

A Decision Support System for Classification and Representation of Fuzzy Soil Boundaries

Mette Balslev Greve and Mogens Humlekrog Greve

Abstract

We present a way of classifying/representing map unit transition zones. High-resolution data (EM38, DEM, aerial photos) are used for classification of soil map unit boundaries. Boundaries are repositioned and the width of the transition zone determined. Multiple grids expressing the changing transition zone width are a flexible tool for modeling spatial distribution of soil types and -attributes.

Keywords: Expert system, decision support, soil unit boundaries, EM38, DEM, aerial photos, soil database, soil survey, fuzzy boundaries

Introduction

On choropleth soil maps all soil variation is sited exactly on the map unit delineation, which in the real world is not the case. The fuzziness of soil map delineation is not normally taken into account - except perhaps for general statements in the survey report - and no procedure for classifying or coding map unit delineation is described in soil survey manuals.

The aim of this study is to develop a GIS-based decision support system for fuzzyfication/smoothing of soil maps using a range of auxiliary data. The output is a data structure showing soil types/map units and the transition zones between them.

We illustrate the process using the following data sources:

Aerial photographs, AP: In this survey, panchromatic B/W aerial photos have been used in a scale of 1:10.000. For post-processing purposes, the aerial photos were scanned in high resolution and rectified.

Digital Elevation Model, DEM: the DEM was generated on the basis of an airborne laser scanning. The scanner system records the distance in a point for every 6m². The points are then interpolated to a DEM in a 2x2 meter grid with a relative accuracy in elevation of approximately 10 cm.

Soil Electrical Conductivity, SEC: A survey was performed using the non-contact conductivity meter EM38. The use of a map of EM38 data during survey operations enables the surveyor to delineate map units manually with very high accuracy, because of the very high correlation between the EM38 map and clay content.

Soil database: A soil database with information on modal soil horizons was used in this study. The database consists of a soil series table and a horizon table. The horizons are defined on the basis of a set of diagnostic properties that are combined into soil series. A soil series has three layers with fixed depths: 0-30 cm, 30-70 cm and 70-120 cm. The diagnostic properties of the system are: geology, drainage, presence of CaCO₃, simplified pedology and soil texture. Soil series are distinguished from one another by differences in one or more of the diagnostic properties in one of

the three layers. The diagnostic properties are equally important. The properties of the diagnostic horizons were determined on the basis of a statistical analysis of the nationwide soil profile database.

The Method

The mapping area was subdivided into areas with uniform soil conditions and as the DEM, SEC and AP were interpreted and strategic augerings were made, the information on soil type was registered in a field computer. The delineation was drawn directly on the computer screen. The area was subdivided into 664 mapping units.

For the development of the method in question, we selected a small area (the test area) in the south part of the research area (Fig. 1). The strip was selected in an area with good relation between the landscape and the soil.

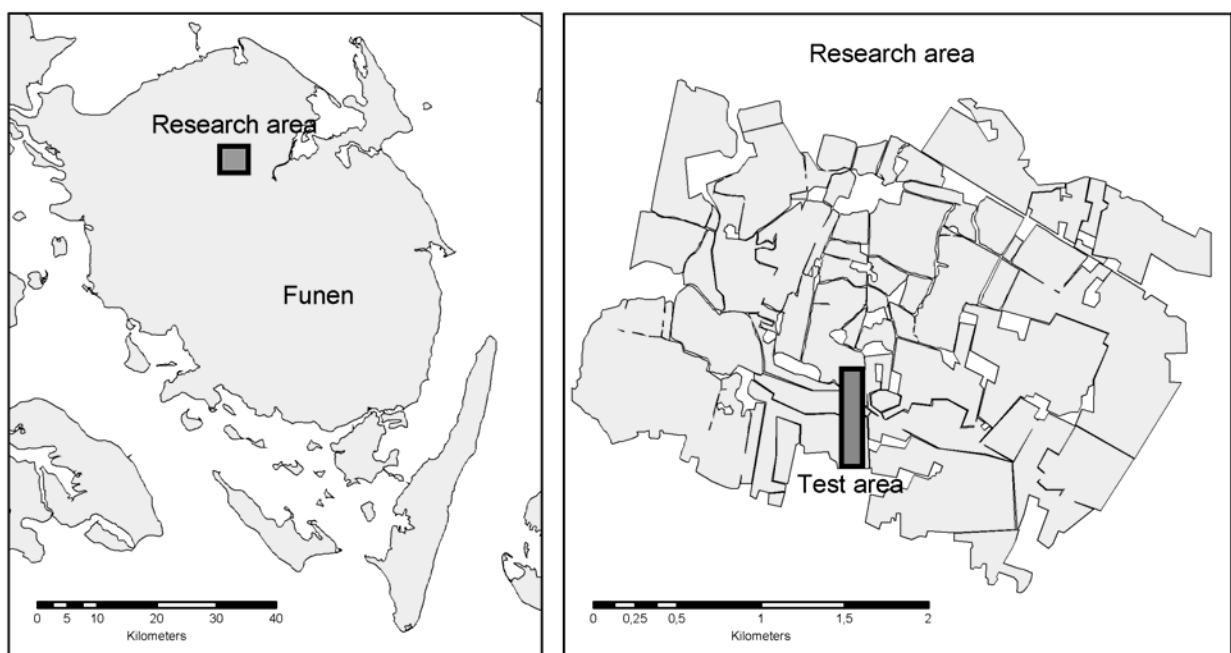


Fig. 1, Research and test area

Fig. 2 shows all the available data in the test area in the small strips. The large strip contains the map units. In the central part of the area the gley- and histosols are clearly visible on the AP and the profile curvature, but not on the SEC. The SEC and the slope of SEC seem very useful for describing the delineations between the sandy and clay soils.

The auxiliary data and the polygons from the soil map were clipped to map the test area. All raster data was resampled to 2*2 m pixels, and slope maps in the same resolution were calculated for all three types of auxiliary data. For the DEM, we also calculated the 2. derivative to express the profile curvature. For each map unit the information on drainage, topsoil TOC (total organic carbon) and texture were simplified into a small number of classes illustrated on Fig. 3. The map unit boundaries were converted to a dataset containing only the boundary lines.

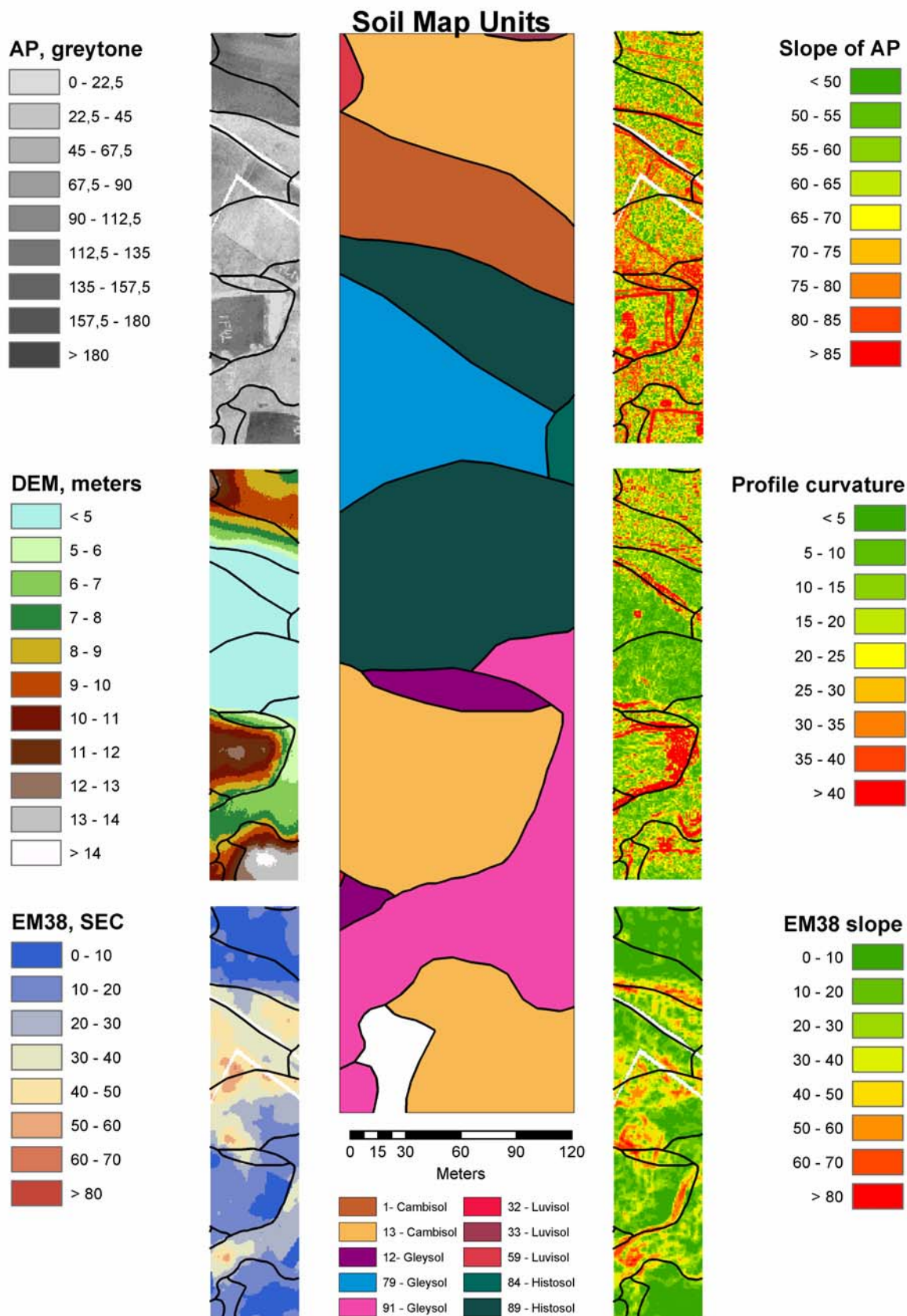


Fig. 2, The different datasources in the test area

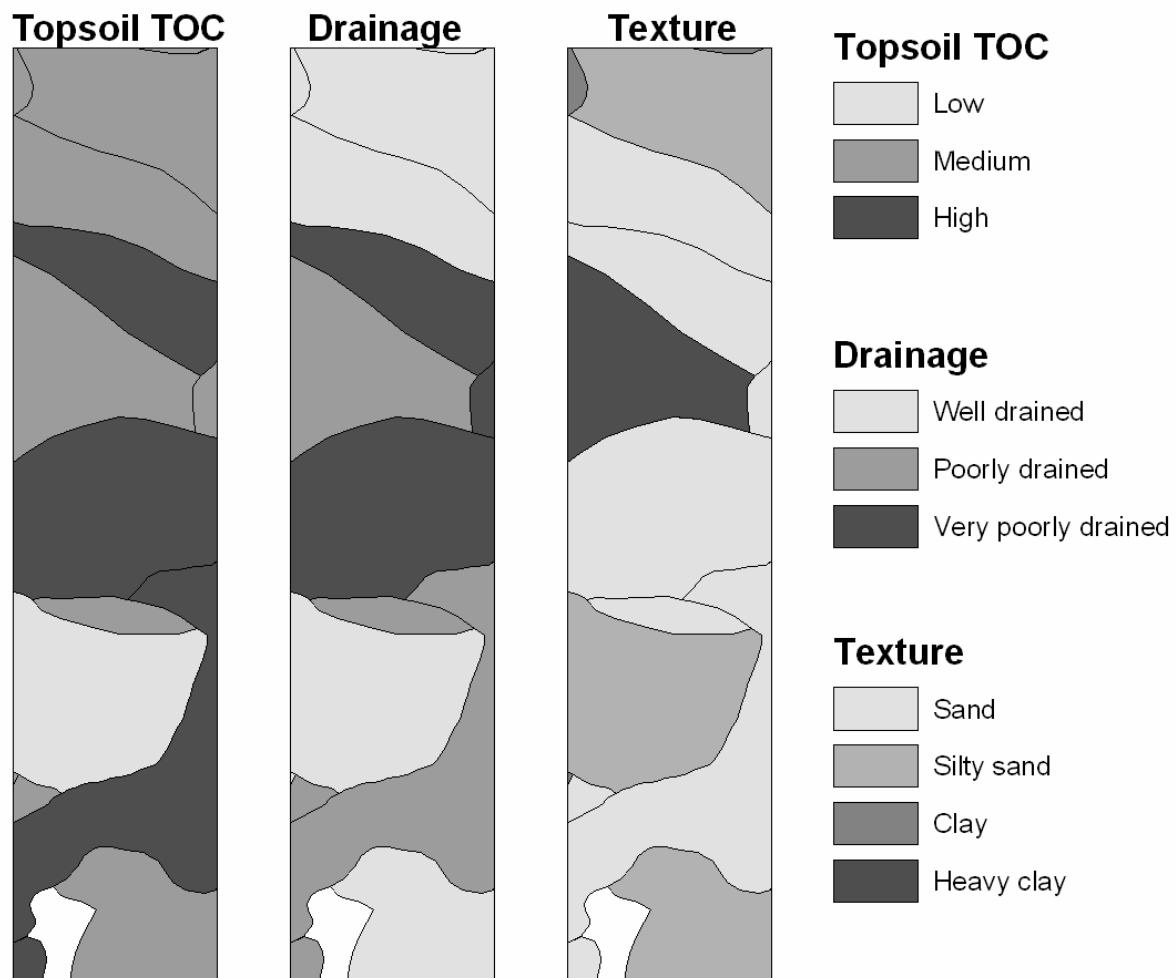


Fig. 3, The diagnostic properties in the map units

During the survey, high correlation between the aerial photography and topsoil TOC was found. Topography is known to have high correlation to drainage condition and as stated above, SEC has a high correlation with texture. By calculating the slope in high-resolution raster data, this is an indication of the spatial rate of change in the data, or, in other words, a map of how fast the values in the high-resolution data changes. Therefore, the auxiliary data can be used as indicators of how wide a transition zone between two soil types is. The data we used in this study were:

- Aerial photograph: Slope - indicates changes in topsoil TOC
- Digital elevation model: Profile curvature – indicates changes in soil drainage
- Soil electrical conductivity: Slope – indicates changes in soil texture

We have registered the topsoil TOC, the drainage condition and the texture for each of the map units. We then need to find out on what parameters the adjacent polygons differ, to determine which data set/indicator grid should be used for calculating the boundary width. Each line is assigned an indicator grid as sketched below.

In this example we set up priorities between the auxiliary data, so that difference in topsoil TOC take priority over drainage condition, which then take priority over texture. This means, that if two adjacent polygons differ on both topsoil TOC and drainage, only the aerial photograph will be used to calculate the boundary width, as topsoil TOC take priority over drainage. The user of the system can change the priorities.

The simple calculation is then, that if the polygons differ on topsoil TOC, the indicator for the line between the polygons will be set to the aerial photograph. If the polygons have the same topsoil TOC, but differ in drainage class, the indicator will be set to the profile curvature. If the polygons only differ on the texture, the slope of the SEC-grid will be used for calculating the boundary width.

Moving the lines

The use of the auxiliary data for the initial compilation of the soil map during the soil survey is crucial for the method, as the delineations between mapping units has to be located at the place with the highest spatial rate of change to make use of this criteria in creating the transition zone. Very often the delineation is still a little off, and may be located just beside the high slope. We created an algorithm for moving the lines, so that they would be placed over the steepest cells in the indicator grid, where the rate of change in soil properties is the highest.

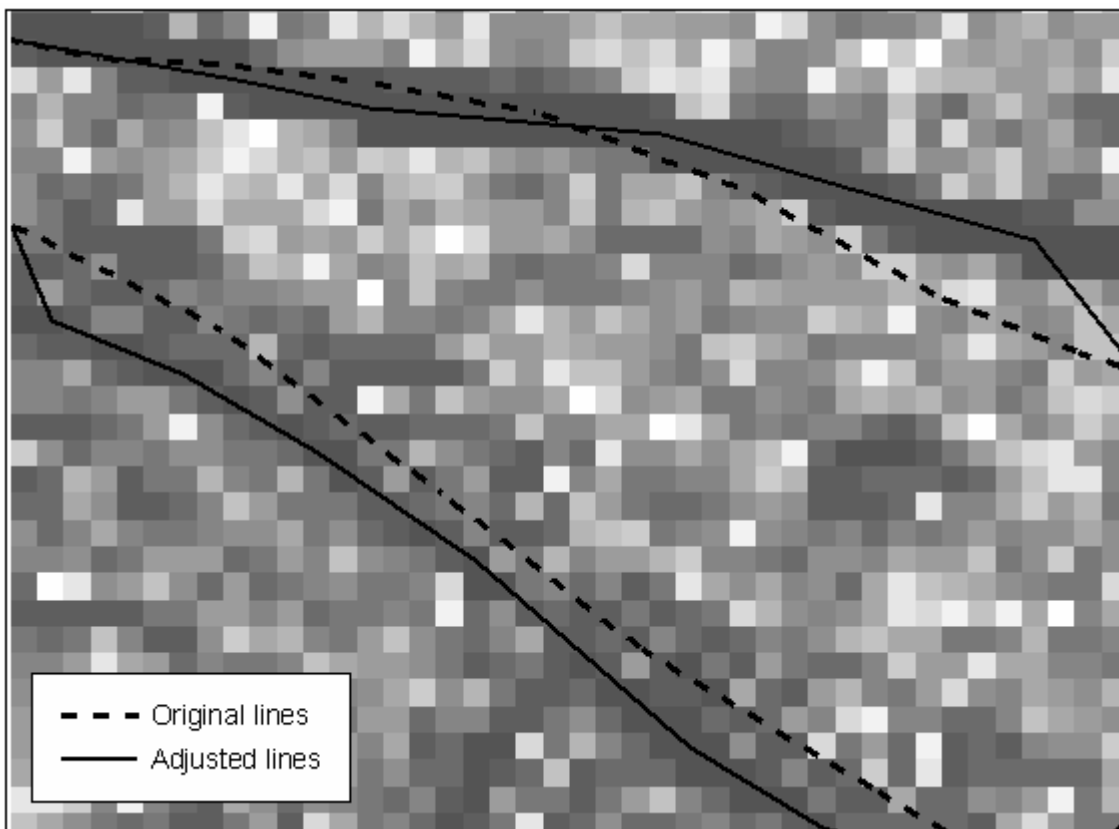


Fig. 4, The map unit boundaries are repositioned

We had to select a maximum distance, that each line should be moved. In our case the distance could be small, as the map unit boundaries had been digitized knowing all the indicator grids. The distance was set to ten meters. We decided not just to move the line, but also alter the shape. This would be done by moving each vertex along the line to the cell in the indicator grid with the highest

value in a ten-meter buffer zone around the vertex. The indicator grids were resampled to 3*3 m pixels for this purpose to smooth out high single-cell values.

We wanted the vertices of each line to be evenly spaced. If the vertices were to close, they would move to the same cell anyhow, and if they were to far apart, we might miss the steep cells on the long stretch. To even-space the vertices we densified the lines, meaning, that extra vertices were added, so that there would be no more than five meters between two vertices. Then we simplified the lines, only keeping vertices that were at least 20 meters apart (twice the maximum distance, that the vertex could move).

A new line shape file containing the adjusted lines was the output of this step, Fig. 4.

Calculating the boundary width

After the lines have been moved/adjusted, the slope values from the indicator grids are extracted in the cells that the lines run through. The values are then converted to a boundary width.

Low slope values result in broad transition zones and high slope values in narrow transition zones between the map units. The slope readings in the boundaries are translated to a transition zone width using a linear function assessed by the survey expert.

The output here is a map showing the transition zone widths (as numbers) in each line grid cell.

For each of the modal soil types in the map unit theme, a layer in the MultiGrid structure is compiled. The layers are obtained in this manner: Each of the soil mapping regions from the choropleth map is converted to a grid (a region can contain more than one polygon, but the polygons have the same modal soil). The delineations of the regions are buffered according to the map of transition zone width; giving a varied size buffer around the regions.

Now, the actual compilation of the layer in the MultiGrid takes place.

1. Cells, which fall inside the region, but outside the buffer, are given the value 1.
2. Cells, which fall both outside the region and outside the buffer, are given the value 0.
3. Cells, which fall inside the region and inside the buffer, are given gradual values between 1 and 0.5 with 0.5 directly on the 'old' boundary.
4. Cells, which fall outside the region, but inside the buffer, are given gradual values between 0.5 and 0.

The sum of the layers (one for each soil mapping region) is calculated and used for normalizing the MultiGrid for the layers to sum to 1.

The MultiGrid can now be used for spatial modeling of soil parameters, see Fig.5. On Fig. 5.1 the importance of the indicator grid is very well illustrated, as the southern delineation is very abrupt and the northern delineation very fuzzy. Compare this to Fig. 2 – the southern delineation is visible on all the indicator grids, while the northern delineation cannot be identified on any of the indicator grids.

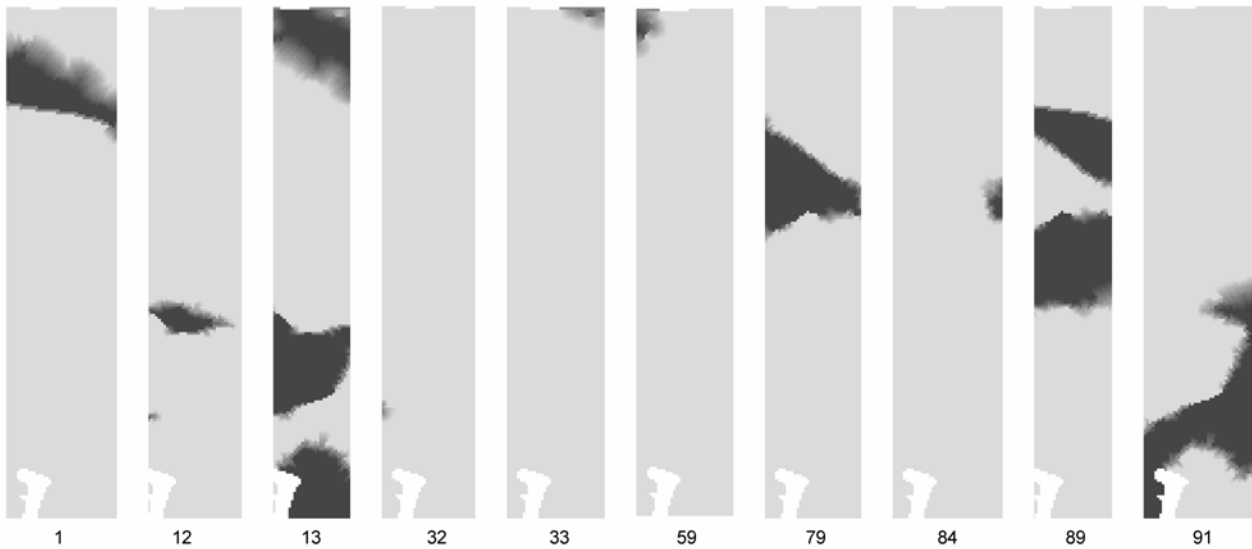


Fig. 5, The layers in the MultiGrid

The steps described in this section were assembled to form a decision support system for detailed (D1 or Order 0) soil mapping. The system was developed for ArcView 3.2.

The user interface for the system, Fig. 6, allow the user to specify up to three priority indicator grids using drop down lists of all the grids in the active view. A corresponding field in the map unit theme must be selected for all indicator grids, e.g. for a grid showing profile curvature (derived from the DEM), a field in the map unit theme must exist, that carry information about a drainage class.

Fig. 6, User interface for the system

The user also inputs the field that the soil categories are listed in, the maximum distance to move the lines, and the minimum and maximum soil boundary width in meters.

Conclusion

The aim of this study was to develop a GIS-based decision support system for fuzzyfication of existing soil maps using auxiliary data. In previous work we have used the spatial rate of change of the EM38 surface to calculate map unit fuzziness. One of the problems with the method presented in that paper was the one-dimensionality of the EM38 sensor and the multi-dimensionality of the soil. This problem is suggested solved by linking the soil database to the various ancillary dataset through a decision support system. The system analyzes the soil database and look for differences in soil across the map unit delineation. From our soil database it was possible to derive information on differences in: total organic carbon (TOC), soil texture and soil drainage.

The decision on which of the auxiliary data grids should be used for the calculation in each case has to be taken - in our case we used topsoil color from the AP when there was a difference in top soil TOC, we used SEC in the case of differences in soil texture, and DEM when there was difference in soil drainage. When soils differ in more than one of the soil parameters, the relative importance has to be decided. We used the following priority:

1. Priority is TOC
2. Priority is drainage
3. Priority is soil texture

If there is no applicable soil database available, one could - instead of using priorities - generate a grid containing the maximum value from all the indicator grids.

Other types of ancillary data e.g. yield data and a range of proximal and remote sensors can also be used as input data for soil survey purposes.

The method is in this example used on soil map units established by the aid of both AP and SEC maps and the delineation are drawn close to the optimal for our purpose, compared to soil maps compiled without the aid of high quality ancillary data. The system can be used on existing soil maps if ancillary data could be obtained for the area, but this is not demonstrated in this study. In the case of using the application on existing soil maps, it would probably be necessary to move the line more than the ten meters in this study.

The concept of fuzzyfication of soil map delineations is relevant only on relatively detailed scale, survey order 1-4, since problems with map unit impurity increases with survey order and the relative importance of the transition zones decreases.

The decision support system for classification/fuzzyfication of soil boundaries is a pioneer way of using auxiliary data in detailed soil survey. Combined with a soil database of modal soils and the concepts and ideas behind the system, the possibilities of utilizing the system are numerous.

The short description of the system is that the ancillary data sets are used for fuzzyfication of the map units by calculation the spatial rate of change of the ancillary data on the map unit delineation and transforming this into a transition zone where each pixel contains a value representing the membership of a soil type. The system aids in selecting the best grid for the fuzzyfication by

analyzing the soil database across map unit delineation. Furthermore the ancillary data are used for making more precise map unit delineation by moving the delineation to the area with the highest spatial rate of change in the selected grid.

The system is under constant development, and especially the area of assigning priorities to the indicator grids need further intention – maybe better results could be obtained by combining the grids.

The code (written in Avenue for ArcView 3.2, but hopefully converted to use in ArcGIS during the next year) can be e-mailed upon request.

Author information

	Primary author	Coauthor
Name	Mette Balslev Greve	Mogens Humlekrog Greve
Title	GIS Software Engineer	Soil Scientist
Organization	Danish Inst. of Agricultural Sciences	Danish Inst. of Agricultural Sciences
Address	P.O.Box 50, 8830 Tjele-DK	P.O.Box 50, 8830 Tjele-DK
Telephone	+45 8999 1801	+45 8999 1734
Fax	+45 8999 1819	+45 8999 1819
E-mail	MetteB.Greve@agrsci.dk	MogensH.Greve@agrsci.dk