Precision Agriculture (2022) 23:1333–1353
https://doi.org/10.1007/s11119-022-09887-2

A loss function to evaluate agricultural decision-making under uncertainty: a case study of soil spectroscopy

T. S. Breure[1,2] · S. M. Haefele[2] · J. A. Hannam[1] · R. Corstanje[1] · R. Webster[2] · S. Moreno-Rojas[3] · A. E. Milne[2]

Accepted: 29 January 2022 / Published online: 12 March 2022
© The Author(s) 2022

Abstract
Modern sensor technologies can provide detailed information about soil variation which allows for more precise application of fertiliser to minimise environmental harm imposed by agriculture. However, growers should lose neither income nor yield from associated uncertainties of predicted nutrient concentrations and thus one must acknowledge and account for uncertainties. A framework is presented that accounts for the uncertainty and determines the cost–benefit of data on available phosphorus (P) and potassium (K) in the soil determined from sensors. For four fields, the uncertainty associated with variation in soil P and K predicted from sensors was determined. Using published fertiliser dose–yield response curves for a horticultural crop the effect of estimation errors from sensor data on expected financial losses was quantified. The expected losses from optimal precise application were compared with the losses expected from uniform fertiliser application (equivalent to little or no knowledge on soil variation). The asymmetry of the loss function meant that underestimation of P and K generally led to greater losses than the losses from overestimation. This study shows that substantial financial gains can be obtained from sensor-based precise application of P and K fertiliser, with savings of up to £121 ha^{-1} for P and up to £81 ha^{-1} for K, with concurrent environmental benefits due to a reduction of 4–17 kg ha^{-1} applied P fertiliser when compared with uniform application.

Keywords Geostatistics · Precision agriculture · Variable-rate application · Proximal soil sensing · X-ray fluorescence

✉ T. S. Breure
timo.breure@rothamsted.ac.uk

[1] Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

[2] Rothamsted Research, Harpenden AL5 2JQ, UK

[3] Gs Growers Ltd, Ely CB7 5TZ, UK

Introduction

It is estimated that globally 9–14 million tonnes of phosphorus (P) leaches from fields into watercourses (Beusen et al., 2016; Chen et al., 2016), largely as a result of excessive application of fertiliser. This excessive use of fertiliser is causing substantial environmental harm. To minimise such harm, fertiliser needs to be applied more precisely than at present, varying across fields to meet crop requirements but no more. This accords with the aims of precision agriculture.

Precision agriculture (PA) aims to produce sufficient crops sustainably for society's needs while minimising costs to the producer and harm to the environment. This involves the management of spatial and temporal variation within fields; it requires intense information. Growers who adopt PA must consider the need, value, costs and possible other sources of information to identify whether they can use it to improve their efficiency and reduce environmental impact. The associated uncertainty of the information they acquire for this purpose affects both its value and consequently their view of this approach to management. Current decision-making on variable rate application (VRA) of fertiliser is based primarily on yield responses as functions of management inputs in agronomic trials (Pringle et al., 2004a, 2004b). To vary fertiliser intelligently, however, growers need to have detailed maps of nutrient status or fertiliser requirement for their fields.

Obtaining detailed soil information to support VRA requires intense sampling. Although the required sampling density depends on specific circumstances, in particular the degree of short-distance variation in the soil, a general consensus is that around 10 observations per hectare are required. Both the sampling and laboratory analysis of soil collected are laborious and time-consuming. The whole process by conventional wet chemical analysis is too expensive for the sizes of samples required to map soil variation accurately within fields. Most growers therefore apply uniform dressings of fertiliser based on average estimates of nutrient concentrations made from a small number of samples. Recent developments in spectroscopy offer an affordable and effective alternative with instruments designed for use both in the laboratory and in the field (Li et al., 2015). X-ray fluorescence (XRF) spectroscopy has been available for several decades, and more recently infrared (IR) reflectance spectroscopy has become feasible for analysing and predicting soil properties on a large number of samples cheaply (Bellon-Maurel & McBratney, 2011; Viscarra Rossel & Webster, 2012). The net result is substantially cheaper than chemical analysis alone and enables surveyors to obtain sufficient detail on the spatial variation of soil properties affordably (Viscarra Rossel & Bouma, 2016).

All measurement embodies some degree of error, however, and so soil data also have their associated uncertainty. Nevertheless this study considered the errors in standard chemical analyses by modern equipment small enough to be ignored. Spectroscopic estimation introduces yet another source of error. That is, calibration equations that describe the relationship between the wet chemistry measures and the soil spectra also have an associated error. Finally, measurements of the soil cannot be made at all locations, and so interpolation is necessary to estimate the soil variables between measured locations. This interpolation, usually done by kriging, has an associated prediction error. These errors accumulate through the whole procedure and are embodied in the error variances of the final estimates. Ignoring the resultant uncertainty can lead to false inferences from the data and hence faulty decision-making (Cherry et al., 2008; Goovaerts, 2001; Heuvelink, 2018).

Given the ease with which spectroscopy can replace conventional chemical methods for analysing soil, it is important to know how the errors it introduces affect the final spatial predictions and in turn the potential financial advantage of precise fertiliser management compared with uniform application. This is because errors carry with them costs. Over-estimation of a plant nutrient concentration in the soil, say P, would lead to a grower's applying too little fertiliser and to loss of potential yield and income. Under-estimation of the concentration would lead to the grower's over-fertilizing, spending unnecessarily on fertiliser, to the point of spending more than earned in increased yield of crop.

Yates (1981) set out the principles by which one might assess the balance between costs of survey and benefits that would accrue from greater accuracy. The aim is to minimise the sum of sampling costs and expected losses due to errors. For this Yates defined a loss function and suggested how it might be minimised. The loss function is a generic approach studied not only in the context of soil survey (Lark & Knights, 2015) but also in environmental protection (Goovaerts, 1997) and soil management (Faechner et al., 2000). Because overestimation and underestimation incur losses for different reasons the loss function may be asymmetric. Given a loss function and an error distribution for the information, one can make a decision that minimises expected losses (e.g. Faechner et al., 2000; Goovaerts, 1997).

This study is concerned with the prediction of available P and K for precise management of fertiliser in horticultural crops. The main objective is to investigate the potential for soil spectral methods (near- and mid-infrared and X-ray fluorescence) to predict how these soil variables change within well-managed fields and so determine the effect of prediction errors on the expected loss. It was hypothesised that, once uncertainty in predictions has been accounted for, soil spectroscopy could adequately predict the spatial variability to justify variable rate application of P and K fertiliser. For this, soil samples from four fields in the Cambridgeshire Fens of the UK were considered. Measurements of available P and K estimated from soil spectra were used to predict how the concentrations vary across the fields and to compute the associated error variances from interpolation. The expected losses associated with varying applications of fertiliser given the error variance of the predictions are computed. The expected losses are compared with the losses that would have accrued if estimates from wet chemistry had been used to determine a single application rate per field. General conclusions are drawn about the effect of uncertainty in nutrient status on economic and environmental losses and practical considerations are given with regard to implementing the loss function.

Methods

Data

Data were obtained in sample surveys of four fields in the Fen district of Cambridgeshire, England, in 2018 and 2019. The region was originally dominated by peat, much of which has oxidized since the land was drained in the seventeenth century. The land surface is now 1 to 2 m lower than it was except for the natural sinuous drainage channels containing mineral sediment. These former channels, known locally as 'rodhams', have become elevated features in the landscape (Hodge et al., 1984), and they are clearly distinguishable on LiDAR (light detection and ranging) imagery. The LiDAR raster (2 m × 2 m resolution) from the British Environment Agency was used as a basis for sampling. The sampling

design of Field 1 (8.2 ha) was based around a 30-m square grid, with three transects (on alternate rows of the grid) more intensely sampled at 6-m intervals. The designs for Field 2 (16.9 ha), Field 3 (5.1 ha) and Field 4 (8.9 ha) were computed for initial numbers of points (121, 107 and 100, respectively) by spatial coverage sampling (Walvoort et al., 2010). Each point lay in the centre of its Dirichlet tile. All tiles in each field were of equal area, ensuring spatial coverage of the entire field. This led to an approximate grid with an interval of around 30 m. A sub-sample of 36 (Field 2), 26 (Field 3) and 32 (Field 4) of these points were selected with balanced sampling (Grafström & Lilics, 2019) on the spatial coordinates and LiDAR. At each location of these sub-samples, another sampling point 6 m away at a random orientation was added to estimate the short-scale spatial variance. In all fields extra sample points were also added to ensure coverage of the range of soil conditions and LiDAR. The decision for the location of these extra points was based on the LiDAR survey and satellite imagery showing variation in soil colour. In all, the numbers of sampling points for the fields were 256 (Field 1), 161 (Field 2), 138 (Field 3) and 142 (Field 4). Supplementary Fig. 1 shows the field boundaries with the sampling points. Three soil cores of topsoil (0–0.25 m) were taken within a 0.5 m × 0.5 m quadrat at each sampling location. These three cores were bulked and mixed for spectroscopic measurements. A subset of 30 samples from each field was taken for further laboratory analysis. The subset was selected from the total sample by balanced sampling on the spatial coordinates and LiDAR data.

Available P was measured by the standard Olsen method (Olsen et al., 1954) and a SANplus continuous colorimetric flow analysis (Skalar Analytical BV, Breda, Netherlands). Available K was determined in an ammonium-nitrate (NH_4NO_3) extract and Inductively Coupled Optical-Emission Spectroscopy (Optima 7300 DV, Seer Green, UK). The soil samples were dried in air and milled, and sub-samples were pressed into small wells (6 mm across and approximately 1 mm deep) and placed in a Tensor II spectrometer (Bruker, Ettlingen, Germany). The absorbance spectrum in the range 9998–3999 cm^{-1}, i.e. the near-infrared (NIR), of each sub-sample was measured with a resolution of 1 cm. Each sub-sample's mid-infrared (MIR) spectrum in the range 4000–600 cm^{-1} was recorded on the same instrument with a resolution of 2 cm^{-1}. A DP-6000 Delta Premium portable X-ray fluorescence (pXRF) (Olympus Ltd, Center Valley, USA) was used to scan the soil samples. The pXRF features a Rh X-ray tube operated at 10–40 keV with a high resolution (< 165 eV) silicon drift detector. The pXRF was set to scan for 30 s at both 10 and 40 keV. The pXRF was set up in an instrument stand, and samples where placed on the aperture in a sample cup covered with a Prolene Thin Film (Chemplex Ind, Florida, USA). Potential drift in the XRF analyser was reduced by scans of a stainless steel 316-alloy clip containing 16.13% Cr, 1.78% Mn, 68.76% Fe, 10.42% Ni, 0.20% Cu, and 2.10% Mo tightly fitted over the aperture prior to the measurements on each aliquot. The pXRF samples where measured in three replicates on one aliquot, near- and mid-infrared spectra were measured on three aliquots of each soil sample. Further analysis was done on the mean spectra of those three measurements.

Spectral processing and calibration

The raw spectra were pre-processed first by the Savitzky–Golay filter (Savitzky & Golay, 1964) and then transformed to their first derivatives. The H_2O bands (7900–5587 and 6849–5102 cm^{-1}) were removed from the NIR spectra (Bowers & Hanks, 1964). The region of 4464–4115 cm^{-1} was removed from the MIR spectra to account for the CO_2 peak at 4248 cm^{-1} (Sandford & Allamandola, 1990). The 10 keV XRF spectra were

subset to the range of 0.5–7.8 keV, the 40 keV XRF spectra to the range of 0–24.4 keV. The 10- and 40-keV spectra were then combined.

Calibration was done by partial least squares (PLS) regression with the kernel algorithm on the derivative spectra. Separate models were used for Fields 3 and 4, whereas a single model was used on pooled data for Fields 1 and 2 because they were close to one another. The number of components to be included in the model were selected as follows. First, the mean squared error (MSE) between the known values and the predictions was computed by leave-one-out cross-validation (LOO-CV). The standard deviation of the LOO-CV residuals was also computed. To minimise over-fitting, models were computed for a maximum of 15 components. Subsequently, the model was chosen that included the fewest components, yet lay within the MSE's standard deviation of the model that had the smallest error overall (Hastie et al., 2009, Sect. 7.10). The minimum number of components to be included was set to 1.

Preliminary analysis showed that whilst XRF tends to give the most accurate predictions of P and K, this was not always the case, and therefore all three sets of spectra were combined. The PLS predictions from NIR, MIR and XRF matrices for each property were used for an ordinary least-squares (OLS) multiple regression, known as the Granger–Ramanathan averaging method (Granger & Ramanathan, 1984). The model underlying the OLS regression in its general form is

$$\mathbf{y}_{\text{chem}} = w_0 + w_1\mathbf{z}_1 + \cdots + w_t\mathbf{z}_t + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{y}_{\text{chem}}$ is a vector of observed values (as measured by wet chemistry), $\mathbf{z}$ is a vector of PLS predictions, the $w_i, i = 1, 2, \ldots, t$, are weighting coefficients of the t individual predictors included in the regression and $\boldsymbol{\varepsilon}$ is the error term. This equation was solved for the intercept (w_0) and the t coefficients for each of the spectral matrix combinations ($\mathbf{z}$). The intercepts correct for bias if one of the individual predictors is biased.

Each prediction has associated with it an error, and these errors were treated as ones arising from the use of the OLS regression model. The error variance of these predictions, of which there are n, is hereafter referred to as

$$\text{var}\left[\varepsilon_y\right] = \frac{1}{n}\sum_{i=1}^{n}\left\{y_{\text{chem}}(i) - \widehat{y_{\text{spec}}}(i)\right\}^2 \tag{2}$$

where $y_{\text{chem}}(i)$ are the assumed true values measured by wet chemistry and $\widehat{y_{\text{spec}}}(i)$ are the predicted values from the spectra. The error variance has been propagated through into the geostatistical model. The reader is referred to Supplementary Fig. 2 for further details on the PLS and Granger–Ramanathan model-averaging results.

Geostatistical analysis

The predictions from the spectra together with the spatial coordinates and LiDAR heights were then used to predict the concentrations of each nutrient by the empirical best linear unbiased predictor (E-BLUP) (Lark et al., 2006). This is effectively a combination of kriging the spectral estimates with external drift (the LiDAR height) and universal kriging. It combines additively fixed effects (e.g. the unknown mean, the LiDAR height and the coefficients of any trend) and random effects (the spatially correlated random variation). Its model is

$$Y_{\text{krig}}(\mathbf{x}) = \alpha H(\mathbf{x}) + \sum_{j=0}^{J} \beta_j f_j(\mathbf{x}) + \varepsilon(\mathbf{x}) \tag{3}$$

Here α is the coefficient of the height, H, measured by LiDAR, the β are unknown coefficients, $\mathbf{x} \equiv x_1, x_2$ are the coordinates of a position in the field, the $f_j(\mathbf{x})$ are typically first- or second-order polynomials, and the $\varepsilon(\mathbf{x})$ represents the residuals from the fixed effects, i.e. the LiDAR height and the spatial trend. The residuals are assumed to be second-order stationary random variables, jointly normally distributed with zero means and $n \times n$ covariance matrix $\mathbf{C}_{\text{d}}$ with variogram $\gamma(\mathbf{h})$:

$$\gamma(\mathbf{h}) = \frac{1}{2}\text{E}\left[\{\varepsilon(\mathbf{x}) - \varepsilon(\mathbf{x}+\mathbf{h})\}^2\right] \tag{4}$$

where $\mathbf{h}$ is the lag in distance and direction between any two points. The random variation was treated as isotropic, so that $\mathbf{h}$ becomes a scalar in distance only: $h = |\mathbf{h}|$. The variogram of $\varepsilon(\mathbf{x})$ was examined by the method of moments. In all cases the random residuals could be successfully described by the isotropic exponential variogram model:

$$\begin{aligned} \gamma(h) &= c_0 + c_1\left\{1 - \exp\left(-\frac{h}{a}\right)\right\} \text{ for } 0 < h \\ &= 0 \text{ for } h = 0 \end{aligned} \tag{5}$$

or the isotropic spherical variogram model:

$$\begin{aligned} \gamma(h) &= c_0 + c_1\left\{\frac{3h}{2r} - \frac{1}{2}\left(\frac{h}{r}\right)^3\right\} \text{ for } 0 < h < r \\ &= c_0 + c_1 \text{ for } h \geq r \\ &= 0 \text{ for } h = 0 \end{aligned} \tag{6}$$

Here c_0 and c_1 are variances, respectively the nugget and sill of the correlated variance, and r (the range of the spherical function) and a are distance parameters. Parameters for a plausible model can be found by maximum likelihood (ML) or maximization of the likelihood of the residuals given the data (REML). Lark et al. (2006) and Webster and Oliver (2007) give the derivation of the equations in full. The REML estimation method is preferred, because it reduces bias in random effects parameters due to the uncertainty in the fixed effects parameters. However, the ML may be compared between models with different fixed effects structures, but such a comparison is not valid for REML. Therefore, the ML method was used first to select the fixed effects.

Preliminary investigations (visual inspection and marginal plots) suggested that the estimated soil properties, $\widehat{y_{\text{spec}}}$, have long-range trends across the fields. These were considered as fixed effects. Estimated soil properties also vary systematically with elevation as recorded by LiDAR, so this was considered as another fixed effect. Each trend variable (up to quadratic terms in eastings and northings, and elevation) was added in turn and tested whether its addition was significant by a log-likelihood ratio test. A chi-squared p -value from the log-likelihood ratio of 0.05 was taken as threshold and any smaller value ($p \leq 0.05$) was treated as evidence that additional trend parameters should be included. After the selection of fixed effects, the fixed effects and random effects in Eq. (3) were re-estimated using residual maximum likelihood (REML). If c_0 was less than var$\left[\varepsilon_y\right]$ then it was set equal to var$\left[\varepsilon_y\right]$ and Eq. (3) was solved again. Next, the final variograms were used for universal kriging. This provided the predictions and their kriging variances.

The linear mixed models were cross-validated by LOO-CV. The linear mixed models were re-estimated for each iteration to diminish bias in parameter values (Hastie et al., 2009, Sect. 7.10). The LOO-CV of the linear mixed-models were evaluated with the mean- and median-standardized squared prediction errors (SSPEs) (Lark, 2000).

Theory of the loss function

The loss function, $L(F)$, is defined as the difference in profit that results from applying a given amount of fertiliser F compared with the economic optimal amount F_0:

$$L(F) = \Phi\left(F_0\right) - \Phi(F) \tag{7}$$

where the profit $\Phi(F)$ is the difference between the income from the crop (price of the crop × yield) and the cost of the fertiliser:

$$\Phi(F) = M \times \text{Yield} - V \times F \tag{8}$$

where M is the price of the crop (£ t^{-1}) and V is the cost of the fertiliser (£ kg^{-1}). It was assumed that the yield is given by the dose–response equation:

$$\psi = \alpha + \eta R^{\xi F + S} + \nu(\xi F + S) \tag{9}$$

where ψ is the yield, S is the concentration of the nutrient in the soil, F is the applied fertiliser (kg ha^{-1}), ξ is the increase in nutrient concentration (mg kg^{-1}) in the soil for every 1 kg ha^{-1} fertiliser applied, and α, η, ν and R are parameters, then the optimum amount of fertiliser can be calculated from this and is given by

$$F_0 = \ln\left(\frac{B/\xi - \nu}{\eta R^S \ln R}\right)/\xi ln R \tag{10}$$

where $B = V/M$, known as the break-even ratio.

By definition, the loss given by Eq. (7) is zero when the optimum amount of fertiliser is applied. However, computing the optimum amount of fertiliser to apply relies on one's knowing the nutrient status, S, in the soil, and one cannot know it precisely.

Given the probability distribution, $g(s)$, of the nutrient status S the optimum fertiliser rate that maximizes the expected profit can be computed:

$$F_{\text{opt}} = \ln\left(\frac{B/\xi - \nu}{\eta \ln R \int_0^\infty R^s g(s)\mathrm{d}s}\right)/\xi \ln R \tag{11}$$

This also minimises the expected loss function, E[$L(F)$], which was defined here as the difference between the profit where S is known without error and associated optimum fertiliser, F_0 in Eq. (10), is applied

$$\mathrm{E}[L(F,S)] = \Phi\left(F_0, S\right) - \int_0^\infty \{\Phi(F,s)\} g(s)\mathrm{d}s \tag{12}$$

Table 1 Dose—response equation parameters as relevant to Eq. 9

Soil nutrient	α	η	ν	R	ξ
P	142.15	−145.8	−0.776	0.98	0.15
K	63.3	−63.3	0	0.98	0.52

Parameterization and analysis of the loss function

All fields sampled were cultivated for lettuce, and so loss functions were computed associated with this crop. The dose response curve for P was derived from Prasad et al. (1988) and for K from Greenwood et al. (1980). Parameters were estimated for the linear plus exponential functions with the Gauss–Newton algorithm in GenStat (VSN International, 2021). It was assumed that for every 1 kg of P added in fertiliser 0.18 kg becomes available to the crop (Muhammed et al., 2017), for every 1 kg of K added in fertiliser, 0.62 kg becomes available to the crop (Blake et al., 1999). Furthermore, it was assumed that the added nutrients are contained in the top 0.25 m of the soil (the sampling depth). The value of 480 kg m^{-3} for bulk density of this peat soil was taken from Milne et al. (2006). Given the support of the kriged predictions (2 m × 2 m), it follows that an addition of 1 kg fertiliser per ha leads to an increase in the concentration of this layer of 0.15 mg available P kg^{-1} and 0.52 mg available K kg^{-1}, equal to ξ in the dose–response Eq. (9) (see details in Supplementary Material). Greenwood et al. (1980) listed a mean base nutrient concentration of 69 mg available K kg^{-1} for the unfertilised soil in their study, which was used as an additive component. A profit margin (M) of £90 per tonne of lettuce per hectare was assumed. The prices of fertiliser (V) were taken as £0.36 per kg P fertiliser and £0.29 per kg K fertiliser. Table 1 lists the parameter values of the dose–response equations for P and K.

The profitability of variable-rate application (VRA) was assessed based on kriged maps by computing the total expected loss (Eq. 12) across each field. This was done by comparing the expected loss for each field between that from VRA, $\mathrm{E}\left[L\left(F_{\mathrm{opt}}\right]\right)$, and a uniform application (UA) based on wet chemistry alone, $\mathrm{E}\left[L\left(F_{\mathrm{UA}}\right]\right)$. For this purpose, five samples per field were selected in a W-shape from the locations at which the nutrients had been measured by wet chemistry. These five samples were used to compute an average soil nutrient concentration per field that was used to determine the uniform rate of fertiliser application.

Software

Analysis was done with base R commands as well as the following R packages as implemented in RStudio: data handling with the **tidyverse** (Wickham et al., 2019) package, computation of the sampling designs using the **spcosa** (Walvoort et al., 2010), **BalancedSampling** (Grafström & Lilics, 2019) and **SpatialEco** (Evans, 2019) packages, spectral processing using **prospectr** (Stevens & Ramirez-Lopez, 2013), partial least squares regression using **pls** (Bjørn-Helge et al., 2019), Granger–Ramanathan averaging using **GeomComb** (Weiss & Roetzer, 2016), model-based geostatistics using the **geoR** package (Ribeiro & Diggle, 2018) and handling of spatial objects using the **raster**

Table 2 Fixed effects and parameters estimated by REML of the variograms, H stands for LiDAR (elevation height in metres), and x_1 and x_2 are the spatial coordinates

Field	Soil nutrient	Fixed effects	Variogram type	Variogram parameters			
				c_0	c_1	r/m	a/m
1	P	$x_1, x_2, x_1^2, x_2^2, x_1x_2$	Sph	23	59	172	–
	K	$H, x_1, x_2, x_1^2, x_2^2, x_1x_2$	Sph	4100	1784	104	–
2	P	H, x_1, x_2	Exp	22	50	–	35
	K	H, x_1, x_2	Exp	3200	5551	–	64
3	P	$H, x_1, x_2, x_1^2, x_2^2, x_1x_2$	Exp	39	38	–	59
	K	$x_1, x_2, x_1^2, x_2^2, x_1x_2$	Exp	1100	1413	–	64
4	P	H	Exp	15	151	–	12
	K	$H, x_1, x_2, x_1^2, x_2^2, x_1x_2$	Exp	720	1565	–	12

The variogram parameters are the nugget variance (c_0) the sill of the correlated variance (c_1), and r and a are the distance parameters

Sph spherical, *Exp* exponential

(Hijmans, 2020) and **rgdal** (Bivand et al., 2020) packages. Graphics were created with base R functions and the package **ggplot2** (Wickham, 2016).

Results

Uncertainty in kriging predictions from soil properties estimated by spectroscopy

The nugget variances, c_0, were underestimated for the following variogram models: available K (Fields 1, 2, 3, and 4) and available P (Fields 1 and 2). The nugget parameter was therefore set equal to $\text{var}\left[\varepsilon_y\right]$ and Eq. (3) was solved again to account for the under-estimation of the error. Data of both P and K in all four fields were fitted with a linear trend model as fixed effects (Table 2). Fitting trend coefficients, as expected, resulted in smaller semivariances than their equivalents of the original variables, i.e. the difference between black discs and circles (Fig. 1). The LOO-CV results of the mixed model variograms accorded overall with expectations for both P and K and all four fields (Supplementary Fig. 3).

Loss function on variable-rate fertiliser application

The fitted dose–response curves for P and K affected the profit $\Phi(F)$ and the loss function $L(F) = \Phi\left(F_0\right) - \Phi(F)$ differently because of their asymmetry characteristics and by association the expected loss, i.e. loss from perfect knowledge, $\text{E}[L(F)]$ in Eq. (12). For available P, the dose–response curve declines linearly in yield for large values of P (Fig. 2).

That is to say over-application of fertiliser results in financial losses because too much fertiliser is applied and yield is diminished. This effect of over-application of P fertiliser to soil with a large concentration of P fertiliser is greater than a similar over-application of K fertiliser; for K where the dose–response curve increases to an asymptote, and so the associated prediction of profit and loss converge at large rates of applied fertiliser (Figs. 3 and

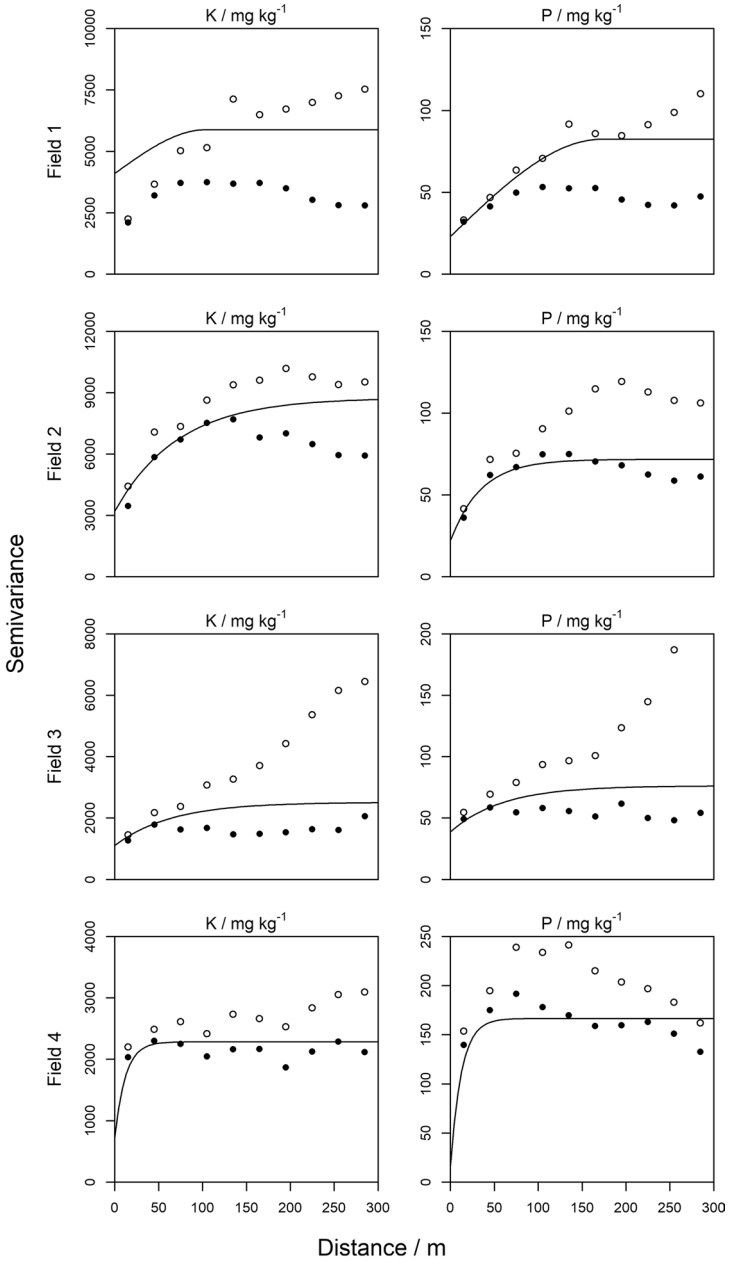


Fig. 1 Linear mixed model variograms for all fields, ∘ refer to the experimental variograms of the original variable, • refer to the experimental variograms of the residuals from the trend model, the solid black line refers to the final model fitted by restricted maximum likelihood procedures

4). For both P and K, larger uncertainty of the predicted nutrient content (large σ^2_{krig}) increases the expected loss and reduces the profit.

As expected, the fertiliser rate that minimises the expected loss (F_{opt}) is greater than the optimal fertiliser rate when the soil variable is known without error (F_0) for all fields, again because of the asymmetry of the loss functions. The asymmetry means that overestimation generally leads to larger losses than does underestimation of soil P and K. Error variance in the estimates of P and K in the soil consequently leads to larger recommended applications of fertiliser than if one had perfect knowledge ($\sigma^2_{krig} = 0$) (Fig. 5).

The optimal P fertiliser application varied substantially in all fields. The variation was less pronounced for K, particularly in Field 4. Across Field 1, most kriged estimates of K fall on the asymptote of the dose–response curve, hence $F_0 = 0$ for a large part of the field (Fig. 6A). The σ^2_{krig} increases the probability that kriging estimates fall below the asymptote, however. In those situations application of fertiliser becomes necessary to minimise the expected loss (Fig. 6B). These observations also hold for Fields 2 and 4. The kriged estimates of available K in Fields 1, 2 and 4 were larger than the range of the dose–response curve (Fig. 7). Consequently, applying no fertiliser for a major portion of the field was more profitable (Fig. 6B, D and H).

The total expected loss on a field-basis was less for variable-rate P and K application ($\mathrm{E}\left[L\left(F_{opt}\right)\right]$) than the total loss of blanket fertiliser application arising from the wet chemical analysis ($\mathrm{E}\left[L\left(F_{UA}\right)\right]$) (Table 3). There was a financial incentive of VRA of P fertiliser across all fields (ranging from £7–£121 ha^{-1}) and for K fertiliser across three fields (ranging from £6–£81 ha^{-1}). That is, for available K the difference between $\mathrm{E}\left[L\left(F_{opt}\right)\right]$ and $\mathrm{E}\left[L\left(F_{UA}\right)\right]$ was small in Field 4. Less P fertiliser was used on a field-basis under VRA (F_{opt}) than with uniform application (F_{UA}) in Fields 1, 2 and 3 (Table 3). Within those fields, total P fertiliser use was reduced by VRA with 4–17 kg ha^{-1} compared with uniform application. Most K fertiliser would be used on a field-basis in all cases under VRA (F_{opt}) (Table 3).

Discussion

Error approximation and the estimation of the variogram

For this study it was assumed that the error was homogeneous across the sample locations that were predicted by spectroscopy. This allowed for the more generic assumption that the prediction library already existed (for example, derived from a spectral library). If, however, predictions were obtained from a single sampling campaign, the error variance of locations with associated wet chemistry data and those without could be separated in a subsequent geostatistical analysis, as Delhomme (1978) suggested.

Error propagation is rarely taken into account in soil surveys based on spectroscopy. Ramirez-Lopez et al. (2019), listed two other studies in which the propagation of errors was reported in the last 10 years (Brodský et al., 2013; Viscarra Rossel et al., 2016). Somarathna et al. (2018) and Ellinger et al. (2019) also propagated errors from infrared spectral into predictions of soil carbon. Error propagation is important for two reasons. First, Somarathna et al. (2018) found that acknowledging the measurement error, in this case $\mathrm{var}\left[\varepsilon_y\right]$, reduces uncertainty in spatial predictions (as supported by Clark et al., 2010). The extent to which this has an effect will depend on the complexity of the target variable's spatial variation and the geographical extent of the study. Second, acknowledgement of the uncertainty (and its minimisation) is necessary to detect small rates of change in the soil property of interest by monitoring over time (Viscarra Rossel et al., 2016). This study is

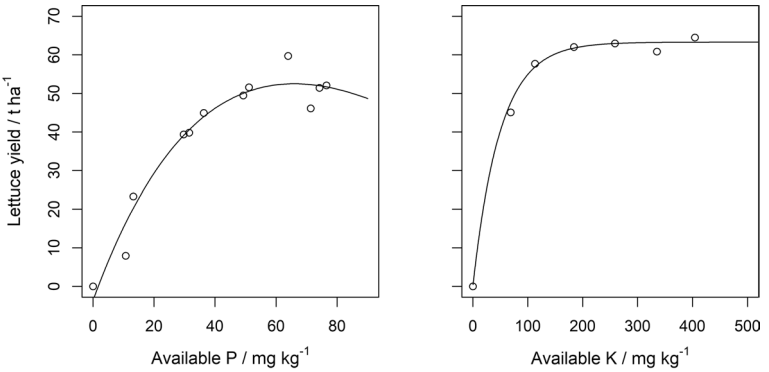


Fig. 2 Dose–response curves for available P (exponential + linear) and available K (exponential) fitted based on data from Prasad et al. (1988) (P) and Greenwood et al. (1980) (K). See Table 1 in the main text for parameter values

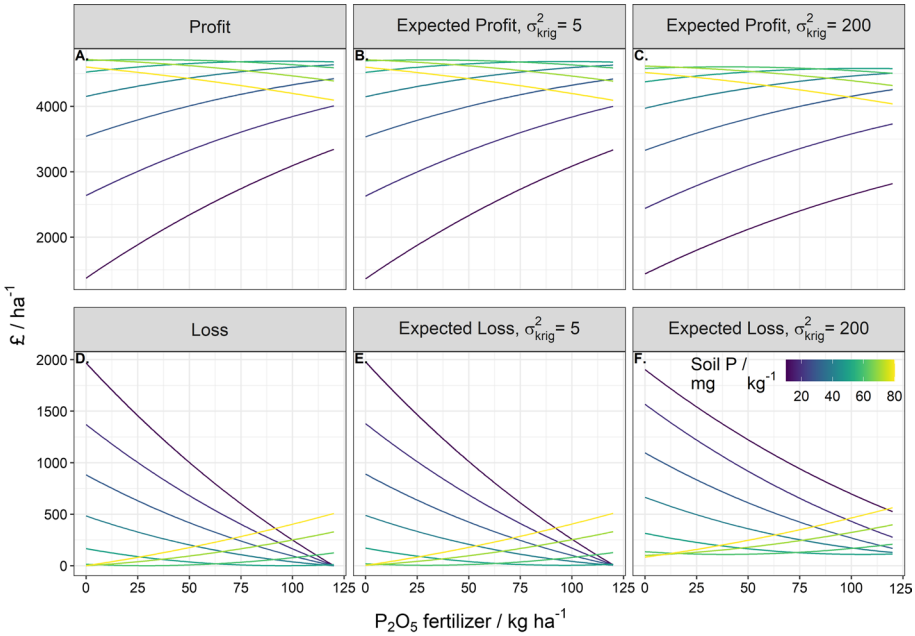


Fig. 3 Profit and loss under zero error variance, expected profit and loss under an error variance of 5 mg kg^{-1} and an error variance of 200 mg kg^{-1} for a range of estimated soil P values from 10 to 80 mg kg^{-1}. The range of P fertiliser applied spans 0 to 120 kg ha^{-1}

concerned with variation in space, and it is this variation that determines whether variable rate application is relevant.

One merit of kriging is that in addition to providing unbiased predictions it minimises the squared-errors of those predictions, which are known. The predictions are best in that sense. Kriging, like other forms of regression, smooths: unobserved small

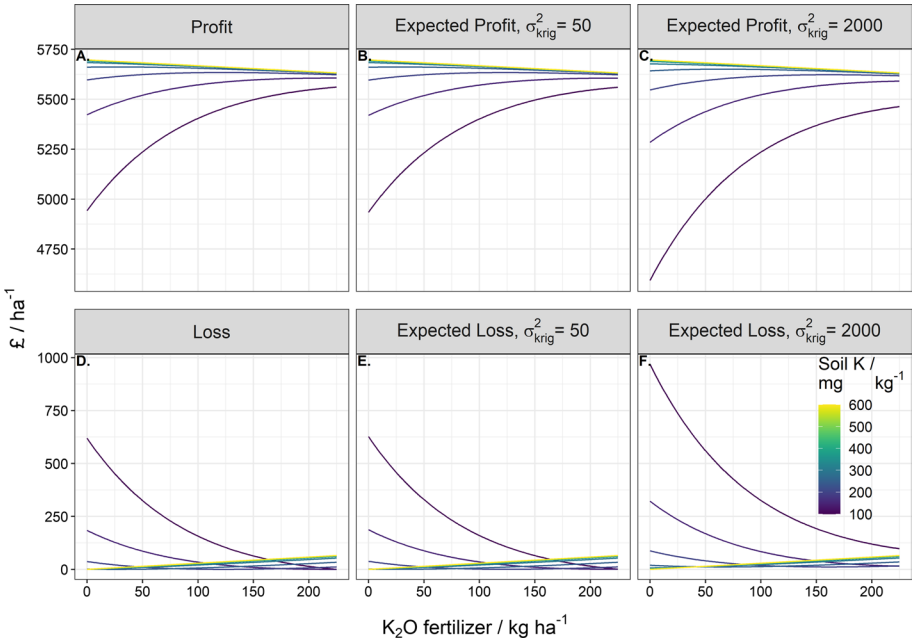


Fig. 4 Profit and loss under zero error variance, expected profit and loss under an error variance of 50 mg kg^{-1} and an error variance of 2000 mg kg^{-1} for a range of estimated soil K values from 100 to 600 mg kg^{-1}. The range of K fertiliser applied spans 0 to 225 kg ha^{-1}

values are over-estimated, and large ones under-estimated (Webster & Oliver, 2007). Smoothing thus leads to underestimation of the spatial variance, a consequence one needs to consider in using the predictions for determining variable rates of fertiliser.

LiDAR was included as a fixed effect in six of the eight LMMs, and all LMMs except one included geographic trends (i.e. in the spatial coordinates) in the model. Kriging within REML is based in the assumption of second-order stationarity of the random part of the process. That is why the fixed effects of trend and LiDAR were separated out and also why each field was treated separately. Having an exhaustive covariate allowed one to do that and to approximate the uncertainty of the target variable more accurately than otherwise (Lark, 2009).

Data requirements and estimation of the loss function

The loss function requires certain requirements of, or assumptions about, data. For example, the soil's bulk density was estimated from general knowledge in the area (see method Sect. 3.4). It is known that the density of soil on the rodhams differs from that of the peaty soil between them. Even in the best scenario, these estimates embody an error which should ideally be accounted for. Similarly, the modelled response of the crop contains error and this should be incorporated in the framework, although this was not included for reasons of clarity.

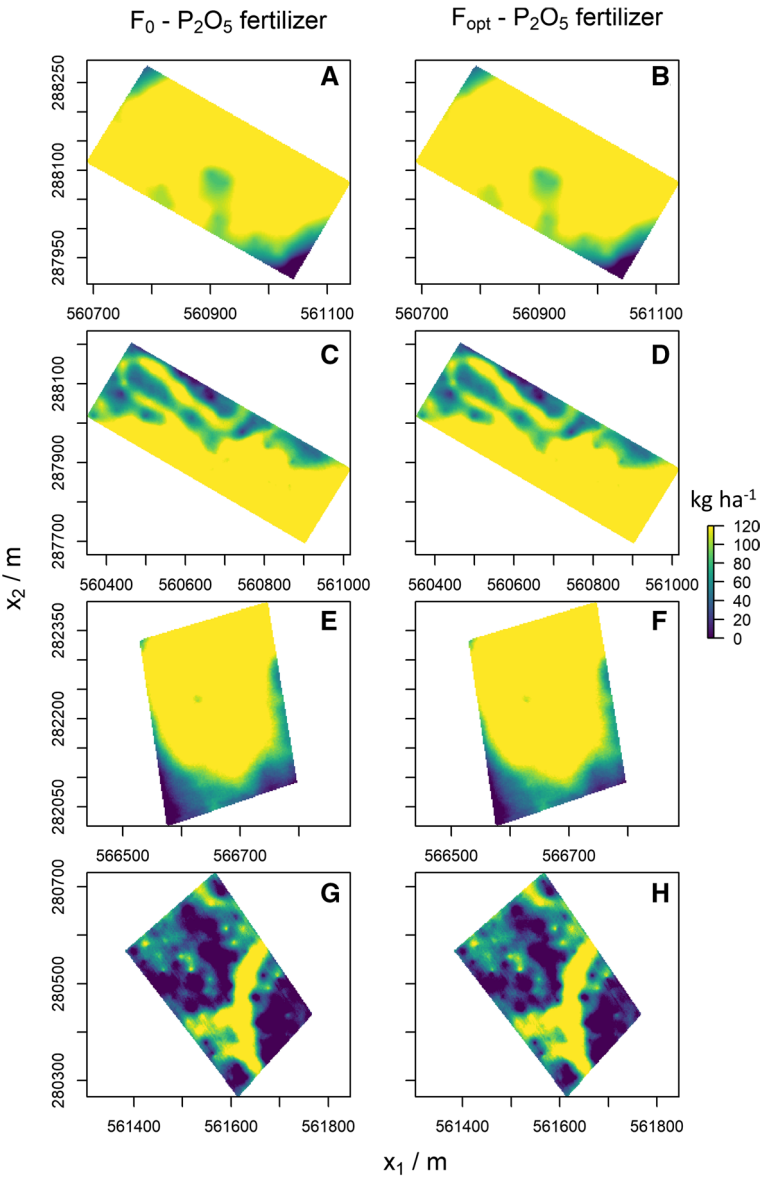


Fig. 5 Optimum P fertiliser application with perfect knowledge of soil P (F_0), optimum application when accounting for σ^2_{krig} in the estimate of soil P (F_{opt})

The loss function to estimate the value of variable-rate application

Based on the differences in E[$L(F)$] between F_{opt} and F_{UA}, there appears to be little financial incentive for variable-rate application of K fertiliser in Field 4 (ΔE[$L(F)$] = 1, Table 3). The difference in E[$L(F)$] between F_{opt} and F_{UA} is larger for Fields 1, 2 and 3 (with ΔE[$L(F)$] equal to 7, 6 and 81, respectively). For available P, most kriged estimates

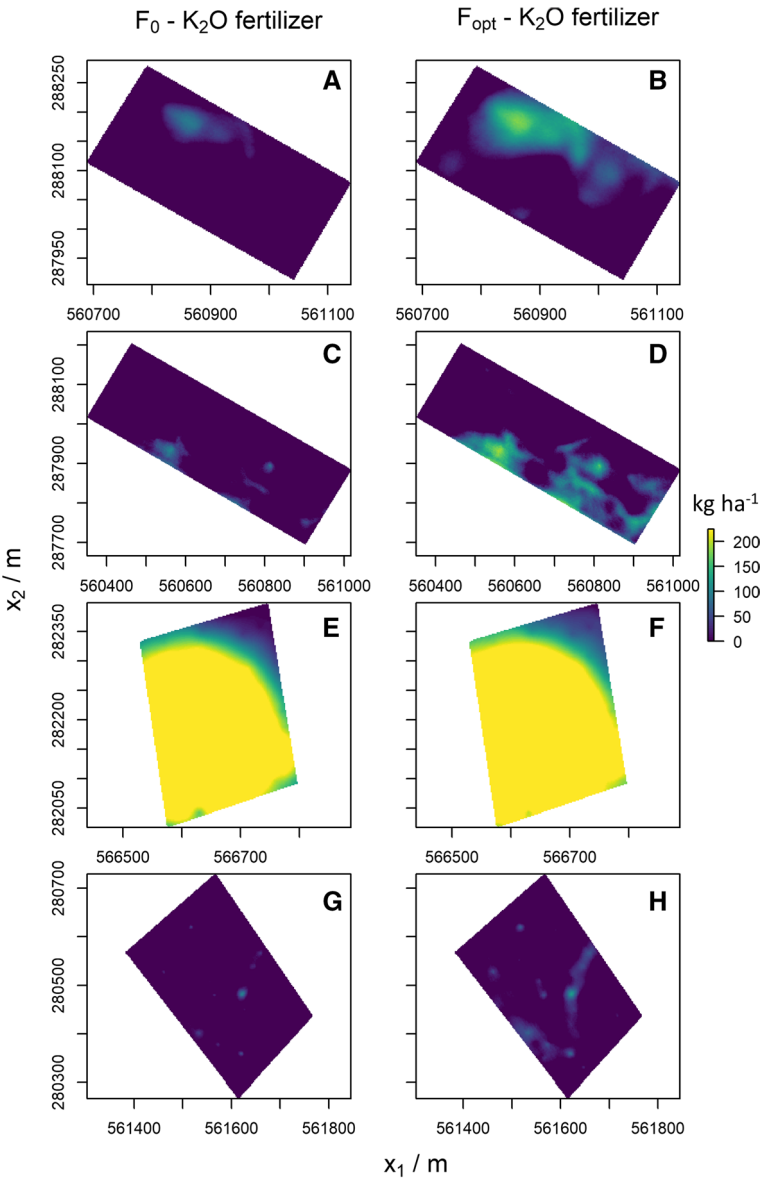


Fig. 6 Optimum K fertiliser application with perfect knowledge of soil K (F_0), optimum application when accounting for σ^2_{krig} in the estimate of soil K (F_{opt})

lay in the linearly increasing range of the dose–response curve, and there is a financial incentive to implement variable rate application for all fields.

The difference in total K fertiliser used between F_0 and F_{opt} was especially large for Fields 1, 2 and 3, which can be explained by the large nugget variance (c_0) and sill (c_1) (Fig. 1 and Table 2). Field 4 has the largest E[$L(F)$] values for P under VRA (Table 3), which can be attributed to large values of σ^2_{krig} (Fig. 7), a short distance parameter (a) and

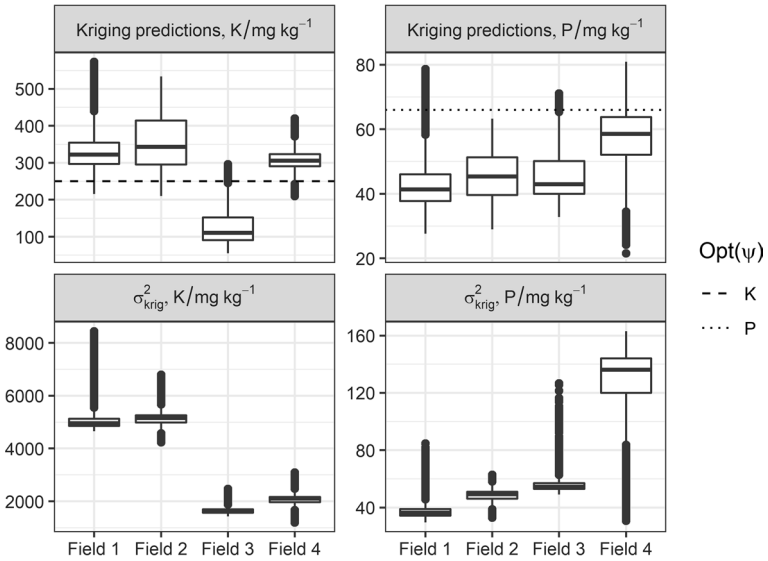


Fig. 7 Box-plots of kriging predictions and the kriging variance (σ^2_{krig}), by field for available P and K, horizontal lines represent the nutrient value for which the maximum yield, Opt(ψ), is obtained on the fitted dose–response curve

Table 3 Fertiliser used (kg ha^{-1}) for perfect knowledge (F_0), variable-rate application (F_{opt}) and uniform application based on wet chemistry samples (F_{UA})

Field	Area (ha)	Soil nutrient	Fertiliser (kg ha^{-1})			Expected loss (£ ha^{-1})		
			F_0	F_{opt}	F_{UA}	$E[L(F_{opt})]$	$E[L(F_{UA})]$	$\Delta E[L(F)]$
1	8.2	P	113	113	120	31	38	7
2	16.9		102	103	120	38	52	14
3	5.1		103	104	61	33	154	121
4	8.9		47	53	57	94	145	51
1	8.2	K	4	27	0	16	23	7
2	16.9		3	24	0	13	19	6
3	5.1		197	205	144	69	150	81
4	8.9		0	4	0	6	7	1

Expected loss (from perfect knowledge) for variable-rate application, $E[L(F_{opt})]$, and uniform application, $E[L(F_{UA})]$, with $\Delta E[L(F)]$ being the difference in expected loss given by $E[L(F_{UA})] - E[L(F_{opt})]$

large sill (c_1) in the variogram (Fig. 1 and Table 2). Additionally, the smaller applications of P fertiliser under VRA than under uniform application for Fields 1, 2 and 4 means that VRA poses less environmental damage than uniform application would; that is, there would be less P lost from the soil to pollute waterways and cause eutrophication.

The large expected loss under uniform application of P in Fields 3 and 4 can be attributed to a biased estimate of the mean concentration of soil P across the field and hence blanket fertilizer recommendation, F_{UA}. Sampling by a W-design has been found to be equivalent to random sampling in the estimation of a mean concentrations of nutrients

(Marchant et al., 2012). However, because the samples that make up the W-design were chosen a posteriori (the samples come from the set that were analysed by wet chemistry and these were selected purposely to span the range in the field) the mean estimate of available P and K is suspected to be biased.

It is further noted that the expected loss under uniform application of K fertiliser is large for Field 3. The large expected loss is likely due to the effect of uncertainty being more pronounced for smaller values of available K (Fig. 4D–F). Additionally, the uncertainty in available K predictions was relatively larger than the range of available K in the calibration set (Supplementary Fig. 2).

Overall, the expected loss, E[$L(F)$], and hence F_{opt} was found to depend on (a) the kriging variance, (b) the ranges of P and K for which the dose–response curves were calibrated, (c) the range of estimated values in the fields and (d) the asymmetry of the loss function. These factors need to be properly quantified, parameterized and accounted for in the loss functions so that farmers can make their decisions with confidence, while taking into account uncertainty.

Implications of the loss function approach on decision-making

Although quantification of uncertainty (based on data and current models) allows one to make statements with confidence, it can also identify where the effort of reducing uncertainty will result in the largest gains. The loss function has enabled scientists and managers to decide how much field-work and analysis is required to answer specific questions in environmental monitoring. For example, it is used to optimise the size of samples for survey; see Yates (1981), Faechner et al. (2000), Goovaerts (1997), Lark and Knights (2015) for examples. Relevant decisions within a sampling campaign involve (a) where and when to take samples, (b) what measurements to make on the samples, and (c) with what accuracy to take these measurements.

The loss function framework provides a method to assess the quality of predictions from spectroscopy beyond specific metrics such as R^2 and investigate whether the accuracy is "sufficient" to address relevant questions. In this particular study to test the hypothesis whether soil spectroscopy could adequately predict the spatial variability to justify variable rate application of P and K fertiliser.

For example, if sampling, handling and spectroscopy costs are less than the difference between $\mathrm{E}\left[L\left(F_{\text{opt}}\right)\right]$ and $\mathrm{E}\left[L\left(F_{\text{UA}}\right)\right]$ then VRA is worthwhile. These costs depend amongst other things on both the total number of samples and the number of samples for calibration. The costs of spectroscopy per sample will also decrease with increased size of field. The smallest estimate of cost was for Field 2 (the largest in area); it was £49 ha^{-1} for P and £47 ha^{-1} for K. Although these costs fall within the range of cost savings of VRA compared with UA (£7–121 ha^{-1}), they are larger than the savings for Field 2 specifically. However, this study investigated a best-case scenario for spectroscopy and did not optimize the data-collection for cost-effectiveness. One could reduce these costs by measuring the reflectance spectra of the soil surface on the run in the field. So far, however, trials to estimate P and K in the soil from visible–NIR in the field by Cozzolino et al. (2013), Daniel et al. (2003), Kuang et al. (2012) and Ji et al. (2014) have produced disappointing results. Reports of R^2 values lie in the range 0.09–0.87 for predictions of available P and 0.03–0.87 for available K. No reports were found that support within-field estimation of available P and K from MIR and XRF spectroscopy, which holds potential for further research.

Within six out of eight fields in this study, the cost-effectiveness of VRA was primarily driven by increases in yield, in some cases at the cost of applying more fertiliser compared to uniform application. The excess use of potash fertiliser does not pose a direct threat to the environment. The overuse of phosphate fertiliser on the other hand in Field 3, might minimise economic loss at the cost of the environment. However, one could argue that increased efficiency gained by VRA also has potential to reduce the total land area used to reach equivalent yields. Setting aside agricultural land in areas vulnerable to leaching is a recognized strategy to manage phosphorus concentrations at the catchment scale (Schoumans et al., 2014). In that case one would need to ensure that leaching or artificial drainage is not the driving force behind small concentrations of P at the field-scale in the first place (Baveye & Laba, 2015). The results further showed that the environmental benefit of fertiliser savings from VRA (in the range of 4–17 kg ha^{-1}) was not strictly accompanied with a large increase in profit (Field 1). Hence, in order to account for the environmental benefits of precise fertiliser application, the results suggest that costs of P leaching (e.g. remediation) need to be included in sustainable phosphorus management strategies aimed at precise fertiliser application. These costs of P leaching can be substantial. For example, if the excess fertiliser finds its way into water bodies, the cost to water companies has been estimated to range between 75 and 114 million pounds sterling per year for England and Wales (Pretty et al., 2003). Costs associated with eutrophication range among others from restorative measures such as dredging, treatment of drinking water (including removal of algal toxins), loss of important species and ecological damage generally (Pretty et al., 2003). The presented framework for estimating financial losses could be adapted to place a larger penalty on over-application of phosphorus fertiliser based on the environmental costs of leaching. It could provide a stepping stone towards fulfilling the requirement to quantify the economic and environmental benefits of sustainable phosphorus management (Brownlie et al., 2021).

Conclusions

The results show that there was an economic incentive for precise fertiliser application of both phosphorus and potassium fertiliser once the uncertainty in soil's nutrient concentrations estimated from sensors was accounted for. Given that growers need to subtract the costs of sampling and sample analysis from their gross income, further study should use the loss function to define an optimum where both uncertainty of information and the effort to collect the data by sampling and analysis are minimised.

To quantify society's benefits of precise fertiliser application holistically, however, environmental costs need to be taken into consideration. The results showed that environmental benefits occurred from precise fertiliser application even though no large increase in profit was gained. These findings have implications for policies aimed at sustainable management of phosphorus fertilisers. That is, it is recommended that the loss function could be adapted to include environmental costs of P leaching to assist in quantifying both the economic and environmental benefits of precise fertiliser application.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11119-022-09887-2.

Funding Funding was provided by Biotechnology and Biological Sciences Research Council (Grant Nos. BBS/E/C/000I0320, BBS/E/C/000I0330, BBS/E/C/000I0100).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Baveye, P. C., & Laba, M. (2015). Moving away from the geostatistical lamppost: Why, where, and how does the spatial heterogeneity of soils matter? *Ecological Modelling, 298,* 24–38. https://doi.org/10.1016/j.ecolmodel.2014.03.018

Bellon-Maurel, V., & McBratney, A. (2011). Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils—Critical review and research perspectives. *Soil Biology & Biochemistry, 43*, 1398–1410. https://doi.org/10.1016/j.soilbio.2011.02.019

Beusen, A. H. W., Bouwman, A. F., Van Beek, L. P. H., Mogollón, J. M., & Milddelburg, J. J. (2016). Global riverine N and P transport to ocean increased during the 20th century despite increased retention along the aquatic continuum. *Biogeosciences, 13*, 2441–2451. https://doi.org/10.5194/bg-13-2441-2016

Bivand, R., Keitt, T., & Rowlingson, B. (2020). *rgdal: Bindings for the `Geospatial' Data Abstraction Library*. R package version 1.5-12. Retrieved December 8, 2020, from https://CRAN.R-project.org/package=rgdal

Bjørn-Helge, M., Wehrens, R., & Hovde Liland, K. (2019). *pls: Partial least squares and principal component regression*. R package version 2.7-1. Retrieved February 5, 2020, from https://CRAN.R-project.org/package=pls

Blake, L., Mercik, S., Koerschens, M., Goulding, K. W. T., Stempen, S., Weigel, A., Poulton, P. R., & Powlson, D. S. (1999). Potassium content in soil, uptake in plants and the potassium balance in three European long-term field experiments. *Plant and Soil, 216*, 1–14. https://doi.org/10.1023/a:1004730023746

Bowers, S. A., & Hanks, R. J. (1964). Reflection of radiant energy from soils. *Soil Science, 100*, 130–138. https://doi.org/10.1097/00010694-196508000-00009

Breure, T. S., Milne, A. E., Webster, R., Haefele, S. M., Hannam, J. A., Moreno-Rojas, S., & Corstanje, R. (2021). Predicting the growth of lettuce from soil infrared reflectance spectra: The potential for crop management. *Precision Agriculture, 22*, 226–248. https://doi.org/10.1007/s11119-020-09739-x

Brodský, L., Vašát, R., Klement, A., Zádorová, T., & Jakšík, O. (2013). Uncertainty propagation in VNIR reflectance spectroscopy soil organic carbon mapping. *Geoderma, 199*, 54–63. https://doi.org/10.1016/j.geoderma.2012.11.006

Brownlie, W. J., Sutton, M. A., Reay, D. S., Heal, K. V., Hermann, L., Kabbe, C., & Spears, B. M. (2021). Global actions for a sustainable phosphorus future. *Nature Food, 2*, 71–74. https://doi.org/10.1038/s43016-021-00232-w

Chen, M., & Graedel, T. E. (2016). A half-century of global phosphorus flows, stocks, production, consumption, recycling, and environmental impacts. *Global Environmental Change, 36*, 139–152. https://doi.org/10.1016/j.gloenvcha.2015.12.005

Cherry, K. A., Shepherd, M., Withers, P. J. A., & Mooney, S. J. (2008). Assessing the effectiveness of actions to mitigate nutrient loss from agriculture: A review of methods. *Science of the Total Environment, 406*, 1–23. https://doi.org/10.1016/j.scitotenv.2008.07.015

Clark, I. (2010). Statistics or geostatistics? Sampling error or nugget effect? *Journal of the Southern African Institute of Mining and Metallurgy, 110*(6), 307–312.

Cozzolino, D., Cynkar, W. U., Dambergs, R. G., Shah, N., & Smith, P. (2013). In situ measurement of soil chemical composition by near-infrared spectroscopy: A tool toward sustainable vineyard management. *Communications in Soil Science and Plant Analysis*, *44*, 1610–1619. https://doi.org/10.1080/00103624.2013.768263

Daniel, K. W., Tripathi, N. K., & Honda, K. (2003). Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Australian Journal of Soil Research, 41*, 47–59. https://doi.org/10.1071/SR02027

Delhomme, J. P. (1978). Kriging in the hydrosciences. *Advances in Water Resources, 1*, 251–266. https://doi.org/10.1016/0309-1708(78)90039-8

Ellinger, M., Merbach, I., Werban, U., & Ließ, M. (2019). Error propagation in spectrometric functions of soil organic carbon. *The Soil, 5*, 275–288. https://doi.org/10.5194/soil-5-275-2019

Evans, J. S. (2019). *spatialEco*. R package version 1.2-0. Retrieved October 3, 2019, from https://github.com/jeffreyevans/spatialEco

Faechner, T., Pyrcz, M., & Deutsch, C. V. (2000). Soil remediation decision making in presence of uncertainty in crop response. *Geoderma, 97*, 21–38. https://doi.org/10.1016/S0016-7061(00)00024-0

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.

Goovaerts, P. (2001). Geostatistical modelling of uncertainty in soil science. *Geoderma, 103*, 3–26. https://doi.org/10.1016/S0016-7061(01)00067-2

Grafström, A., & Lisic, J. (2019). *Balanced sampling: Balanced and spatially balanced sampling*. R package version 1.5.5. Retrieved March 1, 2021, from https://cran.r-project.org/web/packages/BalancedSampling/index.html

Granger, C. W. J., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting, 3*, 197–204. https://doi.org/10.1002/for.3980030207

Greenwood, D. J., Cleaver, T. J., Turner, M. K., Hunt, J., Niendorf, K. B., & Loquens, S. M. H. (1980). Comparison of the effects of potassium fertilizer on the yield, potassium content and quality of 22 different vegetable and agricultural crops. *The Journal of Agricultural Science, 95*, 441–456. https://doi.org/10.1017/S0021859600039496

Hastie, T., Friedman, J., & Tibshirani, R. (2009). *The elements of statistical learning: Data mining, inference and prediction, 2nd printing*. New York: Springer.

Heuvelink, G. B. M. (2018). Uncertainty and uncertainty propagation in soil mapping and modelling. In A. B. McBratney, B. Minasny, & U. Stockmann (Eds.), *Pedometrics* (pp. 439–461). Dordrecht: Springer.

Hijmans, R. J. (2020). *raster: Geographic data analysis and modeling*. R package version 3.3-13. Retrieved December 8, 2020, from https://CRAN.R-project.org/package=raster

Hodge, C. A. H., Burton, R. G. O., Corbett, W. M., Evans, R., & Seale, R. S. (1984). *Soils and their use in eastern England*. Soil survey of England and Wales Bulletin No 13. Harpenden: Lawes Agricultural Trust.

Ji, W. J., Shi, Z., Huang, J. Y., & Li, S. (2014). In-situ measurement of some soil properties in paddy soils using visible and near infrared spectroscopy. *PLoS ONE, 9*(8), e105708.

Kuang, B., Mahmood, H. S., Quraishi, M. Z., Hoogmoed, W. B., Mouazen, A. M., & van Henten, E. J. (2012). Chapter four - sensing soil properties in the laborarity, *in situ*, and on-line: A review. *Advances in Agronomy, 114,* 155–223. https://doi.org/10.1016/B978-0-12-394275-3.00003-1

Lark, R. M. (2000). Estimating variograms of soil properties by the method-of-moments and maximum likelihood. *European Journal of Soil Science, 51*, 717–728. https://doi.org/10.1046/j.1365-2389.2000.00345.x

Lark, R. M. (2009). Kriging a soil variable with a simple nonstationary variance model. *The Journal of Agricultural, Biological and Environmental Statistics, 14*, 301–321. https://doi.org/10.1198/jabes.2009.07060

Lark, R. M., Cullis, B. R., & Welham, S. J. (2006). On spatial prediction of soil properties in the presence of a spatial trend: The empirical best linear unbiased predictor (E-BLUP) with REML. *European Journal of Soil Science, 57*, 787–799. https://doi.org/10.1111/j.1365-2389.2005.00768.x

Lark, R. M., & Knights, K. V. (2015). The implicit loss function for errors in soil information. *Geoderma, 251–252*, 24–32. https://doi.org/10.1016/j.geoderma.2015.03.014

Li, S., Shi, Z., Chen, S. C., Ji, W. J., Zhou, L. Q., Yu, W., & Webster, R. (2015). In situ measurements of organic carbon in soil profiles using vis–NIR spectroscopy on the Qinghai—Tibet Plateau. *Environmental Science & Technology, 49*, 4980–4987. https://doi.org/10.1021/es504272x

Marchant, B. P., Dailey, A. G., & Lark., R. M. (2012). Cost-effective sampling strategies for soil management. Home-Grown Cereals Authority Research and Development Project Report No, 485. *HGCA*, London. https://cereals.ahdb.org.uk/media/252469/pr485.pdf

Milne, R., Mobbs, D. C., Thomson, A. M., Matthews, R. W., Broadmeadow, M. S. J., Mackie, E., Wilkinson, M., Benham, S., Harris, K., Grace, J., Quegan, S., Coleman, K., Powlson, D. S., Whitmore, A. P., Sozanska-Stanton, M., Smith, P., Levy, P. E., Ostle, N., Murray, T. D., Van Oijen, M., & Brown, T. (2006). UK emissions by sources and removals by sinks due to land use, land use change and forestry activities. Report, April 2006. Centre for Ecology and Hydrology. (CEH: Project Report Number C02275), Wallingford, United Kindom. http://nora.nerc.ac.uk/3370/.

Muhammed, S. E., Marchant, B. P., Webster, R., Whitmore, A. P., Dailey, G., & Milne, A. E. (2017). Assessing sampling design for determining fertilizer practice from yield data. *Computers and Electronics Agriculture, 135*, 163–174. https://doi.org/10.1016/j.compag.2017.02.002

Olsen, S., Cole, C., Watanabe, F., & Dean., L. (1954). Estimation of available phosphorus in soil by extraction with sodium bicarbonate. USDA Circular Nr 93, US Government Printing Office, Washington, DC.

Prasad, M., Spiers, T. M., & Ravenwood, I. C. (1988). Target phosphorus soil test values for vegetables. *New Zealand Journal of Experimental Agriculture, 16*, 83–90. https://doi.org/10.1080/03015521.1988.10425619

Pretty, J. N., Mason, C. F., Nedwell, D. B., Hine, R. E., Leaf, S., & Dils, R. (2003). Environmental costs of freshwater eutrophication in England and Wales. *Environmental Science & Technology, 37*, 201–208. https://doi.org/10.1021/es020793k

Pringle, M. J., Cook, S. E., & McBratney, A. B. (2004). Field-scale experiments for site-specific crop management Part I: Design considerations. *Precision Agriculture, 6*, 617–624. https://doi.org/10.1007/s11119-004-6346-1

Pringle, M. J., McBratney, A. B., & Cook, S. E. (2004). Field-scale experiments for site-specific crop management. Part II: A geostatistical analysis. *Precision Agriculture, 6*, 625–645. https://doi.org/10.1007/s11119-004-6347-0

Ramirez-Lopez, L., Wadoux, A.M.J.-C., Franceschini, M. H. D., Terra, F. S., Marques, K. P. P., Sayão, V. M., & Demattê, J. A. M. (2019). Robust soil mapping at the farm scale with vis-NIR spectroscopy. *European Journal of Soil Science, 70*, 378–393. https://doi.org/10.1111/ejss.12752

Ribeiro Jr, P. J., & Diggle, P. J. (2018). *geoR: Analysis of geostatistical data.* R package version 1.7-5.2.1. Retrieved February 5, 2020, from https://CRAN.R-project.org/package=geoR

Sandford, S. A., & Allamandola, L. J. (1990). The physical and infrared spectral properties of CO_2 in astrophysical ice analogs. *The Astrophysical Journal, 355*, 357–372. https://doi.org/10.1086/168770

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*, 1627–1639. https://doi.org/10.1021/ac60214a047

Schoumans, O. F., Chardon, W. J., Bechmann, M. E., Gascuel-Odoux, C., Hofman, G., Kronvang, B., Rubæk, G. H., Ulén, B., & Dorioz, J.-M. (2014). Mitigation options to reduce phosphorus losses from the agricultural sector and improve surface water quality: A review. *Science of the Total Environment, 468–469*, 1255–1266. https://doi.org/10.1016/j.scitotenv.2013.08.061

Somarathna, P. D. S. N., Minasny, B., Malone, B. P., Stockmann, U., & McBratney, A. (2018). Accounting for the measurement error of spectroscopically inferred soil carbon data for improved precision of spatial predictions. *Science of the Total Environment, 631–632*, 377–389. https://doi.org/10.1016/j.scitotenv.2018.02.302

Stevens, A., & Ramirez-Lopez, L. (2013). An introduction to the prospectr package. R package version 0.1.3. Retrieved February 5, 2020, from https://CRAN.R-project.org/package=prospectr

Viscarra Rossel, R. A., & Bouma, J. (2016). Soil sensing: A new paradigm for agriculture. *Agricultural Systems, 148*, 71–74. https://doi.org/10.1016/j.agsy.2016.07.001

Viscarra Rossel, R. A., Brus, D. J., Lobsey, C., Shi, Z., & McLachlan, G. (2016). Baseline estimates of soil organic carbon by proximal sensing: Comparing design-based, model-assisted and model-based inference. *Geoderma, 265*, 152–163. https://doi.org/10.1016/j.geoderma.2015.11.016

Viscarra Rossel, R. A., & Webster, R. (2012). Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. *European Journal of Soil Science, 63*, 848–860. https://doi.org/10.1111/j.1365-2389.2012.01495.x

VSN International. (2021). Genstat for Windows 21st Edition. VSN International, Hemel Hempstead, UK. Web page: Genstat.co.uk.

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 & Geosciences, 36*, 1261–1267. https://doi.org/10.1016/j.cageo.2010.04.005

Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists* (2nd ed.). Chichester: Wiley.

Weiss, C. E., & Roetzer, G. R. (2016). *GeomComb: (Geometric) forecast combination methods. R package version 1.0.* Retrieved March 3, 2020, from https://CRAN.R-project.org/package=GeomComb

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. Retrieved February 14, 2020, from https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., McGowan, L. D. A., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Milton Bache, S., Muller, K., Ooms, J., Robinson, D., Paige Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. and Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*, 1686. https://doi.org/10.21105/joss.01686

Yates, F. (1981). *Sampling methods for censuses and surveys* (4th ed.). London: Griffin.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



https://dspace.org/

A loss function to evaluate agricultural decision-making under uncertainty: a case study of soil spectroscopy

Breure, Timo S.; Haefele, Stephan M.; Hannam, Jacqueline A.; Corstanje, Ronald; Webster, R. ...

2022-03-12
Springer
https://doi.org/10.1007/s11119-022-09887-2
Breure TS, Haefele SM, Hannam JA, et al., (2022) A loss function to evaluate agricultural decision-making under uncertainty: a case study of soil spectroscopy, Precision Agriculture, Volume 23, Issue 4, August 2022, pp. 1333–1353

Downloaded from a DSpace repository. https://dspace.org/