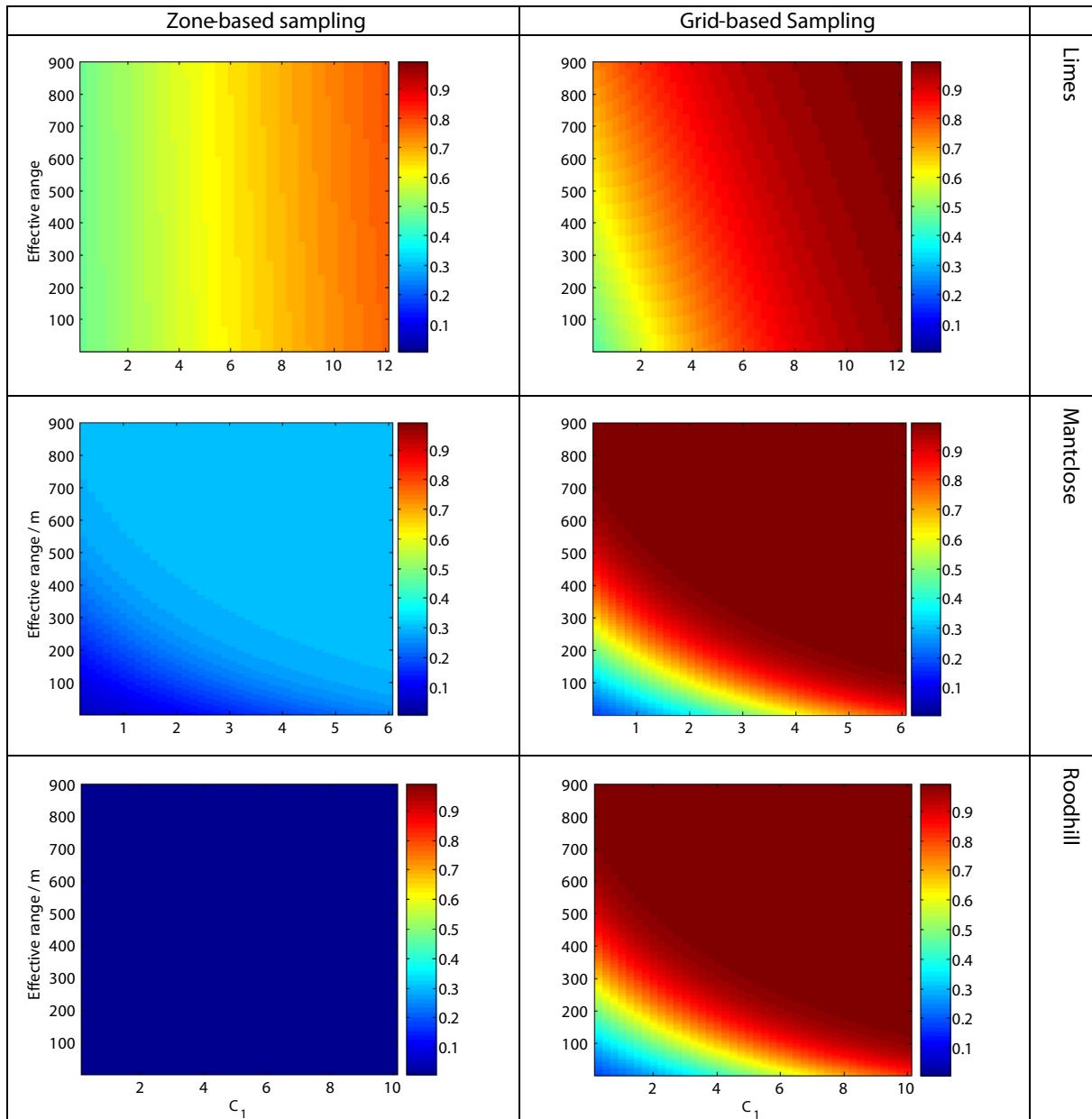


Fig. 6b. Maps Limes, Mantclose and Roodhill showing the probability that zone-based and grid-based sampling are more profitable than field-based sampling.

Table 5The percentage variance accounted for and model parameters for the multiple linear regression models (Eq. (17)) fitted to the $\hat{z}_{\text{zone}} - \hat{z}_{\text{field}}$ and $\hat{z}_{\text{grid}} - \hat{z}_{\text{field}}$.

Field name	Model parameters				Percentage variance accounted for
	b_0	b_1	b_2	b_3	
<i>Zone sampling</i>					
Bassdown	-1.660	-0.0116	-1.079	0.0168	0.0
Easton Right	2.117	-0.00997	-0.843	0.00362	0.3
Gravelpit ^a	0.54	0.00007	0.711	-0.00191	1.5
Limes	0.884	-0.00276	-0.761	0.000096	6.9
Mantclose ^a	6.21	-0.0044	-3.373	0.00874	13.0
Roodhill ^a	6.89	0.012	-0.261	0.01548	1.1
<i>Grid sampling</i>					
Bassdown	0.139	0.0121	-0.363	0.00785	0.0
Easton Right	5.762	0.00653	0.629	0.0123	0.4
Gravelpit	-0.0935	0.00198	0.397	0.000888	12.6
Limes	1.796	0.00439	0.441	-0.00131	6.0
Mantclose	3.72	-0.00554	-1.104	0.00173	8.4
Roodhill	3.14	0.00735	0.119	0.00828	3.0

^a Model fitted to realizations only where estimated P was limiting.

for. For Gravelpit, Mantclose and Roodhill the regression model was fitted to the subset of realizations where \hat{z}_{field} or \hat{z}_{zone} were limiting. The models fitted to $\Delta_{\text{zone}} - \Delta_{\text{field}}$ explained very little of the variation in the data. The variation in $\Delta_{\text{grid}} - \Delta_{\text{field}}$ was better explained, although the value for Bassdown were still small. The realizations for this site were generated from the model with a small effective range and the nugget to sill ratios of the realizations were in general larger than other sites (at least 37% larger, data not shown) indicating a relatively large component of unstructured variance.

Figs. 6a and 6b shows the probability that Δ_{zone} and Δ_{grid} are larger than Δ_{field} . In all cases the probability increases with both c_1 and the effective range (3a). This is also true of $\Delta_{\text{zone}} - \Delta_{\text{field}}$ for Gravelpit and Roodhill, but these sites are dominated by simulations where \hat{z}_{field} and \hat{z}_{zone} were not limiting, and so the effects are negligible.

The multiple linear regressions showed that the effective range and c_1 explained very little of the variation in the excess fertilizer data, $\hat{z}_{\text{zone}} - \hat{z}_{\text{field}}$ and $\hat{z}_{\text{grid}} - \hat{z}_{\text{field}}$ (Table 5).

4. Discussion and conclusion

We have investigated, through simulation, the cost-effectiveness of three sampling strategies commonly used to guide fertilizer recommendations. We aimed to see if the variation captured in yield monitor data could be used to determine which sampling strategy would be best in any given situation. The first option was the conventional single bulked sample for a whole field. The second was to delineate spatially coherent management zones from the yield data, sample the soil of each zone independently (zone-based sampling) and apply fertilizer (P in our case) in accord with the P content of the soil measured. The third option was to estimate the magnitude of the variation of P across the field, and judge from this how fertilizer might be applied at a variable rate. This required grid-based sampling.

Despite the expected increases in yield in response to additions of fertilizer, we know that there are situations in which there are inverse relations between nutrient concentration and yields (see, for example, Lake et al., 1997, for wheat and Cox et al., 2003, for soya beans). It seems that in those situations the larger crops have extracted more of the nutrients and depleted the stock in the soil.

We have shown that the advantages of using grid-and zone-based sampling strategies over field-based ones vary from field to field. In our simulations, on average, the grid-based sampling performed better than the zone-based sampling. This was largely

because of the variation in the concentration of P, and the mean concentrations, did not differ sufficiently from one zone to another (even after scaling down the P measures used to model the variation in soil P by 0.5). This is likely to be so for many arable fields in the UK because so much P has been added to them in the last 150 years. Other investigators (Mallarino and Wittry, 2004; Flowers et al., 2005; Sawchik and Mallarino, 2007) similarly found that grid-based sampling was better than the zone-based approach for estimating the likely responses of field crops to added fertilizer. Factors such as changes in water availability (due to variation in drainage or soil texture) or local infestations of weeds or disease will often be the main cause of large variations. In practice, many farmers would be able to explain the observed differences between proposed management zones and so be able to predict whether the zones were determined by differences in nutrient availability. This would provide valuable information on whether zone-based sampling was sensible.

The maps in Fig. 6a and 6b show that the probability of grid-based sampling's being more profitable increases with both increases in effective range and in c_1 . Larger values of c_1 imply large differences in nutrients, and a farmer might wish to apply fertilizer differentially in accord. This is feasible in practice only if the effective range is also large, and for two reasons. One is the difficulty of varying the application at a fine scale; the other is the cost of sampling and soil analysis on grids fine enough to map the concentration of the nutrient in the soil.

Notice that the probability maps in Fig. 6a and 6b change from field to field. Note also that they are based on current prices of wheat and fertilizer and sampling costs, which are subject to variation, and that our models might have introduced bias in profitability. Indeed, farmers will typically measure several soil variables at once (e.g. P, K, Mg and pH). Soil sampling may reveal anyone of these variables (or a combination of) to be the dominant cause of yield variation. Accurate quantification of the benefits of one sampling approach over another are therefore difficult, and so the absolute values shown in Figs. 6a and 6b should not be applied in other contexts, although, the principles hold. As variation in yield at scales appropriate for management increases, so does the likelihood that grid-based sampling will be more profitable than a single field estimate.

Farmers and their advisors are building increasingly large data sets of crop yields, soil properties and previous management decisions. They should find these data valuable for managing their fields more effectively both to improve productivity and to reduce unnecessary inputs. The principles described here could be integrated into a decision-support system to help farmers decide when

variable-rate application of fertilizer might be cost effective and if so what sampling strategy to apply. This is the first step, then the farmer must decide how to vary the fertilizer spatially. The reliable prediction of how much fertilizer to apply depends not only on the quantification of the soil nutrient supply but also the yield potential of the crop and the efficiency of nutrient uptake by the crop (Kindred et al., 2015). Predicting yield potential and the efficiency of nutrient uptake are substantial topics of research that can involve sensor technology, empirical analysis and simulation modelling (Dhital and Raun, 2016; van Wart et al., 2013; Kindred et al., 2015). In our study we followed the practices of agricultural advisors in the UK and considered zone-based and grid-based sampling schemes. The grid-based sampling scheme was on a 100-m grid, but in practice the variogram could be further used to optimize the grid sampling according to the cost of samples and the apparent dominant scale of spatial variation in the yield data.

Acknowledgement

This research was funded by the AHDB cereals and oilseeds project RD-2012-3785, using facilities funded by the Biotechnology and Biological Sciences Research Council (BBSRC). B. Marchant's contribution is published with the permission of the executive director of the British Geological Survey (Natural Environment Research Council). We also thank Simon Griffin from SOYL for kindly providing us with the data.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.compag.2017.02.002>.

References

- Córdoba, M.A., Bruno, C.I., Costa, J.L., Peralta, N.R., Balzarini, M.G., 2016. Protocol for multivariate homogeneous zone delineation in precision agriculture. *Biosyst. Eng.* 143, 95–107.
- Cox, M.S., Gerard, P.D., Wardlaw, M.C., Abshire, M.J., 2003. Variability of selected soil properties and their relationships with soybean yield. *Soil Sci. Soc. Am. J.* 67, 1296–1302.
- Dhital, S., Raun, W.R., 2016. Variability in optimum nitrogen rates for maize. *Agron. J.* 108, 2165–2173.
- Diker, K., Heermann, D.F., Brohdahl, M.K., 2004. Frequency analysis of yield for delineating zones. *Precision Agric.* 5, 435–444.
- Dray, S., Legendre, P., Peres-Neto, P.R., 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbouring matrices (PCNM). *Ecol. Model.* 196, 483–493.
- Dunn, J.C., 1977. Indices of fuzziness and the detection of clusters in large data sets. In: Gupta, M.M. (Ed.), *Fuzzy Automata and Decision Processes*. North-Holland, New York, pp. 271–283.
- Fleming, K.L., Westfall, D.G., Bausch, W.C., 2000. Evaluating management zone technology and grid soil sampling for variable rate nitrogen application. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), *Proceedings of the 5th International conference on precision agriculture and other source management*. ASA, CSSA, SSSA, Madison, WI, USA.
- Flowers, M., Weisz, R., White, J.G., 2005. Yield-based management zones and grid sampling strategies. *Agron. J.* 97, 968–982.
- Frogbrook, Z.L., Oliver, M.A., Salahi, M., Ellis, R.H., 2002. Exploring the spatial relations between cereal yield and soil chemical properties and the implications for sampling. *Soil Use Manage.* 18, 1–9.
- Fu, W., Zhao, K., Jiang, P., Ye, Z., Tunney, H., Zhang, C., 2013. Field-scale variability of soil test phosphorus and other nutrients in grasslands under long-term agricultural managements. *Soil Res.* 51, 503–512.
- Hawkins, E., LeBarge, G., Watters, H.D., Fulton, J., Culman, S., 2016. Developing a strategy for precision sampling. Ohio State University Extension. <<http://agcrops.osu.edu/newsletter/corn-newsletter/2016-37/developing-strategy-precision-soil-sampling>>.
- Johnston, A.E., 2005. Phosphorus nutrition in arable crops. In: Sims, J.T., Sharpley, A. N. (Eds.) *Phosphorus: Agriculture and the environment*. Agronomy Monograph No 46. ASA-CSSA-SSSA, Madison, WI, USA, pp. 495–519.
- Johnston, A.E., Goulding, K.W.T., 1988. Rational potassium manuring for arable cropping systems. *J. Sci. Food Agric.* 46, 1–11.
- Kindred, D., Milne, A.E., Webster, R., Marchant, B., Sylvester-Bradley, R., 2015. Exploring the spatial variation in the fertilizer-nitrogen requirement of wheat within fields. *J. Agric. Sci.* 153, 2421.
- Kravchenko, A.N., 2003. Influence of spatial structure on accuracy of interpolation methods. *Soil Sci. Soc. Am. J.* 67, 1564–1571.
- Lake, J.V., Bock, G.R., Goode, J.A. (Eds.), 1997. *Precision Agriculture: Spatial and Temporal Variability of Environmental Quality*. John Wiley & Sons, Chichester.
- Lark, R.M., 1998. Forming spatially coherent regions by classification of multivariate data: an example from analysis of maps of crop yield. *Int. J. Geogr. Inf. Sci.* 129, 83–98.
- Lark, R.M., Wheeler, H.C., Bradley, R.I., Mayr, T.R., Dampney, P.M.R., 2003. Developing a Cost-Effective Procedure for Investigating Within-Field Variation of Soil Conditions. Project Report No 296. HGCA, London.
- Mallarino, A.P., Wittry, D.J., 2004. Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, pH, and organic matter. *Precision Agric.* 5, 131–144.
- Marchant, B.P., Dailey, A.G., Lark, R.M., 2012. Cost-Effective Sampling Strategies for Soil Management. Project Report No 485. Home-Grown Cereals Authority. At <<http://cereals.ahdb.org.uk/media/252469/pr485.pdf>> (last accessed November 2015).
- McCormick, S., Jordan, C., Bailey, J.S., 2009. Within and between-field spatial variation in soil phosphorus in permanent grassland. *Precision Agric.* 10, 262–276.
- Milford, G.F.J., Johnston, A.E., 2007. Potassium and nitrogen interactions. In: *Proceedings No 615*. International Fertiliser Society, Leek, UK, pp. 1–22.
- Milne, A.E., Webster, R., Ginsburg, D., Kindred, D., 2012. Spatial multivariate classification of an arable field into compact management zones based on past crop yields. *Comput. Electron. Agric.* 80, 17–30.
- Mylavarapu, R.S., Wonsok, D.L., 2014. Nutrient Management Series, Soil Sampling Strategies for Precision Agriculture. University of Florida, Institute of Food and Agricultural Sciences. <<http://edis.ifas.ufl.edu/pdffiles/SS/SS40200.pdf>>.
- NRM Laboratories (2017). Soil Nutrient Status, Soil Summary 2015-2016. NRM Laboratories, Bracknell, UK. <<http://www.nrm.uk.com/files/documents/NRM-Soil-Summary-2015-2016-Opt03b.pdf>> (last accessed Feb 2017).
- Oliver, M.A., Kerry, R., 2013. Sampling and Geostatistics for Precision Agriculture. International Fertiliser Society, Colchester, UK.
- Oliver, M.A., Webster, R., 1989. A geostatistical basis for spatial weighting in multivariate classification. *Math. Geol.* 21, 15–35.
- Oliver, M.A., Webster, R., 2014. A tutorial guide to geostatistics: computing and modelling variograms and kriging. *Catena* 113, 56–69.
- Patterson, H.D., Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545–554.
- PDA, 2011. Soil Analysis: Key to Nutrient Management Planning, Leaflet 24. The Potash Development Association, York, UK.
- Peralta, N.R., Costa, J.L., Balzarini, M., Franco, M.C., Córdoba, M., Bullock, D., 2015. Delineation of management zones to improve management of wheat. *Comput. Electron. Agric.* 110, 103–113.
- Sawchik, J., Mallarino, A.P., 2007. Evaluation of zone soil sampling approaches for phosphorus and potassium based on corn and soybean response to fertilization. *Agron. J.* 99, 1564–1578.
- Serrano, J.M., Pea, J.O., Marques da Silva, J.R., Shahidian, S., 2011. Spatial and temporal stability of soil phosphate concentration and pasture dry matter yield. *Precision Agric.* 12, 214–232.
- Stafford, J.V., Lark, R.M., Bolam, H.C., 1999. Using yield maps to regionalize fields into potential management units. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), *Proceedings of the 5th International Conference on Precision Agriculture and Other Source Management*. ASA, CSSA, SSSA, Madison, WI, USA, pp. 225–237.
- Syres, 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. Bulletin 18. FAO, Rome. <<ftp://ftp.fao.org/docrep/fao/010/a1595e/a1595e00.pdf>> (last accessed November 2015).
- van Wart, J., Kersebaum, K.C., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. *Field Crops Res.* 143, 34–43.
- Webster, R., Oliver, M.A., 2007. *Geostatistics for Environmental Scientists*. John Wiley & Sons, Chichester.
- Zhang, C., Zhao, M., Hiu, L.G., Chen, F.X., 2009. A PDA-based system for farmland data collection and transmission. In: Yu, F., Liu, Y., Yue, G. (Eds.), *2009 International Symposium on Intelligent Information Systems and Application Proceedings*. Academic Publications, Oulu, pp. 162–165.