Geoderma 199 (2013) 99–105

Contents lists available at SciVerse ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Key variables for the identification of soil management classes in the aeolian landscapes of north–west Europe

Marc Van Meirvenne *, Mohammad Monirul Islam, Philippe De Smedt, Eef Meerschman, Ellen Van De Vijver, Timothy Saey

Research Group Soil Spatial Inventory Techniques, Department of Soil Management, Ghent University, Coupure 653, 9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 1 September 2011
Received in revised form 16 February 2012
Accepted 22 July 2012
Available online 15 November 2012

Keywords:
Apparent electrical conductivity
Boundary line analysis
EM38-MK2
Gamma ray
pH
Yield modelling

ABSTRACT

At present, spatially very detailed data sets can be obtained about soil, landscape and crop variability. However, there is a need to select independent key properties to identify management classes needed for precise land management. In a previous study performed in the European loess belt, topsoil pH, apparent electrical conductivity (ECa) and elevation were identified as key properties. In this study we enlarged the number of soil properties by including gamma ray measurements and employed a similar methodology to a field in the sand belt of northern Europe. Based on a principal component analysis we identified the same three variables as key properties. This was surprising given the big differences in landscape topology and pedogenesis between the loess and sand areas. These three key variables were used to delineate management classes using a fuzzy *k*-means with extragrade classification procedure. This classification was evaluated by mapping the wheat grain yield in the year 2006. A multiple regression model could be constructed that predicted yield from ECa and elevation well ($R^2_{adj} = 0.88$). To analyse the influence of ECa on crop yield in depth a boundary line analysis was conducted. The boundary line could be modelled with an excellent R^2_{adj} of 0.98. It was concluded that ECa, elevation and pH are generic key variables for the delineation of management classes of the aeolian landscapes of north–west Europe. Given its integral nature and strong relationship with crop performances, the authors plea to upgrade ECa from a “secondary” (proxy) source of information to a “primary” variable which can be used directly as a basis for detailed soil mapping of the bulk soil.

1. Introduction

As recent as two decades ago soil inventory studies searched for optimal methods to complement sparse soil data with the best available interpolation techniques, resulting in advances in pedometrics and geostatistics (Goovaerts, 1997). To overcome the limitations of spatially scarce data, the focus shifted to measurement technologies creating dense data sets of secondary soil property information e.g. soil apparent electrical conductivity (ECa) (McBratney et al., 2000). Progress in sensing technology for precision agriculture and water management resulted in advanced data capture techniques, providing high-resolution information about crops (a.o. yield maps) and landscapes (a.o. elevation maps) (Schellberg et al., 2008). As a result, we now have access to a wealth of detailed data on soil, landscape and crop properties. However, these rich data sets often duplicate information, potentially masking the complex interactions that may exist among a few independent variables (Van Meirvenne, 2006). Therefore, the challenge is to select the key variables which are essential for the faced problem.

One of the applications of detailed soil inventories is precision land management based on management classes (Moral et al., 2010). These classes are assumed to delineate zones where management can be implemented in a homogeneous way (Pedroso et al., 2010). Given the strong interactions between crops and climate, and the instability of the nature of such interactions between years, it is better to rely on time-stable land properties to create management classes (Corwin and Lesch, 2005). Ideally these properties should be captured quickly at a detailed scale, i.e. with proximal or remote sensors.

Although early research on the concept of site specific management of soils based on varying soil conditions within a field dates back to the 1990s (Doolittle et al., 1994; Jaynes et al., 1993; Larson and Robert, 1991; Mulla, 1991; Mulla and Schepers, 1997; Sudduth et al., 1995), one of the first efforts to identify key properties for the delineation of management zones in Europe was made by Vitharana et al. (2008). These authors focused on a field located within the European aeolian loess belt, i.e. a more or less continuous zone of silty soil in an undulating landscape and east–west orientated throughout Europe (Haase et al., 2007). The WRB (World Reference Base for soil resources) Major Soil Groups of this belt comprise generally Luvisols and Albeluvisols (European Soil Bureau Network, 2005). In their study 12 soil and topographic variables, including top- and subsoil

* Corresponding author. Tel.: +32 9 2646 056; fax: +32 9 2646 247.
E-mail address: marc.vanmeirvenne@ugent.be (M. Van Meirvenne).

0016-7061/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.geoderma.2012.07.017

textural fractions, organic carbon and pH, were considered. A principal component analysis showed that three could be identified as the most independent ("orthogonal"): pH, ECa and elevation. These were called "key variables" for the delineation of potential management classes in the European loess area (Vitharana et al., 2008). The significance of these key variables was tested by using them to define management classes (using a fuzzy *k*-means procedure) and evaluating these classes in terms of crop performances.

The aim of this study was to use the procedure described by Vitharana et al. (2008) but to extend the study to another important European aeolian soil-landscape: the north European sand belt. This zone stretches from Great Britain over Belgium and the Netherlands to Germany and Poland (Jungerius and Riksen, 2010) and is situated largely north of the loess belt. Generally, the soil can be classified as a Podzol (European Soil Bureau Network, 2005), although agricultural tillage has mostly destroyed the typical profile development in the top decimetres. Similar to the loess belt, the sand area is generally considered to be rather homogeneous, as both belts were created by fluvio-eolian activities associated with the Weichselian glaciation (Derese et al., 2009). Given the distinctly different pedological processes (podsolisation instead of eluviation/illuviation of clay) our working hypothesis was that in the sand area other key variables would be identified for the definition of land management classes. Therefore we extended our study to more soil variables than those used in the study of Vitharana et al. (2008). The final evaluation of the created classes will be conducted using crop yield data.

2. Study site and measured variables

A 4.2 ha arable field located in the sand area of northern Belgium (with central co-ordinates: 3° 37′ 22.43″ E and 51° 07′ 00.03″ N), which is part of the greater sand belt of northern Europe, was selected (Fig. 1). Maize, potatoes and wheat are the main crops used in a typical crop rotation for the area.

Five types of geo-referenced information were collected on this field: (1) topographic information, (2) five soil properties analysed from soil auger samples at two depths, (3) four sets of detailed ECa measurements obtained from an electromagnetic induction (EMI) sensor, (4) four concentrations of radionuclides in the topsoil measured with a gamma ray sensor, and (5) wheat grain yield. Since wheat yield was considered as the target variable it will be used to validate the created management classes.

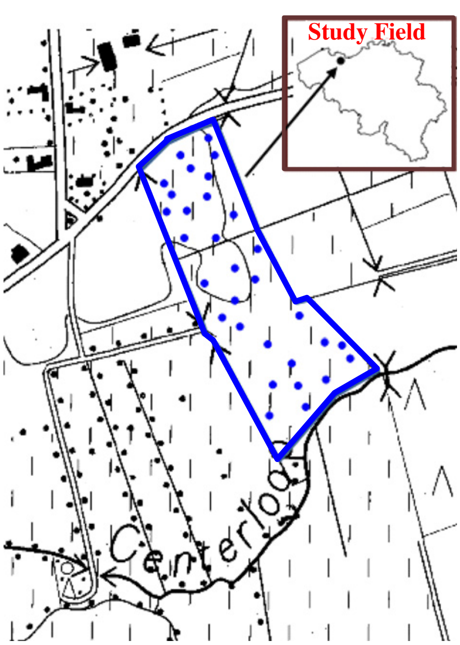

Fig. 1. The study field delineated on a topographic map of the area with its approximate position within Belgium (insert frame) and 30 soil sampling locations (dots within the field).

2.1. Topography

Detailed data on the elevation of the field were obtained by *L*ight *D*etection *a*nd *R*anging (LIDAR) using an airborne laser scanner (OC-GIS Vlaanderen, 2003). On average, one observation per 4 m^2 was taken with average horizontal and vertical errors of 0.14 m and 0.20 m, respectively. These data were used to build a digital elevation model (DEM) of the field at a 5 m resolution using ordinary kriging. Table 1 shows that the elevation of the study field is low (6–7 m) and rather uniform (max. difference is only 1.8 m). Given the flat topography of the field no other topographic indices (like the wetness index or the stream power index) were considered, contrary to Vitharana et al. (2008).

2.2. Soil samples

In May 2009 soil samples were taken at 30 locations chosen randomly out of 100 stratified random grid positions projected over the field. Two depth intervals were sampled: 0–0.3 m (referred to as topsoil *t*) and 0.6–0.9 m (subsoil *s*) (Fig. 1). At each location a pooled sample was obtained from 3 augerings within a 1 m radius. All samples were air dried, sieved through a 2 mm sieve and analysed for organic carbon (OC) (%), pH (in a 1 N KCl solution) and three textural fractions (clay: 0–2 μm, silt: 2–50 μm and sand: 50–2000 μm) by using the conventional methods [i.e. Walkley and Black (1934), pH electrode and pipette-sieve method, respectively].

From Table 1 it can be seen that the average texture of both the topsoil and subsoil is sand loam [USDA (United States Department of Agriculture) texture triangle], which is characteristic for these fluvio-aeolian deposits. However, a considerable variation is present, especially in the silt fraction with the largest coefficient of variation

Table 1
Some descriptive statistics of all variables (*n* = number of observations, CV = coefficient of variation).

Variables	*n*	Mean	Minimum	Maximum	Variance	CV (%)
Topography						
Elevation (DEM) (m)	1867	6.9	5.8	7.6	0.1	5
Soil sample properties						
Topsoil (0–0.3 m)						
Sand (%)	30	73	64	87	55.4	10
Silt (%)	30	19	8	28	37.8	32
Clay (%)	30	8	4	11	2.6	21
OC (%)	30	1.3	1.0	1.9	0.05	18
pH-KCl	98	5.3	4.6	5.6	0.09	6
Subsoil (0.6–0.9 m)						
Sand (%)	30	71	53	95	290.1	24
Silt (%)	30	25	2	44	278.5	67
Clay (%)	30	4	2	6	0.9	23
OC (%)	30	1.0	0.7	1.6	0.05	24
pH-KCl	30	6.1	5.2	7.6	0.45	11
EMI measurements						
ECa-H.5 (mS m^{-1})	87,673	18	7	39	6.8	15
ECa-H1 (mS m^{-1})	87,673	29	8	48	13.6	13
ECa-V.5 (mS m^{-1})	87,670	28	7	48	13.7	13
ECa-V1 (mS m^{-1})	87,670	33	11	55	19.4	13
Gamma-ray measurements						
^{40}K (Bq kg^{-1})	30	299	248	380	1154.7	11
^{238}U (Bq kg^{-1})	30	13	8	21	7.9	22
^{137}Cs (Bq kg^{-1})	30	14	10	19	5.1	16
^{232}Th (Bq kg^{-1})	30	3	0	6	3.7	64

(CV) of the three fractions, and for all textural fractions this variation was the largest in the subsoil (with CVs being twice the topsoil CVs for sand and silt). The OC content was moderate (on average 1.3%) in the topsoil, and was still relatively large in the subsoil (on average 1.0%) which reflects the podsolisation process of organic matter leaching in the topsoil and accumulating in the subsoil. For both layers the CV of OC was around 20%. The pH indicated acid conditions, both in the top- and subsoil, with the subsoil being somewhat less acidic (according to the farmer no lime was added for at least 3 decades).

2.3. EMI measurements

Detailed measurements of the soil ECa were collected with the EMI sensor EM38-MK2 (Geonics Ltd., Canada) mounted on a sled and pulled by an all-terrain vehicle which drove at a speed of approximately 5–7 km h^{-1}. The EM38-MK2 consists of two receiver coils at 0.5 and 1.0 m distances from which measurements were taken every second both in horizontal (ECa-H.5 and ECa-H1) and vertical orientations (ECa-V.5 and ECa-V1) on 20 May 2010. The theoretical depth of influence (DOI — i.e. conventionally the depth below the sensor at which 70% of the cumulative influence of the signal is obtained) of these configurations is: 0.38 m for the H.5 orientation, 0.75 m for both the H1 and V.5 orientations (but with a different distribution of the depth sensitivity) and 1.50 m for the V1 orientation (McNeill, 1980). A Trimble AgGPS332, with an Omnistar satellite correction, provided differential GPS (DGPS) measurements with a pass-to-pass accuracy of approximately 0.10 m. Measurements were taken along parallel lines with an in-between distance of 2 m while the vehicle driving was supported by a Trimble Lightbar Guidance System. All measurements were standardised to the reference temperature of 25 °C according to Sheets and Hendrickx (1995):

$$\mathrm{ECa}_{25} = \mathrm{ECa}_T\left(0.4470 + 1.4034\mathrm{e}^{-T/26.815}\right) \tag{1}$$

with ECa_{25} being the standardised ECa at 25 °C and ECa_T the ECa values at soil temperature T (°C). During the survey T was recorded by a bimetal sensor pushed in the soil to a depth of 0.25 m. In the remaining part of this paper all ECa values refer to ECa_{25}.

Table 1 contains the statistics of the ECa measurements. The mean values show an increasing value with increasing DOI: from 17 mS m^{-1} for H.5 over 28 mS m^{-1} for both H1 and V.5 to 33 mS m^{-1} for V1. The CV of H.5 was slightly larger compared to the three other ECa data sets. This frequently occurs because the most shallow measuring coil configuration (H.5) has often more noise in arable land due to its bigger sensitivity to artefacts located within the ploughing layer. Since ECa measurements were found in many studies (e.g. Saey et al., 2009) to be most strongly influenced by soil moisture and the concentration of clay related colloids, the larger ECa values of the V1 coil orientation indicate that the soil becomes wetter with an accompanied increase in clay content below the sampling interval of the auger samples.

2.4. Gamma-ray measurements

On 11 June 2009, the gamma ray detector system "The Mole" (The Soil Company, the Netherlands) was used to measure the concentrations of the (semi)naturally occurring ^{40}K, ^{238}U, ^{137}Cs and ^{232}Th radionuclides of the top-soil (Bq kg^{-1}) of the study field. This was done at the same locations where the 30 pooled soil samples were taken. The system consists of a CsI crystal detector and the measured gamma spectra were analysed by the full spectrum analysis method using a chi-squared algorithm to fit a set of standard spectra to the measured spectrum. The details of the spectra processing methodology were described by Van Egmond et al. (2010). The presence of these radionuclides is strongly related with the mineralogy of the soil material (^{40}K, ^{238}U, and ^{232}Th) or with some potentially human induced contamination (^{137}Cs). The summary statistics of the gamma-ray measurements are shown in Table 1. The largest concentration was found for ^{40}K, with much smaller values for ^{238}U, ^{137}Cs and ^{232}Th. In general these values are very low, reflecting the quite uniform mineralogy of these deposits which contain mainly quartz mineral in the sand and silt fractions. ^{40}K is more strongly related with clay minerals, but since the clay content of the topsoil is quite low (on average 8%), the ^{40}K in this soil is also relatively small.

3. Identification of key variables

The co-located elevation and ECa measurements at the soil sampling locations were selected using a search algorithm of 0.25 m radius. In this way a data set of 19 variables at 30 locations was created composed of: elevation, top- and subsoil sand, silt, clay, OC and pH, four ECa measurements and four radionuclides. To identify orthogonal (i.e. independent) combinations of variables (termed factors) we used a principal component analysis (PCA). In this way we could also identify the most dominant variable of each factor, which we considered to be a key variable. To avoid spurious correlation due to the compositional nature of the textural fractions (they sum to 100%) only one fraction was considered. Vitharana et al. (2008) selected the clay fraction because in the Luvisols elluviation and illuviation of clay are dominant pedological processes. However, in Podzols the dominant pedological processes strongly depend on the permeability of the soil matrix. Therefore we considered it more appropriate to select the sand fraction. So, a matrix of 15 variables and 30 locations was the input to a PCA.

The applicability of the data matrix for a PCA was evaluated by the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy (the KMO value should be between 0.5 and 1) and the Bartlett's test of sphericity which checks the significance of the correlation between the variables. We conducted a PCA on the correlation matrix and the selection of the retained factors was based on the explained variance by each principal component (PC or factor) and a plot of their eigenvalues (a screeplot). To optimise the interpretation of the retained factors a varimax rotation was applied. From each of the retained factors, a key variable was identified based on the largest factor loading after checking the communality for the retained factors. This multivariate analysis was performed with SPSS v.19 (SPSS Inc., USA).

For our 30×15 matrix the KMO value was 0.71 and the correlation matrix was found to be significantly different from an identity matrix by the Bartlett's test ($\chi^2 = 739.8$, $p < 0.05$). The first factor (PC1) explained 30.1% of the total variance, the second (PC2) explained 27.1% and the third (PC3) explained 15.3%. So together these three components explained 72.5% of the total variance within the data. It was decided to retain only these three factors.

The communalities and loadings of the variables of the three rotated PCs are given in Table 2. Most of the variables had high communalities, only for subsoil pH (pH_s), ^{137}Cs and ^{232}Th this value was below 0.5. So their contribution to the first three PCs was limited. The variance of some variables, like the four ECa measurements, was almost entirely accounted by the first three PCs (communalities > 0.9).

The loading plots of the first and second and the first and third PCs are shown in Fig. 2. The first PC was dominantly associated with the four ECa measurements, of which ECa-H.5 had the largest loading (0.97). The ECas measured with the other coil configurations were strongly similar, with ECa-V1 having the lowest loading (0.90). None of the other 11 variables was strongly related to PC1 (topsoil sand Sa_t had the largest absolute loading of −0.36). Elevation (DEM) is the strongest contributor to the second PC with a loading of 0.85. However, also both top- and subsoil sand (Sa_t and Sa_s) and OC (OC_t and OC_s) showed a strong association with PC2 although in an inverse relationship with elevation. The loadings of

Table 2
Principal component loadings of the rotated first three principal components and the communalities of each variable (of the variables with a communality exceeding 0.5 the largest absolute loading is in bold).

Variables	Communality of first 3 PCs	Principal component (PC) loadings		
		PC1 (30.1%)	PC2 (27.1%)	PC3 (15.3%)
Topography				
Elevation (DEM)	0.83	−0.05	**0.85**	−0.32
Soil sample properties				
Topsoil				
Sand	0.76	−0.36	**−0.74**	−0.29
OC	0.78	−0.28	**−0.82**	−0.16
pH-KCl	0.72	0.09	0.06	**0.85**
Subsoil				
Sand	0.77	−0.33	**−0.78**	−0.23
OC	0.78	−0.28	**−0.82**	−0.16
pH-KCl	0.36	−0.14	−0.01	0.58
EMI measurements				
ECa-H.5	0.95	**0.97**	0.05	0.01
ECa-H1	0.97	**0.95**	0.27	0.08
ECa-V.5	0.97	**0.94**	0.28	0.07
ECa-V1	0.92	**0.90**	0.31	0.12
Gamma-ray measurements				
^{40}K	0.69	0.26	**0.74**	0.27
^{238}U	0.52	0.19	0.26	**0.64**
^{137}Cs	0.39	−0.14	−0.11	−0.67
^{232}Th	0.48	0.02	0.63	0.02

these variables on PC2 were very similar (between −0.74 and −0.82). Moreover, ^{40}K had a large loading on PC2: 0.74, confirming its strong association with clay (being often in an opposite position of sand). The third PC represented mainly topsoil pH (pH_t) with the largest loading on PC3 of 0.85. Also subsoil pH (pH_s) (0.58) and ^{238}U (−0.67) were associated with PC3.

The result of the PCA identified three key variables: ECa-H.5, elevation and pH_t. Surprisingly, these are the same variables as identified by Vitharana et al. (2008) for the loess area.

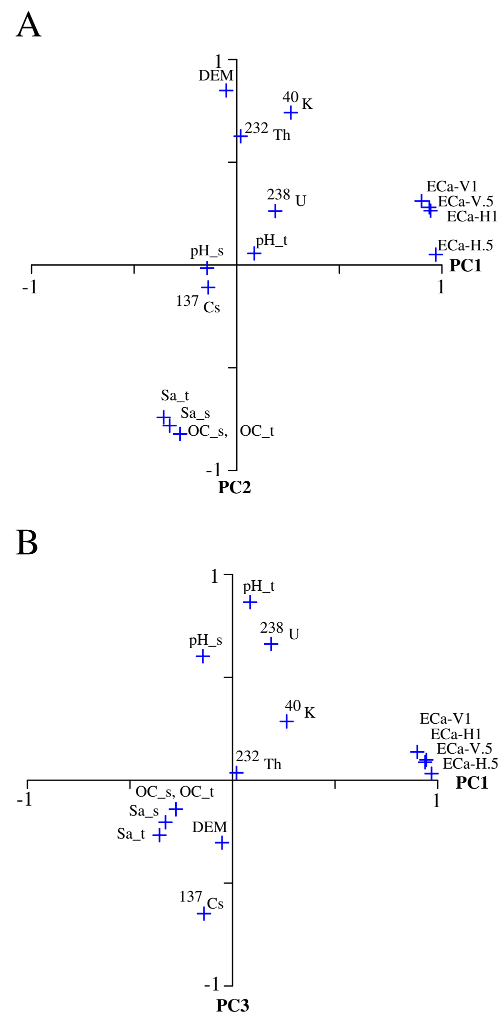

Fig. 2. Rotated loading plots of the first and second principal components (A) and the first and third principal components (B). Symbol definitions are given in the text.

4. Maps of key variables

Each of the three key variables, ECa-H.5, elevation and pH_t, was interpolated with ordinary kriging (Goovaerts, 1997) to create a map with a resolution of 1 m by 1 m. For this purpose, the topsoil of 70 additional locations was sampled and analysed for pH resulting in 100 measured locations. All three variograms were best modelled by an omnidirectional spherical model (not shown). The kriged maps of ECa-H.5, DEM and pH_t are given in Fig. 3.

Fig. 3A reveals several distinct patterns in the spatial behaviour of the first key variable, i.e. ECa-H.5. In general the ECa values are larger in the northern part and lower in the southern part of the field. This general pattern is crossed by three distinct linear features. A rather wide (about 10–15 m) line of higher ECa values crosses the field more or less halfway in an east–west orientation. This feature is an extension of a former field track still present to the west of the field but no longer running through the study field (compare with Fig. 1). The farmer declared that some 40 years ago this track was removed (by his father) and since then both parts were merged into one field. The other two linear features with elevated ECa values are smaller in width. One runs more or less parallel with the former track halfway in the northern part of the field, the other runs almost perpendicular to the former track in the southern part. Both represent former ditches which were also removed by the farmer's father. So until about 40 years ago this field was actually split into four smaller fields each covering an area of approximately 1 ha and managed differently. Note also that along most borders deviating ECa values were measured, potentially representing compaction due to agricultural traffic. It is surprising that although the ECa-H.5 has the shallowest DOI it still records these subsoil features, illustrating the relative nature of the DOI parameter. It is our experience that such former anthropogenic influences are commonly found in this old agricultural area. The DEM in Fig. 3B shows that the field is the highest in the northern part with the elevation decreasing gradually towards the southern side. The former track and ditches are still visible as slight differences onto the micro-relief (which is hardly visible on the field). Within the southern part the western half is lower than the eastern. The southern and western borders of the field have the lowest elevation because they are close to open ditches. The major drainage channel of the area (the “Centerloop” in Fig. 1) runs along the southern border. The interpolated pH_t (Fig. 3C) shows a more patchy pattern with values fluctuating around the mean of 5.3. In general the southern part is slightly more acidic, while the northern part is less acidic.

5. Management classes

We used a fuzzy *k*-means classification for the three key variables to create distinctly different classes. These can be considered to represent stable management zones. However, in contrast to Vitharana et al. (2008) we preferred to use a fuzzy *k*-means with extragrade

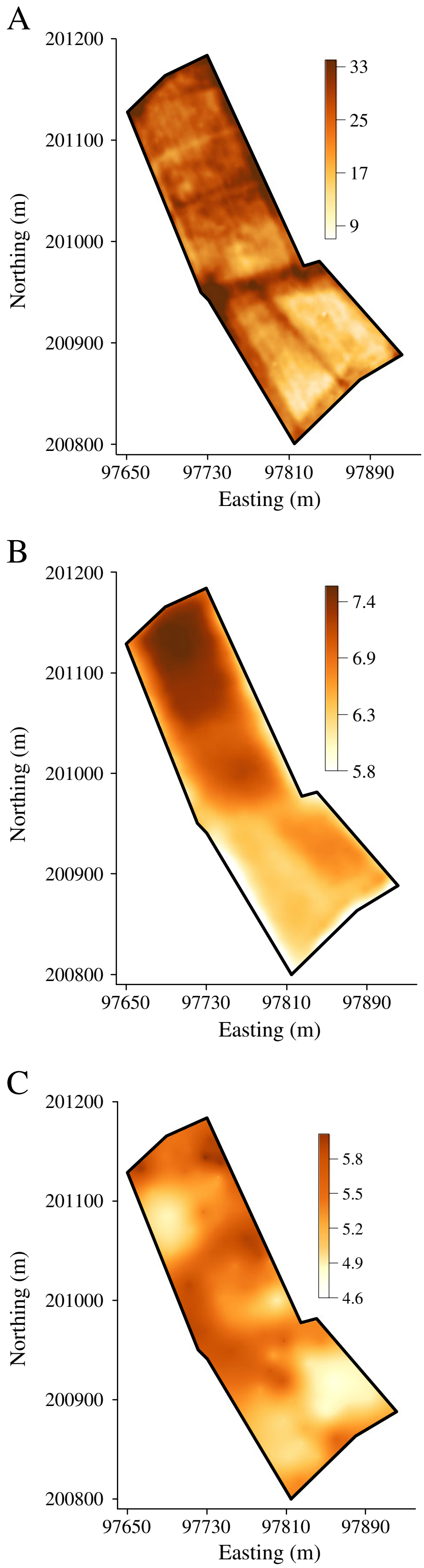

Fig. 3. Kriged maps of the three key variables: ECa-H.5 (mS m^{-1}) (A), elevation (DEM in m) (B) and topsoil pH-KCl (C). Coordinates in all maps are according to the Belgian Lambert72 projection.

classification procedure (de Gruijter and McBratney, 1988) which is a modification of the fuzzy *k*-means classification (Bezdek, 1981). The fuzzy *k*-means classification produces a continuous grouping of objects by assigning partial class membership values, which is to be preferred for grouping properties in the soil continuum. The procedure with extragrades recognises the objects that might not fit well in any of the classes formed (containing the 'intragrades') and places those in an additional outlier group, the 'extragrades'. To allow the discrimination between 'extragrades' and 'intragrades' the 'fuzzy k-means with extragrades' algorithm requires an extra parameter to be chosen: the extragrade exponent α. To obtain memberships that were neither too fuzzy nor too hard, the fuzziness exponent was fixed to the conventional value of 1.35 (Odeh et al, 1992) and to obtain an average extragrade membership α was taken to be very small (2×10^{-7}) (McBratney and de Gruijter, 1992). The fuzziness performance index (FPI) and the modified partition entropy (MPE) were used to guide the classification (McBratney and Moore, 1985). FPI estimated the degree of fuzziness generated by a specified number of classes. MPE estimated the degree of disorganisation created by a specified number of classes. The optimum number of classes is determined when these two measures are minimal. We used the FuzMe 3.0 software (Minasny and McBratney, 2006).

The optimal number of classes was found to be two with an extragrade class. By including the extragrade class, the influence of outliers on the classification was reduced which resulted in more compact and more stable classes. The results of this classification were generalised by removing a few small island areas (the largest covering 12 m^2) to obtain more contiguous zones. The result is shown in Fig. 4A. The two classes formed two continuous zones of about the same area in the northern and southern parts of the field. The extragrade class formed an intermediate zone between the two classes and some elongated zones along the western border of the field. From a comparison with the maps of Fig. 3, it can be concluded that class 1 groups the areas with the largest ECa-H.5, the highest elevation (and thus the smallest sand and OC content) and the highest pH_t. The reverse holds true for class 2. The extragrade class groups intermediate values but includes also some extremes like the band with higher ECa-H.5 values along the western border of the southern part and some areas with low pH in the northern part.

6. Wheat yield

In the autumn of 2005 winter wheat was sown on the study field and on 29 July 2006 it was harvested with a commercial combine (CNH New Holland CX880, Italy) equipped with a DGPS and a yield monitor. The growing conditions were close to normal for most of the winter and spring seasons, but during the ripening period (June–July) weather conditions were exceptionally dry and hot. Therefore the crop matured earlier than usual which implied that grain was harvested earlier than in most other years. Post-processing of the grain yield data included standardisation to 15% moisture content and the filtering of about 3% erroneous data (like corrections for the time delay in grain flow at the start or end of driving lines, spatially duplicated records, etc.). The retained data were block kriged at a resolution of 5 by 5 m (Fig. 4B). The average yield of the field was 9.2 t ha^{-1}, but in general yields were higher (9–10 t ha^{-1}) in the northern part of the field and lower (7.5–8.5 t ha^{-1}) in the southern part. However, the visual similarity between the yield map and the delineated management classes is not straightforward. Therefore, the 6762 yield data were grouped according to the two management classes and the extragrade class and Table 3 shows the result. As could be expected from Fig. 4, class 1 produced the highest average yield (9.9 t ha^{-1}), while class 2 produced on average 1.3 t ha^{-1} less. The extragrade class took an intermediate position with 8.8 t ha^{-1}. The variation around these mean values was similar for classes 1 and 2 (around 0.55 t ha^{-1}) but it was clearly larger for the extragrade class (0.71 t ha^{-1}). This indicates the effect of an

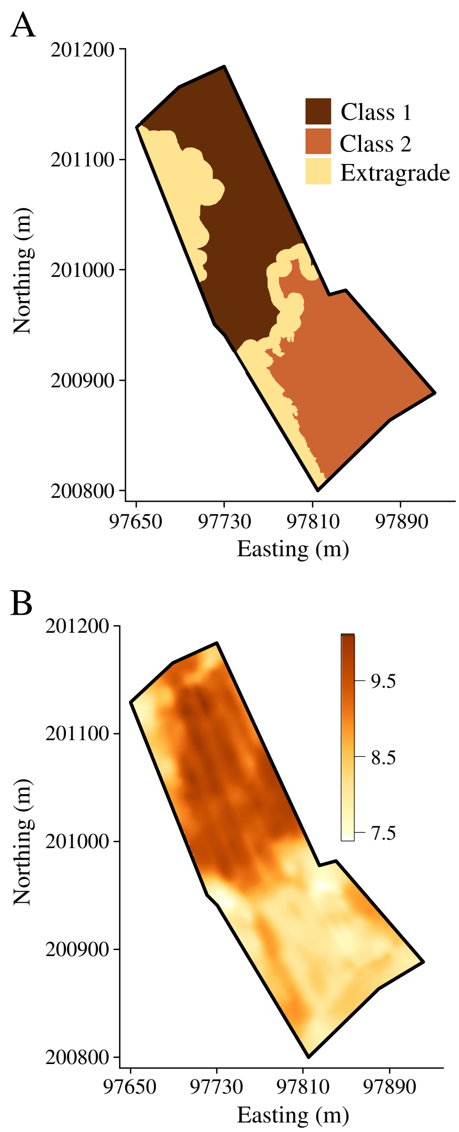

Fig. 4. Management classes obtained through a fuzzy *k*-means with extragrade classification of the three key variables (A) and kriged wheat yield (t ha^{-1}) (B).

extragrade class: the creation of rather homogeneous classes by grouping outliers in a heterogeneous extragrade.

7. Relationship between crop yield and key variables

The ability of the key variables to predict wheat grain yield for 2006 was evaluated by a stepwise multivariate regression analysis. Measurements falling within the extragrade class were excluded from this analysis to avoid instabilities in the modelling. This resulted in the model:

$$\mathrm{Grain}_{2006} = -0.324 + 0.175 * (\mathrm{ECa} - \mathrm{H.5}) + 1.009 * \mathrm{DEM} - 0.00217 * (\mathrm{ECa} - \mathrm{H.5})^2 \quad (2)$$

Table 3
Mean grain yield for 2006 per management class and its standard deviation (t ha^{-1}).

Management class	*n*	Mean	Standard deviation
Class 1	3646	9.9	0.54
Class 2	2353	8.6	0.57
Extragrade	763	8.8	0.71

with $R^2_{adj} = 0.88$, $p < 0.001$ and Grain_{2006} in t ha^{-1}. The high R^2_{adj} indicates the very strong correlation between grain yield and both ECa-H.5 and elevation, but ECa-H.5 clearly was the most influencing variable. Topsoil pH did not significantly influence crop yield in this growing year.

Given the strong influence of ECa-H.5 on crop yield, we analysed its impact through a boundary line analysis (Kitchen et al., 1999). We used the procedure described by Shatar and McBratney (2004) in which the 10% highest yield data are selected by splitting the cloud in bins. The idea is that within each bin the highest yield data have ECa (or combined soil properties) as the most limiting factor, contrary to lower yields which have other limiting factor(s) (e.g. diseases). By concentrating on the highest yield data, the relationship between yield and ECa was modelled by a quadratic curve (Fig. 5):

$$\mathrm{Grain}_{2006} = 6.537 + 0.232 * (\mathrm{ECa} - \mathrm{H.5}) - 0.00331 * (\mathrm{ECa} - \mathrm{H.5})^2 \quad (3)$$

with $R^2_{adj} = 0.98$ and Grain_{2006} in t ha^{-1}.

Model (3) and Fig. 5 indicate that in 2006 the highest grain yields were very strongly related to ECa. Especially between 10 and 25 mS m^{-1}, yield increased strongly. Beyond 25 mS m^{-1} grain yield levelled off around 10.5 t ha^{-1}. Therefore, crop productivity for the 10% highest yield was strongly driven by variations in soil ECa.

8. Conclusions

Based on a PCA of the correlation matrix from 15 soil and topographic variables, three variables were identified as key variables for this field located in the sandy belt of northern Europe: ECa-H.5, elevation and pH_t. Surprisingly, these are the same three variables identified by Vitharana et al. (2008) for a field in the European loess belt, although not all variables that these authors considered are the same as in this study and the topographic and pedogenetic processes are strongly different.

The identified key properties allowed defining two stable management classes using a fuzzy *k*-means with extragrade procedure. The extragrade was found useful to isolate deviating points and extreme values, resulting in the creation of classes with less variability.

The created management classes were spatially continuous and therefore suitable for practical management by the farmer. They differed distinctly in soil and topographic properties. A classification of crop yield (wheat grain of the year 2006) according to the two management classes confirmed the significance of the classification. Crop yield was distinctly higher in one class compared to the other. The extragrade class contained on average an intermediate wheat yield which was, however, more variable.

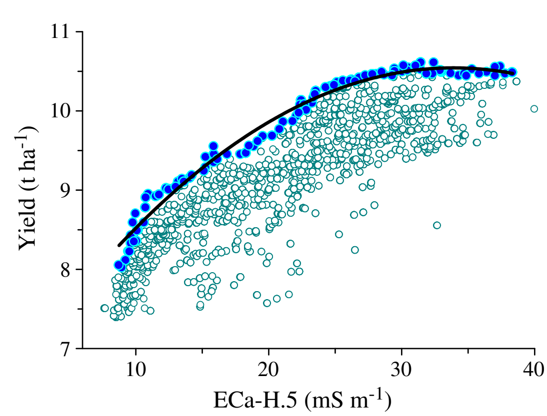

Fig. 5. Relationship between ECa-H.5 and wheat yield. Filled circles represent the points selected by the boundary line procedure described in the text and were used to fit the curve (representing the boundary line). Empty circles represent points which can be considered to have another yield limiting factor than ECa-H.5.

The relationship between crop performance and the key properties was modelled by a stepwise multiple regression and surprisingly pH_t did not seem to be significantly influencing crop production for the year 2006. However, ECa-H.5 was very strongly related to crop yield. A boundary analysis confirmed this finding and an excellent ($R^2_{adj} = 0.98$) quadratic model could be fit to the upper yield data across the range of ECa-H.5 values.

When the study of Vitharana et al. (2008) and this study are combined, it appears that topsoil ECa, elevation and pH are strong potential candidates as generic key properties to delineate management classes in the north-western European aeolian landscapes. ECa and elevation are often related to physical and biological soil properties like texture, organic matter content and moisture availability, while pH is an integral measure for the soil chemical status. In this context, these three variables are strongly complimentary. With current proximal soil sensing technology these three variables can be recorded in a spatially detailed and continuous way, e.g. ECa with EMI or electrical resistivity, elevation by LIDAR or organic matter by near-infrared field spectroscopy (Stenberg et al., 2010) and pH by on-the-go electrochemical sensors (Adamchuk et al., 2007).

Given the many sources which report that proximally sensed ECa is an excellent integrator of the bulk soil status (e.g. Corwin and Lesch, 2005; Kitchen et al., 1999; Saey et al., 2009; Sudduth et al., 1995), the authors propose to upgrade ECa from a secondary proxy for other soil properties, to a primary soil variable. In general, we envisage a future for several types of proximally sensed information as a basis for regional 'multiple purpose' soil maps.

Acknowledgements

E. Loonstra (The Soil Company, the Netherlands) is thanked for conducting the gamma ray measurements. C. De Cuyper is thanked for granting repeated access to his field and his kind cooperation as a farmer.

References

Adamchuk, V.I., Lund, E.D., Reed, T.M., Ferguson, R.B., 2007. Evaluation of an on-the-go technology for soil pH mapping. Precision Agriculture 8, 139–149.
Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Premium Press, New York.
Corwin, D.L., Lesch, S.M., 2005. Apparent soil electrical conductivity measurements in agriculture. Computers and Electronics in Agriculture 46, 11–43.
de Gruijter, J.J., McBratney, A.B., 1988. A modified fuzzy k-means method for predictive classification. In: Bock, H.H. (Ed.), Classification and Related Methods of Data Analysis. Elsevier, Amsterdam, pp. 97–104.
Derese, C., Vandenberghe, D., Eggermont, N., Bastiaens, J., Annaert, R., Van den haute, P., 2009. A medieval settlement caught in the sand: optical dating of sand-drifting at Pulle (N Belgium). Quaternary Geochronology 5, 336–341.
Doolittle, J.A., Sudduth, K.A., Kitchen, N.R., Indorante, S.J., 1994. Estimating depths to claypans using electromagnetic induction methods. Journal of Soil Water Conservation 49, 572–575.
European Soil Bureau Network, 2005. Soil Atlas of Europe. Luxembourg.
Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.
Haase, G., Haase, D., Ruske, R., Jäger, K.D., Alterman, M., 2007. Loess in Europe — spatial distribution in a scale 1:2500000. Quaternary Science Review 26, 1301–1312.
Jaynes, D.B., Colvin, T.S., Ambuel, J., 1993. Soil Type and Crop Yield Determinations from Ground Conductivity Surveys. Paper 933552. American Society of Agricultural Engineers, St. Joseph, MI.
Jungerius, P.D., Riksen, M.J.P.M., 2010. Contribution of laser altimetry images to the geomorphology of the Late Holocene inland drift sands of the European Sand Belt. Baltica 23, 59–70.
Kitchen, N., Sudduth, K., Drummond, S., 1999. Soil electrical conductivity as a crop productivity measure for claypan soils. Journal of Production Agriculture 12, 607–617.
Larson, W.E., Robert, P.C., 1991. Farming by soil. In: Lal, R., Pierce, F.J. (Eds.), Soil Management for Sustainability. Soil Water Conservation Society, Ankeny, IA, pp. 103–112.
McBratney, A.B., de Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science 43, 159–175.
McBratney, A.B., Moore, A.W., 1985. Application of fuzzy sets to climate classification. Agricultural and Forest Meteorology 35, 165–185.
McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97, 293–327.
McNeill, J.D., 1980. Electromagnetic terrain conductivity measurement at low induction numbers. Technical Note TN-6. Geonics Limited, Missisauga, Ontario, Canada.
Minasny, B., McBratney, A.B., 2006. FuzME version 3. Australian Centre for Precision Agriculture. The University of Sidney, Australia.
Moral, F.J., Terron, J.M., Marques da Silva, J.R., 2010. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. Soil and Tillage Research 106, 335–343.
Mulla, D.J., 1991. Using geostatistics and GIS to manage spatial patterns in soil fertility. In: Kranzler, G. (Ed.), Proceedings of Symposium on Automated Agriculture for the 21st Century. American Society of Agricultural Engineers, St. Joseph, MI, pp. 336–345.
Mulla, D.J., Schepers, J.S., 1997. Key processes and properties for site-specific soil and crop management. In: Sadler, E.J. (Ed.), The State of Site-specific Management for Agriculture. ASA/CSSA/SSSA, Madison, WI, pp. 1–18.
OC-GIS Vlaanderen, 2003. Digitaal hoogtemodel vlaanderen. Nieuwsbrief GIS-Vlaanderen, 16. VLM, Brussel (in Dutch).
Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1992. Soil pattern recognition with fuzzy-c-means: application to classification and soil–landform interrelationships. Soil Science Society of America Journal 56, 505–516.
Pedroso, M., Taylor, J., Tisseyre, B., Charnomordic, B., Guillaume, S., 2010. A segmentation algorithm for the delineation of agricultural management zones. Computers and Electronics in Agriculture 70, 199–208.
Saey, T., Van Meirvenne, M., Vermeersch, H., Ameloot, N., Cockx, L., 2009. A pedotransfer function to evaluate soil profile heterogeneity using proximally sensed apparent electrical conductivity. Geoderma 150, 389–395.
Schellberg, J., Hill, M.J., Gerhards, R., Rothmund, M., Braun, M., 2008. Precision agriculture on grassland: applications, perspective and constraint. European Journal of Agronomy 29, 59–71.
Shatar, T.M., McBratney, A.B., 2004. Boundary-line analysis of field-scale yield response to soil properties. The Journal of Agricultural Science 142, 553–560.
Sheets, K.R., Hendrickx, J.M.H., 1995. Non-invasive soil water content measurement using electromagnetic induction. Water Resource Research 31, 2401–2409.
Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. Advances in Agronomy 107, 163–215.
Sudduth, K.A., Kitchen, N.R., Hughes, D.F., Drummond, S.T., 1995. Electromagnetic induction sensing as an indicator of productivity on claypan soils. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Proceedings of the Second International Conference on Site-specific Management for Agricultural Systems, Minneapolis, MN, March 27–30 1994. ASA/CSSA/SSSA, Madison, WI, pp. 671–681.
Van Egmond, F., Loonstra, E.H., Limburg, J., 2010. Gamma ray sensor for topsoil mapping: the Mole. In: Viscarra Rossel, R.A., McBratney, A.B., Minasny, B. (Eds.), Proximal Soil Sensing. Springer, Dordrecht, pp. 323–332.
Van Meirvenne, M., 2006. Soil inventory in transition: from too few to too many geo-data? In: Hartemink, A. (Ed.), The Future of Soil Science. IUSS, Wageningen, pp. 142–144.
Vitharana, U.W.A., Van Meirvenne, M., Simpson, D., Cockx, L., De Baerdemaeker, J., 2008. Key soil and topographic properties to delineate potential management classes for precision agriculture in the European loess area. Geoderma 143, 206–215.
Walkley, A., Black, I.A., 1934. An examination of Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. Soil Science 37, 29–37.