Soil prediction on a Low Budget? Ask the Expert!

Martin Ameskamp
Institut für Informatik und Praktische Mathematik
Christian-Albrechts-Universität Kiel
Olshausenstr. 40, 24098 Kiel, Germany
Tel.: +49 431 880-4474, Fax: +49 431 880-4054
email: ma@informatik.uni-kiel.de

Abstract

Soil survey generally means finding a model that predicts soil conditions from soil samples and landscape conditions. This 'soil-landscape model' is frequently an unrecorded, *mental* model. We employ a system of *fuzzy rules* to codify the mental soil-landscape model of an experienced soil surveyor – the expert – and use it to predict soil conditions from landscape information such as air photos and relief.

Soil information is represented in a three-dimensional model based on fuzzy classifications over a set of horizon designations. This model allows the representation of continuous transitions between soil horizons. This approach to soil prediction is especially suited for applications like precision farming.

Keywords

GIS, three-dimensional soil modelling, linguistic variables, fuzzy rules, expert systems, precision farming

1 SOIL AND SOIL SURVEY

Soil tends to show a *layered* structure; these 'layers' are known as soil *horizons*. The vertical sequence of horizons in the soil, from the surface to the unaltered parent material, is known as a soil *profile*. There are many systems for describing and classifying soil, but most of them agree in using horizons and profiles as fundamental units (Bridges 1990).

The task of a soil survey or a soil *inventory* is to provide soil information for a specific region. The general method for finding this information is based on looking for correlations between landscape features and a limited number of soil samples, and then using these correlations to predict soil conditions for the whole survey region (Hewitt 1993). These predictions are normally expressed by delineating areas of similar soils which are shown on a choropleth soil map.

Environmental Software Systems Vol.2 R. Denzer, D.A. Swayne & G. Schimak (Eds.)

This process is possible and meaningful for two reasons: First, soil formation is to a large extent influenced by external factors such as climate, relief or the action of organisms, including human activity (Jenny 1941). These factors can frequently be observed directly, most notably in the case of relief. Second, other landscape features such as vegetation or, in the case of tilled land, the colour of the bare topsoil itself, are influenced by soil conditions and thus allow the surveyor to predict soil conditions without sampling everywhere.

These patterns of correlation between soil and landscape are also called *soil-landscape* models: '*...predictive models used to make statements about soil classes and their spatial arrangements or soil properties and their trends from observations of landscape features*' (Hewitt 1993). These soil-landscape models are frequently 'mental' models. Thus, results of the survey can be hard to check or to repeat by other surveyors. In addition, the finished map does not contain any explicit information about the model itself – this is unrecorded and lost.

2 MODELLING SOIL PREDICTION

We have developed an interactive soil modelling system – called TRCS for **T**hree-dimensional **R**ule-based **C**ontinuous **S**oil modelling – that employs a computational representation of the mental soil-landscape model of an experienced soil surveyor to predict soil conditions from widely available landscape information such as relief or air photos. Unlike other researchers who use (geo-) statistical models to infer relations between soil samples and landscape information (e.g. Lagacherie *et al.* 1995; Odeh *et al.* 1994), we proceed from the assumption that experienced soil surveyors can provide valuable information about 'their' area.

The advantages of codifying this information are twofold. First, we get a soil model (this is described below). While not as 'exact' as a full survey, this 'educated guess' combines a large amount of information relevant to local soil conditions in a useful form, and where a survey is not economically feasible, e.g. when applying precision farming methods, it provides a low-cost alternative.

Second, we get access to the soil-landscape model itself. Representing this model provides a way of separating evidence (landscape information) from interpretation (the expert's conclusions). This means that the model can be applied to different sets of landscape data, that the models of different experts can be compared, and that the effects of changes to the model and to landscape information can be observed.

TRCS can best be described in terms of *three* models. One is the soil-landscape model, and the others are a *landscape model* and a *soil model*. Both provide information (landscape and soil information, respectively) as a function of geographical position. The collaboration of these models can be shown by the following example: A soil surveyor might reason: 'There is a depression here, so the soil will be peaty.' This is part of the (mental) soil-landscape model. The soil-landscape model only 'knows' that depressions are marshy. It has no information on the occurrence of depressions in the survey area, in fact, it has no spatial information whatsoever. In contrast, the landscape model provides the information that there is a depression 'here', say at position (x_d, y_d), and the soil model associates the position (x_d, y_d) with the soil information 'marshy'.

3 THE LANDSCAPE MODEL

The TRCS landscape model is based on the raster GIS GRASS (Geographical Resources Analysis Support System, CERL 1993). While TRCS imposes no restrictions on the type of landscape data that can be used, the most important sources of landscape information for cultivated land are the relief and aerial photographs. In Germany, a nation-wide soil rating survey (*'Bodenschätzung'*), conducted in 1934/36 and updated when necessary, provides texture soil information.

Some processing is necessary to use this information in soil prediction: Neither an absolute grey value nor an absolute elevation is very useful for this purpose. In the case of the air photos, georeferencing is usually necessary as well as some radiometric corrections to compensate for global trends. In the case of relief information, various methods have been proposed for an interpretation that yields useful 'features' (see e.g. Moore *et al.* 1993). In the example application presented below, we used a thin-plate spline interpolation method (Mitasova and Hofierka 1993) to compute an elevation raster from digitized contours, and the relief analysis system SARA (Köthe 1994) was used to compute *relative* relief positions (for details see Ameskamp 1997).

Even so, a soil surveyor is more likely to think of the information of, say, an air photo in *fuzzy* terms of 'bright' or 'dark' regions rather than in terms of crisp numerical grey-values. In order to provide access to landscape information in these fuzzy terms, we have provided a fuzzification interface that lets the expert define fuzzy sets that can be used as the values of *linguistic variables* (Zadeh 1975).

4 THE SOIL MODEL

In a modelling system that aims to capture a soil surveyors mental soil-landscape model, interaction should be based on paradigms that are close to the terms in which an expert thinks about soil. It has been argued above that the *horizon* and the *profile* are important to the description of soils, and these concepts are central to TRCS as well.

A horizon class is associated with certain soil qualities and thus can be seen as a carrier of soil information. In addition, the term horizon is also used for a physical contiguous soil volume that matches the definition of a specific horizon class.

TRCS represents soil in formation in terms of horizon classes. While an earlier system (BOGS, see Ameskamp *et al.* 1993) provided a 'crisp' soil model by assigning one horizon class to each point in the (three-dimensional) model range, TRCS extends this to a *continuous* or *fuzzy* soil model by using sets of *weighted* horizon classes (with weights summing to unity). A continuous transition from a 'pure' **Ap** horizon to an equally pure **Bv** horizon can be modelled by a series of sets of weighted horizons of the form $\{(\mathbf{Ap}, w_1), (\mathbf{Bv}, w_2)\}$, with w_1 varying from 0 to 1, w_2 varying from 1 to 0, and $w_2 = 1 - w_1$ everywhere (horizon designations are taken from the German 'Bodensystematik', Arbeitsgruppe Boden 1994). An alternative way of describing a set of weighted classes like this is the (fuzzy) *class vector*, where each component of the vector corresponds to one of the horizon classes used.

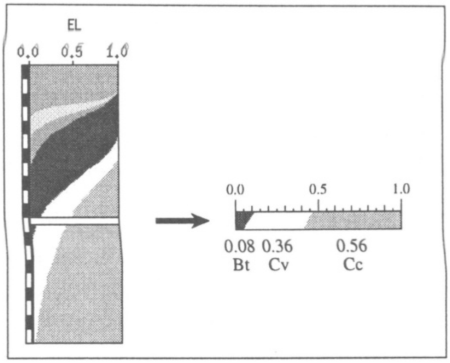


Figure 1 A fuzzy profile

The second important concept is that of the profile: A profile in the sense of a vertical soil sample is the way in which soil is normally encountered in the field. In addition, a profile in the sense of a vertical sequence of horizon classes defines the the *type* of a soil in most classification systems. In TRCS, following the *fuzzification* of horizon information from the crisp horizon class to the fuzzy class vector, a *fuzzy profile* is used to present model information for one point in the plane, from the surface to the parent material (usually at a depth of about 2 m), as a vertical sequence of class vectors. Figure 1 shows an example of a fuzzy profile describing an 'eroded Para Brown Earth'.

5 THE SOIL-LANDSCAPE MODEL

In addition to providing a vehicle for the output of soil model information, the second important use of the fuzzy profile in TRCS is in the soil-landscape model. This model, which must represent correlations between the continuous domains of landscape and soil, should itself be fuzzy, and we implement it using a system of *fuzzy rules* that are similar to those used in fuzzy logic control (see e.g. Mamdani 1993).

The general form of these rules is 'if *landscape* is *L* then *soil* is *p*'. A specific example is: 'if *relief* is *midslope* and *texture* is *loamy* and *image* is *medium* then *soil* is *S*', where *relief*, *texture*, and *image* are the names of linguistic variables; *midslope*, *loamy*, and *medium* are fuzzy sets or lingusitic labels defined for these parameters; and **S** is a name for a fuzzy profile ('Pseudogley' in this case, a soil affected by frequent waterlogging). Using fuzzy profiles in the then-part (*consequent*) of soil-landscape rules keeps these rules very close to the terminology of the soil surveyor.

Also, the profile bridges the 'dimension gap' between two-dimensional landscape information and three-dimensional soil information through the inclusion of expert knowledge on the third dimension, namely the vertical sequence of horizon class vectors expressed by the profile.

The evaluation of the fuzzy soil-landscape rule above for a point (x,y) in the plane proceeds as follows: Querying the GIS at (x,y) yields three (numerical) values. The minimum of the membership grades of these values in the fuzzy sets *midslope*, *loamy*, and *medium*, also called the 'degree of firing' of the rule, is a measure for how well landscape conditions at (x,y) match the antecedent of the rule. The result of evaluating the rule at (x,y) is the consequent profile (**S** in the example) together with the degree of firing.

In general, a TRCS soil-landscape model consists of more than one rule, and several rules may have a positive degree of firing for any one point (x,y). In this case, the weighted profiles are aggregated into a new fuzzy profile. The aggregation method, which is described in detail in Ameskamp (1997), is based on the horizon structure of the fuzzy profiles. Figure 2 illustrates the components of TRCS and their collaboration.

6 AN EXAMPLE

The example application shown in this section is for an area in the eastern part of Schleswig-Holstein, about 25 km south-east of Kiel (lat 54° 12′N, long 10° 17′E). Figure 3a shows a 400 m × 400 m section of the air photo for this area, and Figure 3b shows relative relief positions for the same area. A system of 15 rules was evaluated on a 40 × 40 × 40 grid of points, resulting in a three-dimensional raster of class vectors for 12 horizons classes. Figure 4 shows 40 % isosurfaces for three of these horizons, generated with SGI Explorer. The **M** horizon (colluvia, recent erosion deposits), is primarily found at footslopes and can be seen to form a 'ring' round the depression in the middle of the image. There's a plug of **nH** (peat) in the middle of the depression, and **Bt** horizons (clay-enriched subsoil) are found near 'hill tops' that correspond to bright areas in the air photo.

7 SUMMARY

The modelling system TRCS presented in this paper provides a way of making an important part of the experience of a soil surveyor 'operational'. In connection with widely available landscape information, TRCS computes a three-dimensional continuous soil model. This model, which combines different sources of information into tangible soil structures, is a low-cost alternative to a full soil survey, especially in situations where a very high accuracy or detail of soil information is not required, such as precision farming applications (Olesen 1995). Pedotransfer functions can be defined that translate the fuzzy soil model into a two-dimensional map of relevant soil parameter estimates (Lamp and Ameskamp 1997).

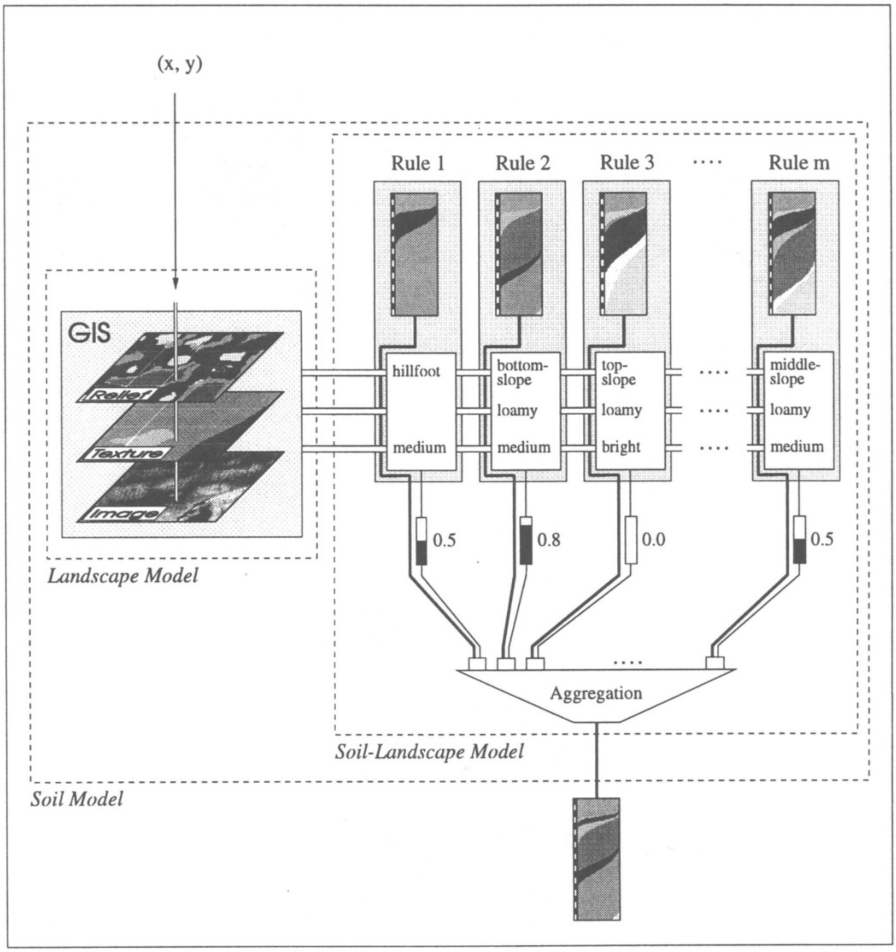
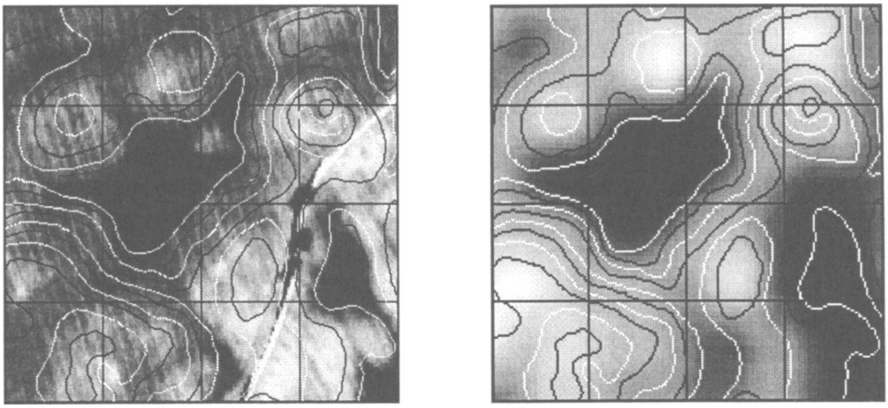


Figure 2 Overall structure of the TRCS modelling system

While relief and air photos are the principal sources of landscape information in many cases, TRCS can make use of *any* information that can be stored in a GIS, such as historical landuse (often available from old maps) or geology. Also, especially in precision farming, yield measurements can be used to improve an initial soil model.

REFERENCES

Ameskamp, M., Prey, T., and Lamp, J. (1993). Ein integriertes System zur Konstruktion von dreidimensionalen Bodenmodellen mit Hilfe der Fernerkundung. *Mitteilungen der Deutschen Bodenkundlichen Gesellschaft*, **69**, 13–16.



(a) (b)

Figure 3 Air photo and relative relief position (white = hill top, black = depression) for Kuehren Application. Grid lines are 100 m apart.

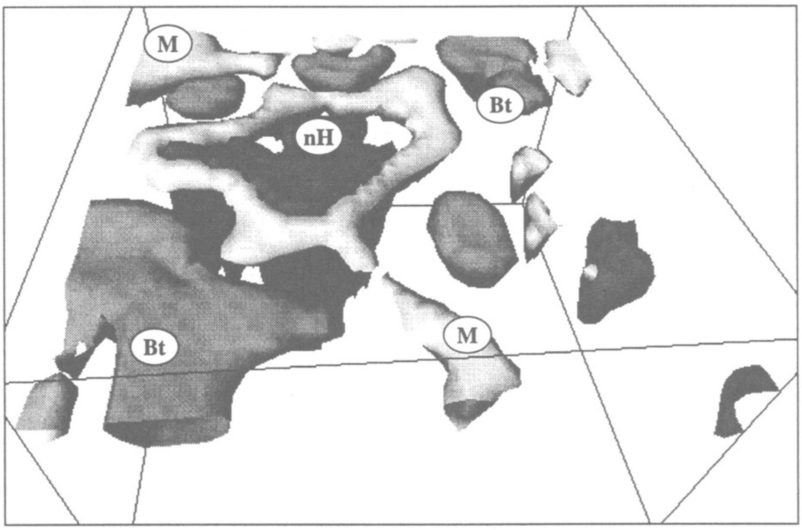


Figure 4 40 % isosurfaces for **M**, **Bt**, and **nH** horizons.

Ameskamp, M. (1997). *Three-dimensional Rule-based Continuous Soil Modelling.* Ph.D. thesis, Engineering Faculty, Kiel University, Germany. Technical Report No. 9701, Computing Department, Kiel University.

Arbeitsgruppe Boden (1994). *Bodenkundliche Kartieranleitung.* Schweizerbart, Stuttgart, fourth edition.

Bridges, E. (1990). Soil horizon designations. Technical Paper 19, International Soil Reference and Information Centre, Wageningen.

CERL (1993). *GRASS 4.1 User's Reference Manual.* United States Army Corps of Engineers Construction Engineering Research Laboratories (USACERL), Champaign, Illinois.

Hewitt, A. (1993). Predictive modelling in soil survey. *Soils and Fertilizers*, **56**, 305–314.

Jenny, H. (1941). *Factors of soil formation.* McGraw-Hill, New York. Reprinted 1994, Dover Publications, New York.

Köthe, R. and Lehmeier, F. (1994). *SARA: System zur Automatischen Relief-Analyse. Benutzer-Handbücher (Morphometrie, Version 2.22; Morphographie, Version 1.15; Ausgabe, Version 1.6).* Geographisches Institut der Universität Göttingen.

Lagacherie, P., Legros, J., and Burrough, P. (1995). A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. *Geoderma*, **65**, 283–301.

Lamp, J. and Ameskamp, M. (1997). Definition and use of functional soil horizons as key information in spatial Land Information Systems. In *Land-Information Systems: Developments for planning the sustainable use of land resources, Hanover, 20.-23. November 1996*. To be published by the European Commission.

Mamdani, E. (1993). Twenty years of fuzzy control: Experiences gained and lessons learnt. In *2nd IEEE International Conference on Fuzzy Systems*, pages 339–344.

Mitasova, H. and Hofierka, J. (1993). Interpolation by regularized splines with tension: II. Application to terrain modeling and surface geometry analysis. *Mathematical Geology*, **25**(6), 657–669.

Moore, I., Gessler, P., Nielsen, G., and Peterson, G. (1993). Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal*, **57**, 443–452.

Odeh, I., Gessler, P., McKenzie, N., and McBratney, A. (1994). Using attributes derived from digital elevation models for spatial prediction of soil properties. In *Proceedings of Resource Technology, Melbourne*, pages 451–463.

Olesen, S., editor (1995). *Proceedings of the Seminar on Site Specific Farming, Århus, March 1995*. Published as SP report no. 26, November 1995, Volume 3. Danish Institute of Plant and Soil Science, Tjele.

Zadeh, L. (1975). The concept of a linguistic variable and its application to approximate reasoning. I - III. *Information Sciences*, **8**, 199–250, 301–357; **9**, 43–80.

8 BIOGRAPHY

Martin Ameskamp received a Diploma in Computing and a Ph.D. (Dr. rer. nat.) from Kiel University. He is currently a teaching/research assistent at the Computing Department in Kiel. Research interests include the application of 'fuzzy' concepts to GIS and related fields, and visualization methods for multi-dimensional data.