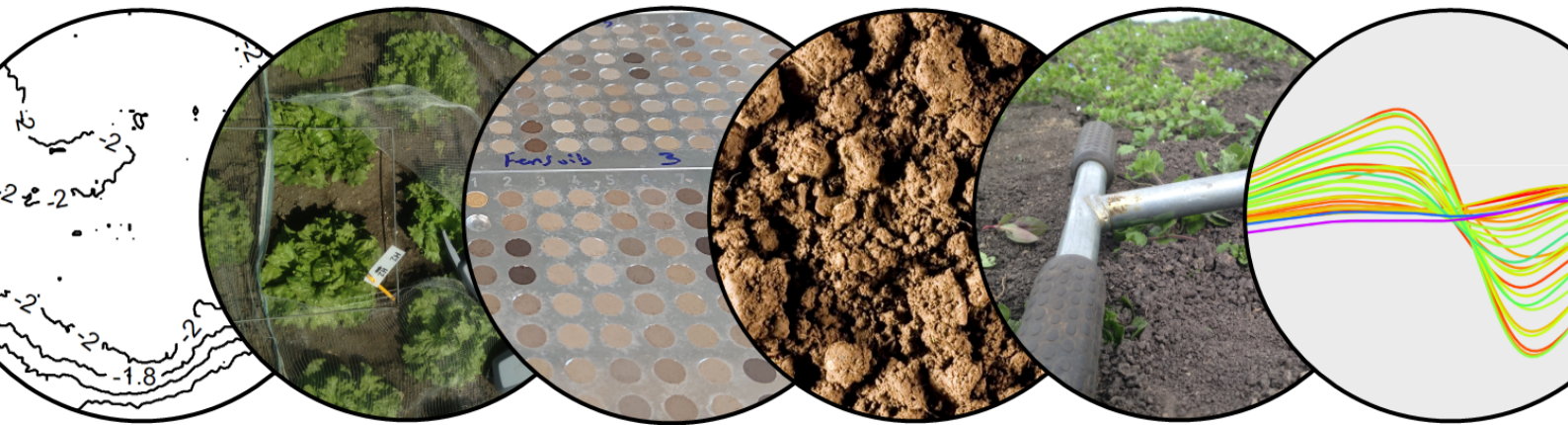


# Quantifying the effect of uncertainty in an applied soil spectroscopy context

A loss function approach

**Timo Samuel Breure**



A thesis presented for the degree of  
Doctor of Philosophy

Cranfield University  
Rothamsted Research  
United Kingdom  
25-06-2021

# Abstract

Advances in sensor technology and agricultural equipment should enable farmers to improve the precision management of nutrients, water and adjusted crop density, but the decision of when to use what sensor and how many measurements to take is still ad hoc. Hence, a systematic approach to sensor use for the determination of soil conditions is necessary and could potentially reduce input use. Consequently, the aim of this thesis is to investigate the uncertainty associated with soil property predictions by spectroscopy in relation to its cost-effectiveness for soil management. To achieve this aim, four case study fields were considered in the Cambridgeshire fens (UK). A total of 747 locations were sampled for top-soil across the fields and spectral measurements were made in the visible- (V), near- (N) and mid- (M) infrared (IR) and X-ray fluorescence (XRF) regions. A subset of the soil samples has been analysed for available P, exchangeable K, pH, organic C, total N and their particle size fractions. The data collected from these fields was used to address four main topics relevant to the application of spectroscopy in soil science.

**Chapter 2** is concerned with the prediction of crop growth from soil NIR and MIR spectra. Crop data derived from air-borne imagery was predicted by both i.) a direct approach that used the soil spectra themselves and ii.) an indirect approach that used soil properties estimated by calibrating the spectra to reference measurements. Results show that estimated soil N, P, K and pH were significant predictors of the crop data within the indirect approach, indicating potential for the use of soil spectral data to inform precision management. Although the direct approach is advantageous for accuracy, it does not provide information on how soil properties can be managed to affect crop performance. The study concludes that there is potential for associating crop response with soil reflectance spectra for improved input management.

**Chapter 3** asks to what extent the effort associated with spectral measurements can be reduced at the cost of prediction accuracy associated with soil property estimates. For this purpose, the magnitude of loss in accuracy was contrasted, relative to field-scale predictions based on milled samples, by either reduced sample processing or the use of existing spectral libraries. Additionally, the predictions were performed for multiple sensors to assess whether their combined effect could minimise the loss in accuracy resulting from reduced sample processing. The study shows that reduced sample processing and spectral libraries have potential to reduce time and cost implications for predicting soil organic carbon, clay and pH from NIR and MIR spectra. Available P and K can only be predicted with moderate accuracy from the milled field-scale samples. Combined predictions from multiple sensors generally led to equal prediction accuracy or a small improvement compared to separate NIR or MIR predictions. The loss of accuracy is specific to the combination of soil property and sensor analysed. The results provide insight



into the expected differences in prediction accuracy and which factors need to be taken into consideration to reduce effort for developing field-scale calibrations.

**Chapter 4** is concerned with quantifying the effect of accounting for uncertainty in soil property predictions from spectroscopy when making decisions about soil management. By accounting for uncertainty, it was tested whether spatial predictions of available P and K were sufficiently accurate to justify the precise application of P and K fertiliser. The effect of uncertainty (compared to using the mean kriging predictions) was quantified as an expected loss under both uniform and precise fertiliser regimes of P and K. Results show that for all four fields, there is an economic incentive for precise fertiliser application of P compared to uniform application. In the case of K, economic advantages were found in two fields. The results also indicate that in general, consideration of uncertainty led to risk-averse fertiliser application. The magnitude of the expected losses and the difference in loss between precise and uniform application were found to be dependent on (a) the kriging variance, (b) the range of the dose-response curve in terms of available P and K, (c) the range of estimated P and K values within the fields and (d) the asymmetry of the loss function. Because reduced application of P fertiliser is not only linked to economic benefits, we conclude that environmental benefits, such as reducing eutrophication of watercourses from reduced P fertiliser applications, should be included in the loss function.

**Chapter 5** is concerned with the uncertainty in soil available P and K estimates from spectroscopy as a function of total- and calibration sample size. The effect of uncertainty on precise fertiliser management was quantified by the difference in profit from applying fertiliser using the estimates of soil nutrient concentration, accounting for uncertainty, relative to the profit that would be gained from applying fertiliser informed by the true variation of available P and K. This difference in profit was denoted as the expected loss. Based on the observed variation in P and K in three experimental fields, 100 realisations per field were simulated for an *in silico* experiment. For each simulation, the fields were sampled and a calibration error was added. After kriging was performed, the fertiliser requirement needed to minimise the expected loss associated with predictions was computed together with the expected profit when data acquisition costs were accounted for. Results show that calibration sample size outweighed the effect of total sample size on the uncertainty associated with predictions. Equally, for the same calibration set size, there were large differences in the kriging variance between total sample sizes. The expected loss showed diminishing returns on investment suggesting that there is an optimum sample size. However, the expected profit in our simulations was dominated by the costs of sampling and spectroscopy. Consequently, no combination of the total- and calibration sample sizes considered would result in a financial gain and could thus be considered optimal. In case costs can be substantially reduced, spectral methods offer a promising method for informing variable rate management. The loss function approach is concluded to be an adequate method to assess whether spectroscopy is effective for informing soil management and should be applied in further case-studies to gain more robust insight in the value of applied soil spectroscopy.

# Contents

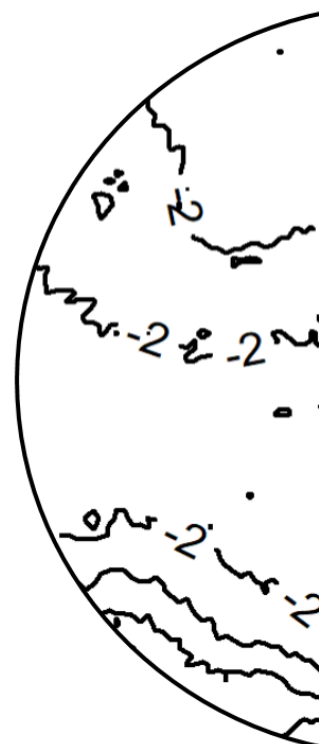
<b>1</b>	<b>Introduction</b>	<b>vi</b>
1.1	Overview . . . . .	1
1.2	Background . . . . .	2
1.2.1	Principles of spectroscopy . . . . .	2
1.2.2	Main research topics in soil spectroscopy . . . . .	3
1.2.3	Contributing sources of uncertainty . . . . .	5
1.2.4	The loss function . . . . .	7
1.3	Aim and objectives of the research . . . . .	8
1.4	Thesis structure . . . . .	10
<b>2</b>	<b>Predicting crop size from soil spectra : the potential for management</b>	<b>11</b>
2.1	Introduction . . . . .	13
2.2	Methods and materials . . . . .	15
2.2.1	The fields and their sampling . . . . .	15
2.2.2	Laboratory analysis . . . . .	15
2.2.3	Chemical analysis . . . . .	15
2.2.4	Spectroscopy . . . . .	17
2.2.5	Statistical analysis . . . . .	17
2.3	Results . . . . .	21
2.3.1	Summary statistics and qualitative description . . . . .	21
2.3.2	PLSR diagnostics . . . . .	22
2.3.3	Relations between lettuce diameter and predicted soil properties	22
2.3.4	PLSR loadings and their interpretation . . . . .	26
2.4	Discussion . . . . .	28
2.4.1	Lettuce diameter prediction directly from soil spectra . . . . .	28
2.4.2	Precision of soil property predictions from soil spectra . . . . .	28
2.4.3	Can values of soil properties predicted from the spectral data be used to predict lettuce diameter? . . . . .	29
2.4.4	Can soil spectra help growers with management? . . . . .	31
2.5	Conclusions . . . . .	32
<b>3</b>	<b>Quantifying uncertainty from reducing field-scale calibration effort</b>	<b>34</b>
3.1	Introduction . . . . .	36
3.2	Methods . . . . .	37
3.2.1	Formation of spectral libraries . . . . .	39
3.2.2	Sample air-drying, sieving and milling . . . . .	41
3.2.3	Wet chemistry analysis . . . . .	41

3.2.4	Spectroscopy . . . . .	41
3.2.5	Data-subsetting and selection of spiking subset . . . . .	43
3.2.6	Partial least squares regression and model validation . . . . .	43
3.2.7	Model-averaging of PLSR predictions and their evaluation . . . . .	44
3.3	Results & discussion . . . . .	44
3.3.1	The selection of representative samples for calibration . . . . .	44
3.3.2	The effect of sample processing on soil property prediction accuracy using (V)NIR, MIR and (V)NIRMIR . . . . .	46
3.3.3	The use of spectral libraries compared with field-scale calibration models . . . . .	52
3.3.4	Contrasting time and cost implications of spectroscopy predictions from spectral libraries and samples with reduced processing . . . . .	54
3.4	Conclusions . . . . .	55
<b>4</b>	<b>A loss function to evaluate agricultural decision-making under uncertainty</b>	<b>56</b>
4.1	Introduction . . . . .	58
4.2	Methods . . . . .	60
4.2.1	Data . . . . .	60
4.2.2	Spectral processing and calibration . . . . .	61
4.2.3	Geostatistical predictions . . . . .	62
4.2.4	Theory of the loss function . . . . .	63
4.2.5	Parameterization and analysis of the loss function . . . . .	64
4.2.6	Software . . . . .	65
4.3	Results . . . . .	65
4.3.1	Uncertainty in kriging predictions from soil properties estimated by spectroscopy . . . . .	65
4.3.2	Loss function on variable-rate fertiliser application . . . . .	65
4.4	Discussion . . . . .	69
4.4.1	Error approximation and the estimation of the variogram . . . . .	69
4.4.2	Data requirements and estimation of the loss function . . . . .	70
4.4.3	The loss function to estimate the value of variable-rate application . . . . .	71
4.4.4	Implications of the loss function approach on decision-making . . . . .	73
4.5	Conclusions . . . . .	74
<b>5</b>	<b>Investigating the effect of uncertainty under different sample sizes</b>	<b>76</b>
5.1	Introduction . . . . .	78
5.2	Methods . . . . .	79
5.2.1	Case study area, sampling, wet chemistry analysis and spectroscopy . . . . .	79
5.2.2	Simulated variation in available P and K . . . . .	80
5.2.3	Estimating the error from the calibration regression . . . . .	81
5.2.4	Modelling procedure to compute kriging predictions for different sample sizes . . . . .	83
5.2.5	The loss function and variable costs of the data acquisition . . . . .	84
5.2.6	Overview of the modelling procedure . . . . .	86
5.2.7	Software . . . . .	86

5.3	Results . . . . .	87
5.3.1	Error variance as a function of the number of calibration samples	87
5.3.2	Prediction uncertainty under different sample sizes . . . . .	87
5.3.3	Expected loss without accounting for data acquisition costs . .	89
5.3.4	Expected profit when data acquisition costs are taken into account . . . . .	90
5.4	Discussion . . . . .	91
5.4.1	Uncertainty in soil properties predicted from spectroscopy at the field-scale . . . . .	91
5.4.2	The expected loss from informing fertiliser management based on spectroscopy estimates . . . . .	94
5.4.3	How cost-effective is spectroscopy at the field-scale? . . . . .	94
5.5	Conclusions . . . . .	95
<b>6</b>	<b>Synthesis</b>	<b>96</b>
6.1	Introduction . . . . .	97
6.2	Overview of findings . . . . .	97
6.3	Future research . . . . .	99
6.3.1	Different methods to account for uncertainty . . . . .	99
6.3.2	Reflections on soil spectroscopy . . . . .	100
6.3.3	Further development of the loss function . . . . .	101
6.3.4	Generic reflection on this thesis research . . . . .	103
6.4	Conclusions . . . . .	104

# Chapter 1

# Introduction



## 1.1 Overview

Soil formation, and its subsequent variation in geographic space, is the result of multiple deterministic processes that take place over various time-scales (Webster, 2000). The collection of data on soil variation is undertaken to make an inventory or monitor changes over time. There are many reasons for which practitioners wish to measure properties of the soil. Agronomists might want to inform management decisions by using a map of soil variation in major nutrients relevant to crop growth. Policy makers require information about the soil's heavy metal content in order to decide whether land remediation is necessary. Scientists monitor CO<sub>2</sub> emissions from soils under different treatments which will inform policy-making on sustainable soil management practices. Decision-making based on data that describes soil-variation accurately allows for informed interventions that are aligned with the expectation of the respective management outcome.

Measurements of soil, however, have an associated uncertainty. Within the decision-making context, social, economic and environmental costs can be incurred as a consequence of not properly accounting for uncertainty in soil measurements. For example, over- or under application of fertiliser will lead to economic losses due to reduced yield or eutrophication of watercourses. The soil's toxicity might have adverse effects on human health when the decision is made to not remediate land when, in reality contamination thresholds are breached. Failure to accurately describe the soil's greenhouse gas emissions under different treatments might lead to inappropriate interventions to mitigate its role in global warming.

Generally, uncertainty of data on soil variation can be reduced by making more measurements. In the case of estimating a mean value over a given area, the uncertainty can be reduced by increasing the number of measurements taken or by the implementation of more efficient sampling methods such as stratified- or balanced sampling. In the case of mapping, the uncertainty associated with the prediction of a soil property can be reduced by increasing the density of measurements taken across the area of interest or by the inclusion of appropriate covariates. A consideration to take into account, however, is that a large number of samples have higher associated costs which may make obtaining precise enough information impracticable. Increased accuracy of soil data is, thus, naturally associated with more effort and higher costs. Soil spectroscopy holds a promising solution to the challenge of generating large sample sizes since it is relatively inexpensive, non-destructive, does not require hazardous chemicals and has capacity to measure multiple soil properties from a single measurement (Viscarra Rossel et al., 2006). Soil spectroscopy can thus serve as a means to provide higher spatial and temporal resolution soil datasets at lower costs compared to traditional laboratory methods. Given the potential of spectroscopy, the number of studies on spectroscopy in the soil science and related peer-reviewed literature has increased exponentially since 1990 (Guerrero et al., 2010).

In this thesis I investigate the uncertainty related to predictions made from soil spectroscopy and develop methods to account for this uncertainty within the context of decision-making for soil management.



## 1.2 Background

This section outlines the main principles of soil spectroscopy followed by the remaining challenges in its application, which constitute the primary motivation for the research presented in this thesis.

### 1.2.1 Principles of spectroscopy

The main principle of soil spectroscopy is that electromagnetic radiation will induce an interaction with the analyte (soil aliquot) to be monitored (such as absorption, fluorescence, emission or diffraction) which is characterized by a spectrum.

The radiation-analyte interaction is specific for the wavelength region that is projected by the spectrometer. For near-infrared (NIR) and mid-infrared (MIR) radiation, absorbance is measured which represents overtones and combinations of fundamental frequencies of organic and inorganic molecules. That is, the atoms within a molecule absorb quanta of energy from the photon by changes in the energy state of their electrons; inducing a periodic displacement of the atoms relative to one another and consequently stretching or bending of that molecule. Most of the fundamental frequencies, the energy at which the first excited stage of a molecule and its periodic displacement take place compared to its ground state, occur in the MIR. Overtones and combinations of fundamental vibrations, particularly those of hydrogen atoms, occur in the NIR (Spragg, 2000; Stenberg & Viscarra Rossel, 2010). Absorption in the visible (V) radiation region induces electronic excitations within atoms and molecules, due to the unsaturation and/or the presence of non-bonded electrons in the absorbing molecules. The principle across V-, NIR- or MIR regions is that a molecular bond will only absorb radiation with an energy quantum that corresponds to the two energy levels of a specific bond. Since the energy quantum is directly related to the wavelength or frequency, the spectra can be used for analytical purposes (Stenberg & Viscarra Rossel, 2010).

The main signal to be measured in x-ray fluorescence (XRF) spectroscopy is the emitted radiation from the analyte (or rather, the atom) as it changes from an excited electronic energy level to its corresponding ground state. Primary x-rays consist of energy levels that exceed the binding energy of the bound inner electrons of an atom, leading to the ejection of an electron from the inner orbitals (typically the K, and L lines, where K is closest to the nucleus). Consequently, the atom is unstable and will regain its ground state through the transfer of an electron from a high-energy outer orbital to the vacancy of the inner electron shell. The process of outer- to inner orbital electron transfer emits energy known as secondary X-ray photons, i.e. fluorescence (Kramar, 2000). The energy of element-specific secondary X-rays produced by a sample is measured over a set time, giving the energy intensity that results in the final spectrum.

This thesis is concerned with X-ray fluorescence (0.01–10 nanometer (nm) range), visible-light (350–700 nm), near-infrared (700–2500 nm) and mid-infrared (2500–25000 nm) spectroscopy. Given the principles of spectroscopy, spectra are in essence a proxy of the soil variable of interest such as, for example, organic carbon. Thus, to quantitatively predict organic carbon content, spectra have to be calibrated by relating them to a known (sub)set of reference samples through an empirical relation (i.e. regression). The (sub)set of known reference samples is

required to be 'representative' of the target set for which organic carbon will be predicted. A representative sample can take on several meanings; e.g. coverage of geographical/pedological/spectral variation. An optimal sample for calibration purposes is 'representative' such that it covers the spread of the target variable. In a sense, it might be biased when compared to the target population mean for which the calibration model is used to predict. Within this thesis, the term 'representative' is used in the context of calibration. For further discussion on 'representative' samples, please see Kruskal and Mosteller (1979). The procedure of using a subset to predict soil physical (e.g. particle size distribution), chemical (e.g. major nutrients) and biological (e.g. microbial C) properties has been reported extensively within the scientific literature (Sinfield, 2010; Kuang et al., 2012;). Work has also been done on the use of soil reflectance spectra to determine crop characteristics directly. These include predictions of grain yield in rice (Van Groenigen et al., 2003) and plant N uptake (Börjesson et al., 1999; Stenberg et al., 2005; Terhoeven-Urselmans, 2008; Wetterlind et al., 2008). However, there are no studies that seek to explain how the variance in crop-yield metrics from soil properties, estimated by reflectance spectra, can inform soil management.

### 1.2.2 Main research topics in soil spectroscopy

The capacity of soil spectroscopy to predict from a representative subset, at a lower cost, with simpler and less laborious methods compared to traditional wet laboratory analysis has the potential to facilitate the generation of larger datasets. Three main topics have been the subject of recent research on the implementation of sensor technology in soil science.

The first major topic considers the effect of combining various sensor technologies for the prediction of soil properties. A wide array of sensors have been studied. These include electrical conductivity ( $EC_a$ ), electromagnetic induction (EMI), ground-penetrating radar, electro-chemical sensors and gamma-ray spectroscopy. The general expectation is that combining different sensors will create more robust predictions. Also, the confidence in predictions is increased since multiple independent measurements are made on the same soil aliquot. Furthermore, each sensor has a unique measurement support leading to a higher coverage of the soil domain in question (Adamchuck & Viscarra Rossel, 2010). However, studies that investigate the combinations of (V)NIR and MIR have been inconclusive reporting both limited benefits (Clairotte et al. 2016; Viscarra Rossel et al., 2006) and general improvements (Viscarra Rossel et al., 2006; Knox et al., 2015; Johnson et al., 2019; Ng et al., 2019). Other studies examined the combination of either (V)NIR or MIR with XRF spectroscopy (Towett et al., 2015; O'Rourke et al., 2016a; O'Rourke et al., 2016b; Zhang and Hartemink, 2019; Benedet et al., 2020). In most cases, the benefits of multiple sensors have been found to differ between soil properties. However, few studies have investigated all three sensors combined (O'Rourke et al. 2016b) and the majority focused on large-scale databases rather than field-scale data sets.

A second major topic of research is on the development of spectral libraries. Spectral libraries are datasets that contain paired analytical and spectroscopy measurements of soil properties that can be used to establish a calibration. The main principle is to build libraries that are representative of the soil variation

in the area of interest, potentially allowing for the prediction of soil properties without performing a site-specific calibration. Notable efforts have been undertaken for a wide variety of geographical scales, spectroscopy instruments and analytical methodologies in the development of regional, continental and even global spectral libraries (Shepherd and Walsh, 2002; Brown et al., 2007; Viscarra Rossel et al., 2008; Viscarra Rossel & Webster, 2012; Stevens et al., 2013; Viscarra Rossel et al., 2016). Over the years, however, limitations to the generalisability of spectral libraries have been pointed out, such as differences between analytical instrumentation and laboratory conditions or under-representation of soil-variation (Viscarra Rossel et al., 2008; Wetterlind and Stenberg, 2010; Ge et al., 2011; Guerrero et al., 2014; Guerrero et al., 2016; Grunwald et al., 2018). While recognising the limitations of using spectral libraries, research has been directed at methodologies to address these issues. One example, first proposed by Brown (2007) is to “spike” a spectral library. Spiking consists of adding a small subset of samples representative of the target site to an existing spectral library (see also: Viscarra Rossel et al., 2009; Guerrero et al., 2010). Another methodology is to select observations from a spectral library that are similar to the target site. Examples range from discretization by geographic extent (Sudduth and Hummel, 1996; Seidel et al., 2019) or strata (e.g. soil class) (Shepherd & Walsh, 2002; Seidel et al., 2019; Moura-Bueno et al., 2020) to metric approaches that assess spectral characteristics and select a representative subset based on similarities between spectra (Shenk et al., 2000; Ramirez-Lopez et al., 2013; Lobsey et al., 2017).

The third major subject covers the portability of various sensor technologies. One reason for which portability proves interesting is that it provides a means to maximise the spatial density at which measurements are taken, consequently reducing the spatial uncertainty of soil measurements (Webster & Oliver, 2007). A second reason for which portability proves interesting is the reduction in effort spent on sampling, handling and laboratory analysis (Reeves, 2010). For example, measurements within the field (in-situ) avoid the process of having to air-dry, sieve and grind the aliquot to get the most accurate reading. Consequently there have been studies that reported on mobile sensor platforms (Christy et al., 2004; Taylor et al., 2006; Adamchuck & Christenson, 2007; Christy, 2008; Adamchuck, 2011) or the use of handheld-held spectrometers within the field (Stevens et al., 2008; Gras et al., 2014; Webster et al., 2016; Hutengs et al., 2019). An important consideration with in-situ measurements is that the accuracy in the soil property predictions can be reduced due to effects of particle-size, aggregation and water content on spectroscopy measurements. A number of studies have researched these effects, for example, studies analysed the effect of different particle sizes (Nduwamungu et al., 2009; Le Guillou et al., 2015; Coutinho et al., 2019; Wijewardane et al., 2020) and soil water content (Bogrekci and Lee, 2006; Minasny et al., 2011; Ji et al., 2015). These studies reported that the effect of particle-size, aggregation and water content with in-situ spectral measurements can be substantial compared to conventional laboratory measurements.

Although the successful application of spectroscopy for soil property predictions is supported by peer-reviewed evidence, generalisability and cost-efficiency related issues hamper its widespread implementation. Soil spectroscopy at the field-scale adds large costs to the overall procedure due to the required high sample density and costs for wet chemistry analysis to establish a robust calibration. The three

areas of research listed here can potentially enable more accurate predictions of soil variables while reducing the overall effort of applied soil spectroscopy. However, the uncertainty in soil property estimates from the use of multiple sensors, spectral libraries and samples under reduced processing at the field-scale has been relatively understudied (Reeves, 2010).

### 1.2.3 Contributing sources of uncertainty

The previous section outlined major topics within the field of soil spectroscopy and efforts that are undertaken to gain accurate predictions while minimising the effort of data acquisition. This section outlines factors that contribute to the accuracy of predictions and how sources of error are introduced within the collection and use of soil property data from spectroscopy.

#### Prediction errors from soil spectroscopy

The first source of error that occurs in the process of soil spectroscopy is the analytical error of the laboratory methods that are used to establish reference values. Although considered to be small within a single laboratory, the inter-laboratory variation has been shown to be larger than expected (Pleijssier, 1986). The second source of error is related to the spectral measurements. Specifically, the inherent variation of the instrument in predicting the target. These variations are dependent on the instrument used and on the conditions of the soil aliquot. Sources of error include scattering effects, leading to non-linearities in the radiation-matter interaction (e.g. absorption) (Stenberg & Viscarra Rossel, 2010), and the effect of sample heterogeneity relative to the measurement support of the spectrometer. For MIR, the diameter of the sampling beam lies in the range of 2–3 mm, for VNIR this is greater than 5 cm whereas for XRF this is approximately 2 cm (Soriano-Disla et al., 2014). The penetration depth of the spectrometers is typically  $\leq 2$  mm for VNIR, MIR and XRF (Kalnicky and Singhvi, 2001; Bänninger et al., 2006) although other sensor technologies mentioned previously (e.g. EC<sub>a</sub> or GPR) can measure at depths up to 60 m and 2 m respectively (Oliver, 2010; Liu et al., 2016). The surface area and penetration depth of a sensor can depend on instrument set-up and specific soil conditions (e.g. mineralogical-, water-content and particle size variation) (Bowers & Hanks, 1965; Wu et al., 2009a; Wu et al., 2009b). In general, the measurement error is minimised by taking a representative sub-sample which has been sieved to  $< 2$  mm particle size and then ground to  $< 100$   $\mu\text{m}$  hence reducing the intra- and inter-particle variation (Soriano-Disla et al., 2014). Additionally, replicate measurements on a single aliquot are averaged to minimise the measurement error from the spectrometer. The third source of error is associated with the calibration model, which is usually formed by regressing the predictor variables (soil spectra) against the response variables (reference values measured in the laboratory).

#### Spatial prediction errors

Often a practitioner requires a map of the soil property of interest. The fact that we can only take a limited number of soil samples creates a necessity for interpolation, i.e. we must predict values in between measurement locations. A large number

of physical processes contribute to the formation of soil. Given the complexity of these often cross-dependent deterministic processes, soil variation often appears to be random (Webster, 2000). Geostatistics involves model-based predictions of variable,  $z$ , in space at locations  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  by treating the target variable,  $z(\mathbf{x})$ , as outcome of a spatial random process  $Z(\mathbf{x})$ .

Geostatistical prediction is done by kriging, which involves weighted averaging of the observed values of a property, within a neighbourhood (Webster & Oliver, 2007). After Webster & Lark (2012), the kriging prediction at location  $\mathbf{x}_0$  is given by:

$$\hat{Z}(\mathbf{x}_0) = \sum_{i=1}^N \lambda_i z(\mathbf{x}_i) \quad (1.1)$$

where the sum of the weights,  $\sum_{i=1}^N \lambda_i$ , is equal to 1. Kriging minimises the squared errors of predictions at the locations where we have taken measurements, thus leading to a model that is most closely in line with the observed data (Webster & Oliver, 2007). The model that describes the global mean and its covariance is described by:

$$Z(\mathbf{x}) = \mu + \varepsilon(\mathbf{x}) \quad (1.2)$$

where,  $\mu$  represents the mean,  $\varepsilon$  is the random process with covariance  $\text{Cov}[\varepsilon(\mathbf{x}_i) - \varepsilon(\mathbf{x}_j)]$ , as the mean of  $\varepsilon(\mathbf{x})$  is considered equal to 0. Subsequent modelling of the spatial variation within a geostatistical framework then is supported by two main assumptions. One being the assumption of first-order stationarity, i.e.  $\mu$  is considered constant across the target area. Second-order stationarity assumes that the covariance between two locations is solely dependent on the distance that separates these two locations. Thus based on the second-order stationarity assumption we assume:

$$\text{Cov}[\varepsilon(\mathbf{x}_i), \varepsilon(\mathbf{x}_j)] = C(\mathbf{x}_i - \mathbf{x}_j) \quad (1.3)$$

where  $C(\cdot)$  is a function of the lag,  $\mathbf{h}$ , which is the distance in two or three dimensions between two locations. The covariance function diminishes to zero for increasing lag distances (i.e. the covariance between two locations reduces for larger lags).

There are instances in which the assumption of second-order stationarity does not hold and an assumption of intrinsic stationarity is required, i.e. the variance of the difference between two locations ( $\text{var}[\varepsilon(\mathbf{x}) - \varepsilon(\mathbf{x} + \mathbf{h})]$ ) is stationary, which can be described by half its expected value:

$$\text{var}[\varepsilon(\mathbf{x}) - \varepsilon(\mathbf{x} + \mathbf{h})] = \text{E}[\{\varepsilon(\mathbf{x}) - \varepsilon(\mathbf{x} + \mathbf{h})\}^2] = 2\gamma(\mathbf{h}) \quad (1.4)$$

The spatial variance between the residuals  $\varepsilon$  is then generally described by one of several standard parametric functions,  $\rho(\mathbf{h}; \theta)$ . For a stationary Gaussian model, the variogram ( $\gamma(\mathbf{h})$ ), is given by:

$$\gamma(\mathbf{h}) = c_0 + c_1 \{1 - \rho(\mathbf{h}; \theta)\} \quad (1.5)$$

The correlation function is denoted by  $\rho(\mathbf{h}; \theta)$ , of which its parameters ( $\theta$ ) are estimated from observations. The intercept term ( $c_0$ ) is the conditional error

variance. The sill ( $c_1$ ) is the variance of the observation process, which indicates the limit up till which the underlying process is spatially correlated. The lag distance up till which the underlying process is spatially correlated is referred to as the range of the variogram.

The error introduced by model-based geostatistics is known as the kriging variance ( $\sigma_k^2$ ) and can be further divided into i.) local uncertainty (i.e. the marginal error at a particular location) and ii.) spatial uncertainty (i.e. estimation of the joint error at multiple point-locations) (Goovaerts, 2001).

### The effect of sampling on error

The kriging error variance depends on the number of samples and the sample design. Generally it is assumed that for a given number of samples this variance is minimised by implementing relatively even geographical coverage at spacings shorter than the spatially correlated range. When soil sensors are used to estimate a soil property, a second optimisation is required to develop a representative calibration set for the prediction of the target property from the spectra. In this case, a calibration sampling design is required that spans the range of the property (or feature) space (Viscarra Rossel et al., 2011).

Thus, the sampling design is a dual-objective process. That is, samples have to be representative of both geographic- and feature space. According to a review by Viscarra Rossel et al. (2011), the majority of the peer-reviewed literature considers either geographical space sampling or feature space sampling designs while not covering both. Optimum designs for sampling a calibration set have been implemented by response-surface sampling (Lesch, 2005) and conditional Latin hypercube sampling (Minasny & McBratney, 2006). Several studies addressed both geographic and feature space sampling (Minasny et al., 2007; De Gruijter et al., 2010; Adamchuk et al., 2008, 2011). These studies on sampling in the context of soil sensing did not take into account the value associated with reducing uncertainty versus the costs associated with sampling which has been considered more recently by de Gruijter et al. (2018) and Shaw et al. (2016). Further research is needed to develop optimal sampling designs for soil spectroscopy that account for the cost of the data collection effort and the value of the information to the research question at hand.

#### 1.2.4 The loss function

As described above, there are a range of error sources affecting the acquisition of soil data and subsequent decision-making. Quantifying the uncertainty of these soil predictions is not a goal in itself, however. Generally, there is a trade-off between the accuracy and cost of sampling. Next to the laboratory analysis of reference samples, the effort of sampling in itself constitutes arguably the largest cost and effort of data collection in soil spectroscopy (de Gruijter et al., 2016). Correct acknowledgement of uncertainty can inform this trade-off by evaluating the accompanying risk of any decision-making based on soil data (Goovaerts, 2001).

Recent studies within the soil spectroscopy literature have acknowledged the need to propagate errors and assess uncertainty in estimates (Brodský et al., 2013; Viscarra Rossel et al., 2016; Somarathna et al., 2018; Ellinger et al., 2019; Ramirez-Lopez et al., 2019). However, only few studies have been published that



assess the effect of uncertainty in the context of decision-making (Fachner et al., 2000; Gebbers & de Bruin, 2010). There is a need to assess cost-reducing measures within spectroscopy balanced with the level of accuracy required for effective decision-making in soil management.

The costs of sampling and data analysis for increased accuracy can be described in a loss function framework. The loss is defined as the costs that are incurred due to making a decision based on soil estimates that are uncertain. The rationale behind this approach is to maximise the value of the information acquired (e.g. a soil map) given the cost of the data collection procedure. Yates (1949) first described this in the context of sampling effort for a monitoring survey. Uncertainty in data that leads to deviation from the true value of  $z$  is expressed as the loss which is to be minimised for a given criterion. The error is defined by  $z - \hat{z}$ , which is expressed by the expected error value that can be used to define an optimum to minimise the loss from uncertainty in our data. For example, in the context of precise application of fertiliser, the loss can be due to under-application of fertiliser and consequent yield loss ( $\hat{z} < z$ ) or over-application and consequent financial loss from unused fertiliser, or alternatively cause loss to the environment by eutrophication ( $\hat{z} > z$ ) (Lark & Knights, 2015). Incurred costs can be different for over- or underestimation and hence the loss function may be asymmetrical. The loss function can be used to assess whether the magnitude of uncertainty is such that management interventions have their desired effects. If the latter is not the case, an additional reconnaissance survey might be required. Although the loss function has been used in the context of sampling for soil variables (Yates, 1949; Ramsey et al., 2002; Boon et al., 2011; Marchant et al., 2013; Lark & Knights, 2015) thus far it has not been applied in the context of soil spectroscopy.

## 1.3 Aim and objectives of the research

Spectroscopy has the potential to provide detailed information on the variation of soil. However, adoption in an applied context is hampered by the failure to properly account for uncertainty and issues of cost-efficiency.

**The overall aim of this thesis** is to explore the practical use of soil spectral measurements to inform soil monitoring and management. The first step is to quantify the uncertainty that is introduced under the various stages involved in predicting soil properties by spectroscopy. This uncertainty can then be explicitly accounted for in a loss function framework that can be used to investigate how the expected loss can be minimised under different configurations of spectroscopy measurements, calibration methods and sampling designs. This aim is addressed through the following research questions:

1. Given the inherent uncertainty associated with spectral predictions, can soil spectra predict crop performance indicators that inform management?

It is widely accepted that soil properties affect crop performance. The first question that arises is to what extent the soil spectra can predict crop performance indicators and if so, to what extent this could be used to inform management. Predictions of crop performance can be done in two distinct ways. The first approach is the direct route, in which the soil spectra by themselves are used to predict crop performance indicators. The second

approach is an indirect route, in which soil spectra are used to predict soil variables which are then in turn used to predict crop performance indicators. Within Chapter 2 I assess these two approaches to use soil spectra for the prediction of crop performance to inform precise agricultural management.

2. To what extent can the effort associated with spectral measurements be reduced and what is the effect of this reduction on the uncertainty of soil predictions?

Field-scale calibration of soil spectral measurements is generally considered to result in the most accurate predictions. This approach however, involves a high expense of laboratory analysis and sample processing. Within Chapter 3, I investigate two approaches to reduce effort of field-scale calibrations and quantify the impacts these approaches have on the uncertainty in predictions.

The first approach is concerned with the reduction of sample processing. I investigate the difference in accuracy between soil property predictions from spectroscopy measurements taken on in-situ, unprocessed (i.e. fresh), air-dried or milled soil samples. The second approach is the use of spectral libraries to reduce the number of reference values needed compared to a bespoke field-scale calibration. Both approaches increase uncertainty associated with predictions compared with a field-scale calibration on milled soil samples. In order to counterbalance the increase in uncertainty I explored whether multiple sensors could minimise the prediction error when we construct calibration models from spectra measured on different sample conditions or the use of spectral libraries.

3. What are the advantages of accounting for uncertainty in soil property predictions when making decisions about soil management?

Using accurate and detailed soil maps has potential to improve field management in terms of precise fertiliser applications. This is known as precision agriculture (PA), the principal aim of which is to maintain or improve current crop production levels while mitigating losses of agrochemicals to the environment (Blackmore et al., 2003; Oliver, 2010). Thus, a fundamental premise to PA is to acknowledge soil variation on a sub-field scale, characterize it and inform decision-making on farm management. Farmers who adopt PA, however, 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 acquired affects both its value and consequently whether sub-field soil variation maps are seen as a reliable information source for decision-making in a PA context. A soil map with high uncertainty imposes risk on farming outcomes and the environment by inadequate management decisions.

Within Chapter 4, I quantify the uncertainty in soil nutrient estimates from spectroscopy. A loss function framework is then developed to make the uncertainty in the estimates explicit. Uniform and precise fertiliser application regimes are contrasted in terms of incurred costs and fertiliser used. Fertiliser usage and incurred costs under uncertainty provide insight in the economic and environmental benefits of field-scale spectroscopy to inform precise soil nutrient management.

4. Can we use the loss function framework to inform the data-acquisition process and sampling configurations for soil property prediction by spectroscopy?

Within Chapter 5, I explore how the data acquisition procedure in the context of a soil spectroscopy survey affects uncertainty in predictions.

Associated uncertainties of predicted soil properties can be reduced by increasing the density of measurements taken across the area of interest. Equally the number of samples used as a subset for calibration will determine the accuracy with which the overall sample population is predicted. Increase in both the density and calibration samples has an associated cost which may make obtaining precise enough information impracticable.

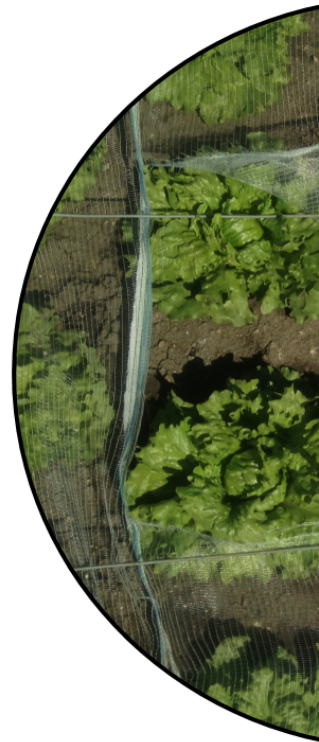
I formulate a loss function based on the costs of sampling, calibration and the incurred costs from uncertainty in our soil estimates. The loss function is in turn used to identify the effect of sampling configuration on prediction uncertainty and its consequences for soil management.

## 1.4 Thesis structure

The thesis is organised in six chapters, including this introduction chapter. The main chapters of this thesis include published (Chapters 2), submitted (Chapters 3 and 4) and unpublished (Chapter 5) manuscripts. Chapters 2–5 follow the outline of these manuscripts and hence can be read by themselves. Appendices differ from the originals that accompanied the publications to accommodate additional work done over the course of my PhD. The bibliographies of all chapters have been combined into a single list of references at the end of this thesis.

## Chapter 2

# Predicting crop size from soil spectra : the potential for management



## Abstract

*How well could the growth of a leafy crop be predicted from soil reflectance spectra and how might a grower manage the crop in the light of those predictions? Topsoil from two fields was sampled and analysed for various nutrients, particle-size distribution and organic carbon concentration. Crop measurements (lettuce diameter) were derived from aerial-imagery. Reflectance spectra were obtained in the laboratory from the soil in the near- and mid-infrared ranges, and these were used to predict crop performance by partial least squares regression (PLSR). Individual soil properties were also predicted from the spectra by PLSR. These estimated soil properties were used to predict lettuce diameter with a linear model (LM) and a linear mixed model (LMM): considering differences between lettuce varieties and the spatial correlation between data points. The PLSR predictions of the soil properties and lettuce diameter were close to observed values. Prediction of lettuce diameter from the estimated soil properties with the LMs gave somewhat poorer results than PLSR that used the soil spectra as predictor variables. Predictions from LMMs were more precise than those from the PLSR using soil spectra. All model predictions improved when the effects of variety were considered. Predictions from the reflectance spectra, via the estimation of soil properties, can enable growers to decide what treatments to apply to grow lettuce and how to vary their treatments within their fields to maximize the net profit from the crop.*

Based on:

Breure, T., 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. *Prec. Agr.* 22, 226–248. <https://doi.org/10.1007/s11119-020-09739-x>.

## 2.1 Introduction

Leafy crops such as lettuce and brassicas are important commercial crops in the UK. The value of these crops depends on quality indicators such as size and weight. Growers size their crops either by direct observation in the field or from air-borne imagery, which has become an established practice to connect crop phenotype with marketability and crop management decisions within the growing season (Bauer et al., 2019; Valente et al., 2020). Lettuces are sampled at frequent time-intervals for fresh-weight, head-weight and head-diameter which determine their market value and time of harvest. Lettuces that do not reach a desired size are not harvested; indeed, frequently large parts of fields are deemed not worth harvesting.

The growth of the lettuce can be restricted by stresses such as shortage of nutrients and water, low temperature, adverse weather, and pests and diseases. Within-field variation of lettuce growth in the United Kingdom is often a result of the soil's varied capacity to provide water and nutrients. If growers have dense information on soil variation within their fields, they are likely to be sufficiently well informed to make two main decisions. First, they should be able to recognize a priori where their crops will not reach a saleable quality and so where not to waste time and resources on production. Second, they should be equipped to decide how best to vary fertiliser and irrigation spatially (precision application) to maximize growth without applying excess of either. In both cases production would be more profitable and less harmful to the environment. Chemical analysis of soil by conventional wet chemistry is expensive and time-consuming. The densest affordable sampling in commercial conditions is one soil sample per ha (Muhammed et al., 2017). That has generally been adequate to estimate mean values and average fertiliser requirements. It is too coarse, however, for mapping the variation within individual fields in a way that enables growers to vary their applications of fertilisers and water rationally. Recent advances in reflectance spectroscopy could enable growers and their advisors to obtain affordable useful information on soil variation at resolutions sufficient for precision agriculture. However, estimated soil properties from reflectance spectra need to be sufficiently accurate to explain variance in crop performance and so inform management decisions.

The utility of near- and mid-infrared reflectance spectra from the soil to predict crop performance and aid management was investigated. One route, which avoids any issues of poor predictions of soil nutrients, is to examine the direct relation between the spectral data and the crop response. A strong relation could tell the grower where to expect good growth and where it is worth planting the lettuce (and where not to plant). It does not tell the grower which soil properties might be causing variation, however, and what action he or she should take to enhance yield. The second route is to predict soil chemical properties from the spectral data and identify which of those properties explain the variance in growth of the lettuce. This route, though less direct and perhaps less accurate because of the issues related to predicting soil nutrients, has greater potential for management; it should enable the grower to vary the management in accordance with the variation of individual soil properties that affect growth. To the authors' knowledge, there are no studies that seek to explain variance in crop-yield metrics from soil properties estimated by reflectance spectra, however, work has been done using soil reflectance spectra directly to determine crop characteristics. These include predictions of grain yield



in rice (Van Groenigen et al., 2003) and plant N uptake (Börjesson et al., 1999; Stenberg et al., 2005; Terhoeven-Urselmans, 2008; Wetterlind et al., 2008).

This study investigated both methods of predicting crop performance from near- and mid-infrared soil reflectance spectra, i.e. directly from the soil spectra and indirectly using soil properties estimated by reflectance spectra (Fig. 2.1).

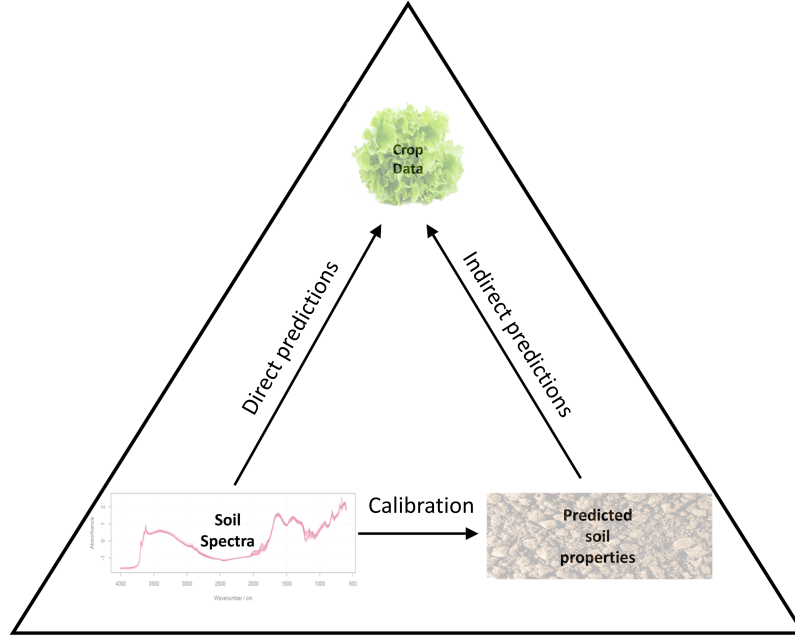


Figure 2.1: Framework for predicting crop data (lettuce diameter in cm) from soil spectra. The diameter of the lettuce may be predicted (i) directly from soil spectral measurements (the left have edge of the triangle) using partial least squares regression (PLSR) or (ii) by first predicting relevant soil properties from the soil spectra and using these in a linear model (LM) or linear mixed model (LMM) to predict lettuce diameter.

The following questions were addressed;

1. Can the diameters of lettuce be predicted directly from the soil-spectral data?  
And if so, how well?
2. How accurately can important soil properties be predicted from soil-spectral data?
3. Can values of soil properties predicted from the spectral data be used to model and predict lettuce diameter?

The results are used to discuss whether soil spectral measurements could be used in practice to help the grower (i) decide where to grow lettuce and/or (ii) manage the nutrition, irrigation or planting density of the crop.

## 2.2 Methods and materials

### 2.2.1 The fields and their sampling

The case-study is located in the Fenlands of eastern England. The soil there is generally fertile and is well suited for their growth, though within individual fields there is substantial soil variation, mainly in particle-size distribution and organic matter content. The soil on the flat land is rich in organic matter and varies from silty clay to sand on somewhat (up to 0.5 m) higher sinuous narrow strips known locally as ‘rodhams’. It is known that the crop differs in its response on the flat land from that on the rodhams to uniform management.

After consulting growers in the Fenland region, two fields were chosen for this study. Field 1, covering 10.5 ha, is near the village of Prickwillow (52°27′58.65″N; 000°21′51.02″E). The sampling design was based around a 30-m square grid, with three transects (on alternate rows) more intensely sampled at 6-m intervals. The sampling strategy was designed to provide good coverage of soil conditions and to enable us to assess the magnitude and spatial scale(s) of variation in soil properties. Field 2 lies adjacent to Field 1 and covers 18.2 ha. The design was computed for 121 sample locations such that each point lay in the centre of its Dirichlet tile all of which have the same area. This was done using the **spcosa** package (for more detail, see Walvoort et al., 2010) which led to an approximate grid with an interval of 30 m. A further 36 of these points were selected with the **BalancedSampling** package (Grafström & Lisic, 2019), balanced on the spatial co-ordinates and elevation. At each location of the spatially balanced subset another sample point 6 m away at a random orientation was added. The random orientation was computed with the **SpatialEco** package in R (Evans, 2019). In both fields, additional sample locations were added to ensure that the full range of soil conditions and elevation were encompassed based on predictions from the LiDAR survey and satellite imagery showing variation in soil colour. In all, 256 samples were taken from Field 1 and 161 samples from Field 2. Fig. 2.2 shows the fields with the sample locations.

At each sample location, three cores of topsoil (0–0.25 m) were taken using an auger within a 0.5 m by 0.5 m quadrat. These cores were bulked for laboratory analysis and spectral study.

### 2.2.2 Laboratory analysis

#### Sample preparation

The soil samples were dried in air, passed through a 2-mm sieve and milled. The analyses on these samples are described in detail below. For spectroscopy measurements, samples were placed in a stainless-steel cup together with a disc. The samples were then milled for 35 s at 960 rpm in a TEMA Machinery Ltd mill (Northants, UK).

### 2.2.3 Chemical analysis

Thirty sub-samples for each field were selected by balanced sampling on the spatial co-ordinates and elevation from the sieved samples for chemical analysis and

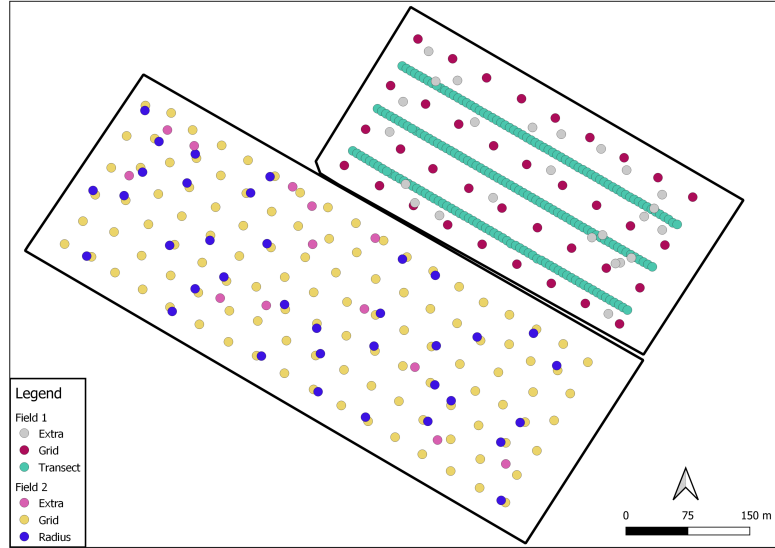


Figure 2.2: The two fields sampled. Field 1 includes the three transects sampled at 6-m intervals (Transect). The 30-m square grid (Grid) and additional sample points capture variation in soil colour and elevation from satellite imagery and LiDAR data (Extra). Field 2 includes the central points of Dirichlet tiles that all cover an equal area, leading to an approximate grid (Grid). The sub-sample of points that contains a paired sample within a 6-m radius (Radius) and additional sample points capture variation in soil colour and elevation from satellite imagery and LiDAR data (Extra)

particle-size distribution. Chemical properties and particle-size distribution of the soil samples were measured as follows.

Total carbon (C) and nitrogen (N) were determined by Dumas combustion in a TruMac Combustion Analyser from LECO Corporation (Stockport, UK).

Exchangeable potassium ( $K^+$ ), calcium ( $Ca^{2+}$ ), magnesium ( $Mg^{2+}$ ) and sodium ( $Na^+$ ) were determined in an ammonium nitrate extract (10 g of 2 mm sieved soil in 1 M ammonium nitrate) by an Optima 7300 DV Inductively Coupled Plasma-Optical Emission Spectrometer (ICP-OES) (Seer Green, UK).

Available phosphorus (P) was measured by the standard Olsen method in a sodium bicarbonate extract (5 g of 2 mm sieved soil in 0.5 M  $NaCO_3$  (Olsen et al., 1954) with a SANplus continuous colorimetric flow analysis from Skalar analytical BV (Breda, The Netherlands).

Sulphur (S) was measured in a potassium phosphate extract from 5 g soil in 25 ml solution. From the filtrate 9.5 ml was stabilized with 0.5 ml nitric acid ( $\approx 68\%$ ) and analysed by ICP-OES.

The soil pH was measured in a suspension of 5 g 2-mm sieved soil to 12.5 g deionized water and measured with a thin semi-micro sealed combined pH electrode from Fisher scientific (Loughborough, UK).

Particle-size fractions were determined by laser diffraction on a L-960 particle-size analyser from Horiba scientific Ltd (Northampton, UK) in the AfsIS spectral laboratory at Rothamsted Research. We set the upper limit for clay to 9  $\mu m$

because we did not remove the organic matter. This is as recommended by Konert & Vandenberghe (1997) and Fisher et al., (2017). The intervals were therefore clay:  $<9\ \mu\text{m}$ , silt:  $9\text{--}50\ \mu\text{m}$ , sand:  $>50\ \mu\text{m}$ .

### 2.2.4 Spectroscopy

Each sub-sample of milled soil was pressed into a small well (6 mm across and approximately 1 mm deep) and placed in a Tensor II spectrometer from Bruker scientific (Ettlingen, Germany) in the AfSIS spectral laboratory at Rothamsted Research. Its reflectance spectrum in the range 1000–2500 nm, i.e. the near infrared (NIR), was measured with a resolution of 1 nm and converted to wave-number units ( $\text{cm}^{-1}$ ) by division by  $10^{-7}$ . Subsequently, the moisture bands were removed in two regions: ( $7900\text{--}6849\ \text{cm}^{-1}$ ) and ( $5587\text{--}5102\ \text{cm}^{-1}$ ), respectively. All spectral measurements were replicated three times for each sample.

Each sub-sample's mid infrared (MIR) spectrum in the range  $4000\text{--}600\ \text{cm}^{-1}$  ( $2500\text{--}16\ 666\ \text{nm}$ ) was recorded on the same instrument with a resolution of  $2\ \text{cm}^{-1}$ . The atmospheric  $\text{CO}_2$  bands were removed in the region  $2430\text{--}2240\ \text{cm}^{-1}$ . The reflectance,  $R$ , in both regions have been transformed to optical density (i.e. absorbance,  $A$ ) as  $A = \log_{10}(1/R)$ .

Spectra were smoothed to remove noise using the Savitzky–Golay filter (Savitzky & Golay, 1964) with a third-order polynomial in a moving window of 11. Subsequently, the spectra were standardized (by subtracting the means and dividing the result by its standard deviation) after which the 1st derivatives of the spectra were computed. For the NIR region, a filter length of 31 wavebands was used (i.e. the spacing between points over which the derivative is computed). For the MIR region, a filter length of 11 and a segments size (i.e. the range over which the points are averaged) of 8 wavebands were used. Subsequently, the NIR and MIR spectra were combined into a single matrix used in the subsequent modelling. Processing was done using the **prospectr** package (Stevens & Ramiro-Lopez, 2013).

### 2.2.5 Statistical analysis

#### Partial least squares regression (PLSR)

The first aim was to predict the diameters of the lettuce from the soil spectra—i.e. taking the route along the left-hand arm of the triangle in Fig. 2.1. Initially this was done for each field and each variety of lettuce separately. This was because the lettuce crops were grown and measured at slightly different times, and it was unknown whether the different varieties would respond differently to the soil conditions. The second aim was to predict the properties of the soil as measured by wet chemistry from the reflectance spectra. For this part of the exercise, the data from both fields together were treated as a single set. The combined data provide a wider range across the soil properties than if each field were considered separately.

Both cases are in the general sense a common problem in statistics: we have a set of predictor variables,  $\mathbf{x} \equiv \{x_1, x_2, \dots, x_m\}$ , and we wish to use the set to predict a target variable,  $y$ . The task might at first seem to be one of straightforward multiple regression. However, two features make that solution impracticable: (a) the spectral estimates are strongly correlated with one another, and (b) there are more of them, i.e. more variates, than there are units (lettuces). One feasible solution now popular

in such circumstances is partial least squares regression (PLSR). This method finds a few orthogonal factors that maximize the covariance between the predictors and the target variable, or variables if there are more than one.

Let  $\mathbf{X}$  be an  $n \times m$  matrix of  $n$  units (lettuces or quadrats) and  $m$  variates (spectral estimates) and let  $\mathbf{y}$  be the vector of length  $n$  of measurements (diameters of the lettuces). Then define

$$\begin{aligned} \mathbf{X} &= \mathbf{S}\mathbf{L}^T + \mathbf{E} \\ \text{and} \quad \mathbf{y} &= \mathbf{U}\mathbf{q}^T + \mathbf{f} . \end{aligned} \tag{2.1}$$

In these equations  $\mathbf{S}$  is an  $n \times p$  matrix of factor scores and  $\mathbf{L}$  is the corresponding orthogonal  $m \times p$  matrix of loadings for the predictors in which  $p \ll m$ . In like manner,  $\mathbf{U}$  is an  $n \times p$  matrix of factor scores for the target variable and  $\mathbf{q}$  is the corresponding vector of loadings. The matrix  $\mathbf{E}$  and vector  $\mathbf{f}$  are error terms, which are assumed to be independent and identically distributed. These equations are solved in such a way as to maximize the covariance between  $\mathbf{S}$  and  $\mathbf{U}$ .

In this way, the number of spectral predictors is reduced while maximizing the effectiveness of those retained. The retained components predict directly the diameters of the lettuces. In the indirect route, retained components predict the soil properties as determined by wet chemistry. The method is not quite as straightforward as multiple regression, and a final selection of the number of components retained was determined by leave-one-out cross-validation and calculation of the mean squared error (MSE) of prediction. In general, the MSE initially decreases sharply as a function of the number of components retained and then increases as a result of over-fitting. The number of components for which the MSE was least was kept. In this paper, Lin's concordance correlation coefficient (Lin, 1989) was used to get a measure of the distance from the predicted data relative to the 1:1 line. A value closer to 1 indicates a higher measure of both accuracy and precision relative to observed values.

### Multiple regression

Subsequently, linear mixed models were computed with lettuce diameter as a response and soil properties derived from the spectra by PLSR as predictor variables. Although lettuce size can be directly predicted from soil spectra, as above, those spectra do not tell growers how they might manage the land differentially to achieve some desired size of lettuce. For that, they would like to know the soil's nutrient status, carbon content and particle-size distribution. This case therefore, takes the route along the right-hand side of the triangle (Fig. 2.1) to answer the questions: are derived soil properties and lettuce diameters (after controlling for different lettuce varieties) related? And if so, how strong are those relations, which soil properties are deemed important for predicting the size of lettuce, and how good is the prediction along that route? However, it important to note here that the coefficients of the multiple regression reflect both the covariance between predictor variables as well as their relation to the dependent variable. Thus, an additional experimental field trial would be required to assess which soil properties determine lettuce size. The multiple regression accounted for possible different regression coefficients for the different varieties of lettuce. For each variety

we separately fit a regression model of the form:

$$Y_i = \boldsymbol{\beta}^T \mathbf{x}_i + \varepsilon_i \quad (2.2)$$

Here  $Y_i$  is the diameter of the  $i$ th lettuce, vector  $\boldsymbol{\beta}$  comprises the regression coefficients of the fixed effects  $\mathbf{x}_i$  which comprise a vector where the first entry is one and the following represent the soil properties for observation  $i$  and  $\varepsilon_i$  is an independent residual error. We assume that  $\varepsilon_i$  is a random variable drawn from a normal distribution with zero mean and variance  $\sigma^2$ , thus:

$$\varepsilon_i \sim N(0, \sigma^2) .$$

We set up the regression equations based on this model and solved them with maximum likelihood (ML) estimation. In the event we noticed that the residuals, the  $\varepsilon_i$ , appeared to be spatially correlated with a variogram defined by

$$\gamma(\mathbf{h}) = \frac{1}{2} E [\{\varepsilon(\mathbf{z}) - \varepsilon(\mathbf{z} + \mathbf{h})\}^2] , \quad (2.3)$$

in which  $\gamma(\mathbf{h})$  is the semivariance of  $\varepsilon$  for points separated by the vector  $\mathbf{h}$ , with  $\mathbf{z} \equiv \{z_1, z_2\}$  representing the coordinates in the two spatial dimensions.

Most variograms of crop yields and soil properties have fairly simple forms in which  $\gamma(\mathbf{h})$  increases from some small value at short separating distances to a constant or asymptote as the distance increases. To choose a suitable model we estimated  $\gamma(\mathbf{h})$  by the method of moments, thus:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2m(\mathbf{h})} \sum_{k=1}^{m(\mathbf{h})} \{\varepsilon(\mathbf{z}_k) - \varepsilon(\mathbf{z}_k + \mathbf{h})\}^2 , \quad (2.4)$$

in which  $\varepsilon(\mathbf{z}_k)$  and  $\varepsilon(\mathbf{z}_k + \mathbf{h})$  are the residuals at places  $\mathbf{z}_k$  and  $\mathbf{z}_k + \mathbf{h}$  separated by the vector  $\mathbf{h}$  and for which there are  $m$  paired comparisons. The variation appeared isotropic, and so we treated the vector  $\mathbf{h}$  as scalar  $h$  in distance only. By varying  $h$  an ordered series  $\hat{\gamma}(h)$  were obtained and these were graphed. The graphs that contained the ordered series of  $\hat{\gamma}(h)$  (experimental variograms) were used to assess whether the fitted parameter values were sensible. All could be described by the popular exponential model:

$$\begin{aligned} \gamma(h) &= c_0 + c_1 \left\{ 1 - \exp \left( -\frac{h}{a} \right) \right\} \quad \text{for } h > 0 \\ &= 0 \quad \text{for } h = 0 . \end{aligned} \quad (2.5)$$

In this equation  $c_0$  and  $c_1$  are variances, respectively the nugget and sill of the correlated variance, and  $a$  is the distance parameter. The equation describes second-order stationarity and so has equivalent covariances,  $\text{cov}(h) = c_0 + c_1 - \gamma(h)$ , for incorporation into our prediction model. For convenience model parameters were designated by  $\boldsymbol{\theta} \equiv \{c_0, c_1, a\}$ . To account for the spatial correlation of the residuals we replace  $\varepsilon_i$  by  $\eta_i$ , which is drawn from a variance–covariance matrix  $\boldsymbol{\Xi}$  of error variables, i.e.

$$\eta_i \sim N(0, \boldsymbol{\Xi}) ,$$



and depends on the spatial coordinates of the lettuce  $i$ ; it is now spatially dependent.

The problem now is to estimate both the coefficients of the fixed effects, i.e. the  $\alpha_j$  and  $\beta_j$ , and the parameters of the random terms in  $\boldsymbol{\theta}$ . We do so by the method of residual maximum likelihood (REML) introduced by Patterson & Thompson (1971) and described in the current context by Lark & Cullis (2004). Briefly, the equation we have to solve is

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \boldsymbol{\Xi} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Xi}^{-1} \mathbf{y}, \quad (2.6)$$

in which the matrix  $\mathbf{X}$  contains all the data for the predictors plus a column of 1s and  $\boldsymbol{\beta}$  includes the intercepts of equations (2.2), and vector  $\mathbf{y}$  contains the corresponding measured diameters of the lettuces. The elements of matrix  $\boldsymbol{\Xi}$  are obtained by maximization of the likelihood

$$\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{\theta} | \mathbf{y}) = c - \frac{1}{2} \ln |\boldsymbol{\Xi}| - \frac{1}{2} (\mathbf{y} - \boldsymbol{\beta}^T \mathbf{X}) \boldsymbol{\Xi}^{-1} (\mathbf{y} - \boldsymbol{\beta}^T \mathbf{X}). \quad (2.7)$$

where  $c$  is a constant.

### Selection of variables

Many soil properties can affect the growth of crop plants. Some of them are typically strongly correlated with one another, and to include all in regression equations could lead to spurious results from over-fitting. To avoid such an outcome, soil properties were selected that are most likely to affect the growth of lettuce. These included first the concentrations of the nutrients N, P and K. Total C was omitted as it was directly related to total N. Particle-size distribution is important because it is closely related to the soil's capacity to hold water, so its measurement can serve as a proxy for water-holding capacity. The soil's pH can be important, and so it was added to the list of predictors. Finally, Mg was added because many crops suffer magnesium deficiency in the UK. The final list was as follows.

1. Total N, exchangeable  $\text{K}^+$ , Olsen P.
2. Particle size, pH.
3. Exchangeable  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Na}^+$ , available S.

These variables were added one at a time in the regression in that specific order (forward selection) as fixed effects using maximum likelihood (ML), since the residual likelihood is a direct function of the number of fixed effects within the model. The log-likelihood of models fitted by REML with different fixed effects are not comparable. As the properties were added in this stepwise fashion, the updated model was tested against the previous one by a log-likelihood ratio test (Woelf 1957). A chi-squared  $p$ -value of 0.05 was taken as threshold and any smaller value ( $p \leq 0.05$ ) as evidence that an additional coefficient explained sufficiently more of the total variance to justify inclusion of that property. Once a final set of soil properties were selected as coefficients, the model was refitted by solving Equation 2.6 using REML for unbiased estimates of variance and covariance parameters. The modelling procedure described above was computed for each field individually. The partial least squares regression was done using the **ppls** package (Mevik et al., 2019). The linear models were computed using the **nlme** package in R (Pinheiro et al., 2019). Linear mixed models and the spatial correlation structures were computed using the **geoR** package (Ribeiro & Diggle, 2018).

## 2.3 Results

### 2.3.1 Summary statistics and qualitative description

Three sampling locations in Field 1 and two in Field 2 had lettuce diameters less than 50 mm and were evidently outliers. These individuals seem to have failed to establish after transplantation and were therefore removed before analysing the data formally. Tables 2.1 and 2.2 summarise the statistics of the crops with outliers excluded. Most of the measured soil properties had near-normal distributions. The pH and P were somewhat skewed, but as residuals from the PLSR were approximately normally distributed for pH and P these were not transformed. The distribution of available S and that of its residuals, on the other hand, were strongly skewed, and its concentrations were transformed to logarithms (Table 2.3).

Table 2.1: The summary statistics of observed and predicted lettuce diameters for  $n$  lettuce, along with the number of components used in the prediction and the mean squared error (MSE) and Lin’s concordance correlation coefficient (CCC) of the leave-one-out (LOO) cross-validation for that number of components – Field 1

Variety	$n$	Observed			Predicted			Nr comp	MSE	CCC
		Mean	Std dev.	Range	Mean	Std dev.	Range			
All varieties	218	14.3	3.49	15.5	14.3	2.45	15.7	7	6.15	0.66
Etude	74	12.9	3.63	15.5	12.9	3.09	15.3	4	3.59	0.84
Challenge	71	13.8	2.67	10.5	13.8	2.06	8.34	5	2.87	0.74
Glassica	73	16.2	3.24	14	16.2	2.29	10.1	4	5.23	0.66

The spectral signatures of all soil samples were similar; all had smaller absorbance features in the near-infrared (NIR) than in the mid-infrared region. The NIR includes predominantly weak overtones and fundamental vibrational bands for H–N, H–C and O–H bonds. Absorption bands within the near infrared (NIR) frequently overlap, which makes it difficult to interpret the spectra directly. The MIR is characterized by fundamental frequencies (no overlap) and directly relates to mineral and organic compounds. For example, kaolinite in the Si–O stretching region between 1200 and 1000  $\text{cm}^{-1}$  and carbon functional groups in C–H aliphatic bonds (e.g.  $-\text{CH}$ ,  $-\text{CH}_2$ ) between 3000 and 2850  $\text{cm}^{-1}$  (Viscarra Rossel et al., 2006; Du & Zhou, 2009; Viscarra Rossel et al., 2011; Du et al. 2015). Loadings of the PLSR decomposition can be used for qualitative interpretation of the NIR and MIR spectra. Peaks are caused by the response variable, whereas troughs indicate interference of different soil components (Haaland & Thomas, 1988).

Table 2.2: The summary statistics of observed and predicted lettuce diameters for  $n$  lettuce, along with the number of components used in the prediction and the mean squared error (MSE) and Lin’s concordance correlation coefficient (CCC) of the leave-one-out (LOO) cross-validation for that number of components – Field 2

Variety	$n$	Observed			Predicted			Nr comp	MSE	CCC
		Mean	Std dev.	Range	Mean	Std dev.	Range			
All varieties	106	20.6	5.01	27.3	20.6	3.51	15.1	5	12.6	0.66
Challenge	59	19.1	4.18	18.8	19.1	2.90	10.4	1	8.93	0.65
Glassica	26	19.6	3.47	13.5	19.6	2.39	7.56	1	6.11	0.64
Yucaipa	21	26.1	5.11	18.7	26.1	4.89	18.7	6	2.07	0.96

### 2.3.2 PLSR diagnostics

The appropriate number of components for predicting lettuce diameter and soil properties varied substantially according to variety and target property (Fig. S1). For example, the MSEs of  $K^+$  and clay increased with numbers of components retained. This result is likely to be caused by over fitting. The MSE for P diminished gradually to a minimum at 11 components, whereas the MSE for total N diminished rapidly to a minimum at four. Tables 2.1, 2.2 and 2.3 report the chosen numbers of components.

As expected, the mean of the predicted lettuce diameters was close to the observed diameters; the standard deviations of the predictions were smaller than that of the observations. The scatter plots of predictions against the measured diameters (top 4 panels Fig. 2.3) show that different varieties relate differently to the soil spectra. The predictions of soil properties were generally close to true values (Fig. 2.4). The log ratio of sand and clay over silt were computed to provide two independent variables of particle size. The means of the predictions were also close to the observations, with again standard deviations of the predictions smaller than those of the observations. The predictions of  $Mg^{2+}$ , P and pH had larger ranges than the observed ranges. Nevertheless, these differences are small and fall within the range of the residual mean squared error, leading to a slight over-prediction.

### 2.3.3 Relations between lettuce diameter and predicted soil properties

The diameters of the lettuce were related positively to  $Ca^{2+}$ ,  $Mg^{2+}$ ,  $Na^+$ , total N and total C in both fields and for all varieties (Figs. 2.5, 2.6). There were consistent negative relations with pH in both fields, probably because there are patches of relatively acid peat in the fields. The correlation coefficient between pH and total C was  $-0.83$  based on the wet chemistry measurements (not shown). In Field 2,

Table 2.3: The summary statistics of observed and predicted soil wet chemistry data for the 60 calibration samples, along with the number of components used in the prediction, the mean squared error (MSE) and Lin's concordance correlation coefficient (CCC) of the leave-one-out (LOO) cross-validation for that number of components.

Property	Observed			Predicted			Nr comp	MSE	CCC
	Mean	Std dev.	Range	Mean	Std dev.	Range			
Total C/%	12.31	3.67	14.2	12.3	3.64	13.9	6	0.42	0.99
Total N/%	0.84	0.25	0.96	0.84	0.25	0.96	6	0.026	0.99
Ca <sup>2+</sup> /mg kg <sup>-1</sup>	7271	417.8	4958	7271	1063	3786	1	400	0.93
K <sup>+</sup> /mg kg <sup>-1</sup>	347	112	514	347	98	487	9	54	0.87
Mg <sup>2+</sup> /mg kg <sup>-1</sup>	419	117	480	419	113	485	10	29	0.97
Na <sup>+</sup> /mg kg <sup>-1</sup>	48.8	21.4	86.6	48.8	19.5	76.9	6	8.6	0.91
Mn <sup>2+</sup> /mg kg <sup>-1</sup>	0.51	0.34	1.65	0.52	0.32	1.18	5	0.12	0.94
P/mg kg <sup>-1</sup>	41.5	11	61.4	41.5	10.5	61.6	11	3.2	0.95
S/mg kg <sup>-1</sup>	13.3	13.8	69.1	12.3	11.4	67	6	6.6	0.86
pH	7.09	0.58	2.26	7.09	0.58	2.31	12	0.038	1
Sand/%	30.6	5.13	25.1	30.6	4.83	17.2	4	1.7	0.94
Clay/%	36	4.06	29.7	36	3.83	17.2	4	1.3	0.94
Silt/%	33.3	2.19	9.35	33.3	1.81	7.23	3	1.2	0.81

lettuce diameter related positively to Mn<sup>2+</sup>, available S and to log(sand/silt). The relation between lettuce diameter and P was weakly negative, possibly due to a bias from the PLSR predictions. As expected from the soil-forming history in the region, the soil properties appear to co-vary with elevation (see Figs. 2.5, 2.6). The final fitted LM (2.2) relating the predicted soil variables to lettuce diameter comprised total N, K<sup>+</sup> ( $p \leq 0.0001$ ), log(clay) ( $p \leq 0.0001$ ) and pH ( $p = 0.003$ ) for Field 1 as coefficients; and variety, total N, K<sup>+</sup> ( $p = 0.0187$ ) and P ( $p = 0.001$ ) for Field 2.

Further investigation of the LM residuals suggested that the exponential variogram model (2.5) would describe their spatial autocorrelation well in both fields. This model was therefore included in the regression (2.2). The final model for Field 1 retained variety, total N, K<sup>+</sup> ( $p = 0.0114$ ) and log(clay) ( $p = 0.0004$ ) as coefficients. The final model for Field 2 retained variety, total N, K<sup>+</sup> ( $p = 0.0195$ )

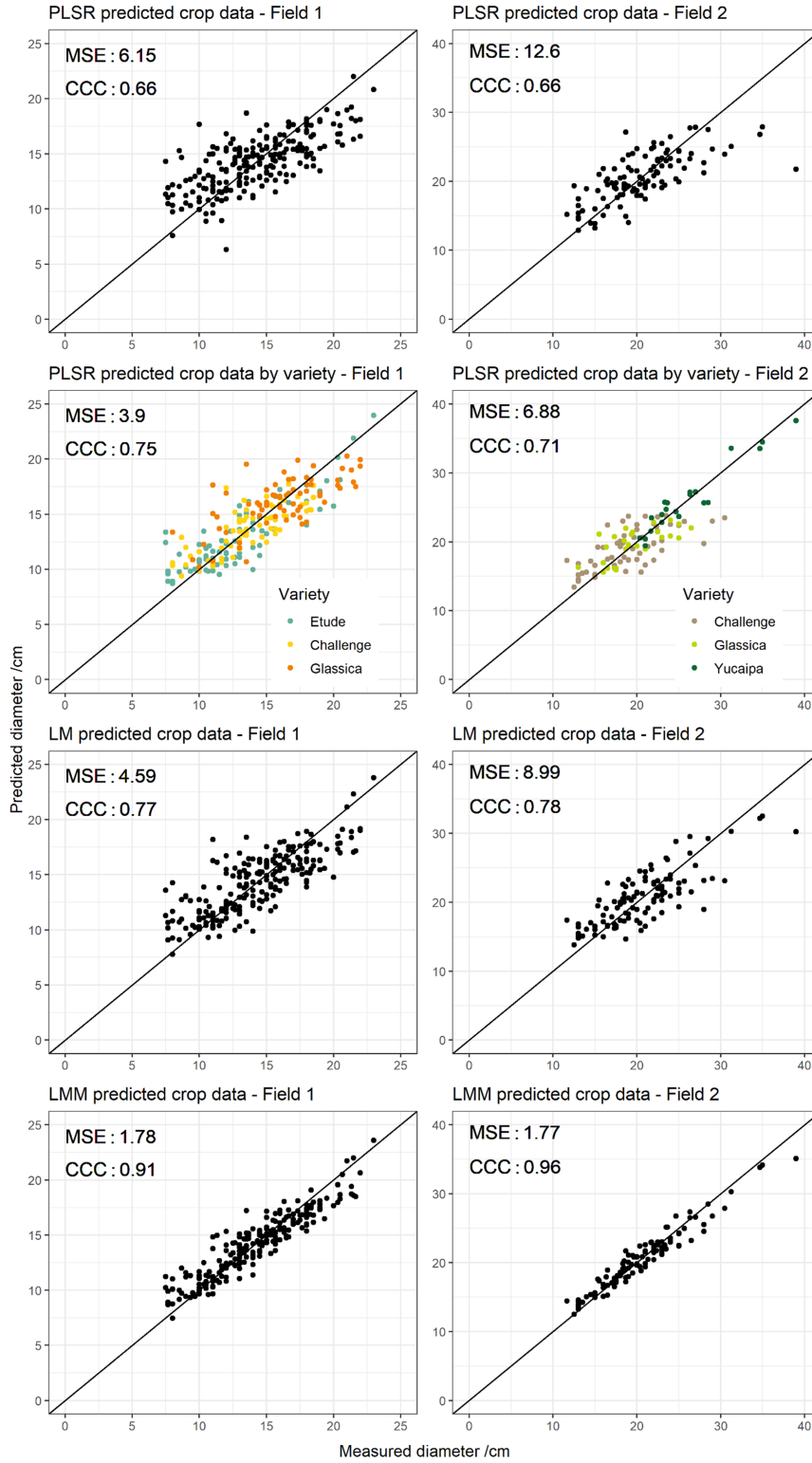


Figure 2.3: Predicted versus measured lettuce diameter for both fields. The top two rows show predictions from partial least squares regression (PLSR) from the soil spectra, where the second row takes lettuce variety into account. The bottom two rows show predictions from the multiple regression models (LM) and the linear mixed models (LMM) from the IR predicted soil properties

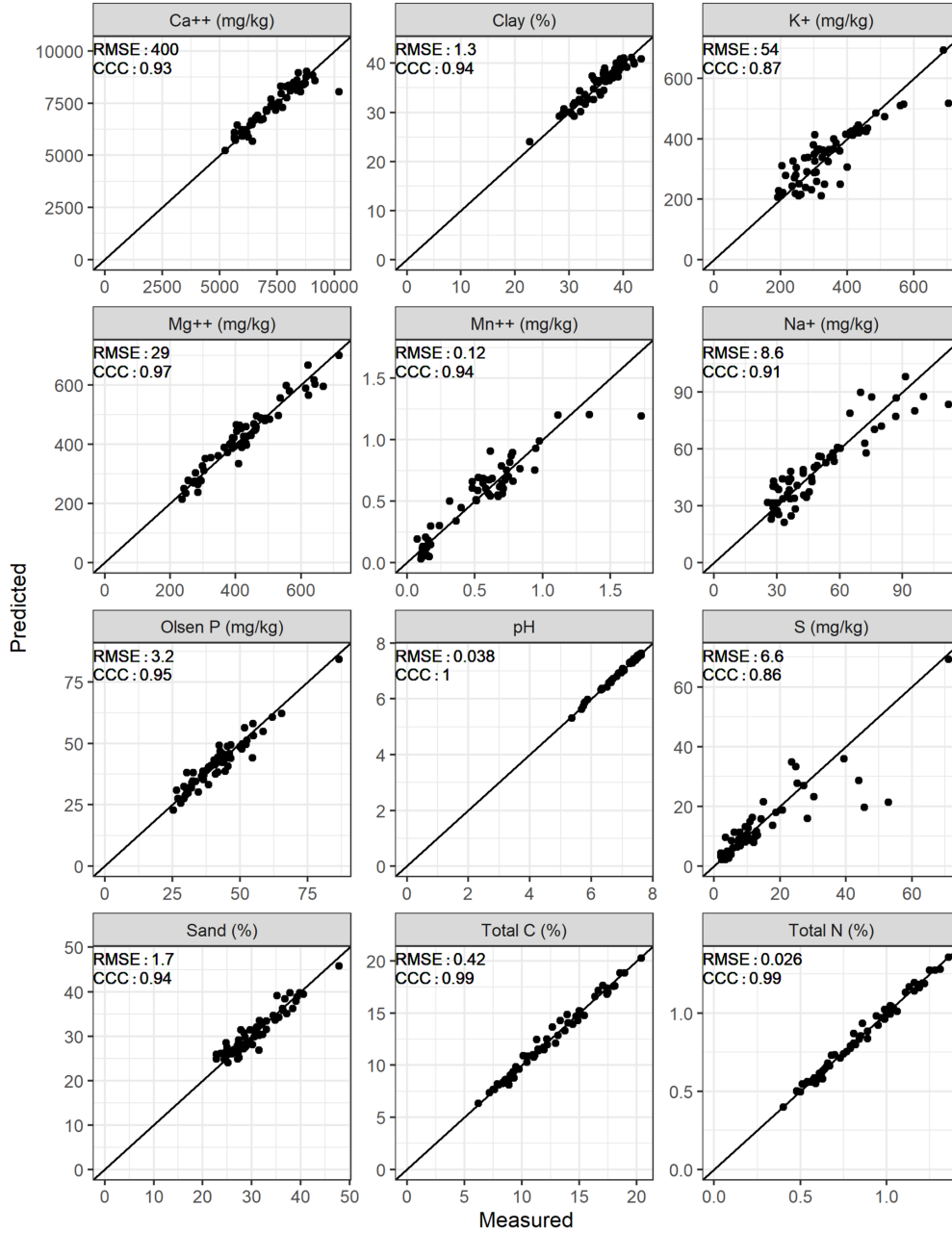


Figure 2.4: Leave-one-out predicted against measured soil properties for the 60 calibration samples from both fields by partial least squares regression (PLSR) with the soil spectra. Metrics include the root mean squared error (RMSE) and Lin's concordance correlation coefficient (CCC)

and  $P$  ( $p = 0.0013$ ) as coefficients (See Tables S1 and S2 in supplementary material for fixed effects coefficients).

The variograms associated with covariance models for the spatial autocorrelation term for Eq. 2.5 are in Fig. 2.7. The parameters for Field 1 are  $c_0 = 2.9$ ,  $c_1 = 2.8$  and  $a = 17.2$  m. Those for Field 2 are  $c_0 = 4.2$ ,  $c_1 = 6.8$  and  $a = 29.2$  m. The nugget variance ( $c_0$ ) in Field 1 is approximately half of the total variance ( $c_0 + c_1$ ). For Field 2, it comprises somewhat less than half of the total. The nugget includes both measurement error and very short-range, unresolved, spatial variation. The effective limit of spatial correlation, i.e. the effective range (approx  $3a$ ), in Field 2 is almost

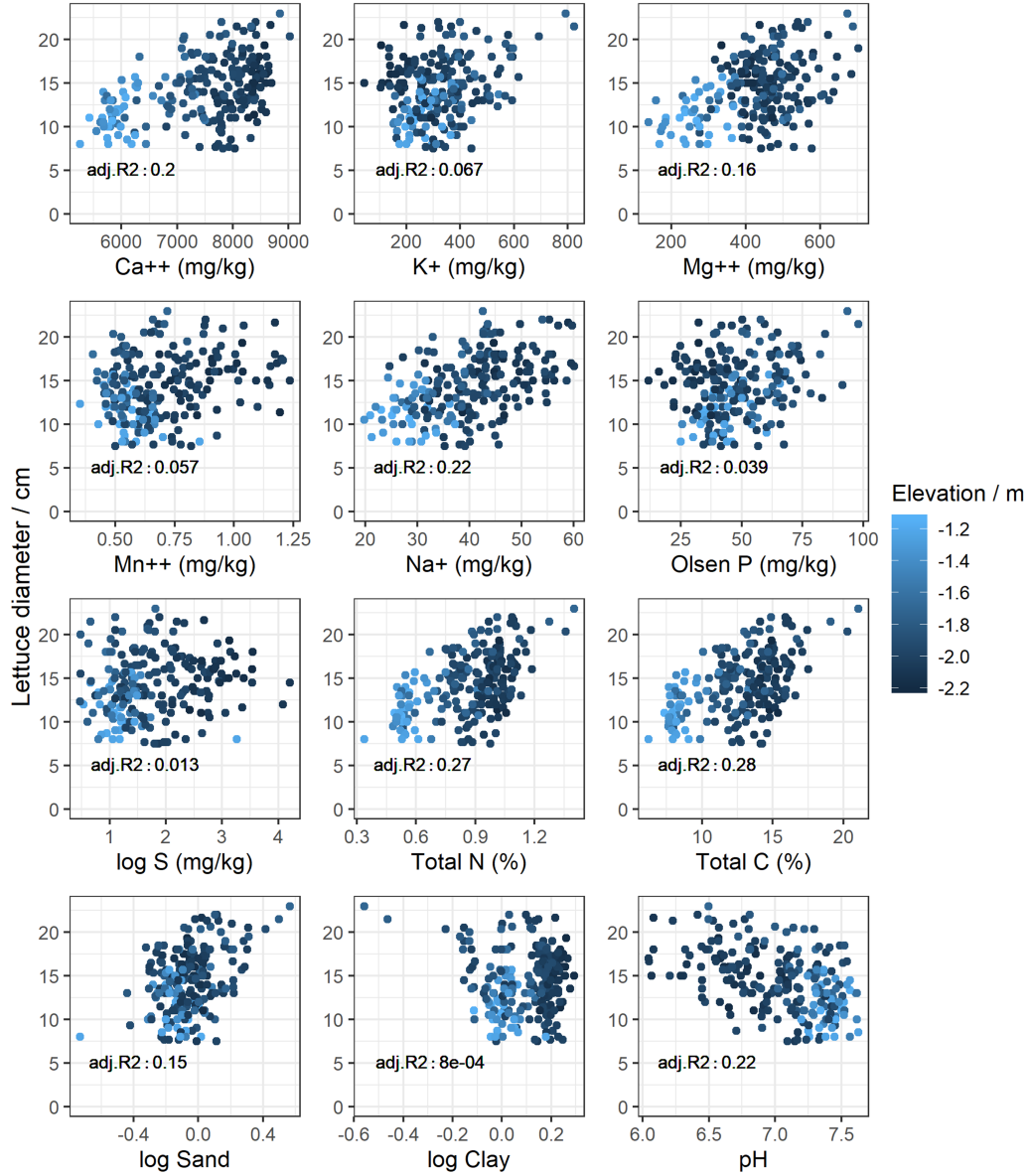


Figure 2.5: Measured lettuce diameter coloured by elevation (LiDAR) against partial least squares regression (PLSR) predicted soil properties—Field 1. Adj.  $R^2$  is the coefficient of determination from linear regression between lettuce diameter and the PLSR predicted soil property, adjusted by the degrees of freedom

twice that in Field 1. The mean-squared errors and Lin's concordance correlation coefficients for the LM and LMM predictions were computed (bottom 4 panels Fig. 2.3).

#### 2.3.4 PLSR loadings and their interpretation

As previously described, the loadings of the PLSR decomposition can be used for qualitative interpretation of the spectra. Peaks are caused by the soil property used as the response variable, whereas troughs indicate interference of different soil properties measured by the spectra. For each property included as a coefficient in the

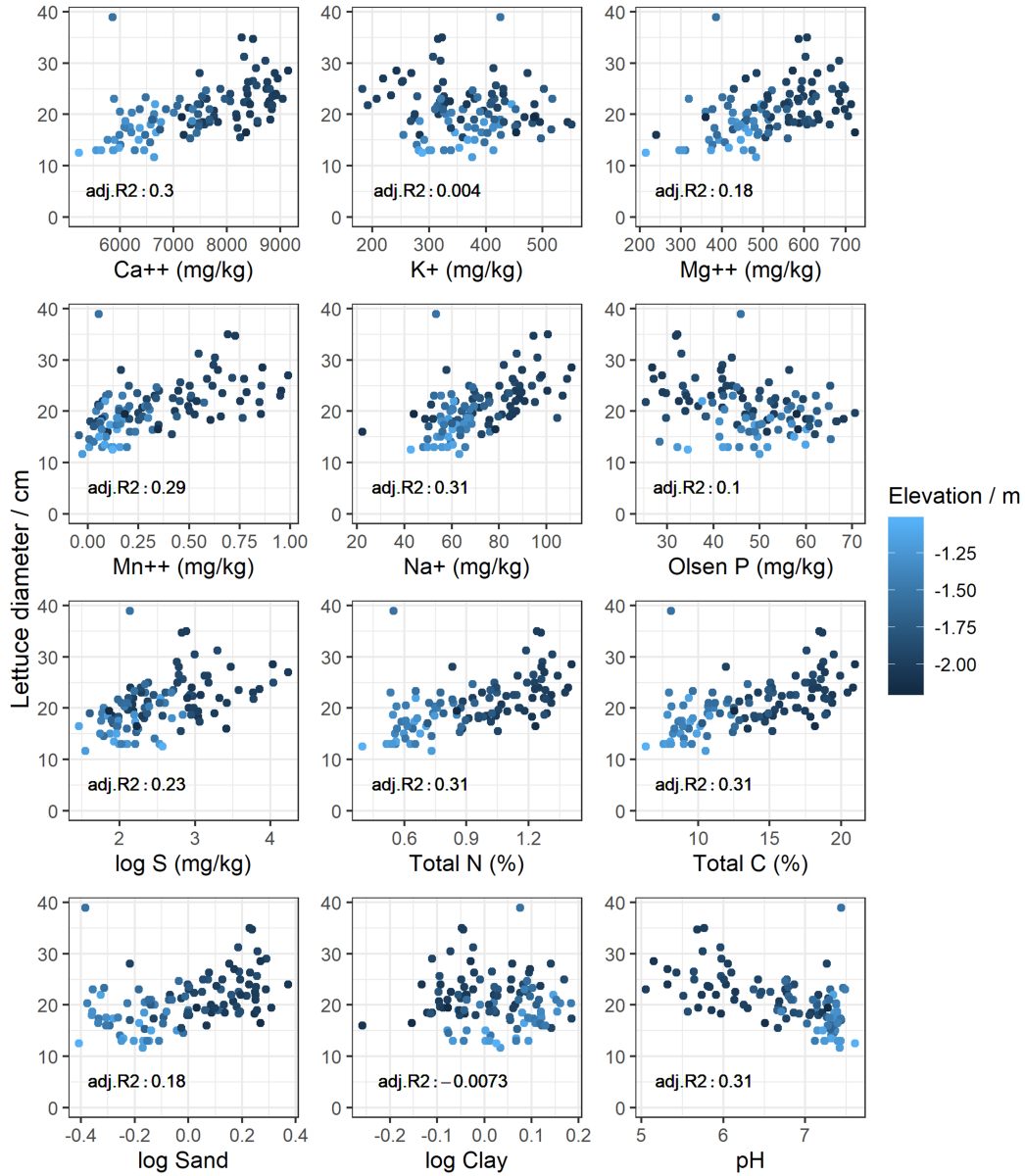


Figure 2.6: Measured lettuce diameter coloured by elevation (LiDAR) against partial least squares regression (PLSR) predicted soil properties–Field 2. Adj.  $R^2$  is the coefficient of determination from linear regression between lettuce diameter and the PLSR predicted soil property, adjusted by the degrees of freedom

LMM, the loadings of the first component were plotted against wave number. The loadings of the first component in the final fitted PLSR models to predict lettuce diameter from the soil spectra were also plotted. The loadings from the total N PLSR model and to some extent those of the  $K^+$  PLSR model align better than the other soil properties with the loadings from the models that predict lettuce diameter from the soil spectra (Fig. 2.8). Closer loading alignment indicates that the same wave numbers explain an equal amount of variance from the response variable.



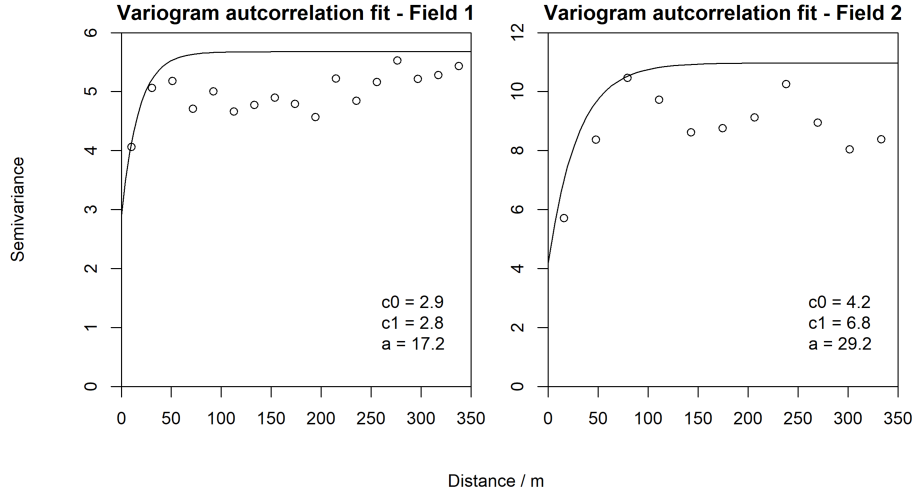


Figure 2.7: Fitted exponential model for the spatial autocorrelation term in the linear mixed model (LMM) for Fields 1 and 2, where ‘c0’ and ‘c1’ are the nugget and sill of the correlated variance and ‘a’ is the distance parameter

## 2.4 Discussion

The aim of this study was to discover whether soil spectral measurements could be used to predict variation in the diameter of lettuce grown in commercial fields, both directly from the soil spectra and via predictions of soil properties.

### 2.4.1 Lettuce diameter prediction directly from soil spectra

For both fields, the predictions of lettuce diameter from the soil spectra were close to observed values. These predictions were even better when separate models were fitted to each variety (top 4 panels Fig. 2.2). As the varieties were grown in blocks, the observed effect of variety could also be an effect of management that is independent of the soil. For example, some parts in Field 2 were not covered by the irrigation system. Furthermore, varieties associated with parts of the field with larger variation in soil properties were better predicted by the soil spectra (see concordance correlation coefficients in Tables 2.1, 2.2). Wetterlind et al. (2008) found that plant N uptake could only be modelled accurately for fields that showed a large range in soil organic matter and texture. The fields in this study are characterized by large variation in soil properties and hence can fulfil this requirement *reference*(Figs. S2 and S3).

### 2.4.2 Precision of soil property predictions from soil spectra

Overall, the predictions for soil properties were good (Fig. 2.4). The reported errors might be optimistic for some properties because leave-one-out cross validation tends to over-estimate the accuracy and precision in PLSR (Viscarra Rossel, 2008). The errors (expressed as root mean squares, RMSEs) proved to be similar to those found by other investigators: see Viscarra Rossel et al. (2016) for a review of 51 studies in which soil organic carbon was predicted from reflectance spectra with RMSEs ranging from 0.1 to 1.1%. However, the comparisons between results from this study and those reported in the literature need to be viewed with caution because the

variances depend to some extent on the concentration of the target variable: smaller concentrations tend to be predicted with smaller errors. For example, Hutengs et al. (2018) reported RMSEs of 0.14–0.24% for SOC in the range 0.62–2.70%, whereas those from this study were 0.54% (from the leave-one-out cross validation) for the range 6.21–20.41%. Soil properties that had larger RMSE in their predictions are not spectrally active (e.g. S, P and  $K^+$ ). This accords with other studies in which the prediction of non-spectrally active properties are functions of their correlation with soil organic matter and particle-size distribution. Predictions for these properties are hence less robust than SOC and particle-size fractions (Du & Zhou, 2009 and references therein). See Table S3 in the supplementary material for the correlations between laboratory reference values.

The numbers of components included in the PLSRs are akin to those reported in the literature (namely from 3 to 9, Yang & Mouazen, 2007; Wang et al., 2015; Hutengs et al., 2018). They rarely exceed 12, and they were fewer for soil properties that have a direct relation to molecular bonds in the MIR (e.g. C and clay). More components are generally retained for larger sets of spectral data where there are more pronounced differences in lithology, climate and other soil forming factors—see for example Dangal et al. (2019) and Lopo et al. (2016). The PLSR loadings from the first component indicate that the NIR region explains little of the variance and much less than those in the MIR region. This holds true for both the PLSR models to predict lettuce diameter and the models to predict soil properties. These findings accord with the literature. The loadings depend on the soil properties of interest and are unique for each study; nevertheless, the MIR region generally leads to more robust calibration than does the NIR (Viscarra Rossel et al., 2006; Yang & Mouazen, 2007).

### 2.4.3 Can values of soil properties predicted from the spectral data be used to predict lettuce diameter?

The LMs that related lettuce diameter from predicted soil properties performed reasonably well (MSE: 4.59 for Field 1 and MSE: 8.99 for Field 2). The prediction performance of the LM implies that it captures a large amount of the explanatory power of the IR spectra. However, the PLSR from the soil spectra alone predicted the lettuce diameter more precisely than did the LMs. It seems that the IR spectra capture more information about the soil relevant for crop growth than the soil properties included in the LM. This effect aligns with studies that compare crop predictions from IR spectra with crop predictions from laboratory reference values, in which the first outperforms the latter (Börjesson et al., 1999; Wetterlind et al., 2008).

The predictions from the LMMs (Eq. 2.5) were more precise than those from the soil spectra (by PLSR alone) with a difference in mean squared error (MSE) of 21 mm<sup>2</sup> for Field 1 and of 51 mm<sup>2</sup> for Field 2. This is because the LMMs account for the spatial structure in the lettuce diameters through the random term in the model. These results are somewhat misleading because the prediction of each lettuce relies on its spatial auto-correlation with the other lettuces in the field. In practice, growers would not be able to predict lettuce size at the beginning of the season with this model because they would not have these other measurements. Therefore, although the auto-correlation in the model is the necessary model giving

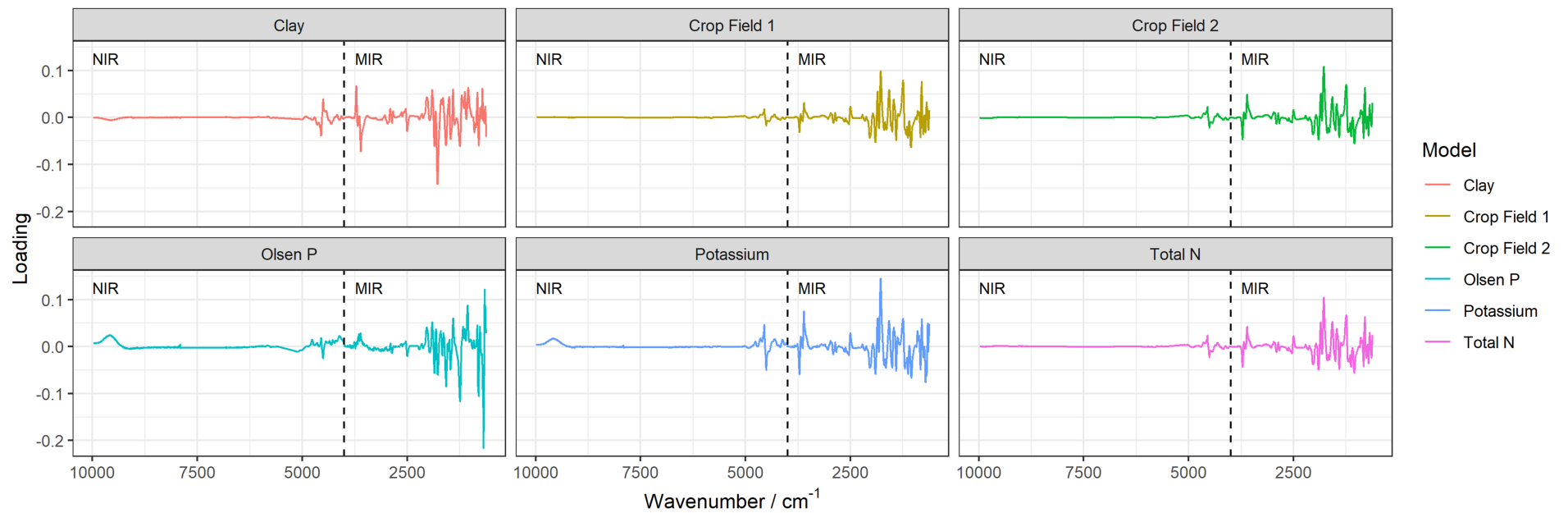


Figure 2.8: Loading values from the first component plotted as a function of wavenumber for PLSR models relevant in the modelling of lettuce diameter

the sampling design, it is not of practical use to the grower.

#### 2.4.4 Can soil spectra help growers with management?

This study showed two different ways of using the soil spectra. First, PLSR was used to predict lettuce diameter directly from the soil spectra. This does not allow for precise fertiliser or irrigation management as there is no information in how soil properties affect the lettuce response in the model. However, the predictions could be of practical use to management when the areas predicted to be high or low-yielding are consistent over seasons (which may not always be the case, Milne et al., 2012; Diacono et al., 2013 and references therein; Kindred & Sylvester-Bradley, 2014) or if the grower has prior knowledge on why some areas are high or low-yielding. In this case, the management can be adapted based on an understanding of the causes of poor yields. For example, variable-rate planting or deciding not to crop certain areas if the soil spectra indicate likely poor yields that will lead to in-field yield waste. In this particular case, when lettuce are predicted to be less than 0.1 m in diameter (a size deemed too small by the growers) the grower may choose not to plant there.

The second approach was to predict lettuce diameter from estimated soil properties using a LM. Although predictions were poorer than the direct prediction by using the reflectance spectra, this approach gives growers more information and so could help them decide how to vary the application of fertiliser and irrigation within each field. These models are relevant when estimated soil properties are used for precise fertilization or irrigation. A remaining question is how to predict the exact amount of each nutrient needed by the crop at each place in the field and when to apply it (Baveye & Laba, 2015; Kindred et al., 2017). The grower therefore needs prior knowledge on which soil properties influence crop growth for each specific part of each field. With this understanding, potential environmental impacts of farming can be minimised and profits maximised. This can be achieved by either not planting in low-yielding areas (and therefore no cost of fertiliser, herbicide and irrigation), or by managing inputs more precisely so that the economic return in crop response exceeds the amount spent on inputs. The LM and LMM reported here showed that total N, P, K and pH were significant predictors of lettuce size, indicating that variable rate application based on these properties could be used to advantage in lettuce production. Panagapoulos et al. (2006) demonstrated such an approach by creating “lettuce production capability” maps from kriged soil properties and identifying localised areas where the soil could be treated to improve yield. This illustrates the potential utility of predicting the variation of soil properties from IR spectra for the precision management of lettuce (in particular fertiliser and management of soil pH).

Soil spectra offer great promise for the precision management of crops but collecting the soil samples from the field and processing them (i.e. drying and milling) adds to the expense for a farmer. Therefore in practice, field-based spectral measurements are likely to be more attractive than spectral measurements made in the lab. Currently there has been limited exploration of field versus lab-based prediction errors with portable MIR spectrometers for soil properties other than soil carbon constituents (Ji et al., 2016a; Hutengs et al., 2019). Differences between field and lab-based MIR predictions range from 0 to 45% whereas the increase in

error associated with VNIR techniques is reported to be as much as 57% (Ji et al., 2016a; Hutengs et al., 2019). Field based prediction for pH, organic matter and total nitrogen have been explored using (V)NIR with increases in prediction error of 14%, 27% and 22% compared to lab-based spectral measurements, respectively (Ji et al., 2016b). Most of the few existing studies on macro nutrient prediction using portable (V)NIR/MIR spectroscopy show limited success (Wenjun et al., 2014; Ji et al., 2016a). Poor prediction performance is commonly attributed to the absence of distinct features in the IR spectrum and varying relationships between total and available element content (Kuang et al., 2012; Pätzoldt et al., 2019). Consequently, the prediction models for available major and trace nutrients from soil spectra often prove to be less robust than those developed for particle size fractions and soil organic carbon. An exception is the study by Mouazen & Kuang (2016) with an 18% error increase for field-based soil available P predictions compared to predictions from spectra measured in the laboratory. These accurate predictions can probably be attributed to the large number of calibration samples used by the authors. Given that the linear models showed K and P to be important predictors of lettuce diameter, and that most studies show poor prediction of these variables from field-based measurements, further development of sensor technology is required for field-based measurements to be of practical use for this study's methodology.

In the study reported here, properties of 30 samples from each field were measured by wet chemistry and the values were used to calibrate the models. Samples of this size are more than a commercial grower could expect to take in fields of 10.5 and 18.2 ha. In practice, there will be a trade-off between the number of calibration samples, with their associated costs, and accuracy of the prediction. Optimization of the sampling design, sample processing and the number of replicate measurements are examples of other factors that affect the accuracy. Optimization of sampling design depends not only on the sizes of samples to provide reference data but also on good coverage of the conditions within the field (Ramirez-Lopez et al., 2019 and references therein) or parent material (Sila et al., 2016). Once errors in soil predictions have been properly estimated these must be propagated through to the predictions of crop response. Only when these are properly accounted for could one estimate the true value of measuring the soil spectra for precision application of fertiliser and irrigation (Ramirez-Lopez et al., 2019).

Thus, the relevance of using the soil spectra directly or via estimated soil properties will depend on the situation. This study showed that under optimal conditions, there is potential for associating crop response to soil reflectance spectra. This association can be made directly from the soil reflectance spectra or by a regression that uses soil property values estimated by reflectance spectra.

## 2.5 Conclusions

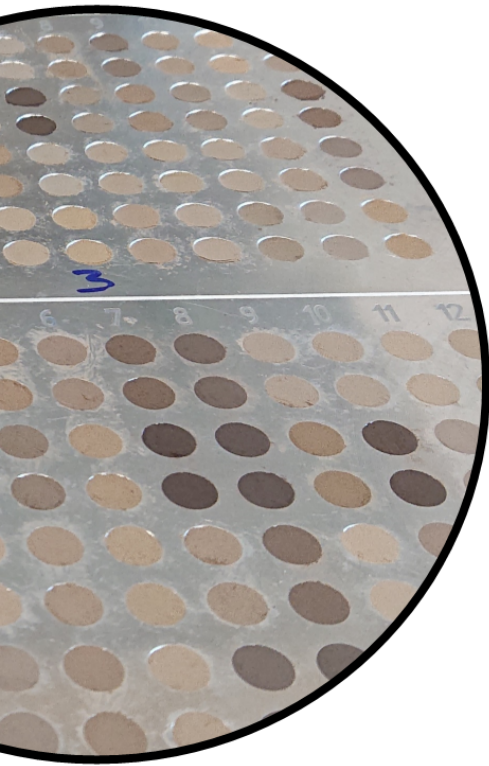
Reflectance spectra from soil in the near- and mid-infrared range were related to the diameters of lettuce grown in two fields in the Fenland region of England. They led to reasonably precise predictions of lettuce diameter and therefore are of value to the grower. The partial least squares regression (PLSR) that used soil spectra as response variables showed a mean squared error (MSE) of 39 mm<sup>2</sup> for Field 1 and 68.7 mm<sup>2</sup> for Field 2. Predictions of lettuce diameter that used linear models with

the PLSR estimated soil properties gave somewhat poorer results, with a difference in MSE for Field 1: 6.9 mm<sup>2</sup> and Field 2: 21.2 mm<sup>2</sup>). Predictions from the linear mixed models were more precise than those from the raw spectra (by PLSR alone) with a difference in MSE of 21.2 mm<sup>2</sup> for Field 1 and of 51 mm<sup>2</sup> for Field 2.

The spectra were related strongly to soil properties that determine crop growth, specifically, nitrogen (measured as total N), available phosphorus (P), exchangeable potassium (K<sup>+</sup>), clay content and pH. Using the values of the soil properties estimated from the reflectance spectra to predict the sizes of the lettuce was somewhat less precise than direct prediction from the spectra. The advantage to the grower of the indirect prediction is the gain in knowledge about which soil properties are important. This enables the grower to adapt the management to the soil. Precise indirect prediction is only feasible with a suitable calibration dataset that captures the variability of the underlying soil.

## Chapter 3

# Quantifying uncertainty from reducing field-scale calibration effort



# Abstract

*The prediction accuracy of soil properties by proximal soil sensing has made their application more practical. However, in order to gain sufficient accuracy, samples are typically air-dried and milled before spectral measurements are made. Calibration of the spectra is usually achieved by making wet chemistry measurements on a subset of the field samples and local regression models fitted to aid subsequent prediction. Both sample handling and wet chemistry can be labour and resource intensive. We consider two approaches to reduce these expenses for predictions made from visible-near-infrared ((V)NIR), mid-infrared (MIR) spectra and their combination. First, we considered reducing the level of processing of the samples by comparing the effect of different sample conditions (in-situ, unprocessed, air-dried and milled). Second, we explored the use of existing spectral libraries to inform calibrations (based on milled samples from the UK National Soil Inventory) with and without ‘spiking’ the spectral libraries with a small subset of samples from the study fields. Prediction accuracy of soil organic carbon, pH, clay, available P and K for each of these approaches was evaluated on samples from agricultural fields in the UK.*

*Available P and K could only be moderately predicted with the field-scale dataset where samples were milled. Therefore this study found no evidence to suggest that there is scope to reduce costs associated with sample processing or field-scale calibration for available P and K. However, the results showed that there is potential to reduce time and cost implications of using (V)NIR and MIR spectra to predict soil organic carbon, clay and pH. Compared to field-scale calibrations from milled samples, we found that reduced sample processing lowered the ratio of performance to inter-quartile range (RPIQ) between 0% and 76%. The use of spectral libraries reduced the RPIQ of predictions relative to field-scale calibrations from milled samples between 54% and 82% and the RPIQ was reduced between 29% and 70% for predictions when spectral libraries were spiked. The increase in uncertainty was specific to the combination of soil property and sensor analysed. We conclude that there is always a trade-off between prediction accuracy and the costs associated with soil sampling, sample processing and wet chemical analysis. Therefore the relative merits of each approach will depend on the specific case in question.*

Based on:

Breure, T.S., Prout, J.M., Haefele, S.M., Milne, A.E., Hannam, J.A., Moreno-Rojas, S., Corstanje, R., 2021. Comparing the effect of different sample conditions and spectral libraries on the prediction accuracy of soil properties from near- and mid-infrared spectra at the field-scale. *Soil & Tillage – Accepted*



## 3.1 Introduction

Farmers are interested in the spatial variation of soil properties because this helps them explain the variation in crop performance and so infer appropriate interventions. Mapping subfield soil variation in the traditional manner (i.e. analysed by wet chemistry analysis) is usually deemed too expensive to obtain the accuracy required for precision agriculture (Muhammed et al., 2017). Improvements in technology, mean proximal and remote soil sensing (for example using visible (V), near-infrared (NIR) and mid-infrared (MIR) spectroscopy) now offers an alternative and less resource demanding approach to predict soil variation than measurements based on wet chemistry (Viscarra Rossel & Bouma, 2016). Due to the reduced labour and monetary inputs, soil spectroscopy can be implemented at finer sampling scales than traditional sampling. For example, at a 10-m scale, which is reported to be necessary to characterise spatial and temporal soil variability for site-specific management (McBratney et al., 1996). Despite the practical advantages of using soil spectral measurements over wet chemistry, issues of efficiency still need to be addressed before wide-scale adoption is practical (Reeves, 2010). These largely relate to reducing sample processing and using spectral libraries to minimise resource input under the constraint that to be practically useful, however, they should maintain accuracy near to that of laboratory methods (Viscarra Rossel & McBratney, 1998).

The common methodology of soil preparation before the measurement of soil reflectance spectra includes air-drying and sieving ( $<2$  mm) and for MIR milling ( $<100$   $\mu\text{m}$ ). Minimizing the sample processing can reduce the accuracy in the soil property predictions due to effects of particle-size, aggregation and water content on spectroscopy measurements. A number of studies have researched these effects for (V)NIR/MIR soil spectroscopy. For example, studies analysed the effect of different particle sizes (Nduwamungu et al., 2009; Le Guillou et al., 2015; Coutinho et al., 2019; Wijewardane et al., 2020), soil water content (Bogrekci & Lee, 2006; Minasny et al., 2011; Ji et al., 2015) and in-field (V)NIR measurements (Stevens et al., 2008; Gras et al., 2014) on soil spectroscopy predictions.

Within the soil IR spectroscopy discipline, there have been efforts to develop spectral libraries (a point-dataset with paired reflectance and wet chemistry measurements) at local, regional, continental (Shepherd & Walsh, 2002; Viscarra Rossel & Webster, 2012; Stevens et al., 2013) and even global scales (Viscarra Rossel et al., 2016). Where traditional soil survey data already exists, creating a spectral library has the potential to minimise the effort of developing field-scale calibrations.

Ideally, existing literature would be consulted to infer a quantified effect on prediction accuracy of using either reduced sample processing or spectral libraries to minimize calibration expenses. However, comparison across literature is hampered by differences in case-study characteristics (e.g. overall variance of soil properties and their counterparts in the calibration and validation set) and methods (e.g. number of samples with wet chemistry used in the calibration, chemometric models considered, (cross-)validation techniques used etc.). For example, due to increased availability of portable MIR spectrometers, recent studies have explored the accuracy of in-situ MIR measurements. However, the comparison between these studies is not straightforward as there are differences in the number of replicate measurements

taken (Webster et al., 2016; Hutengs et al., 2019), whether or not in-situ means <2 mm sieved soils (Webster et al., 2016), the MIR spectrometer may be different (Dhawale et al., 2015; Ji et al., 2016; Webster et al., 2016) and the range of wave numbers captured can vary (Dhawale et al., 2015; Ji et al., 2016).

Prediction accuracy from spectral libraries at a local-scale have also been shown to be affected by different instruments or laboratory conditions, under-representation of the local soil type and differences in lithology, climate and other soil forming factors (Wetterlind & Stenberg, 2010; Ge et al., 2011; Guerrero et al., 2014). To overcome some of the limitations of using spectral libraries, Brown (2007) developed an approach to compute adequate models for new local target sites by adding local samples to a spectral library, which has been described as ‘spiking’ (Viscarra Rossel et al., 2009). Spiking has been shown to improve prediction accuracy (Guerrero et al., 2010; ; Wetterlind & Stenberg, 2010; Guerrero et al., 2014; Guerrero et al., 2016; Seidel et al., 2019)

We are unaware of any previous research that examined the effects of in-situ sensing, particle size variation, aggregation and soil water content on spectral measurements for both (V)NIR and MIR spectroscopic predictions within a single study. Furthermore, contrasting the effect on prediction accuracy of reduced sample processing with that of spectral libraries on a single dataset will advance our understanding of when one approach could be preferable over the other. To that end, we explored the following questions:

- What is the difference in accuracy between soil property predictions from (V)NIR, MIR and (V)NIRMIR spectroscopy measurements taken on in-situ, unprocessed (i.e. fresh), air-dried or milled soil samples?
- If we were to use a spectral library rather than a field-scale calibration, does subset selection from a national spectral library by region or pedological characteristics minimise the prediction error?
- What can we learn from these findings to determine the best way in which to reduce laboratory, sampling preparation and handling efforts whilst minimising the loss in prediction accuracy?

## 3.2 Methods

Using soil samples from four fields within the Cambridgeshire fens in the UK, we evaluated two approaches for reducing the expense associated with soil variable predictions made from visible-near-infrared ((V)NIR), mid-infrared (MIR) spectra and their combination. The first considers reducing effort related to sample processing (Fig. 3.1 A) and the second by using regional and stratified soil spectral libraries (with and without spiking) (Fig. 3.1 B). The sample processing steps ranged from standard laboratory processing of soil for spectral analyses (air-dried, sieved and milled) to a gradual reduction of the laboratory processing effort (removing milling, sieving and air-drying) to taking spectral measurements in the field without sample processing. Wet chemistry was conducted on the field-scale dataset, which was then split into calibration (75%) and validation (25%) subsets. For the reduced sample processing analysis, the calibration samples were used to

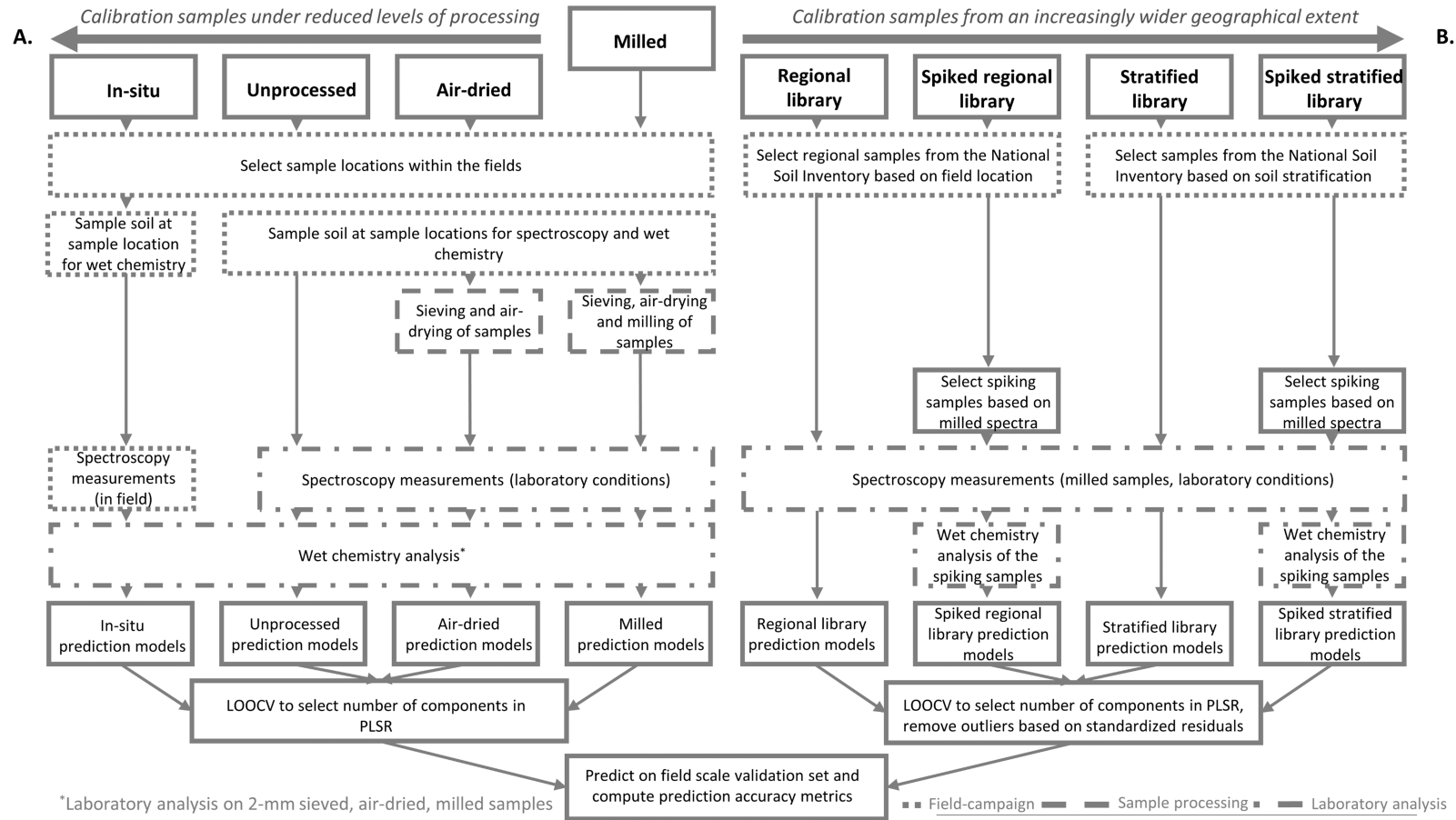


Figure 3.1: A schematic showing two strategies to reduce the effort required to make predictions about soil properties from visible-near-infrared ((V)NIR) and mid-infrared (MIR) reflectance spectra. The first is to reduce the processing of the soil samples used for calibration and those used for prediction (in-situ, unprocessed and air-dried). The second uses soil spectra from a National Soil Inventory (NSI). In this case the soils for prediction must be air-dried, sieved and milled to accord with those from the NSI. They can be chosen according to how representative they are, in this study based on geographic location (Regional library) or soil type (Stratified library). In both cases we also consider “spiking” the library set with soils from the field-scale dataset for which we wish to predict soil properties. PLSR stands for partial-least squares regression, the method of regression used in this study. LOOCV stands for leave-one-out cross-validation, the procedure used in this study to select the final model for prediction.

develop regression models between the measured soil properties and the spectra for each processing method.

We used two methods to subset a national soil inventory (NSI) into spectral libraries. Samples were selected i.) in close geographical extent to the field-scale dataset (the regional library) and ii.) by the two soil types found at the field site (the stratified library). Representative samples from the field-scale calibration subset were selected to spike the regional and stratified libraries. Regression models were developed for the regional and stratified libraries (and spiked versions). All prediction models were applied to the field-scale validation subset and model accuracies computed. Details of data collection, processing and analysis of the various datasets are presented below.

### 3.2.1 Formation of spectral libraries

#### Field-scale dataset

The four experimental fields used in this study make up the field-scale dataset and are located within the Cambridgeshire fens, south-east of England (UK). The fens contain complex soils which are a combination of peat with underlying alluvial and marine silts that became elevated features in the landscape due to lowering of the peat surface by oxidation and wind erosion (Hodge et al., 1984). The two soil types present are classified according to the World Reference Base taxonomy as a drainic sapric Histosol (dr sa HS) and a mollic Gleysol (mo GL) (IUSS Working Group WRB 2015). Field 1, covering 8.2 ha, with British National Grid reference: TL607880, lies adjacent to Field 2 which covers 16.9 ha. Field 3 lies 8.3 km south-west, covering 5.1 ha and Field 4 lies 7.5 km south of Field 1 and Field 2, covering 8.9 ha. Three soil cores of topsoil (0–25 cm) were taken within a 0.5 m  $\times$  0.5 m quadrat at 25 sampling locations in each field. The fields were sown with lettuce which do not have a substantial root system, and any previous thatch layer was mixed in by tillage. For each sampling location, the three soil cores were bulked and mixed for laboratory analysis and spectral study, described in detail below (see also Breure et al., 2020). Direct spectral measurements of the soil surface were also taken at the sample locations in each field. Given the restricted number of samples for each field, we considered them as a single dataset. Since three locations had incomplete measurements we continued the analysis with a field-scale dataset where  $n = 97$ .

#### Spectral library subsetting

We formed two spectral libraries. The samples that make up the two spectral libraries are a subset of the National Soil Inventory (NSI) dataset of England and Wales (McGrath & Loveland, 1992). The topsoil samples (0–15 cm) were taken as part of a 5 km  $\times$  5 km grid-based soil survey from 1979–1983. A full description of the survey methods, analytical methods and available data is given in the LandIS database ([www.landis.org.uk](http://www.landis.org.uk); Proctor et al., 1998).

The two spectral libraries were selected according to two different methodologies, and we refer to these as the regional library and stratified library (Fig. 3.2):

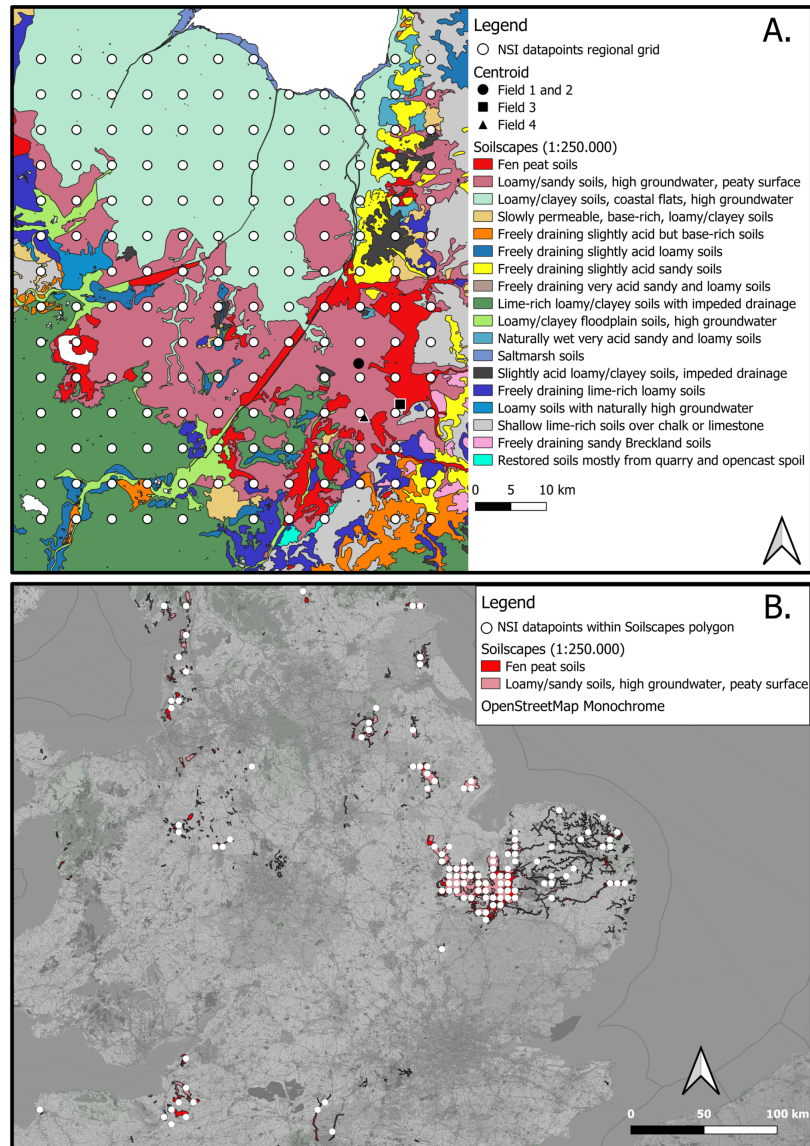


Figure 3.2: A.) National Soil Inventory (NSI) samples selected with the regional approach overlaid on the SoilScapes (1:250,000) dataset with the centroids of the study fields, B.) NSI samples selected with the stratification approach, overlaid on the SoilScapes polygons of the two major soil types that occur in the study fields.

**i. Regional library** Based on the SoilScapes dataset (1:250,000 scale) (Farewell et al., 2011) a regional grid was selected around the case study area. We classified our field-scale dataset by two soil type descriptions: ‘Fen peat soils’ and ‘Loamy and sandy soils with high groundwater and a peaty surface’. We placed the regional grid such that it was centred around these two soil types within our case-study area and encompassed the fields sampled (Fig. 3.2A). The grid size was 65 by 55 km. The furthest distance from a field to a grid node was 68 km and the closest distance 4.5 km. The total number of samples was 159.

**ii. Stratified library** The NSI dataset was stratified by the two dominant soil

types found within the case-study area using the SoilScapes dataset. We extracted all of the NSI samples where the centroid of the cell they represent lay within either a ‘Fen peat soil’ or ‘Loamy and sandy soils with high groundwater and a peaty surface’ polygon (Fig. 3.2B). The total number of samples was 109.

### 3.2.2 Sample air-drying, sieving and milling

The soil samples for the field-scale dataset were air-dried for 7 days, aggregates were crushed in a roller mill and passed through a  $<2$  mm sieve. The samples were then placed in a stainless steel cup together with a disk and milled for 35 seconds at 960 rotations per minute in a TEMA Machinery Ltd mill to a powder ( $<100$   $\mu\text{m}$ ) (Northants, UK).

The soil samples of the NSI, which we used to make spectroscopy measurements, were stored as a powder ( $<100$   $\mu\text{m}$ ) in plastic sample bottles in the Rothamsted archive.

### 3.2.3 Wet chemistry analysis

The following laboratory methods were used for the samples from the field-scale dataset. Total carbon (C) (%) was determined by combustion (Dumas method) using an elemental analyser by LECO (TruMac Combustion Analyser, Michigan USA). Total C was assumed to represent total organic C as these soil types are unlikely to contain substantial amounts of carbonates. Available potassium (K) ( $\text{mg kg}^{-1}$ ) was determined by ammonium nitrate extraction and Inductively Coupled Plasma-Optical Emission Spectrometer (ICP-OES) (MAFF, 1986). Available phosphorus (P) ( $\text{mg kg}^{-1}$ ) was measured by the standard Olsen method (Olsen et al., 1954). The pH was measured in a 1:2.5 ratio of  $\text{H}_2\text{O}$ . Particle-size fractions (%) were determined by laser diffraction (Breure et al., 2021a).

For the NSI the values held in LandIS for the first sampling of the NSI were used. The laboratory methods for the soil properties of interest were measured as follows: Organic carbon (%) by loss-on-ignition for soils that were estimated to contain more than 20% organic carbon (Avery & Bascomb, 1982), otherwise by dichromate digestion (Kalembasa & Jenkinson, 1973). Extraction methods for extractable potassium (K) and phosphorus (P) were standardized by their volume rather than their weight (MAFF, 1988). Extractable K ( $\text{mg L}^{-1}$  of soil) was determined from a filtered ammonium nitrate extract with flame photometry (MAFF, 1986). Extractable P ( $\text{mg L}^{-1}$ ) by the standard Olsen method. Soil pH was measured in a 1:2.5 ratio of  $\text{H}_2\text{O}$ . Clay content ( $\% < 2$   $\mu\text{m}$ ) was measured using the pipette method on  $< 2$  mm mineral (peroxide-treated) soil (for further details see McGrath & Loveland, 1992).

### 3.2.4 Spectroscopy

#### (V)NIR measurements

Whereas spectroscopy measurements for the NSI dataset were taken only on milled samples (as this is the condition of the available stored samples in the NSI), the spectroscopy measurements for the field-scale dataset were taken on in-situ,

unprocessed, air-dried and milled samples (Fig. 3.1). The VNIR spectra from in-situ, unprocessed, and air-dried samples were taken using an ASD FieldSpec 4 spectrometer (Malvern Panalytical Inc., Westborough USA) in the range of 350–2500 nm with a resolution of 3 nm at 700 nm and 10 nm at 1400- and 2100 nm. In-situ measurements were taken with the ASD contact probe after we removed the rubber o-ring and placed a Prolene Thin Film (Chemplex Industries Inc., Florida USA) across the glass sampling interface to avoid contamination. In-situ measurements were taken where the top-soil appeared dry and we placed the ASD contact probe on the soil surface, ensuring good contact, without plant residues or stones to take spectral measurements. In-situ measurements were taken at three different locations within the 0.5 m  $\times$  0.5 m quadrat used for soil sampling. The measurements on unprocessed samples were taken on the fresh bulked sample, before air-drying, sieving and milling.

These samples were placed within a petri-dish and measured with the ASD contact probe. The fresh, unprocessed, bulked samples did not show aggregation and were rather moldable due to their high volumetric water content (ranging from 20–45%), resulting in a relatively smooth surface once the sensor was placed on the sample due to compression. Replicates were taken at three different locations within the petri-dish. The measurements on the air-dried samples were performed on the bulked samples before sieving and milling. The bulked sample varied from aggregates that were approximately 5 cm in width to aggregates reduced to powdery soil, the stone content was negligible. A subsample was (re)poured in triplicate into a glass vial and measured with the ASD Muglight. The milled soil samples were pressed into a small well in replicates of three (6 mm across and approximately 1 mm deep) and placed in a Tensor II spectrometer (Bruker scientific, Ettlingen Germany) in the AfSIS spectral laboratory at Rothamsted Research. Its absorbance spectrum in the range 9997–3999  $\text{cm}^{-1}$  (1000–2500 nm), i.e. the near infrared (NIR), was measured with a resolution of 4  $\text{cm}^{-1}$  (1 nm). The reflectance,  $R$  of the ASD FieldSpec4 was transformed to optical density (i.e. absorbance,  $A$ ) as  $A = \log_{10}(1/R)$  to align with the Tensor II measurements. All triplicate measurements were averaged.

The samples from the NSI database were also measured using the Tensor II instrument, spectroscopy measurements were taken on two aliquots of the sample and were averaged.

### **MIR measurements**

We took in-situ, unprocessed and air-dried MIR measurements using the Agilent 4300 FTIR spectrometer (Agilent Technologies, Santa Clara USA) in the range of 4000–650  $\text{cm}^{-1}$  (2500–15 385 nm) with a resolution of 4–16  $\text{cm}^{-1}$  (15–62 nm). The in-situ MIR measurements were taken at the same locations as the in-situ VNIR measurements. Equally to the (V)NIR measurements, the MIR measurements for the unprocessed and air-dried samples were taken within a petri-dish and replicate measurements were taken at three different locations. For the milled soil samples, each sub-sample’s mid infrared (MIR) spectrum in the range 4000–600  $\text{cm}^{-1}$  (2500–16 666 nm) was recorded on the Tensor II with a resolution of 4  $\text{cm}^{-1}$  (16.6 nm). The same well plates with soil aliquots prepared for NIR measurements were used by switching the light source on the Tensor II to MIR. The procedure was repeated for the measurements on the NSI samples.

### Spectral pre-processing

All spectra were smoothed to remove noise using the Savitzky–Golay filter (Savitzky & Golay, 1964) with a third-order polynomial in a moving window of 11. Subsequently, all spectra have been transformed to their standard normal variate and were subject to 1<sup>st</sup> order derivatization. The derivatives for the (V)NIR spectra were computed with a filter length of 11 (i.e. the spacing between points over which the derivative is computed), a segment size of 31 (i.e. the range over which the points are averaged). Subsequently, two column regions in the (V)NIR spectra were removed as these correspond to moisture absorption bands: (7900–6849 cm<sup>-1</sup>) and (5587–5102 cm<sup>-1</sup>), respectively (Bowers & Hanks, 1965). For the MIR spectra, we used a filter length of 11 and a segment size of 8. The atmospheric CO<sub>2</sub> bands were removed in the region 2430–2260 cm<sup>-1</sup> for the MIR spectra (Sandford & Allamandola, 1990).

#### 3.2.5 Data-subsetting and selection of spiking subset

The field-scale dataset was split into a dataset for calibration (75%) and validation (25%). We followed the standard procedure where samples are selected to span the range of soil variation anticipated across the samples. This was done using the Kennard-Stone sampling algorithm on the euclidean distance of the 1<sup>st</sup> derivative (V)NIR/MIR spectra from milled samples, to select a subset of 75% that represented the field-scale dataset spectrally (Kennard & Stone, 1969).

The spiking methodology comprised two main steps. Firstly, we took 10% of our milled calibration field-scale dataset as a spiking subset ( $n = 7$ ). Again these were chosen by Kennard-Stone sampling. Second, we applied additional weighting to the spiking subset when we regressed the spectra to laboratory reference values. Weighting was applied by adding the spiking subset  $m$  times, where  $m$  was the ratio between the size of the spectral library and the spiking subset (Table 3.2) (Guerrero et al., 2014).

#### 3.2.6 Partial least squares regression and model validation

Partial least squares methods were used to regress the absorbance measurements against the wet chemistry reference values. The partial least squares (kernel) algorithm selects orthogonal components that maximize the covariance between the predictor (spectral matrix) and the response (wet chemistry data). We performed a leave-one-out cross validation with the calibration dataset to gain the root mean squared error (RMSE). To avoid over fitting we allowed our models to have a maximum of fifteen components. The number of components retained was equal to the model that gave the lowest RMSE in the cross-validation. For more robust comparison across literature studies, we additionally include the ratio of performance to inter-quartile range (RPIQ). This method provides a standardized metric using the inter-quartile range of the observed data and is recommended by Bellon-Maurel et al. (2010) as suitable for IR spectroscopy predictions on skewed response variables. It is described by:

$$\text{RPIQ} = \frac{q_{y_i}(3) - q_{y_i}(1)}{\sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{N}}} \quad (3.1)$$



where  $y$  and  $\hat{y}$  are the observed and predicted data for the  $i^{th}$  observation,  $N$  the total number of samples,  $q_{y_i}(3)$  the 3<sup>rd</sup>- and  $q_{y_i}(1)$  the 1<sup>st</sup>-quantile of the observed data. We further computed the prediction bias as the mean of  $(y - \hat{y})$ .

The predictions from different sample conditions and those from spectral libraries were then evaluated on the field-scale validation set. PLSR residuals of the spectral libraries were evaluated for each individual model to assess for the presence of outliers in the spectral library due to subsetting of the NSI by stratification or region. After evaluating the PLSR standardized residuals from the spectral libraries (both with and without spiking) for each soil property, we removed data points that we considered to be outliers. Cut-off values of  $-3$  and  $3$  for the standardized residuals were used to remove observations.

### 3.2.7 Model-averaging of PLSR predictions and their evaluation

Combining predictions from multiple sensors can lead to better accuracy. The PLSR predictions from (V)NIR, MIR 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 OLS regression in its general form is:

$$\mathbf{Y} = w_0 + w_1\mathbf{z}_1 + \dots + w_k\mathbf{z}_k, \quad (3.2)$$

where  $\mathbf{Y}$  is a vector of random observed values,  $\mathbf{z}$  is a vector of PLS predictions, the  $w_i$ ,  $i = 1, 2, \dots, k$ , are weighting coefficients of the  $k$  individual predictors included in the regression. This equation was solved for the intercept ( $w_0$ ) and the  $k$  coefficients for each of the spectral matrix combinations ( $\mathbf{z}$ ). The intercepts correct for bias if one of the individual predictors is biased.

To evaluate predictions from in-situ, unprocessed, air-dried and milled samples, we compared models from (V)NIR, MIR and their model average (V)NIRMIR. Since model-averaging gave consistent equal- or improved predictions from milled samples, we evaluated the spectral library predictions on the model averaged (V)NIRMIR predictions only.

Analysis was done using the following R packages: spectral processing using **prospectr** (Stevens & Ramirez-Lopez, 2013) and partial least squares regression using **pls** (Bjørn-Helge et al., 2019). Kennard-Stone sampling using **resemble** (Ramirez-Lopez & Stevens, 2016), Granger–Ramanathan averaging using **GeomComb** (Weiss & Roetzer, 2016). Graphics were created with **ggplot2** (Wickham, 2016) and maps using QGIS3 (QGIS development team, 2019).

## 3.3 Results & discussion

### 3.3.1 The selection of representative samples for calibration

Soil spectroscopy is applied under the assumption that the calibration dataset is representative of the target population. It is therefore important that the calibration set spans the range of wet chemistry values in the validation dataset. This was the case for all of the soil properties we considered except for available P

and K, which had a slightly lower minimum in the validation data compared to the calibration data (Table 3.1).

Table 3.1: Summary statistics of the field-scale dataset that was used to regress laboratory reference values to soil spectra with partial least squares methods. Kennard-stone sampling was performed on the combined (V)NIRMIR spectra to select 75% of the samples for calibration and 25% for validation (see method section 2.5)

Dataset	Property	$n$	Mean	Median	Std dev.	Min	Max	Range	Skew
Field-scale	Organic C/g kg <sup>-1</sup>		12.99	12.94	3.44	6.26	20.41	14.15	-0.02
	pH		7.31	7.48	0.46	5.37	7.77	2.40	-2.29
	Clay/%	97	35.80	36.57	4.69	22.70	44.63	21.94	-0.38
	P/mg kg <sup>-1</sup>		46.17	44.65	12.28	25.32	86.70	61.38	0.72
	K/mg kg <sup>-1</sup>		283.34	278.93	120.41	86.65	705.22	618.57	0.86
Calibration	Organic C/g kg <sup>-1</sup>		13.06	13.15	3.43	6.26	20.41	14.15	-0.04
	pH		7.32	7.49	0.46	5.37	7.74	2.37	-2.41
	Clay/%	73	35.59	36.19	4.74	22.70	44.63	21.94	-0.46
	P/mg kg <sup>-1</sup>		47.42	45.83	13.15	27.12	86.70	59.58	0.68
	K/mg kg <sup>-1</sup>		289.46	274.93	125.87	91.97	705.22	613.25	0.97
Validation	Organic C/g kg <sup>-1</sup>		12.79	12.45	3.54	6.58	18.08	11.50	0.05
	pH		7.27	7.45	0.47	5.68	7.77	2.09	-1.84
	Clay/%	24	36.46	37.26	4.55	30.14	44.38	14.24	-0.07
	P/mg kg <sup>-1</sup>		42.34	44.22	8.25	25.32	54.20	28.88	-0.77
	K/mg kg <sup>-1</sup>		284.94	298.82	104.40	86.65	513.25	426.60	0.14

The spectral libraries subsetting from the NSI captured the range of wet chemistry data in the field-scale dataset (Table 3.1 and Table 3.2). However, for all soil properties the distribution differed between the field-scale dataset and the spectral libraries. The spiking subset selected by the Kennard-Stone algorithm encompassed the complete range of the calibration dataset for organic carbon only. A comparable, but incomplete, range was selected for clay, available P and K (Table 3.1 and Table 3.2). The pH distribution was not well captured in the spiking

subset, with a range of 0.81 in the spiking subset compared to the range of 2.37 in the calibration dataset.

Table 3.2: Summary statistics of the spiking subset and the spectral libraries used to regress laboratory reference values to soil spectra with partial least squares methods

Dataset	Property	$n$	$m$	Mean	Median	Std dev.	Min	Max	Range	Skew
Spiking	Organic C/g kg <sup>-1</sup>			13.39	15.29	5.07	6.26	20.41	14.15	-0.15
	pH			7.41	7.51	0.27	6.82	7.63	0.81	-1.32
	Clay/%	7	-	33.90	36.66	6.90	22.70	40.04	17.34	-0.63
	P/mg kg <sup>-1</sup>			49.27	42.30	20.38	28.35	86.70	58.35	-1.11
	K/mg kg <sup>-1</sup>			295.82	262.77	190.39	115.41	690.04	574.63	1.08
Regional library	Organic C/g kg <sup>-1</sup>	158	22	5.83	2.80	7.24	0.70	56.40	55.70	3.28
	pH	158	22	7.20	7.50	0.74	4.60	8.20	3.60	-1.34
	Clay/%	127	18	30.76	28.70	14.59	3.00	73.20	70.20	0.36
	P/mg kg <sup>-1</sup>	158	22	37.12	30.50	26.74	4.00	162.00	158.00	1.94
	K/mg kg <sup>-1</sup>	158	22	335.04	305.00	261.13	28.00	2776.00	2748.00	5.23
Stratified library	Organic C/g kg <sup>-1</sup>	109	15	12.19	9.20	10.29	0.70	56.40	55.70	1.59
	pH	108	15	6.53	6.90	1.12	3.60	8.00	4.40	-0.74
	Clay/%	62	8	29.28	26.50	16.26	3.00	73.20	70.20	0.36
	P/mg kg <sup>-1</sup>	108	15	31.91	29.00	20.91	2.00	120.00	118.00	1.33
	K/mg kg <sup>-1</sup>	108	15	249.09	205.00	182.47	21.00	1066.00	1045.00	1.46

### 3.3.2 The effect of sample processing on soil property prediction accuracy using (V)NIR, MIR and (V)NIRMIR

As expected, the effort of sample processing and homogenisation, i.e. air-drying and milling, led to the best predictions in all soil properties. The RPIQ values for organic carbon, clay and pH predictions from milled samples compare favourably to existing literature (Figs. 3.3–3.5). For example, several studies list RPIQ values that range from: 2.49–3.6 for organic carbon, 1.55–2.25 for pH and 3.88–6.4 for clay (Nocita et al., 2014; Terra et al., 2015; Clairotte et al., 2016; Hermansen et al., 2016; O’Rourke et al., 2016). We note that most of the studies listed predicted soil properties at

a different geographical scale, hence comparison needs to be viewed with caution. Although the RPIQ metric allows for better comparison, variances are dependent on the concentration of the target variable which in turn depends on geographical extent and soil variation present. Available P and K could be moderately predicted only under milled sample conditions (Fig. 3.6 and Fig. 3.7).

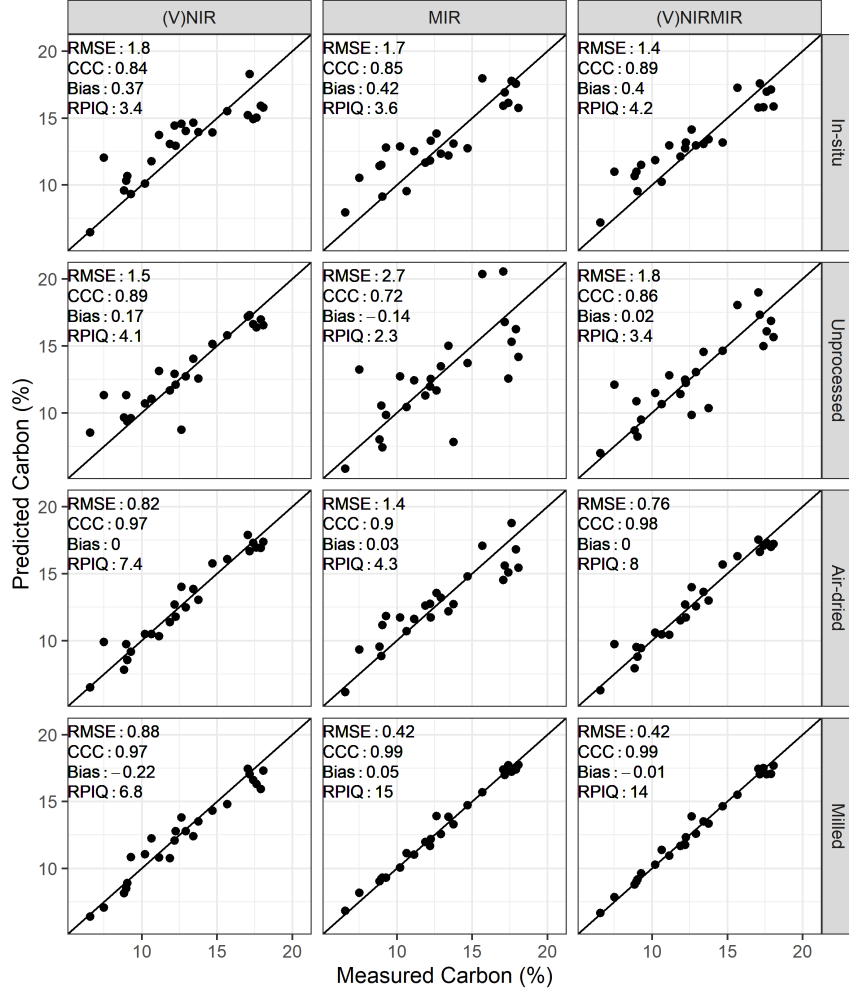


Figure 3.3: Predicted versus measured organic carbon by partial least squares regression under different soil conditions (in-situ, unprocessed, air-dried and milled) for (V)NIR, MIR, and (V)NIRMIR. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range. Prediction models for the top three rows are based on spectra made by handheld spectrometers whereas the models in the bottom row are based on benchtop spectrometer data (details in method section 2.4).

### Contrasting (V)NIR, MIR and (V)NIRMIR predictions

Model-averaged (V)NIRMIR led in most cases to either equal or consistent, albeit small, improvement in predictions compared to (V)NIR and MIR predictions by themselves (Figs. 3.3–3.7). The limited benefit of (V)NIRMIR compared to MIR for milled samples has been previously observed by Clairotte et al. (2016) in their study on soil organic carbon. Our results indicated that model-averaging did

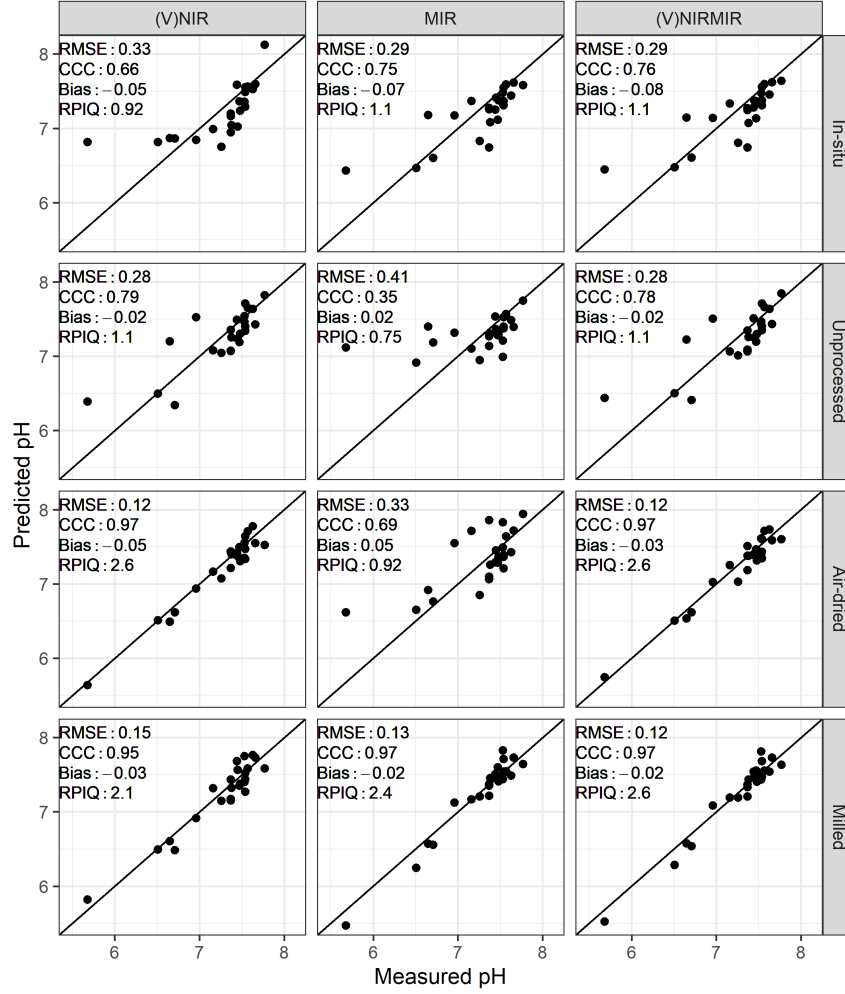


Figure 3.4: Predicted versus measured clay by partial least squares regression under different soil conditions (in-situ, unprocessed, air-dried and milled) for (V)NIR, MIR, and (V)NIRMIR. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range. Prediction models for the top three rows are based on spectra made by handheld spectrometers whereas the models in the bottom row are based on benchtop spectrometer data (details in method section 2.4).

improve accuracy for organic carbon predictions from in-situ and air-dried samples in particular (Fig. 3.3). Granger-Ramanathan averaging ensured in most cases that the (V)NIRMIR predictions were more accurate or comparable compared to the best predictions of either (V)NIR or MIR, this has also been shown in other studies (O'Rourke et al., 2016). We further note that model-averaging improved prediction by reducing bias, demonstrated in our study for available K predictions from milled samples (Fig. 3.7).

For in-situ and milled sample conditions, predictions of organic carbon, pH and clay based on MIR measurements outperformed (V)NIR predictions. (V)NIR predictions of organic carbon, pH and clay outperformed those from MIR for unprocessed or air-dried conditions (Fig. 3.3, Fig. 3.4, Fig. 3.5). Over all four sample conditions, predictions from milled samples did not always guarantee the best accuracy. (V)NIR predictions from air-dried samples of pH were equal to the most accurate predictions from milled samples ( $RPIQ = 2.6$  for both). Clay, pH

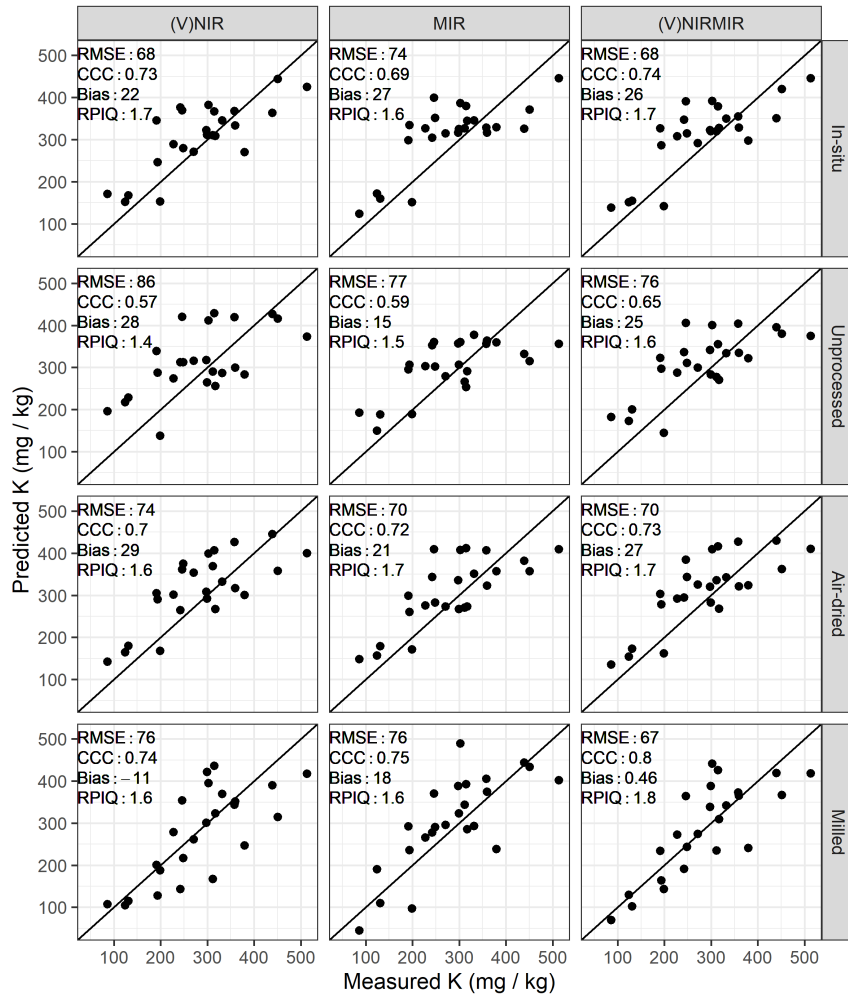


Figure 3.5: Predicted versus measured pH by partial least squares regression under different soil conditions (in-situ, unprocessed, air-dried and milled) for (V)NIR, MIR, and (V)NIRMIR. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range. Prediction models for the top three rows are based on spectra made by handheld spectrometers whereas the models in the bottom row are based on benchtop spectrometer data (details in method section 2.4).

and organic carbon predictions from air-dried samples outperformed those from milled samples for the (V)NIR range only (Figs. 3.3–3.5).

Nduwamungu et al. (2009) did not find improvements in predictions from <2 mm soil samples for the NIR range. Both Le Guillou et al., (2015) and Wijewardane et al. (2020) reported that predictions from fine ground samples always outperformed those from non-fine ground for the MIR region. Results in this study align with the literature. For the (V)NIR range, milling did not strictly show improvement in predictions. However, milling always led to the most accurate predictions from MIR spectra.

### Spectrometer differences and sample heterogeneity

Observed differences in prediction accuracy cannot be solely attributed to sample conditions because the spectra from handheld spectrometers are not directly

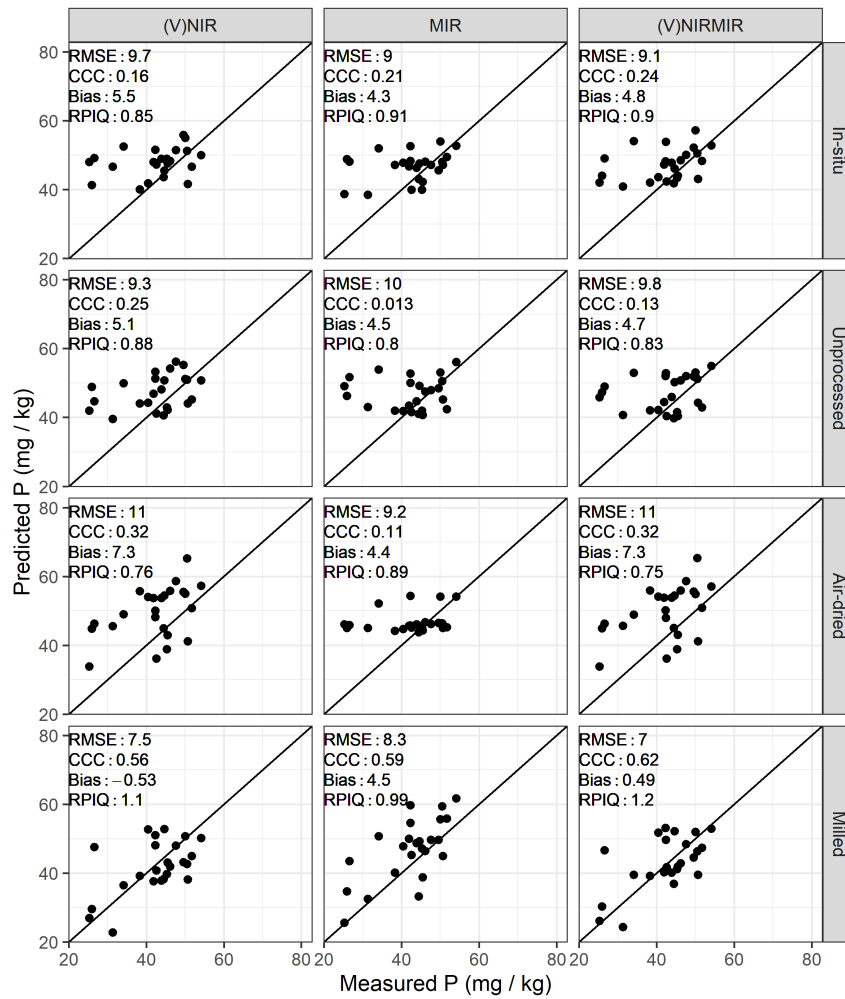


Figure 3.6: Predicted versus measured available P by partial least squares regression under different soil conditions (in-situ, unprocessed, air-dried and milled) for (V)NIR, MIR, and (V)NIRMIR. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range. Prediction models for the top three rows are based on spectra made by handheld spectrometers whereas the models in the bottom row are based on benchtop spectrometer data (details in method section 2.4).

comparable to those measured by benchtop spectrometers.

MIR predictions from in-situ, unprocessed and air-dried samples underperformed compared to MIR predictions from milled samples. This can be explained in part by the small support size of the FTIR 4300 sampling interface (2–3 mm), which results in problems to scan a representative area of the soil sample (Reeves 2010). Ji et al. (2016) found that small-scale soil heterogeneity and electrical noise contributed up to 50% of the total prediction error of soil properties in their in-situ MIR study. Hutengs et al. (2018) found that the MIR portable spectrometer used in this study measured spectra with the same accuracy as a DRIFT benchtop spectrometers (equivalent to the Tensor II in this study) for milled samples, particularly when replicate measurements with the handheld instrument were taken at different locations to account for the small support size of the sampling interface.

For the (V)NIR spectra, predictions of organic carbon and pH from air-dried samples outperformed those from milled samples. Several studies reported that

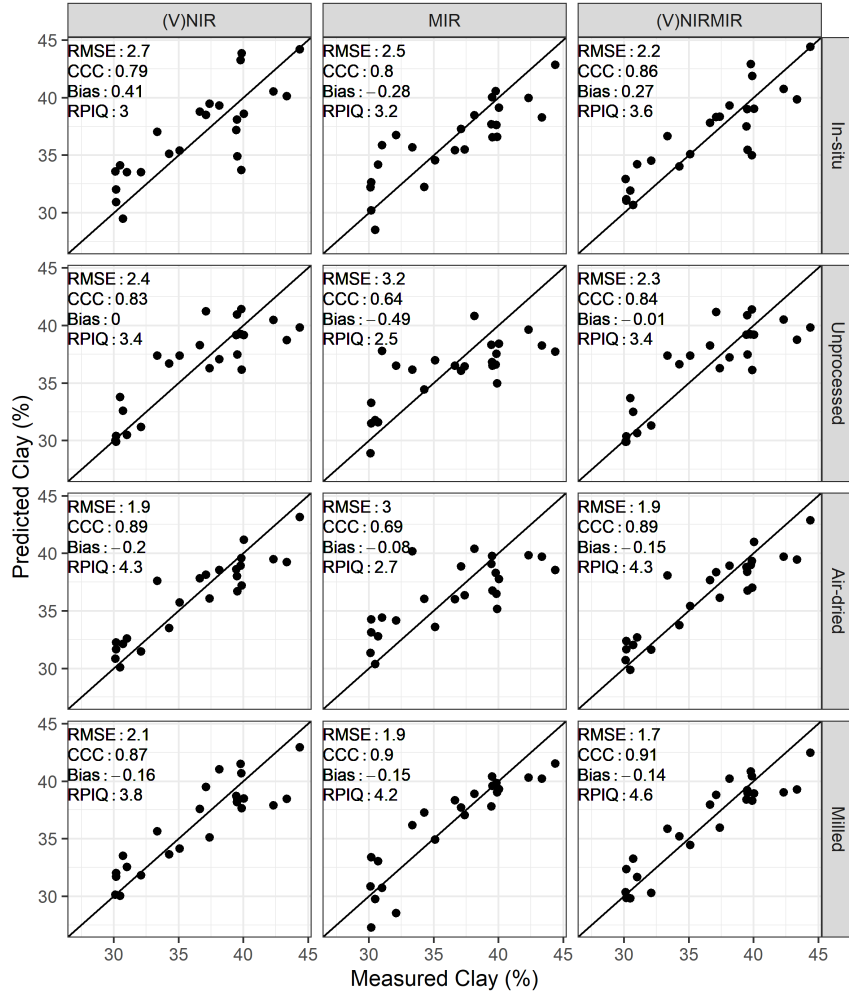


Figure 3.7: Predicted versus measured available K by partial least squares regression under different soil conditions (in-situ, unprocessed, air-dried and milled) for (V)NIR, MIR, and (V)NIRMIR. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range. Prediction models for the top three rows are based on spectra made by handheld spectrometers whereas the models in the bottom row are based on benchtop spectrometer data (details in method section 2.4).

(V)NIR predictions within the field outperformed those under lab conditions (Stevens et al., 2008; Gras et al., 2014). Stevens et al. (2008) explain their results are due to the dryness of the soil, soil roughness and vegetation cover associated with their in-situ measurements. Spectral processing to mitigate confounding effects is also mentioned as a potential contributing factor to good predictions from in-situ reflectance measurements. Although the benchtop spectrometer used to collect milled (V)NIR spectra has a reduced wavelength range (excluding the VIS region), no large differences between benchtop and laboratory (V)NIR spectrometers have been reported when compared on the same sample conditions (Hodge & Sudduth 2012; Knadel et al., 2013; Lopo et al., 2016). However, the usefulness of the VIS region, particularly for organic matter predictions, has been pointed out in multiple studies (Fystro et al. 2002; Islam et al., 2003). Conversely, Chang et al. (2001) and Dunn et al. (2002) reported poor predictions for organic matter in their studies from VIS. The VIS region also relates to texture, structure,



moisture and mineralogy. It appears that the soil's brightness as a predictor for soil properties has an application within limited geological types/parent materials (Stenberg & Viscarra Rossel et al., 2010). One other factor that could contribute to increased prediction accuracy is that (V)NIR measurements from the handheld spectrometer tend to smooth effects of sample heterogeneity on the spectrum, since measurements were averaged over a larger surface area (10 mm spot size for the contact probe and 12 mm for the mug light) (ASD Accessories User Manual).

Although the predictions from unprocessed samples were the least accurate, the effect of soil moisture content was not as large as we expected based on the range of volumetric water content (20–45%). Soil moisture reduces total reflectance, particularly for the H<sub>2</sub>O bands, where the magnitude of this relation changes between different soil types (Bowers & Hanks, 1965; Minasny et al., 2009). Although this effect generally reduces robustness of a calibration, in our case-study the timing of sampling might have enhanced a distinction between spectra from the two soil types due to their difference in water holding capacity.

#### 3.3.3 The use of spectral libraries compared with field-scale calibration models

Across the variables considered, unsurprisingly the field-scale calibration model led to the best predictions (Fig. 3.8). Comparing our regional and stratified spectral libraries (without spiking), the regional library performed best for organic carbon whereas the stratified library performed best for pH. Organic carbon predictions from the unspiked regional library showed good precision (i.e. they captured the range) but poorer accuracy, i.e. large RMSE and bias (Fig. 3.8). Predictions for pH from the unspiked stratified library showed moderate precision and accuracy (Unspiked stratified: RPIQ = 1.2, Bias = 0.02). Clay, available P and K showed poor results for spectral libraries without spiking (Fig. 3.8).

Our poor clay predictions contrast with the literature, for example, Waiser et al. (2007) state that kaolinite, smectite and muscovite minerals have distinct spectral features which leads to accurate predictions. In our study, the poor clay predictions were likely due to different laboratory methods and the high organic matter content of the soil: samples from the National Soil Inventory were pre-treated with H<sub>2</sub>O<sub>2</sub> and measured by the pipette method. The field-scale dataset samples were not pre-treated due to their high organic carbon content and were measured by laser diffraction. The confounding effect of different laboratory methods on prediction performance of spectral libraries has been mentioned within the literature previously (Viscarra Rossel et al., 2016). Differences between laboratory methods also occurred for available P and K, in particular the fact that the NSI extractions were standardised by volume of soil rather than their weight (McGrath & Loveland, 1992). Given the organic nature and therefore low bulk density of the two soil types in the case study area, the standardisation by volume will have affected the comparison to the field-scale dataset (standardised by weight). Additionally, available P and K are known to have weak or no spectral features in the IR region (Kuang et al., 2012).

Spectral library predictions of organic carbon and clay improved substantially once spiked (Fig. 3.8). Clay predictions were still poor from the spiked regional library compared to the field-scale dataset with a large bias (Spiked regional: Bias

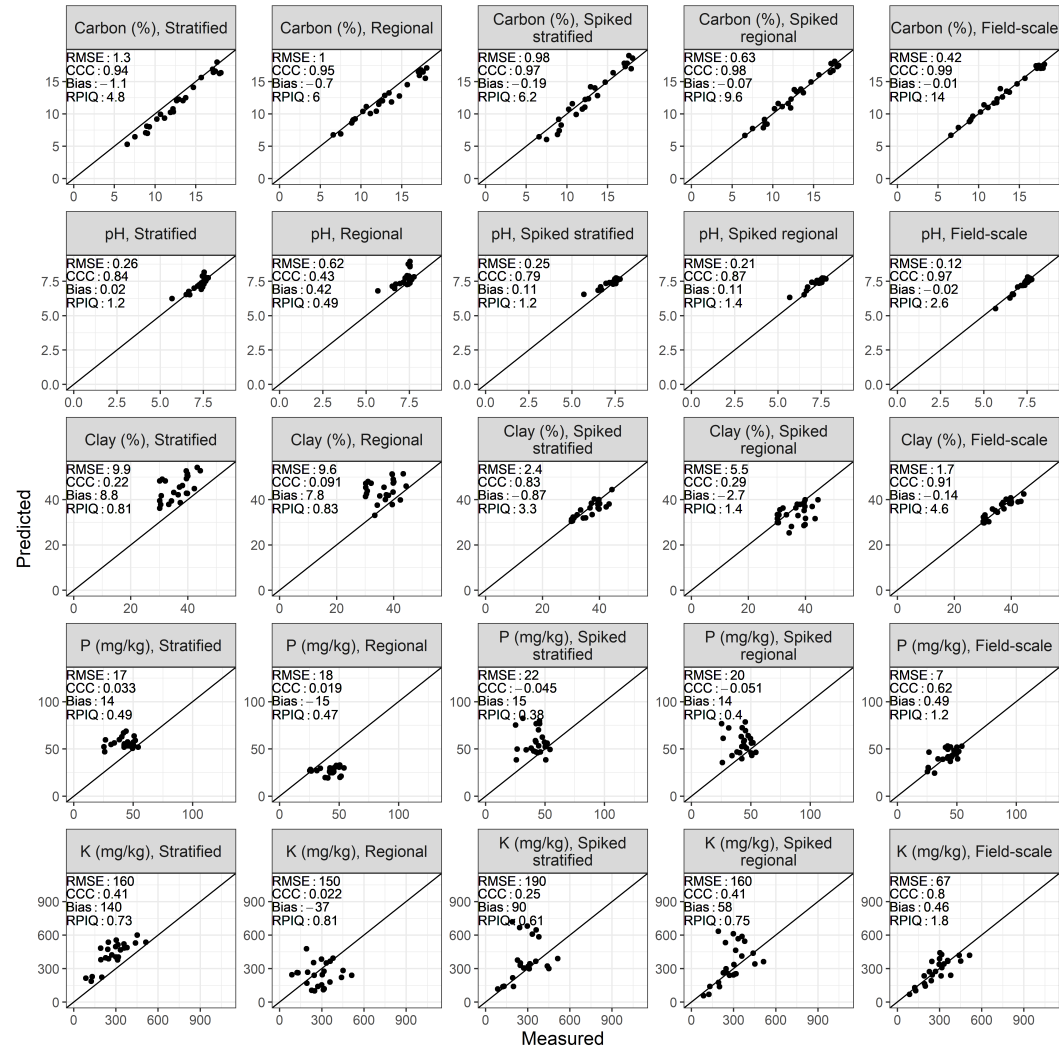


Figure 3.8: Predicted versus measured organic carbon, pH, clay, available P and K from the spectral libraries and spiked spectral libraries and the field-scale dataset. Models and predictions were performed using milled samples and by model-averaging predictions from (V)NIR and MIR spectra. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range.

= 2.7, Field-scale: Bias = 0.02). Spiking only improved predictions for pH from the regional library (Unspiked regional : RPIQ = 0.49, spiked regional: RPIQ = 1.4). Once spiked, organic carbon and pH predictions from the regional library outperformed the stratified library, suggesting that geographical representation, rather than soil type in the spectral library is more representative of the relation between these properties and spectral reflectance. This potentially reflects a regional soil signature caused by a specific land use or management in the area (the case-study area is used for outdoor horticulture). Clay predictions from the spiked stratified library outperformed those from the spiked regional library. Clay is unaffected by local management and more closely related to soil lithology and parent material, which could explain the better representation of the stratified library for this property.

### 3.3.4 Contrasting time and cost implications of spectroscopy predictions from spectral libraries and samples with reduced processing

The decision whether to reduce sample processing or use spectral libraries depends on the soil property of interest. In our analysis we found that neither reduction in effort would allow accurate prediction of P or K but both showed promise for predictions of organic carbon, clay and pH. For these variables, the choice of which approach to use in practice will depend on the accuracy required, the number of prediction samples needed and the costs associated with field sampling, preparation and handling and laboratory costs. This will be case study specific, but here we place the relative differences in uncertainty in the context of the data acquisition process. For example, our results showed that RMSEs for organic carbon from spiked spectral libraries (RMSE = 0.63–0.98) were lower compared to the lowest RMSEs under in-situ (RMSE = 1.4) and unprocessed (RMSE = 1.5) sample conditions. The lowest RMSE for organic carbon predictions from air-dried samples (RMSE = 0.76) lay in between the spiked stratified and regional library predictions. However, the use of spiked spectral libraries still requires sampling a field-scale dataset where samples need to be air-dried, sieved and milled so they are comparable to the samples from the library.

Prediction accuracy under in-situ, unprocessed and air-dried conditions was good but calibration samples had to be analysed with wet chemistry data compared to the spectral library approach, where wet chemistry data was already available. In some situations, the cost of a greater number of samples to be processed and analysed by wet chemistry could be offset by reduced hours spent on field sampling (in-situ) or handling of the samples (unprocessed and air-dried) (Fig. 3.1). For example, for (V)NIR predictions only, there was no loss in accuracy for organic carbon, pH and clay predictions from air-dried samples. Hence, the benefits of milling became redundant. Similarly there is a trade-off between the two approaches in terms of laboratory, sampling and handling costs occurs for clay predictions. In our study, the best clay predictions under in-situ and unprocessed sample conditions were roughly equal to those from the stratified spiked spectral library (in-situ : RPIQ = 3.6, unprocessed : RPIQ = 3.4, spiked stratified : RPIQ = 3.3). Clay predictions from air-dried conditions approximated those of the milled field-scale dataset (air-dried : RPIQ = 4.3, milled : RPIQ = 4.6).

It should of course be considered whether the additional loss in accuracy affects the value of the final dataset created from soil property predictions. For example, predictions with reduced accuracy can be of practical use depending on the available budget and purpose of the analysis. An error of 0.12 units of pH (predictions from air-dried or milled samples) in determining liming requirements for an agricultural field could lead to an erroneous under or over application of 1.5 t limestone per ha<sup>-1</sup>. Whether variable rate liming under this condition is cost-effective compared with a field average will depend on specific circumstances of the subfield variation and the price of limestone. Equally this question can be asked for predictions from different sample conditions or spectral libraries that showed a higher error variance.

## 3.4 Conclusions

This study contrasted the magnitude of loss in accuracy, relative to field-scale predictions on milled samples, by either reduced sample processing or the use of spectral libraries. The results show that there is potential to reduce time and cost of using near- and mid-infrared spectra to predict soil organic carbon, clay and pH. We found that reduced sample processing lowered the ratio of performance to inter-quartile range (RPIQ) by 0% to 76%. The use of spectral libraries reduced RPIQ of predictions by 54% to 82% and was reduced in the range of 29% to 70% for predictions when spectral libraries were spiked. The reduction in uncertainty was specific to the combination of soil property and sensor analysed. We conclude that the decision about which approach to use will depend on the case-study in question because implications of cost and uncertainty will vary from case to case. This study provides insight into the expected differences in prediction accuracy and which factors need to be taken into consideration to reduce effort for developing field-scale calibrations with near- and mid-infrared soil spectra.

## Chapter 4

A loss function to evaluate  
agricultural decision-making under  
uncertainty



# Abstract

*The excessive use of fertiliser in food production systems has caused substantial negative environmental impacts. To minimise these impacts, the application of fertiliser needs to be more precise, which requires detailed information about soil variation. Modern sensor technologies can offer this information at significantly lower costs than wet-chemistry analysis. However, in order to incentivise farmers to adopt these technologies, we must acknowledge and account for uncertainties in predicted soil-nutrient content to make this approach risk neutral or positive to yield and income. Here we present a framework that accounts for the uncertainty and determines the cost-benefit of soil-nutrient data from sensors to economically rationalise the precise application of fertilisers. For four fields, we determined the uncertainty associated with variation in soil P and K predicted from sensors. Using published fertiliser dose–yield response curves we then quantified the effect of estimation errors from sensor data on expected financial losses. 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 lead to greater losses than the losses from overestimation. We demonstrate that substantial financial gains can be obtained from sensor based precise application of P and K fertiliser, with savings for our fields of up to £76 ha<sup>-1</sup> for P and up to £73 ha<sup>-1</sup> for K, with concurrent environmental benefits due to a reduction of 13–25 kg ha<sup>-1</sup> applied P fertiliser when compared with uniform application. Our results showed that the framework can be used to account for uncertainty in sensed soil data, so providing the most cost effective and environmentally sustainable soil and fertiliser management approach for farmers.*

Based on:

Breure, T.S., Haefele, S.M., Hannam, J.A., Corstanje, R., Webster, R., Moreno-Rojas, S., Milne, A.E. 2021. A loss function to evaluate agricultural decision-making under uncertainty: a case study of soil spectroscopy. Precision Agriculture – *Under review*

## 4.1 Introduction

Annually, an estimated global total of 9–14 million tonnes of phosphorus (P) leaches into watercourses (Beusen et al., 2016; Chen et al., 2016). This excessive use of fertiliser in food production systems has caused substantial negative environmental impacts. To minimise these impacts, the application of fertiliser needs to be more precise, varying across the field according to crop requirements and soil supply. This aligns with the ambition 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 and so it requires intense information (Blackmore et al., 2003). Farmers 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 what is 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, farmers need to have detailed maps of nutrient status or fertiliser requirement for their fields.

Obtaining detailed soil information to support VRA requires intense sampling. A general consensus is that around 10 observations per hectare are required, though McBratney et al. (1996) suggested that sampling at 10-m intervals might be necessary to map nutrient status for such management. 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. 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 numerous 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 measurements embody some degree of error, however, and so soil data have associated uncertainty. Nevertheless we have considered the errors in standard chemical analyses by modern equipment are small enough to be ignored. Spectroscopic estimation introduces yet another source of error. That is, calibration equations which describe the relationship between the wet chemistry measures and the soil spectra will 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 (Goovaerts, 2001; Cherry et al. 2008; Heuvelink, 2018).

Given the ease with which spectroscopy can replace conventional chemical methods for analysing soil and the potential financial savings, it is important to know how the errors it introduces affect the final spatial predictions. This is because errors carry with them costs. Over-estimation of a plant nutrient concentration in the soil, say P, would lead to a farmer's applying too little fertiliser and to loss of potential yield and income. Under-estimation of the concentration would lead to the farmer's over-fertilizing, spending unnecessarily on fertiliser, to the point of spending more than earned in increased yield of crop. If the excess fertiliser finds its way into water bodies and causes pollution there is a further cost to water companies, which has been estimated to range between 75 and 114 million pounds sterling per year for England and Wales (Pretty et al., 2003). Although P is chemically adsorbed by soil particles, it can leach to the environment after long and excess application (Carpenter et al., 1998; Smil, 2000). Algal growth is limited by the N:P ratio, and loss of P or N from fields into water courses accelerates eutrophic algal growth (Redfield, 1958). 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).

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 mining (Journel, 1984). Because overestimation and underestimation incur losses for different reasons the loss function may be asymmetrical. Given a loss function and an error distribution for the information, one may make a decision that minimises expected losses (e.g. Journel, 1984; Goovaerts, 1997).

Here we are concerned with the prediction of available P and potassium (K) for precise management of fertiliser in horticultural crops. We set out to discover 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. For this we considered soil samples from four fields in the Cambridgeshire Fens of the UK. Concentrations of available P and K estimated from soil spectra were used to predict how concentrations vary across the fields and to compute the associated error variances from interpolation. We determine the expected losses associated with varying applications of fertiliser given the error variance of our predictions. The expected loss is compared with the losses should we have used our estimates from wet chemistry to determine a single application rate per field. From this we draw some general conclusions about the effect of uncertainty in our soil nutrient status on economic and environmental losses and draw some practical considerations for implementing the loss function.



## 4.2 Methods

### 4.2.1 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 17th century. The land surface is now 1 to 2 metres 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. We used the LiDAR raster (2 m  $\times$  2 m resolution) from the British Environment Agency as a basis for our 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 & Lisic, 2019) on the spatial coordinates and LiDAR. At each location of these sub-samples, we added another sampling point 6 m away at a random orientation 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. S4) shows the field boundaries with the sampling points. Three soil cores of topsoil (0–25 cm) were taken within a 0.5 m  $\times$  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 ( $\text{NH}_4\text{NO}_3$ ) 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  $\text{cm}^{-1}$  (1000–2500 nm), 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  $\text{cm}^{-1}$  (2500–16 666 nm) was recorded on the same instrument with a resolution of 2  $\text{cm}^{-1}$ . 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 an 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 seconds at both 10 and 40 keV. The pXRF was set up in an instrument stand,

and samples were 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 were 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.

### 4.2.2 Spectral processing and calibration

The raw spectra were pre-processed first by the Savitzky–Golay filter (Savitzky & Golay, 1964) and then transformed to standard normal variates and their first derivatives to give four sets of spectra, including the original raw spectra. The H<sub>2</sub>O bands (7900–5587 and 6849–5102 cm<sup>-1</sup>) were removed from the NIR spectra (Bowers & Hanks, 1965). The region of 4464–4115 cm<sup>-1</sup> was removed from the MIR spectra to account for the CO<sub>2</sub> peak at 4248 cm<sup>-1</sup> (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. We selected the number of components to be included in the model as follows. First, we computed the mean squared error (MSE) between the known values and the predictions by leave-one-out cross-validation (LOO-CV). The standard deviation of the LOO-CV residuals was also computed. To minimise over-fitting, we computed models for a maximum of 15 components. Subsequently, we chose the model 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, section 7.10).

Preliminary analysis showed that whilst XRF tends to give the most accurate predictions of P and K, this was not always the case, and so in this study we chose to combine all three sets of spectra. 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 OLS regression in its general form is

$$\mathbf{Y} = w_0 + w_1\mathbf{z}_1 + \cdots + w_k\mathbf{z}_k, \quad (4.1)$$

where  $\mathbf{Y}$  is a vector of random observed values,  $\mathbf{z}$  is a vector of PLS predictions, the  $w_i$ ,  $i = 1, 2, \dots, k$ , are weighting coefficients of the  $k$  individual predictors included in the regression. This equation was solved for the intercept ( $w_0$ ) and the  $k$  coefficients for each of the spectral matrix combinations ( $\mathbf{z}$ ). The intercepts correct for bias if one of the individual predictors is biased. Given the variation in prediction accuracy between different spectral processing methods, Granger–Ramanathan averaging was subsequently performed for a second time. The vector  $\mathbf{z}$  in Equation 4.1 now takes the OLS predictions from the four sets of infrared and XRF spectra (raw, Savitzky–Golay filtered, standard normal and first derivative).

Each prediction has associated with it an error, and we treated these errors as ones arising from the use of the regression model. We computed the prediction error by re-predicting on the calibration data. Effectively, this is equivalent to the residual

mean-squared deviations, which is biased and likely to be over-optimistic about the calibration model's performance. However, we decided to implement this approach for two reasons. First, we minimised over-fitting models by choosing ones with the smallest MSEs of the LOO-CV and taking into account the standard deviation of residuals of the LOO-CVs. Second, predictions from the LOO-CVs overestimate the MSEs in the presence of outliers. Since we use all the calibration data to predict all locations in the sampling design, we computed the MSEs using fresh prediction on the calibration set. The error variance of these predictions, of which there are  $n$ , is hereafter referred to as

$$\text{var}[\widehat{\varepsilon}_y] = \frac{1}{n} \sum_{i=1}^n (y_i - \widehat{y}_i)^2 ,$$

and it has been propagated through into the geostatistical model.

### 4.2.3 Geostatistical predictions

Following Lark et al. (2006) we model the spatial variation using the empirical best linear unbiased predictor (E-BLUP). This combines in an additive way fixed effects (e.g. the unknown mean, coefficients of a trend) and random effects (the spatially correlated random variation), thus:

$$Y(\mathbf{x}) = \sum_{j=0}^J \beta_j f_j(\mathbf{x}) + \varepsilon(\mathbf{x}) . \quad (4.2)$$

Here the  $\beta$  are unknown coefficients, the  $f_j(\mathbf{x})$  are typically first- or second-order polynomials, and the  $\varepsilon(\mathbf{x})$  represents the residuals from the trend, which are assumed to be second-order stationary random variables, jointly normally distributed with zero means and  $n \times n$  covariance matrix  $\mathbf{C}_d$  with variogram  $\gamma(\mathbf{h})$ :

$$\gamma(\mathbf{h}) = \frac{1}{2} \text{E}[\{\varepsilon(\mathbf{x}) - \varepsilon(\mathbf{x} + \mathbf{h})\}^2] , \quad (4.3)$$

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

$$\begin{aligned} \gamma(h) &= c_0 + c_1 \left\{ 1 - \exp\left(-\frac{h}{a}\right) \right\} \quad \text{for } 0 < h \\ &= 0 \quad \text{for } h = 0 . \end{aligned} \quad (4.4)$$

Here  $c_0$  and  $c_1$  are variances, respectively the nugget and sill of the correlated variance, and  $a$  is the distance parameter. 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 & 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, we first use the ML method to select the fixed effects for our model.

Preliminary investigations (visual inspection and marginal plots) suggested that the estimated soil properties,  $y$ , have long-range trends across the fields. We consider these as fixed effects. Estimated soil properties also vary systematically with elevation as recorded by LiDAR, so we consider this as another fixed effect. We added in each trend variable (fixed effect) in turn and tested whether its addition was significant by a log-likelihood ratio test. We took a chi-squared  $p$ -value from the log-likelihood ratio of 0.05 as threshold and treated any smaller value ( $p \leq 0.05$ ) as evidence that additional trend parameters should be included. After the selection of fixed effects, we re-estimated the fixed effects and random effects in Equation (4.2) using residual maximum likelihood (REML). If  $c_0$  was less than  $\text{var}[\hat{\varepsilon}_y]$  then it was added to the diagonal of the covariance matrix  $\mathbf{C}_d$  and Equation (4.2) was solved again. Next, we used the final variograms for universal kriging. This provided us with predictions and their kriging variances.

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

#### 4.2.4 Theory of the loss function

The loss function,  $L(F)$ , for a given application of fertiliser,  $F$  ( $\text{kg ha}^{-1}$ ), is defined as the difference in profit that results from applying a given amount of fertiliser  $F$  compared with the economic optimal amount  $F_0$ :

$$L(F) = \Phi(F_0) - \Phi(F) , \quad (4.5)$$

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

$$\Phi(F) = M \times \text{Yield} - V \times F , \quad (4.6)$$

where  $M$  is the price of the crop ( $\text{£ t}^{-1}$ ),  $V$  is the cost of the fertiliser ( $\text{£ kg}^{-1}$ ) and yield in  $\text{t ha}^{-1}$ .

We assume that the yield is given by the dose-response equation:

$$\text{Yield} = \alpha + \eta R^{\xi F + S} + \nu(\xi F + S) , \quad (4.7)$$

where  $S$  is the concentration of the available nutrient in the unfertilised soil,  $F$  is the applied fertiliser ( $\text{kg ha}^{-1}$ ),  $\xi$  is the increase in nutrient concentration ( $\text{mg kg}^{-1}$ ) in the soil for every  $1 \text{ kg ha}^{-1}$  fertiliser applied, and  $\alpha$ ,  $\eta$ ,  $\nu$  and  $R$  are parameters, then the optimum amount of fertiliser can be calculated from this and is given by

$$F_0 = \ln \left( \frac{B/\xi - \nu}{\eta R^S \ln R} \right) / \xi \ln R , \quad (4.8)$$

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

By definition, the loss given by Equation (4.5) 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 concentration,  $S$ , in the soil, and it is unlikely that one would know it precisely.

Given the error distribution,  $g(s)$ , of the nutrient concentration  $S$  we can compute optimum fertiliser rate that maximizes the expected profit

$$F_{\text{opt}} = \ln \left( \frac{B/\xi - \nu}{\eta \ln R \int_0^\infty R^S g(s) ds} \right) / \xi \ln R . \quad (4.9)$$

This also minimises the expected loss function,  $E[L(F)]$ , which we define here as the difference between the profit where  $S$  is known without error and associated optimum fertiliser,  $F_0$  in Equation (4.8), is applied

$$E[L(F)] = \Phi(F_0) - \int_0^\infty \{\Phi(F)\} g(s) ds . \quad (4.10)$$

#### 4.2.5 Parameterization and analysis of the loss function

All fields sampled were cultivated for lettuce, and so we computed loss functions 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). We assumed that for every 1 kg of P added in fertiliser 0.18 kg becomes available to the crop (Muhammed et al., 2017), and for every 1 kg of K added in fertiliser, 0.62 kg becomes available to the crop (Blake et al., 1999). Furthermore, we assumed that the added nutrients are contained in the top 25 cm of the soil (the sampling depth). We took from Milne et al. (2006) the value of 480 kg m<sup>-3</sup> for bulk density of this peat soil. Given the support of our 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<sup>-1</sup> and 0.52 mg available K kg<sup>-1</sup>, equal to  $\xi$  in the dose-response Equation (4.7). Greenwood et al. (1980) listed a mean base nutrient concentration of 69 mg available K kg<sup>-1</sup> for the unfertilised soil in their study, which was used as an additive component. We assumed a profit margin ( $M$ ) of £90 per tonne of lettuce per hectare. The prices of fertiliser ( $V$ ) were taken as £0.36 per kg P fertiliser and £0.29 per kg K fertiliser. Table 4.1 lists the parameter values of the dose-response equations for P and K.

Table 4.1: Dose-reponse equation parameters as relevant to Equation 4.7

Soil property	$\alpha$	$\eta$	$\nu$	R	$\xi$
P	142.15	-145.8	-0.776	0.98	0.15
K	63.3	-63.3	0	0.98	0.52

We assessed the profitability of variable-rate application (VRA) based on kriged maps by computing the total expected loss (Equation 4.10) across each field. We did so by comparing the expected loss for each field between that from VRA,  $E[L(F_{\text{opt}})]$ , and a uniform application (UA) based on wet chemistry alone,  $E[L(F_{\text{UA}})]$ .

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

## 4.3 Results

### 4.3.1 Uncertainty in kriging predictions from soil properties estimated by spectroscopy

The distribution of available K was strongly positively skewed, and the data were therefore transformed to logarithms to stabilize the variances. The correction as described by Laurent (1963) was used to estimate kriging predictions and their variances back to the original scale (Equation 1 in Suppl. material). The nugget variances,  $c_0$  in Equation (4.4), were underestimated for the following variogram models: available K (Fields 3 and 4) and available P (Fields 1 and 2). We therefore added  $\text{var}[\hat{\varepsilon}_y]$  to the diagonal of the covariance matrix  $\mathbf{C}_d$  and solved Equation (4.2) 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 4.2). 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. 4.1). The LOO-CV results of the mixed model variograms accorded overall with expectations for both P and K and all four fields (Suppl. Fig. S5).

### 4.3.2 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  $L(F) = \Phi(F_0) - \Phi(F)$  differently because of their asymmetry characteristics and by association the expected loss, i.e. loss from imperfect knowledge,  $E[L(F)]$  in Equation (4.10). For available P, the dose–response curve declines linearly in yield for large values of P (Fig. 4.2). That is to say over application of fertiliser results in financial losses due to excess fertiliser application and reduced yield. This makes effects of large application rates with large soil P fertiliser have a more dramatic impact than for K where the dose–response curve is characterised by an asymptote and so the associated prediction of profit and loss converge at large rates of applied fertiliser (Figs 4.3 and 4.4). For both P and K, larger uncertainty of the predicted soil nutrient content (high  $\sigma_k^2$ ) increases the expected loss and reduces the profit.

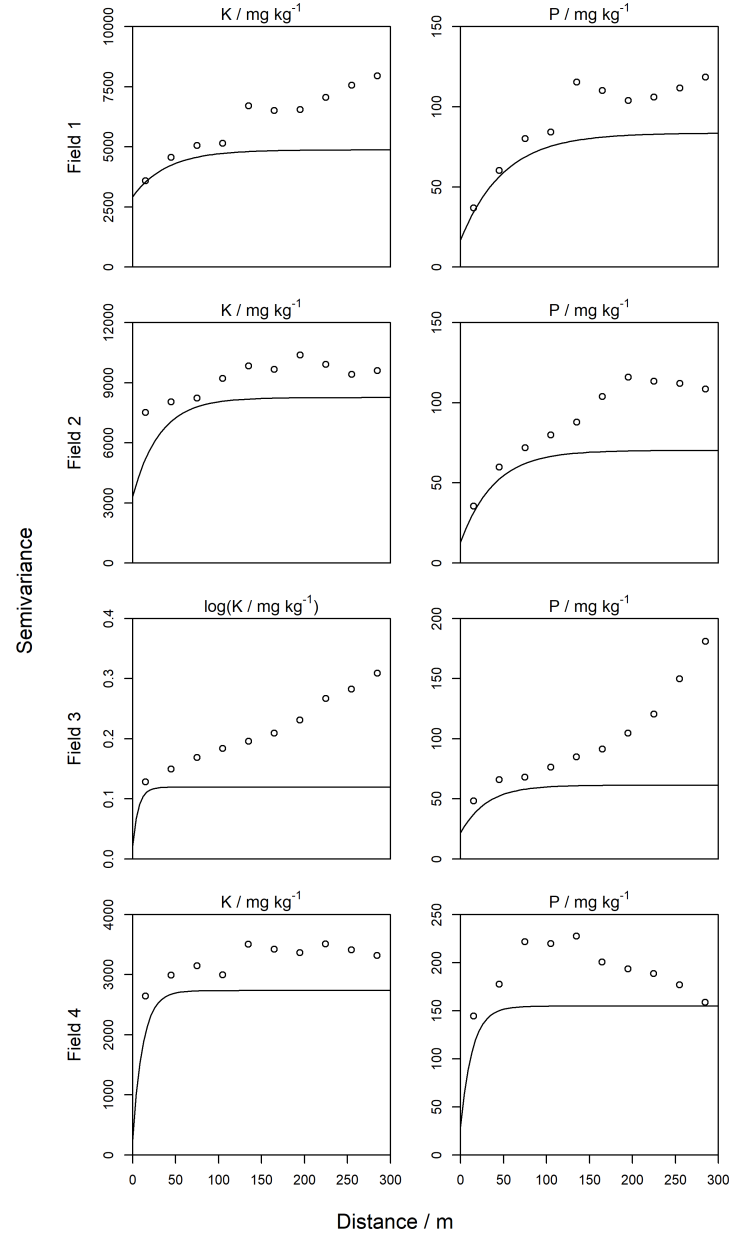


Figure 4.1: Linear mixed model variograms for all fields,  $\circ$  refer to the experimental variograms of the original variable,  $\bullet$  refer to the experimental variograms of the residuals from the trend model, — refers to the final model fitted by restricted maximum likelihood procedures.

As expected, the fertiliser rate that minimises the expected loss ( $F_{\text{opt}}$ ) 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 ( $\sigma_k^2 = 0$ ) (Figs 4.5 and 4.6).

The optimal P fertiliser application varied substantially in all fields. The variation was less pronounced for K, particularly in Field 1. Across the entire Field 1, all the kriging estimates of K fall on the asymptote of the dose–response

Table 4.2: Fixed effects and parameters estimated by REML of the exponential variograms, L stands for LiDAR (elevation), and  $x_1$  and  $x_2$  are the spatial coordinates. The variogram parameters are the nugget variance ( $c_0$ ) the sill ( $c_1$ ) and the range ( $a$ ) respectively.

Field	Soil property	Fixed effects	Variogram parameters		
			$c_0$	$c_1$	$a$
<b>1</b>	P/mg kg <sup>-1</sup>	$x_1, x_2, x_1^2, x_2^2, x_1x_2$	17	67	50
	K/mg kg <sup>-1</sup>	$x_1, x_2, x_1^2, x_2^2, x_1x_2$	2930	1942	40
<b>2</b>	P/mg kg <sup>-1</sup>	L, $x_1, x_2$	13	57	39
	K/mg kg <sup>-1</sup>	L, $x_1, x_2$	3337	4917	32
<b>3</b>	P/mg kg <sup>-1</sup>	L, $x_1, x_2, x_1^2, x_2^2, x_1x_2$	22	40	30
	log(K/mg kg <sup>-1</sup> )	L, $x_1, x_2, x_1^2, x_2^2, x_1x_2$	0.023	0.097	6.62
<b>4</b>	P/mg kg <sup>-1</sup>	L	30	125	14
	K/mg kg <sup>-1</sup>	$x_1, x_2, x_1^2, x_2^2, x_1x_2$	260	2473	12

curve, hence  $F_0 = 0$  for all locations (Fig. 4.6A). The  $\sigma_k^2$  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. 4.6B). The kriged estimates of available K in Fields 1, 2 and 4 were larger than the range of the dose–response curve (Fig. 4.7). Consequently, applying no fertiliser for a major portion of the field was more profitable (Fig. 4.6B, 4.6D and 4.6H). The total expected loss on a field-basis was less for variable-rate P and K application ( $E[L(F_{\text{opt}})]$ ) than the total loss of blanket fertiliser application arising from the wet chemical analysis ( $E[L(F_{\text{UA}})]$ ) (Table 4.3). There was a financial incentive of VRA of P fertiliser across all fields (ranging from £12–£76 ha<sup>-1</sup>) and for K fertiliser across two fields (ranging from £15–£73 ha<sup>-1</sup>). That is, for available K the difference between  $E[L(F_{\text{opt}})]$  and  $E[L(F_{\text{UA}})]$  was small in Fields 1 and 4. Less P fertiliser was used on a field-basis under VRA ( $F_{\text{opt}}$ ) than with uniform application ( $F_{\text{UA}}$ ) in Fields 1 and 2 (Table 4.3). Within those fields, total P fertiliser use was reduced by VRA with 13–25 kg ha<sup>-1</sup> compared with uniform application. Most K fertiliser would be used on a field-basis in all cases under VRA ( $F_{\text{opt}}$ ) (Table 4.3).



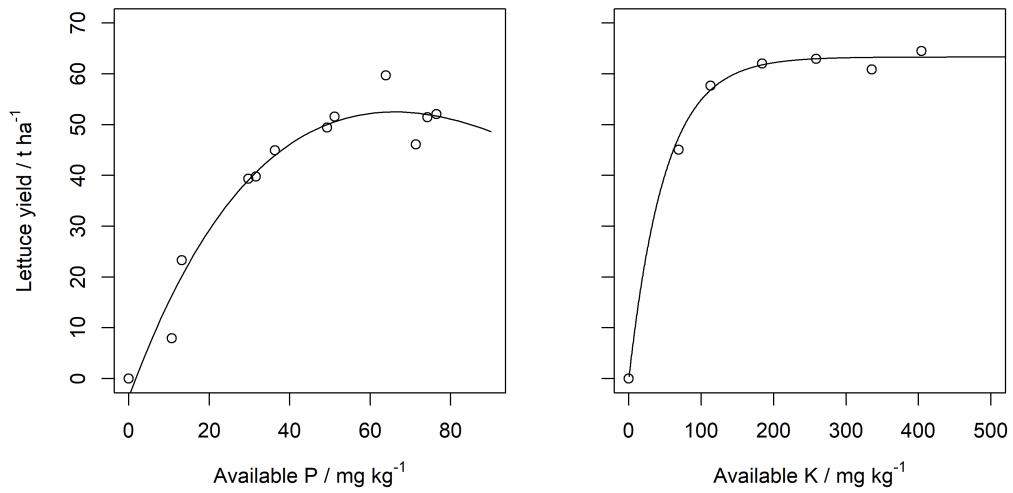


Figure 4.2: Dose-response curves for available P (exponential + linear) and available K (exponential) fitted based on data from Prasad et al. (1988) (P) and Greenwood et al. (1980) (K). See Table 4.1 in the main text for parameter values.

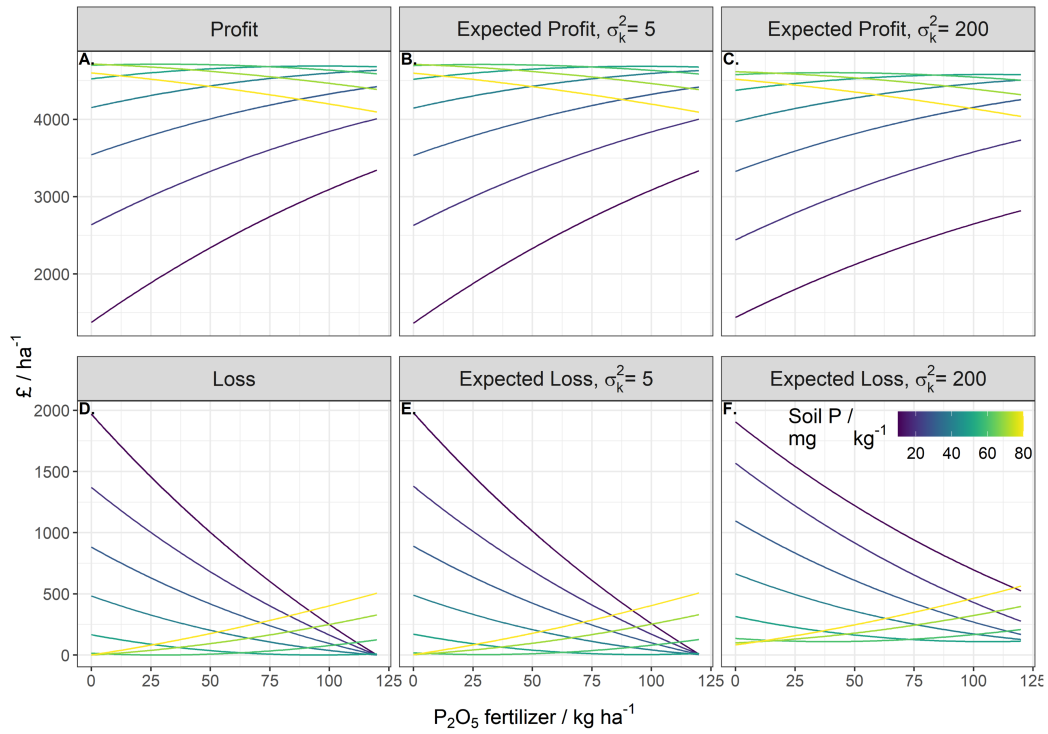


Figure 4.3: Profit and loss under zero error variance, expected profit and loss under an error variance of 5 mg kg<sup>-1</sup> and an error variance of 200 mg kg<sup>-1</sup> for a range of estimated soil P values from 10 to 80 mg kg<sup>-1</sup>. The range of P fertiliser applied spans 0 to 120 kg ha<sup>-1</sup>.

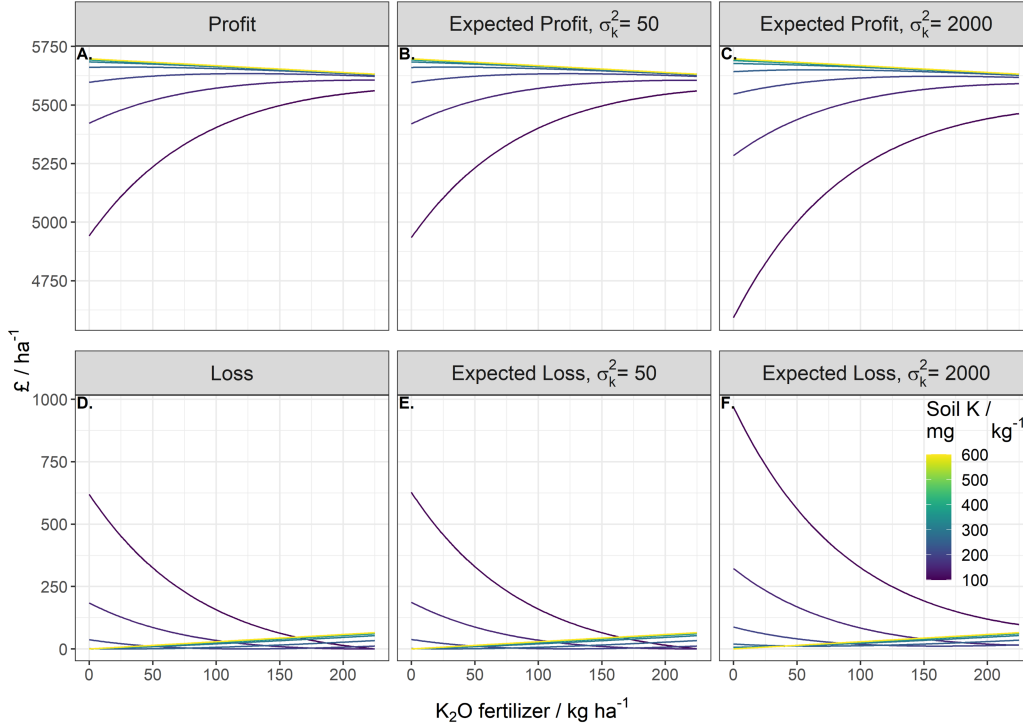


Figure 4.4: Profit and loss under zero error variance, expected profit and loss under an error variance of 50 mg kg<sup>-1</sup> and an error variance of 2000 mg kg<sup>-1</sup> for a range of estimated soil K values from 100 to 600 mg kg<sup>-1</sup>. The range of K fertiliser applied spans 0 to 225 kg ha<sup>-1</sup>.

## 4.4 Discussion

### 4.4.1 Error approximation and the estimation of the variogram

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 our case  $\text{var}[\hat{\varepsilon}_y]$ , 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 minimization) is necessary to detect small rates of change in the soil property of interest by monitoring over time (Viscarra Rossel et al., 2016). Here we are concerned with variation in space, and it is this variation that determines whether variable rate application is relevant.

In four of the eight LMMs we included LiDAR as a fixed effect, 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 we separated out the fixed effects of trend

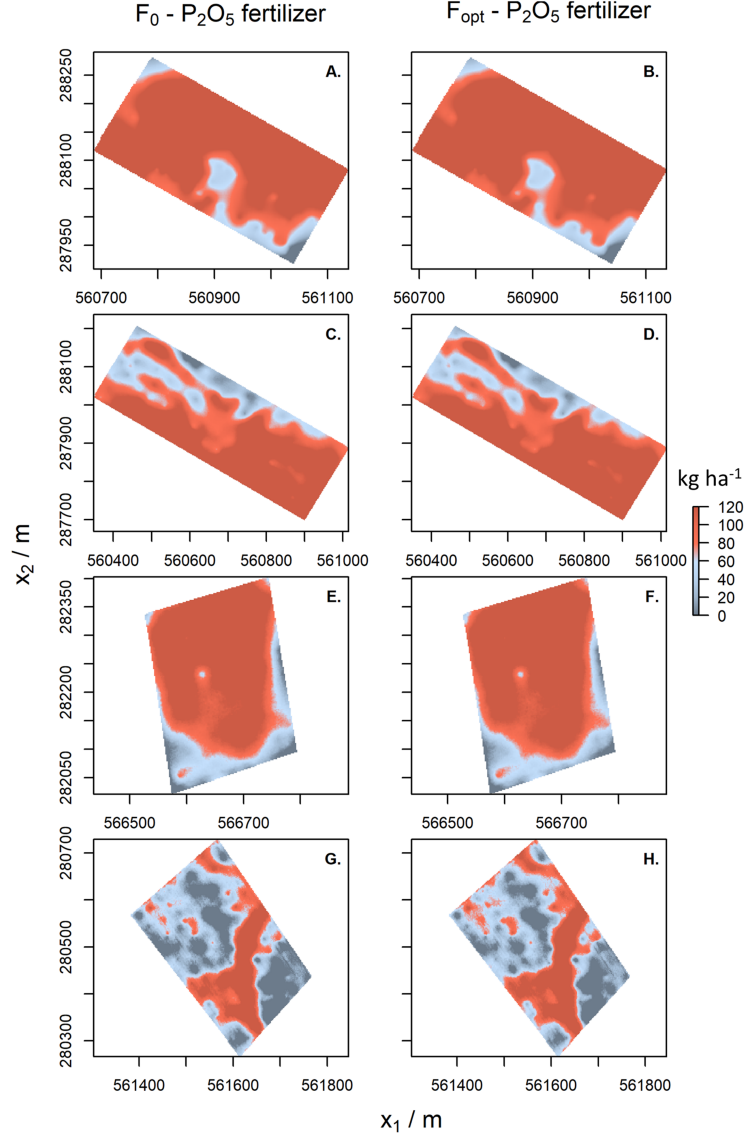


Figure 4.5: Optimum P fertiliser application with perfect knowledge of soil P ( $F_0$ ), optimum application when accounting for  $\sigma_k^2$  in our estimate of soil P ( $F_{\text{opt}}$ )

and LiDAR and also why we treat each field 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).

#### 4.4.2 Data requirements and estimation of the loss function

Use of the loss function imposes constraints on the required data. For example, we estimated the soil's bulk density from general knowledge in the area (see method section 3.4). We know 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 in our description we did not include this for reasons of clarity.

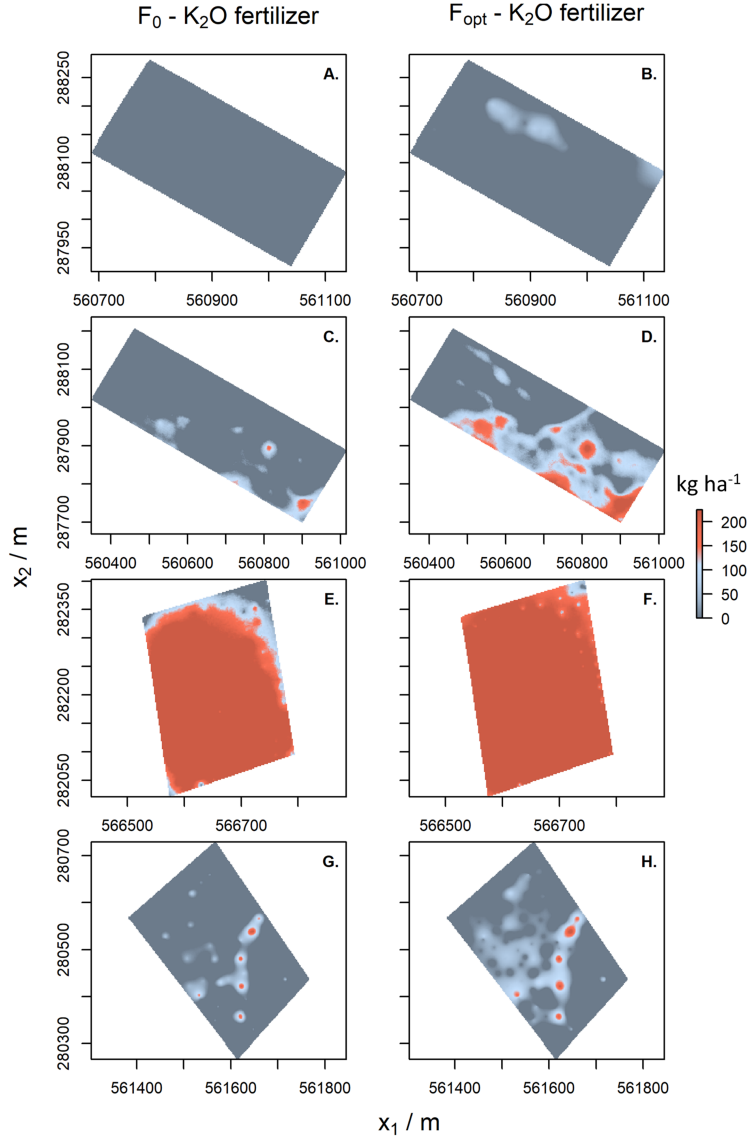


Figure 4.6: Optimum K fertiliser application with perfect knowledge of soil K ( $F_0$ ), optimum application when accounting for  $\sigma_k^2$  in our estimate of soil K ( $F_{\text{opt}}$ )

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

Based on the differences in  $E[L(F)]$  between  $F_{\text{opt}}$  and  $F_{\text{UA}}$ , there appears to be little financial incentive for variable-rate application of K fertiliser in Fields 1 and 4. The difference in  $E[L(F)]$  between  $F_{\text{opt}}$  and  $F_{\text{UA}}$  is larger for Fields 2 and 3. For available P, most kriged estimates 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, although Field 1 does not show a large difference (Table 4.3).

The difference in total K fertiliser used between  $F_0$  and  $F_{\text{opt}}$  was especially large for Field 2, which can be explained by the high nugget variance ( $c_0$ ) and sill ( $c_1$ ) (Fig. 4.1 and Table 4.2). Field 4 has the largest  $E[L(F)]$  values for P (Table 4.3), which can be attributed to large values of  $\sigma_k^2$  (Fig. 4.7), a short range parameter ( $a$ ) and large sill ( $c_1$ ) in the variogram (Fig. 4.1 and Table 4.2). Additionally, the smaller applications of P fertiliser under VRA than under uniform application for Fields 1

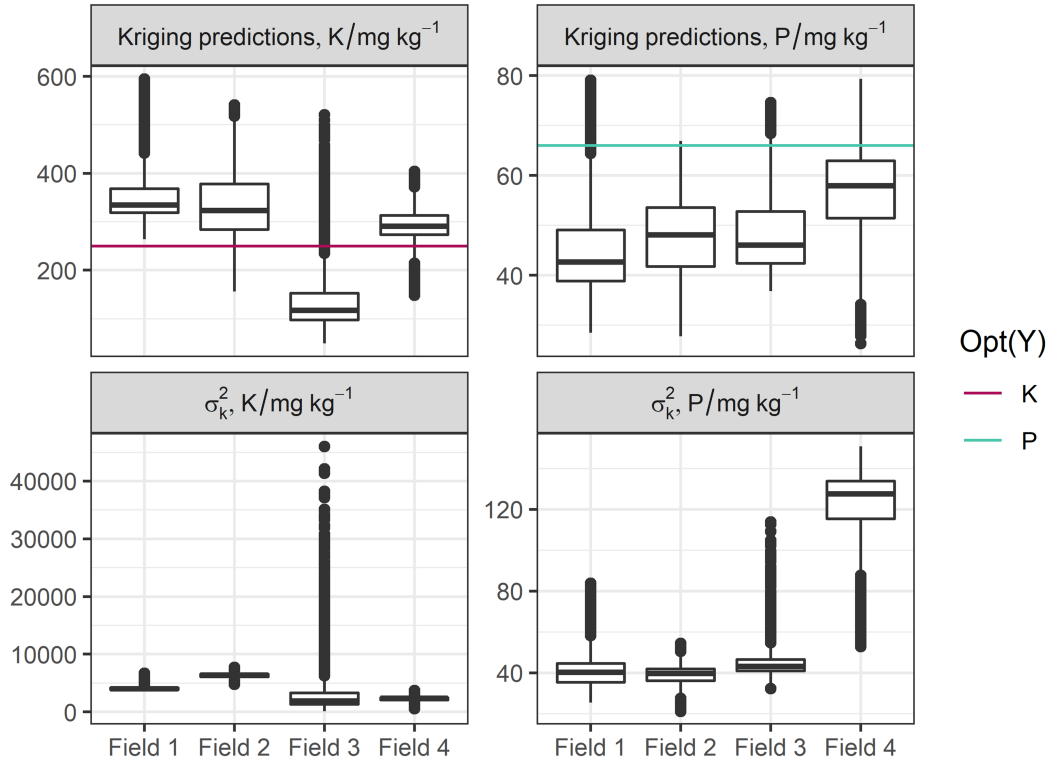


Figure 4.7: Box-plots of kriging predictions and the kriging variance ( $\sigma_k^2$ ), by field for available P and K, horizontal lines represent the nutrient value for which the maximum yield,  $\text{Max}(Y)$ , is obtained on the fitted dose-response curve.

and 2 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  $F_{\text{UA}}$ . 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 our W-design were chosen *a posteriori* (the samples come from the set that was analysed by wet chemistry and these were selected to span the range in the field) we suspect that mean estimates of available P and K are biased.

We further note that the expected loss under uniform application of K fertiliser is large for Field 3. The large expected loss can be attributed to the combination of a small range relative to the dose-response curve combined with a long-tailed distribution of kriging variances caused by back-transformation from the logarithmic scale.

Overall, the expected loss,  $E[L(F)]$ , and hence  $F_{\text{opt}}$  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.

Table 4.3: The estimated mean nutrient value for each field by wet chemistry ( $\mu_{\text{UA}}$ ), fertiliser used per hectare for perfect knowledge ( $F_0$ ), variable-rate application ( $F_{\text{opt}}$ ) and uniform application based on wet chemistry samples ( $F_{\text{UA}}$ ). Expected loss (from perfect knowledge) for variable-rate application,  $E[L(F_{\text{opt}})]$ , and uniform application,  $E[L(F_{\text{UA}})]$

Field	Area / ha <sup>-1</sup>	Nutrient	$\mu_{\text{UA}}$ / mg kg <sup>-1</sup>	Fertiliser / kg ha <sup>-1</sup>			Expected loss / £ ha <sup>-1</sup>	
				$F_0$	$F_{\text{opt}}$	$F_{\text{UA}}$	$E[L(F_{\text{opt}})]$	$E[L(F_{\text{UA}})]$
1	8.2	P	40	106	107	120	33	45
2	16.9		37	94	95	120	30	54
3	5.1		55	97	98	62	34	110
4	8.9		56	54	61	57	90	137
1	8.2	K	337	0	3	0	9	9
2	16.9		302	6	44	0	21	36
3	5.1		189	200	221	144	44	117
4	8.9		312	6	17	0	8	11

#### 4.4.4 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 optimize the size of samples for survey; see Yates (1981), Journel (1984), Goovaerts (1997), Lark & 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  $E[L(F_{\text{opt}})]$  and  $E[L(F_{\text{UA}})]$  then VRA is worthwhile. These costs could be reduced by measuring the reflectance spectra of the soil surface on the

run in the field. It would then be feasible to obtain data at 10-m intervals, as McBratney et al. (1996) stated to be necessary to characterize within-field soil variation accurately. 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. We found no reports to support within-field estimation of available P and K from MIR and XRF spectroscopy.

Within six out of eight fields in our case-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 K fertiliser does not pose a direct threat to the environment. The overuse of P fertiliser on the other hand, particularly in Field 3, might minimize economic loss at the cost of the environment. However, it could be argued that increased efficiency gained by VRA also has potential to reduce the total land area used to reach an equivalent level of yield outputs. Setting aside agricultural land in areas vulnerable to leaching is a recognized strategy to manage P levels at the catchment scale (Schoumans et al. 2014). In that case it would need to be ensured that leaching or artificial drainage is not the driving force behind low P levels at the field-scale in the first place (Baveye & Laba, 2015). Furthermore, our results further showed that the environmental benefit of fertiliser savings from VRA was not strictly accompanied with an increase in profit (Field 1). Hence, in order to account for the environmental benefits of precise fertiliser application, our results suggest that costs of P leaching (e.g. remediation) need to be included in sustainable P management strategies aimed at precise fertiliser application. The loss function framework presented could be adapted to place a larger penalty on over-application of P fertiliser based on the costs of leaching. Consequently, the loss function framework can provide a stepping stone towards fulfilling the requirement to quantify the economic and environmental benefits of sustainable P management (Brownlie et al., 2021).

## 4.5 Conclusions

This study has shown that there was an economic incentive for precise fertiliser application of both P and K fertiliser once we accounted for uncertainty in soil-nutrient concentrations estimated from sensors. Given that farmers 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.

In order to quantify the societal benefits of precise fertiliser application holistically, however, environmental costs need to be taken into consideration. Our results showed environmental benefits occurred from precise fertiliser application even though no large increase in profit was gained. These findings have implications for policy aimed at sustainable management of P fertilisers. That is, we recommend 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.

Furthermore, the loss function we defined can be extended to other studies within

the environmental sciences. We discussed the use of the loss function for robust decision-making based on uncertain data, the identification of gaps in knowledge, to assess the accuracy of information in answering questions and as a method to assess where reduction in uncertainty leads to the largest gains.



## Chapter 5

### Investigating the effect of uncertainty under different sample sizes



---

## Abstract

*Few studies to date have investigated the effect of uncertainty in soil property estimates from spectroscopy on soil management. In this study we considered the implications for variable rate application of phosphorus (P) and potassium (K) fertiliser. First the uncertainty in soil available P and K estimates from spectroscopy was quantified as a function of the calibration set size at the field-scale. Based on the observed variation in P and K in three experimental fields, we simulated 100 realisations per field for an in silico experiment. To simulate the process of sampling soil and predicting fertiliser requirement, we performed sampling on our simulated fields using a spatial coverage design with 20 close-pairs. We added a calibration error to each sample value to simulate the error associated with spectroscopic prediction. Kriging was used to estimate the variation in the soil property of interest. We then computed the fertiliser requirement needed to minimise the expected loss associated with predictions, where the expected loss is defined as the difference in profit between applying fertiliser based on the estimated soil nutrient concentration accounting for uncertainty relative to the profit that would be gained from fertiliser application given the true value of available P and K. We also considered the expected profit when accounting for sampling costs. Results showed that calibration sample size outweighed the effect of total sample size on the uncertainty associated with predictions. Equally, for the same calibration set size, there were large differences in the kriging variance between total sample sizes. When data acquisition costs were disregarded, the expected loss for available P was particularly affected by the total sample size. For available K, the calibration sample size had a predominant effect on the expected loss. The expected loss showed diminishing returns on investment suggesting that there is an optimum sample size. However, the expected profit was dominated by the costs of sampling and spectroscopy, indicating that currently using spectral methods to inform fertiliser management is not cost effective. That is, no combination of the total- and calibration sample sizes considered would result in a financial gain and could thus be considered optimal. Should costs substantially reduce then spectral methods offer a promising method for informing variable rate management. We conclude that the loss function approach is an appropriate method to assess whether soil spectroscopy is a cost-effective means to inform soil management. We further suggest its application in different case-studies to gain more robust insight in the value of applied soil spectroscopy.*

Based on:

Breure, T.S., Haefele, S.M., Corstanje, R., Hannam, J.A., Moreno-Rojas, S., Milne, A.E. 2021. Quantifying the effect of prediction uncertainty from soil spectroscopy for different total- and calibration sample sizes in the context of soil management.  
– *In preparation*

## 5.1 Introduction

Sustainable management of the soil requires explicit knowledge about the spatial variation of its properties. It is not feasible to measure soil properties everywhere and hence soil samples are taken from a limited set of locations. Interpolation methods (for example inverse distance weighting or kriging) are then used to predict the values of the variables at unmeasured locations. Kriging provides the best unbiased prediction at each unsampled location together with its associated error variance (Cressie, 1990). A prediction based on kriging is a weighted averaging of the observed values for the soil property within a neighbourhood. The weights are determined from a model that describes the spatial dependency of the variable of interest. For robust estimates of model parameters that describe the spatial dependency, a minimum sample size of 100 observations has been suggested, although this depends on the underlying complexity of the soil variation and its geographical extent (Webster & Oliver, 1992; Lark, 2000). Given the costs and time expenditure of traditional soil analyses, the number of samples required for kriging is often impractical.

To overcome this problem, there have been many applications of soil spectroscopy as a method to characterise soil chemical, physical and biological properties (Guerrero et al., 2010). This application is not only motivated by the fact that spectroscopy is relatively inexpensive compared to traditional laboratory methods, but also that it is non-destructive (in the case of *in-situ* measurements), does not require hazardous chemicals and has the capacity to predict multiple soil properties from a single measurement (Viscarra Rossel et al., 2006). Soil spectroscopy has the potential to provide cost-effective prediction of soil properties allowing for more samples to be processed than traditional methods and so potentially providing more information about soil variation.

To predict a soil property of interest using spectroscopy the spectra must be calibrated. This is usually done by undertaking wet chemistry measurements on a representative subset of samples and deriving a statistical model (known as the calibration model) to relate the spectra to the wet chemistry measures. The calibration model is then used to predict the values for the total sample population of interest. This model is subject to uncertainty, however, and is potentially biased. For both accurate and precise predictions, a calibration set that includes coverage of property (or feature) space is required (Viscarra Rossel et al., 2011; Schmidt et al., 2014). Violation of this central assumption in the use of soil spectroscopy can lead to model instability and introduce bias and error in estimates (Bellon-Maurel et al., 2011). The number of samples used for calibration will also affect the uncertainty associated with predictions and the cost of soil data acquisition by spectroscopy.

The two sources of uncertainty described above (kriging error and calibration error) can usually be reduced by increasing sampling effort. The kriging error can be reduced by increasing the sampling density. The calibration error can be reduced by increasing the number of calibration samples. The effort of sampling in itself constitutes a large cost of a survey by soil spectroscopy as well as the laboratory analysis of the calibration samples (deGruijter et al., 2018). According to a review by Viscarra Rossel et al. (2011), the majority of the peer-review literature on soil spectroscopy considers either geographical space sampling or feature space sampling designs while not covering both (van Groenigen & Stein, 1998; de Gruijter,

2002; de Gruijter et al., 2006). Sampling methods for a calibration set have been implemented by response-surface sampling (Lesch, 2005) and conditional Latin hypercube sampling (Minasny & McBratney, 2006). Several studies addressed both geographic and feature space sampling (Minasny et al., 2007; De Gruijter et al., 2010, 2016; Adamchuk et al., 2008, 2011; and Shaw et al., (2016). Except for de Gruijter et al. (2018) and Shaw et al. (2016), these studies on sampling in the context of soil sensing did not take into account the value associated with reducing uncertainty versus the costs associated with sampling.

One method to address both the costs and accuracy of a survey effort is through the loss function framework described first by Yates (1949). Yates (1949) described the expected loss in accuracy as a function of sampling effort, which consequently allows one to determine the sampling effort required to minimise the expected loss accounting for the costs associated with the sampling. It is referred to as an expectation since the estimate itself has an associated probability distribution. The concept of the loss function was further developed by Lark & Knights (2015). The expected loss is described by a function that determines the costs incurred from decision-making based on an estimated value, given its deviation from the true value. Lark & Knights (2015) noted that in general the loss function is asymmetric since the consequences of over- and underestimation are different in kind and magnitude. Depending on the loss function's asymmetry, there might be a slight preference towards either over- or underestimation. Consequently, the optimum value might actually not be equivalent to the largest expected financial gain.

Within this study we considered soil spectral predictions of potassium (K) and phosphorous (P) to inform variable rate fertiliser management. The costs incurred within this decision-making context are associated with a.) the loss in yield due to insufficient amount of fertiliser and b.) the financial loss due to overapplication of fertiliser. Our aim was to quantify the uncertainty in the prediction of available P and K from spectral measurements for different total- and calibration sample sizes. We then investigated the effect of this uncertainty on decision-making for soil management using the loss function framework. For that purpose, we identified the following research questions:

- How does total sample size and total number of calibration samples affect prediction accuracy?
- How does total sample size and total number of calibration samples affect expected loss when sampling costs are not accounted for?
- How does total sample size and total number of calibration samples affect expected profit when sampling costs are accounted for?

## 5.2 Methods

### 5.2.1 Case study area, sampling, wet chemistry analysis and spectroscopy

Our simulated fields were based on data obtained in sample surveys of three 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 17th century. Due to the oxidization, the underlying alluvial and marine silts have become exposed showing complex sub-field soil variation.

To characterise the variation in P and K across the study fields we took soil samples from across each field. The sampling designs, wet chemistry methodology, spectroscopy and calibration are described in detail in Breure et al. (2021a), hence we summarise them briefly here. 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), and Field 3 (8.9 ha) were computed by spatial coverage sampling (Walvoort et al., 2010). From these initial points we selected a subset with balanced sampling (Grafström & Lisic, 2019) on the spatial coordinates and elevation (measured by LiDAR). The number of samples in this subset was 36 for Field 2 and 32 for Field 3. At each location of these sub-samples, we added another sampling point 6 m away at a random orientation to estimate the short-scale spatial variance.

Spectroscopic measurements were taken on each of the soil cores. This comprised near-infrared (NIR) and mid-infrared (MIR) measurements using a Tensor II spectrometer (Bruker, Ettlingen, Germany), and X-ray fluorescence (XRF) spectra, measured by a DP-6000 Delta Premium portable X-ray fluorescence (pXRF) spectrometer (Olympus Ltd, Center Valley, USA). The pXRF samples were 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 the three measurements. In terms of spectral pre-processing, we applied the Savitzky–Golay filter (Savitzky & Golay, 1964) on the raw spectra and then transformed them to standard normal variates and their first derivatives.

In order to predict a given variable from soil spectra, we developed a calibration model by regression of the soil spectra (using partial least squares regression) onto wet chemistry measurements made on a subset of the soil cores (for details see Breure et al. 2021a). In each field we selected 30 locations to be measured by wet chemistry using a balanced sampling approach on the coordinates and elevation (from LiDAR). These samples measured by wet chemistry constituted our calibration set.

### 5.2.2 Simulated variation in available P and K

Our aim was to simulate the process of sampling across a field to predict the soil properties from spectral measurements, selecting a subset of these samples for calibration, and then predicting the spatial variation in P and K from point spectral measurements to inform fertiliser management. By accounting for sampling effort when optimising the profit associated with this decision, we can determine the number of samples and the calibration set that would have been optimal for each of our simulated fields. To assess this, we used random realisations of the true variation in each field as follows.

The random realisations were based on the kriged map of true soil property predictions. Before we could krig, a variogram model was required to describe the spatial dependence between sample locations. The main assumption of the variogram is that the mean and variance are stationary across the area of interest. Based on preliminary plots this assumption did not seem to hold since available P and K varied systematically with the coordinates and elevation. We therefore fitted a linear mixed model to the data whereby the trend factors were fixed effects and

the spatial autocorrelation was captured in the random term (see Lark & Cullis, 2004). For the random term a variogram model form needs to be selected *a priori*. An initial investigation suggested that the exponential variogram model fitted well so we adopted that for our analysis. This model is given by the equation:

$$\begin{aligned}\gamma(h) &= c_0 + c_1 \left\{ 1 - \exp\left(-\frac{h}{a}\right) \right\} \quad \text{for } h > 0 \\ &= 0 \quad \text{for } h = 0 .\end{aligned}\tag{5.1}$$

In this equation  $c_0$  and  $c_1$  are variances, respectively the nugget and sill of the correlated variance, and  $a$  is the distance parameter.

To test the significance of the coordinates (eastings, northings and an interaction term) and elevation as trend parameters, we added each in turn and performed a log-likelihood ratio-test. Fitting was done by Maximum Likelihood to allow for the comparison between models with a different number of fixed effect parameters. A chi-squared  $p$ -value of 0.05 from the log-likelihood ratio was taken as significant evidence that the trend parameters should be included. Once we came to a final set of fixed effects, both the fixed and random effects (variogram parameters) were estimated by residual maximum likelihood (REML). REML is generally preferred since it reduces bias in the random effect parameters due to uncertainty in the fixed effect parameters (Lark et al., 2006). The estimates of the random- and fixed effect parameters are listed in Table 5.1.

In the later step where we generated realisations of the field and sample in silico, we adopted an automated process for fitting the variogram and kriging (see below). Therefore, to reduce the chances of problematic fitting we removed the trend from the raw data and considered only the residuals when creating our in-silico realisations of the field and sampling process. The trend surface was only readded at the step where we determined the fertiliser requirements. To create our in-silico fields we kriged the residuals onto a 2 m  $\times$  2 m grid (Webster & Oliver, 2007). Kriging naturally smooths the true variation in the field, therefore to simulate more realistic variation we added a component of noise to our simulated data by sampling from  $\text{Norm}(0, \sqrt{c_0})$ , where  $c_0$  is uncorrelated variance, known as the nugget variance of the fitted variogram. The outcome of these realisations overcome the conditional bias introduced by kriging and portray a more realistic of the global variation (Goovaerts, 1997; Deutsch & Journel, 1998). We considered these simulated data the underlying true values of the soil variable and denominated this true value at any given location as  $S$ .

### 5.2.3 Estimating the error from the calibration regression

The next step in our investigation was to formulate an equation that described the prediction error variance as a function of the number of calibration samples ( $n_c$ ). Given our limited samples with wet chemistry reference values for each field, we pooled all samples used for calibration. The pooled set was used to select calibration samples by the conventional method of using the Kennard-Stone algorithm on a matrix of combined NIR, MIR and XRF soil spectra. This method allows for selecting samples with a uniform distribution over the predictor space based on their Euclidean distances (Kennard & Stone, 1969). The number of calibration samples was varied from 20 samples to 80 samples by increments of 6 samples. We used the

Table 5.1: Exponential variogram parameters (random effects) and trend parameters (fixed effects) used for the simulation of trend residuals in each field, where  $c_0$  is the nugget variance,  $c_1$  the sill,  $a$  the distance parameter,  $\beta_0$  the intercept,  $\lambda$  elevation as measured by LiDAR and  $x$  and  $y$  are coordinates.

Soil property	Field	Random effects			Fixed effects						
		$c_0$	$c_1$	$a$	$\beta_0$	$\lambda$	$x$	$y$	$x^2$	$xy$	$y^2$
P	1	17	68	53.2	77.2	-	-0.01	-0.33	3.11e-05	-7.42e-05	8e-04
	2	13	57.3	38.5	5.89	-6.52	0.03	0.083	-	-	-
	3	29.7	125	13.6	4.80	-34.38	-	-	-	-	-
K	1	2939	1942	39.5	648.6	-	-0.29	-2.97	1e-03	0.001	0.007
	2	3337	4918	31.5	-51.42	-113.9	0.095	0.59	-	-	-
	3	260	2473	11.9	356.3	-	-0.61	-0.45	0.003	0.001	0.001

calibration dataset for predicting the remaining validation set. The mean-squared error was used as an accuracy metric to describe the error variance, denoted by  $\sigma_{n_c}^2$ .

#### 5.2.4 Modelling procedure to compute kriging predictions for different sample sizes

We simulated the process of predicting the soil nutrient status from spectral measurements as follows. First, we simulated the process of selecting the sampling points for spectral measurement in the field. For this we computed a spatial coverage sampling design for  $N_t$  samples. A spatial coverage design allows sample locations to be evenly spread across the domain of interest to minimise the maximum interpolation error (Webster & Oliver, 2007). In order to gain a robust estimate of the soil variation in space, sampling locations should include distances shorter than the range at which the variable is spatially correlated. The requirement of including the spatially correlated range can be met more effectively by including paired observations within close distance for a subset of the design (Lark & Marchant, 2018). Thus, a subset of 20 locations were selected by balanced sampling on the eastings, northings and LiDAR and an additional sample point added 6 m away from each in a random direction. At each sample location in our in-silico experiment, we assumed the soil variable is predicted from spectroscopy. To account for the error associated with this, we sampled from the normal distribution  $\text{Norm}(0, \sigma_{n_c})$  and added that to the simulated true value of the soil to give our observed predicted value. These observed values were then kriged as follows.

Visual inspection is unsuitable for the number of variograms to be estimated within our analysis, hence we implemented the following approach to gain initial parameter estimates for an exponential variogram. First, we computed the bounding box around the spatial coverage sampling design. The initial estimate of the distance parameter (a) was one tenth of the diagonal of the bounding box. We used half the diagonal as the maximum distance for the experimental variogram. Second, we computed the omnidirectional experimental variogram using the method of moments (Webster & Oliver, 2007). The minimum of the experimental variogram's semivariance was taken as the initial estimate for the nugget parameter ( $c_0$ ). The mean value of the median and maximum semivariance was used as the initial estimate of the sill parameter ( $c_1$ ). The variogram parameters were then estimated by REML. In cases where the estimated nugget parameter was lower than the known error from calibration,  $\sigma_{n_c}^2$  the model was refitted with ( $c_0 = \sigma_{n_c}^2$ ), similarly  $c_0$  was restricted to not fall below zero. The variogram model was then used to perform ordinary-kriging. Once we obtained the kriging predictions and their error variances at each location, we re-added the trend to the predictions. This procedure was repeated for 100 realisations for each soil variable (available P and K) and field and for each combination of total sample size ( $N_t$ ) and calibration set size ( $n_c$ ). For total sample size  $N_t$  we considered values between 100 and 520 and for the calibration set size ( $n_c$ ) we considered values between 20 and 80.



### 5.2.5 The loss function and variable costs of the data acquisition

The deviation in estimates of available P and K compared to the true value was then assessed within the context of decision-making on fertiliser management. We quantified the effect of error in the estimates by a loss function,  $L(F)$ , for a given application of fertiliser,  $F$ , which is generically defined as the difference in profit that results from applying a given amount of fertiliser  $F$  compared to an economically optimal amount of fertiliser  $F_0$ :

$$L(F) = \Phi(F_0) - \Phi(F) , \quad (5.2)$$

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

$$\Phi(F) = M \times \text{Yield} - V \times F , \quad (5.3)$$

where  $M$  is the price of the crop (£ t<sup>-1</sup>) and  $V$  is the cost of the fertiliser (£ kg<sup>-1</sup>). Given our focus on precise fertiliser application, we assume that the yield is given by the dose-response equation:

$$\text{Yield} = \alpha + \eta R^{\xi F + S} + \nu(\xi F + S) , \quad (5.4)$$

where  $S$  is the concentration of the nutrient in the soil,  $F$  is the applied fertiliser (kg ha<sup>-1</sup>),  $\xi$  is the increase in nutrient concentration (mg kg<sup>-1</sup>) in the soil for every 1 kg ha<sup>-1</sup> fertiliser applied, and  $\alpha$ ,  $\eta$ ,  $\nu$  and  $R$  are parameters. Equation 5.4 describes the generic dose response curve function. Since the case-study area was used for the cultivation of lettuce we derived relevant dose response curves from the literature for both P and K for this crop (Greenwood et al., 1980; Prasad et al. 1988). We 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, we assumed that the added nutrients are contained in the top 25 cm of the soil (the sampling depth). We took from Milne et al. (2006) the value of 480 kg m<sup>-3</sup> for bulk density of this peat soil. Given the support of our kriged predictions (2 m  $\times$  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<sup>-1</sup> and 0.52 mg available K kg<sup>-1</sup>, equal to  $\xi$  in the dose-response Equation (5.4). Greenwood et al. (1980) listed a mean base nutrient concentration of 69 mg available K kg<sup>-1</sup> for the unfertilised soil in their study, which was used as an additive component. We assumed a profit margin ( $M$ ) of £90 per tonne of lettuce per hectare. The prices of fertiliser ( $V$ ) were taken as £0.36 per kg P fertiliser and £0.29 per kg K fertiliser.

Based on Equation 5.3, we then calculated the economically optimum amount of fertiliser, which is given by:

$$F_0 = \ln \left( \frac{B/\xi - \nu}{\eta R^S \ln R} \right) / \xi \ln R , \quad (5.5)$$

where  $B = V/M$ , known as the break-even ratio. By definition, the loss given by Equation (5.2) is zero when the optimum amount of fertiliser is applied. However, computing the optimum amount of fertiliser to apply relies on an exact estimate  $\hat{S}$

whereas predictions from kriging have an associated error distribution,  $f(s)$ . Given the error distribution, we computed the optimum fertiliser rate that maximizes the expected profit:

$$F_{\text{opt}} = \ln \left( \frac{B/\xi - \nu}{\eta \ln R \int_0^\infty R^s f(s) ds} \right) / \xi \ln R . \quad (5.6)$$

The application of  $F_{\text{opt}}$  minimises the expected loss function,  $E[L(F)]$ , which we define here as the difference between the profit where  $S$  is known without error and the profit under the fertiliser application based on the kriged soil nutrient value,  $\hat{S}$ .

$$E[L(F)] = \Phi(F_0, S) - \Phi(F, \hat{S}) . \quad (5.7)$$

We computed  $E[L(F)]$  for two scenarios. One where the fertiliser regime,  $F$  given  $\hat{S}$  equals the application of  $F_0$  (Equation 5.5) and the second where we account for the uncertainty in our estimate of  $\hat{S}$  and  $F$  equals the application of  $F_{\text{opt}}$  (Equation 5.6).

As explained in the introduction, a major constraint to accurate predictions are the costs involved of spectroscopy at the field scale. Major components are the costs of sample sieving, milling, weighing, wet chemistry for calibration, spectroscopy and the sampling campaign itself. The costs with regards to the total number of samples,  $N_t$ , and the number of samples used in calibration,  $n_c$ , were approximated as a simple linear function formulated by the variable costs of a field-work survey, spectroscopy, sample processing, sample handling and analytical measurements.

$$C(N_t, n_c) = \Omega N_t + \delta n_c \quad (5.8)$$

Based on the sampling procedure for our case-study fields, we estimated the variable costs of sampling at £5.7 per sample. We assumed the same variable costs for all sample sizes, given the relatively small areal extent of the fields. Spectroscopy costs were equivalent for both  $N_t$  and  $n_c$  and consist of both milling and loading sample plates for the benchtop spectrometer on a rate of 60 samples for one full working day of technical staff, which are included in the variable costs for the total number of samples ( $\Omega$ ). Based on a salary of £135 per day this would be equivalent to £2.25 per sample. However, given the separate measurement procedure for the XRF spectra, we doubled this value, to £5.50.

The costs associated with  $n_c$  were approximated as a function of the sample handling (sieving and weighing) and the wet chemistry costs, denoted as  $\delta$ . Based on our sample processing experience, we assumed that sieving and weighing would take up 20 minutes per sample (one and a half working day for 60 samples to sieve and weigh), equivalent to £3.38 per sample. The costs of analysing available Potassium (K) by ammonium-nitrate extraction and ICP-OES was estimated at £14.70 per sample. The analysis for plant available Phosphorus (P), by the Olsen method was estimated at £16.30 per sample.

We expected that the sampling costs would exert a strong influence on the expected profit. Hence, we applied a scaling factor to the data acquisition costs ( $C(N_t, n_c)$ ) to explore to what degree sampling costs would need to reduce to make variable rate fertiliser supported by spectroscopy financially viable. We applied a scaling factor of 5%, 1% and 0.5% of the original value to assess its effect on the the expected profit.

### 5.2.6 Overview of the modelling procedure

The entire procedure described above is summarised in Table 5.2. Based on this procedure, we quantified the uncertainty in predicted soil P and K as a function of calibration sample size  $n_c$  and total sample size  $N_t$ . We then assessed the expected loss associated with precise fertiliser application under uncertainty, excluding data acquisition costs. Subsequently, we calculated the expected profit accounting for costs of the data acquisition procedure. The differences in expected profit would determine which combination of total- and calibration sample sizes leads to the largest financial gain under a precise fertiliser application regime under current estimated costs. Finally we explored the impact of reducing the costs associated with the data acquisition procedure. Together these results determine: a.) the effect of increased sampling activity on prediction uncertainty b.) the value of information under these different sampling designs and c.) whether soil spectroscopy is cost-effective given the conditions in this case-study.

Table 5.2: Overview of the modelling procedure.

Step	Action
i	Simulate true residuals based on trend parameters from REML fit (Table 5.1)
ii	Spatial coverage sampling for $N_t$ locations
iii	Select subset of spatial coverage samples by balanced sampling
iv	Random sample from subset neighbour locations within 6 m buffer distance
v	Sample from $\text{Norm}(0, \sigma_{n_c})$ to add calibration error to samples
vi	REML estimation of exponential variogram parameters
vii	Universal kriging to gain kriged predictions and their variances
viii	Add trend surface to kriging predictions
ix	Estimate $F_0$ and $F_{\text{opt}}$ given $\hat{S}$ and $f(s)$ (Eq. 5.6)
x	Compute the expected loss ( $E[L(F)]$ ) excluding the costs of data collection (Eq. 5.7)
xi	Subtract the expected profit ( $\Phi(F_{\text{opt}})$ ) (Eq. 5.3) with the data acquisition costs (Eq. 5.8)

### 5.2.7 Software

Analysis was done with base R commands as well as the following R packages as implemented in RStudio: data handling with the **sf** and **tidyverse** packages (Pebesma, 2018; Wickham et al., 2019), computation of the sampling designs using the **spcosa** (Walvoort et al. 2010), **BalancedSampling** (Grafström & Lisic, 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** and **georob** packages (Ribeiro & Diggle, 2018; Papritz, 2020) and handling of spatial objects using the **raster** (Hijmans, 2020) and **rgdal** (Bivand et al., 2020) packages. Graphics were created with base R and the package **ggplot2** (Wickham, 2016).

## 5.3 Results

### 5.3.1 Error variance as a function of the number of calibration samples

Based on the best performing accuracy metrics from the different validations of the PLS calibration regression, we selected the results from the 1st derivative spectra for available K and the standard normal variate spectra for available P. The relationship between the number of calibration samples and the error variance was described by a simple exponential function:

$$\sigma_{n_c}^2 = A + b\omega^{n_c} \quad (5.9)$$

where  $n_c$  is the number of calibration samples used for the regression, and  $A$ ,  $b$  and  $\omega$  are model parameters.

For available K the function reaches its asymptote after around 30 calibration samples, indicating a limit to how accurately available K can be predicted. The error variance for available P shows a more gradual decrease where it appears to reach the asymptote around 80 calibration samples (Fig. 5.1). The estimated parameters

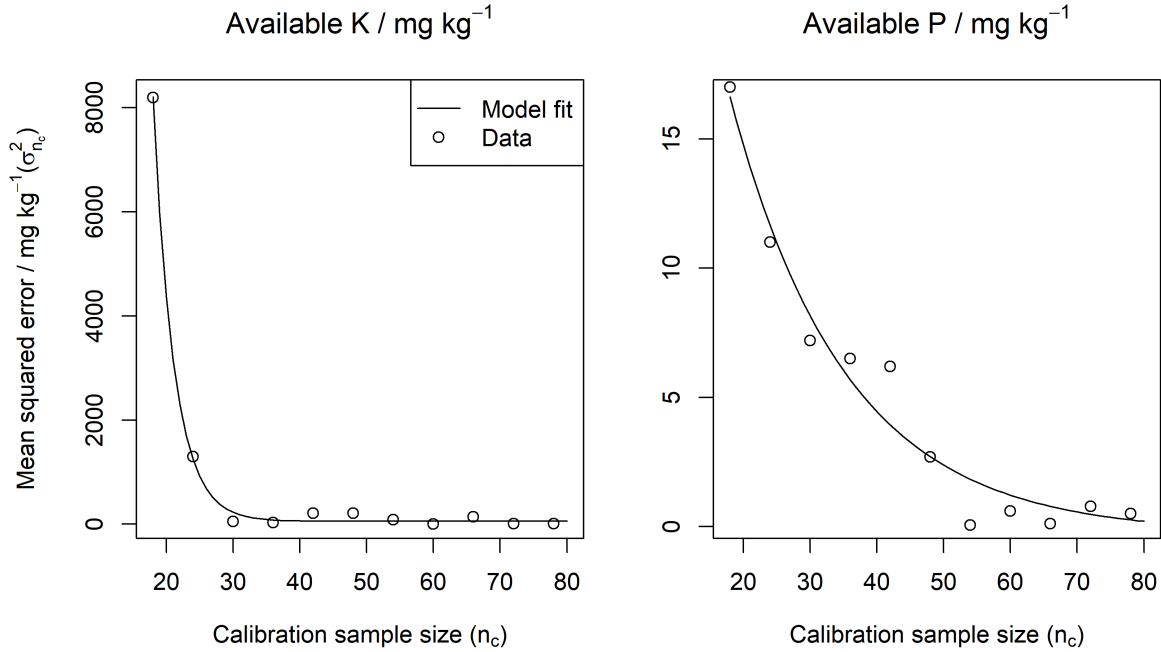


Figure 5.1: The mean squared error ( $\sigma_{n_c}^2$ ) as a function of calibration sample size ( $n_c$ ) as described by Equation 5.9 and the data from

for the exponential equations are given in Table 5.3.

### 5.3.2 Prediction uncertainty under different sample sizes

For available P, an increase in the number of calibration samples ( $n_c$ ) showed a decrease in the mean kriging variance (Fig. 5.2). For the total sample size ( $N_t$ ), there was a similar trend in the kriging variance, following the exponential function

Table 5.3: Parameters for the exponential equation that describes the error variance  $\sigma_{n_c}^2$  as a function of the number of calibration samples ( $n_c$ ) used in regression

Soil property	A	b	$\omega$
P	-0.26	47.93	0.94
K	54.6	2603595	0.73

fitted in the preliminary step (see Equation 5.9 and Fig. 5.1). For a given value of  $n_c$ , the kriging variance is more accurately estimated once  $N_t$  increases. For available K, there is a sharp decrease in the mean kriging variance from 20 to 35 calibration samples, reflecting where Equation 5.9 reaches its asymptote (Fig. 5.2). Similarly

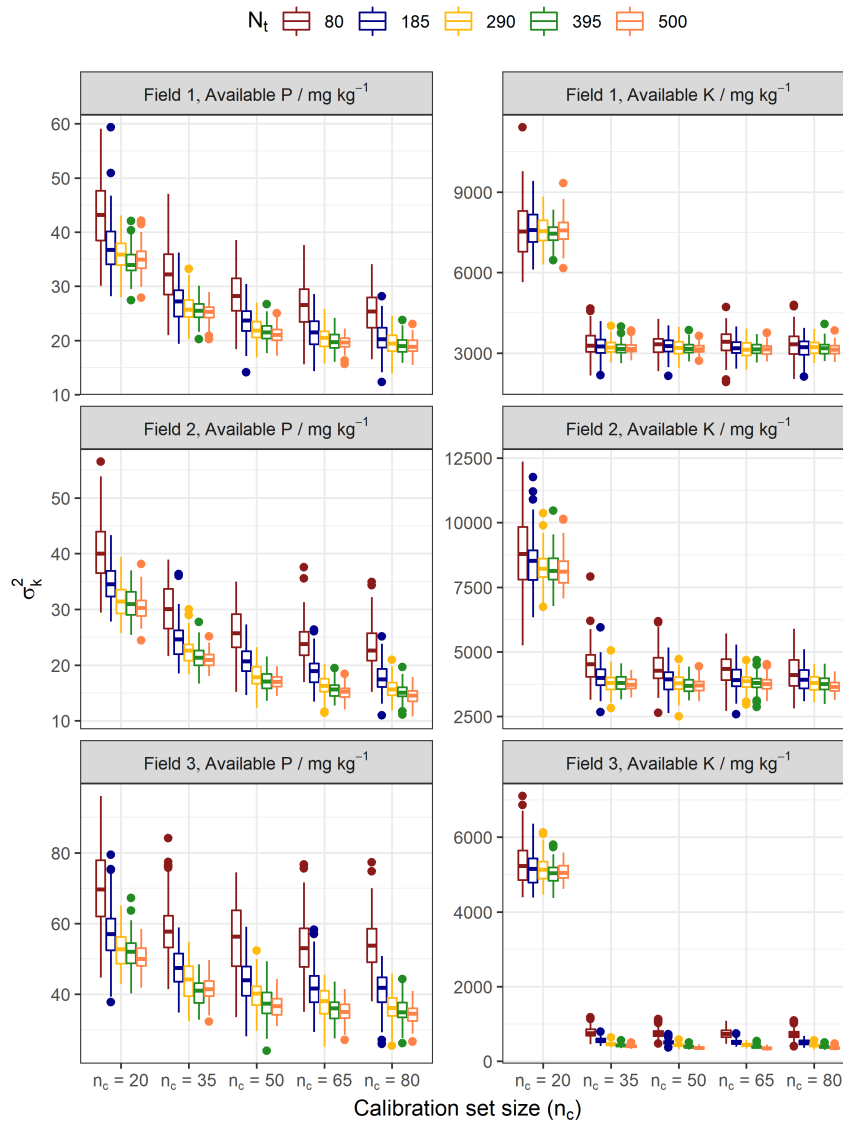


Figure 5.2: Kriging variance ( $\sigma_k^2$ ) distributions for the in-silico simulation results as a function of total sample size ( $N_t$ ) and calibration sample size ( $n_c$ )

to the kriging variance, the estimated nugget parameter declines as  $n_c$  increases. However, nugget variance as a function of  $N_t$  exhibits different behaviour. There are no substantial differences in the mean value of the estimated nugget parameter for different values of  $N_t$ . Across both soil properties and all fields, the variance of the estimated nugget parameter distribution reduces for an increasing number of total sample size ( $N_t$ ) (Fig. 5.3).

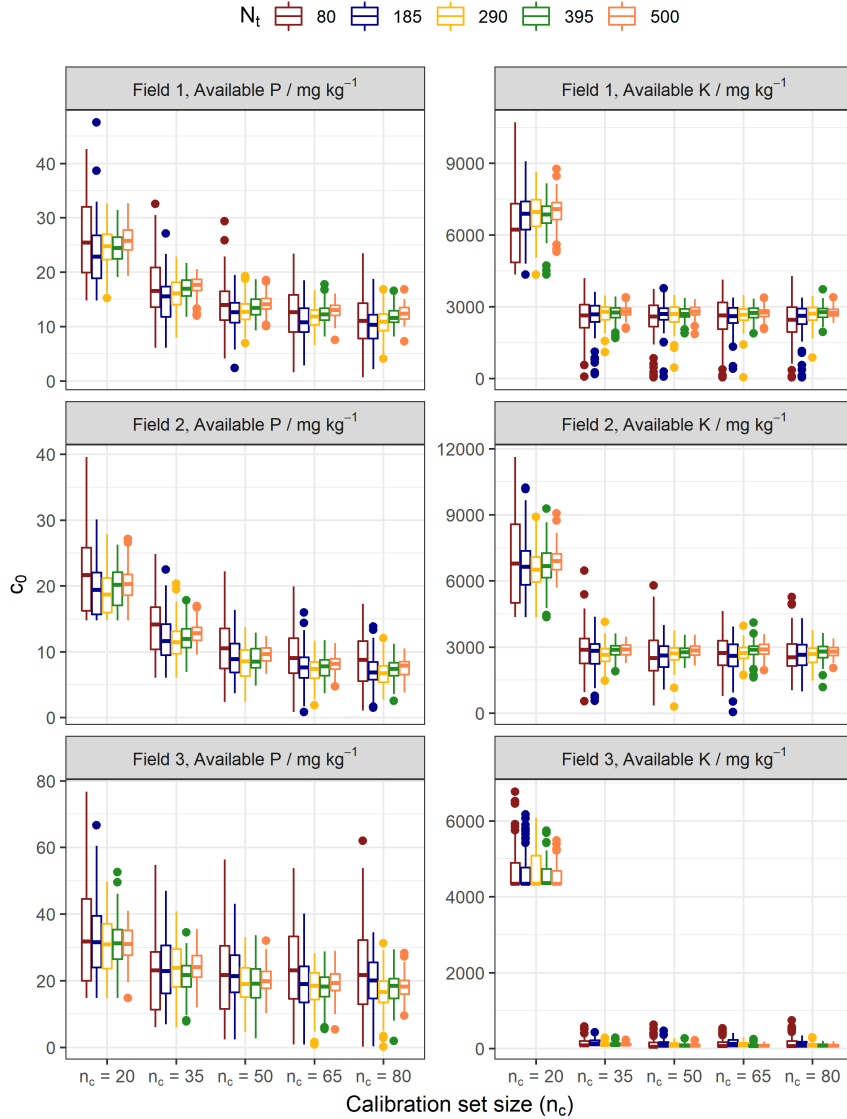


Figure 5.3: Nugget variance ( $c_0$ ) distributions for the in-silico simulation results as a function of total sample size ( $N_t$ ) and calibration sample size ( $n_c$ )

### 5.3.3 Expected loss without accounting for data acquisition costs

The mean expected loss ( $E[L(F_{\text{opt}})] / \mathcal{L} \text{ ha}^{-1}$ ) shows significant differences between the smallest and largest values of total sample size ( $N_t$ ) across all fields for available P. The expected loss for available is less sensitive to the number of samples used for calibration ( $n_c$ ) (Fig. 5.4). For available K, there is a sharp decrease in the expected

loss from 20 to 35  $n_c$  samples. For all fields, the increasing  $N_t$  has less sensitivity on the expected loss. In fields 2 and 3, when  $n_c = 35$  or larger, the expected loss shows a small decrease as  $N_t$  increases.

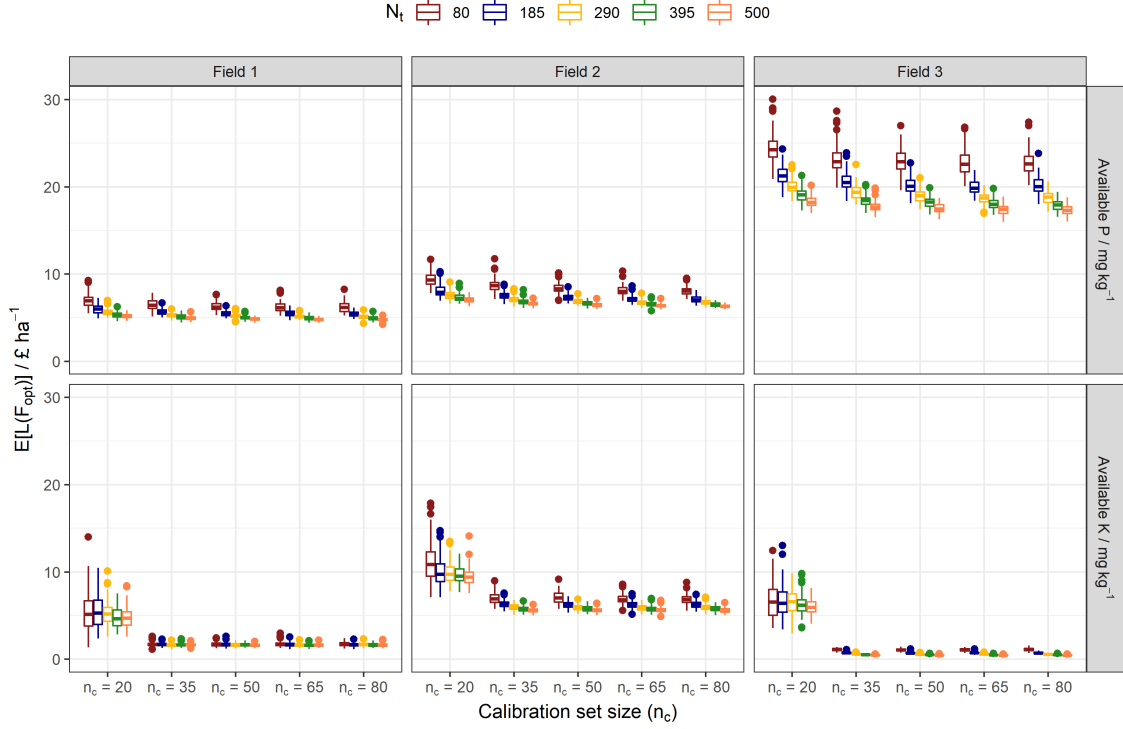


Figure 5.4: Expected loss ( $E[L(F_{\text{opt}})]$ ) distributions for the in-silico simulation results as a function of total sample size ( $N_t$ ) and calibration sample size ( $n_c$ ).

Within our computation of ( $E[L(F_{\text{opt}})] / \text{£ ha}^{-1}$ ) we accounted for the uncertainty in the soil estimates. The probability distribution described by the kriging variance is integrated within the denominator of Equation 5.6 to compute the optimum amount of fertiliser under uncertainty. Generally, accounting for uncertainty reduced the expected loss (compared to  $E[L(F_0)]$ ), although the effect was marginal (Supp. Fig. S9 and S10). Field 2 for available K proved an exception as the reduction of the expected loss by applying  $F_{\text{opt}}$  compared to  $F_0$  was more pronounced (Suppl. Fig. S10).

### 5.3.4 Expected profit when data acquisition costs are taken into account

For each field, the expected profit ( $\Phi(F_{\text{opt}}) / \text{£ ha}^{-1}$ ) from P and potassium fertiliser declines linearly when total- and calibration sample size increases (Fig. 5.5). The linear relation shows that the sampling costs predominate (Equation 5.8) over the potential increases in profit based on the non-linear dose-response curve (Equation 5.4). The slope differs for Field 2 as the overall sampling costs are spread over a larger area.

The last step in our analysis was to apply a scaling factor to the data acquisition costs (of both total- and calibration sample size, Equation 5.8), to assess at which expense the implementation of spectroscopy would become financially viable. These

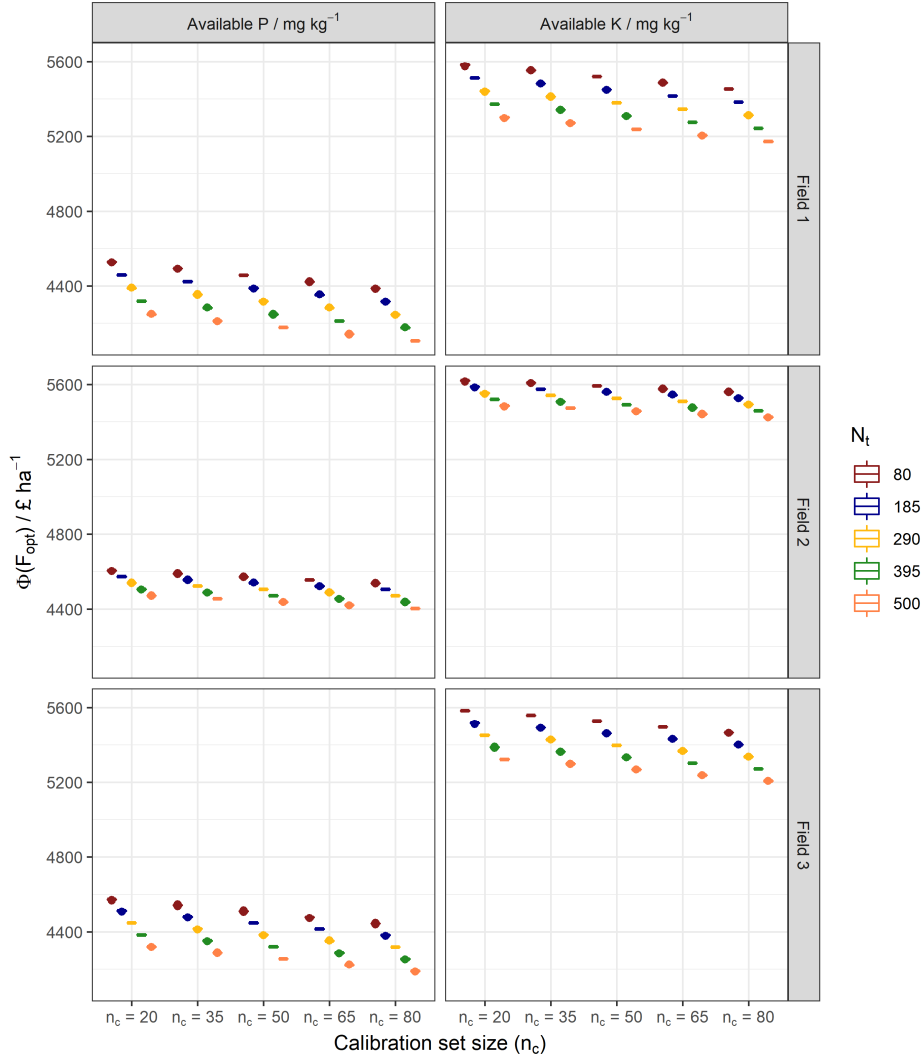


Figure 5.5: Expected profit ( $\Phi(F_{\text{opt}})$ ) distributions for the in-silico simulation results as a function of total sample size ( $N_t$ ) and calibration sample size ( $n_c$ ).

results are shown in Fig. 5.6 where the distribution of  $\Phi(F_{\text{opt}})$  over the range of  $n_c$  was plotted against total sample size ( $N_t$ ). The results showed that for available P and K across all fields, the data acquisition costs would have to fall below 5% of their assumed value in our case study to eliminate the decline in ( $\Phi(F_{\text{opt}})$ ) as function of  $N_t$ . In the scenario of 0.5% of the current data acquisition costs ( $C(N_t, n_c)$ ), sampling by spectroscopy leads to an increase in expected profit in all cases except for available K in Field 1 (Fig. 5.6). We note that the y-axes in Fig. 5.6 do not start at 0 and indeed the effect of sampling on the expected profit is marginal.

## 5.4 Discussion

### 5.4.1 Uncertainty in soil properties predicted from spectroscopy at the field-scale

Our analysis showed that the number of calibration samples has a large effect on the kriging variance. The Kriging variance is less sensitive to total sample size for the



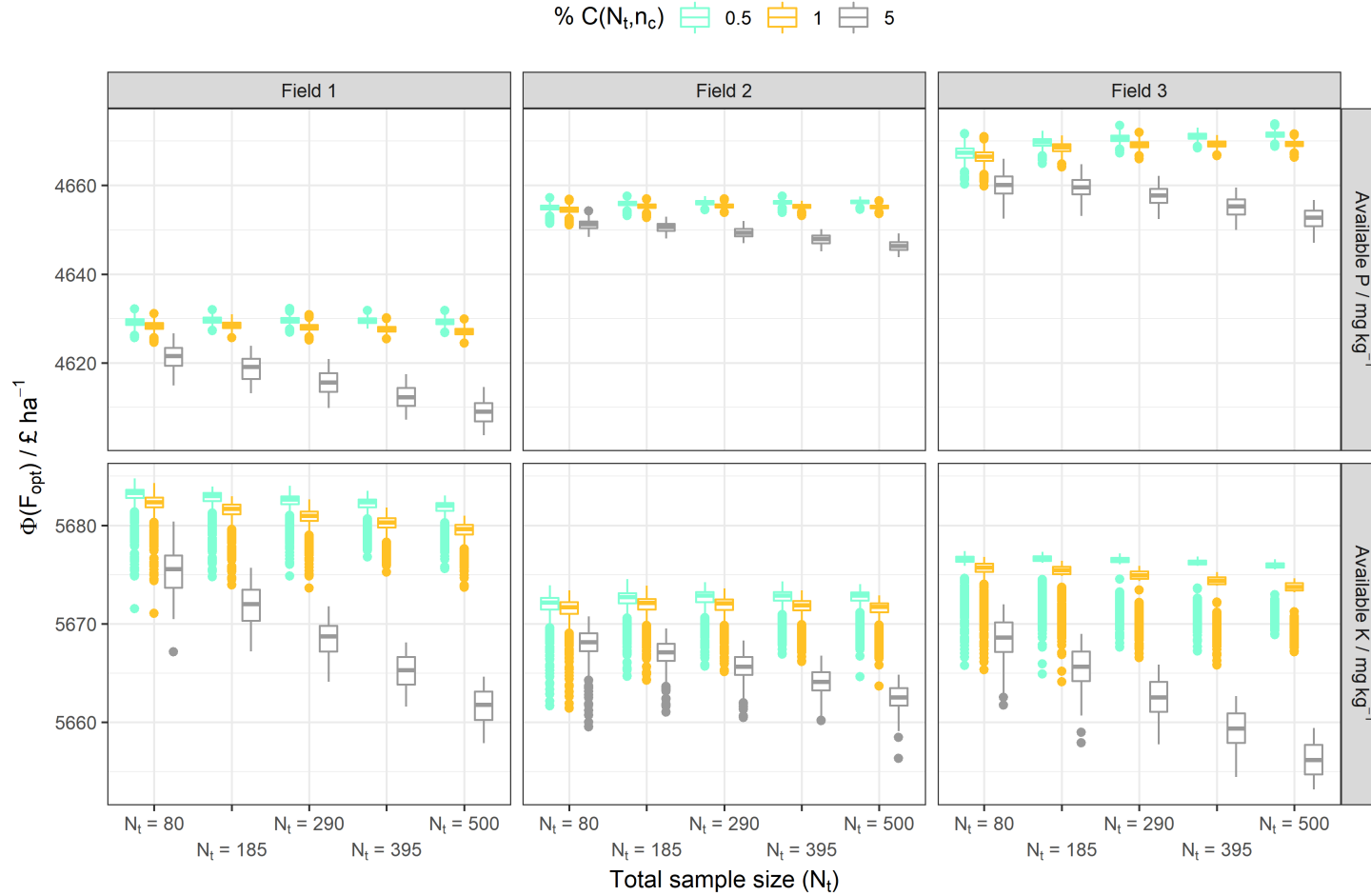


Figure 5.6: Expected profit ( $\Phi(F_{\text{opt}})$ ) distributions for the in-silico simulation results as a function of total sample size ( $N_t$ ) over the range of calibration sample size ( $n_c$ ). Colors indicate the scaling factor applied to the data acquisition costs ( $C(N_t, n_c)$ ), given in % of the original data acquisition costs.

range of sample sizes we selected. These findings contrast to the literature, where the contribution of the spectroscopic model prediction error to the uncertainty of soil property estimates was found to be small relative to the spatial uncertainty (Brodský et al., 2013; Viscarra Rossel et al. 2016). However, these studies were concerned with the prediction of soil organic carbon, which has distinct spectral features in the IR region (Kuang et al., 2012) as opposed to available P and K. Ramirez-lopez et al. (2019) propagated the calibration error through in their mapping of particle-size fractions and exchangeable calcium content and showed that the contribution of the calibration error variance was relatively large, resulting in enhanced smoothing of the kriging predictions due a high nugget variance.

Furthermore, the relationship between uncertainty introduced by the calibration error and the spatial uncertainty is likely to be dependent on the underlying soil variation and the number of samples considered. We considered a range in total sample size from 100 (and above) since this number is generally considered the minimum required to gain a robust estimation of the variogram. Other interpolation approaches (e.g. inverse distance weighting or trend surface analysis) could be considered to allow for fewer total samples. However, these methods do not quantify uncertainty in the soil estimates. As a result, uncertainty in the estimates cannot be accommodated for which adds increased risk to the decision-making. Within the methodological framework used in this study, the distribution over different geostatistical realisations could provide a measure of uncertainty when different interpolation methods are used (Goovaerts, 1997). Geostatistical realisations of the soil variable are drawn from the joint distribution given the spatial variation at sample locations (in this case the predictions from soil spectra). Hence, the use of these realisations would have supported a more realistic approximation of the variance in predictions compared to drawing from the uncorrelated variance component,  $\text{Norm}(0, \sqrt{c_0})$ , only (see Method section 5.2.2).

Across varying numbers  $N_t$ , sampling designs included a fixed number of close-pair points. Close-pair points ensure better estimation of the spatial covariance parameters (Lark & Marchant, 2018; Wadoux et al., 2019). Further investigation can reveal how a variation in the close-pair points will affect the estimation of the variogram. That is, sampling design effects are specific to the underlying variogram model, observed in this study. Depending on the nugget variance of the original variogram (Table 5.1), the effect of total sample-size on the nugget variance parameter was reduced for larger values of  $N_t$ . We attribute this to the fact that the nugget parameter estimate approximated the true underlying short-scale variance.

Another consideration regarding the total sample size is its effect on estimating the underlying trend. The total sample size was computed by a spatial coverage sample design, that allows for robust estimation of the trend parameters (Brus, 2019). Since we removed the trend surface prior to ‘sampling’, the effect of the total sample size on the trend estimation has been ignored. In an actual soil survey however, the difference in how well the trend has been estimated is known to have a large effect on subsequent kriging predictions and the representation of associated uncertainty (Lark, 2009). Consequently we would expect the discrepancy of expected loss between different total sample sizes to increase as the trend surface is approximated with increasing accuracy.

### 5.4.2 The expected loss from informing fertiliser management based on spectroscopy estimates

Overall, the results of the expected loss showed that soil spectroscopy could provide accurate field-scale estimates of available soil nutrient concentration. The expected loss was in the range of  $4.2 - 30 \text{ £ ha}^{-1}$  for available P and  $0.4 - 17.9 \text{ £ ha}^{-1}$  for available K. These values are negligible compared to the mean profit per hectare and thus indicate that estimates for soil spectroscopy are sufficiently accurate to pose minimum risk when used to inform soil management decisions. Furthermore, in both cases (for available P and K) there were diminishing returns on investment for increased sample sizes, indicating that there will be an optimum number for both total and calibration samples. We note however, that the magnitude of the expected loss and resulting calculated optimum are determined by the formulation of the loss function. For example, the true values for available K in fields 1 and 2 were generally above the asymptote of the dose-response curve (Suppl. Fig. S8). Consequently, omitting fertiliser application for large parts of the Field resulted in the largest financial gain. Equally, the asymmetry in the loss function might explain the contrast in the expected loss between  $F_0$  and  $F_{\text{opt}}$  for available K in Field 2 (Suppl. Fig. S9, S10). Given the asymptote in the dose response curve for K fertiliser, risk-averse over-application of fertiliser under uncertainty leads to a lower expected loss.

### 5.4.3 How cost-effective is spectroscopy at the field-scale?

Our results show that under current costs of data acquisition including the sampling procedure, the implementation of spectroscopy was not cost-effective. These findings were supported by a linear decrease in expected profit for a large number of  $N_t$  and  $n_c$  samples. For soil spectroscopy to become cost-effective, the current costs need to reduce by at least 95%.

Breure et al. (2021b) explored the expected loss associated with precise and blanket fertiliser application of P and K for the same fields used in this study. They concluded that the difference in the expected loss between these two fertiliser regimes could be indicative of the allowable expense for a field-survey. The differences in expected loss between these two fertiliser regimes lay in the range of  $\text{£ } 15\text{--}47 \text{ ha}^{-1}$  for available P. The differences in the expected loss for available K lay in the range of  $\text{£ } 0\text{--}15 \text{ ha}^{-1}$ . Given Equation 5.8, the lowest sampling costs in this study are  $\text{£ } 49 \text{ ha}^{-1}$  and  $\text{£ } 47 \text{ ha}^{-1}$  for available P and K, respectively. These values approximate the difference in expected loss between precise and blanket fertiliser of P. However, the lowest sampling cost ( $\text{£ } 49$ ) is based on Field 2 which is almost twice the size of Field 3 that showed a difference of  $\text{£ } 47 \text{ ha}^{-1}$  between precise and blanket fertiliser application.

Our findings hold true under the current case-study assumptions of crop- and fertiliser price, data acquisition costs and the formulation of the loss function. Further studies are required to elaborate on these assumptions. For example, we did not consider a scaling effect of the sampling costs per sample relative to the total sample size. Within a larger geographical area, the variable costs per sample will scale with an increase in total sample size due to reduced travel-time between locations (Lark & Knights, 2015). Equally, the economy of scale might be applicable to the number of samples analysed by wet chemistry. That is, for a larger number of

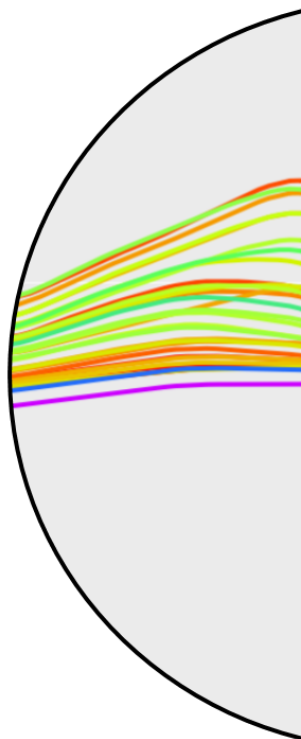
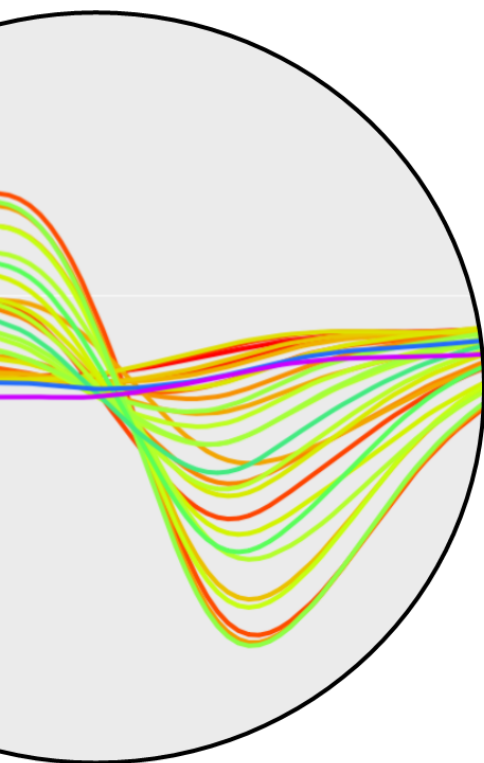
samples a laboratory might charge a lower price per sample. Our results showed a marked decrease in the sampling costs due to the field size. Further work is needed on the scaling effects both in terms of the study area's spatial extent and associated economics. We also assumed that yield is uniquely constrained by the crop's uptake of available P and K. In reality, it is unlikely that crop yield is unambiguously constrained by available P and K concentration only. In reality, crop yield is affected by additional factors such as other soil properties, water availability, disease pressure and weather conditions over the growing season. For example, Lark et al. (2020) developed boundary line models to represent the joint variation of soil nutrient status and crop yield. Their results suggested that the application of P fertiliser should be tailored to variations in soil pH. Where yield is considered dependent on the interaction between two soil variables, the ability of soil spectroscopy to estimate multiple soil properties from a single spectrum might enhance its cost-effectiveness.

## 5.5 Conclusions

Our results showed that the uncertainty in soil property predictions was predominantly determined by the number of samples used for calibration. No combination of total- and calibration sample sizes that we considered would render soil spectroscopy a cost-effective method to inform fertiliser regimes. Although using estimates from spectroscopy led to a relatively small expected loss, the costs of data acquisition dominated the expected profit under the ranges of total- and calibration sample sizes considered. However, the expected loss from using estimates of available soil nutrients from spectroscopy to inform fertiliser applications showed a diminishing return on investment when the costs of data collection were ignored. This suggests that an optimum sample size exists, in case data acquisition costs could be reduced or offset. Overall, our findings showed that the loss function approach can be successfully used as a method to investigate the value of soil spectroscopy for informing decisions on soil management. We therefore suggest an extension to different case-studies to investigate the generalisability of our findings.

# Chapter 6

## Synthesis



## 6.1 Introduction

The overall aim of this thesis was to investigate the uncertainty associated with soil spectral measurements and its effect on soil monitoring and management. Within the literature review, I identified the need to quantify the uncertainty that is introduced under the various stages involved in predicting soil properties by spectroscopy. Another research gap that I identified was the effect of uncertainty on agronomic decision-making. I also discussed the potential of a loss function framework to investigate how the expected loss can be minimised under different configurations of spectroscopy measurements, calibration methods and sampling designs.

The research gaps identified in Chapter 1 were translated into four main questions central to my PhD project. Within this final Chapter, I discuss how my PhD research has contributed to answer these four questions formulated at the start of my project. The research outcomes are also compared with the existing literature and I suggest directions for future research. Finally, I provide the main conclusions based on my PhD research project.

## 6.2 Overview of findings

The first research question was: can soil spectra be used to predict crop performance indicators, given that spectral measurements are inherently uncertain? To investigate this question, I looked at two distinct approaches to predict crop data using soil spectra and what their implications are for management. The first approach used the soil spectra directly to predict the crop data. The second approach involved the calibration of the spectra to estimate soil properties, which were in turn used to predict the crop data. Results showed that the direct approach led to more accurate predictions of crop data compared to the indirect approach. However, estimated soil N, P, K and pH were significant predictors of the crop data along the indirect approach, confirming the utility of soil spectral data to inform precision management. The difference in accuracy between the two approaches quantifies the additional uncertainty by using soil estimates to predict crop growth. Although the direct approach was more accurate, this did not supply an mechanistic understanding on how soil properties might be affecting crop performance. This suggests that direct predictions might be used to inform variable-rate planting when crop performance is consistently linked to soil spectra over multiple growing seasons. The predictions of soil properties can be used to inform precise applications of fertiliser and liming. The study concludes that under optimal conditions, there is potential for associating crop response to soil reflectance spectra.

The second research question was to what extent the effort associated with spectral measurements be reduced in light of its effect on the uncertainty of soil predictions? Chapter 3 involved a study that contrasted the magnitude of loss in accuracy, relative to field-scale predictions on milled samples, by either reducing the sample processing steps or through the use of spectral libraries. Additionally, the predictions were performed for multiple sensors to assess whether their combined effect could minimise the loss in accuracy resulting from reduced sample processing. This chapter showed that both reduced sample processing and spectral libraries potentially could reduce time and cost implications for predicting soil organic carbon,

clay and pH from near- and mid-infrared spectra. Available P and K could only be predicted with moderate accuracy from the milled field-scale samples. The spectral library approach proved more sensitive to specific conditions of its use. For example, different laboratory methods used for the spectral library and the field-scale dataset introduced error. Another finding was that the most effective method with which samples were selected from a national soil database, proved to be specific for each soil property. Combined predictions from multiple sensors generally were of equal accuracy or improved slightly compared to NIR and MIR predictions by themselves. Particularly, results showed that the combined effect was beneficial for the prediction of soil organic carbon from samples measured in the field or under air-dried conditions. The loss of accuracy was specific to the combination of soil property and sensor analysed. The results provided insight into the expected differences in prediction accuracy and which factors need to be taken into consideration to reduce effort for developing field-scale calibrations. I concluded that the appropriate method to use will depend on the accuracy required, the number of prediction samples needed and the costs associated with field sampling, preparation, handling and laboratory costs.

The main research question with regards to Chapter 4 asked: what are the advantages of accounting for uncertainty in soil property predictions when making decisions about soil management? By accounting for uncertainty, I could test the hypothesis whether spatial predictions of available P and K were sufficiently accurate such that precise application of P and K fertiliser was economically and environmentally justified. The effect of uncertainty (compared to using the true values, i.e. the mean kriging predictions) was quantified as an expected financial loss under both uniform and precise fertiliser regimes. For all four fields, there was an economic incentive for precise fertiliser application of P compared to uniform application. In the case of K, economic incentives existed in two out of the four fields. The magnitude of the expected loss and the difference in loss between precise and uniform application was found to be dependent on (a) the kriging variance, (b) the range of the dose-response curve in terms of available P and K, (c) the range of estimated P and K values within the fields and (d) the asymmetry of the loss function. Although uncertainty led to risk-averse fertiliser application in general, less P would be applied under a precise fertiliser regime compared to uniform application for two fields. However, results showed that the environmental benefit of fertiliser savings from precise applications was not always accompanied with a large increase in profit. Hence, it was concluded that the environmental benefits, such as reducing eutrophication of watercourses, from reduced P fertiliser applications should be included in the loss function.

The research question relevant to Chapter 5 asked whether the loss function framework could be used to inform the data acquisition process and sampling for soil property prediction by spectroscopy. To investigate this, I considered the implications of uncertainty in soil nutrient estimates from spectroscopy for informing precise application of P and K fertilizer. Different total- and calibration sample sizes were considered and used to investigate the variation in the kriging variance, expected profit and the expected loss. Furthermore, I applied a scaling factor to the data acquisition costs to explore to what degree sampling costs would need to reduce to make variable rate fertilizer supported by spectroscopy financially viable.

The loss function analysis showed different effects for the sample sizes considered

between soil properties and fields. Although using estimates from spectroscopy led to a relatively small expected loss, the costs of data acquisition dominated the expected profit under the ranges of total- and calibration sample sizes considered. However, the expected loss from using estimates of available soil nutrients from spectroscopy to inform fertiliser applications showed a diminishing return on investment when the costs of data collection were ignored. This suggests that an optimum sample size exists, in case data acquisition costs could be reduced or offset. I concluded that the loss function approach is an appropriate method to assess whether soil spectroscopy is a cost-effective means to inform soil management. Furthermore, I suggested its application in different case-studies to gain more robust insight in the value of applied soil spectroscopy.

## 6.3 Future research

### 6.3.1 Different methods to account for uncertainty

The quantification of uncertainty in soil estimates played a central role within this thesis. Within this section I discuss the knowledge gaps revealed during this thesis research and alternative approaches that can be considered to quantify uncertainty.

One major dependency within uncertainty analysis is deciding which variables are considered uncertain. An example within this thesis can be found in the context of fitting a linear mixed effect model to distinguish which soil properties affect crop growth (Chapter 2). Li et al. (2009) found ignoring measurement error of covariates in maximum likelihood estimation for spatial mixed effect models leads to attenuated regression coefficients and inflated variance components. Hence, within the context of Chapter 2, attenuated regression coefficients could lead to false inferences about which soil properties explain crop growth. For example, error within the georeferencing of crop data might lead to erroneous parameter estimates and hence false inferences from the model. Thus, further studies should investigate the relation between soil spectra and crop data that accounts for the uncertainty in covariate data used in the computation of linear mixed models.

Deciding which variables are considered uncertain equally affects model-based geostatistics. The issue can be illustrated by the example of nugget variance estimation under different calibration predictions. Within this thesis, variogram parameters were estimated based on field-scale predictions only. However, when a spectral library is used, the error source of analytical methods for the soil reference values are generally unknown (Viscarra Rossel et al., 2016). Although the analytical measurement error can be considered small within a single laboratory, the inter-laboratory variation has been shown to be larger than expected (Pleijzier 1986). The results in Chapter 3 confirmed these differences by the increase in calibration error of clay, which was attributed to discrepancy between the analytical method used in the spectral library and the field-scale dataset. Chapters 4 and 5 showed that it was necessary to fix the nugget parameter for a realistic approximation of the calibration error. Given that kriging estimates are most sensitive to unreliable estimates of the nugget variance (Brooker, 1986), it is important to approximate the calibration error (and thus the nugget variance) correctly. Alternative methods for the estimation of covariance parameters can be useful in case the calibration error is unknown. Markov Chain Monte Carlo



(MCMC) methods to estimate spatial covariance parameters have been found to approximate the nugget closer to observed short-range variation in the field (Minasny et al. 2011). Parameter estimation by MCMC methods also allows for inference of the error variance in data from different sources (Samarathna et al., 2018).

Within Chapter 5, we assumed that the calibration errors were spatially independent (Conforti et al. 2015; Viscarra Rossel et al. 2016). However, high concentrations of available P and K could have a larger error due to the smoothing effect of the calibration regression. Within the case-study area in this thesis, high values are associated with pockets of peat within the field and calibration errors would thus exhibit spatial autocorrelation. Different error models might be applicable to quantify the calibration error as a function of the prediction range. For example, treating the error in a multiplicative context (i.e. as the ratio between the prediction and the true value) allows for systematic dependence of the error on the range of the variable (Tian et al. 2013).

#### 6.3.2 Reflections on soil spectroscopy

The second subject this thesis is concerned with is soil spectroscopy. Within this section I discuss the knowledge gaps apparent from research and discuss further investigations that are required.

Chapter 2 reported on the importance of soil N, P, K and pH in the prediction of crop performance across the case-study fields. These soil properties are also relevant to agricultural decision-making in terms of precise fertiliser and liming requirements. A literature review by Sternberg et al. (2010) states that the predictions of some soil properties, such as pH and the contents of plant nutrients, cannot be predicted consistently (Stenberg et al., 2010). The principle explanation for this is that available P and K have weak or no spectral features in the IR region (Kuang et al., 2012). Consequently, they are dependent on their correlation with spectrally active components such as soil C fractions. Another factor is that some nutrients are more mobile than others (e.g. N). Prediction of available N using IR-spectroscopy is challenging in general, because of the volatilization of ammonia during the drying of the soil samples (Wang et al. 2015). In a scenario where a larger set of calibration samples is required, in-situ measurements provide an alternative to reduce costs. Although Chapter 2 reported low prediction accuracies for available P, K and pH under in-field conditions, in-situ XRF measurements were not considered. Current literature does not report studies on in-field sensing of available P and K using a combination of (V)NIR, MIR and XRF spectrometers. Thus further research should investigate how in-situ predictions perform for P, K and pH by using a combination of (V)NIR, MIR and XRF sensors on the field-scale.

Another research gap is the use of XRF spectral libraries for field-scale predictions. Although the performance between sensors is partially case-study specific, inclusion of XRF spectra has been shown to substantially improve prediction accuracy for some soil properties, notably particle size fractions and heavy metal content (Kalnicky and Singhvi, 2001; Carr et al., 2008; Zhang & Hartemink, 2019; Benedet et al., 2020). Hence, research could explore the use of XRF spectral libraries for soil property predictions at the field-scale.

A second aspect on spectral libraries is that, to reduce complexity, I did not

consider a full factorial analysis in Chapter 2. As mentioned in Chapter 1, studies have contrasted approaches based on similarities between spectra to subset a national soil database with a geographic extent/stratification approach (e.g. Seidel et al. 2019). However, few studies have contrasted this with predictions at the field-scale under reduced sample processing conditions (Chapter 2). Another extension to this could be to contrast field-scale prediction under reduced sample processing with predictions based on transfer functions developed using spectral libraries. A transfer function allows for the use of spectral libraries measured on dried and milled soil samples to predict on samples where the spectra have been measured in the field. Literature has reported that standardisation methods can mitigate the effect of soil moisture on spectra measured in the field (Minasny et al., 2011; Ji et al., 2016). However, these methods require a subset that contains spectra from both milled and in-situ samples to establish the transfer function. Thus, a relevant question is how does this contrast in predication accuracy and reduced effort compare to making predictions from in-field measurements (under dry conditions) or spectral libraries combined with spiking (Chapter 2)?

Throughout this thesis, standard methodologies reported in the literature have been used for pre-processing spectra. There is no consensus on the spectral pre-processing methods that lead to the most accurate calibration predictions of soil properties (Dardenne, 2002; Stenberg & Viscarra Rossel, 2010). Within this thesis, the use of XRF spectra, as opposed to the factory calibration on the instrument was required to gain good prediction accuracy (O'Rourke et al. 2016a, 2016b; Zhang & Hartemink, 2019). Within the subsequent analysis of the raw spectra (Chapter 4 and 5) I assumed that similar spectral pre-processing methods used in IR spectroscopy applied to XRF spectra. However, these spectral processing methods have not taken into account the different physical processes that underlie XRF spectroscopy compared to IR. That is, secondary excitations have been recorded to vary in their intensity by a factor of up to 20 and are strongly matrix dependent, requiring correction (Kramar 2000). Matrix effects can be introduced due to physical differences such as the soil's mineralogy, particle size, surface irregularities and moisture content. Equally, chemical matrix effects refer to the interference of certain elements on the XRF spectrum in the presence of other elements. For example, X-ray intensities of Zn are more strongly absorbed in the presence of Fe (Horta et al., 2015). It follows that an IR pre-processing method such as a standard normal variate can be biased towards an element with a high concentration in the sample. Zhang & Hartemink (2019) found little effect of different processing methods applied on XRF spectra in their predictions of particle size, pH, total C and total N. However, further research is required in how pre-processing affects predictions for different soil properties in a wider variety of soil types.

### 6.3.3 Further development of the loss function

I used the loss function to assess soil spectroscopy and the effect of its associated uncertainty on decision-making. Within this section I discuss the knowledge gaps that I identified relevant to the loss function framework and how they might be investigated.

Within Chapters 4 and 5 I introduced the loss function framework to inform

decision-making based on uncertain soil data. Naturally, the decision-making context of applied soil spectroscopy encompasses more trade-offs than dealt within this research. For example, both Chapters 4 and 5 described the combined use of NIR, MIR and XRF spectra. However, there is a need to account for the costs associated with using different sensors and investigate their merit within a loss function framework. For example, XRF measurements require a separate procedure compared to NIR and MIR which both have been measured with the same spectrometer by changing the light source. Research into the merit of different sensors could also utilise the loss function framework to contrast portable/mobile sensors, that typically provide a higher density of samples predicted with reduced accuracy compared to the standard benchtop spectrometers used in this study. Such studies could build on results from Chapter 5 and can provide valuable insight into the trade-off between the cost-effectiveness of mobile sensor platforms compared to their accuracy in the context of decision-making.

Knowledge gaps are not only present in the application of the loss function but also how it has been formulated in the first place. Chapter 4 showed that environmental incentives for precise fertiliser application were not strictly tied to economic incentives. Within Chapter 4, I discussed that the loss function's asymmetry could account for environmental incentives such that a larger penalty is placed on over-application of fertiliser. Such studies can quantify the potential environmental benefits of precise fertiliser applications and perhaps inform the level of required compensation necessary to incentivise sustainable fertiliser practices. That is, generally the farmer is interested in maximising his/her total net income. Some studies reported a lack of economic incentive for precision agriculture practices after taking into account all associated costs (i.e. acquisition of equipment, maintenance and depreciation) (Yang et al., 2001). The fact that environmental and economic incentives for precise fertiliser application do not necessarily coincide (Chapter 4), provides a new perspective to this debate. The inclusion of impacts to the environment is likely to affect the value of precise fertiliser application based on soil spectroscopic information and requires further investigation.

Lastly, the magnitude of the expected loss and difference in loss between the estimates and the true values were found to be dependent on a.) the kriging variance, b.) the range of the dose-response curve in terms of available P and K, c.) the range of estimated P and K values within the fields, and d.) the asymmetry of the loss function. In Chapters 4 and 5 I assumed that yield is unambiguously constrained by a limitation in available P or K, which is not likely to be the case in practice. Further studies could inform the loss function by agricultural field experiments to investigate the effect of precise fertiliser application based on soil spectroscopy. Such an approach would allow for increased certainty that the magnitude of the expected loss is realistic since the loss function has been tailored in line with the soil under study. Within such a study, an extension could be to perform a sensitivity study on the important parameters that define the loss function to gain further insight in the effect of uncertainty on the expected loss.

The loss function analysis in Chapter 5 showed that total sample size (density) exerted a large effect on the expected loss. The *a priori* decision of sampling density requires further research. Kerry and Oliver (2008) described a method to estimate an empirical variogram from covariate data. An empirical variogram allows for the identification of an approximate spatial scale of variation. The range of the spatial

autocorrelation can then in turn be used to determine an appropriate sampling density, as described first by McBratney & Webster (1986). Another approach is to compute an average variogram from a database of variogram parameters (Wadoux et al., 2019). Further studies could investigate both approaches and use the loss function framework to assess the effect of their recommended sampling density on prediction accuracy and decision-making.

A second topic of research that is related to the *a priori* decision on the sampling design, is how to inform the placement of the samples. Whether to adopt design-, model-assisted or model-based sampling, is dependent on the objective of the study. Viscarra Rossel et al. (2016) found that although the total organic C stocks estimated by the three different sampling designs were similar, the variances of both the designs informed by a model were smaller compared to the probability-based approach. The loss function framework could in such cases be used to investigate whether the difference in uncertainty is relevant to the study's objective. In that particular case, what is the minimum variance in organic C stock estimates that allows for the detection a difference in organic C stock in subsequent monitoring? However, model-based estimates are generally associated with higher expenses for sampling, thus the loss function framework might support a study that contrasts different sampling design practices and their accuracy within a decision-making context (e.g. organic C stock change).

#### 6.3.4 Generic reflection on this thesis research

The effect of the soil's natural variability within a case-study on the research findings cannot be reduced by improved model structure (Chapters 2,3,4,5), calibration of the model parameters (Chapters 2,3,4,5) and the resolution of simulations (Chapter 5). The question is then how we can generalise from case-study results to the overall empirical advances from the research presented in this thesis? The Cambridgeshire fens are characterised by large sub-field scale variability due to their specific soil formation history. There are stark contrasts between peat soils and alluvial/marine silts within each field. Sub-field scale variability determines to what extent the use of soil sensors is applicable. A suitable range of calibration is required to gain accurate predictions (Wetterlind et al., 2008). It follows that sub-field scale variability also determines what minimum accuracy of the sensor is required to measure variation in soil properties. That is, the measurement error should be smaller than the short-range variation of the target soil property (de Gruijter et al. 2010). These main characteristics determine if field-scale calibrations are likely to be effective.

The composition of the soil also affects whether certain properties can be accurately predicted. For example, a diverse set of phosphorus binding mechanisms in soil determine its availability for plant uptake. Ahmed et al. (2019) characterised orthoP and molecular mechanisms for the adsorption of P to goethite in a purely chemical system (without spectral interference of soil components). Available P as analysed in this thesis, however, is not a chemically defined component and its relation with spectrally active components is soil type specific. Given that the prediction accuracy of P is dependent on its relation to spectrally active soil components, the variety of P binding mechanisms in soils thus leads to a wide variety of surrogate correlations other soil components and hence variable results

between case-studies (Chang et al., 2001; Pätzoldt et al., 2019). For example, this case-study soil's contain large variation in, generally high, organic matter content. However, other studies have reported indications that available P and K could be predicted accurately only for soils low in organic matter (Hu et al., 2016). Differences in prediction accuracy reported between studies show the wide constraints of the soil's composition and its relative effect on the value of soil spectroscopy. In some cases, these constraints can be approximated for certain soil properties. For example, the mineralogy of soils in Eastern and Southern Africa is constrained by weathering processes (Shepherd and Walsh, 2002). In other cases, constraints are imposed by anthropogenic factors, such as the application of fertiliser. However, fertilisation can also disturb the surrogate correlations on which the prediction accuracy of non spectrally active soil properties is dependent. To determine whether a non-spectrally active soil property can be predicted accurately prior to a soil survey is a difficult and an understudied question. Thus, it is important that the research presented within this thesis should be viewed in context of the soil's variability and spatial characteristics of the case-study.

In addition to the soil's natural variability, another characteristic of a case-study area is its spatial extent. Consequently some of the conclusions might change once applied on a broader geographical scale. For example, wet chemistry analysis costs could be subject to the economy of scale within the loss function framework (Chapter 5). As mentioned within Chapter 5, we also did not consider a scaling effect of the sampling costs per sample relative to the total sample size. Within a larger geographical area, the variable costs per sample will scale with an increase in total sample size due to reduced travel-time between locations (Lark & Knights, 2015).

Given that the effectiveness of spectroscopy is dependent on the soil's natural variability and spatial extent of the case-study area, its application appears to be context-specific for its case-study area. The question on whether soil sensing is viable to implement in the first place is an important and difficult question to answer due to these factors. However, the quantification of uncertainty and understanding its effect in context of the study's objective plays a central role towards a further understanding of the value of soil spectroscopy.

## 6.4 Conclusions

Literature reviews on soil spectroscopy have provided useful insight in the prediction accuracy for a wide range of soil properties under different sample preparation, calibration data collection, spectrum acquisition and assessment of calibration accuracies. The previous sections showed that these methodologies indeed all influence how effectively spectroscopy can estimate soil properties and the magnitude of associated uncertainty. However, to date, few studies have looked beyond accuracy metrics and asked what precision is required within the context of soil management. I believe that this thesis has made a substantial contribution in quantifying and understanding the effect of uncertainty from soil spectroscopy within the context of soil management.

The models formulated within this thesis are statistical and hence give a relativist and scale-dependent measure of the soil's variation. The question of generalisability accompanies most of the studies in soil science, given the complex natural variability of soil. The challenge is to move beyond the characteristics of a case-study. Although

spectroscopy might be able to provide accurate predictions, the required accuracy is context-dependent and thus it is unrealistic to assume it will be effective in all cases. The main two questions to answer are i.) whether soil spectroscopy will be effective, and if so, ii.) how should it be implemented?

As shown, the loss function framework can be used to decide whether predictions from spectroscopy are adequate and when they might not be. The results indicate that spectroscopy as a method can benefit soil management but was restricted in its application by associated costs. These findings emphasise the importance of looking beyond spectroscopy as a technique on its own and assess its relevance for characterising soils and the value of knowledge obtained.

In order to validate research funds allocated to soil spectroscopy, the value of information acquired needs to be assessed in light of the study's objective. The research in this thesis has presented such work within an applied spectroscopy context. Given the centrality of soil to many environmental processes, an exciting variety of topics lie ahead to assess the potential of soil spectroscopy to enhance research.

# References

- Adamchuk, V.I., Christenson, P.T. 2007. Development of an instrumented blade system for mapping soil mechanical resistance represented as a second-order polynomial. *Soil and Till. Res.*, 95, 76–83.
- Adamchuk, V.I., Viscarra Rossel, R.A. 2010. *Development of On-the-Go proximal soil sensor systems*. In: R.A. Viscarra Rossel, A.B. McBratney, B. Minasny (Eds.), *Proximal soil sensing* 15–28. New York, NY: Springer-Verlag.
- Adamchuk, V.I., Viscarra Rossel, R.A., Marx, D.B., Samal, A.K. 2008. Enhancement of on-the-go soil sensor data using guided sampling. In: R. Kholsa (Ed.), *Proceedings of the ninth international conference on precision agriculture*, Denver, Colorado, July 20–23, 2008. Fort Collins: Colorado State University (CD publication, 13 pages).
- Adamchuk, V.I., Viscarra Rossel, R.A., Marx, D.B., Samal, A.K. 2011. Using targeted sampling to process multivariate soil sensing data. *Geoderma*, 163, 63–73.
- Ahmed, A.A., Gypser, S., Leinweber, P., Freese, D., Kühn, O. 2019. Infrared spectroscopic characterization of phosphate binding at the goethite–water interface. *Phys. Chem.*, 21, 4421–4434.
- ASD Accessories User Manual, CrossRef. Accessed at: 16-09-202.
- Avery, B.W., Bascomb, C.L., 1982. Soil survey laboratory methods. *Soil survey technical monograph No. 14*, Harpenden, United Kingdom.
- Bänninger, D., Lehmann, P., Flühler, H. (2006). Modelling the effect of particle size, shape and orientation of light transfer through porous media. *European journal of soil science*, 57(6), 906–915.
- 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.
- Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M., McBratney, A., 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *Trends Analyt. Chem.* 29, 1073–1081.
- 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 Biol. Biochem.* 43, 1398–1410.
- Benedet, L., Faria, W.M., Silva, S.H.G., Mancini, M., Demattê, J.A.M., Guilherme, L.R.G., Curi, N. 2020. Soil texture prediction using portable X-ray fluorescence spectrometry and visible near-infrared diffuse reflectance spectroscopy. *Geoderma*, 376, 114553.
- Beusen, A.H.W., Bouwman, A.F., Van Beek, L.P.H., Mogollón, J.M., Middelburg, J.J. 2016. Global riverine N and P transport to ocean increased during the 20th century despite increased retention along the aquatic continuum. *Biogeosc.*, 13, 2441–2451.
- Bivand, R., Keitt, T., Rowlingson, B. 2020. *rgdal: Bindings for the 'Geospatial' Data Abstraction Library.* R package version 1.5-12. <https://CRAN.R-project.org/package=rgdal>. Accessed at 8 December 2020.
- Bjørn-Helge, M., Wehrens, R., Hovde Liland, K., 2019. *pls: Partial Least Squares and Principal Component Regression.* R package version 2.7-1. <https://CRAN.R-project.org/package=pls>. Accessed at 05-02-2020.
- 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.
- Bogrekci, I., Lee, W.S., 2006. Effects of soil moisture content on absorbance spectra of sandy soils in sensing phosphorus concentrations using UV-VIS-NIR spectroscopy. *Trans. of the ASABE* 49, 1175–1180.
- Boon, K.A., Rostron, P., Ramsey, M.H. (2011). An exploration of the interplay between the measurement uncertainty and the number of samples in contaminated land investigations. *Geostandards and Geoanalytical Research*, 35(3), 353-367.
- Börjesson, T., Stenberg, B., Lindén, B., Jonsson, A. 1999. NIR spectroscopy, mineral nitrogen analysis and soil incubations for the prediction of crop uptake of nitrogen during the growing season. *Plant and soil*, 214(1), 75–83.
- Bowers, S.A., Hanks, R.J., 1964. Reflection of radiant energy from soils. *Soil Sc.*, 100, 130–138
- Breure, T., Milne, A.E., Webster, R., Haefele, S.M., Hannam, J.A., Moreno-Rojas, S., Corstanje, R., 2021a. Predicting the growth of lettuce from soil infrared reflectance spectra: the potential for crop management. *Prec. Agr.* 22, 226–248.
- Breure, T.S., Haefele, S.M., Hannam, J.A., Corstanje, R., Webster, R., Moreno-Rojas, S., Milne, A.E. 2021b. A loss function to evaluate agricultural decision-making under uncertainty: a case study of soil spectroscopy. *Precision Agriculture – Under review*



- 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.
- Brooker, P.I. 1986. A parametric study of robustness of kriging variance as a function of range and relative nugget effect for a spherical semivariogram. *Math. Geol.*, 18(5), 477–488.
- Brown, D.J., 2007. Using a global VNIR soil-spectral library for local soil characterization and landscape modelling in a 2nd-order Uganda watershed. *Geoderma* 140, 444–453.
- 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.
- Brus, D.J., Yang, L., Zhu, A.X., 2019. Accounting for differences in costs among sampling locations in optimal stratification. *Eur. J. of Soil Sc.*, 70(1), 200–212.
- Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface water with phosphorus and nitrogen. *Ec. Appl.*, 8, 559–568.
- Carr, R., Zhang, C., Moles, N., Harder, M. 2008. Identification and mapping of heavy metal pollution in soils of a sports ground in Galway City, Ireland, using a portable XRF analyser and GIS. *Env. Geoch. and Health*, 30(1), 45–52.
- Chang, C.W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R. 2001. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Sc. Soc. of Am. J.*, 65(2), 480–490.
- Chen, M., Graedel, T.E. 2016. A half-century of global phosphorus flows, stocks, production, consumption, recycling, and environmental impacts. *Glob. Environ. Chang.*, 36, 139–152.
- 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. *Sc. Tot. Environ.*, 406, 1–23.
- Christy, C.D., Collings, K., Drummond, P., Lund, E. 2004. A mobile sensor platform for measurement of soil pH and buffering. St. Joseph: American Society of Agricultural and Biological Engineers.
- Christy, C.D. 2008. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Computers and Electronics in Agriculture*, 61, 10–19.
- Clairotte, M., Grinand, C., Kouakoua, E., Thébault, A., Saby, N.P.A., Bernoux, M., Barthès, B.G., 2016. National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. *Geoderma* 276, 41–52.

- Clark, I. 2010. Statistics or geostatistics? Sampling error or nugget effect? *J. South. Afr. Inst. Min. Metall.*, 110(6), 307–312.
- Cressie, N. 1990. The origins of kriging. *Math. Geol.*, 22(3), 239–252.
- Lucà, F., Conforti, M., Castrignanò, A., Matteucci, G., Buttafuoco, G. 2017. Effect of calibration set size on prediction at local scale of soil carbon by Vis-NIR spectroscopy. *Geoderma*, 288, 175–183.
- Coutinho, M.A.N., Alari, F.D.O., Ferreira, M.M.C., do Amaral, L.R., 2019. Influence of soil sample preparation on the quantification of NPK content via spectroscopy. *Geoderma* 338, 401–409.
- Dangal, S.R.S., Sander, J., Wills, S., Ramirez-Lopez, L. 2019 Accurate and precise prediction of soil properties from a large mid-infrared spectral library. *Soil Systems*, 3
- Dardenne, P., Sinnaeve, G., Baeten, V. 2000. Multivariate calibration and chemometrics for near infrared spectroscopy: which method? *J. of Near Infr. Spectr.*, 8(4), 229–237.
- de Gruijter, J.J., McBratney, A.B., Taylor, J. 2010. *Sampling for high resolution soil mapping*. In: R.A. Viscarra Rossel, A.B. McBratney, B. Minasny (Eds.), *Proximal soil sensing* 3–14. New York, NY: Springer-Verlag.
- de Gruijter, J.J., McBratney, A.B., Minasny, B., Wheeler, I., Malone, B.P., Stockmann, U. 2018. *Farm-scale soil carbon auditing*. In: *Pedometrics* 693–720. Springer, Cham.
- Deutsch, C.V., and A.G. Journel. 1998. *GSLIB Geostatistical Software Library and User's Guide*. 2<sup>nd</sup> Ed. Oxford University Press, New York, 119–122.
- Dhawale, N.M., Adamchuck, V.I., Prasher, S.O., Viscarra Rossel, R.A., Ismail, A.A., Kaur, J., 2015. Proximal soil sensing of soil texture and organic matter with a prototype portable mid-infrared spectrometer. *Eur. J. Soil Sc.* 66, 661–669.
- Diacono, M., Rubino, P., Montemurro, F. 2013. Precision nitrogen management of wheat. A review. *Agronomy for Sustainable Development*, 33, 219–241.
- Di Stefano, C., Ferro, V., & Mirabile, S. 2010. Comparison between grain-size analyses using laser diffraction and sedimentation methods. *Biosystems Engineering*, 106, 205–215.
- Du, C., Ma, F., Lu, Y., & Zhou, J. 2015. Soil fertility assessed by infrared spectroscopy. In: Lal, R., & Stewart, B.A., *Soil-specific Farming*. Boca Raton, FL: CRS Press, 155–177.
- Du, C., Zhou, J. 2009. Evaluation of soil fertility using infrared spectroscopy: a review. *Environmental Chemistry Letters*, 2, 97–113.
- Ellinger, M., Merbach, I., Werban, U., Ließ, M. 2019. Error propagation in spectrometric functions of soil organic carbon. *SOIL*, 5, 275–288.

- Evans, J.S. 2019. spatialEco. R package version 1.2-0,  
URL: <https://github.com/jeffrejevans/spatialEco>. Accessed at: 03-10-2019.
- Faechner, T., Pyrcz, M., Deutsch, C.V. 2000. Soil remediation decision making in presence of uncertainty in crop yield response. *Geoderma*, 97(1-2), 21-38.
- Farewell, T.S., Truckell, I.G., Keay, C.A., Hallett, S.H., 2011. The derivation and application of Soilscales: soil and environmental datasets from the National Soil Resources Institute. Cranfield University. <https://www.landis.org.uk>. Accessed at: 08-04-2020.
- Fisher, P., Aumann, C., Chia, K., O'Halloran, N., & Chandra, S. 2017. Adequacy of laser diffraction for soil particle size analysis. *PLoS ONE* 12, e0176510.
- Ge, Y., Morgan, C.L.S., Grunwald, S., Brown, D.J., Sarkhot, D.V., 2011. Comparison of soil reflectance spectra and calibration models obtained using multiple spectrometers. *Geoderma* 161, 202–221.
- Gebbers, R., de Bruin, S. 2010. Application of geostatistical simulation in precision agriculture. In: Oliver, M.A. (Editor) *Geostatistical Applications for Precision Agriculture*. Cham: Springer, 269–304.
- Goovaerts, P, 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Goovaerts, P. 2001. Geostatistical modelling of uncertainty in soil science. *Geoderma*, 103, 3–26.
- Grafström, A., & Lisic, J. (2019) BalancedSampling: Balanced and Spatially Balanced Sampling. R package version 1.5.5.
- Granger, C.W.J., Ramanathan, R. 1984. Improved methods of combining forecasts. *J. of Forec.*, 3, 197–204.  
<https://CRAN.R-project.org/package=BalancedSampling>. Accessed at: 03-10-2019.
- Gras, J.P., Barthès, B.G., Mahaut, B., Trupin, S., 2014. Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils. *Geoderma*, 214–215, 126–134.
- 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. *J. Agr. Sci.*, 95, 441–456.
- Grunwald, S., Congrong, Y.U., Xiong, X. 2018. Transferability and scalability of soil total carbon prediction models in Florida, USA. *Pedosphere*, 28(6), 856–872.
- Guerrero, C., Zornoza, R., Gómez, I., Mataix-Beneyto, J., 2010. Spiking of NIR regional models using samples from target sites: Effect of model size on prediction accuracy. *Geoderma* 158, 66–77.

- Guerrero, C., Stenberg, B., Wetterlind, J., Viscarra Rossel, R.A., Maestre, F.T., Mouazen, A.M., Zornoza, R., Ruis-Sinoga, J.D., Kuang, B., 2014. Assessment of soil organic carbon at local scale with spiked NIR calibrations: effects of selection and extra-weighting on the spiking subset. *Eur. J. Soil Sc.* 65, 248–263.
- Guerrero, C., Wetterlind, J., Stenberg, B., Mouazen, A.M., Gabarron-Galeote, M.A., Ruiz-Sinoga, J.D., Zornoza, R., Viscarra Rossel, R.A., 2016. Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy? *Soil Tillage Res.* 155, 501–509.
- Haaland, D.M., & Thomas, E.V. 1988. Partial least-squares methods for spectral analyses: 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry*, 60, 1193–1202.
- Hastie, T., Friedman, J., Tibshirani, R. 2009. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd printing. Springer, Stanford, CA.
- Hermansen, C., Knadel, M., Moldrup, P., Greve, M.H., Gislum, R., de Jonge, L.W., 2016. Visible-Near-Infrared spectroscopy can predict the clay/organic carbon and mineral fines/organic carbon ratios. *Soil Sc. Soc. Am. J.* 80(6), 1486–1495.
- Heuvelink, G.B.M. 2018. Uncertainty and uncertainty propagation in soil mapping and modelling. In: McBratney, A.B., Minasny, B., Stockmann, U. (editors) *Pedometrics*. Cham: Springer, 439–461.
- Hijmans, R.J. 2020. raster: Geographic Data Analysis and Modeling. R package version 3.3-13. <https://CRAN.R-project.org/package=raster>. Accessed at 8 December 2020.
- 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.
- Hodge, A.M., Sudduth, K.A. 2012. Comparison of two spectrometers for profile soil carbon sensing. *ASABE*, 121338240.
- Horta, A., Malone, B., Stockmann, U., Minasny, B., Bishop, T. F. A., McBratney, A. B., Pozza, L. 2015. Potential of integrated field spectroscopy and spatial analysis for enhanced assessment of soil contamination: a prospective review. *Geoderma*, 241, 180–209.
- Hu, G., Sudduth, K.A., He, D., Myers, D.B., Nathan, M.V. 2016. Soil phosphorus and potassium estimation by reflectance spectroscopy. *Trans. of the ASABE*, 59(1), 97–105. <https://doi.org/10.13031/trans.59.11048>.
- Hutengs, C., Ludwig, B., Jung, A., Eisele, A., Vohland, M., 2018. Comparison of portable and bench-top spectrometers for mid-infrared diffuse reflectance measurements of soils. *Sensors* 18, 993.

- Hutengs, C., Seidel, M., Oertel, F., Ludwig, B., Vohland, M., 2019. In situ and laboratory soil spectroscopy with portable visible-to-near infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma* 355, 113900.
- IUSS Working Group WRB, 2015. World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.
- Ji, W., Adamchuck, V.I., Biswas, A., Dhawale, N.M., Sudarsan, B., Zhang, Y., Viscarra Rossel, R.A., Shi, Z., 2016. Assessment of soil properties *in situ* using a prototype portable MIR spectrometer in two agricultural fields. *Biosyst. Eng.* 152, 14–27.
- 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), article e105708.
- Ji, W., Viscarra Rossel, R.A., Shi, Z., 2015. Accounting for the effects of water and the environment on proximally sensed vis-NIR soil spectra and their calibrations. *Eur. J. Soil Sc.* 66, 555–565.
- Johnson, J.M., Vandamme, E., Senthilkumar, K., Sila, A., Shepherd, K.D., Saito, K., 2019. Near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for assessing soil fertility in rice fields in sub-Saharan Africa. *Geoderma*, 354, 113840.
- Journel, A., 1984. mAD and conditional quantile estimators. In: Verly, G., David, M., Journel, A.G., Marechal, A. (editors) *Geostatistics for Natural Resources Characterization*. D. Reidel, Dordrecht, 261–270 (Part 1).
- Kalembasa, S.J., Jenkinson, D.S., 1973. A comparative study of titrimetric and gravimetric methods for determination of organic carbon in soils. *J. Sc. Food Agr.* 24, 1085–1090.
- Kalnicky, D.J., Singhvi, R., 2001. Field portable XRF analysis of environmental samples. *J. Hazard Mater.*, 83, 93–122.
- Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. *Technometr.* 11, 137–148.
- Kindred, D.R., Sylvester-Bradley, R., 2014. Using precision farming technologies to improve nitrogen management and empower on-farm learning. *Aspects of Applied Biology*, 127, 173–180.
- Kindred, D.R., Sylvester-Bradley, R., Milne, A.E., Marchant, B., Hatley, D., Kendall, S.L., Clarke, S., Storer, K., Berry, P.M., 2017. Spatial variation in Nitrogen requirements of cereals, and their interpretation. *Advances in Animal Bioscience: Precision Agriculture*, 8:2

- Knadel, M., Stenberg, B., Deng, F., Thomsen, A., Greve, M.H. 2013. Comparing predictive abilities of three visible-near infrared spectrophotometers for soil organic carbon and clay determination. *J. of Near Infrared. Spectrosc.* 21, 67–80.
- Knox, N.M., Grunwald, S., McDowell, M.L., Bruland, G.L., Myers, D.B., Harris, W.G. 2015. Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy. *Geoderma*, 239, 229–239.
- Konert, M., & VandenBerghe, J. 1997. Comparison of laser grain size analysis with pipette and sieve analysis: a solution for the underestimation of the clay fraction. *Sedimentology*, 44, 523–535.
- Kramar, U. 2000. X-ray Fluorescence Spectrometers. In: Lindon, J.C., Tranter, G.E., Holmes, J.L. (Eds) *Encyclopedia of Spectroscopy and Spectrometry*, Academic Press, 2467–2477.
- Kruskal, W., Mosteller, F. 1979. Representative Sampling, III: the Current Statistical Literature. *Int. Stat. Rev.* 47, 245–265.
- Kuang, B., Mahmood, H.S., Quraishi, M.Z., Hoogmoed, W.B., Mouazen, A.M., van Henten, E.J., 2012. Sensing soil properties in the laboratory, in situ and on-line. A review. *Adv. Agr.* 114, 155–223.
- Lark, R.M. 2000. Estimating variograms of soil properties by the method-of-moments and maximum likelihood. *Eur. J. Soil Sci.*, 51, 717–728.
- Lark, R.M. 2009. Kriging a soil variable with a simple non-stationary variance model. *J. Agr., Biol., Env. Stat.*, 14, 301–321.
- Lark, R.M., Cullis, B.R. 2004. Model-based analysis using REML for inference from systematically sampled data on soil. *European Journal of Soil Science*, 55, 799–813.
- 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. *Eur. J. Soil Sci.*, 57, 787–799.
- Lark, R.M., Gillingham, V., Langton, D., Marchant, B.P. 2020. Boundary line models for soil nutrient concentrations and wheat yield in national-scale datasets. *Eur. J. of Soil Sc.*, 71(3), 334–351.
- Lark, R.M., Knights, K.V. 2015. The implicit loss function for errors in soil information. *Geoderma*, 251-252, 24–32.
- Lark, R.M., Marchant, B. 2018. How should a spatial coverage sampling design for a geostatistical soil survey be supplemented to support estimation of spatial covariance parameters? *Geoderma*, 3019, 89–99.
- Laurent, A.G. 1963. The lognormal distribution and the translation method: description and estimation problems. *J. Am. Stat. Assoc.*, 58, 231–235.

- Le Guillou, F., Wetterlind, W., Viscarra Rossel, R.A., Hicks, W., Grundy, M., Tuomi, S., 2015. How does grinding affect the mid-infrared spectra of soil and their multivariate calibrations to texture and organic carbon? *Soil Res.* 53(8), 913–921.
- Lesch, S.M. 2005. Sensor-directed spatial response surface sampling designs for characterizing spatial variation in soil properties. *Comp. and Electr. in Agr.*, 46(1), 153–180.
- Liu, X., Dong, X., Leskovar, D. I. 2016. Ground penetrating radar for underground sensing in agriculture: a review. *International Agrophysics*, 30(4).
- 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. *Env. Sci. Tech.*, 49, 4980–4987.
- Li, Y., Tang, H., Lin, X. 2009. Spatial linear mixed models with covariate measurement errors. *Stat. Sin.*, 19, 1077–1093.
- Lin, L., 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometr.* 56, 255–268.
- Lobsey, C.R., Viscarra Rossel, R. A., Roudier, P., Hedley, C. B. 2017. rs-local data-mines information from spectral libraries to improve local calibrations. *Eur. J. of Soil Sc.*, 68(6), 840–852.
- Lopo, M., Páscoa, R.N.M.J., Graca, A.R., Lopes, J.A. 2016. Classification of vineyard soils using portable and benchtop near-infrared spectrometers: a comparative study. *Soil Sci. Soc. Am. J.* 80, 652–661.
- MAFF 1986. The analysis of agricultural materials. MAFF Reference Book 427, 3<sup>rd</sup> edition. HMSO, London.
- MAFF, 1988. Fertiliser recommendations for agricultural and horticultural crops. MAFF Reference Book 209. HMSO, London.
- 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. Available at: <https://cereals.ahdb.org.uk/media/252469/pr485.pdf>.
- Marchant, B.P., McBratney, A.B., Lark, R.M., Minasny, B. (2013). Optimized multi-phase sampling for soil remediation surveys. *Spatial Statistics*, 4, 1-13.
- McBratney, A.B., Webster, R. 1986. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *J. of Soil Sc.*, 37(4), 617–639.
- McBratney, A.B., Whelan, B.M., Viscarra Rossel, R.A., 1996. Spatial prediction for precision agriculture. In ‘Proceedings of the 3rd International Conference on Precision Agriculture’. Bloomington/Minneapolis, Minnesota. ASA/CSSA/SSSA. 331–42.

- McGrath, S.P., Loveland, P.J., 1992. The soil geochemical atlas of England and Wales. Blackie Academic & Professional, London, 101.
- Mevik, B.H., Wehrens, R., & Liland, K.H. 2019. *pls: Partial Least Squares and Principal Component Regression*. R package version 2.7-1. <https://CRAN.R-project.org/package=pls>. Accessed at: 04-10-2019.
- 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, 278 pp. (CEH: Project Report Number C02275), Wallingford, United Kingdom. <http://nora.nerc.ac.uk/3370/>.
- 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, 80, 17—30.
- Minasny, B., McBratney, A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comp. and Geosc.*, 32(9), 1378–1388.
- Minasny, B., McBratney, A.B., Bellon-Maurel, V., Roger, J.-M., Gobrecht, A., Ferrand, L, Joalland, S., 2011. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma* 167–168, 118–124.
- Minasny, McBratney, A.B., Walvoort, D.J.J. 2007. The variance quadtree algorithm: Use for spatial sampling design. *Comp. and Geosc.*, 33(3), 383–392.
- Minasny, B., McBratney, A.B., Pichon, L., Sun, W., Short, M.G., 2009. Evaluating near infrared spectroscopy for field prediction of soil properties. *Austr. J. Soil Res.* 47(7), 664–673.
- Minasny, B., Vrugt, J.A., McBratney, A.B., 2011. Confronting uncertainty in model-based geostatistics using Markov chain Monte Carlo simulation. *Geoderma*, 163 (3–4), 150–162.
- Moura-Bueno, J.M., Dalmolin, R.S.D., Horst-Heinen, T.Z., Ten Caten, A., Vasques, G.M., Dotto, A.C., Grunwald, S. 2020. When does stratification of a subtropical soil spectral library improve predictions of soil organic carbon content?. *Sc. of the Tot. Env.*, 737, 139895.
- 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. *Comp. Electr. Agr.*, 135, 163–174.
- Nduwamungu, C., Ziadi, N, Tremblay, G.F., Parent, L.E., 2009. Near-infrared reflectance spectroscopy prediction of soil properties: effects of sample cups and preparation. *Soil Sc. Soc. Am. J.* 73, 1896–1903.



- Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., McBratney, A.B. 2019. Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma*, 352, 251–267.
- Nocita, M., Stevens, A., Toth, G., Panagos, P., van Wesemael, B., Montanarella, L. 2014. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least squares regression approach. *Soil Biol. Biochem.* 68, 337–347.
- Oliver, M.A. (Ed.). 2010. *Geostatistical applications for precision agriculture*. Springer Science & Business Media.
- Oliver, M.A., Webster, R. 1990. Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information System*, 4(3), 313-332.
- Olsen, S., Cole, C., Watanabe, F., Dean, L., 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. USDA Circular Nr 93, US Gov. Print. Office, Washington, D.C.
- O'Rourke, S.M., Minasny, B., Holden, N.M., McBratney, A.B., 2016a. Synergistic use of Vis-NIR, MIR and XRF spectroscopy for the determination of soil geochemistry. *Soil Sc. Soc. Am. J.* 80, 888–899.
- O'Rourke, S.M., Stockmann, U., Holden, N.M., McBratney, A.B., Minasny, B. 2016b. An assessment of model averaging to improve predictive power of portable vis-NIR and XRF for the determination of agronomic soil properties. *Geoderma*, 279, 31–44.
- Papritz, A., 2020. *georob: Robust Geostatistical Analysis of Spatial Data*. R Package version 0.3-13. <https://CRAN.R-project.org/package=georob>
- Patterson, H.D., Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58, 545–554.
- Pätzold, S., Leenen, M., Frizen, P., Heggemann, T., Wagner, P., Rodionov, A. 2019. Predicting plant available phosphorus using infrared spectroscopy with consideration for future mobile sensing applications in precision farming. *Prec. Agr.*, 1–25.
- Paul, S.S., Coops, N.C, Johnson, M., Krzic, M., Smukler, S.M. 2019. Evaluating sampling efforts of standard laboratory analysis and mid-infrared spectroscopy for cost effective digital soil mapping at field scale *Geoderma* 356, in press.
- Payne, R.W. (editor) 2018. *The Guide to GenStat Release 19 – Part 2: Statistics*. Hemel Hempstead, UK: VSN International.
- Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R J.*, 10(1), 439–446.

- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team 2018. *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-137, URL: <https://CRAN.R-project.org/package=nlme>. Accessed at: 06-09-2019.
- Pleijssier, L.K. 1986. *The laboratory methods and data exchange programme*. Interim report on the Exchange Round 85-2.
- Prasad, M., Spiers, T.M., Ravenwood, I.C. 1988. Target phosphorus soil test values for vegetables. New Zealand J. Exp. Agr., 16, 83–90.
- 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. Env. Sci. Tech., 37, 201–208.
- Pringle, M.J., Cook, S.E., McBratney, A.B. 2004a. Field-scale experiments for site-specific crop management Part I: Design considerations. Prec. Agr., 6, 617–624.
- Pringle, M.J., McBratney, A.B., Cook, S.E. 2004b. Field-scale experiments for site-specific crop management. Part II: A geostatistical analysis. Prec. Agr., 6, 625–645.
- Proctor, M.E., Siddons, P.A., Jones, R.J.A. , Bellamy, P.H., Keay, C.E. 1998. LandIS – a land information system for the UK. In: Land information systems: developments for planning the sustainable use of land resources, (Eds) Heineke, H.J., Eckelmann, W., Thomasson, A.J., Jones, R.J.A., Montanarella, L., Buckley, B., European Soil Bureau Research Report No. 4 EUR 17729 (EN). Office for Official Publication of the European Communities Luxembourg, 219–233.
- QGIS Development Team 2019. QGIS Geographic Information System. Open Source Geospatial Foundation. URL <http://qgis.org>. Accessed at 05-02-2020.
- R Core Team 2018. *R: A language and environment for statistical computing*. Vienna: Foundation for Statistical Computing. URL <https://www.R-project.org/>. Accessed at: 03-10-2019.
- Ramirez-Lopez, L., Behrens, T., Schmidt, K., Stevens, A., Demattê, J.A.M., Scholten, T. 2013. The spectrum-based learner: A new local approach for modeling soil vis–NIR spectra of complex datasets. Geoderma, 195, 268–279.
- Ramirez-Lopez, L., Stevens, A., 2016. Regression and similarity evaluation for memory-based learning in spectral chemometrics. R package version 1.2.2. <https://CRAN.R-project.org/package=resemble>. Accessed at 05-02-2020.
- 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. Eur. J. Soil Sci., 70, 378–393.
- Ramsey, M. H., Taylor, P. D., Lee, J. C. 2002. Optimized contaminated land investigation at minimum overall cost to achieve fitness-for-purpose. Journal of Environmental Monitoring, 4(5), 809–814.

- Redfield, A.C. 1958. The biological control of chemical factors in the environment. *Am. Sci.*, 46, 205–221. [www.jstor.org/stable/27827150](http://www.jstor.org/stable/27827150).
- Reeves, J.B., 2010. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? *Geoderma* 158, 3–14.
- Ribeiro Jr, P.J., Diggle, P.J. 2018. *geoR: Analysis of geostatistical data*. R package version 1.7-5.2.1. <https://CRAN.R-project.org/package=geoR>. Accessed at 5 February 2020.
- Kerry, R., Oliver, M. A. 2008. Determining nugget: sill ratios of standardized variograms from aerial photographs to kriging sparse soil data. *Prec. Agr.*, 9(1), 33–56.
- Sandford, S.A., Allamandola, L.J., 1990. The physical and infrared spectral properties of CO<sub>2</sub> in astrophysical ice analogs. *The Astrophysical Journal*, 355, 357–372.
- Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analyt. Chem.* 36, 1627–1639.
- Schmidt, K., Behrens, T., Daumann, J., Ramirez-Lopez, L., Werban, U., Dietrich, P., Scholten, T. 2014. A comparison of calibration sampling schemes at the field-scale. *Geoderma*, 232, 243–256.
- Seidel, M., Hutengs, C., Ludwig, B., Thiele-Bruhn, S., Vohland, M., 2019. Strategies for the efficient estimation of soil organic carbon at the field scale with vis-NIR spectroscopy: Spectral libraries and spiking vs. local calibrations. *Geoderma* 354, 113856.
- Schoumans, O.F., Chardon, W.J., Bechmann, M.E., Gascuel-Oudoux, 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. *Sc. Tot. Environ.*, 468–469, 1255–1266.
- Shaw, R., Lark, R.M., Williams, A.P., Chadwick, D.R., Jones, D.L. 2016. Characterising the within-field scale spatial variation of nitrogen in a grassland soil to inform the efficient design of in-situ nitrogen sensor networks for precision agriculture. *Agr., Ecosyst., & Env.*, 230, 294–306. <https://doi.org/10.1016/j.agee.2016.06.004>.
- Shenk, J.S., Westerhaus, M.O., Berzaghi, P. 1997. Investigation of a LOCAL calibration procedure for near infrared instruments. *J. of Near Infr. Spectr.*, 5(4), 223–232.
- Shepherd, K.D., Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. *Soil Sc. Soc. Am. J.*, 66, 988–998.
- Sila, A., Shepherd, K., Pokhariyal, G. (2016) Evaluating the utility of mid-infrared spectral subspaces for predicting soil properties. *Chemometrics and Intelligent Laboratory Systems*, 153, 92–105.

- Sinfield, J. V., Fagerman, D., and Colic, O. 2010. Evaluation of sensing technologies for on- the-go detection of macro-nutrients in cultivated soils. *Comput. Electron. Agric.*, 70, 1–18.
- Smil, V., 2000. Phosphorus in the environment: natural flows and human interferences. *Ann. Rev. Ener. Envir.*, 25, 53–88.
- 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. *Sci. Tot. Envir.*, 631–632, pp. 377–389.
- Soriano-Disla, J.M., Janik, L.J., Viscarra-Rossel, R.A., Macdonald, L.M., McLaughlin, M.J. 2014. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Applied Spectroscopy Reviews*, 49(2), 139-186.
- Spragg, R.A. 2000. IR Spectrometers. In: Lindon, J.C., Tranter, G.E., Holmes, J.L. (Eds) *Encyclopedia of Spectroscopy and Spectrometry*, Academic Press, 1048–1057.
- Stenberg, B., Jonsson, A., Börjesson, T. 2005. Use of near infrared reflectance spectroscopy to predict nitrogen uptake by winter wheat within fields with high variability in organic matter. *Plant and Soil*, 269(1), 251–258.
- Stenberg, B., Viscarra Rossel, R.A. 2010. Diffuse Reflectance Spectroscopy for High-Resolution Soil Sensing. In: Viscarra Rossel, R.A., McBratney, A., Minasny, B. (Eds) *Proximal Soil Sensing*, Springer, 29–47.
- Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and near in- frared spectroscopy in soil science. *Adv. Agron.* 107, 163–215.
- Stevens, A., & Ramirez-Lopez, L. 2013. *An introduction to the prospectr package*. R package Vignette R package version 0.1.3, URL: <https://cran.r-project.org/package=prospectr>. Accessed at: 03-10-2019.
- Stevens, A., Nocita, M., Tóth, G., Montanarella, L., van Wesemael, B., 2013. Prediction of soil organic carbon at the European scale by visible and near infrared reflectance spectroscopy. *PLoS One* 8, e66409.
- Stevens, A., van Wesemael, B., Bartholomeus, H., Rosillon, D., Tychon, B., Ben-Dor, E., 2008. Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma*, 144, 395–404.
- Stevenson, M., Nunes, T., Heuer, C., Marshall, J., Sanchez, J., Thornton, R., Reiczigel, J., Robison-Cox, J., Sebastiani, P., Solymos, P., Yoshida, K., Jones, G., Prikahu, S., Firestone, S., Kyle, R., Popp, J., Jay, M., Reynard, C., 2020. epiR: Tools for the Analysis of Epidemiological Data. R package version 1.0-14. <https://CRAN.R-project.org/packag=epiR>. Accessed at: 05-04-2020.
- Sudduth, K.A., Hummel, J.W. 1996. Geographic operating range evaluation of a NIR soil sensor. *Trans. of the ASABE*, 39(5), 1599–1604.

- Taylor, J.A., McBratney, A.B., Viscarra Rossel, R.A., Minansy, B., Taylor, H., Whelan, B., et al. Development of a multi-sensor platform for proximal soil sensing. 18th World congress of soil science. Philadelphia, PA: World Congress of Soil Science.
- Terhoeven-Urselmans, T., Schmidt, H., Joergensen, R.G., Ludwig, B. 2008. Usefulness of near-infrared spectroscopy to determine biological and chemical soil properties: Importance of sample pre-treatment. *Soil Biol. and Biochem.*, 40(5), 1178–1188.
- Terra, F.S., Demattê, J.A.M., Viscarra Rossel, R.A., 2015. Spectral libraries for quantitative analysis of tropical Brazilian soils: Comparing NIR and mid-IR reflectance data. *Geoderma* 255–256, 81–93.
- Tian, Y., Huffman, G.J., Adler, R.F., Tang, L., Sapiiano, M., Maggioni, V., Wu, H. 2013. Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geoph. Res. Letters*, 40, 2060–2065.
- Towett, E.K., Shepherd, K.D., Sila, A., Aynekulu, E., Cadisch, G. 2015. Mid-infrared and total x-ray fluorescence spectroscopy complementarity for assessment of soil properties. *Soil Sc. Soc. of Am. J.*, 79(5), 1375–1385.
- Van Groenigen, J.W., Mutters, C.S., Horwath, W.R., Van Kessel, C. 2003. NIR and DRIFT-MIR spectrometry of soils for predicting soil and crop parameters in a flooded field. *Plant and Soil*, 250(1), 155–165.
- Viscarra Rossel, R.A. 2008. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, 90, 72–83.
- Viscarra Rossel, R.A., Adamchuk, V.I., Sudduth, K.A., McKenzie N.J., & Lobsey, C. 2011. Proximal soil sensing. An effective approach for soil measurements in time and space. *Advances in Agronomy*, 113, 237–282.
- Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., & Skjemstad, J.O. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131, 59–75.
- Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Dematte, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuck, V., Aichi, H., Barthes, B.G., Bartholomeus, H.M., Bayer, A.D., Bernoux, M., Bottcher, K., Brodsky, L., Du, C.W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Ramirez-Lopez, L., Roudier, P., Rufasto Campos, E.M., Sanborn, P., Sellitto, V.M., Sudduth, K.A., Rawlins, B.G., Walter, C., Winowiecki, L.A., Hong, S.Y., Ji, W., 2016. A global spectral library to characterize the world's soil. *Earth-Sc. Rev.* 155, 198–230.
- Viscarra Rossel, R.A., Bouma, J. 2016. Soil sensing: A new paradigm for agriculture. *Agr. Syst.* 148, 71–74.

- 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.
- Viscarra Rossel, R.A., Cattle, S.R., Ortega, A., Fouad, Y., 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma* 150, 253–266.
- Viscarra Rossel, R.A., McBratney, A.B., 1998. Soil chemical analytical accuracy and costs: implications from precision agriculture. *Austr. J. Exp. Agric.* 38, 765–775.
- Viscarra Rossel, R.A., Webster, R., 2012. Predicting soil properties from the Australian soil visible-near infrared spectroscopic database. *Eur. J. Soil Sci.* 63, 848–860.
- Wadoux, A.M.C., Marchant, B.P., Lark, R.M. 2019. Efficient sampling for geostatistical surveys. *Eur. J. of Soil Sc.*, 70(5), 975–989.
- 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. *Comp. Geosci.*, 36, 1261–1267.
- Wang, Y., Huang, T., Liu, J., Lin, Z., Li, S., Wang, R., Ge, Y. 2015. Soil pH value, organic matter and macronutrients contents prediction using optical diffuse reflectance spectroscopy. *Computers and Electronics in Agriculture*, 111, 69–77.
- Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T., 2007. In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. *Soil Sc. Soc. Am. J.* 71(2), 389–396.
- Webster, R., Lark, R.M. 2012. Field sampling for environmental science and management. Routledge.
- Webster, R., Oliver, M.A. 1992. Sample adequately to estimate variograms of soil properties. *J. of Soil Sc.* 43(1), 177–192.
- Webster, R., Oliver, M.A. 2007. *Geostatistics for Environmental Scientists*. 2nd edition. Wiley & Sons, Chichester, England.
- Webster, G.T., Soriano-Disla, J.M., Kirk, J., Janik, L.J., Forrester, S.T., McLaughlin, M.J., Steward, R.J., 2016. Rapid prediction of total petroleum hydrocarbons in soil using a hand-held mid-infrared field instrument. *Talanta* 160, 410–416.
- Weiss, C.E., Roetzer, G.R. 2016. *GeomComb: (Geometric) Forecast Combination Methods. R package version 1.0*. URL : <https://CRAN.R-project.org/package=GeomComb>. Accessed at: 05-03-2020.

- Wetterlind, J., Stenberg, B., 2010. Near-infrared spectroscopy for within-field soil characterization: small local calibrations compared with national libraries spiked with local samples. *Eur. J. Soil Sc.* 61(6), 823–843.
- Wetterlind, J., Stenberg, B., Jonsson, A. 2008. Near infrared reflectance spectroscopy compared with soil clay and organic matter content for estimating within-field variation in N uptake in cereals. *Plant and Soil*, 302(1), 317–327.
- Wickham, H., 2016. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York. <https://ggplot2.tidyverse.org>. Accessed at: 14-02-2020.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D'A., François, R., et al., 2019. Welcome to the tidyverse. *J. Open Source Softw.*, 4, 1686.
- Wijewardane, N.K., Ge, Y., Sanderman, J., Ferguson, R. 2020. Fine grinding is needed to maintain the high accuracy of MIR spectroscopy for soil property estimation. *Soil Sc. Soc. Am.*
- Wu, C.-Y., Jacobson, A.R., Laba, M., Baveye, P.C., 2009a. Alleviating moisture content effects on the near-infrared diffuse-reflectance sensing of soils. *Soil Sci.* 174 (8), 456–465.
- Wu, C.-Y., Jacobson, A.R., Laba, M., Baveye, P.C., 2009b. Surface roughness and near- infrared reflectance sensing of soils. *Geoderma* 152 (1–2), 171–180.
- Yang, C., Everitt, J.H., Bradford, J.M., 2001. Comparisons of uniform and variable rate nitrogen and phosphorus fertilizer applications for grain sorghum. *Trans. Am. Soc. Agric. Eng.*, 44 (2), 201–209.
- Yang, H., Mouazen, A. M. 2007. Vis/Near- and Mid-infrared spectroscopy for predicting soil N and C at a farm scale. *Infrared Spectroscopy – Life and Biomedical Sciences*, 10, 672–677.
- Yates, F. 1981. *Sampling Methods for Censuses and Surveys*, 4th edition. Griffin, London.
- Zhang, Y., Hartemink, A.E. 2020. Data fusion of vis–NIR and PXRF spectra to predict soil physical and chemical properties. *Eur. J. of Soil Sc.*, 71(3), 316–333.

# Acknowledgements

In June 2017, I opened the door of a small accommodation situated close to St Pancras station. It was sweltering hot that day with over 30°C. The fact that I had obtained the key successfully could be considered a small miracle. It required the involvement of two e-mails, a phone call, picking up a key at a different address and the unlocking three different key storages. A truly 21<sup>st</sup> century experience; to stay in an AirBnB that is managed remotely from an unspecified location. Regardless, I could relax and take my time before I would have my interview at Rothamsted the next morning.

Although I had limited experience in soil science, spectroscopy and spatial statistics, the description of the project intrigued me. It would provide a challenge to charter new territory and acquire valuable research skills. The last few years have indeed been a wild journey of field-, laboratory- and desktop work. First and foremost I would like to write a word of thanks to my supervisory team that has guided me throughout the PhD research project.

Alice, throughout the project I have been inspired by your undistinguishable enthusiasm. To my surprise this didn't even decline after we spent two hours driving to the case-study area when in practice they were working the field, which meant that we had to abandon our plans. You were always a cure to the intimidating aspects of the PhD research, explaining concepts when I did not grasp them. Although it was sometimes hard to keep up with your demands (i.e. my office companions Jonah and Xavier found it funny to hear you come in and ask; "Are you done yet?" at increasingly smaller time increments), it made me take the work seriously in the light of your expectations. Overall, besides many 'fried brains' and 'red herrings', you transferred the excitement and satisfaction of how mathematics can be successfully used to describe and make sense of the natural environment.

Ron, you stimulated me to engage with the broader context of the research. Why do we ask the questions we're asking? What is the novelty of this approach? I appreciated your directness, labelling my writing at times as 'unexciting'. It pushed me to improve my writing, reasoning or analysis in case they were unsystematic (or all of them together?). Although taking the work very seriously you were never shy of making a joke and showed that you thought of the research value beyond the project's boundaries.

Stephan, your background in agronomy and soil science provided a valuable link within the project. Spectroscopy was new for both of us and it was valuable to discuss concepts and inspiring to hear what new projects were taken up in the AfSIS laboratory. Multiple times I invaded into your office, disturbed your work, and asked your advice (sometimes in panic) on a specific methodology about sample processing, laboratory measurements or spectroscopy. However, you were always more than happy to provide advice and allow me to rest ensured.



Jack, your passion for soil science came through within the meetings, field-work and discussions. Unfortunately we missed out on a Dokuchaev style digging of a soil pit and subsequent pedological description its profile. However, I fondly remember your excitement to get out hydrochloric acid in the Cranfield meeting room to test for the presence of  $\text{CaCO}_3$ . You were always there to present a fresh view on the work and gave me the impression that you are generally well organised, something that I should definitely improve on (as I have shown a myriad of times throughout this PhD).

Although officially not entitled as a supervisor I want to extend thanks to Richard. Your eagle eye ensured that not a single typo would escape (all credits go to me for the typo mistakes that are undoubtedly hidden away in this thesis). We had engaging discussions on the content and more often than not I found myself challenged to provide a satisfying answer. By the hand of your revisions, I started using L<sup>A</sup>T<sub>E</sub>X and have become much more aware of using concise, clear and well-written English (think Shakespeare!). I hope to develop these skills further in my future career.

Overall, I consider myself very lucky and can confidently state that conducting a PhD research is ‘easy’ when you have such a diverse, skilled and friendly supervisory team to work with.

Once I settled in at Rothamsted research and had gotten familiar with the project, it turned out that the case-study area held a surprise. That particular area of the UK was the most similar to the Dutch landscapes which I grew up in as it could possibly get. Although it didn’t strictly satisfy the criterium of an ‘exotic’ case-study area, I told myself that this should motivate myself to apply more interesting analysis and research based on the sampling effort.

Thus, team work continued beyond the supervision into the realm of the field and laboratory. I want to thank all the sampling superstars that helped in setting up the designs or handle an auger to collect the vast amount of peat soil; Alice Milne, Beatrice Landoni, Cathy Thomas, Chris Stephens, Hannah McGrath, Helen Metcalfe, Jacqueline Hannam, Javier Espinosa Montiel, Jonah Prout, Nikolaos Vavlas, Robbie Manning-Smith, Sajeev EM, Sergio Moreno Rojas, Stephan Haefele and Xavier Albano. Collecting the number of samples would have been impossible without your help and, although tiring, it turned out to be a fun day every time.

Many people have also assisted in the laboratory work and associated activities with the data collection. Gifty Acquah and Cathy Thomas from the AfSIS laboratory, Chris Hall from the sample processing facilities, Sarah Dunham, Javier Hernandez and Joanna Carter from the SAS group, in particular. I also want to thank the systems modelling group for their feedback, discussions and jokes during the meetings. From the Cranfield side, I want to thank Ian Truckell for the introduction to the Agri-EPI sensors. Furthermore your help was invaluable during the several events where I required the soil sensors last minute, to shoot off into the Cambridgeshire fens again to take measurements. I also thank Toby Waine and Gill Drew for their comments and attention during the annual reviews.

Of course next to work relations there were also a multitude of social interactions at Rothamsted. Notwithstanding tea time, climbing, foot- and basketball and a variety of parties and casual chats. In particular, Ewan, Nikos, Jonah, Xavi, Chris, Javi, Chiara, Emilio, Maca, Sajeev, John, Robbie, Tadesse, Cathy and Gifty your

company was greatly appreciated, needed and overall a lot of fun!

On a similar note, I want to thank my family; Bram, Sophie, Annet and Sander for their invaluable and endless support throughout and allowing me to be where I am today.

Bea, simultaneously going through the PhD has been a life-changing experience. Whether we struggled through a desert of lettuces or were looking for small populations of *Linnun bienn*e along the Pembrokeshire coast, our research in the UK imposed some weird, unexpected adventures. You have naturally been such a strong support for me throughout the project and I hope to have been the same to you. You provide a safe haven to relax, test ideas and share anecdotes beyond all else.

While taking the train out of London into the Hertfordshire countryside it promised to be another warm day. At least it would be less obvious that I was actually nervous. After the interview I went for a walk along the Rothamsted Manor and prepared myself for the long commute to continue my Msc research internship in Bangor, Wales. As it turns out, I was given the chance to work on the project, for which I am grateful. Now, after having spent over four years in the United Kingdom, including my time in Bangor, I will set sail for a return to mainland Europe. A new chapter awaits. Although currently a blank page independent of this thesis, it will be filled based on the foundation of scientific skills that I have acquired over the course of this PhD.

# Funding

Rothamsted Research receives grant aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the United Kingdom. This research was funded by the Rothamsted-Cranfield Soil Agriculture Research and Innovation Accelerator. AEM and SMH are funded by the Institute Strategic Programme (ISP) grants, “Soils to Nutrition” (S2N) grant number BBS/E/C/000I0330, and the joint Natural Environment Research Council (NERC) and Biotechnology and Biological Sciences Research Council (BBSRC) ISP grant “Achieving Sustainable Agricultural Systems” (ASSIST) grant number BBS/E/C/000I0100. Thanks are also extended to G’s Growers for support with the field studies.

# Supplementary material



## Chapter 2 – Supplementary material

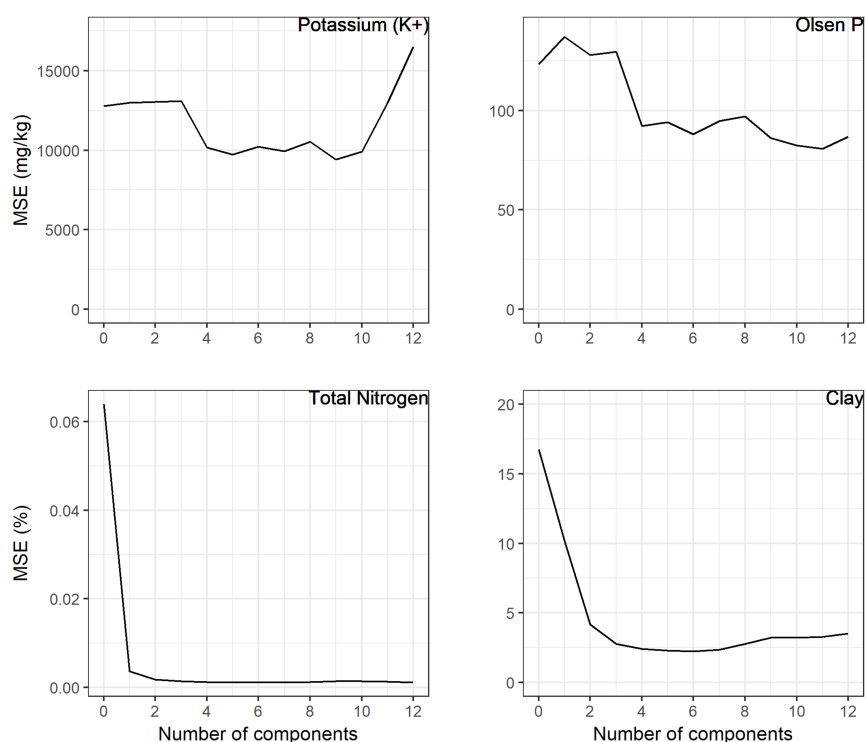


Figure S1: Mean squared error (MSE) as a function of the number of components included within the partial least squares regression (PLSR) model in a leave-one-out (LOO) cross-validation procedure. Four examples of soil properties illustrate the numbers of optimum components for each soil property.

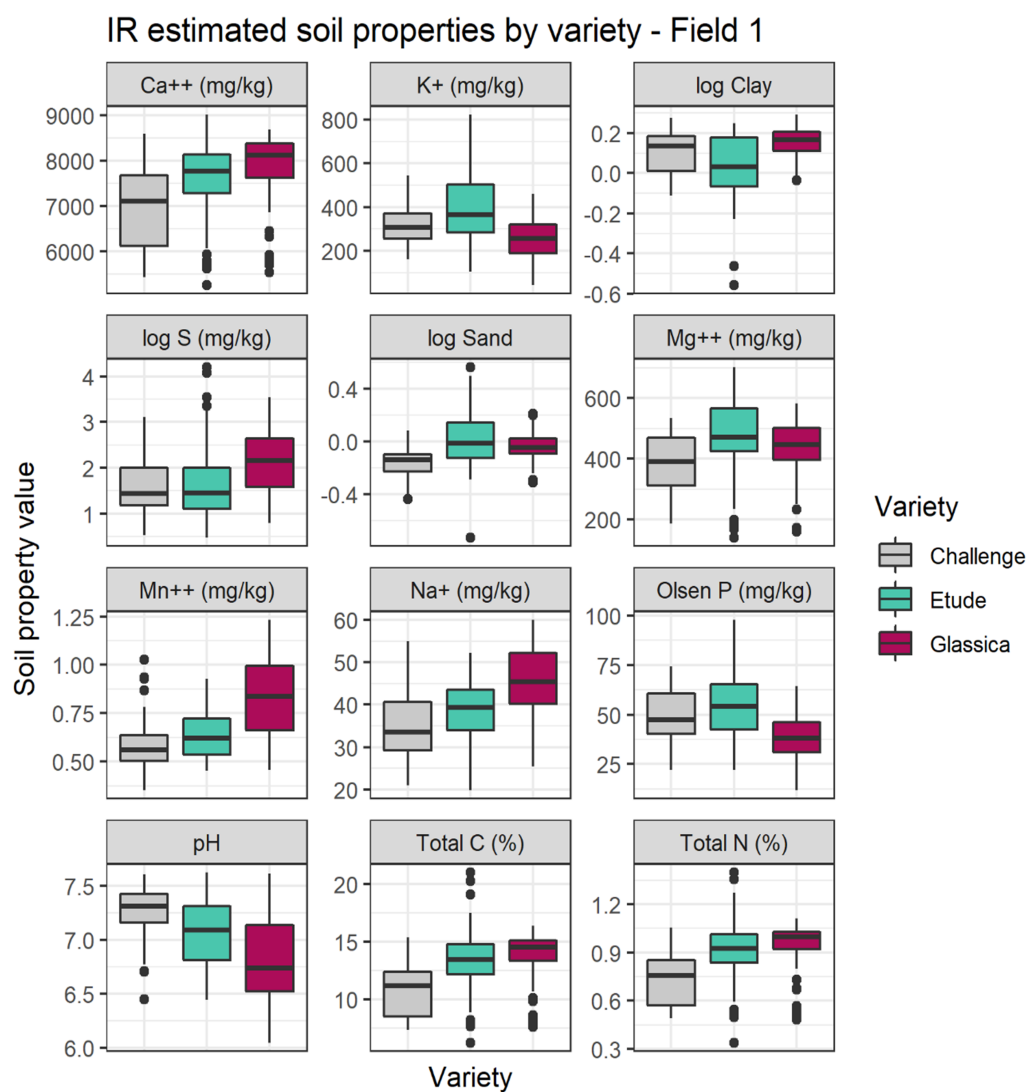


Figure S2: Range of estimated soil properties by lettuce variety for Field 1. Soil properties were estimated from Near and mid-infrared (IR) spectroscopy measurements by partial least squares regression.

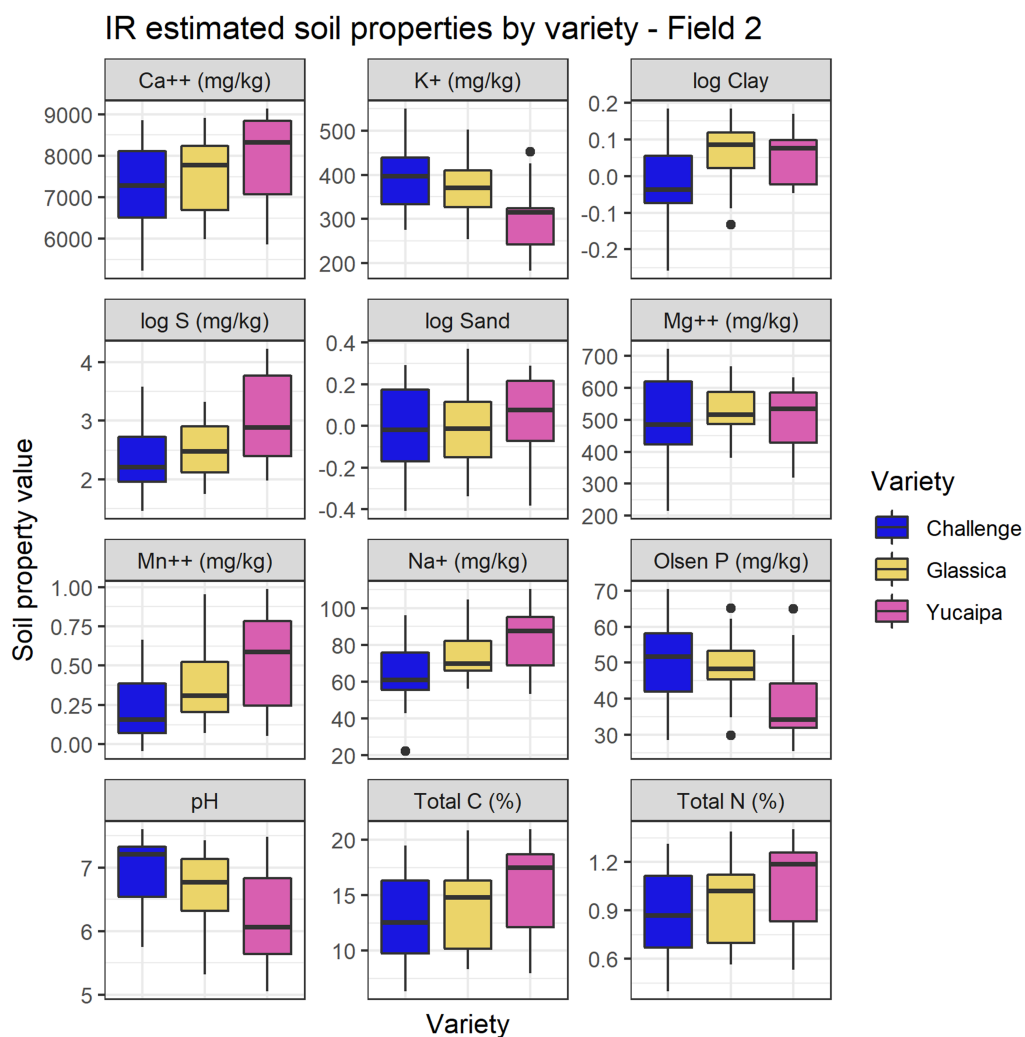


Figure S3: Range of estimated soil properties by lettuce variety for Field 1. Soil properties were estimated from Near and mid-infrared (IR) spectroscopy measurements by partial least squares regression.

Property	Coefficient	
	LM	LMM
Intercept	−29.1	8.44
Total N/%	8.42	2.40
K <sup>+</sup> /mg kg <sup>−1</sup>	0.01	0.01
log(clay)	5.63	7.57
pH	4.60	NA
Variety Etude	66.11	−0.58
Variety Glassica	24.98	−3.79
Total N : Etude	−5.46	3.83
Total N : Glassica	7.88	10.67
K <sup>+</sup> : Etude	−0.01	−0.01
K <sup>+</sup> : Glassica	0	−0.01
log(clay) : Etude	−17.7	−20.4
log(clay) : Glassica	−16.19	−15.20
pH : Etude	8.38	NA
pH : Glassica	−3.78	NA

Table S1: Linear model (LM) and linear mixed model (LMM) fixed effects coefficients – Field 1



Property	Coefficient	
	LM	LMM
Intercept	8.71	8.47
Total N/%	10.24	9.56
K <sup>+</sup> /mg kg <sup>-1</sup>	0.01	0.01
P/mg kg <sup>-1</sup>	-0.04	0
Variety Glassica	-0.52	0.29
Variety Yucaipa	4.08	0.16
Total N : Glassica	-0.61	-0.43
Total N : Yucaipa	-3.40	-2.29
K <sup>+</sup> : Glassica	0	0
K <sup>+</sup> : Yucaipa	0.08	0.09
P : Etude	0.02	0
P : Yucaipa	-0.47	-0.51

Table S2: Linear model (LM) and linear mixed model (LMM) fixed effects coefficients – Field 2

Property	Total C/%	Total N/%	Ca <sup>2+</sup> /mg kg <sup>-1</sup>	K <sup>+</sup> /mg kg <sup>-1</sup>	Mg <sup>2+</sup> /mg kg <sup>-1</sup>	Na <sup>+</sup> /mg kg <sup>-1</sup>	Mn <sup>2+</sup> /mg kg <sup>-1</sup>	P/mg kg <sup>-1</sup>	S/mg kg <sup>-1</sup>	pH	Sand/%	Clay/%	Silt/%
Total C/%	1	.	.	.	.	.	.	.	.	.	.	.	.
Total N/%	1	1	.	.	.	.	.	.	.	.	.	.	.
Ca <sup>2+</sup> /mg kg <sup>-1</sup>	0.89	0.90	1	.	.	.	.	.	.	.	.	.	.
K <sup>+</sup> /mg kg <sup>-1</sup>	0.14	0.14	0.08	1	.	.	.	.	.	.	.	.	.
Mg <sup>2+</sup> /mg kg <sup>-1</sup>	0.74	0.74	0.64	0.42	1	.	.	.	.	.	.	.	.
Na <sup>+</sup> /mg kg <sup>-1</sup>	0.55	0.55	0.36	0.22	0.76	1	.	.	.	.	.	.	.
Mn <sup>2+</sup> /mg kg <sup>-1</sup>	0.44	0.44	0.45	-0.39	-0.03	-0.16	1	.	.	.	.	.	.
P/mg kg <sup>-1</sup>	0.02	0.02	-0.01	0.84	0.22	0	-0.44	1	.	.	.	.	.
S/mg kg <sup>-1</sup>	0.56	0.57	0.61	-0.02	0.50	0.62	-0.01	-0.17	1	.	.	.	.
pH	-0.83	-0.83	-0.68	0.04	-0.59	-0.72	-0.42	0.27	-0.70	1	.	.	.
Sand/%	0.73	0.71	0.52	0.12	0.45	0.44	0.26	0.08	0.30	-0.69	1	.	.
Clay/%	-0.51	-0.49	-0.26	-0.21	-0.35	-0.39	-0.06	-0.16	-0.12	0.50	-0.91	1	.
Silt/%	-0.75	-0.75	-0.74	0.11	-0.40	-0.31	-0.50	0.10	-0.46	0.69	-0.65	0.28	1

Table S3: Correlation between soil properties ( $n = 60$ ) used to regress against near- and mid-infrared reflectance spectra by partial least squares methods

## Chapter 3 – Supplementary material

### Notes on the difference in prediction accuracy metrics between the leave-one-out cross-validation (LOOCV) and the independent validation set

Prior to the study we anticipated that the confounding effects of soil moisture content, particle size variation and aggregation would reduce the generalizability of calibration models. Consequently, one would expect a larger discrepancy between accuracy metrics from the LOOCV and those from the independent validation set for predictions from in-situ, unprocessed and air-dried samples. As a reference, we have included the prediction accuracy metrics for the LOOCV calibration and the independent validation set in table format.

Table S4: Prediction accuracy metrics from *in-situ* samples for the leave-one-out cross-validation on the calibration set (Calibration) and the predictions on the independent validation set (Validation). Ncomp: number of components included in the partial least squares regression. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range.

Condition	Property	Sensor	Ncomp	Calibration			Validation		
				RMSE	RPIQ	Bias	RMSE	RPIQ	Bias
In-situ	Organic C/g kg <sup>-1</sup>	(V)NIR	7	1.80	2.90	0.01	1.80	3.40	0.37
	pH		3	0.35	0.71	0.01	0.33	0.92	-0.05
	Clay/%		8	3.6	1.90	-0.03	2.70	3.00	0.41
	P/mg kg <sup>-1</sup>		1	15.00	1.20	0.22	9.70	0.85	5.50
	K/mg kg <sup>-1</sup>		2	120.00	1.10	0.76	68.00	1.70	22.10
In-situ	Organic C/g kg <sup>-1</sup>	MIR	7	1.90	2.70	-0.01	1.70	3.60	0.42
	pH		7	0.33	0.77	0.00	0.29	1.10	-0.07
	Clay/%		5	3.90	1.80	0.05	2.50	3.20	-0.28
	P/mg kg <sup>-1</sup>		1	17.00	1.10	0.23	9.00	0.91	4.29
	K/mg kg <sup>-1</sup>		2	140.00	0.94	4.79	74.00	1.60	26.60

Table S5: Prediction accuracy metrics from *unprocessed* samples for the leave-one-out cross-validation on the calibration set (Calibration) and the predictions on the independent validation set (Validation). Ncomp: number of components included in the partial least squares regression. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range.

Condition	Property	Sensor	Ncomp	Calibration			Validation		
				RMSE	RPIQ	Bias	RMSE	RPIQ	Bias
Unprocessed	Organic C/g kg <sup>-1</sup>	(V)NIR	8	1.60	3.30	−0.03	1.50	4.10	0.17
	pH		6	0.35	0.71	0.00	0.28	1.10	−0.02
	Clay/%		12	3.20	2.10	0.04	2.40	3.40	0.00
	P/mg kg <sup>-1</sup>		1	16.00	1.10	−0.50	9.30	0.88	5.06
	K/mg kg <sup>-1</sup>		4	120.00	1.00	0.21	86.00	1.40	28.1
Unprocessed	Organic C/g kg <sup>-1</sup>	MIR	7	2.20	2.40	0.01	2.70	2.30	−0.14
	pH		3	0.45	0.56	0.00	0.41	0.75	0.02
	Clay/%		2	3.80	1.80	0.07	3.20	2.50	−0.49
	P/mg kg <sup>-1</sup>		2	20.00	0.89	0.14	10.00	0.80	4.48
	K/mg kg <sup>-1</sup>		2	160.00	0.81	6.00	77.00	1.50	15.50

Table S6: Prediction accuracy metrics from *air-dried* samples for the leave-one-out cross-validation on the calibration set (Calibration) and the predictions on the independent validation set (Validation). Ncomp: number of components included in the partial least squares regression. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range.

Condition	Property	Sensor	Ncomp	Calibration		Validation			
				RMSE	RPIQ	Bias	RMSE	RPIQ	Bias
Air-dried	Organic C/g kg <sup>-1</sup>	(V)NIR	8	0.98	5.20	0.00	0.82	7.40	0.00
	pH		12	0.21	1.20	0.00	0.12	2.60	-0.05
	Clay/%		8	2.60	2.60	0.04	1.90	4.30	-0.20
	P/mg kg <sup>-1</sup>		4	12.00	1.50	-0.57	11.00	0.76	7.29
	K/mg kg <sup>-1</sup>		4	110.00	1.20	-3.06	74.00	1.60	28.70
Air-dried	Organic C/g kg <sup>-1</sup>	MIR	4	2.20	2.30	0.00	1.40	4.30	0.03
	pH		7	0.39	0.63	0.00	0.33	0.92	0.05
	Clay/%		4	4.50	1.50	0.00	3.00	2.70	-0.08
	P/mg kg <sup>-1</sup>		1	16.00	1.10	0.22	9.20	0.89	4.44
	K/mg kg <sup>-1</sup>		2	120.00	1.00	2.78	70.00	1.70	21.30

Table S7: Prediction accuracy metrics from *milled* samples for the leave-one-out cross-validation on the calibration set (Calibration) and the predictions on the independent validation set (Validation). Ncomp: number of components included in the partial least squares regression. RMSE: root mean squared error, RPIQ: ratio of performance to inter-quartile range.

Condition	Property	Sensor	Ncomp	Calibration			Validation		
				RMSE	RPIQ	Bias	RMSE	RPIQ	Bias
Milled	Organic C/g kg <sup>-1</sup>	(V)NIR	9	1.60	3.10	0.02	0.88	6.80	-0.22
	pH		12	0.27	0.94	0.00	0.15	2.10	-0.03
	Clay/%		11	3.00	2.30	-0.01	2.10	3.80	-0.16
	P/mg kg <sup>-1</sup>		15	12.00	1.50	-0.10	7.50	1.10	-0.53
	K/mg kg <sup>-1</sup>		12	110.00	1.20	1.31	76.00	1.60	-11.20
Milled	Organic C/g kg <sup>-1</sup>	MIR	9	1.10	4.60	0.01	0.42	15.00	0.05
	pH		15	0.20	1.30	0.00	0.13	2.40	-0.02
	Clay/%		6	2.10	3.30	0.04	1.90	4.20	-0.15
	P/mg kg <sup>-1</sup>		13	10.00	1.70	-0.14	8.30	0.99	4.50
	K/mg kg <sup>-1</sup>		9	98.00	1.30	-1.38	76.00	1.60	18.40

## Chapter 4 – Supplementary material

With regards to Chapter 4, supplementary figures are provided on the sampling design, the cross-validation results of the linear mixed model variograms, the kriging predictions and their variances that informed the loss function analysis.

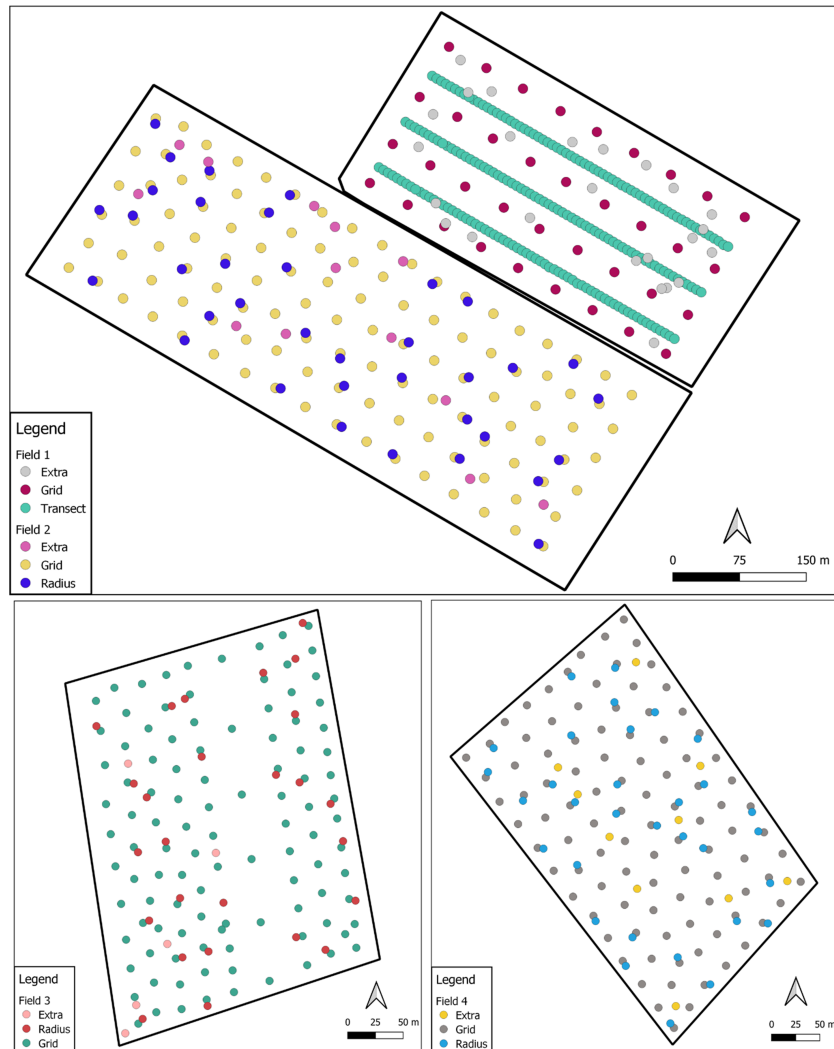


Figure S4: Field boundaries and the sampling points.

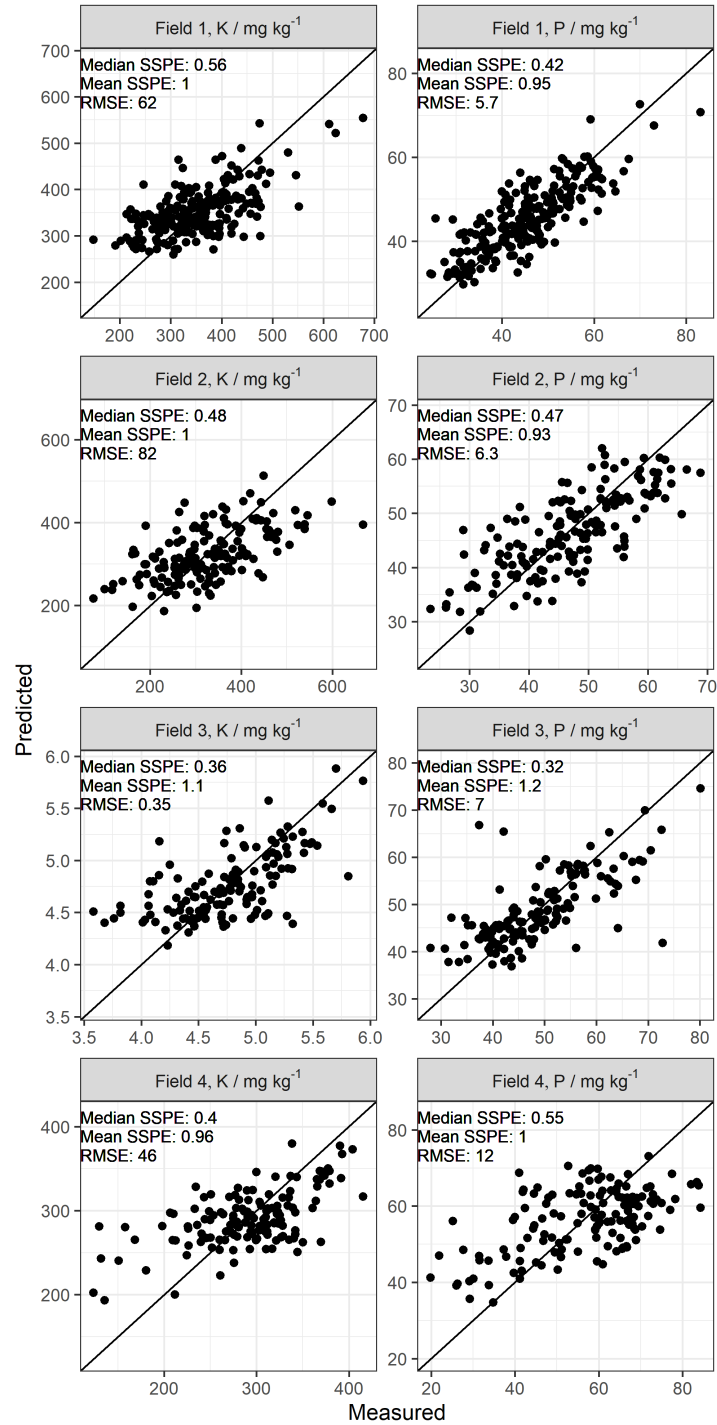


Figure S5: Leave-one-out cross validation of the linear mixed models. The linear mixed model was re-estimated for each iteration to diminish bias in parameter values. Metrics include the standardized squared prediction error (SSPE) and the root mean squared error (RMSE).



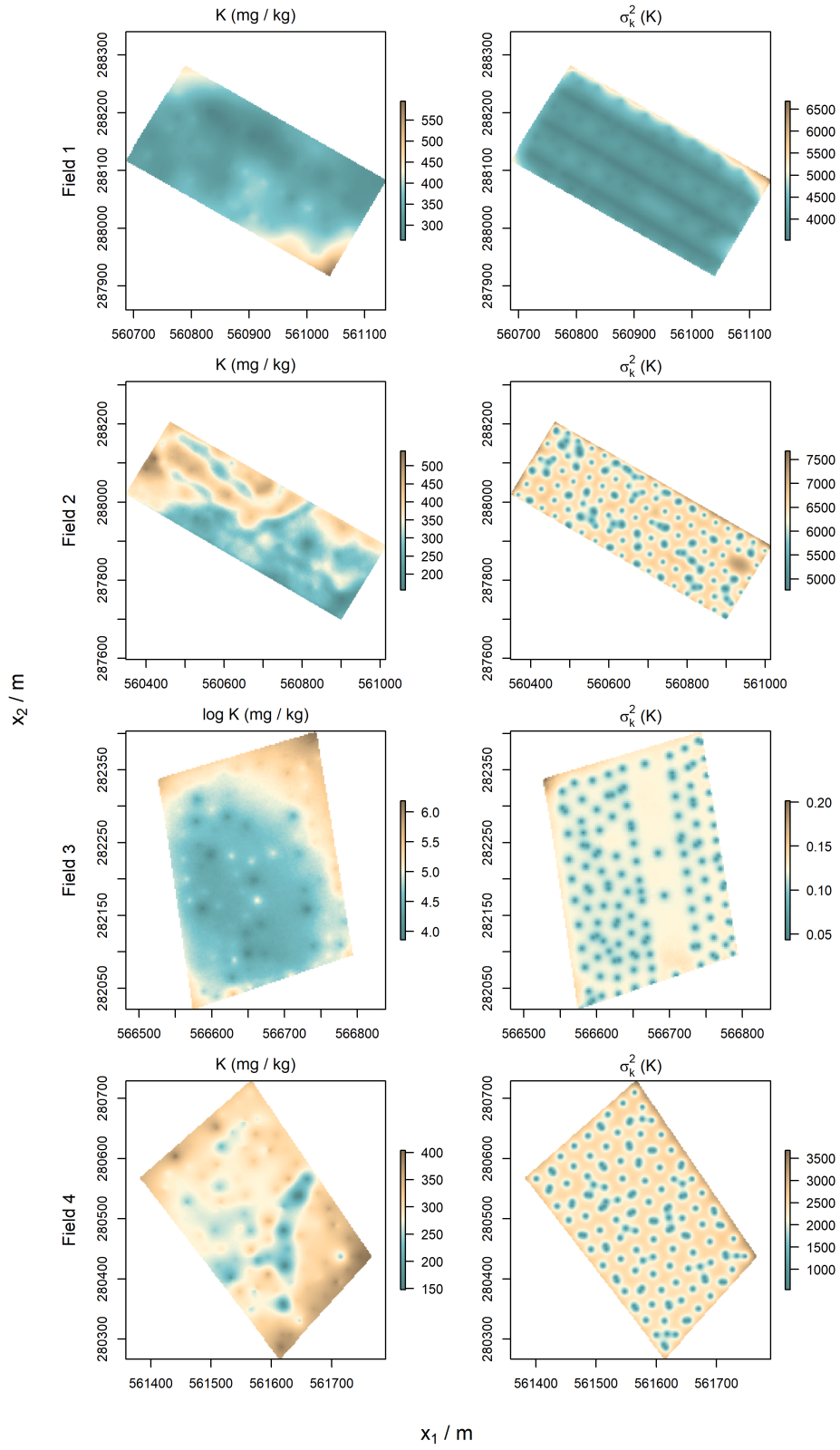


Figure S6: Kriging predictions and their kriging variance ( $\sigma_k^2$ ) by field – available K (mg / kg<sup>-1</sup>)

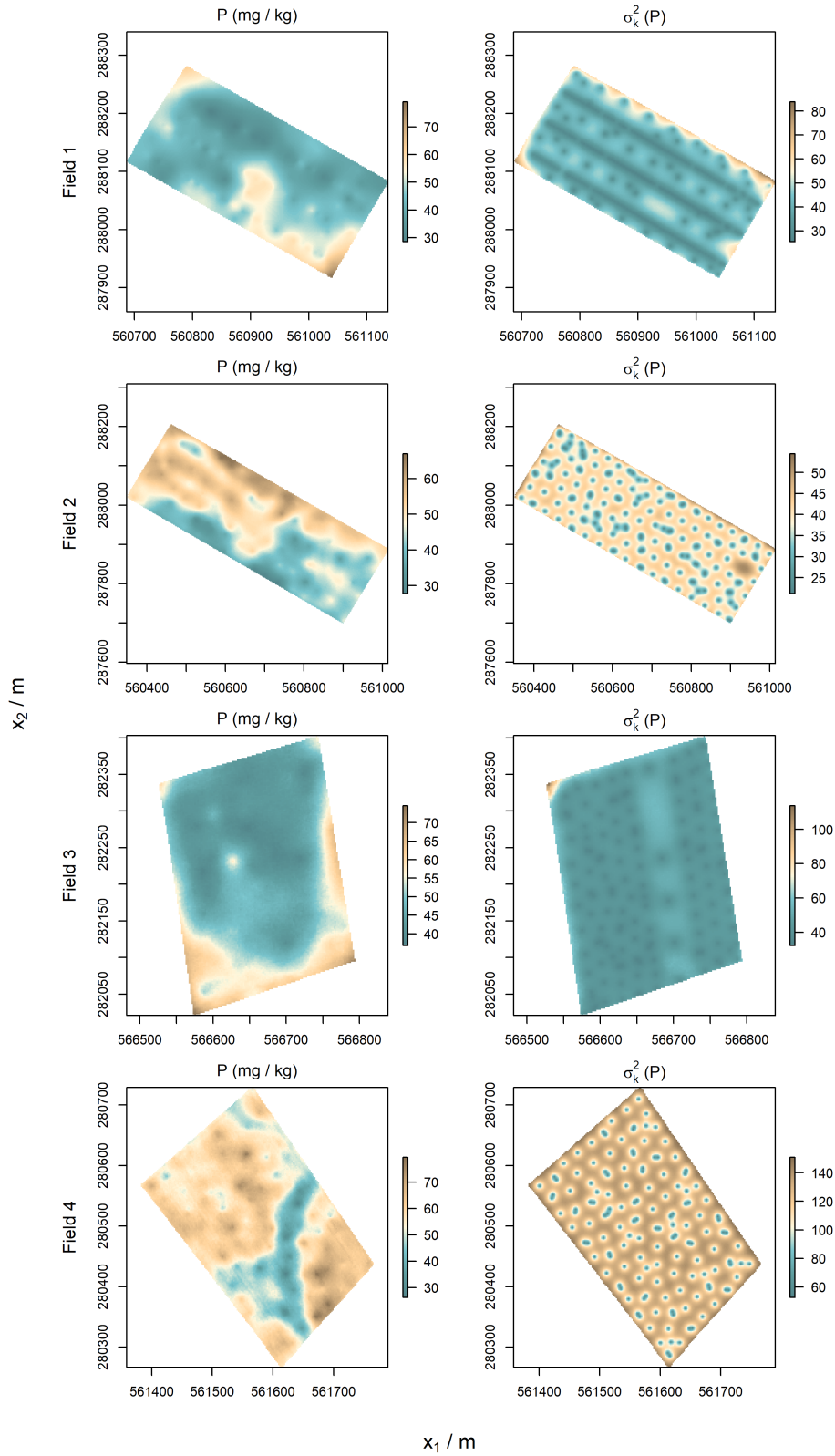


Figure S7: Kriging predictions and their kriging variance ( $\sigma_k^2$ ) by field – available P (mg / kg<sup>-1</sup>)

## Back-transformation of log kriging predictions and variances

The transformation was done following Laurent (1963).

Re-transformed kriging predictions,  $y_{kt}$ , were computed by:

$$y_{kt} = e^{(y_k + 0.5\sigma_k^2)} \quad (6.1)$$

where  $y_k$  and  $\sigma_k^2$  are the kriging predictions and kriging variance respectively, on a logarithmic scale.

The re-transformed kriging variances,  $\sigma_{kt}^2$  were computed by:

$$\sigma_{kt}^2 = e^{2(y_k + \sigma_k^2)} \times e^{\sigma_k^2} - 1 \quad (6.2)$$

## Chapter 5 – Supplementary material

This section contains the supplementary figures as referred to in Chapter 5.

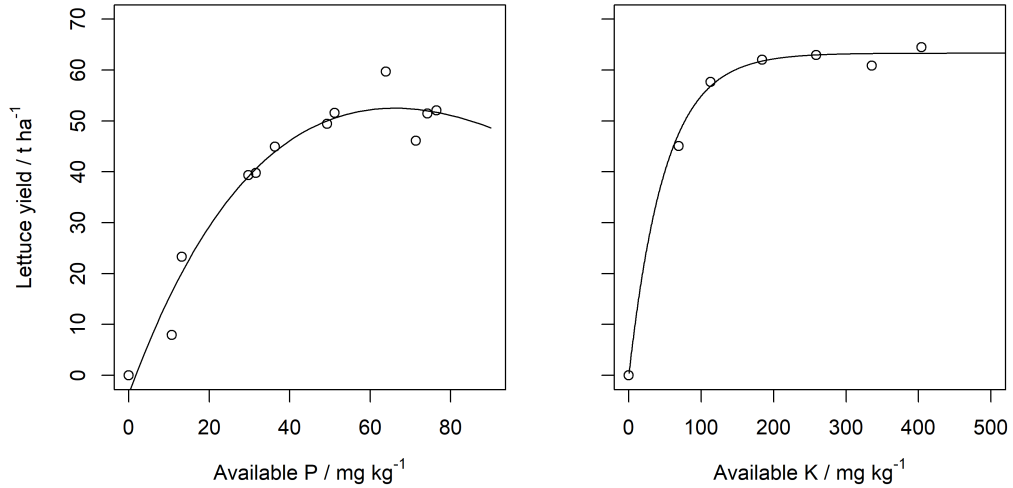


Figure S8: Dose-response curves for available P (exponential + linear) and available K (exponential) fitted based on data from Prasad et al. (1988) (P) and Greenwood et al. (1980) (K). See Table 4.1 in the main text for parameter values.

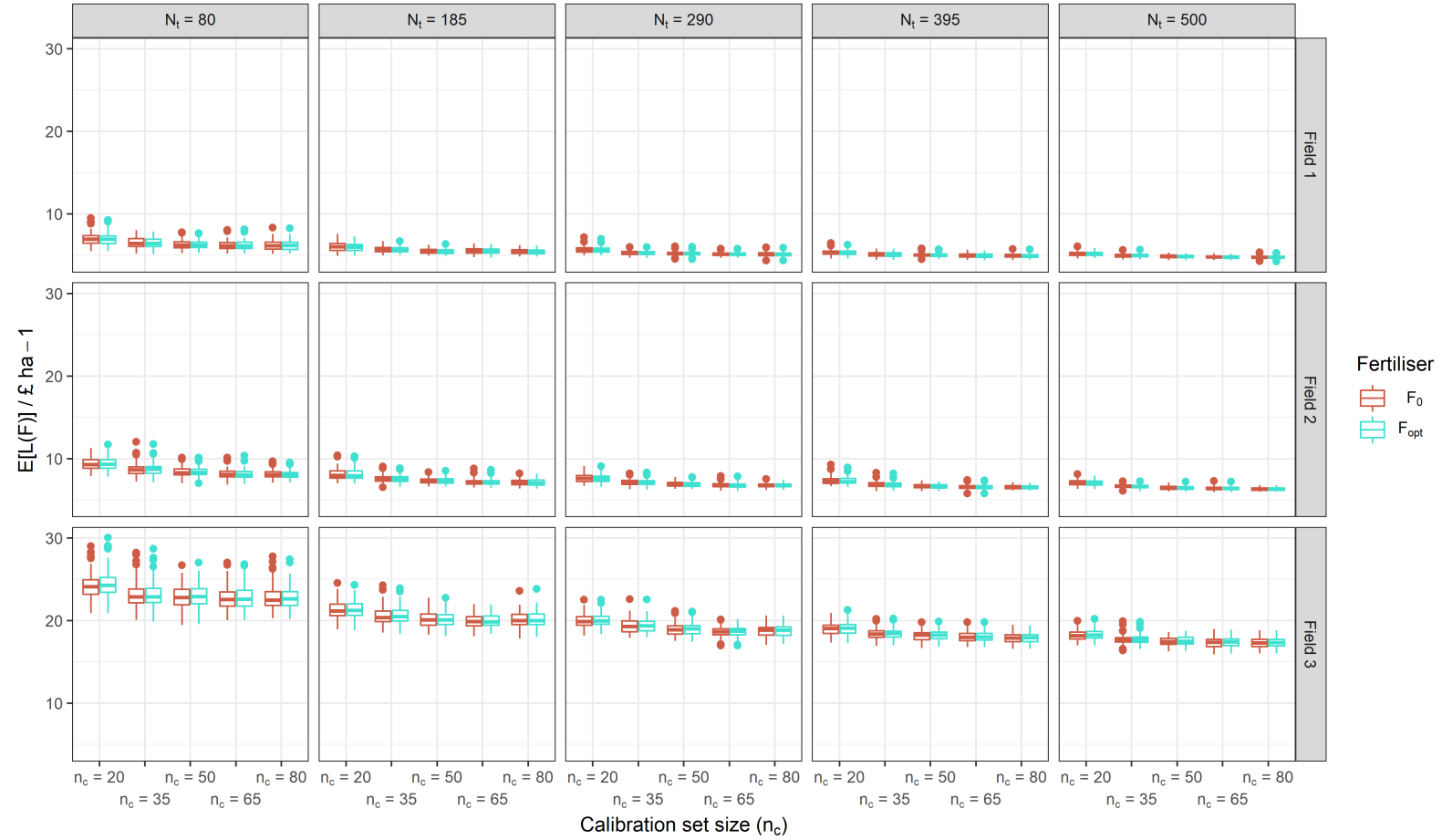


Figure S9: The expected loss ( $E[L(F)]$ ) as a function of total sample size ( $N_t$ ) and calibration sample size ( $n_c$ ) under two different fertiliser regimes. In the first regime, the optimum amount of fertiliser was computed based on the mean kriging prediction ( $F_0$ ). In the second regime, optimum amount of fertiliser was computed that maximises the expected property given the error distribution associated with the kriging prediction ( $F_{\text{opt}}$ ) – available P /  $\text{mg kg}^{-1}$ .

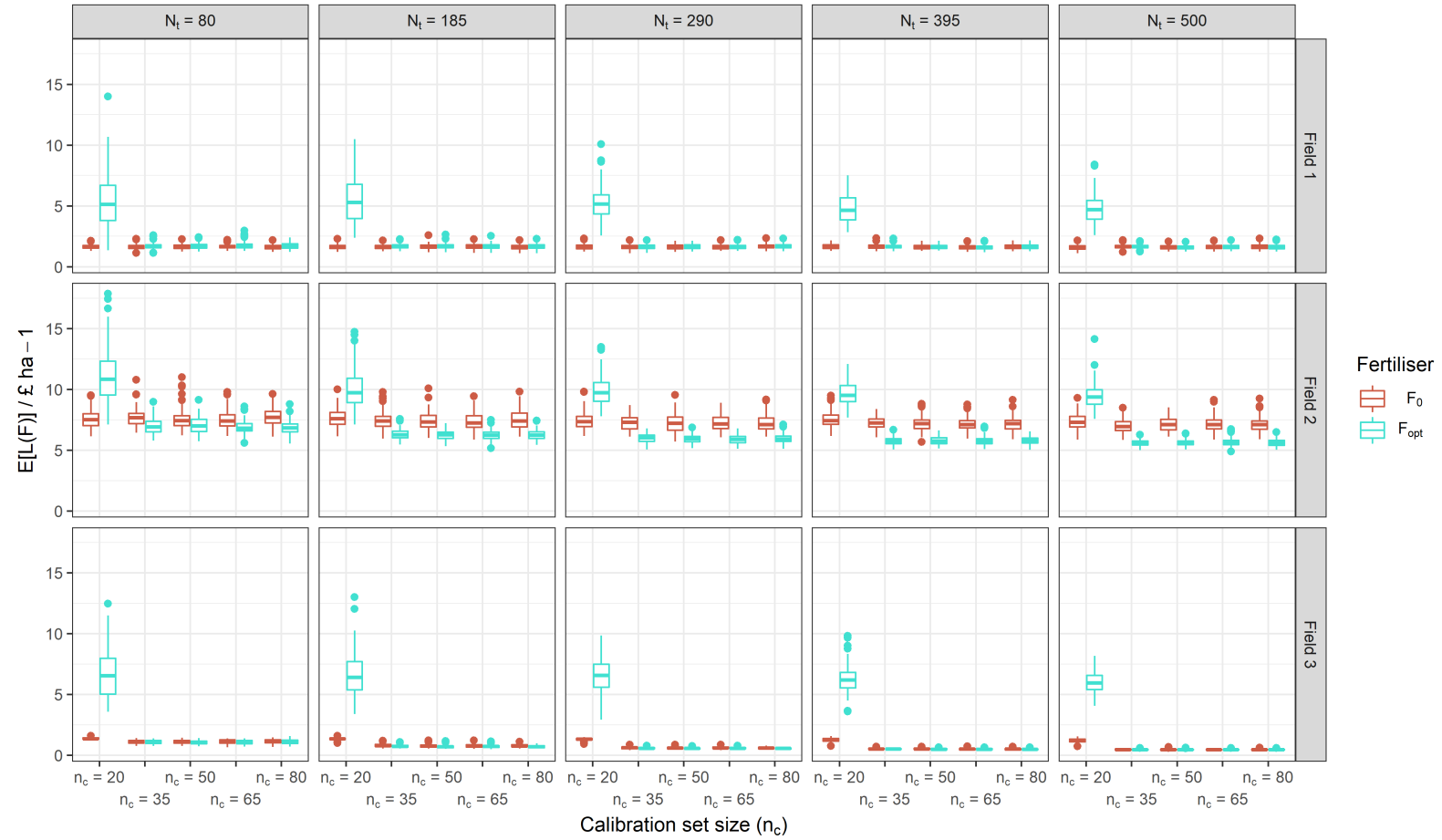
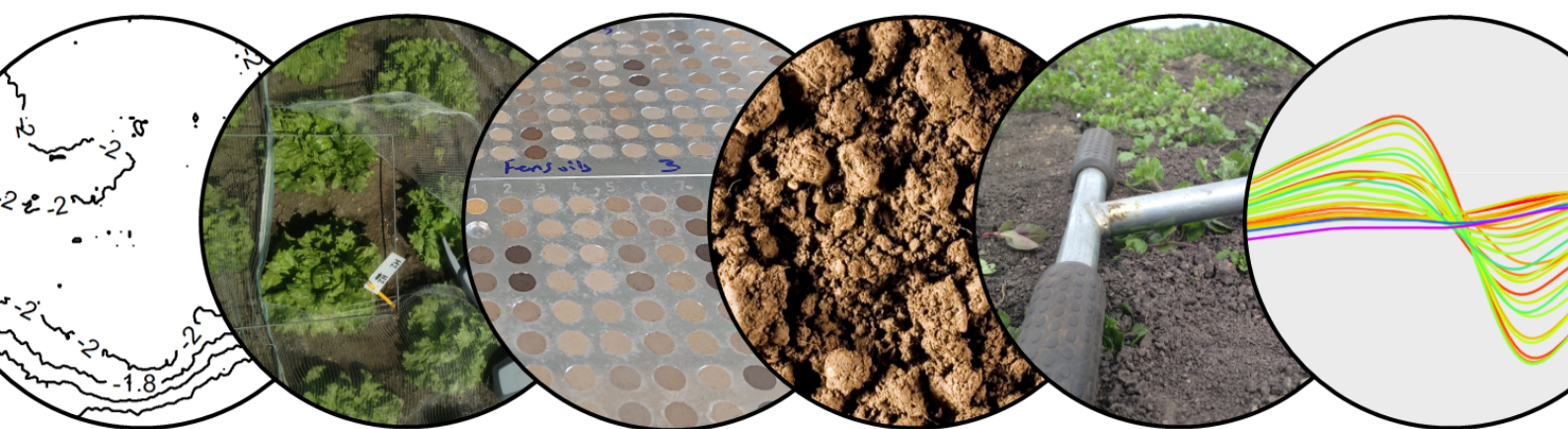


Figure S10: The expected loss ( $E[L(F)]$ ) as a function of total sample size ( $N_t$ ) and calibration sample size ( $n_c$ ) under two different fertiliser regimes. In the first regime, the optimum amount of fertiliser was computed based on the mean kriging prediction ( $F_0$ ). In the second regime, optimum amount of fertiliser was computed that maximises the expected property given the error distribution associated with the kriging prediction ( $F_{\text{opt}}$ ) – available K / mg kg<sup>-1</sup>.



### Colophon

Cover and chapter figures: Timo S. Breure

Cover design: Timo S. Breure

Formatting software: L<sup>A</sup>T<sub>E</sub>X