HGCA

PROJECT REPORT No. 302

EVALUATION OF NON-INTRUSIVE SENSORS FOR MEASURING SOIL PHYSICAL PROPERTIES

APRIL 2003

Price £10.00

PROJECT REPORT No. 302

EVALUATION OF NON-INTRUSIVE SENSORS FOR MEASURING SOIL PHYSICAL PROPERTIES

by

J A KING[1], P M R DAMPNEY[1], R M LARK[2], H C WHEELER[2], R I BRADLEY[3], T MAYR[3], and N RUSSILL[4]

[1]ADAS Ltd., Boxworth Research Centre, Boxworth, Cambridgeshire, CB3 8NN

[2]Silsoe Research Institute, Wrest Park, Silsoe, Bedfordshire, MK45 4HS

[3]National Soil Resources Institute, Cranfield University, Silsoe, Bedfordshire, MK45 4DT

[4]TERRADAT Ltd., Ocean House, Hunter Street, Cardiff Bay, CF10 5FR

This is the final report of a two-year project, which started in October 1999. The work was funded by a grant of £ 104 337 from HGCA (project No. 2243).

CONTENTS

ABSTRACT

Knowledge of soil physical properties has always been important for decisions concerning cropping and crop management inputs, especially the use of fertilisers and lime. Information on the geographical distribution of soils is important if precision farming methods are to be used. HGCA project 2243 has investigated the practical usefulness and applicability of three soil sensor technologies for measuring and mapping soil types within fields, namely Electro-Magnetic Induction (EMI) which measures the apparent electrical conductivity of the soil (EC_a), Ground Penetrating Radar (GPR) and Spectral Reflectance. The work was carried out in 1999-2002 on 4 experimental sites in England on contrasting soil landscapes (limestone, glacio-fluvial outwash, chalk, river terrace). Emphasis was given to studies on EMI.

The EMI sensor was housed in a metal-free cart drawn by an ATV at 10-15kph, with a Geographical Positioning System (GPS). It was simple and easy to use, producing a single data value at each measurement point which was able to identify soil texture variations especially where there was interaction between texture and soil wetness. Although EC_a was overwhelmingly influenced by soil moisture, the data distinguished heavier less permeable soils from those that were more permeable and free draining. Regression analysis showed that subsoil clay and organic matter contents, and topsoil sand and organic matter contents were the main factors influencing the EC_a; topsoil bulk density was also important. Since soil moisture had such a strong impact on the EC_a, direct predictive relationships between EC_a and soil properties could not be derived. This means that some in-field examination of soils will always be needed following an EMI survey to confirm the nature of the soils present in different zones. EMI data was more closely related to topsoil properties than output data from the cluster analysis of sequences of yield maps (HGCA project 2116). This suggests that the EMI sensor is reacting to the properties of the upper layers of the soil whereas the cluster analysis approach will be reacting to the soil conditions experienced by the whole crop root system, commonly well over 1m deep. The pattern of EC_a variation was remarkably stable irrespective of whether the soil was wet or dry at the time of sensing. Using geo-statistical analysis, a between-pass spacing of c.20 m usually achieved an error of less than 25%; this is considered acceptable as a cost-effective and practical approach.

GPR proved to be a slower technique producing data that was difficult to interpret. Readings could not be obtained on clay soils but information on the depth to bedrock or free-water interfaces was obtained on sandy material. Spectral reflectance measurements of the bare soil surface did not have any clear or reliable relationships with topsoil properties. Neither of these techniques are considered to have any short-term potential as a practical or cost-effective method for agricultural soil sensing.

The project has shown that EMI is a reliable method for obtaining information on within-field soil patterns. Future work should develop an integrated use of EMI with other precision farming techniques for gathering information to allow improved crop management decisions both within and between fields.

SUMMARY

INTRODUCTION & AIMS

Knowledge of the physical make-up of soils has always been important for farmers when making decisions concerning crop management inputs such as lime and fertiliser use, cultivation and sub-soiling/drainage. More precise soil knowledge, including information on boundaries between soil types, should result in more accurate use of fertilisers and agro-chemicals. The introduction of precision farming techniques (*e.g.* GPS, yield mapping, machinery capable of variable rate application, sensors) has encouraged farmers to pay more attention to crop variations that exists within a farm and within individual fields. Recent research has confirmed that variation in crop yields is primarily due to the inherent soil type, and that a knowledge of soil type and the location of boundaries between contrasting soil types is essential if these technologies are to realise their potential. The key soil variables, needed for cropping and husbandry decisions are soil physical variables such as texture, depth to bedrock, wetness and organic matter, though key soil chemical properties are also considered (*e.g.* soil mineral nitrogen). A major problem however is that current methods for obtaining information on soil properties are manual and usually requires the input of a specialist soils adviser. Even with such input, small scale variations can be missed unless very intensive (and therefore expensive) soil survey methods are used.

A recent ADAS review highlighted the most promising remote sensing technology that could be employed in arable agriculture to improve efficient soil management. Two sensors showed more potential than others for these variables, namely Electro Magnetic Induction (EMI) and Ground Penetrating Radar (GPR), whilst a third, Spectral Reflectance showed some potential.

The overall aim of this HGCA funded project was therefore: " To evaluate the practical potential of three soil sensor techniques for remotely measuring and mapping important soil physical properties, and to develop robust protocols for use on farms." The three sensors investigated were: EMI, GPR and spectral reflectance, and for each sensor technology the primary objective of the study was; "to identify the soil properties and soil horizons (i.e. soil layers of different characteristics) that can be measured using different sensor configurations, the associated accuracy of measurement and the stability at different times of measurement during the season."

Secondary aims were to identify the best sensor technique, or combination of techniques, for measuring specific target soil variables, develop soundly based protocols for in-field use, and evaluate the cost and ease of use of the sensors in practice.

MATERIALS AND METHODS

SITES & SOILS

Two sites were measured in the cropping year 1999/2000 (Year 1) and two others in the cropping year 2001/02 (Year 2). The two fields used in Year 1 were located at Lodge Farm, Shefford, Bedfordshire and Heydour Lodge Farm, Grantham, Lincolnshire, in fields named "Shagsby 4" (grid reference TL 108400) and "Field 107"(grid reference SK 996373) respectively, and known in this report by the acronyms: ***SHG*** and ***HLF***. The two fields used in Year 2 were located at Shuttleworth Agricultural College, Old Warden, Bedfordshire and Crowmarsh Battle Farm, Benson, Oxfordshire, in fields named "Football field" (grid reference TL 142447)) and "The Clays"(grid reference SU 637915), and known in this report by the acronyms: ***FTB*** and ***CLY*** respectively. All four sites had been predominantly in winter cereals in recent years and several years' yield maps of winter wheat were available.

Soil measurements carried out on site consisted of *in situ* measurements of the soil water regime, and laboratory measurements of physical properties on removed cores so that soil water properties could be calculated using pedo-transfer functions. Measurements of soil texture, bulk density and soil organic carbon, were made on both topsoil and subsoil sections of these cores, and these used to calibrate hand-texturing of further cores taken randomly across each field. Soil moisture measurements were taken at approximately the same time as the EMI scans using a SENTEK DIVINER 2000 soil moisture probe.

ELECTRO-MAGNETIC INDUCTION (EMI) SENSOR

In recent years the technique of measuring the electrical properties of soil (chiefly conductivity) has emerged as a potential tool to help differentiate and map various soil variables. The electrical current needed to measure soil conductivity can be induced electro-magnetically, using principles similar to those in operation in electrical transformers. The electrical coils used can be suspended a few centimetres above the surface, which makes the sensing "remote", and the apparent conductivity (EC_a) measured at any point over the soil surface will be influenced by the textural, moisture and solute components of the soil below it. This opens up the possibility of mapping the EC_a of a field to help evaluate and explain maps of crop yield, or to allow more cost effective mapping of soil type.

A commercially available dual coil EMI system, the EM 38 manufactured by Geonics Ltd, was towed behind an ATV in a lightweight non-metallic cart on each field of the study. This survey was carried out twice in the year at times of field capacity and maximum soil moisture deficit. Using simultaneous GPS measurements, very accurate maps could be recorded with intervals between readings of the order of 3 m. The geostatistical method of kriging may be used to estimate EC_a at points in between passes. Because the mean-squared error of kriged estimates from a given array of data may be computed if the variogram is known it is possible to find the maximum grid spacing such that the error of the kriged EC_a is acceptable.

The primary objective of the geostatistical analysis of these EC_a data was to address the point above, which should help us to answer the basic question 'how widely spaced should the passes of an EMI survey be?'

GROUND PENETRATING RADAR (GPR)

Ground-penetrating radar or GPR allows us to "see" into, below, or through otherwise solid or impenetrable materials. The technique is particularly useful in providing subsoil information where excavation is not immediately possible (*e.g.* under crops or frozen ground), or is inappropriate. The system consists of an antenna (transducer), that acts as both a transmitter and receiver of radar pulses, and a recording/control unit, which can display real-time images of what the device has "seen" from echoes that bounce back from the objects or interfaces within the material. As the pulse or wave propagates through the soil, rock or other medium, it is attenuated, reflected, diffracted, refracted and scattered by the material depending on its physical properties.

Two GPR systems were operated during the project (at *HLF* field in year 1 - a 'Sensors & Software' 'pulseEKKO 1000, and at the *CLY* field in year 2 - a 'Mala Geosciences, Ramac system'). The aim of the surveys was to profile shallow soil layers. To this end the GPR systems were physically dragged at walking pace along four 20 m transects located along selected tramlines within the study areas at *HLF* and *CLY* fields. Measurements were made across each field when the soils were at field capacity (February / March), and also when they were thought to be near maximum soil moisture deficit after harvest (July / August / September) in 2000 (*HLF*) and 2002 (*CLY*).

SPECTRAL REFLECTANCE

Sensing by reflectance of electromagnetic radiation at certain wavebands has found most application for monitoring vegetation and crop cover The measurement of soil variables using spectral reflectance is likely to be restricted to those soil features that occur at or close to the soil surface, as reflectance of incident radiation is not affected by soil properties at depth. Future potential is largely limited to the possibility of being able to sense topsoil organic matter and surface moisture content.

The principle of the system was to traverse a linear array of two-band radiometers systematically over a bare soil surface, and to this end a tractor was equipped with a 24m boom from a conventional agricultural sprayer, and radiometers (Skye Instruments type SKR1800) mounted at 4m intervals across the boom to detect radiation reflected from the soil surface (at 660nm (visible red) and 730nm (near infrared)). The tractor was driven (at 10 – 13 kph) along 24 m spaced tramlines across the *SHG* field during the post harvest period in the autumn of 2002, after the crop stubble had been ploughed in and the soil surface was bare.

Results and Discussion

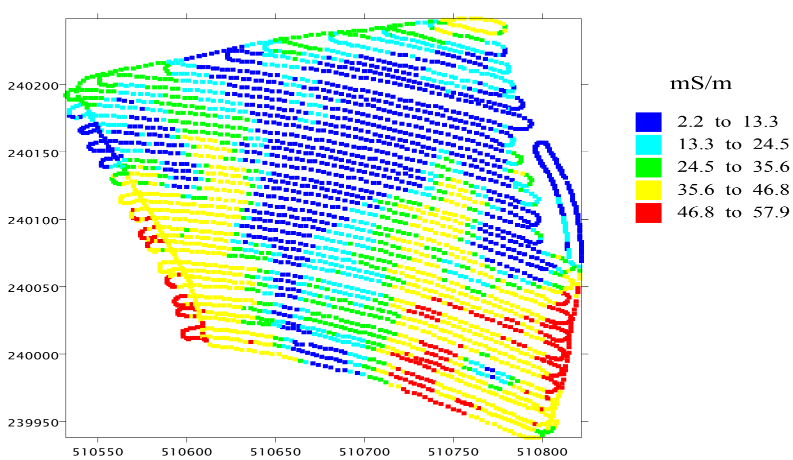
Soil survey

After surveying at *SHG* the soils mapped were of the Evesham and Oxpastures series which are clays and heavy clay loams over clay respectively, Waterstock (a clay loam), and Bearsted and Cottenham series which are sandy loams and sands respectively. At *HLF* the soil mapped were of the Haselor series, a heavy calcareous stony clay, and Elmton, a very shallow sandy clay loam over limestone. Also there were large areas of Cranwell series (a shallow stony sandy loam), and Wilsford (a loamy sand) in valley areas (further soil description in Appendix 1).

Similarly at *FTB* Cottenham series was mapped (a light loamy sand or sandy loam developed over a deep loamy sand to sand parent material) together with sandy loams of Bearsted and clay loams of Ludford series. The heavier clay loam of the Oxpasture series was located at a lower elevation. At the *CLY* site most of the field was mapped as soil of the Wallop series, which is a shallow silty clay loam to silty clay developed over fragmented chalk. There were also smaller areas of Frilsham and Soham (sandy clay loam) (Appendix 1).

Electro-magnetic Induction (EMI) Sensor

The output of the EMI scans can be displayed spatially related to the GPS co-ordinates to obtain apparent electrical conductivity maps of the site, such as that shown below for the *SHG* field, mapped during a period of field capacity (11/02/00).

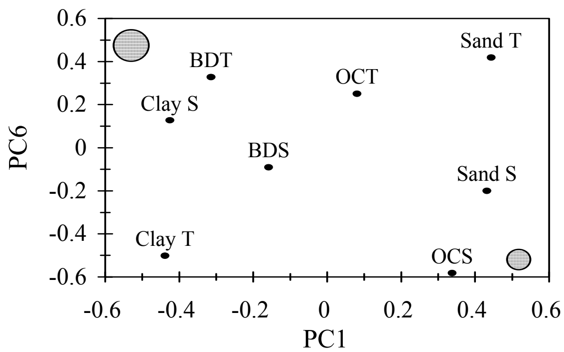


A visual comparison of the above EC_a map with the relevant soil map, indicates a clear distinction of larger EC_a readings in the parts of the field dominated by the heavier clay loam soils of the Evesham and Oxpasture series, compared with smaller values from the lighter soils of Waterstock and Bearsted series. This pattern holds true in general at both wet and dry times of the year suggesting that a large part of the signal from this field is governed by the clay content of the soil. This is also suggested by the fact that the site mean EC_a is only marginally smaller in the summer compared with the winter, and leads to the identification of two distinct classes of response for this field. Site hydrology also shapes the pattern however, as the valley feature in the Waterstock soil to the top left of the figure shows as an area of marginally higher readings, and is probably due to subsoil moisture. Similar maps for the other sites can be related to changes in soil type and hydrology across each site (see Technical Detail section), and also demonstrate a basic stability of pattern across wet and dry seasons. More information about the spatial variability of soil features across each site can be gained however, by employing basic and more advanced geo-statistical analyses.

Summary statistics of the data from each site revealed that whilst some were normally distributed other were skewed and better analysed after log-transformation, and indeed one site, *SHG*, was bi-modally distributed and had to be analysed as two data-sets. Kriging analysis showed that the usually employed estimator of the variogram, "Matheron's", led to over-estimates in most cases (ascertained from the confidence interval of the median value of the standardised cross-validation error), and that other, robust, estimators would be more appropriate. Validation of the analyses showed that in most cases "Dowd's" estimator gave the best estimate of the variogram, but each data-set should be analysed separately to ascertain the most appropriate model and estimator to be used. However, using the best estimator, the required spacing between passes of the EMI instrument to give a target estimation error of either 10 or 25 % of the mean could be calculated for both point and block kriging. This showed that at most sites < 5 m spacing would be required for 10% error by point kriging, but 15 – 24 m would be suitable for 25% error in either vertical or horizontal modes of usage. For block kriging (10 m block) a spacing between 10 & 20 m would be sufficient at 10 % error. If only Matheron's estimator was used, then a requirement for almost twice as many passes to achieve the same level of error would be necessary. However, no single spacing could be recommended for all sites, modes and conditions.

Generally speaking there was found to be only minor changes in these spacing estimates over time for each data-set, most notably at *HLF* where the overall mean conductivity changed most. The robustness of spatial patterns was analysed by two methods, co-kriging of the change in EC_a between the two seasons, and also by cluster analyses performed on sets of EC_a readings on the two dates. For virtually all sites and modes of use the cluster centres showed that relative change in EC_a was minor and relative differences were maintained, highlighting the stability of the basic pattern of EC_a variation across sites between seasons.

The effect of the measured soil physical variables on EC_a readings (kriged values from within 25 m of the same locations as soil measurements) was initially analysed across all sites on a combined data-set by regression of EC_a on principal components of the soil variables. The principal component analysis showed the individual nature of the *CLY* site, which was thereafter analysed separately. Principal component analysis of the soil variables, followed by multiple regression analyses, and the partial regression analysis of EC_a on each soil variable in turn allowed us to interpret how the soil variables contribute to the overall EC_a reading. By this method it was shown that topsoil sand content and bulk density had significant effects on EC_a as did subsoil bulk density organic matter and clay contents. Of these the most significant effect came from topsoil sand content followed by subsoil bulk density, and that these effects are more apparent in the spring when measured at field capacity. The effects (for all fields other than *CLY*) can be summarised for the two most influential principal component vectors (PC1 & PC6) by the diagram below, where a large EC_a is indicated by the large grey circle in the top-left corner of the diagram and a small EC_a by the small grey circle in the bottom-right corner. Thus, particularly large EC_a are expected from soils with large clay contents, particularly in the subsoil, large bulk densities in the topsoil, small sand contents, particularly in the subsoil and small organic carbon contents in the subsoil. Further evidence comes from the next two most influential vectors (PC2 &PC8), which indicate that larger organic carbon content in the topsoil will be associated, other factors being equal, with larger EC_a, as will low sand content in the topsoil. At site *CLY* similar analyses showed that organic carbon in both topsoil and subsoil was more influential, as was subsoil sand content and bulk density, in determining the EC_a.



GROUND PENETRATING RADAR (GPR)

The measurements generated pseudo depth section profiles of the transects at the *HLF* site in year 1, which allowed a fairly deep "view" into the soil to be made. This was because of the sandy nature of the upper soil material with a low electrical conductivity. Radar reflections were measured down to around 4m below ground level, but showed zones of differential penetration where reduced depth may be due to more clay-rich soil. Dipping reflectors in the deeper parts of some profiles were interpreted as possible bedrock strata beneath the soil layer, and within the soil layer, there were numerous reflection events indicative of more prominent soil layers and point source reflectors caused by underground services / drains, buried rocks or voids. Unfortunately, at the *CLY* site in year 2, due to limitations caused by high clay contents in the topsoil which prevents the penetration of the radar signal, exploration was severely restricted to a few centimetres and no useful data could be obtained.

SPECTRAL REFLECTANCE

The reflectance measurements made during the autumn of 2002, were also mapped across the site (*SHG*) for 'visible red' wavelength readings, at 660 nm. Visible red was selected as this tends to be more indicative of changes in soil colour and type than the ratio of red/infra-red that is used in vegetation indices (and is more influenced by moisture and surface structure). The resulting map did not really suggest the same pattern seen in maps of EC_a for this site or the soils. A separate principal component analysis was conducted on the soil data for *SHG* field for reflectance measurements, and the regression of visible red reflectance on static soil properties showed that only topsoil clay content has a significant partial effect. Organic carbon content (which is often important in determining VR reflectance of soil material) had only a small weighting in the dominant principal component vector. This may be because, at the time of measurement it was observed that there was a good deal of variation in the structure of the soil surface, which would influence visible red reflectance because of differing amounts of shadow created by aggregates of different size and shape.

CONCLUSIONS

ELECTRO-MAGNETIC INDUCTION SENSOR

It is necessary to evaluate critically, and validate, the random function models that we assume underlie our data in a geo-statistical analysis. Matheron's estimator of the variogram is the most efficient statistically, and analysis of data on the original scale is always to be preferred since it avoids complications associated with back-transformation of the final results. Alternatives may be considered in the following sequence: **i.** A data transformation should be considered; **ii.** Robust estimators of the variogram should always be considered for use on the original or transformed data (robust estimators assume normality); **iii.** When data have a complex distribution such as a bimodal, the possibility that there are two or more distinct regions requiring a separate spatial analysis should be considered.

Differentiating variations in texture

The evidence from the *HLF* data-sets showed clearly the potential of EMI techniques to distinguish between soil types based on clay or sandy loam textures (Haselor and Cranwell). At *SHG* the main soil types were all clays or clay loams, and the main determinant between them was the amount of clay in the upper subsoil and/or the interaction between texture and soil hydrology. In this case EMI distinguished the heavier less permeable soils (Evesham and Oxpastures) from more permeable and freely draining soil (Waterstock and Cottenham), during both wet and dry times of the year.

Where the definition of a pattern to EC_a variation across a field was chiefly caused by soil type, this pattern remained remarkable stable across seasonal fluctuations in the moisture regime. The principal component analysis on the variations in EC_a with soil physical variables clearly showed that subsoil clay and organic matter contents and topsoil sand and organic matter contents are major determinants of the variability of conductivity across a site. Topsoil bulk density also proved of importance. Since bulk density and clay content are important in determining EC_a values then the measurements will be informative about soil hydrological conditions.

Differentiating variations in soil hydrology

The ability of EMI to record differences in the water content of soils can be useful in one of two ways. Firstly, it can identify regions within fields that behave differently hydrologically to those in the rest of the field, such as valley features at both *HLF* and *SHG*. Similarly, it also locates reserves of water deeper in the profile that are not apparent from surface topography, such as at *HLF*. The vertical mode of operation proves more apt at this usage, and the difference in readings made during different seasons highlights regions of maximum and minimum change more clearly, when geo-statistical and cluster analysis techniques are employed. Secondly, the general ability of fine textured soils to hold water enables changes in depth of soil profile to be assessed, when changes in the depth of underlying parent material are not visible from the surface (*CLY*).

The soil moisture content itself proved surprisingly less useful in explaining variability in the EC_a measurements across fields. It is thought that this may be because the amounts held in the soil pore space is itself governed by static soil properties such as texture and bulk density, and all our analysis shows is that the actual water content provides no additional explanation of the variation in EC_a. However, the presence of water does compound the differences due to these features and make spatial differentiation easier.

<u>Suggested protocols for the use of EMI</u>

The manner in which EMI is best used will to some extend be guided by the reason for the survey, to delineate soil type boundaries, identify soil management zones, map saline or droughty areas or others. However, some more general points to consider are offered here to enable the suspected features of interest to be highlighted more securely in any survey. Some *a priori* knowledge of the soils on site is essential for the capabilities of the EMI instrument to be fully exploited.

Firstly, the orientation of the instrument can help to either emphasise features in the upper soil profile in the horizontal mode or in the upper subsoil in the vertical mode. If changes in clay content of the topsoil, or the presence of impeded water due to compaction or panning at the base of the subsoil are the sort of features of interest, then the horizontal mode of operation should be used. If changes in subsoil texture, depth of soil, or the presence of deeper moisture reserves are under consideration, then the vertical orientation would be better. If little is known of the soil problems or changes, but it is hoped to identify management zones in relation to yield maps, then the vertical mode is probably the better general purpose option.

Secondly, considering the commercial collection of EC_a data to generate maps for farmers. It is clear from these results that a single spacing between passes will not be optimal for all sites, although a spacing of 10 to 20m would not be wildly unsuitable at any site studied here, and indicates that in a growing crop the use of tramlines is not unreasonable. In the longer term one possibility is to automate the robust analysis procedure used above. Four or five passes could be made in a field at a narrow spacing of about 6 m, then after a pause during which the data are analysed, the optimal spacing could be identified and the rest of the field surveyed at a density planned to ensure that the final map is of adequate precision.

It is assumed that the EMI scanner will be housed in a non-metallic cart of some form and towed behind an ATV over a bare or recently established crop with a level surface. Under these circumstances, it should be ensured that the EMI instrument is at least 3 m away from the vehicle or metal components to avoid interference with the signal. A slower speed of around 10–15 km h^{-1} is recommended, depending upon the size of the field and width of passes, because much information is lost as points of measurement become more widely spaced. The instrument can also be used by being walked or towed along single transects across features of interest.

GROUND PENETRATING RADAR

The applicability of GPR to field survey proved to be severely limited by the soil material found on site. At the site, *CLY*, it proved impossible even to obtain a set of readings, due to the fact that the soils had a high clay content in the topsoil that effectively reflected the signal before it had even penetrated the main body of the soil profile. Where the instrument was used on sandy material (*HLF*) rather more information was gained, relating in the main to the location of bedrock or free-water interfaces in the profile. The information which GPR provides and EMI does not, is of course an estimate of the depth in the soil profile at which any features occur (when soil is sandy in nature). It is not however, suitable for mapping the spatial distribution of features in two dimensions

Suggested protocols for the use of GPR

The potential for the use of GPR is rather more limited than EMI, as it is not as yet tow-able behind an ATV in its normal mode of operation, which is to be dragged over the surface of the ground. This necessitates a level surface and slow speed of walking-pace. It is therefore only suitable when no crop is present, and over single transects or short distances. It is also not suitable for sites with a high clay content in the topsoil.

It is suggested that it is better to operate it during summer months when the soil is dry, when trying to identify physical subsoil features such as depth to argillic horizons in the subsoil, bedrock, or the depth to the water-table. It is suggested that it is best used in a limited capacity where such features of interest have already been located by other survey methods and more detail is required.

SPECTRAL REFLECTANCE

Although a map of visible red reflectance seems to offer a coherent pattern of variation across the field it would be unwise to interpret this with regard to soil type. This is because neither multiple regression and partial effects, nor principal components analysis, can adequately identify any soil physical component as a correlating variable. Topsoil clay content did have a partial effect but this was thought to act mainly through its effects on surface structure, which is largely unpredictable.

Suggested protocols for the use of Spectral Reflectance

The use of spectral reflectance techniques to measure soil properties in the field is in its infancy, and is usually allied to crop measurements. At present, vehicle- and air-borne surveys of bare soil by spectral reflectance analysis reveal only limited and quite specific information about the surface soil on a site. As such, they cannot yet really be recommended for use in commercial within-field survey work.

PRACTICAL SENSING

EMI

EMI surveys are currently being offered commercially by several companies, which may also combine it with other agronomic advisory services. Costs are negotiable, but around £20 ha^{-1} can be considered a current (2003) guide price. We consider this technology offers good value for money to growers who experience short-range soil variation within fields, especially when specialist advice on how to interpret and act on the information is available. EMI proves most suitable for targeting such features as:

- Changes in soil type due to texture and differential hydrology.
- Subsoil water reserves in permeable material
- Location of shallow soils and bedrock near the surface.
- Drought prone regions within fields
- Clay subsoil features in otherwise sandy material.

GPR

GPR only proves suitable for targeting specific features that are already suspected, and is a service which is not currently offered commercially in an agronomic context (though specialist geophysical contractors will carry it out). We do not consider it a cost effective technology for soil management decision making, except in highly specialised situations (Cost would be around £650 per day for a maximum of 3 km transect measurement). GPR proves most suitable for targeting such features as:

- Depth of soil profile, and depth to clay rich layers, in sandy materials.
- Depth to water table during dry periods in permeable soil.
- Location and depth to buried pipes, boulders or other hardened point source features.

Spectral reflectance

Spectral reflectance measurements are not currently available for soil management on a commercial basis. If however, soil information is offered as a cost-free addition to crop analysis (which is beginning to be available commercially), then some use could be made of it to guide soil management if guided by expert advice. Spectral reflectance measurements prove most suitable for targeting such features as:

- Surface soil moisture content.
- Changes in features that affect surface soil moisture, such as organic matter.
- Changes in soil type as reflected by topsoil colour.

TECHNICAL DETAIL

AIMS

The overall aim of the project was: " To evaluate the practical potential of three soil sensor techniques for remotely measuring and mapping important soil physical properties, and to develop robust protocols for use on farms."

This was comprised of the four more specific objectives:

1. For each sensor technology, to identify the soil properties and soil horizons (i.e. soil layers of different characteristics) that can be measured using different sensor configurations, the associated accuracy of measurement and the stability at different times of measurement during the season. Emphasis will be given to sensing by EMI and airborne spectral reflectance since currently these techniques are judged to have most potential for cost-effective practical use on farms (King and Dampney, 1999).
2. To identify the best sensor technique, or combination of techniques, for measuring specific target soil variables (e.g. soil depth over rock, topsoil organic matter).
3. To develop soundly based protocols for in-field use.
4. To evaluate the cost and ease of use of the sensors in practice.

INTRODUCTION AND BACKGROUND

A recent ADAS review for the Ministry of Agriculture Fisheries and Food (MAFF) (King and Dampney, 1999) highlighted the most promising sensors that could be employed in arable agriculture to improve efficient soil management.

The main output of the review was the identification of the most promising applications of current and forthcoming sensor technologies, and recommendations for the research needed to test and develop these applications into reliable farm practices. This study follows on from that review, taking the most promising of the technologies identified and assessing their worth in arable crop management alongside other precision farming techniques (for example yield mapping).

THE NEED TO SENSE SOIL PROPERTIES

Knowledge of the physical make-up of soils has always been important for farmers when making decisions concerning crop management inputs such as cultivations, subsoiling/drainage and fertiliser use. However most farmers and advisers are reluctant to spend time examining soil type and condition below the surface. This is partly due to a lack of knowledge of 'what to look for', and partly because it is a laborious job.

Although the National Soil Resources Institute (formerly the Soil Survey and Land Research Centre, and previous to that the Soil Survey of England and Wales) has surveyed all of the country, only 25% of the nation's soils have been mapped at 1:25,000 or 1:63,000 scales. Since the intensity of field soil surveying for these maps is commonly 1 core to 3 or 4 ha, the locations of some important within field soil boundaries are too imprecise. These surveys also take no account of modifications to soil structure resulting from farm operations (*e.g.* soil compaction). Therefore, there are large areas of arable agriculture which have either no existing soil survey information, or where the information is incomplete, at too coarse a scale.

More precise soil knowledge should also result in more accurate use of fertilisers and agro-chemicals. Standard recommendations for lime, NPK fertilisers and herbicides are all adjusted according to soil type, so unless the farmer or adviser can accurately assess the soil type, the use of these inputs is likely to be less than optimal. This will have implications for farm profitability and protection of the environment.

The introduction of precision farming techniques (*e.g.* GPS, yield mapping, machinery capable of variable rate application, sensors) has encouraged farmers to pay more attention to variation that exists within a farm and within individual fields. Several research projects (Dampney *et al.*, 1998; Lark *et al.,* 2003a) have confirmed that variation in crop yields is primarily due to the inherent soil type, and that a knowledge of soil type and the location of boundaries between contrasting soil types is essential if these technologies are going to realise their potential.

The review of King & Dampney (1999) concentrated on those soil variables that may be used for cropping and husbandry decisions, and identified the key soil variables needed for i) cropping strategy, ii) crop and soil management, iii) detection of artificial sub-surface features and iv) non-agricultural purposes. These are primarily soil physical variables such as texture, depth to bedrock, wetness and organic matter though key soil chemical properties are also considered (*e.g.* soil mineral nitrogen).

Two sensors showed more potential than others for these variables, namely Electromagnetic Induction (EMI) and Ground Penetrating Radar (GPR), whilst a third, Spectral Reflectance Measurements, showed some potential. Some background to these technologies is given below.

ELECTRO-MAGNETIC INDUCTION (EMI) SENSING

In recent years the technique of measuring the electrical properties of soil has emerged as a potential tool to help differentiate and map various soil variables (King & Dampney, 1998). Measuring the electrical conductivity of soil/water paste extracts has long been used to evaluate the solute concentration when assessing soil salinity hazard. Until recently however, measurements made at the field scale were the province of geophysical surveys. This was accomplished by inserting four electrodes into the surface in a line (the "Wenner" array), and recording the resistivity (ρ) of the material measured between the inner pair of electrodes (Kollert, 1969). Resistivity is the inverse of the conductivity of a material, and a typical range of electrical conductivities (EC) for component materials of soil, is shown in Table 1.

Table 1. *Typical ranges of electrical conductivity (Kollert, 1969)*

Material	**EC (mS/m)**
Irrigation water (0.7% salt)	<1500
Potable water (0.25% salt)	<550
Clay	10-1000
Sand	0.01-1
Loam	20-200
Topsoil	5-20
Sandstone	0.1-50
Limestone	0.4-20
Crystalline rocks	<0.001-1

The electrical current needed to measure soil conductivity can also be induced electromagnetically, using principles similar to those in operation in electrical transformers (McNeill, 1980). A transmitting coil, that is energised with an alternating current at an audio frequency, is placed on the soil surface. This sets up a magnetic field around it, that induces a weak electrical current in the soil, which in turn generates a second magnetic field a set distance from the transmitting coil. A second receiving coil placed here generates an alternating current in response to, and proportional with that in the transmitting coil, but altered by the electrical conductivity of the soil. If the magnetic field at the transmitting coil is H_t , and that at the receiving coil is H_r , then;

$$H_r/H_t = (\sqrt{-1}(2\pi f)\mu_o \mathrm{EC}a^2) / 4 \qquad (1)$$

where f is the current frequency (Hz), μ_o is the permeability of free space, EC is the ground conductivity (S/m) and a is the inter-coil spacing (m). This can be re-written to obtain a reading of the apparent ground conductivity (EC_a) which is linearly proportional to the ratio of the two magnetic fields (H_r/H_t).

The electrical coils used can be suspended a few centimetres above the surface, which makes the sensing "remote". Using a 1 m inter-coil spacing, the zone of influence is the surface 1.5 to 3.0 m of soil. From Table 1 it can be seen how the apparent conductivity (EC_a) measured at any point over the soil surface will be influenced by the textural, moisture and solute components of the soil below it. As a result of this, some detailed knowledge of the site will always be needed to explain the measurements. Nevertheless, this opens up the possibility of mapping the EC_a of a field to help evaluate and explain maps of crop yield, or to allow more cost effective mapping of soil type. On more uniform sites there is also the potential to make more direct measurements of variables such as soil moisture for management decision making.

The instrument can be used in two modes, with the coils either vertically or horizontally orientated. With the coils in a vertical orientation the signal obtained is influenced by soil material up to 4 m in depth, but the upper 1.5 m of soil depth contributes most to the signal. In horizontal mode however, a slightly shallower layer of material (> 2 m) is penetrated, but the signal response is more strongly influenced by the the upper 50 cm of a soil profile. The explanation of this can be found in McNeill (1980), but by using both modes of operation we may deduce more about the depth profile of soil on a site.

GROUND PENETRATING RADAR (GPR)

Ground-penetrating radar or GPR is one of a number of non-invasive geophysical techniques that allows us to see into, below, or through otherwise solid or impenetrable materials. The technique is particularly useful in providing subsoil information where excavation is not immediately possible (*e.g.* under crops or frozen ground), or is inappropriate.

GPR is a time-domain impulse radar that transmits broad bandwidth pulses into geologic media, and acts as a sounding device very much like depth finders in boats.

The system consists of an antenna (transducer), that acts as both a transmitter and receiver of radar pulses, a recording/control unit, which can display real-time image of what the device has "seen", and connecting cables. GPR is a reflection system that uses non-ionizing electromagnetic waves to probe the material under investigation (Fig. 1), remaining at the surface of the soil and picking up echoes that bounce back from the objects or interfaces within the material.

The operating heart of the system is the control recording unit that sends signals via fibre optic cables to the antenna or transducer which in turn produces a polarised radar signal. The antenna then switches off, and detects reflections, which are then recorded by the control unit. Because the pulses are going out and coming back at light velocities, there is ample time available for the antenna to record all the reflections or return echoes, before sending out another pulse. As the pulse or wave propagates through the soil, rock or other medium, it is attenuated, reflected, diffracted, refracted and scattered by the material depending on its physical properties. (Fig. 1).

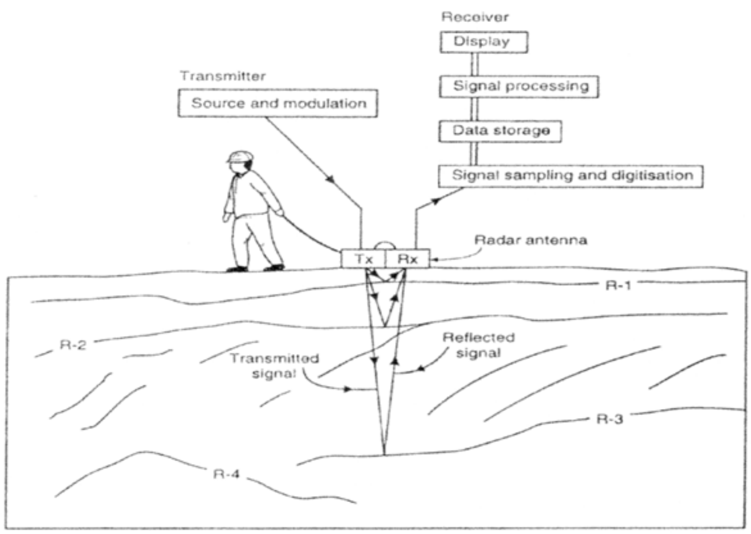


Figure 1. General Arrangement for pulseEKKO 1000 GPR

Two physical conditions of the medium have major impacts on the radar waves and influence the depth attainable in a GPR survey: (1) dielectric properties and (2) conductivity (Table 1). The velocity of EM wave propagation is determined by the dielectric properties of the medium. Just as light (which is a form of EM radiation) is slowed down and refracted when it travels from air into water, radar waves are refracted or bent by material below the antenna. If the dielectric changes underground are abrupt for example going from dry gravel into the water table, the change will appear as an interface or strong horizon on the resulting radar section. Dielectric losses occur in water because the EM energy produces mechanical rotation of the water molecule under the influence of an electrical field. With GPR surveys, the continued loss of energy by the radar pulse in water-saturated material means that penetration depth is greatly limited.

Conductivity of the substrate is the most important factor determining the rate of signal attenuation. Soils or materials with high conductivity (generally > 10 mS m^{-1}) (Table 1) will cause rapid dissipation of the radar pulses *via* the transformation of EM energy into heat when ions are pushed through the medium by the electric field induced by the radar pulse. Highly conductive materials close to the surface will end up reflecting radar energy back to the antenna, and nothing will be seen below this reflector. Clays in the substrate make it highly conductive, and thus signal loss is greatest in clayey soils.

SPECTRAL REFLECTANCE

Sensing by reflectance of electromagnetic radiation at certain wavebands has found most application in monitoring vegetation and crop cover. This application has already been reviewed for MAFF (Dampney *et al*., 1998). Commonly, measurement of specific crop characteristics involve the use of some form of vegetation index, the simplest of which is a ratio of "near infrared" (750 nm) to "red" (650 nm) responses (Steven and Clark, 1990) and includes a component due to soil reflection.

The reflectance from soil surfaces is not simply a function of the colour of the mineral particles, but depends also on its organic matter content, moisture and structure. Jansinski & Eagleson (1989) actually obtained three "soil lines": for soil minerals, soil moisture, and soil shadow respectively. Rondeaux *et al.* (1996) measured the reflectance from 26 soil samples at 660 and 865 nm and found that using the reflectance ratio all soil groups, except those with very high organic matter, fitted a single universal line. The ratio of reflectance at red and NIR wavelengths has thus been shown to be linear for all soil types (Baret *et al*., 1993) and has become known as the "soil line".

The measurement of soil variables using spectral reflectance is likely to be restricted to those soil features that occur at or close to the soil surface. Reflectance of incident radiation is not affected by soil properties at depth. Future potential is largely limited to the possibility of being able to sense topsoil organic matter and surface moisture content. More immediately, the main application is likely to be 'anomaly detection', in which "sensed" information should allow contrasting field areas to be mapped and considered for different management practice (Barnes *et al.*, 1996).

MATERIALS AND METHODS

The soil sensors were evaluated in cereal fields of contrasting soil type that had also been used in recently completed HGCA Project 2116 ('Developing a cost-effective procedure for investigating within-field variation of soil conditions'*)*. The fields were selected as they showed a high level of soil variability. Some variability was quite obvious but some was too subtle to be visible or predictable from the surface.

Sites

Two sites were measured in the cropping year 1999/2000 (Year 1) and two others in the cropping year 2000/01, though the advent of 'Foot and Mouth Disease' (FMD) in February 2001 caused a delay in the completion of measurements until the following season of 2001/02 (Year 2) on the same sites.

The two fields used in Year 1 were located at Lodge Farm, Shefford, Bedfordshire and Heydour Lodge Farm, Grantham, Lincolnshire, in fields named "Shagsby 4" (grid reference TL 108400) and "Field 107" (grid reference SK 996373) respectively, and known in this report by the acronyms; ***SHG*** and ***HLF***.

The two fields used in Year 2 were located at Shuttleworth Agricultural College, Old Warden, Bedfordshire and Crowmarsh Battle Farm, Benson, Oxfordshire, in fields named "Football field" (grid reference TL 142447)) and "The Clays"(grid reference SU 637915), and known in this report by the acronyms; ***FTB*** and ***CLY*** respectively.

All four sites have been predominantly in winter cereals in recent years and were chosen initially for project 2116, because several years' yield maps of winter wheat were available. All fields were sown to winter wheat during the year of study and the yield of that year was also mapped. Topographic details of the fields are discussed below in relation to the maps of the soil boundaries and position of soil sampling (Figs. 3 – 6) in the 'Results and Discussion' section.

Soils

Year 1 - *SHG*
The *SHG* site is characterised by soils of the Bearsted Association. Soil series of Bearsted, Waterstock & Hanslope, have been mapped over the major part of the site, but it also has inclusions of Evesham and Ludford. A brief description of these soils is given in Table 2, and fuller descriptions can be found in Hodge *et al.* (1984). A detailed soil map of this site was made in 1988 from 1 observation per 2 ha (Fig. 2), which also shows the field boundary.

The highest point is towards the lower right-hand side and the land generally falls away towards the topmost left-hand corner. There is a distinct but shallow valley feature running along the left hand edge of the Waterstock map unit. At the left hand edge of the field is a mature woodland and the right-hand boundary is an farm road between fields.

Table 2. Soil profile descriptions in (*SHG*).

Soil series	Topsoil texture	Subsoil texture	Wetness class (Hodge *et al.* 1984)
Evesham	Clay to ~25 cm; stoneless; calcareous.	Clay to > 1.2 m; stoneless; calcareous	III
Hanslope	Clay or clay loam to ~25 cm; slightly stony; calcareous.	Clay to >1.2 m; slightly stony; calcareous.	II
Waterstock	Clay loam to ~25 cm; slightly stony; non-calcareous.	Clay loam to ~80 cm, clay loam or sandy loam to > 1.2 m; slightly stony; non-calcareous.	II
Bearsted	Sandy loam or sandy silt loam to ~25 cm; slightly stony; non-calcareous.	Sandy loam to ~45 cm, loamy sand to ~70 cm over sandy parent material to >1.2 m; slightly stony; non-calcareous.	I
Ludford	Sandy silt loam or clay loam to ~25 cm; slightly stony; non-calcareous.	Clay loam to > 1.2 m; slightly stony; non-calcareous.	I

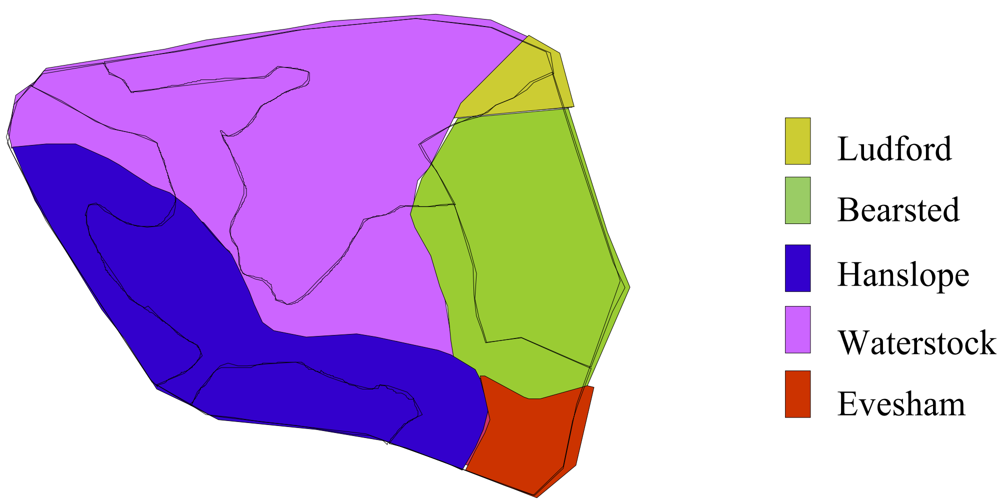


Fig. 2 Soil series mapped within the boundary of *SHG* field.

Year 1 - *HLF*

The soils at *HLF* are characterised by soils of the Elmton Association. A minor part of the site is occupied by Elmton and Aberford soil series (Hodge *et al.*, 1984), which are calcareous clay loams over magnesian limestone at either 25 cm depth (Elmton) or below 55 cm depth (Aberford), but the major part by shallow sandy loams of the Cranwell series.

Year 2 - *FTB*

The *FTB* site was mapped as the Bearsted Association. Soils were Bearsted series on the upper part of the slope, which is a sandy loam over sandy loam and sandy parent material, but also had the wetter clay loam over clay soils of the Oxpasture series at the bottom of the slope in the north of the field.

Year 2 - *CLY*

The *CLY* site was mapped as the Coombe Association over chalk. The site was a fairly uniform well drained shallow clay of the Wallop series (Jarvis *et al*., 1984), but of variable depth as the slope of the field falls away from an underlying chalk ridge in the east.

Soil measurements

Soil measurements carried out on site consisted of *in situ* measurements of the soil water regime, and laboratory measurements of physical properties on removed cores so that soil water properties could be calculated using pedo-transfer functions. Five cores were taken from each of four transects on the four sites, distributed along crop 'tramlines' such that they covered potential changes in soil properties and types. Measurements of soil texture, bulk density and soil organic carbon, were made on both topsoil and subsoil sections of these cores, and these used to calibrate hand-texturing of further cores taken randomly across each field. A further 50 cores were taken at *SHG*, 54 at *HLF*, 61 at *FTB* and 100 at *CLY*. In addition these were supplemented with data taken under HGCA project 2116 (led by Silsoe Research Institute) and also HGCA project 2298 (led by Reading University) on shared sites (*CLY* and *FTB).*

'SENTEK' Diviner tubes (Sentek Pty Ltd., South Australia; distributed by Peter White Water Management, Ipswich, IP6 9JS, UK) were installed into the 20 core-holes (to a depth of at least 1.2 m) on the four transects at each site, for soil hydrology measurements. This allowed soil moisture measurements to be taken at 10cm intervals down to 1m depth.

Soil moisture measurements were taken at approximately the same time as the EMI scans using a SENTEK DIVINER 2000 soil moisture probe. The DIVINER 2000 sensor utilises electrical capacitance to measure soil moisture. A high frequency electrical field is created around each sensor, extending through the access tube into the soil. The volumetric soil moisture content is measured by responses to changes in the dielectric constant of the soil. The capacitance of soil increases considerably with an increase in the number of soil water molecules, which are free to relax as their electric dipoles respond to the capacitor sensors field reversal. This measurement is proportional to capacitance and is also called specific polarisation or electric dipole moment per unit volume.

ELECTRO-MAGNETIC INDUCTION (EMI) SENSOR

Commercially available dual coil EMI systems have been developed such as the EM 38 manufactured by Geonics Ltd (Plate 1). These instruments directly read H_r/H_t, and thus EC_a, (McNeill 1980), and are either carried across a field, or towed behind an ATV in a lightweight non-metallic cart (Plate 2). Using simultaneous GPS measurements, very accurate maps can be recorded with an interval between readings of the order of 3 m.

EMI mapping

In this study, the EM38 (and GPS antenna) was towed in a small non-metallic cart about 3 m behind a "quad-bike" ATV to avoid interference with the GPS signal from the ATV. Measurements were made across each field when the soils were at field capacity (February / March), and also when they were thought to be near maximum soil moisture deficit after harvest (July / August / September) between 2000 and 2002. Measurements were made along transects at 6 m spacing, parallel to crop tramlines, at approximately 3 m intervals when the ATV was driven at 15 kph. At one site (*CLY*) it proved impossible to obtain an adequate GPS signal to complete coverage of the study area within the field, and so a minimum data-set was supplement by 'walking' the instrument along the transects at moisture deficit to ensure a comparative EC_a reading for moisture measurements (as in Plate 1).

Plate 1. The Geonics EM38 being walked along a transect.

Plate 2. The Geonics EM38 housed in a wooden cart with GPS receiver mounted, being towed behind an ATV.

EMI and Geostatistics

The collection of EC_a data in this study was carried out at transect spacings far narrower than would be contemplated for a commercial survey. This was deliberate to enable the progressive degradation from a dense data-set, so that more statistical information could be gathered and the optimum spacing identified. At such densities geostatistics would not normally be required to predict the EC_a for points in the field, but even so, it is likely that geostatistical methods for analysis and estimation will be useful for three reasons, *viz*:

i. In commercial practice the time spent surveying a field will depend in part on the spacing between passes of the sensor. The geostatistical method of kriging may be used to estimate EC_a at points in between passes. Because the mean-squared error of kriged estimates from a given array of data may be computed if the variogram is known (McBratney *et al.,* 1981) it is possible to find the maximum grid spacing such that the error of the kriged EC_a is acceptable. This allows the most cost-effective EC_a survey to be planned.

ii. Our basic data are point measurements of EC_a. In practice we are likely to be interested in the mean EC_a of small regions (blocks) coinciding, for example, with pixels from satellite imagery, or the basic units of a treatment map. Block kriging allows us to 'scale up' point estimates of EC_a to block estimates with known and minimum error.

iii. Other geostatistical methods might be used with sensor data in the future. For example, they might be usefully co-kriged with yield or spectral data.

The primary objective of the geostatistical analysis of these EC_a data was to address point (i) above and to characterise the relationship between the intensity of sampling and the precision with which apparent conductivity may be estimated for points and blocks in the field. This should help us to answer the basic question 'how widely spaced should the passes of an EMI survey be?'

Geostatistical analysis of soil properties is based on the assumption that a variable, z, measured at a location, $\mathbf{x}$, may be treated as a realisation of a random function, denoted by $Z(\mathbf{x})$. The analysis is possible if the random function is intrinsic, that is if

$$\mathrm{E}\left[Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})\right] = 0, \tag{2}$$

and

$$2\gamma(\mathbf{h}) = \mathrm{E}\left[\left\{Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})\right\}^2\right] \tag{3}$$

depends only on the spatial separation or lag $\mathbf{h}$. The function $\gamma(\mathbf{h})$ is the variogram.

An estimate of the variogram is needed for geostatistical estimation (kriging) and the design of optimal sampling strategies (McBratney *et al.*, 1981). The most widely used estimator of the variogram is due to Matheron (1962):

$$2\hat{\gamma}_{\mathrm{M}}(\mathbf{h}) = \frac{1}{N(\mathbf{h})}\sum\left\{z(\mathbf{x}) - z(\mathbf{x}+\mathbf{h})\right\}^2 , \tag{4}$$

where $N(\mathbf{h})$ pairs of observations among the available data are separated by lag $\mathbf{h}$.

This estimator is asymptotically unbiased for any intrinsic random function, but because it is based on squared differences among data it is very sensitive to outlying values of z. A single outlier can distort the estimate of the variogram. This in turn affects calculated kriging variances and so the intensity of sampling designs obtained with the variogram. A number of robust estimators of the variogram have been proposed as alternatives to Matheron's, and Lark (2000a) evaluated three of the major ones. He showed that they could perform well, and also how alternative variograms for a data-set may be assessed.

The initial analyses in this study were conducted on the most closely spaced data-sets for fields *SHG* and *HLF*. Histograms and descriptive statistics of the data were examined with and without log-transformation. The data were then divided at random into a prediction set and a validation set. Variograms of the prediction data (raw and log-transformed) were then evaluated using Matheron's estimator (Equation 3), Cressie and Hawkins (1980) estimator, Dowd's (1984) estimator and that of Genton (1998). Models were fitted to the estimates of the variogram using weighted least-squares.

The prediction data sets were then thinned to a smaller subset for kriging. This was done by stratified random sampling. The reason for doing this was so that the kriged estimates of the validation data were derived from prediction data at different lag distances (so testing the estimated variograms over a range of lags) without creating kriging matrices too large to solve in reasonable time. Kriged estimates were then obtained for each validation datum, by kriging with each variogram. This yields, for each variogram and each validation datum, $z(\mathbf{x})$, a kriged estimate $\hat{Z}(\mathbf{x})$ and a kriging variance $\sigma^2_K(\mathbf{x})$ — an estimate of the error variance of this prediction. From these values a new variable is computed:

$$\theta(\mathbf{x}) = \frac{\left\{z(\mathbf{x}) - \hat{Z}(\mathbf{x})\right\}^2}{\sigma^2_K(\mathbf{x})}. \tag{5}$$

If the intrinsic hypothesis applies and the variogram used in kriging is correct then the expected value of $\theta(\mathbf{x})$ is 1. However, $\theta(\mathbf{x})$ is very susceptible to outliers. Lark (2000a) showed that it is better to evaluate the median of $\theta(\mathbf{x})$, with an expected value of 0.455. This was shown to be a robust measure of the validity of a variogram model in kriging, and the median of $\theta(\mathbf{x})$ was obtained from all the validation data for each variogram. If this median is significantly smaller than 0.455 this suggests that the variogram is overestimated (due probably to outliers) and so gives inflated kriging variances.

Having evaluated alternative variograms of the EC_a data, kriging variances were calculated assuming that EC_a is kriged from data on passes of the sensor at different spacings. This can be done since the kriging variance depends only on the disposition of the sites from which the kriged estimate is derived and the variogram. It was assumed that the data are collected at 3m intervals along each pass. Kriging variances were calculated for the point estimate at a site equidistant between adjacent passes, and for the estimate of a 10m square block centred at the same point.

This procedure was then followed using the other data-sets collected at each site (vertical setting). Because these sets were smaller all the data were used for variogram estimation and the $\theta(\mathbf{x})$ statistic was estimated by cross-validation (see Lark, 2000a).

Seasonal stability of EMI Measurements

An important question about EC_a data is whether the spatial pattern that is revealed in a single data set is likely to be stable over time. If we wish to use EC_a data to delineate management zones are we likely to do better at certain times of year (*e.g.* when the season is at field capacity rather than when there is a substantial deficit)? Some of the results presented here for quantitative analysis of EC_a data and soil measurements suggest that bigger differences are found in spring, but we want a more general method for assessing the stability of patterns, and we have employed two approaches to do this.

In an initial study we compared the measurements of EC_a (vertical mode) at *HLF* on the two dates of sampling, by multivariate geostatistical analysis. We computed robust pseudo cross-variograms of EC_a on the two dates, which describe the spatial variation and joint variation of the two variables. We then used these to estimate change in EC_a by cokriging. Details of these statistical methods are given elsewhere (Lark, 2002).

To avoid the limitations of the linear model of co-regionalisation we used an alternative method. We overlaid the EC_a data sets for the two dates to be compared and sorted them into vectors — by pairing observations from the two dates which are closest together in space (provided they are no further apart than 5 metres). We then performed a fuzzy cluster analysis on these vectors. This method is not described in detail here, but more information is provided by Lark (2001). The aim of the method is to discover distinct clusters in the data – *i.e.* groups of vectors with similar EC_a values at the two dates. If there are regions of the field with (relatively) large values of EC_a data on both dates then these should form a distinct cluster, similarly areas of small EC_a will form a cluster. If there is any complex temporal variation — e.g. areas where EC_a 'flips' from large to small values — then these will also form a distinct cluster themselves.

Relating EMI to Soil Measurements

The objective of the analyses reported here was to obtain insight into the soil properties which contribute to the observed variation in measured EC_a.

The basic method used in this analysis was regression analysis, whereby a response variable (EC_a here) is modelled as a linear combination of input variables, and a random error variable. It would be possible to regress EC_a on a set of soil properties s_1 , s_2, … s_m giving rise to an equation:

$$EC_a = a + b_1\, s_1 + b_2\, s_2 + \ldots + b_m\, s_m + e, \qquad (6)$$

where a and the b s are regression coefficients and e are the error. This equation would be useful for predicting the EC_a for a site from the soil properties s_1 , s_2, … s_m but it would be of little use for generating insight into what causes the variation in EC_a (Webster, 1997). This is because the variables s_1 , s_2, … s_m are not statistically independent of each other. Thus, while the coefficients b represent a least-squares solution to the problem of predicting EC_a from s_1 , s_2, … s_m , we may not interpret the coefficient b_i as a measure of the importance of soil property s_i (e.g. clay content in the topsoil) in physically determining the signal EC_a. This is because the clay content of the topsoil will not be independent of, for example, the bulk density of the topsoil. Both these variables may make distinct physical contributions to the EC_a of the soil, but the regression equation is not able to disentangle these because the variables do not vary independently of each other. Therefore the coefficient which is estimated for clay content will change as other variables are added to, or dropped from, the regression, and so the variable does more or less work accounting for the effects on EC_a of variables not included in the model with which it is correlated.

Given this constraint, we carried out two separate analyses on the EC_a and soil data. In the first we fitted a regression model to predict EC_a from all the measured soil variables. We then fitted a new model corresponding to each of the soil variables in which it alone was dropped from the set of predictors. Each reduced model may then be compared to the full model (using the Akaike information criterion, discussed below) if the model with a term dropped is deemed to give a poorer fit to the observed EC_a measurements, then we may conclude that that variable has a *partial effect* (direct or indirect) on the EC_a , that is to say an effect *which is not correlated with the effects of any terms remaining in the model.* This qualification is critical, if we identify no effect of a variable independent of other terms, this does not necessarily mean that the variable bears no relation to EC_a, it may be that it is of physical importance, but is sufficiently strongly correlated with terms still in the model for its effect to be expressed through them. We may therefore only draw positive conclusions from the analysis.

In the second approach we started by computing the principal components of the soil variables. The principal components of a data set with *m* variables are *m* new variables formed by an orthogonal linear transformation of the original ones, that is to say we can think of the principal components as a rigid rotation of the original variables. While an infinite number of sets of orthogonal transformations of the variables are possible, the principal components are constrained (i) to be un-correlated with each other and (ii) to account for as much of the variability in as few of the principal components as possible. Details of the method can be found elsewhere, here we conducted principal component analysis by finding eigen-values and vectors of the correlation matrix of the original variables which effectively analyses the original variables on a common dimensionless scale.

Having formed principal components of the original soil variables we conducted regression analyses on the principal component scores. Because the principal components are un-correlated we may identify the components which account for most variation in EC_a by considering the *t* statistic for each in turn. We may then examine the coefficients whereby the important principal components are formed from the original soil variables, and see whether they admit an interpretation in terms of the pattern of soil variability which they seem to describe.

It is important to note that the regression was not done by ordinary least squares (OLS). OLS regression is not suitable for the analysis of spatial data sets which have not been collected by random sampling. All these data were collected on systematic grids or transects. In these circumstances OLS can seriously overestimate the significance of a regression relationship.

The ML regression procedure was used instead (Lark, 2000b). The ML estimate of the vector of regression coefficients, **b**, is given by

$$\mathbf{b} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{X}\right)^{-1}\left(\mathbf{X}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{y}\right). \qquad (7)$$

X and **y** are respectively a matrix and a vector of *n* observations of the independent variable(s) (ancillary variables) and the dependent variable (soil property). **A** is a correlation matrix for the *n* errors. This may be specified by a variogram function for the errors, determined by the spatial parameter and the ratio of the sill to the nugget component of the error variance. The maximum likelihood estimate of the error variance, conditional on **A** is given by:

$$\hat{\sigma}^2 = \frac{1}{n}\left(\mathbf{Y} - \mathbf{Xb}\right)^{\mathrm{T}}\mathbf{A}^{-1}\left(\mathbf{Y} - \mathbf{Xb}\right), \qquad (8)$$

and the maximum likelihood estimates of the parameters of the error variogram, the regression coefficients and the error variance is then obtained by a numerical maximisation of the likelihood function:

$$-\log\left|\mathbf{A}\right| - n\log\hat{\sigma}^2. \qquad (9)$$

with respect to the variogram parameters. The significance of the resulting regression model was then tested by computing the Wald statistic:

$$W = \mathbf{b}^{\mathrm{T}}\mathbf{L}\ \ \mathbf{b} \qquad (10)$$

where $\mathbf{L}$ is the Fisher information matrix. For a regression model with *k* predictors, *W* is approximately distributed as χ^2 with *k* degrees of freedom. This tests the null hypothesis that the regression coefficients are all zero.

When regressions of one variable (EC_a here) have been computed on alternative sets of independent variables they cannot simply be compared on their error variances, since increasing the number of predictors almost invariably reduces the error variance. Instead the models were compared on the basis of the Akaike information criterion, AIC, (Akaike, 1973). Selecting from among models fitted to the same data on the basis of the minimum AIC is a parsimonious procedure. Additional predictors are only accepted if the improvement in fit is sufficient to justify them by a likelihood–based rule. Here the AIC was applied by computing the statistic

$$a = \log\left|\mathbf{A}\right| + n\log\sigma^2 + 2k, \qquad (11)$$

where k is the number of independent variables in the prediction. The value of a will be least for the regression with the smallest AIC.

The ML procedure outlined above was used in most cases. However, in the analysis of data sets including the Diviner measurements of soil water the numbers of observations were very small. In these circumstances we used the REML directive in Genstat (Payne *et al.*, 1988) for spatial modelling. This indicated no evidence for spatial correlation of the error term, so we conducted the analysis using OLS regression in Genstat, and computed the corrected AIC for small sample sizes proposed by Bedrick and Tsai (1994).

GROUND PENETRATING RADAR (GPR)

Two GPR systems were operated during the project. At *HLF* field in year 1 measurements were made using a 'Sensors & Software' 'pulseEKKO 1000 GPR' using a 450 MHz antenna, whilst at the *CLY* field in year 2 a 'Mala Geosciences' 'Ramac system' with a 500 MHz shielded antenna was used, to give optimal depth and target resolution (Plate 3).

Plate 3. The Mala Geosciences GPR being dragged across a concrete surface.

Reflectance Measurements of Sub-surface Features

The aim of the survey was to profile shallow soil layers. The simplest means of doing this is to traverse the ground surface with the antenna in direct contact with the soil, and measure the reflected signal. To this end the GPR systems were physically dragged at walking pace (Plate 3) along the four 20 m transects located along selected tramlines within the study areas at *HLF* and *CLY* fields. The transects at *HLF* were located in positions where it was considered changes in subsoil features may be of interest, and at *CLY* they were located continuously so that one long transect was sampled running through the centre of the field from the chalk ridge and down to the lower elevation. Measurements were made across each field when the soils were at field capacity (February / March), and also when they were thought to be near maximum soil moisture deficit after harvest (July / August / September) in 2000 (*HLF*) and 2002 (*CLY*).

Novel Measurements with GPR

At *HLF* field in year one a small one-off survey was conducted using GPR in novel manner. This is known as "air launched mode" when it is hoisted to a fixed height above the ground surface. This was conducted over a small (50 m x 10 m) patch across a valley feature, to investigate the possible differentiation of soil types. The GPR antenna was mounted at a height of 1 m above the soil surface on the back of an ATV and driven a slow speed across the area.

Spectral Reflectance

The principle of the system was to traverse a linear array of two-band radiometers systematically over a bare soil surface so that a regular spaced grid of measurement points could be generated. To this end a tractor was equipped with a 24m boom from a conventional agricultural sprayer. Radiometers were mounted at 4m intervals across the boom to detect radiation reflected from the soil surface. A radiometer with a cosine-correcting filter was mounted at the centre of the boom to measure down-welling radiation in order to compensate the soil radiometers for variation in ambient radiation. The boom could be set at different heights relative to the ground so that the sensing area for each radiometer could be varied. Typically the height was set to give a sensing circle of diameter 0.5m.

The radiometers were Skye Instruments type SKR1800 dual channel sensors fitted with narrow band interference filters centred at 660nm (visible red) and 730nm (near infrared), and the filter bandwidths were 25nm. The radiometer signals were carried by cable to a multi-channel A-D converter expansion board in a PC, mounted in a tractor cab. Sampling of the signals was initiated by an interrupt generated from a GPS receiver at a rate of 1Hz. The GPS receiver aerial was mounted on the tractor cab roof at a measured distance from the centre-line of the boom.

The tractor was driven along 24 m spaced tramlines across the SHG field during the post harvest period in the autumn of 2002, after the crop stubble had been ploughed in and the soil surface was bare. It was also essential that the soil surface was fairly flat with no standing water to avoid extraneous reflections. The file of radiometer measurements generated by traversing each tramline with the scanning radiometry system was imported into a spreadsheet so the measurements could be compensated, corrected, and synchronised with GPS location.

The tractor was driven at approximately 10-13 kph, and the data collection process for a 7 ha field was about 0.5 hours.

RESULTS AND DISCUSSION

SOIL SURVEY

After collection and analysis of the core samples from each site, it became apparent that the soil types dominant on each site did not quite match the prior assumptions about their series and distribution. These revised maps, based on approximately 4 observations per hectare, are reproduced here. Profile descriptions of each soil series mentioned are given in Appendix 1.

SHG site year 1.

Fig. 3 shows the revised soil type distribution for this field and should be compared with Fig.2. The topography of this site was such that the highest point was towards the bottom right hand corner of Fig. 3, with a ridge running along the right hand boundary. A valley feature that was often more moist than the surrounding boundary areas, but not as wet and cloddy as the southern boundary, ran from approximately the middle of the field towards the top left hand corner of Fig. 3.

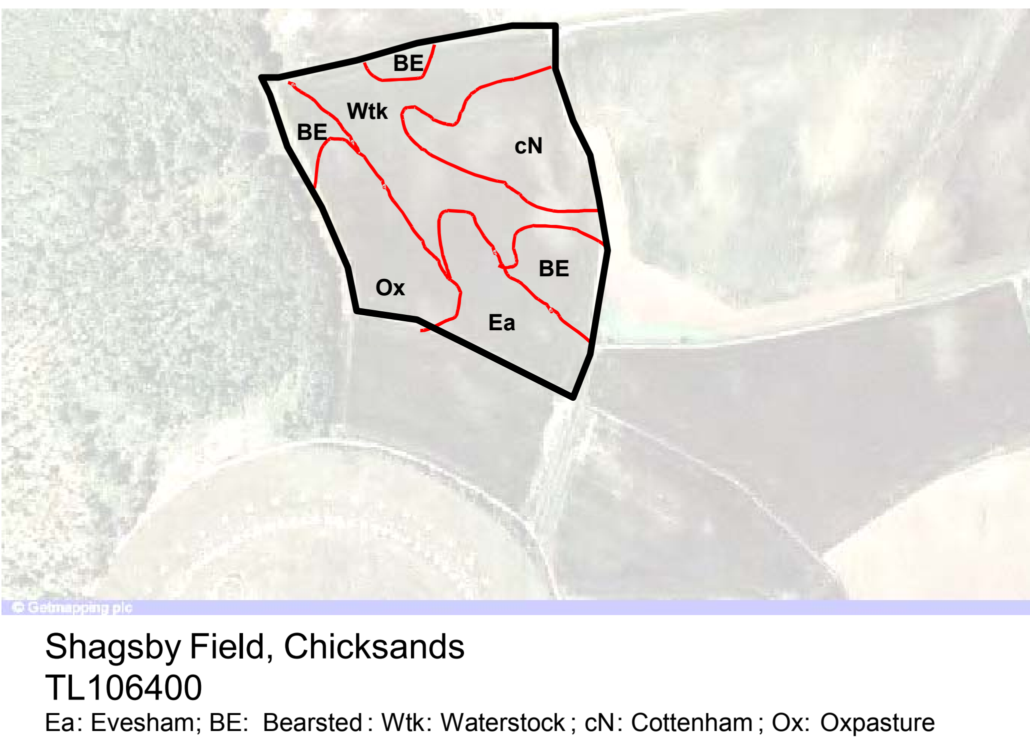
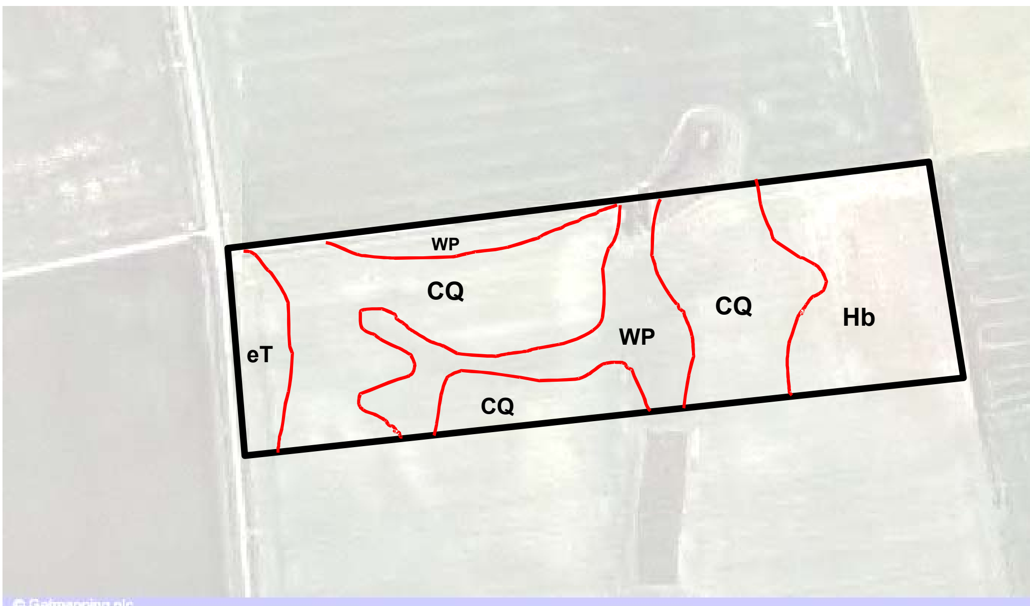


Fig. 3 Soil types at *SHG*.

Evesham and Oxpastures soils are clays and heavy clay loams over clay respectively, whereas Waterstock is a clay loam, and Bearsted and Cottenham soils are sandy loams and sands respectively (Appendix 1). The area receives an annual rainfall of approximately 570 mm, has about 101 days at field capacity and a maximum moisture deficit of 103 mm under wheat.

<u>*HLF* site year 1.</u>

Fig. 4 shows the soil types for the part of this field which was mapped using EMI and GPR instruments. The highest point is towards the top right hand corner of Fig. 4, with a plateau running along the right hand boundary where the Haselor soil is located. A deep valley feature ran the middle of the field (top to bottom) where the Wilford soil is located between Cranwell soils on the slopes to either side. The site rose again to another slightly lower plateau along the left hand edge of Fig. 4.



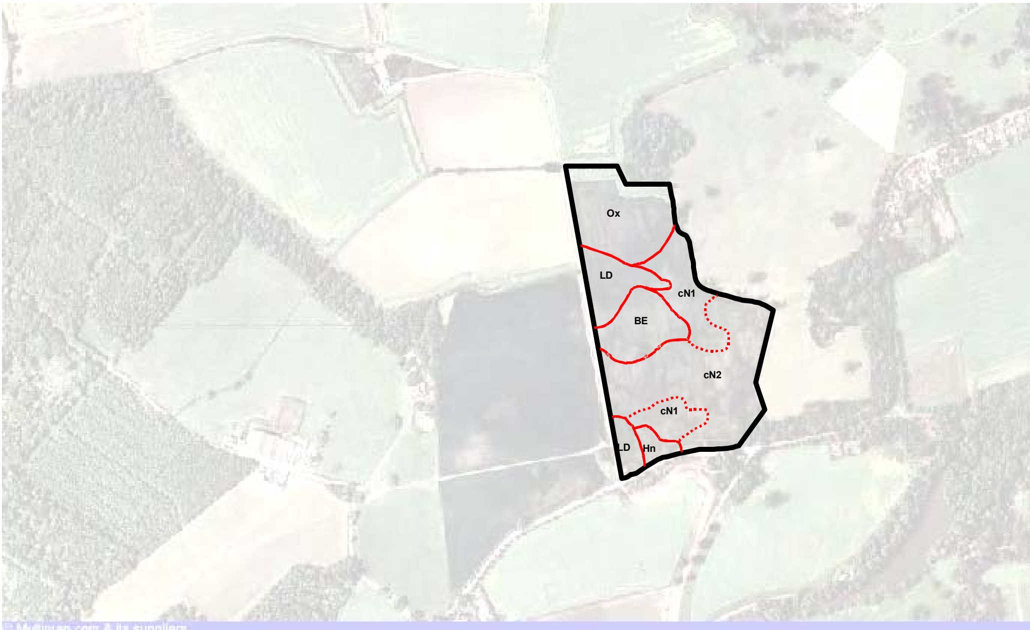
Field 107, Heydour Lodge
TF004372
eT: Elmton; CQ: Cranwell; WP: Wilsford; Hb: Haselor; Ox: Oxpasture

Fig.4   Soil types at *HLF*.

Haselor soils are heavy calcareous stony clay, and Elmton a shallow sandy clay loam over limestone. The Cranwell soil coming off the limestone ridges is a shallow stony sandy loam, and runs into the deep loamy sand of the Wilsford soil in valley areas (further soil description in Appendix 1). The site receives annual rainfall of approximately 652 mm, has about 139 days at field capacity and a maximum moisture deficit of 106 mm under wheat.

<u>*FTB* site year 2</u>.

Fig. 5 shows the soil type distribution for this field, though only the lower half was mapped by EMI. The topography of this site was such that the highest point of the field was along the bottom edge of Fig. 5, sloping gently where the Cottenham soil series is predominant. This slope steepens at the more mixed area of Bearsted and Ludford soils in the middle of the picture, until it levels off again at a lower elevation, where the Oxpasture soil is located.



Football Field, Old Warden
TL142447
BE: Bearsted ; cN1: Cottenham (sandy loam top);
Hn: Hanslope ; cN2: Cottenham (loamy sand top); Ox: Oxpasture

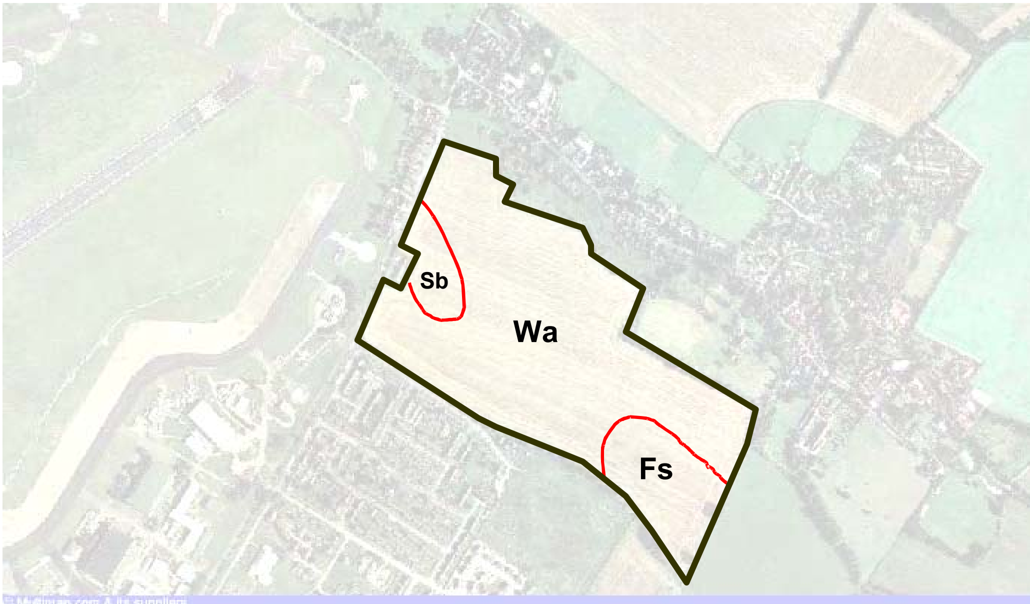
Fig. 5 Soil types at *FTB* .

Cottenham soil is a light loamy sand or sandy loam developed over a deep loamy sand to sand parent material. Sandy silt loams of Bearsted and clay loam Ludford series intergrade on the slope between the Cottenham soils and the heavier clay loam of the Oxpasture soil at the lower level (further soil description in Appendix 1). The site receives rainfall of approximately 555 mm a^{-1}, has about 95 days of field capacity and a maximum moisture deficit of 118 mm under wheat.

CLY site year 2.

Fig. 6 shows the soil type distribution for this field, though the study area comprised only a narrow strip running from the lower edge of the Soham soil series on the top left field boundary to the top edge of the Frilsham soil at the lower right hand edge of the field. The highest point is along the bottom right hand edge of Fig. 6 along a chalk ridge, sloping steadily towards the top left hand edge of the field, and more steeply so into the top corner.

Most of the field is soil of the Wallop series, which is a fairly shallow silty clay loam to silty clay developed over fragmented chalk. The main variation across the field is the depth of this material over chalk. There are smaller areas of sandy clay loam on both the ridge, (Frilsham) and lower plain areas (Soham)(Appendix 1). The site receives annual rainfall of approximately 644 mm, has about 139 days at field capacity and a maximum moisture deficit of 122 mm under wheat.



The Clays, Benson
SU640915
Wa: Wallop; Fs: Frilsham; Sb: Soham

Fig. 6 Soil types at *CLY* .

<u>SOIL PHYSICAL PROPERTIES AND HYDROLOGY</u>

The location of soil moisture measurement tubes (‘Diviner’ tubes) on two sites *SHG* and *CLY* are shown here (Figs 7 & 9) as examples of their distribution across features of interest and/or contrasting soil/management zones in fields.

The distribution across *SHG* (Fig. 7) relative to the revised soil boundaries indicates that tubes 1 – 10 are probably located within heavier clay/clay loam soils, whilst 11 – 15 are in the lighter clay loam and 16 – 20 sandy loam or loamy sand. The moisture contents of the top- and sub-soils during the summer moisture deficit period are shown in Fig. 8 where generally reduced levels in the sandy loam Cottenham series soil are seen in comparison with the rest of the field. The feature that does stand out however, is the somewhat higher levels of moisture in the sub-soil at tubes 6, 12, 13 and 14 on two separate transects spanning three soil types. All however, are in depressed areas of the field; 6 as it slopes towards the left hand edge in Fig. 7 and 12, 13 and 14 in a valley feature running from the middle of the field toward the top left-hand edge. This should be compared with Fig. 12 to see that they lie just within higher reading EC_a areas.

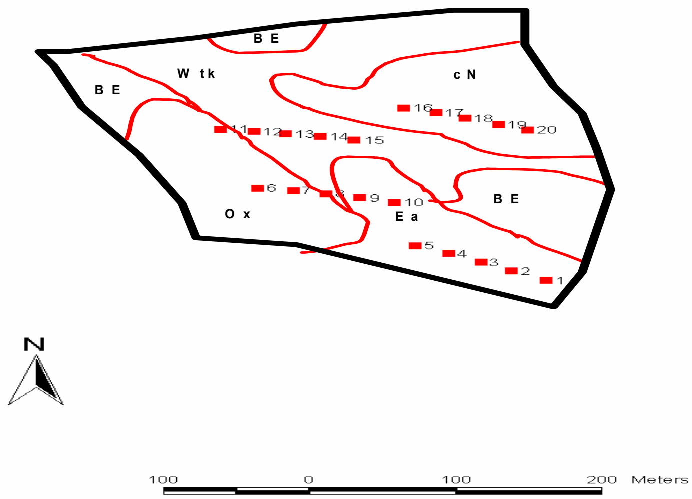


Fig. 7 Location of 'Diviner' tubes and soil coring transects at *SHG*.

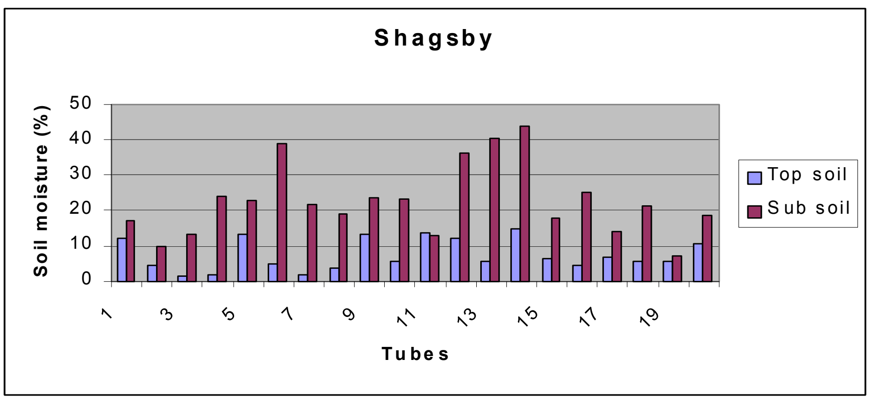


Fig. 8 Soil moisture content (% g/g) for top- & sub-soil at *SHG* on 15/08/2000.

The distribution of tubes across *CLY* is shown in Fig. 9 relative to the soil types indicating that tubes 1 – 20 are probably located within only one soil type (a clay), which runs from a low-lying flat area at 1 to the top of a chalk ridge at 20. The moisture contents of the top- and sub-soils during the summer moisture deficit period are shown in Fig. 10 where very low levels are recorded, but particularly so in the in tubes 5 – 8 and 19 – 21 for the subsoil. Tubes 19 – 21 are on the very top of the ridge where the soil is very shallow and tubes 5 – 8 are also in an area of shallow soil where the underlying chalk comes near to the surface. These areas are identified as small value EC_a areas by the use of the EMI instrument in vertical mode in Fig. 14a.

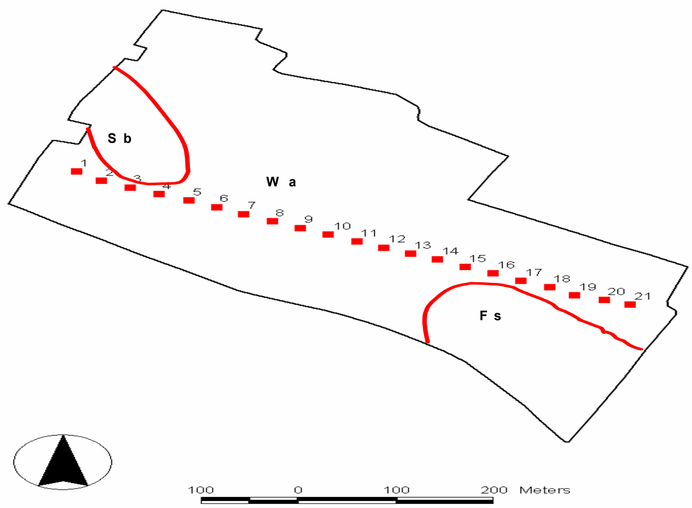


Fig. 9. Location of 'Diviner' tubes and soil coring transects at *CLY*.

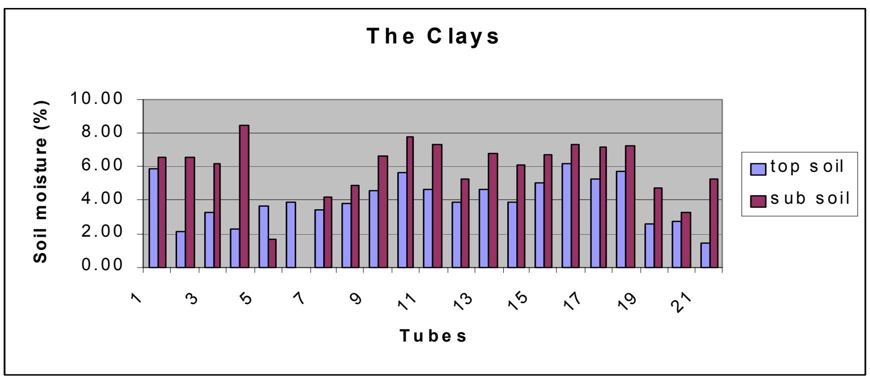


Fig. 10 Soil moisture content (% g/g) for top- & sub-soil at *CLY* on 26/06/2002.

Similarly the tubes in *HLF* identified the much higher moisture contents associated with the clay soil on that site, and also certain tubes on *FTB* had higher moisture contents in the subsoil showing where the clay inclusions were, which led to larger value EC_a readings in some areas in Fig.13a.

ELECTRO-MAGNETIC INDUCTION (EMI) SENSOR

EMI mapping

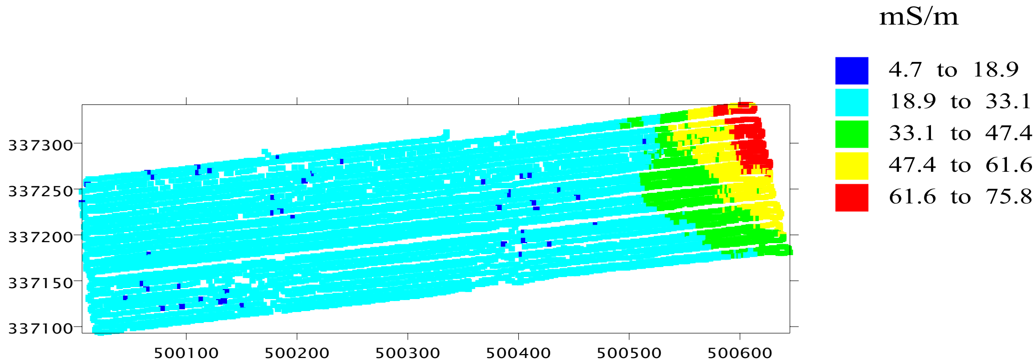
The output of the EMI scans can be displayed spatially related to the GPS co-ordinates to obtain apparent electrical conductivity maps of the site, such as those in Figs. 11 - 15 for the sites studied here.

Year 1

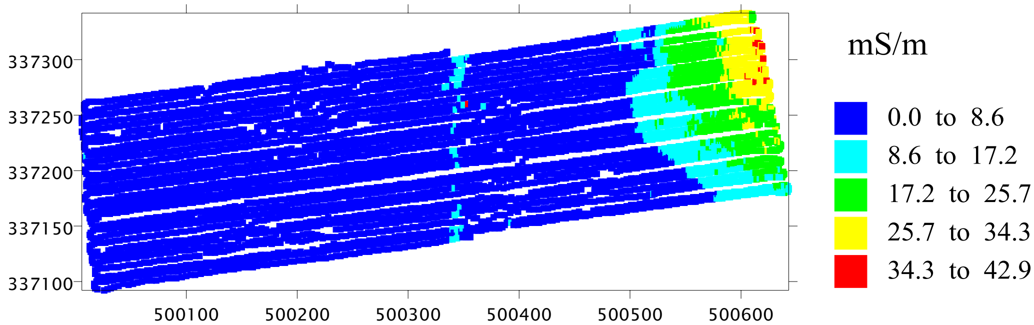
Fig. 11 shows the distribution of EC_a readings across the part of the *HLF* field studied (Fig. 4), on occasions near field capacity (date 1 = 16/03/00) and during harvest when the soil was dry (date 2 = 17/08/00). It should be noted that harvesting was still in progress on part of this field on date 2, which is why EMI surveying could not be extended to the right hand edge of the field in Fig. 11 c&d.

The maps of EC_a readings made on *SHG* field are shown in Fig.12. Again, readings were made at field capacity (date 1 = 11/02/00) and after harvest when the soil was dry (date 2 = 23/08/00), but only with the instrument in the vertical mode of operation. The second set of measurements (Fig. 12b) were made at a higher speed of ATV operation compared with the first set (Fig. 12a) (by a different operator), and serve to show how this has led to a much less dense data-set (459 points instead of 3564; King *et al.,* 2001).
It is instructive to compare visually the range of EC_a values mapped for each site with the relevant soil maps for the same sites. If this is done for *SHG* site (Figs. 3 & 12) then the value of EMI as a technique becomes clear, as there is a clear distinction of larger EC_a readings in the parts of the field dominated by the heavier clay loam soils of the Evesham and Oxpasture series, compared with lower values from the lighter soils of Waterstock and Bearsted series. This pattern holds true in general at both wet and dry times of the year suggesting that a large part of the signal from this field is governed by the clay content of the soil. This is also suggested by the fact that the site mean EC_a is only marginally smaller in the summer compared with the winter (Table 7), and leads to the identification of two distinct classes of response for this field (Fig. 18). Site hydrology also shapes the pattern however, as the valley feature in the Waterstock soil to the top left of Fig. 3 shows as an area of marginally higher readings in Fig. 12, and is probably due to subsoil moisture.

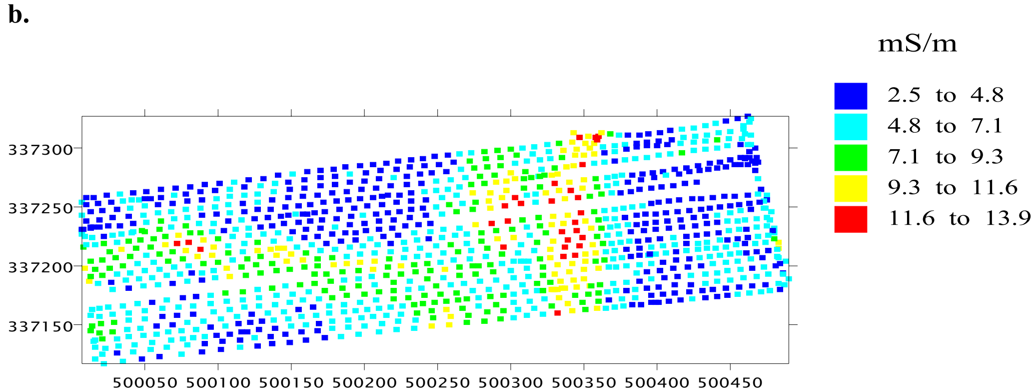
The comparison for the *HLF* site is a little different (Figs. 4 & 11) as there is only a small area of heavier soil and most variation across the site is topographical. At times of both field capacity and maximum moisture deficit, the EC_a readings across the *HLF* site are much smaller than *SHG* (Table 7), reflecting the sandier material and drier conditions. The area of clay loam soil topsoil at the top of the slope to the right of Fig. 11a&b however, showed comparable apparent conductivity to the clay loams on *SHG* (Fig. 12a). There were big differences recorded in the site mean between dates however (Table 7), due mainly to the second date not including the high conductivity clay area (cf. Fig 11a&b and 11c&d). However, the major part of



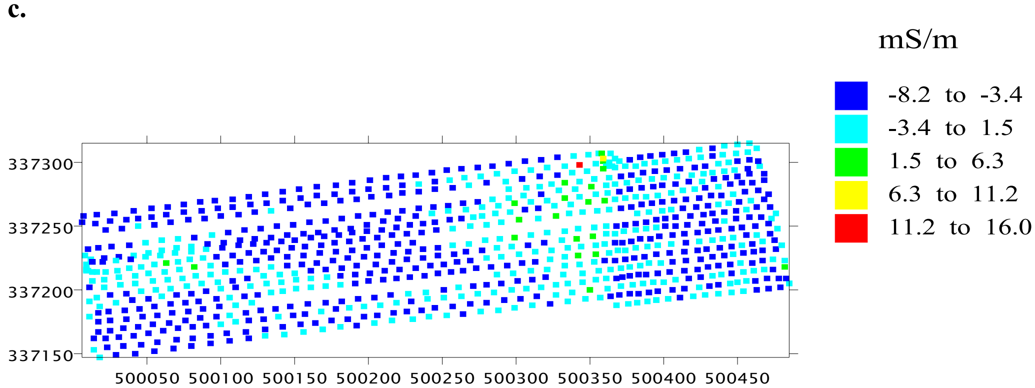
a.



b.

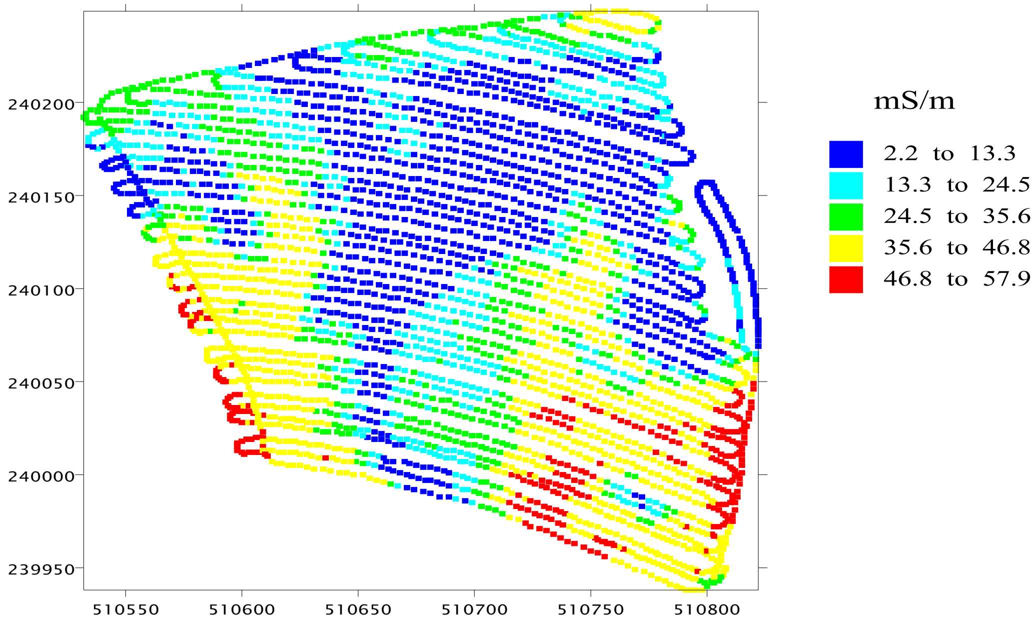


c.

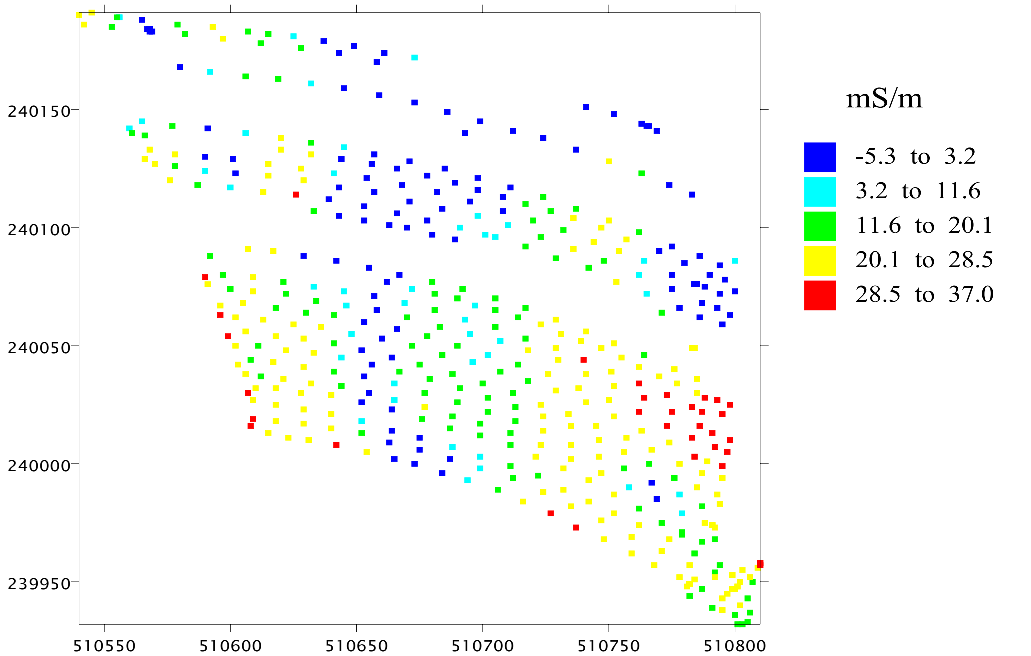


d.

Fig. 11. EMI scans of *HLF* field. **a**. = date 1 (16/03/00) vertical mode, **b** = date 1 horizontal mode, **c** = date 2 (17/08/00) vertical mode, **d** = date 2 horizontal mode. NB. Note from X-axis GPS co-ordinates, that **c** & **d** do not extend as far to the right hand side of the field as **a** & **b**.



a.



b.

Fig. 12. EMI scans of *SHG* field. **a**. = date 1 (11/02/00) vertical mode, **b** = date 2 (23/08/00) vertical mode.

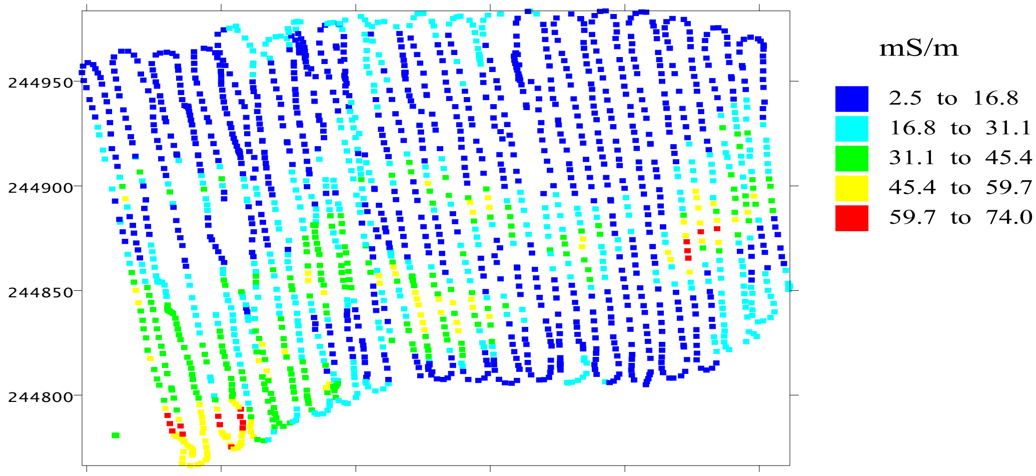
the field was uniformly low reading in both vertical and horizontal modes at field capacity (Fig. 11a&b), and readings were even more reduced during the dry part of the year (Fig. 11c&d) indicating that in this material the main contributor to the signal was moisture. Little pattern was visible in this part of the field in the vertical mode, but the main valley feature was visible in the horizontal mode when at field capacity (Fig 11b). This valley feature was even more apparent by the contrast in readings made during a period of moisture deficit (Fig. 11c&d), as were other less distinct valley features. This time the contrast was more apparent in the vertical mode (Fig. 11c) and coincides with areas of Wilsford soil which is deeper sand material between areas of shallow Cranwell series over limestone outcrops (Fig. 4). In this case the main contributor to the EMI signal is moisture held at depth in this material, and the EC_a is effectively soil depth as well as texture.

It should be noted from Fig. 11d that negative EC_a readings can be obtained from a survey. This occurs because they are not truly absolute readings, taking a zero datum in free air at one particular point in the field. In some cases the underlying geology may influence the zero reading slightly. For this reason comparisons between sites or dates should not be taken as indicating precise differences. Also, we have not computed ratios, or geometric means, of horizonal and vertical mode readings which can be informative (Lark *et al.,* 2003b) since these require an absolute scale.

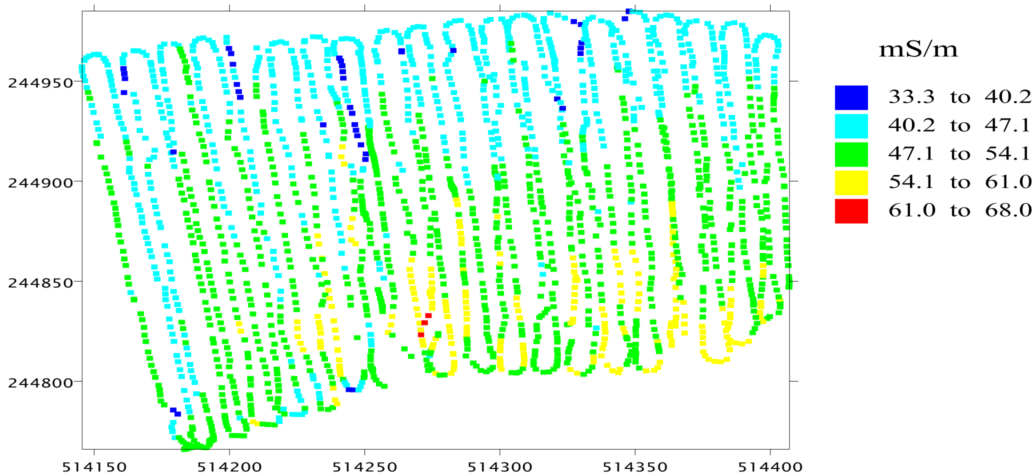
<u>Year 2</u>

Figure 13 displays the EC_a distribution across the *FTB* site at field capacity (date 2 = 21/03/02) and after harvest (date 1 = 19/09/01) when the bulk of the soil profile was dry. On this second date however, there was considerable surface water from rainfall that eventually prevented the horizontal mode readings from being taken (to avoid damage to the site in susceptible areas). The maps in Fig. 3 correspond to the higher part of the field, which is the lower portion of the area outlined in Fig. 5. The top right hand corner of Fig. 13 relates to the pointed middle section of the right hand side of the field in Fig. 5, and the top left approximates to the junction between Bearsted (BE) and Cottenham 2 (cN2) soils along the left hand edge in Fig. 5.

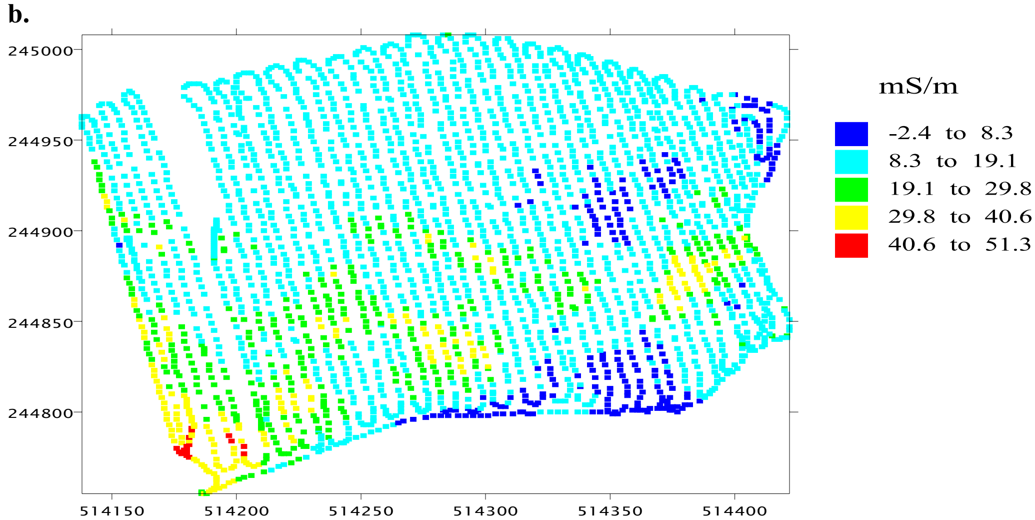
A comparison between the EC_a maps for *FTB* field (Fig. 13) and the soil map (Fig. 5, lower half) reflect the general uniformity of the studied half of the field that is mainly Cottenham series soils. This deep sandy material gives generally small value EC_a measurements in the vertical mode during both moisture deficit (Fig. 13a) and field capacity (Fig. 13c), except for small areas of more clayey materials in the lower left hand corner. However, there is a pattern of patches in the field where conductivity is somewhat higher than the rest, which is not however strongly reflected in the horizontal mode readings made during a moisture deficit period (Fig. 13b). This pattern is consistent over both dates. This rather suggests that there is a change in materials at depth (>2 m) in these areas which is probably in the mineral soil, because of its persistence. It has been noted in the past that there are areas of clay inclusions in the subsoil of this upper field area



a.

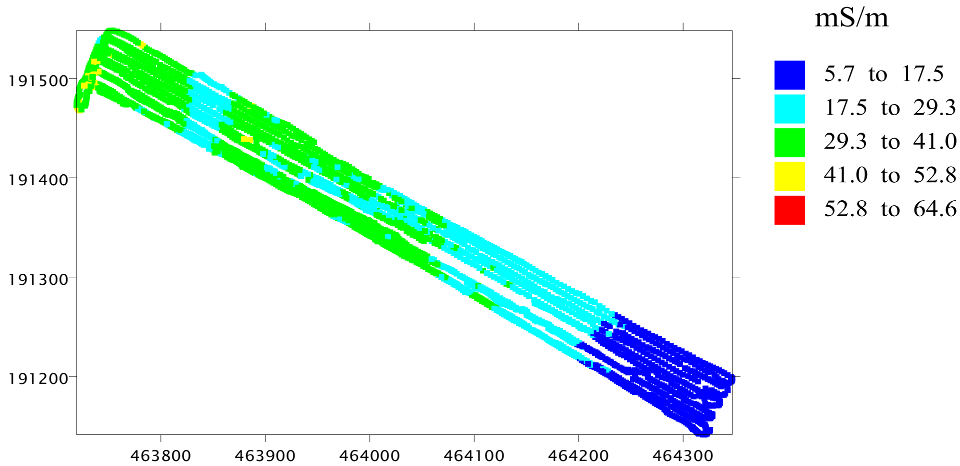


b.

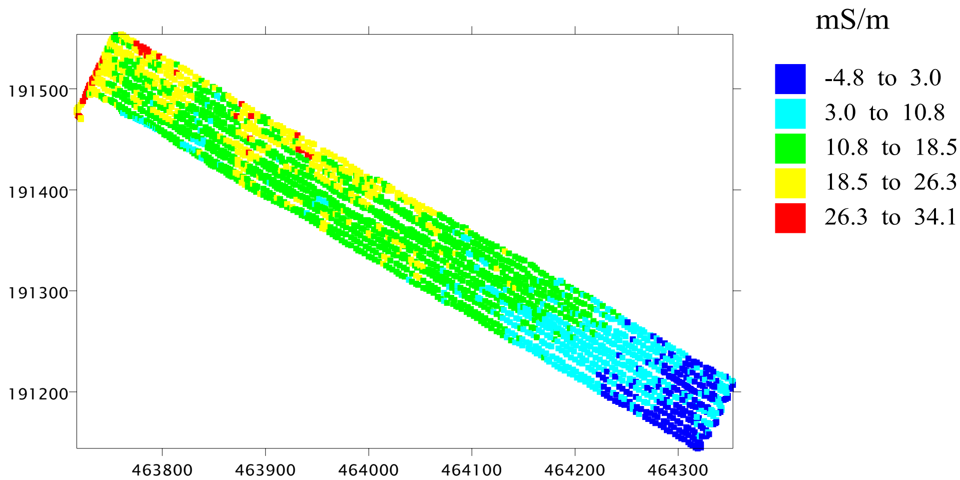


c.

Fig. 13. EMI scans of *FTB* field. **a**. = date 1 (19/09/01) vertical mode, **b** = date 1 horizontal mode, **c** = date 2 (21/03/01) vertical mode.

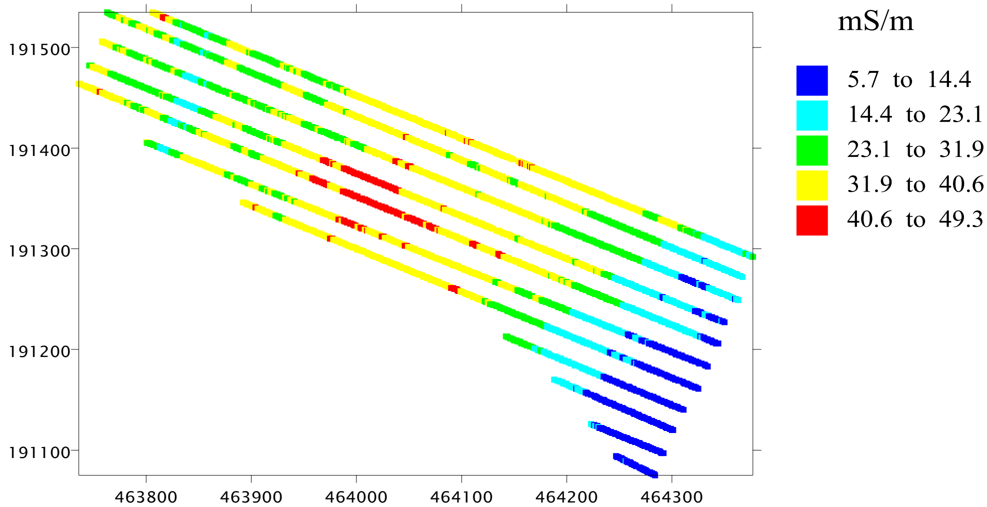


a.

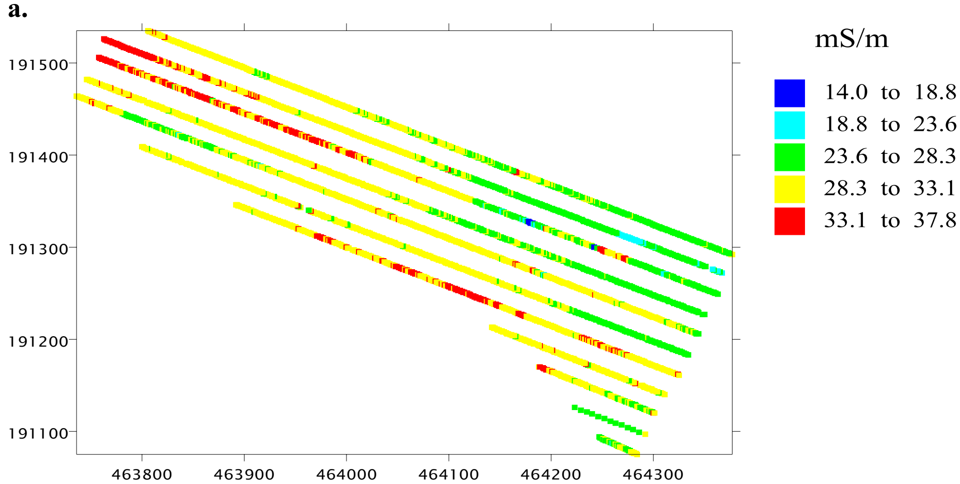


b.

Fig. 14. EMI scans of *CLY* field. **a**. = date 1 (22/08/01) vertical mode, **b** = date 1 (22/08/01) horizontal mode.



a.



b.

Fig. 15. EMI scans of *CLY* field. **a**. = date 2 (06/06/02) vertical mode, **b** = date 2 horizontal mode.

(Frogbrook, *pers. comm*.). Also, the Greensand parent material is very rich in iron, with common ironstone nodules, which may themselves influence apparent conductivity on the site (though more likely as point sources of high conductivity).

It is also interesting to note from the EC_a readings made at FTB field, that those made in the horizontal mode in the moisture deficit period are actually generally rather large and much larger than those in the vertical mode, which is counter to the situation at HLF and to what would normally be expected. This was caused by weather conditions at the time of measurement, which was persistent light rain. This had only wet the surface soil, leaving the soil at depth still dry, and so leading to conditions where moisture content would be highest in the surface soil layers that contribute a higher component of the EMI signal when used in the horizontal mode. Because the soil was very permeable it was still trafficable despite the surface moisture, which however, was not the case under field capacity conditions (see earlier comments on this page as to why horizontal mode readings could not be taken at *FTB* on this occasion).

Despite persistent problems obtaining GPS signals at the *CLY* site sufficient data were obtained to map a considerable part of the field covering the range of soil conditions present. Fig. 14 displays the EC_a distribution across a central portion of the *CLY* site after harvest (date 1 = 22/08/01) when the soil profile was dry. This part of the field corresponds to a strip running from the base of the inset rectangular section of the top left edge shown in Fig. 6 to the bottom right hand edge. On the second date the soil was also dry in the following year (date 2 = 06/06/02) (Fig. 15) and taken at a lower density from a slightly wider area.

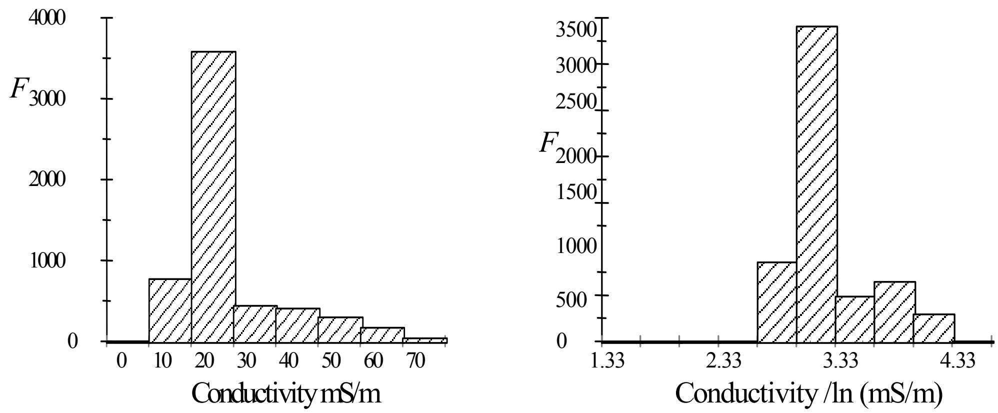
At *CLY* site the major variation in EC_a readings visible in both modes and dates during a deficit period (Figs. 14 & 15) is probably not caused by differences in soil type. The main feature is the patch of reduced EC_a readings at the right hand edge of the study area in Fig. 14 and 15, and it would be tempting to ascribe this to the Frilsham soil series mapped in Fig. 6 near this location, and not the majority Wallop series over the rest of the field. However, both of these soil types have a reasonable amount of clay in the topsoil and similar depths of topsoil, and it is the Frilsham soil that is generally deeper in total profile (Appendix 1). It is not likely therefore, that it would have a lower total apparent conductivity than the Wallop, as seen in Figs. 14 & 15. In this case it more likely that the entire strip surveyed is Wallop series and the variation in EC_a reflects the increasing depth of soil from very shallow areas on the top of the chalk ridge, down to the lower elevation part of the field. The pattern of variation in this part of the field was not indicated by changes in surface topography, and probably shows more minor variation in depth of soil to underlying chalk. This chalk is also variable in hardness, with many softer patches where the material is more like marl than rock and holds considerably more water.

EMI and Geostatistics

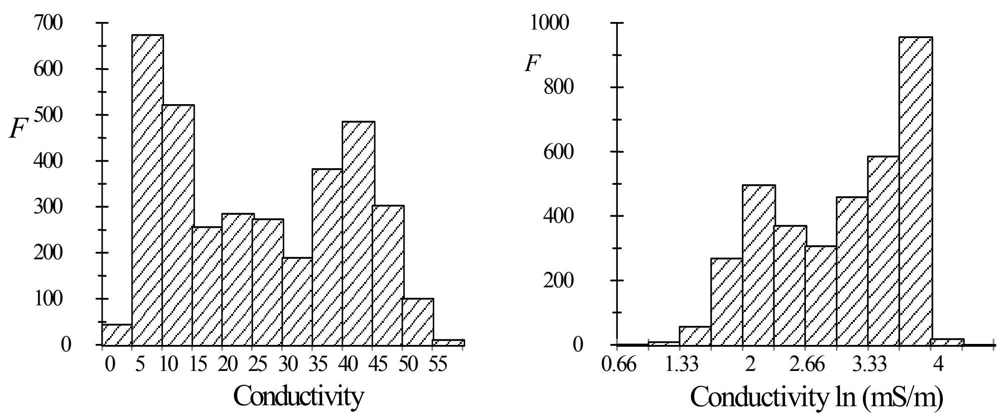
The basic frequency distribution of the EC_a readings (vertical mode) made at the two sites *SHG* and *HLF* in the first year whilst at field capacity (date 1), are shown in Fig. 16 for both raw and log transformed data. The summary statistics for the data from both sites on both occasions (vertical mode of EMI use) are given in Table 3. Log–transformation of the *HLF* (date 1) data caused some reduction in skew. The *SHG* data are bimodal and so the log-transformation makes the data negatively skewed. The date 2 data from *HLF* field were not skew so the log-transformation was not considered.

Figure 17 shows by way of example, variograms obtained with different estimators and fitted models for *HLF* (date 1) data set (untransformed). The most obvious feature of this figure is that Matheron's estimator is giving much larger values for the variogram than do the robust estimators. The Matheron and Cressie-Hawkins variograms were initially fitted with Gaussian models. This is generally to be avoided since it can lead to numerical instability. For this reason, and because the fit at the shortest lags was poor, a power function model was also fitted to the Matheron variogram for lags up to 100m. This is the model shown in the figure, kriging and determination of grid spaces with this model did not involve lag distances greater than 100m.

Tables 4 & 5 show the results of validating these different variogram models (and those fitted to the other data sets) by kriging at the validation data sites and evaluating the $\theta(\mathbf{x})$ statistic. In all cases the median value of $\theta(\mathbf{x})$ for kriging with the variogram model fitted to estimates obtained with Matheron's estimator was substantially below the expected value of 0.45 and the 95% confidence interval. This indicates that the model overestimates the variogram. In the case of the date 1 data set for *HLF* field (not log-transformed) Dowd's estimator gives a median value of $\theta(\mathbf{x})$ very close to the lower limit of the 95% interval. It is clearly the best estimator of the variogram for this data set. In the date 2 data from the site Dowd's estimator is also the best and falls within the 95% interval. Log-transformation has little effect on the results for *HLF*. This suggests that the best model of the data is a dominant normally-distributed process with some outlying values generated by an independent process.



a. *HLF* field date 1 (raw data left, log-transformed right).



b. *SHG* field date 1 (raw data left, log-transformed right).

Figure 16. Frequency (*F*) distribution histograms of the year 1 data-sets (date 1).

Table 3. Summary statistics of Year 1 data (Vertical Mode) raw data (mS/m) and transformed (ln mS/m).

	HLF (date 1)		*HLF* (date 2)	*SHG* (date 1)		*SHG* (date 2)
	mS/m	ln mS/m	mS/m	mS/m	ln mS/m	mS/m
Mean	27.7	3.25	6.25	26.0	3.04	20.7
S.D.	12.0	0.35	2.35	15.0	0.71	10.9
Skew	1.81	1.36	0.07	0.15	-0.45	-0.44

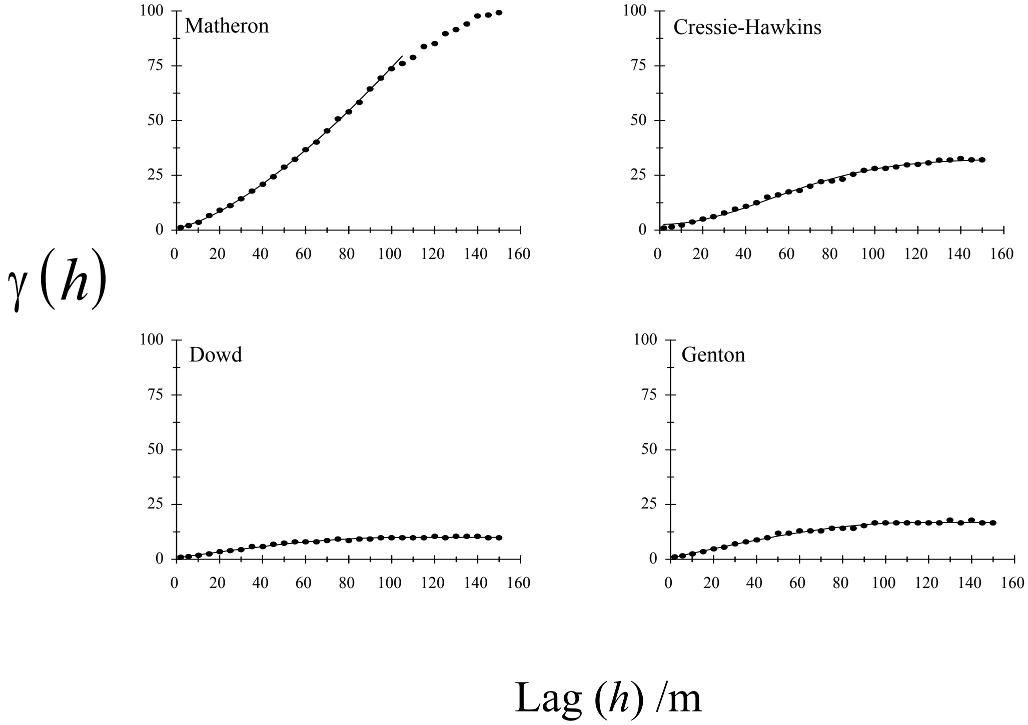


Figure 17. Variograms by different estimators and fitted models for *HLF* field date 1 (raw data)

Table 4 Validation of variogram models for *HLF* 1 and *SHG* 1 (vertical mode) date 1.

	Data set			
	HLF date 1		*SHG* date 1	
95% interval for median $\theta(\mathbf{x})$	0.42–0.49		0.40–0.50	
	mS/m	ln mS/m	mS/m	ln mS/m
Estimator	Median value of $\theta(\mathbf{x})$			
Matheron	0.23	0.35	0.15	0.16
Cressie-Hawkins	0.38	0.37	0.23	0.26
Dowd	0.40	0.39	Model not fitted	
Genton	0.30	0.40	0.22	0.27

In the case of the *SHG* data none of the variograms gave a median value of $\theta(\mathbf{x})$ close to the expectation on either date. This, and the strongly bimodal histogram, suggests that the data cannot be regarded as a realisation of a simple intrinsic random function but are best subdivided into regions with different variograms. For this reason the data were divided into two groups by a simple classification method minimising the within-group variance. Figure 18 shows the classification as a map. Class 1 contains data with larger EC_a values (mean 36 mS/m), the mean in class 2 is 10mS/m.

Table 5 Validation of variogram models for *HLF* 2 and *SHG* 2 (vertical mode) date 2.

	Data set	
	HLF date 2, mS/m	*SHG* date 2 mS/m
95% interval for median $\theta(\mathbf{x})$	0.39–0.52	0.36–0.55
Estimator	Median value of $\theta(\mathbf{x})$	
Matheron	0.28	0.15
Cressie-Hawkins	0.36	0.23
Dowd	0.40	Model not fitted
Genton	0.34	0.23

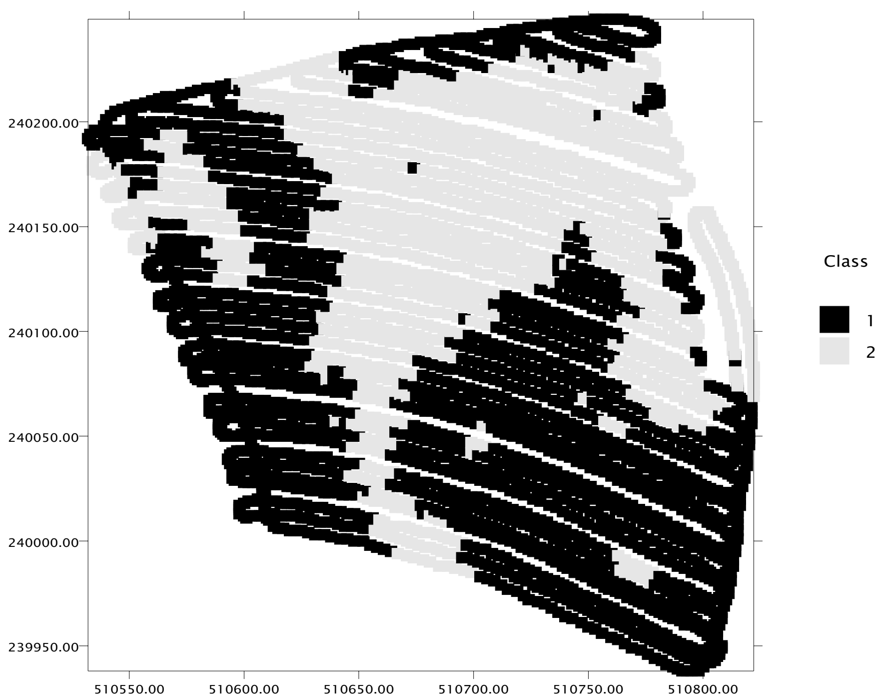


Figure 18. Two classes defined on the date 1 data from *SHG* Field.

The data within each class were then divided into a prediction and classification subset and variograms fitted and validated as above. Table 6 shows the results of the variogram validation. The date 2 set of data for the site was then subdivided according to the classification of the nearest point in the original data set. Variograms were fitted to data from each class within set C2 and compared by cross validation, see Table 7.

Again, the variogram based on Matheron's estimator appears to be an overestimate in all cases, but the robust variogram estimators mostly give acceptable results. Since the un-transformed data appear to be well-modelled by variograms obtained robustly, the log-transformation is not considered further.

Table 6 Validation of variogram models for the two classes in data set *SHG* 1 (vertical mode)

	Shagsby date 1, Class C1		Shagsby date 1, Class C2	
95% interval for median $\theta(\mathbf{x})$	0.39–0.52		0.38–0.53	
	mS/m	ln mS/m	mS/m	ln mS/m
Estimator	Median value of $\theta(\mathbf{x})$			
Matheron	0.34	0.30	0.32	0.31
Cressie-Hawkins	0.42	0.45	0.41	0.40
Dowd	0.47	0.59	0.50	0.43
Genton	0.42	0.89	0.40	0.40

Table 7 Validation of variogram models data set *SHG* 2 according to the classes in data set *SHG* 1 (vertical mode).

	SHG 2, Class C1	*SHG* 2, Class C2
95% interval for median $\theta(\mathbf{x})$	0.34–0.57	0.26–0.65
Estimator	Median value of $\theta(\mathbf{x})$	
Matheron	0.31	0.22
Cressie-Hawkins	0.39	0.31
Dowd	0.42	0.85
Genton	0.41	0.60

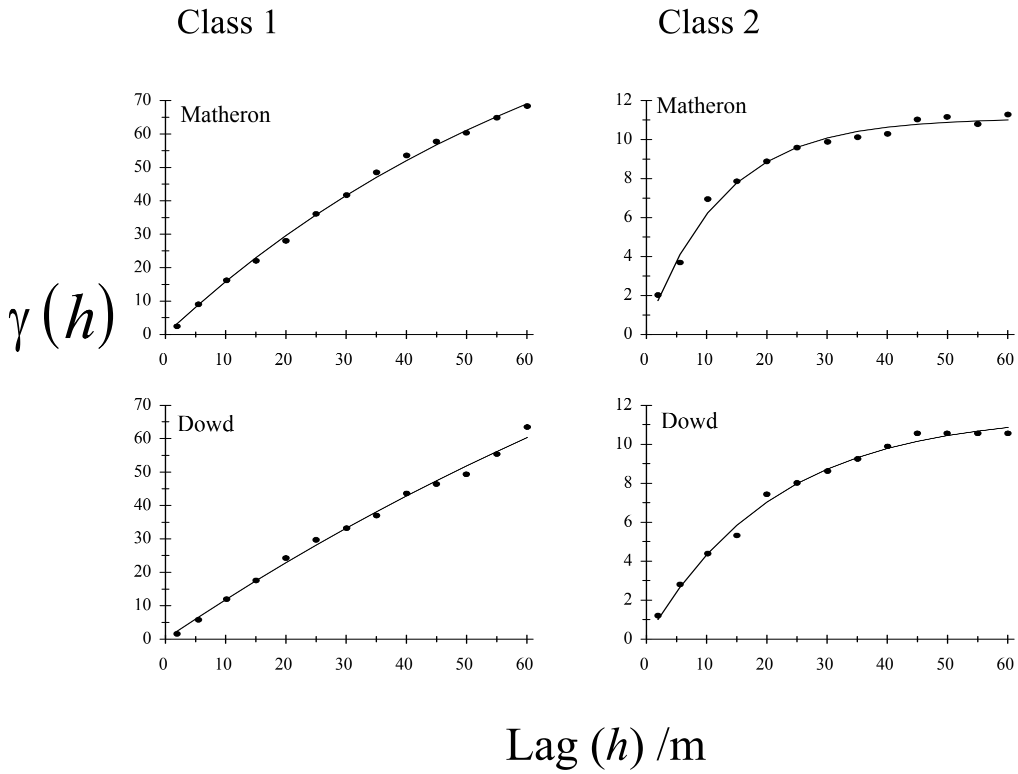


Figure 19. Variograms for (date 1) data on *SHG* field within each class.

Figure 19 shows variograms for the two classes (date 1) obtained by Matheron's and Dowd's estimators in each case. This figure shows clearly one effect of the classification. The variograms within the two classes are very different both with respect to the magnitude of the variability and also the pattern of the spatial dependence which is of shorter range in class 2. These are two distinct patterns of spatial variation that previously had been averaged together into a variogram that was representative of no one part of the field.

An important practical implication of these considerations in spatial analysis is for the optimal spacing between passes of the sensor across a field. Kriging variances were calculated for different spacings as described in the methods section. In Figure 20 are shown the point kriging variances using the variograms obtained by Matheron's estimator and the robust estimator of Dowd (1984) applied to the data sets from *HLF* field (both dates, not log-transformed). These results are also expressed in terms of the $\pm$ 1 RMS error as a percentage of the mean EC_a. Figure 21 shows the same results for data-sets C1 and C2. Here the kriging variances with Matheron's estimator applied to the whole data-set are compared with those for Dowd's robust estimator applied to the two classes separately (assuming that we krige at each point from data within the class that occurs nearest to the target site using the class variogram). The percentage error is also shown.

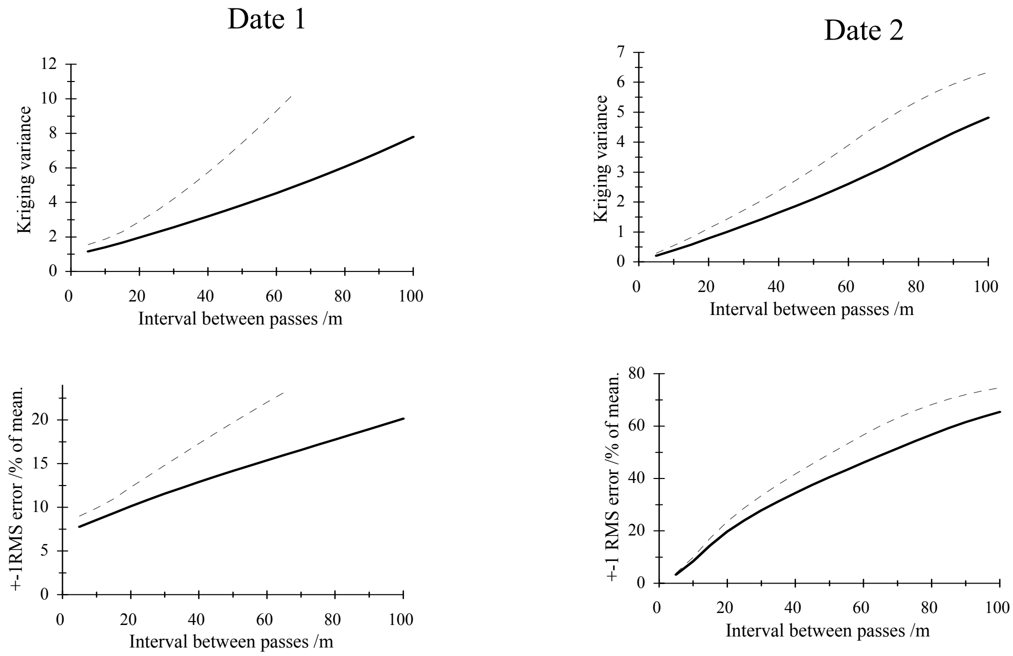


Figure 20. Kriging variances and error % for data from *HLF* field. The broken line is based on Matheron's variogram estimator, the solid line on Dowd's.

Variograms were then obtained for the remaining data sets, using cross validation to select the appropriate estimator. In Table 8 below we present the variogram models used for each data set. In Table 9 we present the spacing between passes of the EMI equipment necessary to give an estimation error (range of ± 1 standard error) which is no more than 10% of the mean of the signal across the whole field. Where this cannot be achieved with spacing wider than 5 m the spacing which limits the error to 25% is quoted.

It is noteworthy that in only two cases (out of 16) did a simple application of Matheron's estimator to EC_a data give a variogram that was supported by subsequent validation. Since the overwhelming majority of geostatistical analyses by environmental scientists only use Matheron's estimator, and only rarely attempt to validate the variogram, this is an important finding. Getting the most out of EC_a technology is likely to require the more sophisticated methods of analysis used here.

The problems with the Matheron estimator have practical consequences. In the case of the date 1 data from *HLF* field (point kriging) the Matheron variogram indicates the need for almost twice as many passes of the sensor across the field as are indicated by the robust variogram estimator, that validation showed to be a much better representation of the variability of the data. In the case of the date 1 data from *SHG* field there is an even bigger difference between the spacings indicated for a target error (25%).

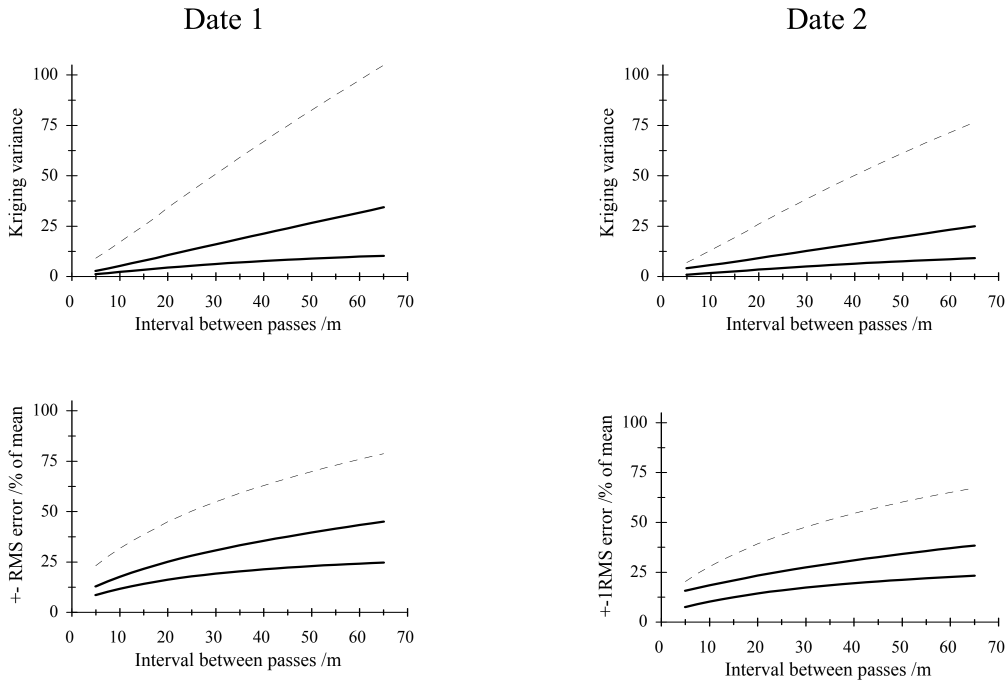


Figure 21. Kriging variances and error % for data from SHG field. The broken line is based on Matheron's variogram estimator (applied to the whole data set), the solid lines on the robust estimators within the two classes.

The temporal differences are of interest. The recommended spacing in Table 9 do not change much over time at the *SHG* site. There is a change on *HLF* field. This latter change is probably largely a result of the marked difference in the mean conductivity between the two dates at this site. There is some reduction in the kriging variances over time, but it is not very pronounced.

There is a marked difference between the indicated spacing of passes over all the sites, and also between the horizontal and vertical modes. At the Grantham site we can use a coarser grid of data than at *SHG*. This reflects differences between the two sites in the kriging variances obtained from the validated robustly estimated variograms. A practical consequence of this is that a single spacing between passes is unlikely to suit all sites, modes and conditions.

Table 8. Validated variogram for each data set (non-transformed).

Field	Date	Mode	Estimator	Model
HLF	1	V	Dowd	Spherical, 0.7 + 9.34 Sp (h|103.8)
	1	H	Dowd	Spherical, 0.0 + 9.5 Sp (h|98.4)
	2	V	Dowd	Spherical, 0.0 + 4.8 Sp (h|81.4)
	2	H	Matheron	Exponential, 1.93 + 5.89 Exp(h|1.23
SHG	1	V	Dowd	(Class 1) Exponential, 0.0 + 179.6 Exp (h|146.7)
	1	V	Dowd	(Class 2) Exponential, 0.0 + 11.54 Exp (h|21.3)
	2	V	Dowd	(Class 1) Exponential, 1.96 + 158.3 Exp (h|194)
	2	V	Dowd	(Class 2) Linear, 0.0 + 0.13 h
FTB	1	V	Dowd	Exponential, 0.0 + 190.5 Exp (h|181.8)
	1	H	C-H	Linear, 1.46 + 0.168 h
	2	V	Dowd	Spherical, 0.0 + 29.3 Sp (h|121.7)
CLY	1	V	Matheron	Linear, 3.79 + 0.173 h
	1	H	C-H	Linear, 4.12 + 0147 h
	2	V	Genton	Exponential, 0.84+295 Exp (h|752)
	3	V	Dowd	Linear, 0.0 + 0.364 h
	3	H	C-H	Power, 0.0 + 0.688 h^0.441

Table 9. Spacing between passes (metres) required to achieve target an estimation error (range of ± 1 standard error) which is no more than 10% of the mean (spacing in brackets achieves a 25% error).

Field	Date	Mode	Point Kriging	Block Kriging (10m block)
HLF	1	V	19	42
	1	H	<5 (15)	11
	2	V	<5 (16)	11
	2	H	<5 (<5)	13
SHG	1	V	<5 (20)	12
	1	V	<5 (24)	11
FTB	1	V	<5 (17)	11
	1	H	>60	>60
	2	V	<5 (24)	13
CLY	1	V	<5 (>60)	32
	1	H	<5(<5)	10
	2	V	<5(44)	17
	3	V	14	21
	3	H	7	17

<u>Seasonal stability of EMI Measurements</u>

The results of the initial multivariate geostatistical analysis of the *HLF* data are given in figures 22 and 23. Figure 22 shows the variograms and pseudo –cross- variogram with fitted models, whilst Figure 23 shows the kriged estimate of change in EC_a, and Figure 24 shows this as a standard variate (i.e. the mean change is subtracted then the result divided by the square root of the kriging variance; where this is larger or smaller than 1.96 there is evidence of a change in EC_a different from the average change).

Over most of this field the change in EC_a appears to be uniform, with only relatively small areas in which the change is significantly smaller than average (blue in Figure 24) and a smaller area still where the change is significantly larger than average. In short the pattern appears to be very stable.

There are some concerns about this approach. First, it assumes that the EC_a measurements on the two dates conform to a linear model of co-regionalisation (Webster and Oliver, 2001). This may be questionable, and the slight strain in the fit of the model (Figure 22) may indicate some non-linearity in the relationship between the EC_a values on the two dates. The second problem is a statistical one. The model of the pseudo cross-variogram has to be fitted with some assumptions when we have few comparisons between the dates over distance zero. These may inflate the cokriging variance so that we underestimate the area over which the change in EC_a differs significantly from zero.

To try and avoid these problems and provide an alternative analysis the previously mentioned cluster analysis was performed. This was done for each pair of data-sets to be compared, to find 2,3,…8 groups of vectors which form compact clusters according to a Euclidean or a Mahalanobis norm. Then the normalised classification entropy for each set of clusters were computed, and selected as the best representation of the data clustering for which there was a distinct local minimum in the NCE plot (Fig. 25). When both the Euclidean and Mahalanobis norms had a minimum in the NCE plot that clustering for which NCE was smallest was selected (Fig. 26). The results of these analyses are illustrated in Appendix 2, where we show the NCE plot in each case, and the cluster centres (*i.e.* a plot of the EC_a values of the ‘typical’ member of each cluster). We also show as a map the cluster with maximum membership at sites across the field, and show that for PSK here as an illustration (Fig. 26).

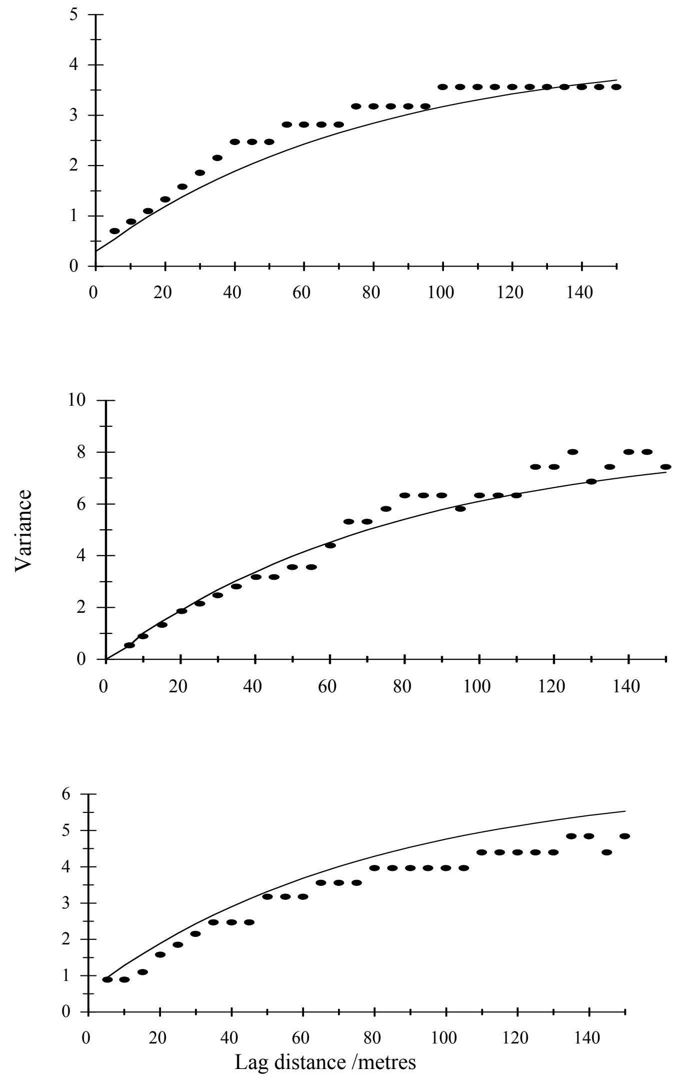


Figure 22 Variograms (top, date 1; middle, date 2) and pseudo cross-variograms (bottom) for EC_a (vertical mode) data collected from *HLF*, with fitted model.

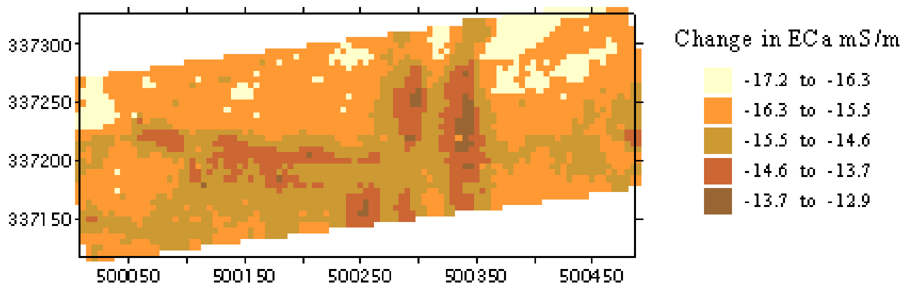


Figure 23 Cokriged estimate of change in EC_a from date 1 to date 2 for *HLF* field (units of mS/m)

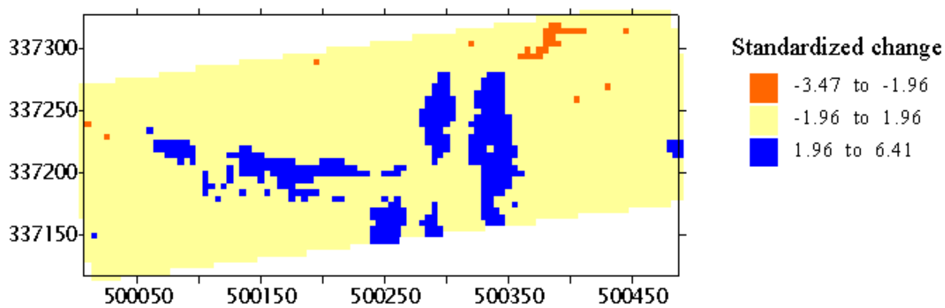


Figure 24 Cokriged estimate of change in EC_a from date 1 to date 2 for *HLF* field expressed as the standardised difference from mean change.

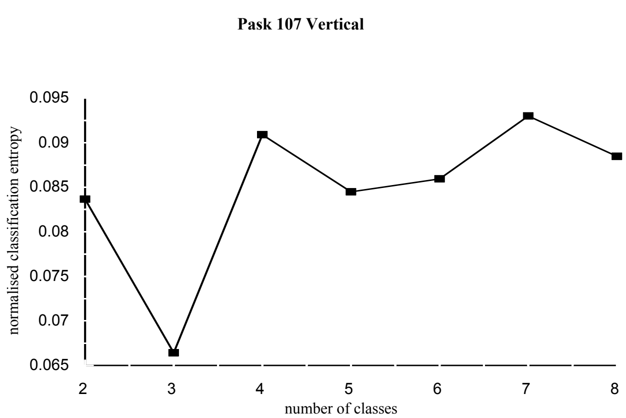


Fig. 25 Plot of the normalised classification entropy (NCE) from a cluster analysis on *HLF* data (vertical mode).

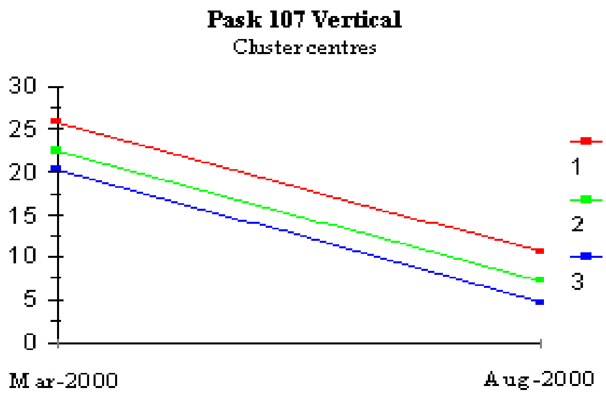


Fig. 26 Plot of the NCE cluster centres for both dates on *HLF* data (vertical mode).

Considering the results for *HLF* (vertical mode), the plot for the cluster centres (Fig. 25) shows that the relative EC_a values of the three clusters are almost identical on the two dates. This is consistent with the result of the multivariate geostatistical analysis on these data, reported above, and gives a similar looking map of the cluster centre classes (Fig. 27) as that given in Fig. 24 for the change in EC_a. On average the change in each cluster class is very similar (Fig. 26), and there is also a similar consistency in horizontal mode.

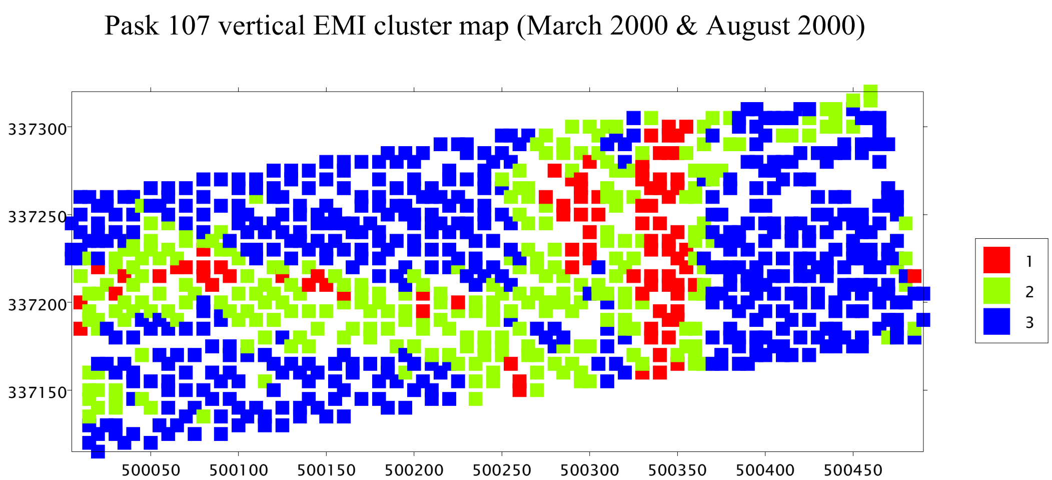


Fig. 27 Spatial plot of the NCE cluster centres for both dates on *HLF* data (vertical mode).

At some other sites the cluster centres are consistent in their *relative* values of EC_a (i.e. they come in the same order) but the difference between the clusters may be smaller on one date than on another. In one case the difference between the clusters is bigger when the absolute values are smaller (*CLY* horizontal mode). In other cases the differences were larger on the date when the absolute values were larger (*CLY*, vertical mode; *SHG*, vertical mode). The most complex pattern of change in EC_a is on *FTB* field (vertical mode), where two clusters are consistently large and small (1 and 2) but a third class changes from EC_a values more or less equal to those in the large class (when overall EC_a is large) to virtually equal to those in the small class (when overall EC_a is small)) (Appendix 2).

Relating EMI to Crop and Soil Measurements

Data manipulation

Particle size (sand, silt, clay %), bulk density and organic carbon have been measured as topsoil and subsoil variables. Available water capacity however, has been derived from these other variables using pedotransfer functions, and so cannot be used in the analyses reported here since they are not in any sense random. We wished to consider all soil properties at both depths in both the analysis of partial effects of each variable and in the principal components analysis. This led to some difficulties because, while more than seventy locations had been sampled on each field, analyses on all variables were only available for a much smaller subset of the sites (Table 10 below).

Table 10. The total number of samples collected from each site and the number on which all variables were measured at all depths.

No.	Field	Total number of samples from each sites	Total number of samples with all variables measured at all depths
1	*HLF*	74	15
2	*SHG*	70	34
3	*FTB*	87	4
4	*CLY*	121	86

For this reason we considered combining the data from all fields into a single analysis. This was done, and a principal component analysis was conducted with the combined soil data set. Figure 28 below shows the data projected onto the first principal components. The numbers refer to the fields (as in the first column of Table 10 above).

It is clear that the *CLY* field (4 in Table 10 and Fig. 28) is very different from the other three. For this reason we decided to combine *HLF*, *SHG* and *FTB* fields into a single data set (56 points) to be analysed together, and to analyse the data from *CLY* separately.

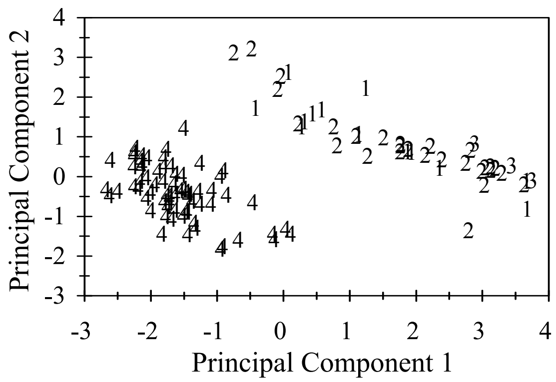


Figure 28 Projection of soil data from all fields onto first two principal components.

EC_a values at each soil sample site were obtained by ordinary point kriging from observations within 25 metres, using the variograms reported previously. In some cases the EC_a data were rather sparse so could not be kriged at the sites. Because we have combined the soil data for three fields, it was necessary to find comparable measurements of EC_a. It was possible to identify vertical mode measurements taken in Spring (February/April) or late summer (August/September) for each of the three fields in the combined data set. However it was not possible to provide a comparable data set in horizontal mode, since, no horizontal mode data were collected on *SHG* field, and on the other fields horizontal mode data collected at comparable times were sparse and could only be obtained by kriging at relatively few locations.

The 'Diviner' water content data, measured on dates close to a set of EC_a measurements, were analysed separately for each field. A regression was computed of EC_a (horizontal and vertical modes) on the water content in the first 30cm and the water content from 30cm to 100cm and any available static soil properties. The same regression was then computed after dropping the water content data, and the two models were compared to test the partial effect of water content.

Data for spectral reflectance (visible red) of the bare soil surface were available from *SHG* field. These were kriged to the soil sampling points, and the same analyses conducted as for EC_a data.

<u>Analyses</u>

<u>Principal Components Analysis</u>

<u>Combined data-set *HLF, SHG & FTB*</u>

Figure 29 shows the cumulative percentage of the variation in the original soil data set accounted for by the principal components, whilst Table 11 shows the latent vectors for each principal component - *i.e.* the loading given to each soil variable to determine the value of each component. Note from Fig. 29 that nearly 90% of the variation is accounted for by the first four principal components.

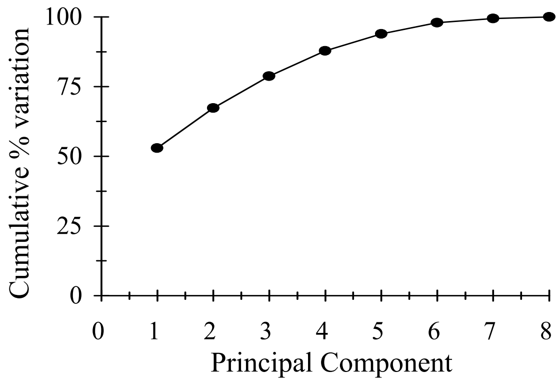


Figure 29. Cumulative percentage of variation accounted for by principal components of soil data.

Partial effects by multiple regression analysis

Table 12 below shows the Akaike information criteria (AIC) for the full regression model (EC_a regressed on all soil properties) and then for models with each soil property alone dropped in turn.

In the regression of spring EC_a (vertical) on static soil properties (Table 12) there are significant partial effects of bulk density (both depths), organic carbon and clay content (subsoil) and sand content (topsoil). The largest effect on the model arises from removing the topsoil sand content as a contributor, and the next largest from removing bulk density in the subsoil. Note again that, if the partial effect of a variable is not significant, this does not mean that it has no physical effect. Only topsoil sand content has a significant partial effect on the late summer measurements of EC_a. This may reflect the smaller overall variability of the signal. As the soil becomes drier overall so differences in conductivity between areas with contrasting moisture characteristics may become less marked.

Table 11 Latent vectors for each principal component. (BD=Bulk density, C=Carbon)

Soil layer	Soil variable	Principle Component							
		PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**
Top	Clay	-0.438	-0.176	0.107	0.082	0.197	-0.501	-0.468	0.5
	Sand	0.444	0.027	-0.129	-0.081	-0.3	0.419	-0.533	0.479
	Organic C	0.081	-0.837	-0.282	0.05	0.38	0.251	0.045	-0.027
	BD	-0.314	0.233	-0.059	-0.733	0.43	0.328	-0.117	-0.009
Sub	Clay	-0.425	-0.238	-0.014	-0.055	-0.508	0.128	-0.447	-0.534
	Sand	0.432	0.159	0.088	0.132	0.465	-0.2	-0.525	-0.481
	Organic C	0.338	-0.212	-0.233	-0.614	-0.258	-0.581	0.062	-0.052
	BD	-0.158	0.293	-0.909	0.224	0.038	-0.091	-0.045	-0.025

Table 12 Partial effects of soil properties in regressions of EC_a data '***a***,' the variable portion of the Akaike information criterion is given for each model. The highlighted (bold text) values of ***a*** are larger than that for the full model, indicating a significant partial effect.

Soil Layer	Variable modelled	EC_a (vertical mode, Spring).	EC_a (vertical mode, Autumn).
	Full model (Wald statistic and *p* value)	117.7, p<0.001	28.93, p<0.001
	a	236.92	151.45
Top	**Clay**	235.61	149.46
	Sand	**250.87**	**154.16**
	Organic C	235.95	150.86
	Bulk density	**239.08**	150.60
Sub	**Clay**	**238.25**	149.66
	Sand	236.04	149.49
	Organic C	**238.08**	149.63
	Bulk density	**245.77**	150.17

Regression on principal components.

Table 13 and 14 below show the results of regression of EC_a on the principal components of the soil variables.

Table 13 Regression model of ECa (vertical , Spring) on principal components of soil properties

Error structure	
Model	Spherical
Spatial dependence ratio	1
Distance parameter /metres	39

Overall Inference

AIC	236.92
Wald	117.73
p	<0.001

PC	Coefficient	variance	t ratio	*p*
1	-5.221	0.375	-8.53	<0.001
2	-4.054	1.104	-3.86	<0.001
3	-1.929	1.399	-1.63	>0.05
4	-0.553	1.827	-0.41	>0.05
5	-1.654	2.541	-1.04	>0.05
6	8.653	3.717	4.49	<0.001
7	-5.961	10.733	-1.82	>0.05
8	-15.3	32.027	-2.7	0.01

Of the principal components of the static soil properties PC1 has the largest effect on EC_a (Spring) - Table 13. The next largest is PC6 then PC2 then PC8. The effects of the other principal components are not significant. To aid interpretation we plot the elements of the latent vectors below. Figure 30 shows the elements of the vectors for PC1 and PC6 (with the largest effects on EC_a). Each point corresponds to a soil property. The larger the value associated with a property (positive or negative) the larger its contribution to the principal component. Thus a soil with large clay content in the topsoil and subsoil and small sand content, large bulk density in the topsoil and small organic carbon content in the subsoil will have a large and negative value of PC1. Since PC1 has a negative regression coefficient in Table 13 we see that such soils are also expected to have a large EC_a (vertical mode, spring). The large circular symbol in Figure 30 shows where in the projection the soils are expected to have a large EC_a, the small symbol in the opposite corner indicates where EC_a will be smallest.

Table 14 Regression model of EC_a (vertical, Autumn) on principal components of soil properties

Error structure

Model	Spherical
Spatial dependence ratio	1
Distance parameter /metres	41

Overall Inference

AIC	151.45
Wald	28.927
p	<0.001

PC	Coefficient	variance	t ratio	*p*
1	-2.686	0.549	-3.63	0.001
2	-1.269	1.284	-1.12	>0.05
3	0.481	1.163	0.45	>0.05
4	1.630	1.458	1.35	>0.05
5	-3.999	2.204	-2.69	0.012
6	5.983	4.026	2.98	0.006
7	-6.835	19.650	-1.54	>0.050
8	-4.957	40.508	-0.78	>0.050

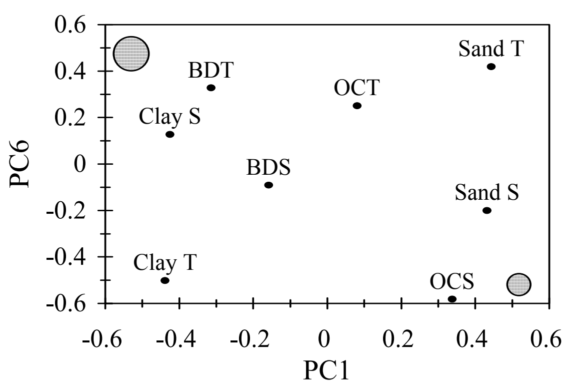


Figure 30 Elements of the latent vectors for PC1 and PC6 of static soil properties in the combined data set (T=Topsoil, S=Subsoil)

Thus, particularly large EC_a are expected from soils with large clay content, particularly in the subsoil, large bulk density in the topsoil, small sand content, particularly in the subsoil and small organic carbon content in the subsoil. Figure 31 similarly presents the latent vectors for PC2 and PC8. This indicates that larger organic carbon content in the topsoil will be associated, other factors being equal, with larger EC_a, as will low sand content in the topsoil.

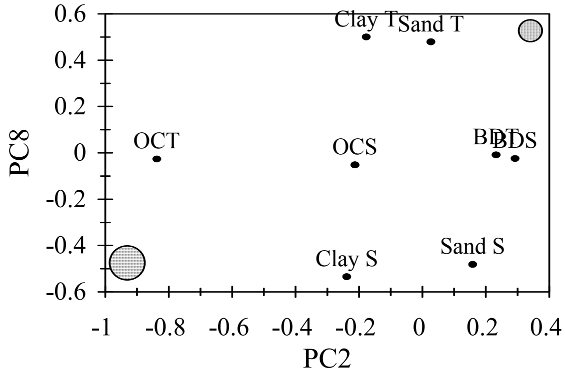


Figure 31 Elements of the latent vectors for PC2 and PC8 of static soil properties in the combined data set.

Principal components 1 and 6 are also the two most important in the regression of the late summer EC_a measurements. PC5 is the only other significant one. The latent vectors, in Table 11, show that the dominant variables in this component are the clay and sand content of the subsoil.

<u>*CLY* data-set</u>

Note that the organic carbon data used here(at site CLY) were obtained at the University of Reading by loss on ignition. Note also that the particle size data used here were determined by hand texturing in the case of the overall regression models. The particle size data used later in the analysis of the Diviner data were determined in the laboratory. These data sets are separated because the laboratory determination was made after removal of calcium carbonate.

The EC_a data used in the overall regression models (not with the Diviner data) are limited to the vertical and horizontal mode data collected on the third and final occasion (6/6/02) since on previous dates coverage of the whole field was limited.

Figure 32 shows the cumulative percentage of the variation in the original soil data set accounted for by the principal components. Note that over 75% of the variation is accounted for by the first four principal components. Table 15 shows the latent vectors for each principal component - i.e. the loading given to each soil variable to determine the value of each component.

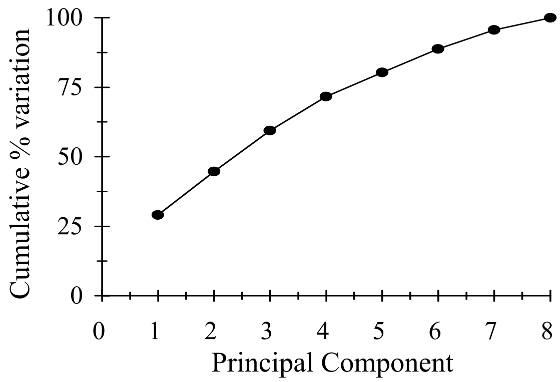


Figure 32. Cumulative percentage of variation accounted for by principal components of soil data from *CLY*.

Table 15 Latent vectors for each principal component from *CLY* analysis.

Soil layer	**Soil variable**	**Principle Component**							
		PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**
Top	Clay	-0.164	0.07	-0.338	-0.879	-0.182	0.014	-0.074	0.208
	Sand	0.404	0.217	-0.348	0.219	-0.369	-0.561	0.16	0.38
	Organic C	-0.328	0.549	0.068	0.255	-0.333	0.5	-0.066	0.394
	Bulk density	0.436	-0.093	-0.321	-0.013	0.333	0.546	0.488	0.226
Sub	Clay	-0.292	-0.388	-0.502	0.22	-0.508	0.172	0.19	-0.374
	Sand	0.418	0.459	0.21	-0.211	-0.327	0.131	0.202	-0.601
	Organic C	-0.163	0.499	-0.577	0.133	0.471	-0.1	-0.195	-0.324
	Bulk density	0.475	-0.167	-0.17	0.064	-0.148	0.283	-0.782	0

Partial effects by multiple regression analysis

Table 16 below shows AIC for the full regression model (EC_a regressed on all soil properties) then for models with each soil property alone dropped in turn for the *CLY* data-set.

Regression on principal components.

Table 17 below show the results of regression of EC_a on the principal components of the soil variables. Note that the regression of the vertical mode EC_a on the principal components was not significant, the results in the table are for the horizontal mode only.

There is only a significant regression of the EC_a in horizontal mode on the soil properties at this site. This could be because in vertical mode, when the signal from deeper parts of the profile dominate, variation in the depth of soil over the chalk has a dominating effect. We do not have a direct measurement of this variable.

In the regression of EC_a (horizontal) on the static soil properties all have a significant partial effect. The strongest effect is of bulk density in the subsoil, in fact when this variable is dropped the overall regression model is no longer significant.

Table 16 Partial effects of soil properties in regressions of EC_a data - ***a***, the variable portion of the Akaike information criterion is given for each model. The highlighted values of ***a*** are larger than that for the full model, indicating a significant partial effect.

Soil Layer	**Variable modelled**	**EC_a (vertical mode, Spring).**	**EC_a (vertical mode, Autumn).**
	Full model (Wald statistic and *p* value)	Not Significant	26.53, $p<0.001$
	a		17.68
Top	**Clay**		**20.14**
	Sand		**19.06**
	Organic C		**17.84**
	Bulk density		**18.04**
Sub	**Clay**		**17.76**
	Sand		**20.85**
	Organic C		**20.56**
	Bulk density		**25.64**

Principal component PC2 of the static soil properties has the strongest effect on the EC_a (horizontal) signal, followed by PC7 and PC8. Plots of these components are shown in Figure 33 below. Note that sand content in the subsoil appears to be positively associated with EC_a (horizontal) here as is organic carbon content at both depths. Bulk density in the subsoil is strongly and negatively associated with EC_a.

Table 17 Regression model of EC_a (horizontal) on principal components of soil properties for *CLY* field.

Error structure

<u>**Model**</u>	Spherical
Spatial dependence ratio	1
Distance parameter /metres	135

<u>Overall Inference</u>

AIC	17.68
Wald	26.53
p	<0.001

PC	Coefficient	variance	t ratio	*p*
1	-0.181	0.013	-1.59	>0.05
2	0.45	0.024	2.9	0.006
3	0.172	0.017	1.32	>0.05
4	0.261	0.027	1.59	>0.05
5	0.123	0.035	0.66	>0.05
6	0.201	0.037	1.04	>0.05
7	0.413	0.031	2.35	0.024
8	-0.414	0.033	-2.28	0.028

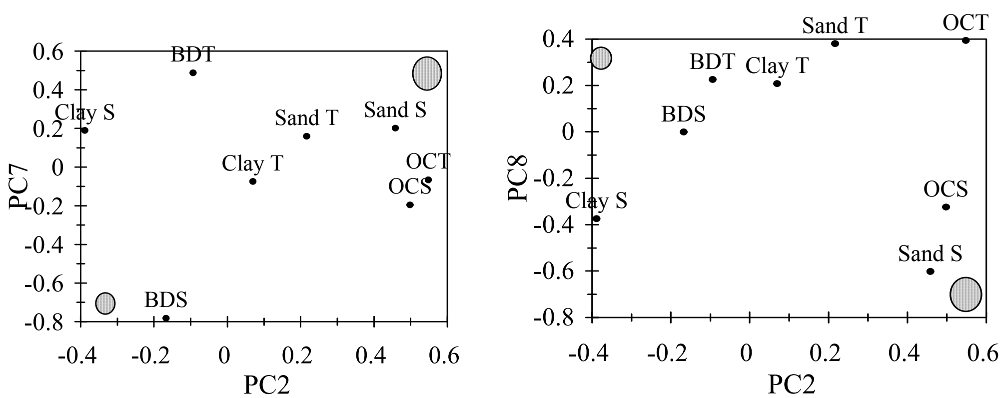


Figure 33   Elements of the latent vectors for PC2 and PC7 (left) and PC2 and PC8 (right) of static soil properties in the combined data set.

<u>Soil hydrology data at "Diviner" tube locations on each site.</u>

Table 18 below shows the results of regression analysis on sets of static soil properties and Diviner measurements of water content, and the effect of dropping the water content measurements from the model. Conclusions here must be tentative because of the relatively small data-sets. Out of the eight data-sets there was only a significant partial effect of the Diviner measurements in three instances (*SHG* field spring measurements, and *FTB* field, vertical-mode measurements in spring and late summer).

Relating Spectral Reflectance of bare soil at *SHG* field to soil data

A separate principal component analysis was conducted on the soil data for *SHG* field for this purpose, and the latent vectors are shown in Table 19 below.

Table 20 shows the partial effects of each soil property in the regression of visible red reflectance (VR) on the measured soil variables, and Table 21 shows the regression of VR reflectance on principal components of the soil properties.

In the regression of visible red reflectance on static soil properties only topsoil clay content has a significant partial effect (Table 20), and when this variable is dropped the overall regression model is not significant. Only PC6 of the principal components of the static soil properties on *SHG* field has a significant contribution in the regression of VR reflectance. Examining the latent vector for PC6 in Table 19 shows that this will have large (negative) values for soils with large clay content in the topsoil and also relatively large sand content in the subsoil. Organic carbon content (which is often important in determining VR reflectance of soil material) has small weightings in this PC

At the time of measurement it was observed that there was a good deal of variation in the structure of the soil surface, which had recently been worked to a seedbed. This may explain why particle size distribution appears to be important, because this will contribute to the structure, and since variation in structure of the soil surface will have a marked effect on its overall bi-directional reflectance properties.

Summary

There is evidence for a complex of separate and correlated effects of soil properties on signals measured by sensors. Soil texture is an important factor, but bulk density and the carbon content may also be important. With optical sensors soil structure may be important too. These results caution against any oversimplification as to which factors determine measured sensor signals, and suggest that a more general conceptual model of how a particular soil profile is likely to respond to the EC_a instrument is probably necessary.

Table 18 Results of regression of EC_a data on soil properties and water content at Diviner tube transects. H= horizontal mode; V = vertical mode of EMI use.

Site	**Mode**	**Date (EC_a)**	**Date (Diviner)**	**Number of static predictors**	**RMS (f)**	***p***	**RMS(d)**	**AICc* (full model)**	**AICc (Diviner data dropped)**
HLF	H	17/8/00	30/6/00	4[a]	**				
	V	17/8/00	30/6/00	4	**				
SHG	V	11/2/00	21/2/00	8[b]	1.12	0.05	53.73	1.36+	47.81
	V	23/8/00	15/8/00	6[c]	30.37	0.03	26.84	120.2	94.35
FTB	V	19/9/01	20/12/01	5[d]	68.08	0.007	133.9	146.36	152.88
	H	19/9/01	20/12/01	5[d]	7.57	0.01	9.09	93.61	88.16
	V	21/03/01	23/3/02	5[d]	30.31	0.006	55.11	131.48	137.17
CLY	H	6/6/02	26/6/02	8[b]	1.25	0.024	3.96	128.35	89.66
	V	6/6/02	26/6/02	8[b]	3.43	0.006	15.6	143.49	110.21

* Bedrick, E.J. and Tsai C-L. 1994. Biometrics 50, 226-231.
**Neither the full model nor the reduced were significant
+In this case the conventional AIC is used since there are too few data

a Topsoil Clay, Sand, OC, BD
b Topsoil and subsoil Clay, Sand, OC, BD
c Topsoil Clay, Sand, OC, BD and subsoil Clay, Sand.
d Topsoil and subsoil Clay, Sand, topsoil BD

Table 19 Latent vectors for each principal component of soil data on *SHG* field.

Soil layer	**Soil variable**	**Principle Component**							
		PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**
Top	Clay	-0.421	-0.115	0.097	0.099	-0.347	-0.59	-0.425	0.375
	Sand	0.434	-0.013	0.048	-0.221	0.045	0.373	-0.696	0.367
	Organic C	0.139	-0.714	0.514	0.434	0.068	0.106	0.04	-0.035
	Bulk density	-0.324	0.26	-0.248	0.74	0.09	0.386	-0.249	-0.005
Sub	Clay	-0.419	-0.161	0.122	-0.326	-0.219	0.269	-0.361	-0.653
	Sand	0.423	0.094	-0.112	0.231	0.25	-0.522	-0.362	-0.53
	Organic C	0.362	0.319	0.239	0.214	-0.796	0.076	0.092	-0.125
	Bulk density	-0.157	0.522	0.761	-0.032	0.347	-0.038	-0.034	0.017

Table 20 Partial effects of soil properties in regressions of VR data - ***a***, the variable portion of the Akaike information criterion (AIC) is given for each model. The highlighted (bold text) values of ***a*** are larger than that for the full model, indicating a significant partial effect.

Soil Layer	**Variable modelled**	**Visible red reflectance.**
	Full model (Wald statistic and *p* value)	20.78, $p = 0.008$
	a	-261.28
Top	**Clay**	**-251.86**
	Sand	-261.71
	Organic C	-263.10
	Bulk density	-261.70
Sub	**Clay**	-262.07
	Sand	-263.17
	Organic C	-262.31
	Bulk density	-262.59

Table 21 Regression model of VR reflectance on principal components of soil properties

Error structure

<u>**Model**</u>	Spherical
Spatial dependence ratio	1
Distance parameter /metres	41
<u>Overall Inference</u>	
AIC	151.45
Wald	28.927
p	<0.001

PC	Coefficient	variance	t ratio	*p*
1	0.0027	4.9E-06	1.22	>0.05
2	0.0013	5.3E-06	0.56	>0.05
3	-0.0022	8.9E-06	-0.74	>0.05
4	-0.0015	1.11E-05	-0.45	>0.05
5	-0.003	1.86E-05	-0.7	>0.05
6	-0.019	3.02E-05	-3.46	0.002
7	-0.0186	9.96E-05	-1.86	>0.05
8	0.0222	0.000219	1.5	>0.05

<u>GROUND PENETRATING RADAR (GPR)</u>

The profiles indicated below (Figs. 34 – 36) are pseudo depth sections with a two way travel time (in nanoseconds) on the left hand Y axis and a distanced profile in metres on the X axis. Additionally, using an assumed GPR velocity of 0.1 m/nanosecond, an interpreted depth axis has been superimposed on the right hand Y axis, and is given in metres below the ground surface.

<u>Reflectance Measurements of Sub-surface Features</u>

<u>*HLF* data year 1</u>

The *HLF* data acquired in year 1 (processed by Earth Science Systems) allows a fairly deep "view" into the soil profile to be made, due to the sandy nature of the upper soil material with a low electrical conductivity. Radar reflections have been measured down to around 4m below ground level, and the traces for transects 1 & 2 (Lines 1 & 2) are given in Fig 34 below, whilst those for transects 3 & 4 (Lines 3 & 4) are given in Fig. 35.

There are zones of differential penetration *e.g.* below 16m along transect 1 (Line 1 Tramline 6; in Fig. 34) where the reduced depth may be due to more clay-rich soil. There appears to be some dipping reflectors in the deeper parts of some profiles. These have been highlighted on the sections in red (Fig. 34; Line 1) and are interpreted as possibly bedrock strata beneath the soil layer. Within the soil layer, there are numerous reflection events indicative of more prominent soil layers (flat or slightly dipping boundaries highlighted in blue) (Fig 34 & 35; Lines 1, 3 & 4) and point source reflectors (green diffraction hyperbolae) (Fig 34 & 35; all Lines) caused by underground services / drains, buried rocks or voids.

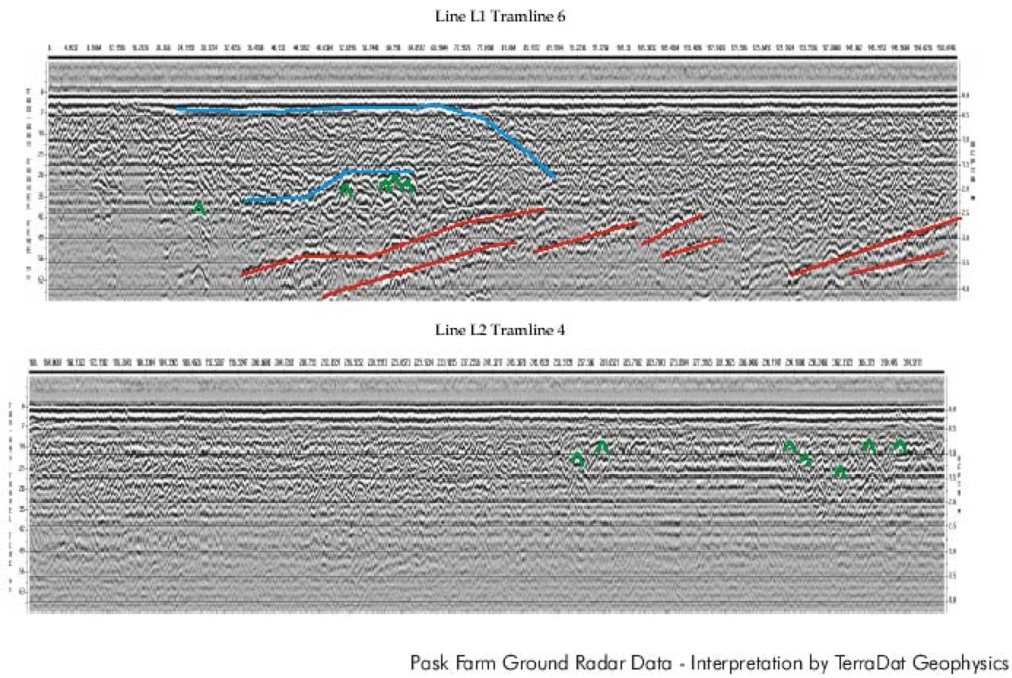


Fig. 34 GPR reflection with depth traces for transects 1 and 2 at *HLF* in year 1. For explanation of coloured lines see text. Total depth shown is 4.0 m.

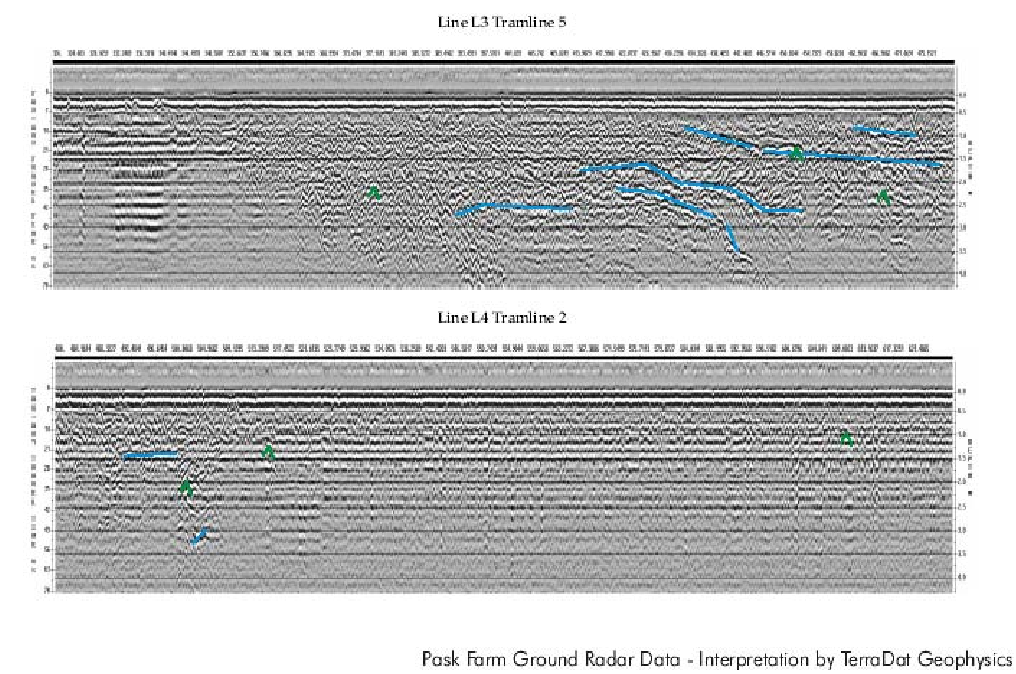
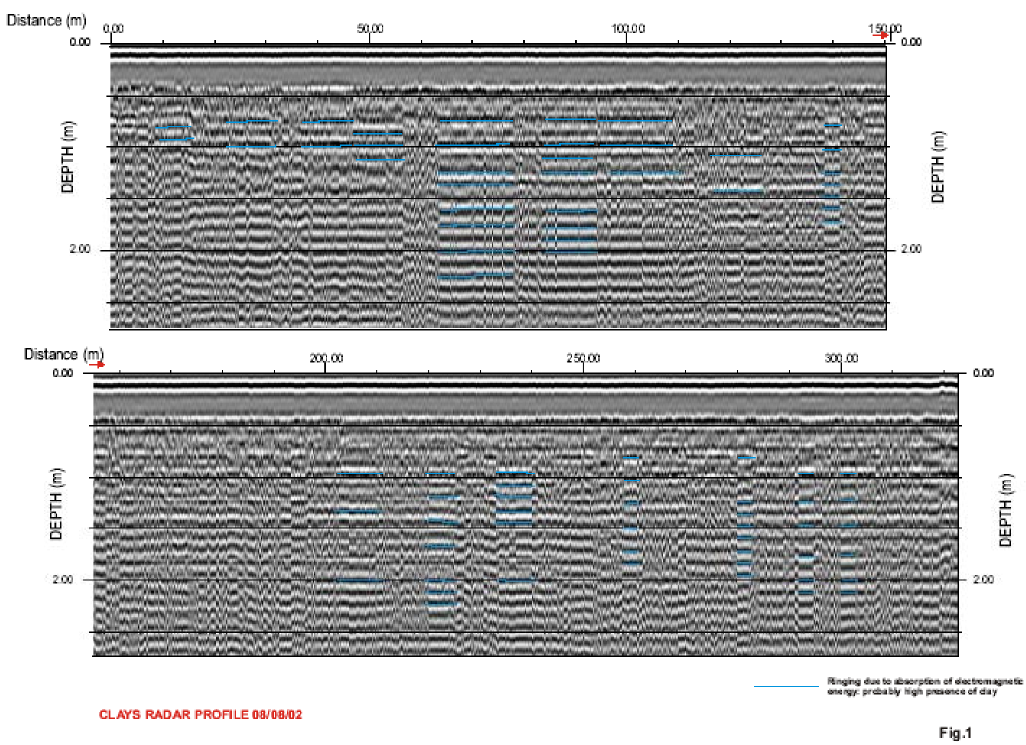


Fig. 35 GPR reflection with depth traces for transects 1 and 2 at *HLF* in year 1. For explanation of coloured lines see text. Total depth shown is 4.0 m.

<u>*CLY* data year 2.</u>

The 320m long profile presented in Figure 36 shows the data acquired at CLY in August 2002. The section represents distance along the surface on the horizontal axis and depth below ground level on the vertical axis. Note that although the scale shows depths down to >2 metres, no useful reflections were obtained.

The black and white traces on the sections represent the amplitude or strength of reflected radar energy. If a subsurface horizon such as a soil layer or geological layer was within the depth range of exploration, a distinct reflector would be observed (as in Figs. 34 & 35 above). Unfortunately, due to limitations described above (clay-rich soil), the depth of exploration was severely restricted to a few centimetres and no useful data could be obtained. The blue lines indicated on Fig. 36 are patches where "ringing" due to the absorption of electromagnetic energy has occurred in the soil, probably caused by a very large clay content at these locations.



<u>Fig. 36.</u> GPR reflection with depth traces for transects 1 – 4 in one continuous line at *CLY* in year 2. For explanation of coloured lines see text. Total depth shown is 2.5 m

Novel Measurements with GPR

The “air launched“ survey carried out over the small area centred over transect 1, and preliminary processing showed that the GPR signal amplitude response could be classified following low pass filtering, spherical exponential control (SEC) gain, and ‘de-wow’ processing. Figure 37 shows a colour coded plan map of the site – time sliced to highlight a depth section between 14-35 nanoseconds two-way travel time (approximates to 1 – 2 m soil depth). In Figure 37 the colour coding ranges from blue where a low amplitude response was received, to red where there was a high amplitude response. In the terrain under consideration, which corresponds roughly to drier sandy material over shallow (25 – 50 cm depth) limestone with pockets of deeper (2 m or more) moister loamy sand between ridges; the blue areas would signify the shallower limestone ridges, and the red the high amplitude response of the deeper more variable loamy sand soil material.

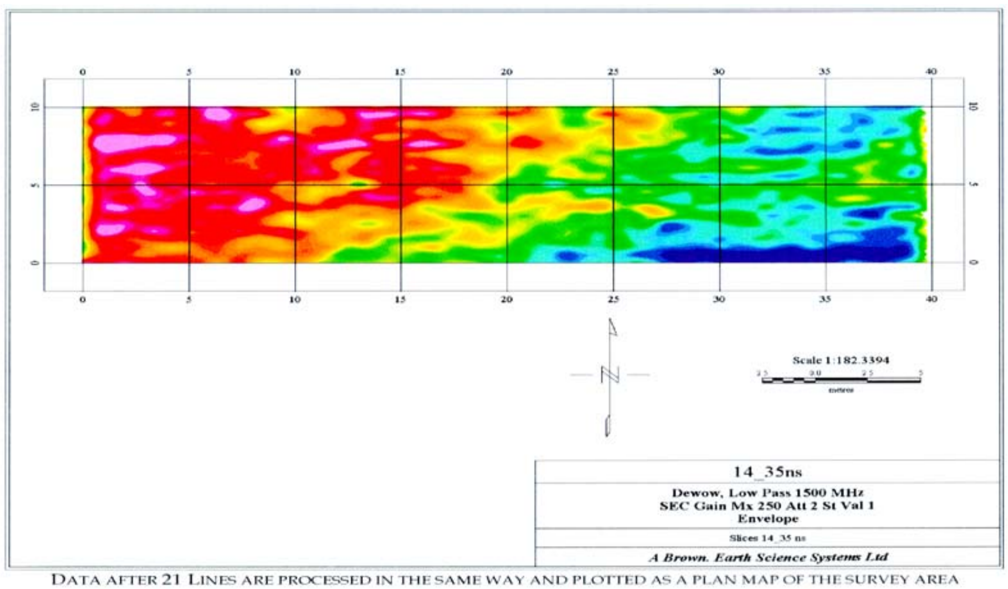


Fig. 37. Colour coded plan of air launched GPR signal of patch over transect 1 at *HLF* field in year 1. Blue = low amplitude response, Red = high amplitude response.

SPECTRAL REFLECTANCE

The spectral reflectance measurements made during the autumn of 2002, can also be mapped across the site. This has been done for the ‘visible red’ wavelength readings, at 660 nm in Fig. 38, which shows clearly the wider spacing of these measurements made down the 24 m tramlines, compared with the 6m spaced EMI measurements in Fig. 12a. Visible red has been mapped as this tends to be more indicative of changes in soil colour and type than the ratio of red/infra-red that is used in vegetation indices (and is more influenced by moisture and surface structure).

An appraisal of Fig. 38 does not really suggest a repeat of the pattern seen in maps of EC_a for this site (Fig 12) or the mapped soils (Fig. 3). The two 'hotspots' seen about mid-field to the left and on the right hand edge, do not really correspond with anything except that they are in the smaller value EC_a region.

It should be remembered that this type of sensing of the surface of the soil is susceptible to the effects of factors such as surface moisture, aspect, roughness of tilth and other factors which have a more indirect relationship with soil type. The conclusion is that the mapping potential of this type of sensing should be limited to surface properties (notably organic matter, see Lark 2000b for example where topsoil organic carbon was significantly related to visible red reflectance). Furthermore, it would be necessary to define a strict protocol for defining appropriate surface conditions before data are collected. It is highly likely that most information from spectral reflectance will come from measurements on a crop where soil induced variation in moisture stress can be detected by the normalised difference vegetation index.

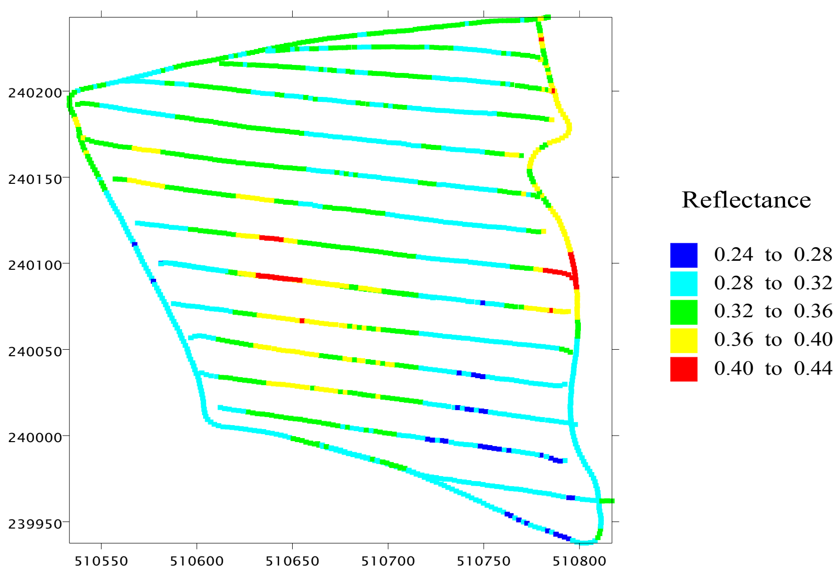


Fig. 38 Visible red (660 nm) reflectance measurements made from a vehicle-borne sensor made at site *SHG* in October 2002.

CONCLUSIONS

ELECTRO-MAGNETIC INDUCTION SENSOR

Soil properties measured; accuracy and temporal stability.

From previous scientific knowledge, soil EC_a in British soils is likely to be governed by two main variables: particle size distribution (texture) and soil moisture. The interpretation of EC_a maps requires some, though not necessarily extensive, knowledge of the soil materials and underlying geology. However, our findings demonstrate that spatial EC_a measurements, and a knowledge of the soils in the area, can be useful for identifying zones within fields that may warrant differential management. They may also be used to guide a more focused soil survey to delineate the boundaries between soil types and to allow more meaningful sampling of soils for soil nutrient analysis or other purposes.

Differentiating variations in texture

The evidence from the *HLF* data-sets showed clearly the potential of EMI techniques to distinguish between soil types based on clay or sandy loam textures . At *SHG* the main soil types were all clays or clay loams, and the main determinant between them was the amount of clay in the upper subsoil and/or the interaction between texture and soil hydrology. In this case, EC_a values distinguished the heavier less permeable soils (Evesham and Oxpastures) from more permeable and freely draining soil (Waterstock and Cottenham), during both wet and dry times of the year. The choice of operating mode proved important in some cases, in that inclusions of clay in a predominantly sandy subsoil at *FTB*, could be detected using the vertical mode, but were barely identified using the horizontal mode.

Where EC_a variation across a field was chiefly caused by soil type, this pattern remained remarkable stable across seasonal fluctuations in the moisture regime. However, a cluster analysis on the changes in readings between two dates can reveal more clearly the distinction between two soil types based on their hydrology (*SHG*, Appendix 2), or the presence of subsoil clay which is less susceptible to a change in signal caused by the hydrology (*FTB*, Appendix 2).

The principal component analysis on the variations in EC_a with soil physical variables clearly showed that subsoil clay and organic matter contents, and topsoil sand and organic matter contents are major determinants of the variability of EC_a across a site (Fig. 30). Topsoil bulk density also proved of important. Since bulk density and clay content seem to be important in determining EC_a values then the measurements are likely to be informative about soil hydrological conditions.

Differentiating variations in soil hydrology

The ability of EMI to record differences in the water content of soils can be useful in one of two ways. Firstly it can identify regions within fields that behave differently hydrologically to those in the rest of the field, such as valley features at both *HLF* (Fig. 11) and *SHG* (Fig. 12). Similarly, it can also locate reserves of water deeper in the profile that are not apparent from surface topography, such as at *HLF* (Fig. 11). The vertical mode of operation proves more apt at this usage, and once again the difference in readings made during different seasons highlights regions of maximum and minimum change more clearly (Figs. 24 & 27), when geo-statistical and cluster analysis techniques are employed.

Secondly, the general ability of clayey soils to hold water enables changes in depth of any clayey layer to be assessed which will not be visible from the surface (*CLY*, Fig. 14 & 15).

The soil moisture content itself proved surprisingly less useful in explaining variability in the EC_a measurements across fields (Table 18). It is thought that this may be because the amounts held in the soil pore space (except at extremes of wetness and dryness) is itself governed by static soil properties such as texture and bulk density. Though this effect does compound the differences due to these features and make spatial differentiation easier (see above section).

The effect of extreme wetness on the signal can be a problem under certain circumstances, as in the case of measurements made at *FTB* in a very permeable soil during a period of high moisture deficit. In this case temporary saturation of the surface soil provided a fairly uniform distribution of EC_a measurements made in the horizontal mode of operation. This masked a spatial pattern which was more apparent in measurements made in the vertical mode. However, this effect could be turned to an advantage, if the location of areas of impeded drainage due to plough pans or compacted topsoil was the purpose of the survey.

Geostatistical considerations

It is necessary to evaluate critically the random function models that we assume underlie our data in a geostatistical analysis. Given the density with which sensor data can be collected there is scope to validate alternative models in a rigorous way using the methodology of Lark (2000).

Matheron's estimator of the variogram is the most efficient statistically, and analysis of data on the original scale is always to be preferred since it avoids complications associated with back-transformation of the final results. Thus if a variogram obtained by Matheron's estimator from the original data is supported by the validation step then this should be used. Alternatives may be considered in the following sequence.

(i) A data transformation should be considered. It may be that a variogram model can be validated for data on a transformed scale, but not on the original scale.

(ii) Robust estimators of the variogram should always be considered and robustly estimated variograms compared to those obtained by Matheron's estimator in any model validation. Robust estimation may be used on the original or transformed data, and both should be tried since the robust estimators assume normality of the underlying process.

(iii) When data have a complex distribution such as a bimodal, the possibility that there are two or more distinct regions requiring a separate spatial analysis should be considered. Regions may be identified from the data, or from other sources (e.g. soil maps). McBratney *et al.* (1991) advocated such a procedure, but given the sparseness of data in most soil studies it has not been widely followed. Again, the validation procedure may be applied to evaluate separate variogram models within the sub-regions.

Recommended protocol for the use of EMI on farms

The research reported here combined with other research and practical experience demonstrates that EMI is likely to be a useful technique for targeting such features as:

- changes in soil type due to texture.
- changes in soil type due to differential hydrology.
- subsoil water reserves in permeable material.
- location of shallow soils and bedrock near the surface.
- drought prone regions within fields.
- clay subsoil features in otherwise sandy material.
- general patterns of soil physical features over wide areas.
- location of buried pipes, old hedgerows, etc.

The manner in which EMI is best used will to some extent be guided by the reason for the survey; e.g. to delineate soil type boundaries, identify soil management zones, map saline or droughty areas. However, there are some general points to consider. Crucially, some *a priori* knowledge of the soils on site is very useful if the capabilities of the EMI instrument are to be fully exploited.

1. The EMI sensor is best used when housed in a non-metallic cart and towed behind an ATV over bare ground or an established crop that is not too tall. Such equipment has been successfully used on cereal crops at tillering stage with no crop damage. Under these circumstances, the EMI instrument (and maybe also the GPS antenna) should be at least 3m away from the vehicle or other metal components to avoid interference of the EMI sensor. An operating speed of around 10–15 kph is recommended, depending upon the size of the field, width of passes and terrain evenness. Since data is recorded at time intervals, the faster the speed of travel, the fewer data-points are recorded. If ground is too rough, excessive ‘bounce’ of the cart can occur which may result in false data. Depending upon the purpose of the survey; the instrument can also be hand held or towed along single transects across features of interest. if GPS is not available, location can be estimated by dead-reckoning, use of landmarks or other survey techniques.

2. The orientation of the EMI sensor can help to either emphasise features in the upper soil profile in the horizontal mode or in the upper subsoil in the vertical mode. If changes in clay content of the topsoil, or the presence of impeded water due to compaction or panning at the base of the topbsoil are the sort of features of interest, then the horizontal mode of operation should be used. If changes in subsoil texture, depth of soil, or the presence of deeper moisture reserves are under consideration, then the vertical orientation would be better. If little is known of the soil problems or changes, but it is hoped to identify management zones in relation to yield maps, then the vertical mode is probably the better general purpose option.

3. Although the patterns revealed by EMI sensing were similar irrespective of whether the equipment was used under wet or dry soil conditions, there are practical advantages to using EMI in the autumn, winter or spring months but not at times following recent heavy rain or when soils are waterlogged. It is important that there is not surplus water in the soil waiting to drain away since this water will affect the EC_a reading yet may not reflect the characteristics of the soil. Provided EMI is used correctly, there is little advantage to be gained in most situations by repeat use of the technique.

4. It is clear from these results that a single spacing between passes will not be optimal for all sites. However, a spacing of about 20 m would be a practically acceptable compromise at any site studied here. Ideally some form of pre-calibration to establish the ideal between-pass spacing would be desirable for a particular field by automating the robust analysis procedure used above. Four or five passes could be made in a field at a narrow spacing of about 6 m, then after a pause while the data are analysed, the optimal spacing could be identified and the rest of the field surveyed at a density planned to ensure that the final map is of adequate precision. Defining 'adequate precision' remains a problem, which will only be solved as more experience is gained in the use of EC_a data, and loss functions become defined for errors of over- or under-estimation of the local EC_a.

EMI surveys are currently being offered commercially by several companies, who may also combine it with other supporting soil investigation or agronomic advisory services. Costs vary, but around £20 per ha is common. Fields of large size can be surveyed rapidly using a cart-based sensor drawn by an ATV with GPS equipment, and timing is flexible so that it can be carried out at times of the year when little or no crop damage is incurred. Fields need only be surveyed once, but to make the best use of resulting maps and data-sets for precision farming, EMI maps must be interpreted in combination with supporting information from conventional, though targeted, soil examination in the field using an auger and spade.

In conclusion, EMI mapping can offer good value for money on farms that have significant soil variation within fields, especially when there is a desire to understand, interpret and manage this variation by use of precision farming methods. It is the best direct method of obtaining soil physical information currently available though cannot be used in isolation of direct soil examination in the field.

GROUND PENETRATING RADAR

Soil properties measured; accuracy and temporal stability.

The applicability of GPR to field survey proved to be severely limited by the soil material found on site. At the site where soil depth was most difficult to visually assess, and barely distinguishable from a fragmented chalk bedrock, *CLY*, it proved impossible even to obtain a set of readings. This was due to the fact that the soils had a high clay content in the topsoil (21–36 % in the top 15 cm) that effectively reflected the signal before it had even penetrated the main body of the soil profile (Fig. 36). This was unfortunate in that the ability to measure the depth of soil is often quoted as an agricultural use for GPR.

Where the instrument was used on sandy material (*HLF*) rather more information was gained, relating in the main to the location of bedrock or free-water interfaces in the profile (Fig. 34). The information which GPR provides and EMI does not, is of course an estimate of the depth in the soil profile at which any features occur (when soil is sandy in nature). However, the operation of GPR sensing is slow and cannot cover the ground as quickly as for EMI. It is not normally used to supply a two dimensional map.

A map of the reflected radar signal can, however, be made if the instrument is used in the 'air launched' manner behind an ATV. The plan obtained (Fig. 37) is superficially similar to those obtainable by EMI, but is not a measure of any easily definable variable, but rather the integrated response to several. Further technological development will be necessary before this becomes a readily used technique.

No protocols are given for the use of GPR to help arable crop management. However, if the use of GPR is developed in future, it could prove most suitable for targeting such features as:

- depth of soil profile in sandy materials.
- depth to clay rich layers.
- depth to water table during dry periods in permeable soil.
- location and depth to buried pipes, boulders or other hardened point source features.

Currently, GPR is only suitable for targeting specific features that are already suspected, and where a depth location is required. It is a service which is not currently offered commercially to farmers, though specialist geophysical contractors are available with the equipment and knowledge to use it (price negotiable but c.£650 per day for measurements on up to a 3 km transect). We do not recommend GPR for use by farmers to obtain soils information, except in highly specialised situations.

SPECTRAL REFLECTANCE

Soil properties measured; accuracy and temporal stability.

Although the project evaluated one map of visible red reflectance which seemed to offer a coherent pattern of variation across the field (Fig. 38) this could not be interpreted with regard to soil type. This is because neither multiple regression and partial effects, nor principal components analysis adequately identified any soil physical component as a correlating variable (Tables 19-21). Topsoil clay content did have a partial effect (Table 20) but this was thought to act mainly through its effects on structure which plays a major role in determining surface roughness and thereby reflectance, but which is largely unpredictable. At present, vehicle- and air-borne surveys of bare soil by spectral reflectance analysis reveal only limited and quite specific information about the surface soil on a site. As such, they cannot yet really be recommended for use in commercial within-field survey work.

The potential use of spectral reflectance techniques to measure soil properties in the field is in its infancy, and is likely to be of more use when carried out on a cropped soil rather than bare soil. Conventional aerial photography has been available for many years and has a proven ability to help map soil characteristics and patterns especially where this are reflected in crop growth (e.g. drought). Archived aerial photographs are an under-utilised resource that farmers could make more use of. Future development of multi-spectral airborne or satellite imagery may provide alternative information sources of value for helping map soil patterns based on patterns of crop growth revealed in these images.

OVERALL CONCLUSIONS AND RECOMMENDATIONS

EMI has been shown to be a reliable method for obtaining information on soil patterns that may occur in fields. Although EMI can not provide quantitative information on soil properties, it can help target field investigations and provide the basis for delineating areas or zones within fields which can be considered to have similar soil properties. Such zones can be used for purposes of soil nutrient sampling or as a basis for reaching decisions on the implementation of variable rate application of lime or fertilisers.

As with any technology, the practical use of EMI by farmers will need to be targeted. Use on soil landscapes which are intrinsically variable (e.g. soils developed on glacio-fluvial material) is likely to be more cost-effective than on other landscapes which are know to have less variability (e.g. many clay-lands). Its use must be combined with manual soil examination but this can be much less and more targeted than would otherwise be possible. Farmers should also consider other potential information sources on soil properties, such as existing archives of aerial photographs that exist going back many years. Although there is a high element of ‘pot luck’ in terms of timing and location of these photographs, ones showing patterns of crop growth can be very informative about soil patterns. Simple aerial photographs can be obtained to order and forthcoming airborne and/or satellite imagery may also provide useful information on soil patterns as reflected by variations in crop growth.

Future work on EMI should be focused on demonstrating the practical integration of EMI with other techniques for the cost-effective gathering of information and making decisions on the management of the spatial variability of crops both within and between fields on a whole farm (i.e. precision farming). Its value, use and cost-effectiveness as part of an integrated approach will then become more clear to farmers.

The potential value of EMI to measure soil mineral nitrogen (SMN) is another potential direction. Nitrate (NO_3) is the main component of SMN and has an important effect on EC_a. The sensitivity of EC_a to changes of SMN is not known, but bearing in mind the crucial importance of SMN for making correct N decisions, the advent of NVZs and the current high cost of soil sampling and analysis for SMN, development of a sensor-based method would be attractive.

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In Petov, B.N. and Csaki, F., (Eds*).* *2nd International Symposium on Information Theory.* Akademia Kiado, Budapest. pp. 267–281.

Baret F., Jacquemoud S. & Hanocq J.F. 1993. The soil line concept in remote sensing. *Remote Sensing Review,* **7;** 65-82.

Barnes, E.M., Moran, M.S., Pinter, P.J. & Clarke, T.R. 1996. Multispectral remote sensing and site-specific agriculture: examples of current technology and future possibilities. *Proceedings of the Third International Conference on Precision Agriculture.* 1996. Ed., Robert, P.C., Rust, R.H. & Larson W.E., Unniversity of Minnesota, St. Paul, Minnesota, USA. Publ., ASA, CSSA, SSSA, Madison, WI. 845-854.

Bedrick, E.J. and Tsai, C.L. 1994. Model selection for multivariate regression in small samples. *Biometrics.* **50;** 226–231.

Cressie, N. & Hawkins, D. 1980. Robust estimation of the variogram. *Mathematical Geology*, **12;** 115–125.

Dampney P.M.R., Bryson R., Clark W., Strang M., & Smith A. 1998. The use of sensor technologies in agricultural cropping systems. A scientific review and recommendations for cost-effective developments. *Report to MAFF contract CE0140.*

Dowd, P.A. 1984. The variogram and kriging: robust and resistant estimators. In: *Geostatistics for Natural Resources Characterization* (eds G. Verly, M. David, A.G. Journel & A. Marechal), Part 1. pp. 91–106. D. Reidel, Dordrecht.

Genton, M.G. 1998. Highly robust variogram estimation. *Mathematical Geology*, **30;** 213–221.

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, UK. 1984. 450pp.

Jansinski M.F. & Eagleson P.S. 1989. The structure of red-infrared scattergrams of semi-vegetated landscapes. *IEEE Transactions Geosciences in Remote Sensing,* **27;**: 441-451.

Jarvis M G, Allen R H, Fordham S J, Hazelden J, Moffat A J, & Sturdy R G. 1984. Soils and their use in South East England. *Soil Survey of England and Wales, Bulletin No. 15*. Harpenden, UK. 1984. 405pp.

King J A, & Dampney P M R. 1999. The potential of remote sensing technologies for measuring soil parameters: A scientific review and R & D recommendations. *Report to MAFF, project No. CE0164.*

King J.A., Dampney P.M.R., Lark M., Mayr T.R. & Bradley R.I. 2001. Sensing soil spatial variability by electro-magnetic induction (EMI): Its potential in precision farming. In, *Proceedings of the Third European Conference on Precision Agriculture (Vol. 1), June 18-20, Montpellier, France.* 419-424. Publ., Agro Montpellier, Ecole Nationale Supérieure Agronomique, 2001.

Kollert R. 1969. Ground water exploration by the electrical resistivity method. *Geophysical memorandum 3/69. ABEM printed matter No. 90081.* ABEM Geophysics & Electronics.

Lark, R.M. 2000a. A comparison of some robust estimators of the variogram for use in soil survey. *European Journal of Soil Science.* **51;** 137-157

Lark, R.M. 2000b. Regression analysis with spatially auto-correlated error: examples with simulated data and from mapping of soil organic matter content. *International Journal of Geographical Information Science.* **14**; 247–264.

Lark, R.M. 2001. Some tools for parsimonious analysis and interpretation of within-field variation. *Soil and Tillage Research* **58;** 99–111.

Lark, R.M. 2002. Robust estimation of the pseudo cross-variogram for co-kriging soil properties. *European Journal of Soil Science.* **53;** 253–270.

Lark, R.M., Wheeler, H.C., Bradley, R.I., Mayr, T. & Dampney, P.M.R. 2003a. Developing a cost-effective procedure for investigating within-field variation of soil conditions. *HGCA Project Report (No. 2116) 158pp.*

Lark, R.M., Kaffka, S.R. & Corwin, D.L. 2003b. Multi-resolution analysis of data on electrical conductivity of soil using wavelets. *Journal of Hydrology*, **272;** 276-290.

McBratney, A.B., Webster, R. & Burgess, T.M. 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. *Computers and Geosciences*, **7;** 331–334.

McBratney, A.B., Hart, G.A. & McGarry, D. 1991. The use of region partitioning to improve the representation of geo-statistically mapped soil attributed. *Journal of Soil Science*, **42;** 513–532.

McNeill D J. 1980. Electromagnetic terrain conductivity measurement at low induction numbers. *Technical Note TN-6, Geonics Ltd., Ontario, Canada.*

Matheron, G. 1962. *Traité de Géostatistique Appliqué*, Tome 1. Memoires du Bureau de Recherches Géologiques et Minières, Paris.

Payne, R.W., Lane, P.W., Ainsley, A.E., Bicknell, K.E., Digby, P.G.N., Harding, S.A., Leech, P.K., Simpson, H.R., Todd, A.D., Verrier, P.J., White, R.P., Gower, J.C., Tunnicliffe Wilson, G. & Paterson, L.J. 1988. *Genstat 5 reference manual.* Clarendon Press, Oxford. 749pp.

Rondeaux G., Steven, M. & Baret F. 1996. Optimization of soil-adjusted vegetation indices. *Remote Sensing Environment,* **55;** 95-107.

Steven M.D. & Clark J.A. 1990. *Application of remote sensing in agriculture*. Publ., Butterworths, London. 1990.

Webster, R. 1997. Regression and functional relations. *European Journal of Soil Science.* **48;** 557–566.

Webster, R. & Oliver, M.A. 2001. *Geo-statistics for Environmental Scientists.* John Wiley & Sons, Chichester.

APPENDIX 1

SOIL PROPERTIES FOR THE SENSOR SITES

THE CLAYS, BENSON, OXFORDSHIRE

Wallop **clayey, lithoskeletal chalk; brown rendzina; wetness class 1**

Brief profile description

0-25cm Ah	Dark brown moderately stony silty clay loam or silty clay; strong fine subangular blocky structure; calcareous.
25-30cm BCu	Dark brown, moderately stony silty clay; weak medium angular blocky structure; calcareous.
30-40cm 2Cu	Fragmented chalk and occasional flints.
At 40cm 2Cr	Weakly bedded chalk.

Frilsham **fine loamy material over lithoskeletal chalk; typical argillic brown earth; wetness class 1**

Brief profile description

0-25cm Ap	Dark brown slightly stony sandy clay loam or clay loam.
25-50cm EBt	Strong brown, slightly stony sandy clay loam or clay loam; moderate medium angular blocky structure.
50-70cm Bt	Brown, slightly stony sandy clay loam or clay loam; moderate coarse angular blocky structure.
At 70cm 2Cu	Hard white chalk rubble.

Soham **fine loamy material over lithoskeletal chalk; typical calcareous brown earth; wetness class 1**

Brief profile description

0-30cm Ap	Dark greyish brown slightly stony sandy clay loam; calcareous.
30-50cm Bw	Brown or strong brown, slightly stony sandy clay loam; weak medium angular blocky structure; calcareous.
50-55cm Cu	Yellowish brown, extremely stony clay loam; very calcareous.

FIELD 107, HEYDOUR LODGE, LINCOLNSHIRE

Elmton	**fine loamy, lithoskeletal limestone; brown rendzina; wetness class 1**

Brief profile description

0-25cm Ap	Brown slightly or moderately stony clay loam or sandy clay loam; calcareous.
At 25cm R	Limestone

Cranwell	**coarse loamy; lithoskeletal limestone; brown rendzina; wetness class 1**

Brief profile description

0-25cm Ap	Brown slightly or moderately stony sandy loam; calcareous.
At 25cm R	Limestone

Wilsford	**sandy; stoneless drift; typical brown sand; wetness class 1**

Brief profile description

0-25cm Ap	Dark brown stoneless sandy loam or loamy sand.
25-55cm Bw	Brown, stoneless loamy sand or sand; weak fine subangular blocky structure.
55-120cm Cu	Yellowish red or brownish yellow, stoneless sand; single grain structure.

Haselor	**swelling clayey material passing to clay with interbedded limestone; typical calcareous pelosol; wetness class 3.**

Brief profile description

0-25cm Ah	Very dark greyish brown slightly stony clay; calcareous.
25-55cm Bw(g)	Olive brown, mottled, slightly or moderately stony clay; strong coarse angular blocky structure; calcareous.
55-85cm Cr	Grey, fine grained limestone bands with interbedded clay shale.

Oxpasture **fine loamy or fine silty drift over clayey material passing to clay or soft mudstone; stagnogleyic argillic brown earth; wetness class 2.**

Brief profile description

0-25cm Ap	Dark brown, slightly mottled, slightly stony silty clay loam or clay loam.
25-45cm Eb(g)	Brown, slightly mottled, slightly stony clay loam; moderate coarse subangular blocky structure.
45-75cm 2Bt(g)	Yellowish brown, mottled, stoneless clay; strong coarse angular blocky structure.
75-100cm 2BCg	Yellowish brown with many grey mottles, stoneless clay or silty clay; strong fine platy, prismatic or massive structure.

FOOTBALL FIELD SHUTTLEWORTH, OLD WARDEN, BEDFORDSHIRE

Bearsted **coarse loamy material passing to sand or soft sandstone; typical brown earth; wetness class 1**

Brief profile description

0-25cm Ap	Dark brown slightly stony sandy loam or sandy silt loam.
25-45cm Bw	Brown, slightly stony sandy loam or sandy silt loam; weak fine subangular blocky structure.
45-70cm BCu	Brownish yellow, slightly or moderately stony sandy loam or loamy sand; weak coarse angular blocky or massive structure.
70-120cm Cu	Pale yellow, stoneless loamy sand; single grain structure.

Cottenham sandy loam top **sandy material passing to sand or soft sandstone; typical brown sand; wetness class 1**

Brief profile description

0-25cm Ap	Dark brown slightly stony sandy loam.
25-60cm Bw	Reddish brown, slightly stony loamy sand; weak coarse subangular blocky structure.
60-100cm Cu	Yellowish red, slightly or moderately stony loamy sand or sand; single grain structure iron stone fragments.

Cottenham loamy sand top **sandy material passing to sand or soft sandstone; typical brown sand; wetness class 1**

Brief profile description

0-25cm Ap	Dark brown slightly stony loamy sand.
25-60cm Bw	Reddish brown, slightly stony loamy sand; weak coarse subangular blocky structure.
60-100cm Cu	Yellowish red, slightly or moderately stony loamy sand or sand; single grain structure iron stone fragments.

Oxpasture **fine loamy or fine silty drift over clayey material passing to clay or soft mudstone; stagnogleyic argillic brown earth; wetness class 3.**

Brief profile description

0-25cm Ap	Dark brown, slightly mottled, slightly stony silty clay loam or clay loam.
25-45cm Eb(g)	Brown, slightly mottled, slightly stony clay loam; moderate coarse subangular blocky structure.
45-75cm 2Bt(g)	Yellowish brown, mottled, stoneless clay; strong coarse angular blocky structure.
75-100cm 2BCg	Yellowish brown with many grey mottles, stoneless clay or silty clay; strong fine platy, prismatic or massive structure.

Hanslope **clayey chalky drift; typical calcareous pelosol; wetness class 2**

Brief profile description

0-25cm Ap	Dark greyish brown slightly stony clay or clay loam; slightly calcareous.
25-60cm Bw(g)	Light olive brown, slightly mottled, slightly stony clay; moderate medium subangular blocky structure; calcareous.
60-100cm BCg	Yellowish brown, mottled, slightly or moderately stony clay; moderate medium angular blocky or prismatic structure; calcareous with chalk stones.

SHAGSBY FIELD, CHICKSANDS, BEDFORDSHIRE

Evesham	**swelling clayey material passing to clay or soft mudstone; typical calcareous pelosol; wetness class 3..**

Brief profile description

0-25cm Ap	Dark greyish brown stoneless clay; calcareous.
25-40cm Bw(g)1	Olive brown, slightly mottled, stoneless clay; moderate medium subangular blocky structure; calcareous.
40-75cm Bw(g)2	Light olive brown, slightly mottled, stoneless clay; strong medium subangular blocky structure; calcareous.
75-120cm BC(g)	Grey, slightly mottled, stoneless clay; massive structure; calcareous.

Bearsted	**coarse loamy material passing to sand or soft sandstone; typical brown earth; wetness class 1**

Brief profile description

0-25cm Ap	Dark brown slightly stony sandy loam or sandy silt loam.
25-45cm Bw	Brown, slightly stony sandy loam or sandy silt loam; weak fine subangular blocky structure.
45-70cm BCu	Brownish yellow, slightly or moderately stony sandy loam or loamy sand; weak coarse angular blocky or massive structure.
70-120cm Cu	Pale yellow, stoneless loamy sand; single grain structure.

Waterstock **fine loamy drift with siliceous stones; gleyic argillic brown earth; wetness class 2.**

Brief profile description

0-25cm Ap	Brown slightly stony clay loam.
25-40cm Bt	Yellowish brown, slightly stony clay loam; moderate medium subangular blocky structure.
40-55cm Bt(g)	Yellowish brown, slightly mottled slightly stony clay loam; moderate medium subangular blocky structure.
55-85cm 2Btg	Pale yellow, mottled, slightly stony clay loam; weak fine angular blocky.
85-120cm 3BCtg	Brown, mottled, slightly to moderately stony sandy loam or clay loam; structureless.

Cottenham sandy loam top **sandy material passing to sand or soft sandstone; typical brown sand; wetness class 1**

Brief profile description

0-25cm Ap	Dark brown slightly stony sandy loam.
25-60cm Bw	Reddish brown, slightly stony loamy sand; weak coarse subangular blocky structure.
60-100cm Cu	Yellowish red, slightly or moderately stony loamy sand or sand; single grain structure iron stone fragments.

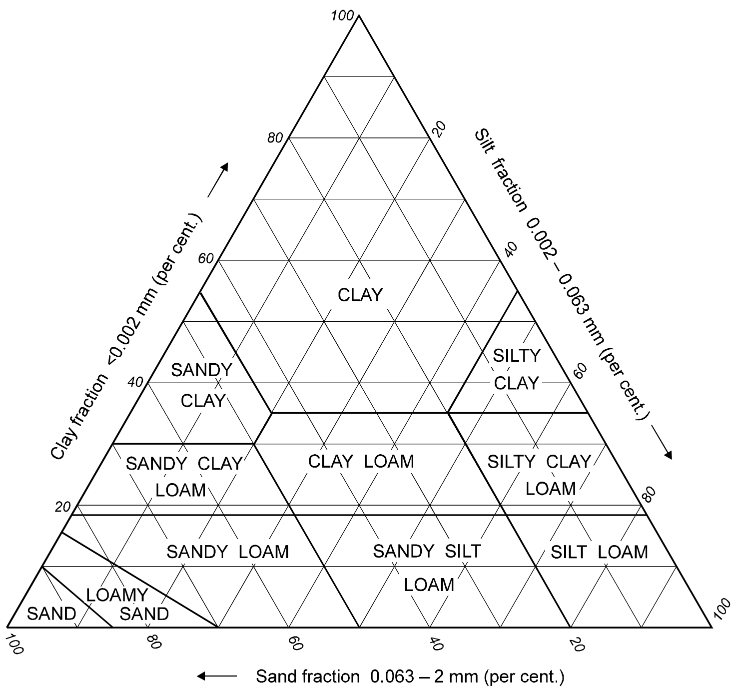
Oxpasture **fine loamy or fine silty drift over clayey material passing to clay or soft mudstone; stagnogleyic argillic brown earth; wetness class 3.**

Brief profile description

0-25cm Ap	Dark brown, slightly mottled, slightly stony silty clay loam or clay loam.
25-45cm Eb(g)	Brown, slightly mottled, slightly stony clay loam; moderate coarse subangular blocky structure.
45-75cm 2Bt(g)	Yellowish brown, mottled, stoneless clay; strong coarse angular blocky structure.
75-100cm 2BCg	Yellowish brown with many grey mottles, stoneless clay or silty clay; strong fine platy, prismatic or massive structure.

Note: although Football and Shagsby fields have a similar range of soil series, in the former the sand is up-slope from the clay whereas the reverse is true in the latter.

The clay content of the material in The Clays depends on the method of analysis - hand texturing gives a higher proportion clay than laboratory measurements because of the dispersion methods used in the analyses.



<u>**TOPSOIL TEXTURE GROUP**</u>

The classes in the triangular diagram are simplified into the following texture groups:

Texture group	Texture classes
Sandy	Sand, loamy sand
Coarse loamy	Sandy loam, sandy silt loam
Fine loamy	Clay loam, sandy clay loam
Coarse silty	Silt loam
Fine silty	Silty clay loam
Clayey	Clay, silty clay, sandy clay

<u>**STONINESS**</u>

NSRI uses the following groupings to relate percent stone content (volume basis) to a descriptive term.

Stoneless	Less than 1%
Very slightly stony (few)	1 to 5%

Slightly stony (common)	6 to 15%
Moderately stony (many)	16 to 35%
Very stony (abundant)	More than 35%

WETNESS CLASS

The duration and degree of water-logging are described by the system of wetness classes grading from Wetness Class 1, well drained, to Wetness Class 6, almost permanently waterlogged within 40 cm depth. The incidence of water-logging depends on soil and site properties, under-drainage and climate. The classes given for each soil assume an appropriate level of under-drainage.

1	Soil profile is not waterlogged within 70 cm depth for more than 30 days[1] in most years[2]
2	Soil profile is waterlogged within 70 cm depth for 30 to 90 days in most years
3	Soil profile is waterlogged within 70 cm depth for 90 to 180 days in most years
4	Soil profile is waterlogged within 70 cm depth for more than 180 days, but not waterlogged within 40 cm depth for more than 180 days in most years
5	Soil profile is waterlogged within 40 cm depth for 180 to 335 days and is usually waterlogged within 70 cm depth for more than 335 days in most years
6	Soil profile is waterlogged within 40 cm depth for more than 335 days in most years

1 The number of days specified is not necessarily a continuous period

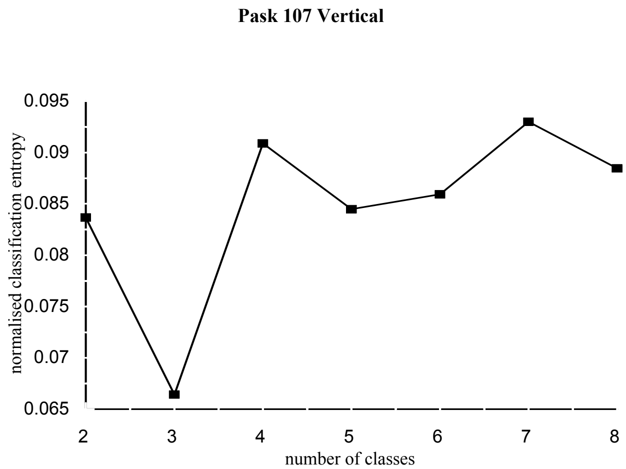
2 *In most years* is defined as more than 10 out of 20 years.

APPENDIX 2.

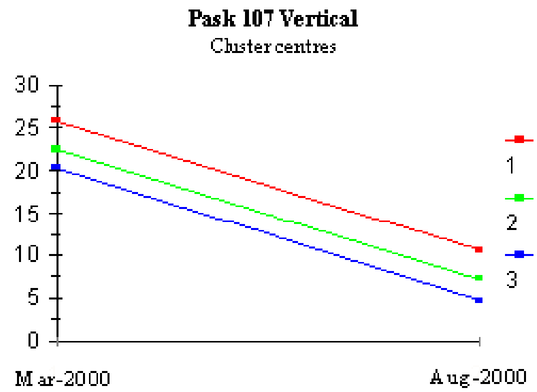
HLF (FIELD 107, PASKS FARM)

Vertical Mode

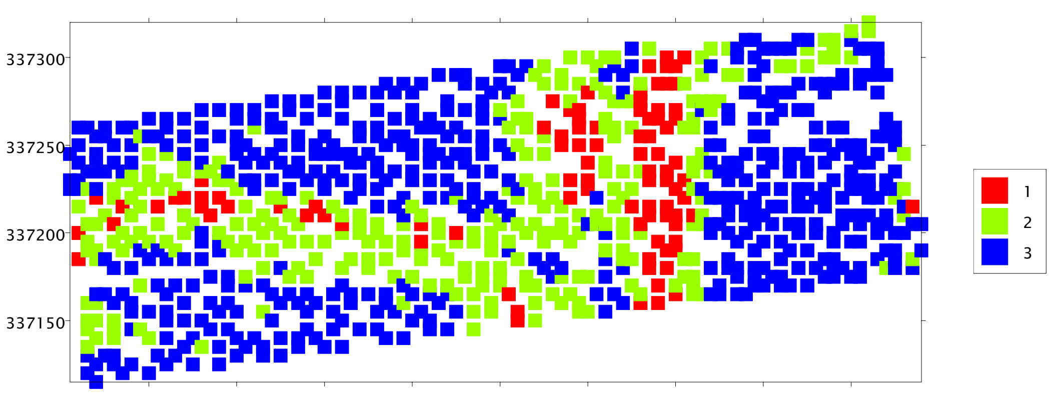
NCE plot, 3 selected



Cluster centres



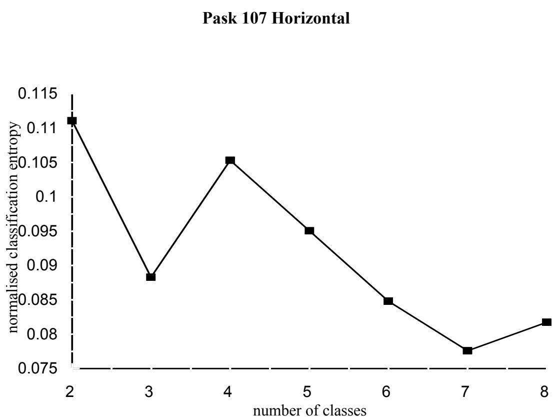
Pask 107 vertical EMI cluster map (March 2000 & August 2000)

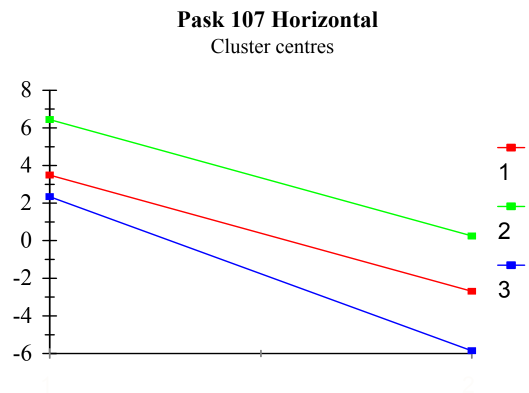


Horizontal Mode

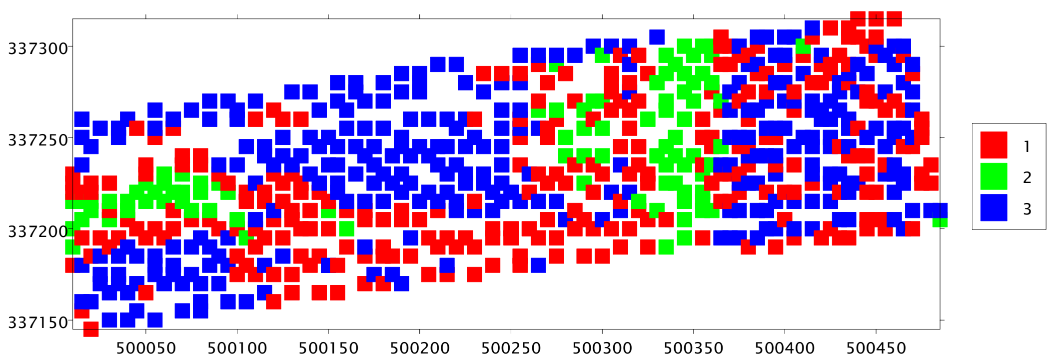
NCE plot, 3 selected

Cluster centres





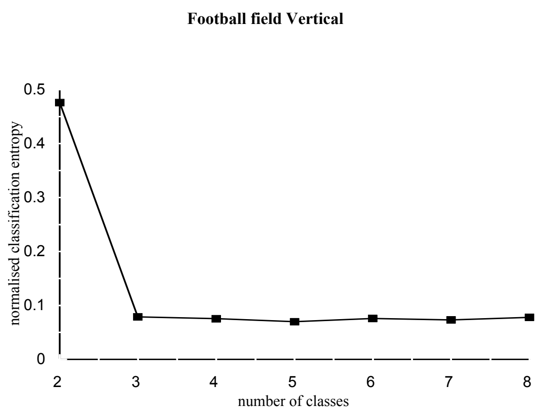
Pask 107 horizontal EMI cluster map (March 2000 & August 2000)



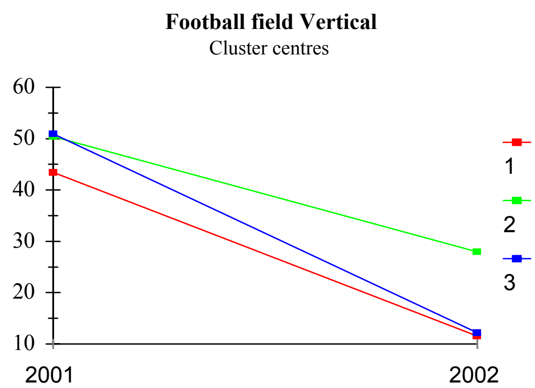
***FTB* (FOOTBALL FIELD, SHUTTLEWORTH FARMS)**

Vertical Mode

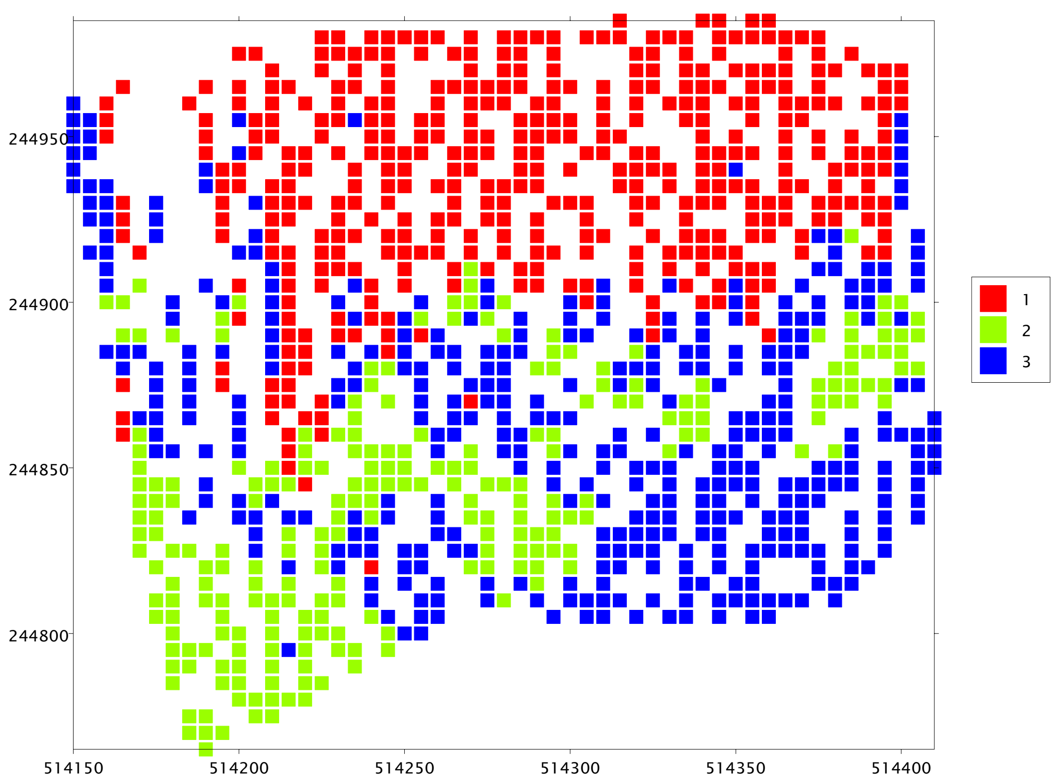
NCE plot, 3 selected



Cluster centres



Football Field Vertical EMI cluster map (2001 & 2002)

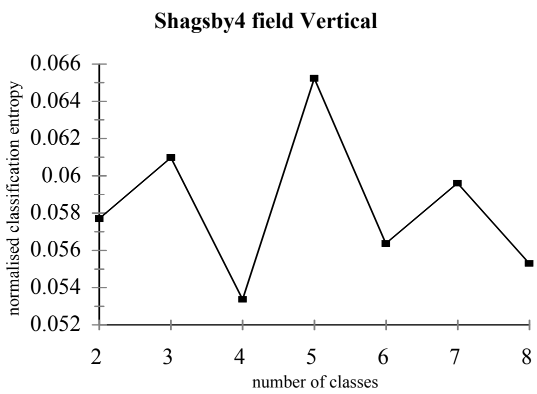


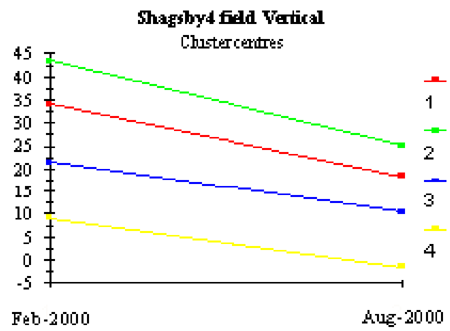
***SHG* (SHAGSBY 4 FIELD, LODGE FARM.)**

Vertical Mode

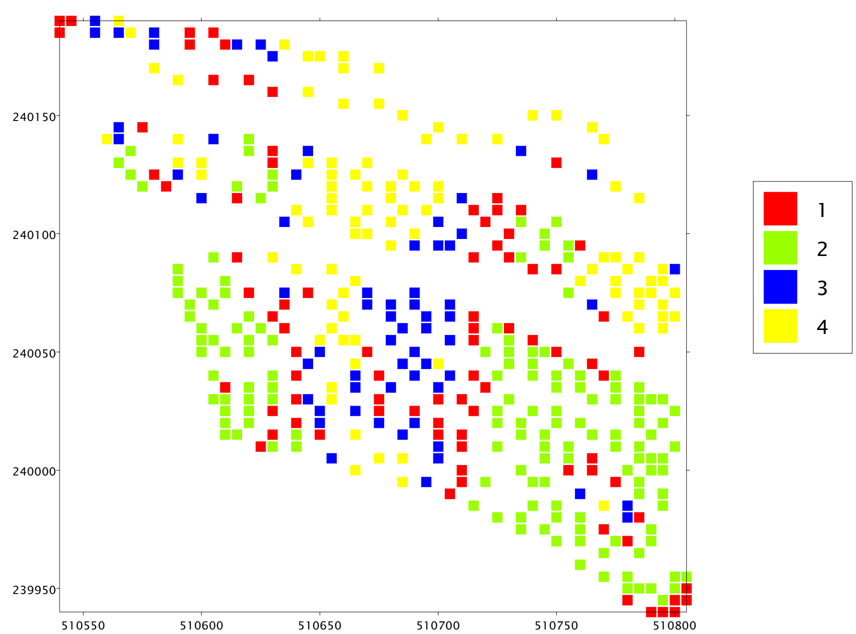
NCE plot, 4 selected

Cluster centres





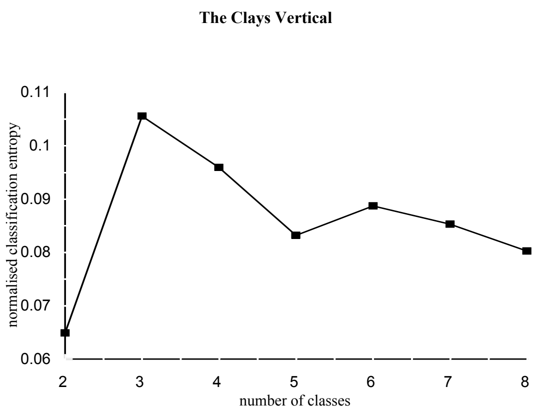
Shagsby4 vertical EMI cluster map (February & August 2000)



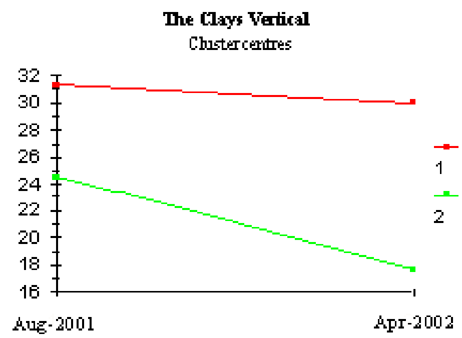
***CLY* (THE CLAYS, CROWMARSH FARM)**

Vertical Mode

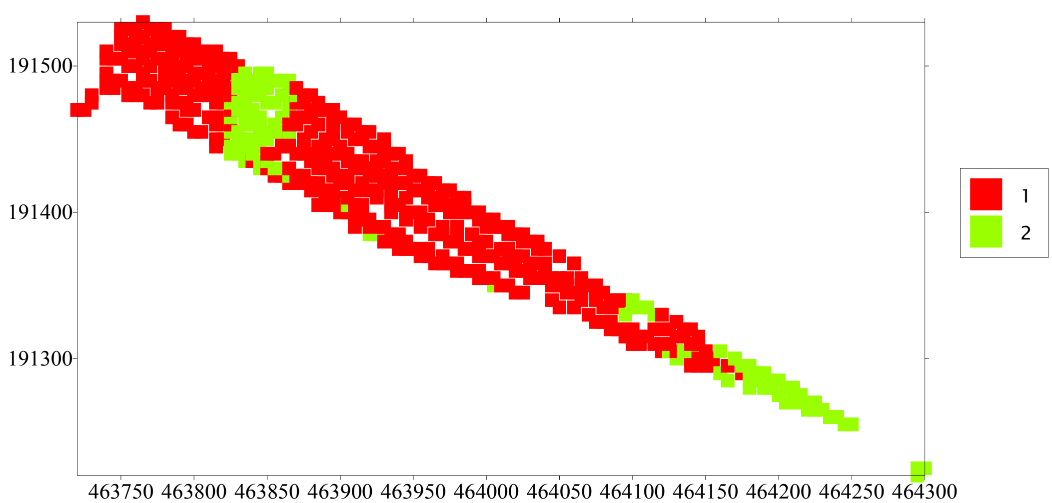
NCE plot, 2 selected



Cluster centres



The Clays vertical EMI cluster map (August 2001 & April 2002)

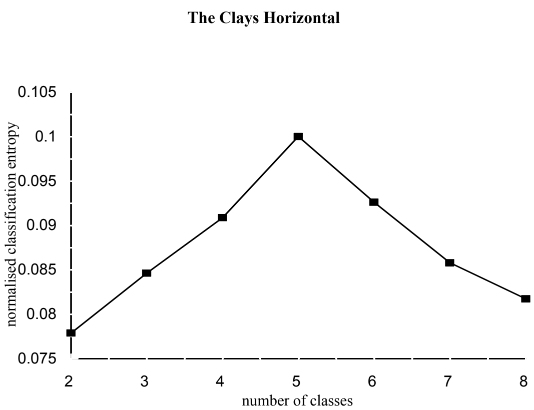


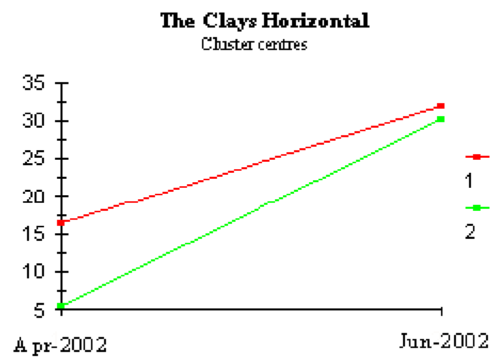
CLY **(THE CLAYS, CROWMARSH FARM)**

Horizontal Mode

NCE plot, 2 selected

Cluster centres





The Clays horizontal EMI cluster map (April 2002 & June 2002)