Processing of soil survey data

Promotor: dr. ir. J. Bouma
hoogleraar in de bodeminventarisatie en landevaluatie,
speciaal gericht op de (sub)tropen

Arnold K. Bregt

Processing of soil survey data

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
dr. H.C. van der Plas,
in het openbaar te verdedigen
op vrijdag 15 mei 1992
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen.

Stellingen

1. De ruimtelijke verschilkansfunctie kan goed worden gebruikt voor het beschrijven van de ruimtelijke structuur van kwantitatieve bodemvariabelen.
   Dit proefschrift

2. Het berekenen van gemiddelden en varianties van nominale variablen, zoals uitgevoerd door o.a. Wilding & Drees (1978), Oliver & Webster (1987) en Agbu et al. (1990), is onjuist.
   Wilding, L.P & L.R. Drees, 1978. Spatial variability: a pedologist's viewpoint. In: Diversity of soils in the tropics. ASA Spec. Publ. 34 ASA and SSSA, Madison, WI. pp. 1-12.
   Oliver, M.A. & R. Webster, 1987. The elucidation of soil pattern in the Wyre Forest of the West Midlands, England. I. Multivariate distribution. Journal of Soil Science 38:279-291.
   Agbu, P.A., D.J. Fehrenbacher & I.J. Jansen, 1990. Soil property relationships with SPOT satellite digital data in East Central Illinois. Soil Science Society America Journal 54:807-812.
   Dit proefschrift

3. Het presenteren van bodemvariabelen in de vorm van de kans op overschrijding van een bepaalde grenswaarde is voor een gebruiker relevant en dient meer te worden toegepast.
   Dit proefschrift

4. De verandering van bodemvariabelen in de tijd krijgt binnen het bodemkarteringsonderzoek te weinig aandacht.

5. De indeling van bodemvariabelen in soil characteristics, soil properties, single land qualities en compound land qualities zoals beschreven door o.a. Bouma en Van Lanen (1987) is zeer moeilijk consequent toe te passen en werkt daardoor niet verhelderend.
   Bouma, J. & H.A.J. Van Lanen, 1987. Transfer functions and threshold values: from soil characteristics to land qualities. In: Beek et al. (Eds.). ITC Publication No 6, Enschede, pp. 106-110.

6. Variabiliteit in bodemgegevens: voor de onderzoeker een lust, voor de gebruiker een last.

7. Bij het opzetten van *geografische* informatiesystemen wordt nog onvoldoende gebruik gemaakt van de gangbare informatica-methoden voor het ontwikkelen van informatiesystemen.

8. Onzekerheid in bodemkundige gegevens vraagt 'zeker onderzoek'.

9. Gezien het feit dat belangrijke wetenschappelijke ontdekkingen in een vakgebied soms door onderzoekers vanuit andere vakgebieden worden gedaan, zou bij het selecteren van wetenschappelijk personeel minder dan tot nu toe op een bij het vakgebied passende opleiding en ervaring moeten worden gelet.

10. De richting van wetenschappelijk onderzoek wordt helaas te veel bepaald door onderwerpen die in de mode zijn.

11. De stelling van Brunsson dat (top)managers van bedrijven wel besluiten kunnen nemen, maar dat ze niet moeten verwachten dat deze besluiten ook zo door de organisatie worden uitgevoerd, geldt in versterkte mate voor onderzoeksorganisaties.

    Brunsson, N., 1991. The Organization of hypocricy: talk, decisions and actions in organizations. John Wiley & Sons, Chichester. 242 pp.

12. De mat-stelling is zowel de beste als de slechtste stelling.

Stellingen behorende bij het proefschrift van A.K. Bregt: Processing of soil survey data.

Wageningen, 15 mei 1992.

Dankwoord

Bij het tot stand komen van dit proefschrift heb ik de hulp en steun gehad van velen:

Hiervoor vanzelfsprekend mijn hartelijke dank

Abstract

Bregt, A.K., 1992. Processing of soil survey data.
Doctoral thesis. Wageningen Agricultural University, Wageningen,
The Netherlands, (X) + 167 pp.

This thesis focuses on processing soil survey data into user-specific information. Within this process four steps are distinguished: collection, storage, analysis and presentation. A review of each step is given, and detailed research on important aspects of the steps are presented.

Observation density, type of soil attributes and selection of observation sites are important aspects in the collection of soil data. The effect of observation density on the accuracy of spatial predictions was investigated in an acid sulphate soil area in Indonesia. It was found that a similar accuracy could be obtained with a marked reduction in observation density.

Most attributes collected in soil survey are on an ordinal measurement scale. Commonly used statistics, such as mean, standard deviation and semivariance, and most spatial interpolation techniques are not permissible for this type of data. Ordinal data from a soil survey in Costa Rica are used to illustrate processing possibilities. For instance, the spatial-difference-probability function was proposed for describing the spatial structure of ordinal data.

Over the past twenty years the storage of soil survey data in information systems has been receiving much attention. Digital storage is essential for rapid analysis of data. The soil information system of The Netherlands is described.

Seven main categories of soil data analysis can be distinguished. Examples of some categories are presented. The differences between interpreted soil maps on scales of 1 : 10 000, 1 : 25 000 and 1 : 50 000 for predicting moisture deficits and changes in crop yield were investigated. No differences in quality were found between the three maps when predicting average values for an area. The best predictions for point locations, however, were obtained with the 1 : 10 000 map.

Also a comparison was made between a thematic map produced by spatial prediction from point data (kriging) and one derived from a general-purpose soil map. The thematic map contain attributes that are important for water movement

in the soil. No significant difference in purity was found between the two maps. When combining soil data with other spatial data a vector to raster conversion of the soil map is often necessary. Several sheets of the soil map of The Netherlands 1 : 50 000 of different complexity were investigated for the magnitude of the rasterizing error. The regression equations determined related map complexity to rasterizing error. The rasterizing error of a complex map may be as high as 20% for a raster cell size of 4 mm * 4 mm.

Two display methods are introduced for the presentation of uncertainty in soil data. The first method yields an isoline map with empirical confidence limits based on the use of kriging and associated estimated kriging variance. The second method yields a map showing the probability that a certain threshold value is exceeded. When presenting soil data in the form of a map, the complexity of the map pattern has an important influence on its readability. Six complexity measures for maps were compared. The fragmentation index was selected as the best measure for evaluating map complexity.

Additional index words: soil survey, ordinal soil data, observation density, spatial variability, uncertainty in soil data, kriging, spatial prediction, presentation of uncertainty, map complexity.

CONTENTS Page

PART I

PART II

PART I

1 General Introduction

Soil plays an important role in activities such as agriculture, land use planning, building, erosion control, environmental protection and nature conservation. All these activities require information about attributes of the soil and soil processes. Due to the different objectives of the activities, different information is required. For example, information is required for agricultural production about soil fertility and possibilities for crop production, whereas environmental protection requires information about the amount and behavior of pollutants in the soil. Contrastingly, nature conservation needs information about soil attributes that allow the development of certain vegetation types. As a result, soil is studied from several viewpoints: soil physics, soil chemistry, soil biology, soil tillage, soil survey.

This thesis deals with soil survey, where the spatial distribution of soil on earth and its attributes are studied. The soil surveyor describes the spatial distribution of soils based on field observations and the use of landscape features, that are related to soil patterns. The obtained survey data, traditionally presented in the form of a soil map and soil survey report, however, does not always meet the requirements of the user. The 'language' spoken by the soil scientist is often not understood by, for instance, the landuse planner or the ecologist (Zinck, 1990). It is not easy for an outsider to know that the soil in The Netherlands with the code Mn25A-VI is suitable for arable farming, and that it is better not to use a soil with the code kVc-II for this purpose. The lack of understanding of the soil terminology by non-soil scientists has been realized for a long time. To bridge this gap, soil scientists have interpreted the soil data for a variety of purposes (e.g. Van Lanen, 1991). Initially, manual procedures are followed and interpretations are quite fixed. Currently, due to the introduction of the computer, more flexible interpretations are possible.

The various steps in the process of transforming collected soil data to user-specific information are presented in Figure 1.

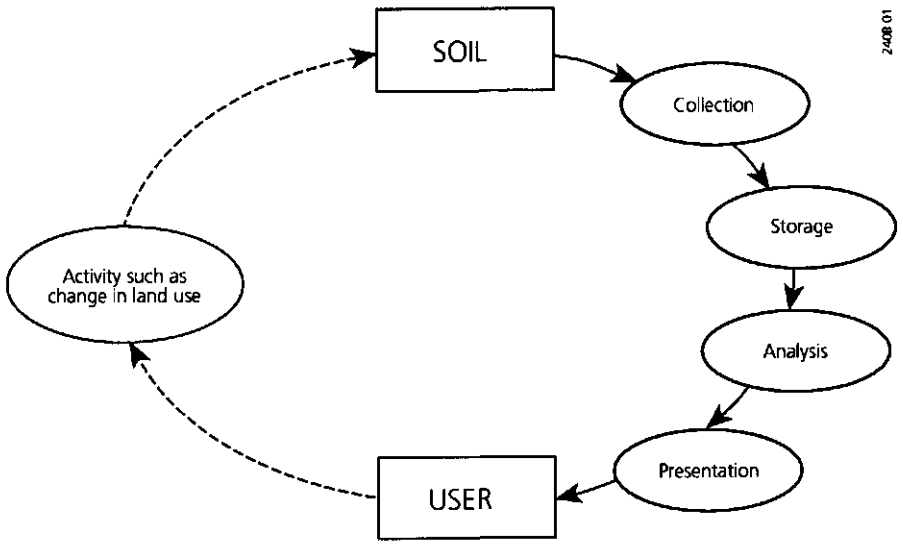


Fig. 1 Steps in the transformation of collected soil data to user-specific information (processing of soil survey data).

The subject studied is the soil in the real world. In the first step, data about the real world are collected (collection). Up to now, the most widely used data collection procedure is the free survey but when quantitative information is required, probability sampling is becoming more popular. In the second step, the collected data are stored (storage), which can be done both on paper as well as in digital form. In this thesis the emphasis is on data storage in digital form. The stored data is analyzed (analysis) in the next step to derive the desired information. The term analysis is used in this context, in a broad sense. It includes both simple selection of data as well as the assessment of land suitability by dynamic simulation modelling. After analyzing the data, the resulting information presented (presentation) can be used to initiate certain actions (or not) which influence the real world such as, erosion control measures or the removal of pollutants.

The user controls the process of data transformation to user-specific information. He asks certain questions and the soil scientist must be able to provide the answers. Certain changes have occurred in the demands. These current

changes can be summarized into three main points as follows:

- The demand for new types of information. The user is interested in other attributes and information about variability within mapping units and also in more information about the behavior of attributes in time.
- The demand for quick information.
- The demand for specification of the quality of information.

These changes in demand have their impact on the way soil survey data are collected, stored, analyzed and presented. This thesis contributes towards new procedures for soil data processing to satisfy user demands.

In Part I, Chapter 2, the various steps in the transformation of soil survey data to user-specific information are discussed in more detail. In Part II, papers which deal with aspects of the described procedures in more detail are presented, and Part III consists of concluding remarks and a summary in Dutch .

REFERENCES

Van Lanen, H.A.J., 1991. Qualitative and Quantitative physical land evaluation: an operational approach. PhD Thesis, Agricultural University, Wageningen.

Zinck, J.A., 1990. Soil survey: epistemology of a vital discipline. ITC Journal 1990-4: 335-351.

2 Processing of soil survey data: a review

2.1 NATURE OF SOIL

The nature of soil is extremely complex. McBratney (1991) describes it as follows:

> *"In any small handful of soil there is a great complexity. Such a volume of soil consists of a framework of solid particles of various sizes and shapes aggregated in an intricate manner and enclosing complementary complex and tortuous cavities within which aqueous solutions and various gases flow and diffuse and in which lives a myriad of organisms. These solids, liquids and gases interact in many complicated ways and when one considers that the soil under a square kilometer of land may consist of several thousand millions of these handfuls then the inordinate complexity of soil is manifest"*

As well as being complex at the scale of one handful, soil vary in both space and time over larger distances. The variation in space may be quite different for the various soil attributes. In a certain area, for instance, although there may be hardly any variation in clay content, a considerable variation in stoniness may be found. Different levels of spatial variation of soil attributes have been observed by, for example, Burrough (1983) and Bouma (1989). When concentrating on a certain point in space there may be a considerable variation and different levels of variation in time for the soil attributes. Soil temperature, for example, has a daily and yearly fluctuation. More static attributes like clay content and organic matter content may not change so rapidly, but over longer periods changes will occur due to soil forming processes.

It is this complex reality of many different soil attributes that we study to obtain the information we need for our activities.

Models for describing soil

The basic unit for describing soil are attribute values at a certain point (x, y, z) in time (Holmgren, 1988), in other words *what* (attribute) appears, *where* (space) and *when* (time). A complete description of reality can be obtained by recording

all attribute values in all possible points in space and time. It is obvious that this it is impossible in practice. A complete description of the reality is also not necessary. For our purposes, building a model from reality that satisfies our information requirements is enough.

There seems to be no general accepted classification for models. We devided on the following classification:

- models describing attributes in space (spatial models);
- models describing attributes in time (temporal models);
- models describing attributes in space and time (spatio-temporal models).

In this thesis, the emphasis is on spatial models that describe soil. In soil survey the spatial models that can be distinguished are:

- discrete spatial model;
- continuous spatial model.

Discrete spatial model

The discrete spatial model is the model most used in soil surveys. Other terms used for this model are choropleth map model (Burrough, 1986) and soil map model (Heuvelink & Bierkens, in prep.). The soil is divided into discrete horizons in the *z* direction (down the profile). In the *x*-*y* direction the soil is divided into a finite number of more or less homogeneous units, which are represented on a soil map as mapping units. In its simplest form, each unit is characterized by a representative profile in which one value is given for each soil attribute:

$$z_{ij} = r_j \qquad (1)$$

where z_{ij} is the value of attribute Z at point *i* in mapping unit *j* and r_j is the representative value of attribute Z in mapping unit *j*.

An extension of the model can be obtained by adding information about the variability within the map unit. This can be simply done by providing a representative profile for the dominant soil type, the associated soil type and the inclusions of each mapping unit. Another possibility is to add range indications to the attribute values. This last type of variability information is standardly provided for the soil survey of the Netherlands, scale 1 : 50 000 (e.g. Ebbers & Visschers, 1983). The variability within the map unit can also be characterized by the mean and variance of soil attributes:

$$z_{ij} = \mu_j + \varepsilon_j \quad (2)$$

where z_{ij} is the value of attribute Z at point *i* in mapping unit *j*, μ_j is the mean of Z in mapping unit *j* and ε_j is a random component with a mean of zero and variance σ^2, the within unit variance. Marsman & De Gruijter (1984), De Gruijter & Marsman (1985) and Visschers (1992) use additional sampling to obtain this type of information.

Continuous spatial model

Continuous spatial models for describing soil variation include trend surfaces, splines and the geostatistical model. The geostatistical model has become the most popular continuous spatial model for describing soil variation. It describes soil variability as a realization of a random function. The model is based on the regionalized variable theory developed by Matheron (1971). In its simplest form, the geostatistical model can be written as:

$$z_i = \mu + \varepsilon_i \quad (3)$$

where z_i is the value of Z at any location *i*, μ is the mean of Z and ε_i is a spatially correlated random component whose variation is defined by the semivariogram.

The first paper applying the geostatistical model for describing the spatial variability of soil attributes appeared in 1978 (Campbell, 1978). Later followed by publications in the United Kingdom (Burgess & Webster, 1980a,b; Webster & Burgess, 1980; Burgess et al., 1981; McBratney et al., 1982; McBratney & Webster, 1983, and in the USA (Gajem et al., 1981; Vieira et al., 1981). Since then, there has been a "boom" in publications on the use of geostatistics for describing soil variability.

Combining both models

Both the soil map and the geostatistical model involve approximations, assumptions and simplifications, and thus they disagree to some extent with reality. The soil map model disregards spatial correlation within and across map units, and ignores fuzziness of some of the unit boundaries. The geostatistical model does disregards sharp boundaries in soil variation. The dimension of time is not incorporated in either model.

According to Bouma (1985) and Voltz and Webster (1990) there is a clear

advantage in combining the two models. This has been done by Stein et al. (1988), Voltz & Webster (1990) and Heuvelink & Bierkens (in prep.). Stein et al. (1988) used the discrete model of the soil map to stratify the survey area. Within each stratum, the geostatistical model was assumed and a semivariogram was calculated for each stratum, after which predictions were made. Voltz & Webster (1990) followed a similar approach. However, instead of a semivariogram for each stratum, they used a pooled semivariogram which was assumed to be valid for all the strata. A disadvantage of the stratified procedure is that it only uses the map unit delineations of the soil map.

2.2 COLLECTION

Data collection is the first step in obtaining information about the soil for a certain area. Questions such as: *"What model do I use for spatial variation ","What data are already available for an area?","Which attributes need to be investigated?", "How many observations are required?" and "What survey or sampling method must I use?"*, need to be answered. The specific answers to these questions depend on the goal for which the soil information will be used. And as there is a great variety of goals, there can be no standardized data collection procedure to suit all possible requirements. Each time the procedure needs to be tailored to the requirements of a certain question.

According to Domburg & De Gruijter (in prep.) questions concerning soil information can be divided in a *'what' question* and a *'where' question.* The *'what' question* demands answers in the form of means, variances and proportions of attributes for a certain area (e.g. parcel, mapping unit). The *'where' question* demands answers in the form of predictions of attribute values at certain locations. In practice, combinations of both types of questions are often found. To these two types of questions a third one can be added: the *'when' question.* This question is related to the behavior of attributes in time, namely, when will the value of an attribute exceed a critical value?

In the following sections attention will be paid to data the collection aspects:

- attributes;
- survey and sampling methods.

Attributes

A great variety of soil attributes can be recorded. For example, the soil information system of the Netherlands (Bregt et al., 1987) stores about 285 different attributes. It is beyond the scope of this thesis to review all possible attributes. Detailed information on this point can be found in, for example, Soil Survey Staff (1975), FAO (1977), Hodgson (1978), Page et al. (1982), Klute (1986).

Attributes are measured or described according to a certain measurement scale. Four different scales can be distinguished: nominal, ordinal, interval and ratio. The characteristics of each scale are presented in Table 1. Attributes recorded on a nominal and ordinal scale are qualitative data, and data on interval and ratio scale are quantitative data. As shown by Bregt et al. (1992b, Chapter 4) the majority of the recorded soil data in a standard soil survey is nominal or ordinal (qualitative). This must to be realized when an analysis of soil survey data is done.

Table 1. Characteristics of measurement scales and useful summary statistics.

Scale	Description	Useful summary statistics	
		central tendency	dispersion
Nominal	determination of equality; data can be placed into classes	mode	-
Ordinal	determination of greater or less; data can be ranked	mode, median	range
Interval	determination of equality of intervals or differences	mode, median, mean	range, standard deviation
Ratio	determination of equality of ratios; measurements have a true zero	mode, median, mean	range, standard deviation, coefficient of variation

Survey and sampling methods

Beckett (1968) distinguished three main types of survey methods: physiographic

survey, free survey and grid survey. They differ in the extent to which the mapped boundaries are based on observations at sampling points, rather than on surface appearance of soil attributes as perceived on the ground, on air photos, or on topographic or geological maps. All three methods focus on the production of a soil map (*'where' question*). Over the past twenty years, however, probability sampling of the soil is being used more and more to obtain accurate information for attributes in an area (*'what' question*). The probability sampling can be in the form of a grid or another sampling scheme. In this section the following division of survey and sampling methods is used:

- physiographic survey;
- free survey;
- probability sampling.

Physiographic survey

The physiographic survey yields a description of the soil in the form of a discrete spatial model (soil map model). In the physiographic survey, the mapped soil boundaries are based on external features of the soil and the landscape perceived by interpretation of air photos or other remote sensing imageries (Dent & Young, 1981; Avery,1987). Field observations are made to describe the soil within each map unit (representative profile). The physiographic survey produces a description of the soil in the form of the discrete spatial model without any information about the variation within the map unit (Eq. 1). The attributes are usually recorded on a nominal or ordinal measurement scale, and the survey costs per square km are relatively low.

Free survey

The free survey also yields a description of the soil in the form of a discrete spatial model. It is the most widely used survey method. According to this method the soil surveyor uses known or observed relationships between soil attributes and visual features of the landscape that can be seen in the field or on air photos. In this way, the surveyor builds a conceptual model about the soil behavior. Using the model he selects each observation point from where the most useful information is likely to be obtained. Soil boundaries are delineated during field work, which means that in contrast to the physiographic survey, field observations are made to locate boundaries.

The free survey yields soil information in the form of a soil map (discrete spatial model) and descriptions of soil profiles at points. Each mapping unit is characterized by a representative profile, which is used for the interpretation of the soil map. Generally more observations are available for a map unit, so some idea can be obtained about the variation within the unit (e.g. Bregt & Beemster, 1989, Chapter 5). More detailed information is obtained by increasing the survey intensity which goes hand in hand with a narrower definition of the mapping units. In the free survey, describing the detailedness of the survey in terms of the scale of the published map is common practice. There is a direct relation between observation density and map scale, which is defined by a fixed number of observations per square cm of the final map. As a rule of thumb, Buringh et al. (1962) mentioned four observations, and Dent (1986), one observation per square cm of the final map. The recorded attributes in a free survey are mainly on a nominal or ordinal measurement scale, and the survey costs per square km are higher than those of the physiographic survey. Point observations which result from a free survey are also used for a continuous description of the soil.

Probability sampling

An essential feature of probability sampling is that the sampling locations are determined by a random procedure. This contrasts with the physiographic and free survey, where the sampling locations are chosen according to the judgement of the soil surveyor based on his experience about the relation between landscape and soil. A result of probability sampling is that statistical theory can be used for an unbiased estimation of, for instance, means and variances of attributes. The reasoning behind the great diversity of possible sampling schemes lies in the need to achieve maximum efficiency and accuracy within constrains such as cost, time, available facilities and so on.

Three basic sampling schemes can be distinguished (Ripley, 1981; Webster & Oliver, 1990):

- random;
- systematic;
- stratified.

In the random scheme the sampling locations are chosen completely at random. Random sampling has some disadvantages. Firstly, the total number of observations

needed to obtain a certain precision in the statements is relatively high. Secondly, clustering of samples can take place, which leads to over- or undersampled areas.

An even coverage of an area can be obtained by systematic sampling, in which points are located at regular intervals on a grid. The location of one grid point, however, must be chosen at random. Systematic sampling is easy to perform. If there are, however, periodicities in the soil population, the results can be biased.

With stratified sampling the area is divided into different strata. Sample locations are chosen within the different strata according to a random (stratified random) or systematic (stratified systematic) procedure.

For the selection of the final sampling scheme it is important to know the question that needs to be answered. Most probability sampling schemes concentrate on the *'what' question*. Marsman & De Gruijter (1984) use transect sampling to quantify map unit composition. If we want to answer the *'where' question*, the scheme will be quite different. Using the geostatistical model for answering this question, sampling must be done in order to:

- estimate the spatial structure of the attribute (semivariogram);
- make predictions for locations.

Sampling schemes for estimating the semivariogram are proposed by Russo (1984), Warrick & Myers (1987) and Pettitt & McBratney (1992). All three procedures proposed attempt to obtain short-range information as well as information over larger distances. Sampling schemes for spatial prediction using kriging have been evaluated by McBratney et al. (1981) and McBratney & Webster (1981). A triangular grid appeared to be the most efficient one. Square grids are nearly as efficient and have the additional advantage that they are easier to lay out in the field.

2.3 STORAGE

Traditionally the collected soil data were stored in archives in paper form. The accessibility of these archives is often quite limited, even by the survey organization themselves. For users outside the survey organization, the published profile descriptions and soil maps were often the only available soil data source.

Detailed descriptions of auger holes are not available or are very difficult to get hold of. At this moment, a lot of collected soil data in the world are still stored in paper archives. Obviously this method of storage does not allow rapid analysis of the data and, furthermore, even hampers its use. It is therefore not surprising that the storage of soil data in computerized information systems has received much attention during the last 20 years.

From a data storage point of view it is useful to divide the collected soil data into point and area data. Point data are detailed soil profile descriptions in a pit, including chemical, physical and mechanical analysis and augerhole descriptions: In soil survey, what we consider to be point data is quite often a line description in the *z* dimension (profile). Soil maps and mapping unit descriptions are regarded as area data.

Database management systems are frequently used for the storage of point data. On personal computers dBASE is quite popular (e.g. Van Waveren & Bos, 1988; Maggogo, 1989) and on mini-computers various software packages are used. Geographical information system (GIS) packages are often used for the storage of area data, where both vector and raster storage of area data are found.

International development

The first steps towards the establishment of soil information systems were made at the beginning of the seventies. On a national level, activities began in The Netherlands (Schelling, 1972), in Canada (Dumanski et al., 1975), in New Zealand (Lee et al., 1976) and in the USA (McCormack and Miller, 1977). At an international level, there have been recommendations for the development of an international soil information system (FAO, 1972). In 1974 an international working group on soil information systems was set up by the International Society of Soil Science (ISSS), and the first meeting of the working group was held in Wageningen, The Netherlands, in 1975 (Bie, 1975). The working group was very active and successive meetings followed in Canberra, Australia, in 1977 (Moore and Bie, 1977), Varna/Sofia, Bulgaria, in 1977 (Sadovski and Bie, 1978), Canberra, Australia, in 1980 (Moore et al., 1981), Paris, France, in 1981 (Girard, 1981) and Bolkesjø, Norway, in 1983 (Burrough and Bie, 1984).

At the start the contributions were mainly focused on the technical and developmental aspects of soil information systems. Later the applications of soil

information systems began to receive more and more attention. Consequently, at the congress of the ISSS in Hamburg (1986) the working group on soil information systems joined the working group on land evaluation to become the working group on land evaluation information systems. The SOTER working group was also established at the Hamburg congress. This working group was assigned the task of developing a world SOils and TERrain database on a scale of 1 : 1 000 000 (Baumgardner & Van de Weg, 1989). In 1991 databases were made available for three map sheets of different regions of 250 000 km^2.

Soil information system of The Netherlands

The first steps towards an integrated earth science information system in The Netherlands were taken in 1972. During the period 1972 until 1976, the WIA-system (WIA = Werkgemeenschap Informatiesysteem Aardwetenschappen) was developed for the input, storage and retrieval of both soil and geological data (Bie & Schelling, 1978). Although parts of this system worked well, a complete integration of data was never achieved. This was chiefly due to the complexity and diversity of the data, the wide scope of the project and the state of the art in information technology and methodology at that time. Digitization of the soil maps of The Netherlands on a scale of 1 : 50 000 was done successfully within this project.

In 1984 a new soil information system (BIS, bodemkundig informatiesysteem) for soil point data was developed. BIS was built using the relational database management system ORACLE. Prototyping was used as the developing methodology (Bregt, 1986). In 1988 the soil maps were converted to ARC/INFO, which was linked with ORACLE. The data stored in the soil information system of The Netherlands are as follows:

- area data:
  . soil map of The Netherlands on a scale 1 : 250 000;
  . soil map of The Netherlands on a scale 1 : 50 000;
  . representative profile for mapping units soil map 1 : 50 000, including land qualities and suitability classes;
- point data:
  . auger hole descriptions;
  . detailed profile descriptions, including data of chemical, physical and mechanical analysis.

The main function of the system is to store data. These data are analysed to answer a great variety of questions.

2.4 ANALYSIS

Soil data are seldom collected and stored in the proper form to answer the user's questions. Consequently, selection and transformation is necessary to yield the right information. All activities undertaken to transform the data in such a way that they become relevant for a user (become information) is called analysis.

Classification of analysis methods

In soil science literature, classification of analysis techniques are scarce, only Burrough (1989a) presented a possible classification. In the GIS literature, a classification of analysis methods can be found in Burrough (1986), Goodchild (1987), Van der Schans (1988), Aronoff (1989), and Burrough (1991). The most generic classification is presented by Van der Schans (1988). This classification is based on the three basic elements of spatial data (attribute, geometry (space), and time). Based on these three elements, seven major categories of analysis can be distinguished:

- attribute analysis;
- geometrical analysis;
- temporal analysis;
- attribute and geometrical analysis;
- attribute and temporal analysis;
- geometrical and temporal analysis;
- attribute, geometrical and temporal analysis.

In this thesis the classification of Van der Schans (1988) is used, as it is in line with the description of the soil presented before (see Section Models describing soil on page 7). Analysis on the basis of time alone and on geometry and time are not or seldom found in soil survey, so they will not be discussed in greater detail.

Attribute analysis

Attribute analysis is important in soil science. Much soil research has been focused on the derivation of new attributes from existing ones. In its general form this can be written as:

$$N = f(A,B,C,...) \quad (4)$$

where N is the new attribute, A,B,C are attributes used to estimate N, and f(...) is a function used to estimate N. The function f(...) can be:

- logical;
- arithmetical.

Logical operations

Logical operations calculate new attributes by evaluating the value of existing attributes. The existing variables are often, but not necessarily, classified data. For logical operations, Bouma & Van Lanen (1987) use the term class pedotransfer functions, and Burrough (1989b) the term threshold models.

An example of such an operation is:

if $A = v1$ and $B = v2$ then $N = v3$

where $v1$, $v2$, $v3$ are attribute values, and A, B, N are attribute names.

Logical operations are particularly suitable for attributes on a qualitative measurement scale. Because a lot of soil survey data are available on this measurement scale (Bregt et al., 1992b), logical operations are quite popular to derive new attributes from existing ones. Most of the examples can probably be found in land evaluation (Van Lanen, 1991). To obtain new attributes (e.g. suitabilities), the land capability classification system for evaluating agricultural land (McRae & Burnham, 1981) and the land suitability evaluation method of the FAO (FAO, 1976) incorporate large numbers of logical operations on attributes. Many such operations are also included in the Dutch system for soil survey interpretation (Haans, 1979).

Classification of a soil profile can also be considered as a logical evaluation of collected attributes, which yield a new attribute: the soil type.

Arithmetical operations

Arithmetical operations calculate new attributes as an arithmetical function(s) (+, -, /, *, logarithms, exponents, and all possible combinations thereof) of the existing

attributes. An example of such an operation is:

$$N = 2.5A + B + 3C \quad (5)$$

where 2.5 and 3 are constants. Other terms used for these types of operations are continuous transfer functions (Bouma & Van Lanen, 1987) and regression models (Burrough, 1989b). The arithmetical operations can only be used for quantitative attributes.

In the area of soil fertility, there are a lot of empirically defined arithmetical functions (parametric procedures) which relate crop yield to soil characteristics, such as clay content and organic matter (McRae & Burnham, 1981; Janssen et al., 1989). In soil physics a number of functions have been defined which relate soil physical characteristics to soil survey attributes. (Ritchie & Crum, 1989; Van Genuchten et al., 1989; Vereecken et al., 1989; Wösten & Van Genuchten; 1988; Wösten, 1990). Examples of functions which relate soil chemical characteristics to soil attributes can be found in Breeuwsma et al. (1986).

Geometrical analysis

Geometrical analysis includes those activities that modify the geometry of soil data such as smoothing of map unit boundary lines, edge matching, change of map projection, map overlay and vector to raster or raster to vector conversion.

Map overlay can be considered as a gemometrical analysis only when attributes of the individual maps are transferred to the resulting map without changes being made. An example is the overlay of the soil map of Europe with the agro-climatic map of Europe (Bulens et al., 1990).

Vector to raster conversion of the soil map is commonly done. However, the error introduced as a result of this analysis can be quite large. Bregt et al. (1991a) showed that by a vector to raster conversion of the soil of The Netherlands even for a relatively small raster-cell size of 4 mm * 4 mm there could be an error of 20% (Fig. 1).

Attribute and geometrical analysis

Attribute and geometrical analysis go frequently hand in hand. The types of analysis belonging to this category are spatial prediction and overlay.

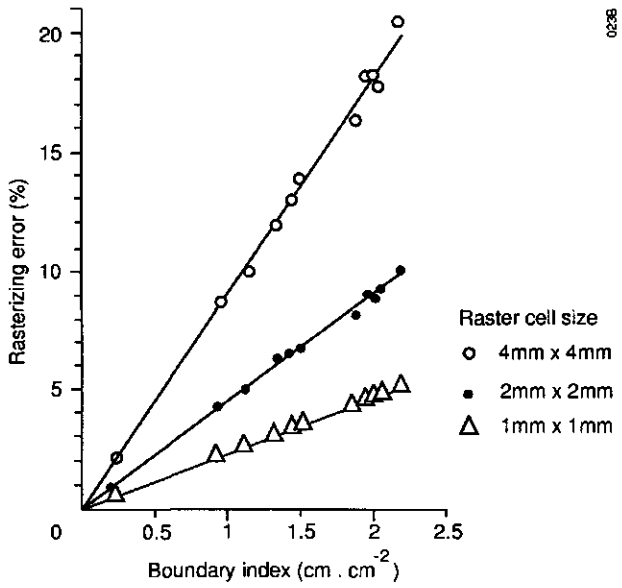


Fig. 1 Relation between rasterizing error and boundary index for eleven map sheets for three raster cell sizes (Bregt et al., 1991)

Spatial prediction

Spatial prediction methods are used to predict the soil attributes at, usually, unvisited locations from point observations. A large number of possible techniques are available (Ripley, 1981; Lam, 1983; Burrough, 1986; Stein, 1991). Most techniques, such as kriging, trend surface analysis and inverse distance interpolation can only be applied to interval and ratio data. Spatial prediction with Thiessen polygons (Davis, 1986) can, however, also be applied to nominal and ordinal data. Kriging has become the most popular mathematical prediction technique in soil survey. It is based on the geostatistical model for soil variation.

Studies comparing the accuracy of different methods for predicting soil attributes have been described by Van Kuilenburg et al. (1982), Bregt et al. (1987), Laslett et al. (1987) and Voltz & Webster (1990). Van Kuilenburg et al. (1982) and Bregt et al. (1987, Chapter 6) compared kriging with soil map predictions. In these case studies no difference was found in the accuracy. Laslett et al. (1987) compared thirteen different spatial prediction methods, all of which showed some deficiencies. Of all the methods investigated, Laplacian smoothing splines and kriging appeared to perform best. Voltz and Webster (1990) used four methods

to predict the clay content of the topsoil in two regions of contrasting physiography, which were: the mean of the mapped class, kriging, cubic spline and kriging within classes. In situations where some abrupt changes in the soil had occurred, the mapped classes performed best. In other situations, kriging and kriging within classes performed somewhat better.

From the presented case studies about the experimental comparison of spatial prediction methods no preferance can be detected, of for example, the geostatistical model (kriging) preferable to the soil map model.

Overlay

Map overlay can also be both an attribute and geometrical analysis, as seen when new geometry and attributes are obtained by an overlay of maps together with an logical or arithmetical operation on the attributes (see Section Attrbiute analysis on page 18).

Attribute and temporal analysis

Attribute and temporal analysis are found in those situations where attribute values are analyzed in time, without taking into account the spatial interaction between attributes values. Typical data used for this type of analysis are time series collected in a monitoring program. Temporal models describe the behavior of attributes in time, examples of which are modelling soil compaction, change of organic matter content and crop growth. Temporal models are focussed on answering the *'when' question.*

In soil survey research on attribute analysis in time did not receive much attention. This will undoubtly change due to the increasing attention for sustainable land use, which asks questions about the behavior of attributes in time.

Attribute, geometrical and temporal analysis

This last category comprises the most complicated analysis. There is a spatial and temporal interaction between attribute values. Spatio-temporal models are frequently used to analyze and describe this interaction, and they are focussed on answering the *'when'* and *'where' question.* Two types of spatio-temporal models are distinguished:

- deterministic model;
- stochastic model.

Deterministic spatio-temporal models attempt to model the most important aspects of the processes in terms of well understood physical and chemical laws. Most models for describing the water-soil-plant interaction are of this type (e.g. WOFOST (Van Diepen et al., 1989), SWATRE (Feddes et al., 1978). These models generally model spatial interaction only in one dimension (down the profile). Most groundwater models are deterministic, they usually model space in three dimensions.

Stochastic spatio-temporal models do not model the outcome of a single event but the average behavior of a large number of events. The output of the model does not describe an exact result but rather one possible realization within a whole range of results. The use of stochastic process models is still in an experimental phase. Up to now the main application field has been in hydrology and solute movement (Serrano et al., 1985a,b,c; Yeh et al., 1985; Jury & Gruber, 1989). They have hardly been used in crop production analysis.

A combination of the stochastic and the deterministic approach is often found. For instance, the spatial variation of attributes is described by a stochastic model (geostatistical model) and the behavior of attributes in time is described by a deterministic model. An example has been presented by Stein et al. (1991). They described the spatial variability of soil attributes by the stochastic geostatistical model. Using the geostatistical model predictions were made for test locations, followed by deterministic simulation of the moisture deficit on the test locations.

Combined analysis

For answering a particular question, often a sequence of analysis is required. This will be illustrated by a study concerning crop production potentials for the European Community (Bulens et al., 1990; Van Lanen et al., 1992). In this study potential- and water limited production of some major crops were determined by the combined use of a geographical information system (GIS) and land evaluation models. The soil map of European Community on a scale of 1 : 1 000 000 (EC, 1985) was used. The projection of the soil map was changed (geometrical analysis) and new attributes (rooting depth, drainage class and water retention data) were

derived for the dominant soil units (attribute analysis). The soil map was overlaid with the agro-climatic and an administrative unit map in order to obtain a land evaluation map (attribute and geometrical analysis). The units on the obtained map represent an unique combination of soil, agro-climatic zone and administrative regions. For the land evaluation units potential- and water limited production for some major crops was determined by the crop growth model WOFOST (Van Diepen et al., 1989) (attribute, geometrical and temporal analysis).

2.5 PRESENTATION

Results of an analysis must be presented to the user. Two basic forms can be distinguished:

- text;
- illustrations.

Text

The results of an analysis can be presented in the form of text. The term text is rather broadly defined in this context, and includes tables, formulas and written text. The advantage of the use of text for presentation is that the statements can be precise (e.g. numbers). Tables too form an effective way for the presentation of information for further computer processing of the obtained results. The disadvantage is that the human mind is not able to process text quickly. Illustrations are far more appreciated as means of communication: "pictures speak louder than words".

Illustrations

At present there is a growing interest in visualization of data for exploring data analysis. Computer workstations are now able to deliver the power to display and manipulate images at reasonable costs (Medyckyj-Scott, 1991). In the visualization of results a distinction can be made such as in graphs, maps and animations.

Graph

A graph is used for visualization of attribute change in one dimension. This type of output is common when changes of attributes in time or in one dimensional space are presented (transect, profile). Line graphs can be used for quantitative data; for qualitative data, diagrams are useful.

Map

The most common way of presenting soil survey information is in the form of a map. There is a great diversity of map types (Ormeling & Kraak, 1990). The measurement scale of the data and the spatial model determine the possible types. Maps most used for presenting soil survey data are chorochromatic map, choropleth map and isoline map.

A chorochromatic map is based on the discrete spatial model for soil variation and visualizes qualitative data. The general purpose soil map is a typical example of a chorochromatic map (Fig. 2).

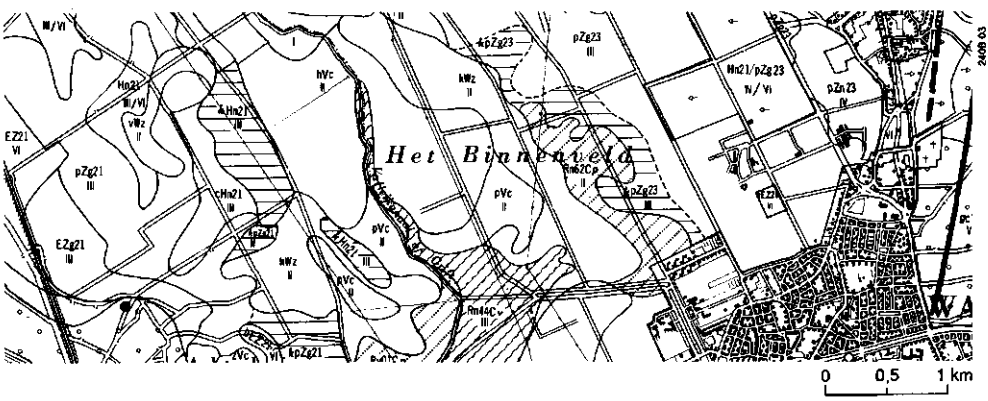


Fig. 2 Chorochromatic map. Fragment of the soil map of The Netherlands.

A choropleth map is also based on the discrete spatial model for soil variation, but now quantitative data are presented. This type of map is often obtained by deriving a quantitative attribute from a soil map, or by classification of a continuous attribute (Fig. 3).

An isoline map presents an attribute that varies continuously over an area (continuous spatial model), and where points of equal value are connected by lines (Dent, 1985). Results of spatial predictions are often presented as isoline maps (Fig. 4)

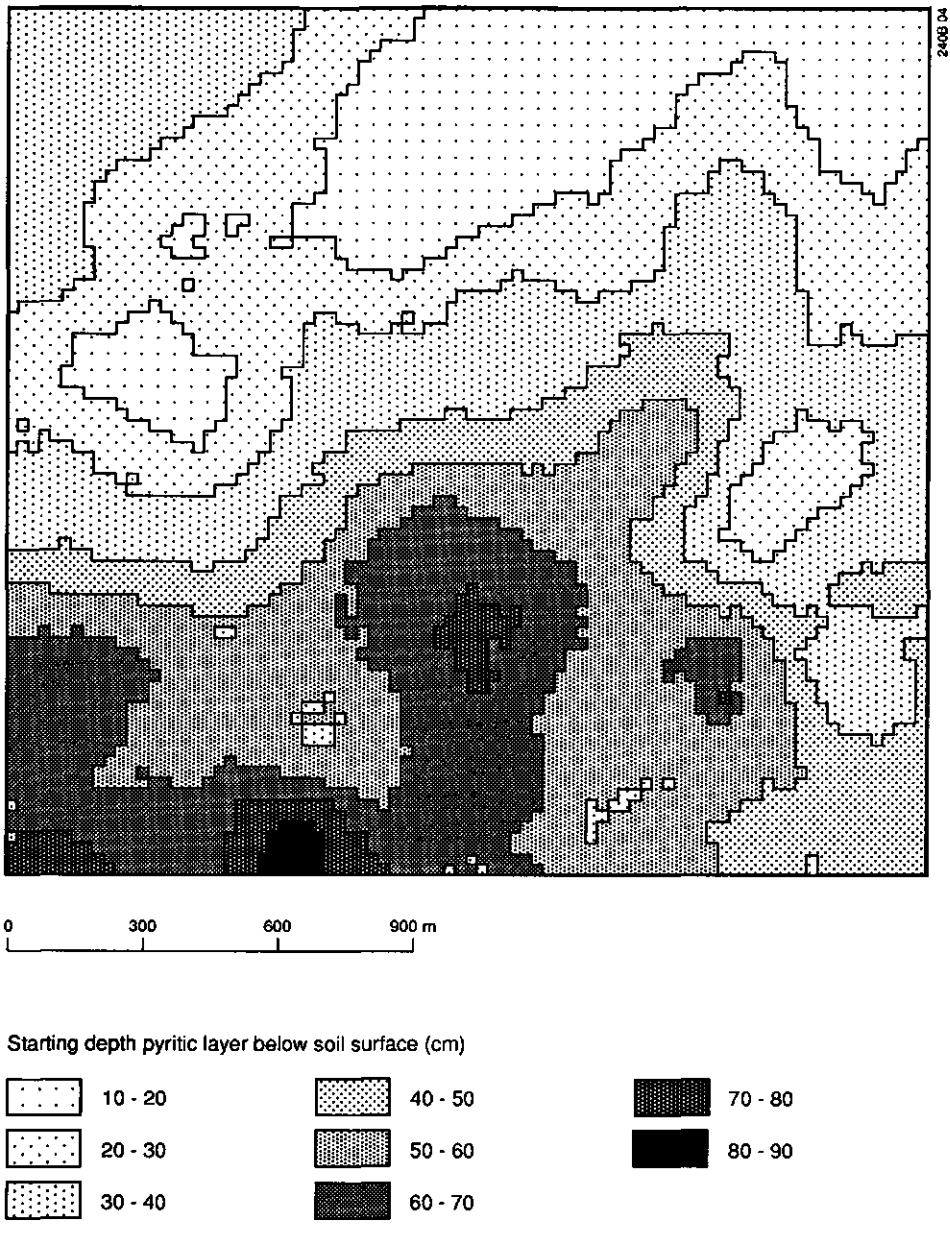


Fig. 3 Choropleth map obtained by classification ofnthe continuous attribute depth to the pyritic layer (Bregt et al., 1992a).

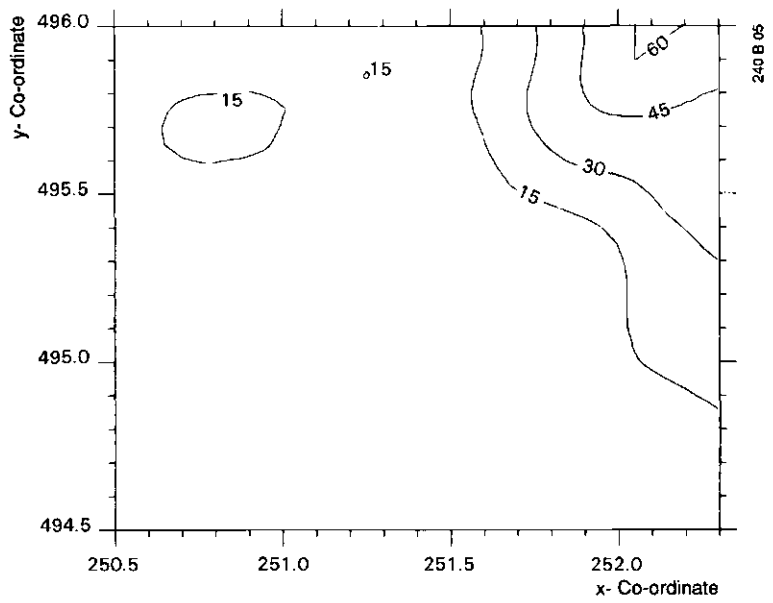


Fig. 4 Isoline map of the attribute moisture deficit (Bregt et al., 1991b).

Three dimensional visualization (or 2.5D depending on one's view) is occasionally used for soil attributes which have only one value in the z direction (depth to groundwater, depth to pyritic layer, rooting depth) (Fig. 5). The most common 3D visualization is, however, a Digital Terrain Model (DTM).

The presentation of uncertainty in data has long been a neglected item in map making. A chorochromatic soil map, for instance, suggests that we are dealing with homogeneous units without any impurities. Recently, the presentation of uncertainty has been receiving more and more attention. A NCGIA (National Center for Geographic Information and Analysis) research initiative about this topic was started in 1991 with a specialist meeting (Beard et al., 1991). New techniques have been developed. For instance, Fisher (1991) presented a computer program which visualizes the impurities of a soil map, and Bregt et al. (1991) (Chapter 8) presented a method for the construction of an isoline map with confidence limits (Fig. 6). Although some progress has been made this field, a lot of research is still needed.

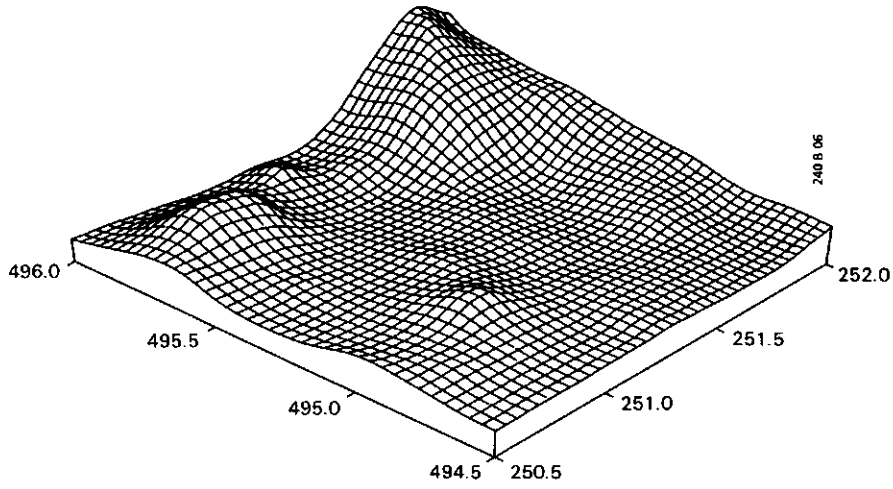


Fig. 5 3D presentation of the attribute moisture deficit (Bregt et al., 1991b).

Animation

The term animation is used for visualization of spatial processes in time (Koussoulakou, 1990). A nice example of visualization of processes in time is given by Wesseling (in prep.). He developed the software programma BALANCE to show the terms of the soil-water balance on a graphic screen. Despite this example, animation is hardly used in soil science. But it is obvious that it has great potential. It is possible, for instance, to visualize alternative soil erosion scenario's by running an erosion simulation model in conjunction with a DTM, and to show where erosion and deposition will occur. And also the movement of water and pollutants in the soil can effectively be visualized by computer animation. The development of these techniques will be technology-driven. It is our task to select the appropriate techniques for our soil applications.

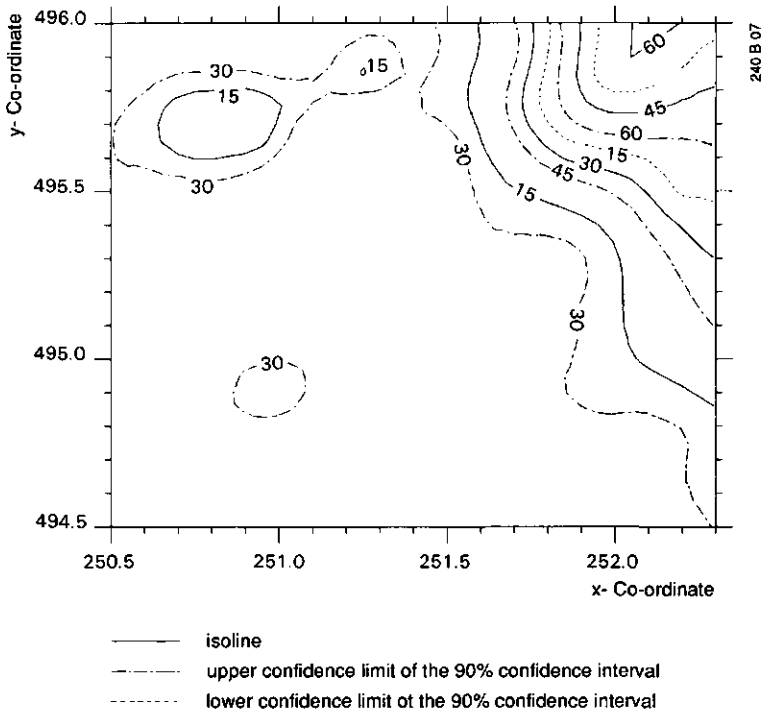


Fig. 6 Isoline map with confidence limits (Bregt et al., 1991b).

REFERENCES

Aronoff, S., 1989. Geographic information systems: a management perspective. WDL Publications, Ottawa.

Avery, B.W., 1987. Soil survey methods: a review. Soil Survey Technical Monograph No. 18, Soil Survey and Land Research Centre, Silsoe.

Baumgardner, M.F & R.F. Van de Weg, 1989. Space and time dimensions of a world soils and terrain digital database. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space and time. Pudoc, Wageningen. p. 35-43.

Beard, M.K., B. Buttenfield & S. Clopham (Eds.), 1991. Visualization of data quality: scientific report for the specialist meeting. NCGIA Technical Report 91-26, Santa Barbara.

Beckett, P.H.T., 1968. Methods and scale of land resource surveys, in relation to precision and cost. In: G.A. Stewart (Ed.), Land evaluation. Macmillam of Australia, Victoria, p. 53-63.

Bie, S.W. (Ed.), 1975. Soil information systems. Proceedings of the meeting of the ISSS working group on soil information systems, Wageningen, The Netherlands, September 1-4, 1975. Pudoc, Wageningen.

Bie, S.W. & J. Schelling, 1978. An integrated information system of soil survey. In: A.N. Sadovski & S.W. Bie (Eds.), Developments in soil information systems. Pudoc, Wageningen. p. 43-49.

Bouma, J., 1985. Soil variability and soil survey. In: D.R. Nielsen & J. Bouma (Eds.), Soil spatial variability. Pudoc, Wageningen, p. 130-145.

Bouma, J., 1989. Land qualities in space and time. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space and time. Pudoc, Wageningen, p. 3-13.

Bouma, J. & H.A.J. van Lanen, 1987. Transfer functions and threshold values: from soil characteristics to land qualities. In: Beek et al. (Eds.). ITC Publication No 6, Enschede, p. 106-110.

Breeuwsma, A., J.H.M. Wösten, J.J. Vleeshouwer, A.M. van Slobbe & J. Bouma, 1986. Derivation of land qualities to assess environmental problems from soil surveys. Soil Science Society American Journal 50: 186-190.

Bregt, A.K., 1986. Development of a soil information system by prototyping. In: M.F. Baumgardner & R. Oldeman (Eds.), Proceedings of an international workshop on the structure of a digital international soil resources map annex database, ISRIC, Wageningen. p. 44-47.

Bregt, A.K., J. Bouma & M. Jellinek, 1987. Comparison of thematic maps derived from a soil map and from kriging of point data. Geoderma 39: 281-291.

Bregt, A.K., J. Denneboom, Y. van Randen & E. Rose, 1987. Gebruikershandleiding bodemkundig informatiesysteem. Rapport 1945, Stichting voor Bodemkartering, Wageningen.

Bregt, A.K. & J.G.R. Beemster, 1989. Accuracy in predicting moisture deficits and changes in yield from soil maps. Geoderma 43: 301-310.

Bregt, A.K., J. Denneboom, H.J. Gesink & Y. van Randen, 1991a. Determination of rasterizing error: a case study with the soil map of The Netherlands. International Journal of Geographical Information Systems 5: 361-367.

Bregt, A.K., A.B. McBratney & M.C.S. Wopereis, 1991b. Construction of isolinear maps of soil attributes with empirical confidence limits. Soil Science Society of America Journal 55: 14-19.

Bregt, A.K., H.J. Gesink & Alkasuma, 1992a. Mapping the conditional probability of soil variables. Geoderma (in press).

Bregt, A.K., J.J. Stoorvogel, J. Bouma & A. Stein, 1992b. Mapping ordinal data in soil survey: a Costa Rican example. Soil Science Society of America Journal (in press).

Bulens, J.D, A.K. Bregt, G.H.J. de Koning, G.H.J. Reinds, C.A. van Diepen & H.A.J. van Lanen, 1990. GIS supporting crop production potential research for the European Communities. In: J. Harts, H.F.L. Ottens, H.J. Scholten & D.A. Ondaatje (Eds), EGIS'90. EGIS Foundation, Utrecht, p. 117-125.

Burgess, T.M. & R. Webster, 1980a. Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. Journal of Soil Science 31: 315-331.

Burgess, T.M. & R. Webster, 1980b. Optimal interpolation and isarithmic mapping of soil properties. II. Block kriging. Journal of Soil Science 31: 333-341.

Burgess, T.M. & R. Webster & A.B. McBratney, 1981. Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. Journal of Soil Science 32: 643-659.

Buringh, P. & G.L. Steur & A.P.A. Vink, 1962. Some techniques and methods of soil survey in The Netherlands. Netherlands Journal of Agricultural Science 10: 157-178.

Burrough, P.A., 1983. Multiscale sources of spatial variation of soil. I. The application of fractal concepts to nested levels of soil variation. Journal of Soil Science 34: 577-597.

Burrough, P.A., 1986. Principles of geographical information systems for land resources assessment. Clarendon Press, Oxford.

Burrough, P.A., 1989a. Modelling land qualities in space and time: the role of geographical information systems. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space and time. Pudoc, Wageningen, p. 45-59.

Burrough, P.A., 1989b. Matching spatial databases and quantitative models in land resource assessment. Soil Use and Management 5: 3-8.

Burrough, P.A., 1991. The development of intelligent geographical information systems. In: J. Harts et al. (Eds.), Proceedings EGIS '91. EGIS Foundation, Utrecht, p. 165-174.

Burrough, P.A. and S.W. Bie (Eds.), 1984. Soil information systems technology. Proceedings of the sixth meeting of the ISSS working group on soil information systems, Bolkesjø, Norway, 28 February - 4 March 1983. Pudoc, Wageningen.

Campbell, J.B., 1978. Spatial variation of sand content and pH within single contiguous delineations. Soil Science Society of America Journal 42: 460-464.

Davis, J.C., 1986. Statistics and data analysis in geology. John Wiley & Sons, New York.

De Gruijter, J.J. & B.A. Marsman, 1985. Transect sampling for reliable information on mapping units. In: D.R. Nielsen & J. Bouma (Eds.), Soil spatial variability. Pudoc, Wageningen, p. 150-163.

Dent, B.D., 1985. Principles of thematic map design. Addison Wesley, Reading.

Dent, D.L., 1986. Acid sulphate soils: a baseline for research and development. Publication 39, ILRI, Wageningen.

Dent, D. & A. Young, 1981. Soil survey and land evaluation. George Allen & Unwin, London.

Domburg, P. & J.J. de Gruijter, 1992. A framework of concepts for soil survey with probability sampling. Report 55, The Winand Staring Centre, Wageningen.

Dumanski, J., B. Kloosterman & S.E. Brandon, 1975. Concepts, objectives and structure of the Canada soil information system. Canadian Journal of Soil Science 55: 181-187.

Ebbers, E. & R. Visschers, 1983. Bodemkaart van Nederland schaal 1 : 50 000. Toelichting bij kaartblad 28 West Almelo. Stichting voor Bodemkartering, Wageningen.

EC, 1985. Soil map of the European Communities, 1 : 1 000 000. European Communities, Luxembourg.

FAO, 1972. Report of the FAO/UNESCO ad hoc consultation on computerized soil data interpretation for development purposes. Rome, April 19-21, 1972.

FAO, 1976. A framework for land evaluation. FAO Soils Bulletin 32, Rome.

FAO, 1977. Guidelines for soil profile description. FAO, Rome.

Feddes, R.A., P.J. Kowalik & H. Zaradny, 1978. Simulation of field water use and crop yield. Simulation monograph. Pudoc, Wageningen.

Fisher, P.F., 1991. Modelling soil map-unit inclusions by Monte Carlo simulation. International Journal of Geographical Information Systems. 5: 193-208.

Gajem, Y.M., A.W. Warrick & D.E. Myers, 1981. Spatial structure of physical properties of a typic torrifluvent soil. Soil Science Society of America Journal 45: 709-715.

Girard, M.C., 1981. Proceedings ISSS working group on soil information systems, Paris, September 14-17, 1981. Department des Sols Nos. 4,5,6. Institut National Agronomique, Paris.

Goodchild, M.F., 1987. A spatial analytical perspective on geographical information systems. International Journal of Geographical Information Systems 1: 327-334.

Haans, J.C.F.M. (Ed.), 1979. De interpretatie van bodemkaarten. Rapport van de werkgroep interpretatie bodemkaarten, stadium C. Rapport 1463, Stichting voor Bodemkartering, Wageningen.

Haans, J.C.F.M., G.L. Steur & G. Heide (Eds.), 1984. Progress in land evaluation. Balkema, Rotterdam.

Heuvelink, G.B.M. & M.F.P. Bierkens, in prep. Combining multiple sources of information to predict quantitative soil properties. Submitted to Geoderma.

Hodgson, J.M., 1978. Soil sampling and soil description. Clarendon Press, Oxford.

Holmgren, G.G.S., 1988. The point representation of soil. Soil Science Society of America Journal 52: 712-716.

Janssen, B.H., F.C.T. Guiking, D. van der Eijk, E.M.A. Smaling, J. Wolf & H. Reuler, 1989. A system for quantitative evaluation of soil fertility and the response to fertilizers. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space and time. Pudoc, Wageningen, p. 185-188.

Jury, W.A. & J. Gruber, 1989. A stochastic analysis of the influence of soil and climate variability on the estimate of pesticide groundwater pollution potential. Water Resources Research 25: 2465-2472.

Klute, A. (Ed.), 1986. Methods of soil analysis, physical and mineralogical methods. American Society of Agronomy, Madison, Wisconsin.

Koussoulakou, A. 1990. Computer-assisted cartography for monitoring spatio-temporal aspects of urban air pollution. Delft University Press, Delft.

Lam, N.S., 1983. Spatial interpolation methods: a review. The American Cartographer 10: 129-149.

Laslett, G.M., A.B. McBratney, A.B. Pahl & M.F. Hutchinson, 1987. Comparison of several spatial prediction methods for soil pH. Journal of Soil Science 38: 325-241.

Lee, R., G. Mew, M.J. Newman & A.R. Gibson, 1976. Computer processing of soil profile data from surveys in New Zealand. Geoderma 16: 201-211.

Magoggo, J.P., 1989. SISTAN: a soil information system for Tanzania. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space and time. Pudoc, Wageningen. p. 105-108.

Marsman, B.A. & J.J. de Gruijter, 1984. Dutch soil survey goes into quality control. In: P.A. Burrough & S.W. Bie (Eds.), Soil information systems technology. Pudoc, Wageningen, p. 127-134.

Matheron, G., 1971. The theory of the regionalized variables and its applications. Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau No. 5.

McBratney, A.B., 1991. Rudiments of soil sampling and data analysis. In: S.M. Bannerman & P. Hagelton (Eds.), Soil Technology-Applied Soil Science. Australian Society of Soil Science, NSW, Sydney. p. 253-284

McBratney, A.B. & R. Webster, 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. II. Program and examples. Computers and Geosciences 7: 331-334.

McBratney, A.B., R. Webster & T.M. Burgess, 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. Computers & Geosciences 7: 335-365.

McBratney, A.B., R. Webster, R.G. McLaren & R.B. Spiers, 1982. Regional variation of extractable copper and cobalt in the topsoil of south-east Scotland. Agronomie 2: 969-982.

McBratney, A.B. & R. Webster, 1983. Optimal interpolation and isarithmic mapping of soil properties. V. Co-regionalization and multiple sampling strategy. Journal of Soil Science 34: 137-162.

McCormack, D.E. & C.R. Miller, 1977. Resource information system of the U.S. Soil Conservation Service. Comm. 5 Publ. Pap. Fig XV Int. Congr. Surv. Stockholm, Sweden.

McRae, S.G. & C.P. Burnham, 1981. Land evaluation. Clarendon Press, Oxford.

Medyckyj-Scott, D., 1991. The presentation of spatial data. Mapping Awareness 5 (No.5): 19-22.

Moore, A.W. and S.W. Bie (Eds.), 1977. Uses of soil information systems. Proceedings of the Australian meeting on the ISSS working group on soil information systems, Canberra, Australia, March 2-4, 1976. Pudoc, Wageningen.

Moore, A.W., B.G. Cook and L.G. Lynch (Eds.), 1981. Information systems for soil and related data. Proceedings of the second Australian meeting of the ISSS working group on soil information systems, Canberra, Australia, 19-21 February 1980. Pudoc, Wageningen.

Ormeling, F.J & M.J. Kraak, 1990. Kartografie: visualisatie van ruimtelijke informatie. Delftse Universitaire Pers, Delft.

Page, A.L. et al. (Eds.), 1982. Methods of soil analysis, chemical and microbiological properties. American Society of Agronomy, Madison, Wisconsin.

Pettitt, A.N. & A.B. McBratney, 1992. Sampling designs for estimating spatial variance components. Applied Statistics (in press).

Ripley, B.D., 1981. Spatial Statistics. John Wiley & Sons, New York.

Ritchie, J.T. & J Crum, 1989. Converting soil survey characterization data into IBSNAT crop model input. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space and time, Pudoc, Wageningen, p. 155-167.

Russo, D., 1984. Design of an optimal sampling network for estimating the variogram. Soil Science Society of America Journal 48: 708-716.

Sadovski, A.N. & S.W. Bie (Eds.), 1978. Developments in soil information systems. Proceedings of the second meeting of the ISSS working group on soil information systems, Varna/Sofia, Bulgaria, 30 May - 4 June 1977. Pudoc, Wageningen.

Schelling, J., 1972. Earth science information system for the Netherlands. In: Report of the FAO/UNESCO ad hoc consultation on computerized soil data interpretation for development purposes. Rome, April 19-21, 1972. FAO, Rome. p. 29-30.

Serrano, S.E, T.E. Unny & W.C. Lennox, 1985a. Analysis of stochastic groundwater problems. Part I: Deterministic partial differential equations in groundwater flow. A functional-analytical approach. Journal of Hydrology 82: 247-263.

Serrano, S.E, T.E. Unny & W.C. Lennox, 1985b. Analysis of stochastic groundwater problems. Part II: Stochastic partial differential equations in groundwater flow. A functional-analytical approach. Journal of Hydrology 82: 265-284.

Serrano, S.E, T.E. Unny & W.C. Lennox, 1985c. Analysis of stochastic groundwater problems. Part III: Approximate solution of stochastic partial differential equations. Journal of Hydrology 82: 285-306.

Soil Survey Staff, 1975. Soil Taxonomy: A basic system of soil classification for making and interpreting soil surveys. USDA-SCS Agric. Handb. 436. Washington, DC.

Stein, A., M. Hoogerwerf & J. Bouma, 1988. Use of soil-map delineations to improve (co-)kriging of point data on moisture deficits. Geoderma 43: 163-177.

Stein, A, 1991. Spatial Interpolation. PhD Thesis, Agricultural University, Wageningen.

Stein, A, I.G. Staritsky, J. Bouma, A.C. Van Eijnsbergen & A.K. Bregt, 1991. Simulation of moisture deficits and areal interpolation by universal cokriging. Water Resources Research 27: 1963-1973.

Van der Schans, R., 1988. Verwerking van de geografische gegevensstroom met steun van de computer. Nr. LUW-LMK-198803, Landbouwuniversiteit, Wageningen.

Van Diepen, C.A., J. Wolf, H. van Keulen & C. Rappoldt, 1989. WOFOST: a simulation model for crop production. Soil Use and Management 5: 16-24.

Van Genuchten, M. Th., F. Kaveh, W.B. Russell & S.R. Yates, 1989. Direct and indirect methods for estimating the hydraulic properties of unsaturated soils. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space and time, Pudoc, Wageningen, p. 61-72.

Van Kuilenburg, J., J.J. de Gruijter, B.A. Marsman, & J. Bouma, 1982. Accuracy of spatial interpolation between point data on soil moisture supply capacity, compared with estimates from mapping units. Geoderma 27: 311-325.

Van Lanen, H.A.J., 1991. Qualitative and quantitative physical land evaluation: an operational approach. PhD Thesis, Agricultural University, Wageningen.

Van Lanen, H.A.J., C.A. van Diepen, G.J. Reinds, G.H.J. de Koning, J.D. Bulens & A.K. Bregt, 1992. Physical land evaluation methods and GIS to explore the crop growth potential and its effects within the European Communities. Agricultural Systems (in press).

Van Waveren, E.J. & A.B. Bos, 1988. ISRIC soil information system. User manual and technical manual. Technical papers 15, ISRIC, Wageningen.

Vereecken, H., J. Maes, J. van Orshoven & J. Feyen, 1989. Deriving pedotransfer functions of soil hydraulic properties. In: J. Bouma & A.K. Bregt (Eds.), Land qualities in space in time. Pudoc, Wageningen, p. 121-124.

Vieira, S.R., D.R. Nielsen & J.W. Biggar, 1981. Spatial variability of field-measured infiltration rate. Soil Science Society America Journal 45: 1040-1048.

Visschers, R, 1992. Upgrading van de bodemkaart van Nederland door steekproeven in kaarteenheden. Rapport 186. The Winand Staring Centre, Wageningen.

Voltz, M., & R. Webster, 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science 41: 473-490.

Warrick, A.W. & D.E. Myers, 1987. Optimization of sampling locations for variogram calculations. Water Resources Research 23: 496-500.

Webster, R.& T.M. Burgess, 1980. Optimal interpolation and isarithmic mapping. III. Changing drift and universal kriging. Journal of Soil Science 31: 505-524.

Webster, R. & M.A. Oliver, 1990. Statistical methods in soil and land resource survey. Oxford University Press, Oxford.

Wesseling, J.G. (in prep.) The software package balance.

Wösten, J.H.M. & M. Th. van Genuchten, 1988. Using texture and other soil properties to predict the unsaturated soil hydraulic functions. Soil Science Society of America Journal 52: 1762-1770.

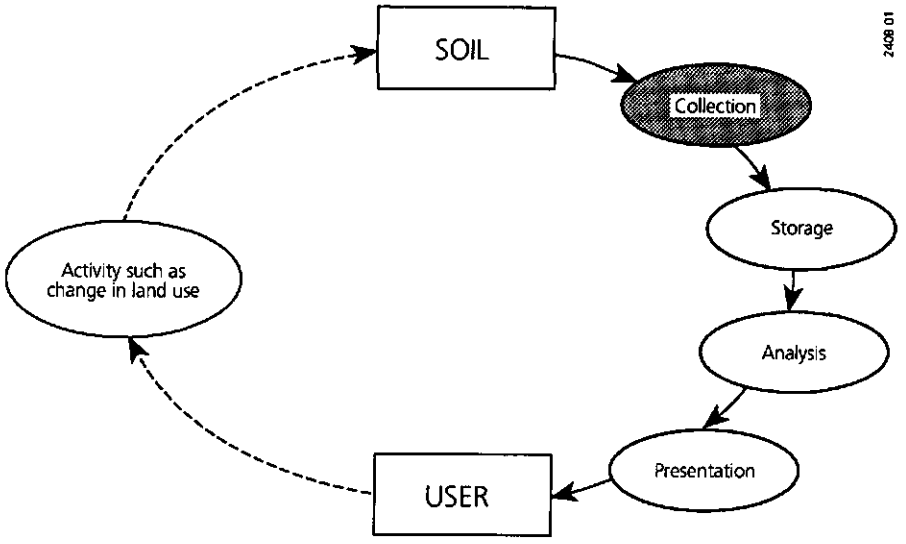
Wösten, J.H.M., 1990. Use of soil survey data to improve simulation of water movement in soils. PhD Thesis, Agricultural University, Wageningen.

Yeh, T.C., L.W. Gelhar & A.L. Gutjar, 1985. Stochastic analysis of unsaturated flow in heterogeneous soils, 3. Observations and applications. Water Resources Research 21: 465-471.

PART II

3 Optimum observation density for mapping acid sulphate soils in Conoco, Indonesia: accuracy and costs

Submitted for publication in Soil Use and Management



Optimum observation density for mapping acid sulphate soils in Conoco, Indonesia: accuracy and costs

A.K. BREGT[1], J.A.M. JANSSEN[1], J.S. van de GRIENDT[1], W. ANDRIESSE[1] and ALKASUMA[2]

1 *DLO The Winand Staring Centre for Integrated Land, Soil and Water Research (SC-DLO), P.O. Box 125, 6700 AC Wageningen, The Netherlands.*

2 *Centre for Soil and Agroclimate Research, Jalan Ir. H. Juanda 98, Bogor, Indonesia.*

ABSTRACT

In a study area of 410 ha (Conoco) located in an acid sulphate soil area in Indonesia, 820 observation points were made in a regular grid. In addition, 75 test observations were made. To determine an optimum observation density in terms of prediction accuracy and costs the variable "depth to the pyritic layer" was used. The number of observations was reduced step-wise by a random procedure in order to obtain lower observation densities. Predictions were made for the test observations using the prediction techniques kriging, inverse distance, local mean and global mean. Prediction errors were calculated for all the step-wise reduced observation densities. No significant differences were found between kriging, local mean and inverse distance. The performance of the global mean was worse than the other techniques. A density of 22 observations per square kilometer was selected as the optimum for the conditions in the Conoco grid area. This density is equivalent to a map scale of 1 : 30 000.

INTRODUCTION

Soil surveys are carried out to obtain information about the distribution of soil characteristics within a given area. These data are presented in the form of soil maps and reports. The soil surveyor uses landscape features and augerings to

characterize the soil. In some areas, landscape features such as landform, topography, vegetation, land use and hydrology are a good indication of the nature of the soil and how and where it changes. In other areas, particularly in tidal coastal plains as dealt with here, the soil pattern cannot easily be discerned from the landscape features. In these situations the soil surveyor has to rely heavily on his augerings for the characterization of the area. (Janssen et al., 1990)

If only field sampling is used for soil mapping, there is a direct relationship between observation density and map scale. This relationship is defined by a fixed number of observations per square cm of the ultimate map. As a rule of thumb Buringh et al. (1962) mentioned four observations and Dent (1986) one observation per square cm of the final map.

The choice of a particular map scale, and related observation density is mostly made a priori. The general expectation is that the higher the observation density the more detailed and accurate the information.

In most cases, however, there is no real basis for the assumption that the a priori choice is the most appropriate for an area. Under certain soil conditions a less dense observation net (implying lower costs) might still be appropriate for achieving a similar accuracy (Stein et al., 1988). In other situations, however, a denser observation net might be necessary for obtaining the desired information (Bos & Van Mensvoort, 1984).

In the study presented, we investigated the relationship between observation density and the accuracy of the obtained information for mapping depth to the pyritic layer. The associated costs were also taken into account. Local mean, global mean, inverse distance and kriging were used as spatial prediction techniques. The study was carried out in an area of 410 ha containing acid sulphate soils located in Pulau Petak, South Kalimantan, Indonesia.

MATERIALS AND METHODS

Soil survey

A detailed soil survey was carried out in the 410 ha sample area (Conoco) within Pulau Petak. Augerings were made in a regular grid of 75 * 75 m. Additional observations near some of the grid points were made at distances of 1, 5, and 25 m

respectively to get an impression of the short-range variation. All these observations will be referred to in this article as observation points. Besides these observation points, 75 extra augerings were made to serve as a test set. These points were randomly distributed over the area and are referred to in this article as test points.

At the observation and test points a number of soil characteristics was recorded and soil samples were taken for analysis in the laboratory. Analysis of the samples included the determining of potential and actual acidity (Konsten et al., 1988).

From the data collected (both field and laboratory data) information was derived such as the depth to the pyritic layer, the maximum total acidity within 50 cm depth, etc. In this study only the variable depth to the pyritic layer has been used.

Survey accuracy

In our study a complete relationship between map scale and observation density was assumed. For the purpose of our study a survey density of two observations per cm^2 map was used. Table 1 shows the number of observations and associated map scales for the Conoco area.

Table 1. Observation density ($obs.km^{-2}$), total number of observations and associated map scale for the Conoco study area (410 ha).

Observation density ($obs.km^{-2}$)	Total number of observations	Map scale
200	820	1:10 000
89	362	1:15 000
50	205	1:20 000
32	130	1:25 000
22	91	1:30 000
16	66	1:35 000
13	51	1:40 000
10	41	1:45 000
8	32	1:50 000

Insight into the effect of observation density on the accuracy of the acquired information was obtained by predicting the value of the test points from the observation points. Lower densities were simulated by reducing the number of observation points. The comprehensive procedure for the determination of the relation between observation density and survey accuracy is presented in Figure 1.

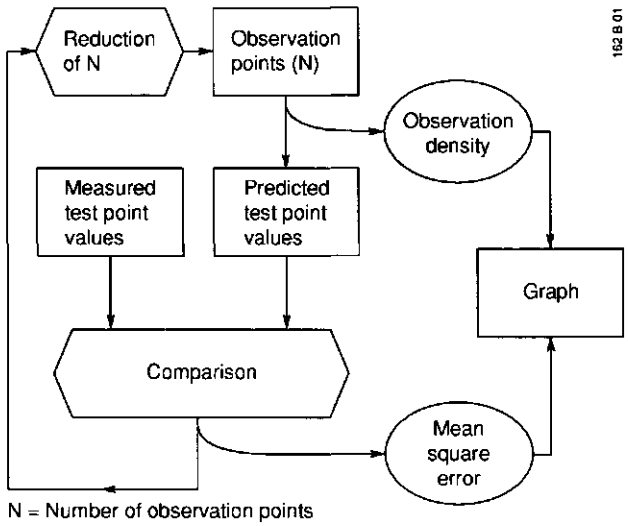


Fig. 1 Procedure for the determination of the relationship between observation density and prediction accuracy (MSE = mean square error).

The procedure necessitated choosing:
- the soil variables to be studied;
- the spatial prediction methods to be applied;
- the calculation of the prediction error and;
- the reductions of the observation points.

Soil variables

We selected the variable depth to the pyritic layer, as it is a significant variable in the characterization of acid sulphate soils, directly relevant to their agricultural use possibilities.

Spatial prediction

Spatial prediction is the estimation of the value of a variable at a specific location from values at other locations:

$$\hat{Z}_p = \sum_{i=1}^{n} W_i.Z_i \qquad (1)$$

where:

n = number of points involved in the prediction;
W_i = weighing factor of point i;
Z_i = the value of the variable at point i;
$\hat{Z}p$ = the predicted value of the variable at point p.

There are several spatial prediction techniques available (Lam, 1983; Burrough, 1986; Davis, 1986; Laslett et al., 1987). They differ in the number of observation points (n) used and in the way the weighing factor (W) is determined. We selected the local mean, global mean, inverse distance and kriging method for our study (Table 2). The semivariogram needed for the kriging methods was estimated from the 820 observation points.

Table 2. Characteristics of selected prediction methods.

Method	Number of points	Weighing factor used in prediction
Local mean	24	1
Global mean	all points	1
Inverse distance	24	1/distance*
Kriging	24	based on spatial structure*

* Sum of the weights is equal to one.

Prediction error

An indication of the accuracy of the spatial prediction can be obtained by comparing the estimated values with the measured values at the test points. The

Mean Square Error (MSE) is often used for this comparison (Puente & Bras, 1986; Stein et al., 1988; Bregt & Beemster, 1989; Voltz & Webster, 1990). It is calculated according to:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{Z}_i - Z_i)^2 \qquad (2)$$

where:

n = the number of test points;
$\hat{Z}_i$ = is the predicted value for a test point;
Z_i = the measured value for a test point.

Observation points

The number of observation points was reduced by randomly extracting points from the initial 820 observation points. To estimate the MSE accurately for each new density, the extraction of points and the prediction to test points were repeated 10 times.

Survey costs

Bie & Beckett (1971) and Dent & Young (1981) present general equations for the estimation of survey costs. They relate survey costs only to the effort of the soil surveyors. Soil sampling and analysis of the samples in the laboratory are omitted in their equations. In our study a cost equation was developed, more suited to our requirements. The costs of the survey for the different observation densities consisted of three main components:

- the first component (C_f) comprised costs that are independent from the observation density. They included the preparation for the field work, data processing and reporting;
- the second component (C_s) was the actual cost of the survey including the salaries and three field allowances for the survey teams (one team comprised two soil surveyors and one assistant) and cost of transport;
- the third component (C_a) referred to the costs for soil analyses in the laboratory.

The last two components were observation density dependent. The total costs

of a survey at a certain density were then calculated from:

$$\text{Total costs} = C_f + \left(\frac{N_0.H}{N_d}.C_s\right) + (N_p.N_s.C_a) \quad (3)$$

where:

C_f = fixed costs;
C_s = survey costs ;
C_a = analysis costs ;
N_o = the number of observations per hectare;
H = the size the area in hectares;
N_d = number of observations per day;
N_p = number of profiles sampled;
N_s = average number of samples per profile.

In our study, the following prices in Rupiah (Rp) were applied (price level 1989/1990; 1700 Rupiah = 1 U.S. Dollar): fixed costs (C_f) 600,000 Rp, survey costs (C_s) 150 000 Rp per team per day, analysis costs (C_a) 5,500 Rp per sample (Muhrizal Sarwani et al., 1990).

Table 3. Observation density, total number of observations and number of observations per day in the Conoco study area.

Observation density (obs.km^{-2})	Total number of observations per day	Number of observations
200	820	10
89	362	10
50	205	10
32	130	9
22	91	9
16	66	9
13	51	8
10	41	8
8	32	8

The number of observations (N_0.H) and the total number of observations per day for a given survey density (N_d) are presented in Table 3. The last naturally decreases with increasing distance between observation points. All profiles were sampled (N_p) in our study averaging 3.5 samples per profile (N_s).

Optimum observation density

The optimum observation density of soil surveys is essentially related to the requirements of the user. If a user is only interested in a general idea about the occurrence of acid sulphate soils in the world, then a soil map at a scale of 1 : 5 000 000 might be adequate for his needs. For local and regional planning much more detailed information is needed. For our purposes it was assumed that the survey data were needed for the identification of settlement areas. Survey scales suitable for this purpose are in the order of 1 : 10 000 to 1 : 50 000 (Dent & Young, 1981).

In this study the optimum observation density was defined as the lowest possible density, where the mean square error being minimal or not statistically different from the mean square error at higher densities. The differences between spatial prediction techniques at the same observation density were assessed with the Student test for paired observation (10 pairs in each comparison). To test the effects of observation density on prediction error for the same prediction technique the Student two sample test was used (10 observations per observation density) (Snedecor & Cochran, 1967). The last test was only performed with the kriging predictions.

RESULTS AND DISCUSSION

Accuracy of prediction

Predictions were made of the depth to the pyritic layer at test points using four different methods. The mean square error (MSE) values of these predictions are presented as a function of observation density in Figure 2.

All methods showed a decrease in MSE with increasing density. In most cases, the MSE for kriging was lower than the MSE for the other prediction methods.

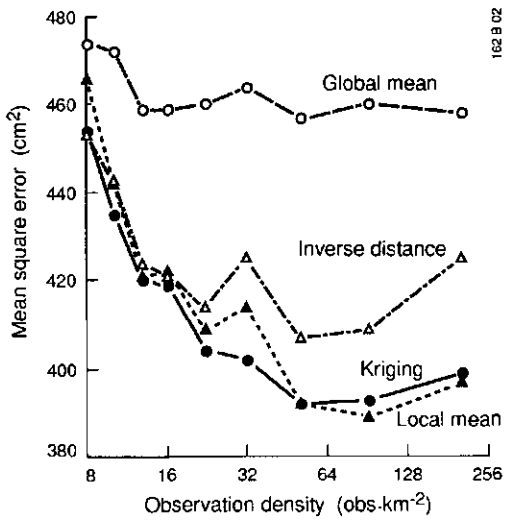


Fig. 2 The Mean Square Error of the four prediction techniques versus observations density (obs.km^{-2}) for Conoco.

Differences with local mean and inverse distance were small and not-significant ($\alpha = 0.05$) across the whole range of observation densities. Differences between kriging and global mean, however, were significant ($\alpha = 0.05$). Remarkably, the conceptually and operationally much more complicated kriging technique did not perform better than the simpler techniques of inverse distance and local mean. This can be explained by the high short-range variability of the variable depth to the pyritic layer. Due to this high short-range variation, the semivariogram (Fig. 3) showed a large 'nugget effect' (300 cm^2). (A semivariogram characterizes the spatial structure of the variable and it is needed for the kriging prediction technique). Besides, the slope of the semivariogram was quite flat. As a result of this type of semivariogram the weights in the spatial prediction (W_i in Eq. 1) of each point became almost equal, and kriging tended towards a local mean prediction. This can also be seen in Figure 2: the kriging method and the local mean method produced almost the same line. In another study area in Pulau Petak (Bregt et al., in press) also a nugget effect of 300 cm^2 was found. Similar large nugget effects of about 300 cm^2-400 cm^2 were also found by Bos and Van Mensvoort (1984) in the Mekong Delta, Vietnam. They calculated semivariograms for the depth to pyrite based on sampling in transects with a minimum spacing of 35 m. A nugget effect of about 300 cm^2 seems to be normal for the depth to the pyritic layer.

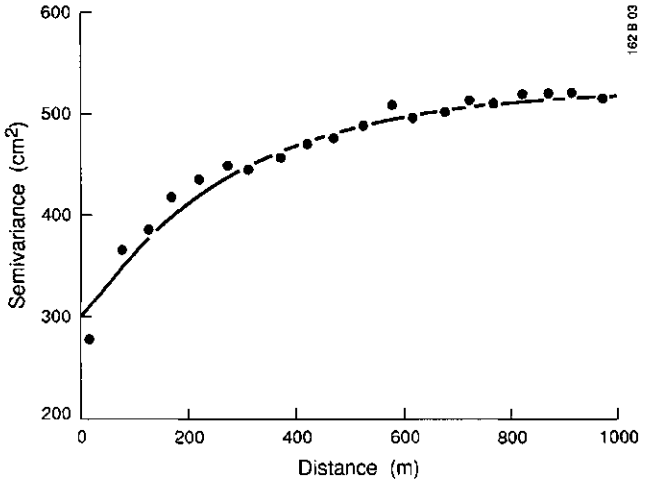


Fig. 3 Semivariogram of the depth to the pyritic layer in Conoco.

Costs

The total costs of the surveys at the various observation densities are presented in Figure 4.

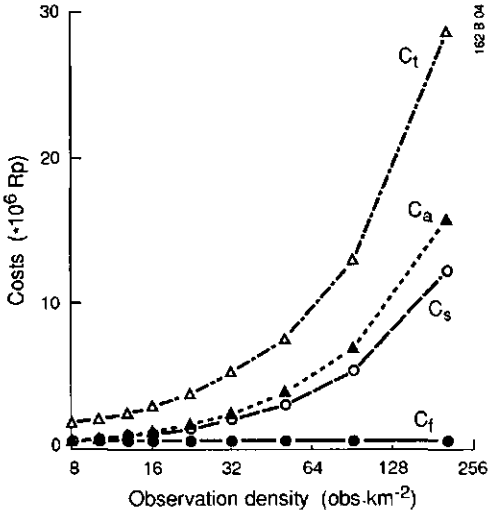


Fig. 4 Costs (in millions Rupiah) versus observation density (obs.km^{-2}) in Conoco. C_t = total costs; C_a = analysis costs; C_s = survey costs; C_f = fixed costs.

The total costs decreased from about 29 million Rp for a density of 200 observations per km^2 (map scale 1 : 10 000) to less than 2 million for a density of 8 observations per km^2 (map scale of 1 : 50 000). This is quite a dramatic decrease in costs.

In Figure 4 the three main cost components are presented separately. It shows that in this case the costs of the soil analysis in the laboratory C_a were slightly higher than the costs of field survey C_s.

Optimum observation density

The optimum observation density was determined by testing ($\alpha = 0.05$) whether the MSEs of the different densities were significantly higher than the MSE for the largest density. The kriging predictions were used solely for this purpose. According to the test, the density of 22 observations per km^2 (map scale 1 : 30 000) appeared to be the optimum for the survey of the variable depth to the pyritic layer. For this density the total costs were 3.9 million Rp, which is 87% less than the costs at a density of 200 observation per km^2 (map scale 1 : 10 000).

For densities of 22 and less, the MSE is significantly increased. It is clear that the optimum is dependent on the users definition. If another confidence level (α) had been chosen for the test, e.g. 0.1 or 0.001, perhaps the optimum density might have been different.

CONCLUSIONS

Results of the study show that there was no significant difference in accuracy between the kriging, local mean and the inverse distance techniques, in predicting the depth to the pyritic layer in the Conoco study area. The global mean proved to be worse than the other techniques.

The optimum density for mapping the depth to the pyritic layer was a density of 22 observations per km^2. At this density an accuracy equal to the accuracy obtained at an observation density of 200 observations per km^2 was achieved. The costs of the survey, however, were 87 % lower than in the last-mentioned case.

REFERENCES

Bie, S.W. & Beckett, P.H.T., 1971. Quality control in soil survey. II The costs of soil survey Journal of Soil Science 22: 453-465.

Bos, J. & Mensvoort, M.E.F. van 1984. Spatial variability of potential and actual acid sulphate soils in the Mekong Delta, Vietnam. Department of Soil Science and Geology, Wageningen, The Netherlands.

Bregt, A.K. & Beemster, J.G.R., 1989. Accuracy in predicting moisture deficits and changes in yield from soil maps. Geoderma 39: 281-291.

Bregt, A.K., Gesink, H.J. & Alkasuma, 1992. Mapping the conditional probability of soil variables. Geoderma (in press).

Buringh, P., Steur, G.G.L. & Vink, A.P.A., 1962. Some techniques and methods of soil survey in The Netherlands. Netherlands Journal of Agricultural Science 10: 157-178.

Burrough, P.A., 1986. Principles of geographical information systems for land resources assessment. Clarendon Press, Oxford.

Davis, J.C., 1986. Statistics and data analysis in geology. John Wiley & Sons, New York.

Dent, D.L., 1986. Acid sulphate soils: a baseline for research and development. Publication 39, ILRI, Wageningen, The Netherlands.

Dent, D. & Young, A., 1981. Soil survey and land evaluation. George Allen and Unwin, London.

Janssen, J.A.M., Hendro Prasetyo, Alkasuma & Andriesse, W., 1990. Landscape genesis and physiography of Pulau Petak as basis for soil mapping. LAWOO/AARD Scientific Report no. 18, ILRI, Wageningen, The Netherlands.

Konsten, C.J.M., Andriesse, W. & Brinkman, R., 1988. A field laboratory method to determine total potential and actual acidity in acid sulphate soils. In: Selected papers of the Dakar symposium on acid sulphate soils, (Ed. H. Dost), ILRI, Wageningen, The Netherlands, p. 106-134.

Lam, N.S., 1983. Spatial interpolation methods: a review. The American Cartographer 10: 129-149.

Laslett, G.M., McBratney, A.B., Pahl, P.J. & Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. Journal of Soil Science 38: 325-341.

Muhrizal Sarwani, Hairunsyah & Janssen, J.A.M., 1990. Management aspects of a field laboratory in support of soil surveys in coastal lowlands; an experience from Pulau Petak, Southern Kalimantan, Indonesia. LAWOO/AARD Scientific Report No. 31, ILRI, Wageningen, The Netherlands.

Puente, C.E. & Bras, R.L., 1986. Disjunctive kriging, universal kriging, or no kriging: small sample results with simulated fields. Mathematical Geology 18: 287-305.

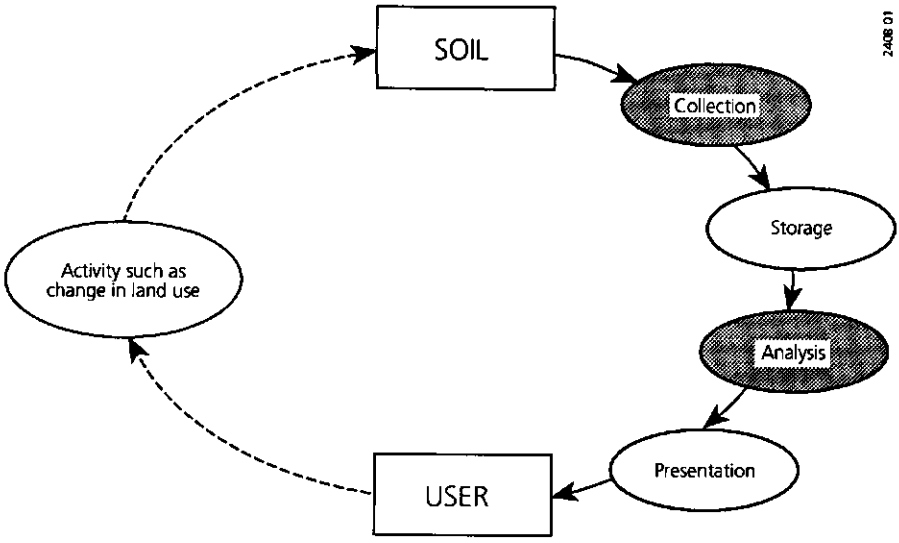
Snedecor, G.W. & Cochran, W.A., 1967. Statistical methods. The Iowa State University Press, Ames, Iowa.

Stein, A., Dooremolen, W. van, Bouma, J. & Bregt, A.K., 1988. Cokriging point data on moisture deficit. Soil Science Society of America Journal 52: 1418-1423.

Voltz, M. & Webster, R. 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science 41: 473-490.

4 Mapping ordinal data in soil survey: a Costa Rican example

Accepted for publication in Soil Science Society of America Journal



Mapping ordinal data in soil survey: a Costa Rican example

A.K. BREGT[1], J.J. STOORVOGEL[2], J. BOUMA[2] and A. STEIN[2]

[1] *DLO The Winand Staring Centre for Integrated Land, Soil and Water Research (SC-DLO), P.O. Box 125, 6700 AC Wageningen, The Netherlands.*

[2] *Dept. of Soil Science and Geology, Agricultural Univ. P.O. Box 37, 6700 AA Wageningen, the Netherlands.*

ABSTRACT

In most soil surveys, and particularly those carried out in developing countries, the bulk of recorded soil data is on an ordinal measurement scale. This determines the type of statistics and techniques for data processing that can be used. For example, with ordinal data the mode, median and range can be used as summary statistics. For describing the spatial structure of ordinal data we propose the spatial-difference-probability function, which is comparable with the semivariogram. In this article, the discussion about ordinal soil data is illustrated with data from a soil survey in Costa Rica. From the same study area three different suitability maps for banana were produced according to a qualitative land evaluation procedure by: (1) interpreting the soil map 1 : 200 000, (2) interpreting the soil map 1 : 50 000, and (3) interpolating point data. The quality of these three maps was tested by looking at the reliability, relevance and presentation of information, using 98 test borings. Reliability was characterized in terms of purity and range. The suitability map based on the soil map 1 : 200 000 was the most reliable one, with an overall purity of 49%. The suitability map produced by interpolating point data was the most relevant one, as defined in terms of the possibility to correctly identify potential locations for banana plantations from the map. Differences in presentation of information were evaluated by comparing the boundary indices of the different maps. The suitability map based on the soil map 1 : 50 000 was the most readable one. Map choice should be based on a consideration of the different quality criteria by the user.

INTRODUCTION

Soil surveys are carried out all over the world, and data obtained are increasingly being stored in geographical information systems. Many of these soil surveys have a general-purpose character and, as a consequence, a great variety of soil variables are recorded in the field by the soil mapper. Continuous soil variables, such as texture and stoniness, are often recorded in classes (qualitative data) as defined by the Soil Survey Manual (Soil Survey Staff, 1951). Neither funds nor time are usually available to make many measurements of soil variables at quantitative levels, such as clay content instead of texture class.

For the use of soil survey data in e.g. land use planning, environmental protection studies and agricultural planning, further processing and interpreting the basic soil survey data are necessary. In recent years, development and use of new processing and interpretation techniques have received much attention in soil science. Application of interpolation techniques and use of dynamic simulation models for quantitative land evaluation is being promoted by many researchers (Beek et al., 1987; Bouma and Bregt, 1989). Almost all mathematical interpolation techniques, however, assume the presence of quantitative data. Such techniques cannot be used when only qualitative soil data are available. Dynamic simulation of soil water regimes and associated crop development require quantitative data as well.

Land evaluation procedures such as those based on the framework for land evaluation (FAO, 1976), however, allow interpretation of qualitative data. Qualitative evaluations are useful to broadly identify areas of land that are relatively suitable for a particular type of land use. With increasing emphasis on quantitative simulation techniques, less attention is being paid in literature to such more qualitative interpretation procedures. This is unfortunate because the latter procedures are still very relevant in surveys where few quantitative data can be generated or where qualitative data are already available.

In this study some implications of having only qualitative data available for processing and interpreting are discussed. The spatial-difference-probability function is proposed for describing the spatial structure of qualitative variables. As an illustration, data from a soil survey in the Atlantic Zone of Costa Rica were used. A suitability classification for banana was carried out, following qualitative land evaluation procedures (FAO, 1976; Soto, 1985). Three different suitability maps were produced: (1) by interpreting the representative profiles of the mapping units of a soil map 1 : 200 000, (2) by interpreting the representative profiles of

the mapping units of a soil map 1 : 50 000, and (3) by interpolating point observations. The quality of the maps obtained was compared in terms of reliabilty, relevance and presentation.

THEORY

When processing recorded soil data, the scale on which the data have been measured is extremely important. Measurement is defined in this context as assigning values to variables. Therefore, both estimating texture in the field and determining it in the laboratory are considered to be measurements. Measurement scales are usually distinguished in terms of being qualitative or quantitative. Qualitative measurement involves naming or grouping of variables. Examples are horizon code and suitability class. Quantitative measurement involves assigning numerical values, which reflect magnitude or amount, to variables. Examples are values of organic-matter and water contents.

Stevens (1946) presented a subdivision of the qualitative scale in terms of nominal and ordinal and of the quantitative scale in terms of interval and ratio. The characteristics of these four measurement scales are summarized in Table 1.

Table 1. Characteristics of measurement scales.

Scale	Description	Useful summary statistics	
		Central tendency	Dispersion
nominal	determination of equality; data can be placed into classes	mode	-
ordinal	determination of more or less; data can be ranked	mode, median	range
interval	determination of equality of intervals or differences	mode, median, mean	range, standard deviation
ratio	determination of equality of ratios; measurements have a true zero	mode, median, mean	range, standard deviation, coefficient of variation

The nominal scale is the lowest one. It involves only the classification or naming of observations. Examples are land use classification and soil classification. The

ordinal scale involves the ranking of observations against each other. Values of observations can be placed in an ordered list, but the distance between the values is without meaning. Examples are soil suitability classes and drainage classes. With the interval scale differences between values can be established. The data do not contain an absolute zero and negative values are possible. Examples are soil temperature and change in yield. Ratios between values of interval variables do not have a real meaning. For instance, 20 °C is not twice as warm as 10 °C, since 0 °C is an arbitrary baseline. The highest measurement scale is the ratio scale. The data on this scale have an absolute origin, negative data values are not possible. Examples are data from laboratory analyses like texture and iron content. Ratios between values of ratio variables have a real meaning. For instance, 200 K is twice as warm as 100 K. Most quantitative soil data are on a ratio scale.

The scale of measurement of the data depends on the measurement method and the nature of the data. Some variables are ratio in nature, e.g. clay content, but they are recorded as ordinal variables due to the measurement equipment used. Less sophisticated measurement equipment generally yields data on a lower measurement scale, as distinguished in Table 1. The scale of measurement partly determines the type of statistics and processing methods to be used for describing and analysing data. For example, calculation of mean and standard deviation is only meaningful for interval and ratio variables. The coefficient of variation is only meaningful for ratio variables. For ordinal variables the mode, median and range are useful summary statistics (Table 1). More information about appropriate techniques for analyzing data of different measurement scales can be found in Shaw and Wheeler (1985).

For describing the spatial structure of soil variables the semivariogram is frequently used (Webster, 1985). It can be applied to both interval and ratio data. For describing the spatial structure of ordinal and nominal data we propose the spatial-difference-probability function (s-d-p(h):

$$s\text{-}d\text{-}p(h) = \frac{1}{N(h)} \sum_{j=1}^{N(h)} i\,(z(x_j),h): \qquad (4)$$

$i(z(x_j),h) = 0$ if $z(x_j)$ equals $z(x_j+h)$
$i(z(x_j),h) = 1$ if $z(x_j)$ unequals $z(x_j+h)$

Where N(h) is the number of pairs of points at distance h, $z(x_j)$ is the value of variable z at location x_j, and $z(x_j+h)$ is the value of variable z at a distance h from x_j.

The s-d-p function describes the probability of encountering different values between point observations at an increasing separation distance. It is estimated for every distance interval by comparing the value of pairs of observations at this distance interval. Like the semivariogram, this s-d-p function can be presented graphically. On the horizontal axis the distance is given, and on the vertical axis the estimated difference probability. The value of the difference probability ranges from 0 (all observations have the same value) to 1 (all observations have different values). The interpretation of this graph is comparable to the interpretation of a semivariogram. If a spatial structure is present, an increase in difference probability is expected with increasing distance between observation points.

For spatial interpolation of point data, several techniques are available. A review was given by Lam (1983) and Burrough (1986). Most techniques, such as kriging, splines, trend surface analysis, Fourier models, and distance-weighting methods, can only be used with interval and ratio data. Spatial interpolation with Thiessen polygons (Davis, 1986) can be applied to nominal and ordinal data.

Soil profile descriptions form the basis for currently available soil data. A widely used system for describing soil profiles is the one presented in the Soil Survey Manual (Soil Survey Staff, 1951). It has been adopted by soil surveys in numerous countries and many national soil survey handbooks have been derived from the Soil Survey Manual (Hodgson, 1978). In Table 2 the distribution of soil profile variables, as defined by the Soil Survey Manual, is presented for the different measurement scales. Most variables are on an ordinal scale. Only 7 of the 61 variables are interval or ratio. Most data obtained from laboratory analyses of soil samples are, however, ratio variables. Despite the large amount of ordinal variables in a profile description, most soil variability studies published so far consider only quantitative variables (Beckett and Webster, 1971; Gajem et al., 1981; Edmonds et al., 1982; Wang, 1982; Edmonds et al., 1985; Ogunkunle 1986). And when ordinal soil data are selected, they are sometimes incorrectly analysed. For instance, Wilding and Drees (1978) calculated coefficients of variation for an ordinal variable such as structure grade and even for a nominal variable such as structure class. Oliver and Webster (1987) calculated means and variances for several ordinal and even nominal soil variables. Agbu et al. (1990) calculated

coefficients of variation for the nominal variables landscape position and slope form.

Table 2. Distribution of soil and related variables over the four measurement scales.

Variables of:	Measurement scale	Number of variables	Variables
site description	nominal	5	landform, land use, climate, physiographic unit, position of site
	ordinal	1	topography
	interval	3	latitude, longitude, altitude
	ratio	0	-
general soil description	nominal	5	temperature regime, classification, moisture regime, microrelief, parent material
	ordinal	4	stoniness, rock outcrop, salt, drainage
	interval	1	water-table
	ratio	2	soil depth, rooting depth
profile description	nominal	15	horizon, structure type, kind of organic matter, type of mottles, color (hue), pores (form, orientation, continuity, distribution), cutans (kind, location), pans (kind, continuity, structure), kind of biological activity
	ordinal	24	texture, color (value, chroma), mottles (abundance, size, contrast, sharpness), structure (grade, size), consistence (dry, moist, wet), pores (size, quantity), porosity, roots (size, quantity), cutans (thickness), rock (size, quantity, degree of weathering), pans (cementation), biological activity (abundance), carbonates
	interval	1	depth
	ratio	0	-

MATERIALS AND METHODS

Study area and survey data

The study area (160 km^2) is located in the cantons Guapiles and Guacimo in the Limon province in Costa Rica. It is situated in the transition zone from the central mountains to the coastal plain. The mean annual temperature in the Limon province is 25 °C, and the mean annual precipitation is 4400 mm. The major land use types are extensive grazing and a few banana plantations. In parts of the area, crop production is limited due to surface stoniness, poor drainage conditions, and steep slopes.

Of the study area a soil map on a scale of 1 : 200 000 was made by interpreting aerial photographs and some field checks. The mapping units were characterized by a representative profile based on 2 or 3 borings per mapping unit. Of the same area also a soil map on a scale of 1 : 50 000 was made. The soils were classified according to Soil Taxonomy (Soil Survey Staff, 1975), mainly as Dystrandepts, Dystropepts and Troporthents. Mapping units of the soil map were delineated in the field by using soil borings and by interpreting landscape features. For each mapping unit a representative profile was defined by the soil mapper. For the 1 : 50 000 soil map 835 soil borings were made.

Land evaluation and map interpretation

One of the objectives of the soil survey was producing a suitability map for bananas, which is the most important crop in the area. The land evaluation procedure used, was locally developed by Soto (1985). By this procedure the suitability is based on eight separate evaluations of ordinal soil and land characteristics. The soil or land characteristic in the lowest suitability class determines the final suitability. Land evaluation criteria are presented in Table 3.

Suitability maps of the area were made by three different procedures. The first map was produced by interpreting the 1 : 200 000 soil map. The data of the representative profiles associated with the mapping units were evaluated and a suitability map was produced (map I). The second map was produced by interpreting the soil map 1 : 50 000. The representative profiles of the soil units were eva-

luated, and subsequently a suitability map was produced (map II). The third map

Table 3. Land evaluation criteria for bananas (Soto, 1985).

Soil and land characteristics	Highly suitable 1	Moderately suitable 2	Marginally suitable 3	Marginally not suitable 4	Not suitable 5
Soil depth (cm)	> 120	90-120	60-90	30-60	0-60
Texture [1]) at 15 cm	3, 4, 5	6, 7, 8, 9	11	2, 10, 12	1
at 45 cm	3, 4, 5	6, 7, 8, 9	11	1, 2, 10, 12	-
at 75 cm	3, 4, 5, 6, 7, 8, 9	11	1, 2, 10, 12	-	-
at 105 cm	3, 4, 5, 6, 7, 8, 9	1, 2, 10, 11, 12	-	-	-
Stoniness [2])	0	1	2	3	≥ 4
Slope [3])	0, 1	-	2	-	≥ 3
Drainage [4])	4	3, 5	2	1, 6	-

[1]) Texture 1: sand; 2: loamy sand; 3: sandy loam; 4: loam; 5: silty loam; 6: silt; 7: sandy clay loam; 8: clay loam; 9: silty clay loam; 10: sandy clay; 11: silty clay; 12: clay

[2]) Stoniness 0: 0%; 1: 0-0.01%; 2: 0.01-0.1%; 3: 0.1-3%; 4: > 3%

[3]) Slope 0: 0%; 1: 0-2%; 2: 2-5%; 3: > 5%

[4]) Drainage 1: poorly drained; 2: somewhat poorly drained; 3: moderately well drained; 4: well drained; 5: somewhat excessively drained; 6: excessively drained

was produced by interpolating point data. For all the individual soil borings the suitability was determined. Ninety-eight borings were selected at random from this data set to serve as a test set. The remaining 737 borings were transformed into a suitabilty map (map III) by interpolation via Thiessen polygons. Maps I and II were produced according to standard interpretation procedures in soil survey and land evaluation. Map I (scale 1 : 200 000) is meant for national planning. Maps II and III are meant for regional planning and identification of potential project areas (scale 1 : 50 000). The total costs for producing maps II and III are more or less the same. The total costs for producing map I are estimated to be a factor 16 lower than the costs for maps II and III.

Quality of the maps

According to Western (1978), the quality of a map is a function of (i) the reliability of the information; (ii) the relevance of the information, and (iii) the presentation of the information.

The **reliability** of the information on a map is characterized by the purity and the homogeneity (Beckett and Webster, 1971; Bie and Beckett, 1973; Marsman and De Gruijter, 1986; Bregt et al., 1987). The purity indicates the degree to which the (suitability) classes, as indicated on the map, agree with the suitability of locations in the field. It was estimated in this study by comparing the values for mapping units with the corresponding values of the 98 test borings within these mapping units. Purities were calculated for each individual suitability class and for the map as a whole. The homogeneity indicates how homogeneous the mapping units are with respect to the suitability class. For interval and ratio data the standard deviation within mapping units can be used as a measure for the homogeneity. For ordinal data the standard deviation cannot be used, but the range provides a good measure for homogeneity. The **relevance** of information as a quality component is very difficult to quantify for a general-purpose soil map. Quite often, only some qualitative indications can be given about the relevance of the information. In the case of a special-purpose map, quantification is easier, since we can test whether the map meets the demand. The objective of our map was to identify potential locations for banana plantations. In this study, we assumed that their establishment is potentially feasible in areas with suitability classes 1 and 2. Detailed investigations must reveal the right locations. The map which identifies potential areas correctly and which, on the other hand, correctly eliminates unsuitable areas is considered to be a relevant one. Four situations can be distinguished:

(1) the map indicates "suitable", and the area is "suitable" in reality;
(2) the map indicates "suitable", but the area is "not suitable" in reality;
(3) the map indicates "not suitable", but the area is "suitable" in reality;
(4) the map indicates "not suitable", and the area is "not suitable" in reality.

Situations one and four represent correct decisions. Situations two and three represent wrong decisions. Situation three is the worst one, because a suitable area will not be identified as such. In the case of situation two the potentially suitable area will be declared not suitable after detailed investigations. The map with the lowest

values for situations two and three is the most relevant one. The magnitude of the four situations were characterized for the three maps by using the 98 test borings.

The quality of the information **presentation** is a function of the cartographic methods used and of the complexity of the map pattern. As indicated by Monmonier (1974) and Dent (1985), the complexity of the map pattern has a strong influence on the readability of the map. Various complexity measures were compared by Bregt and Wopereis (1990). From the measures presented, the boundary index (Bregt et al., 1989) was used in this study to evaluate different types of cartographic presentations in a quantitative manner. The readability of the map increases if the boundary index decreases.

RESULTS AND DISCUSSION

Spatial structure

In Fig. 1 the s-d-p graph for suitability classes is presented which is based on all borings. The estimated difference probability is almost constant with increasing distance, implying no spatial structure within the investigated distance range. The estimated difference probability of the shortest distance interval (250 m) is high (0.65), indicating a high short-range variability. A possible explanation for this high short-range variability can be found in the land evaluation procedure used. By this procedure eight separate evaluations of soil and land characteristics were performed (Table 3). The soil or land characteristic in the lowest suitability class determines the final suitability.

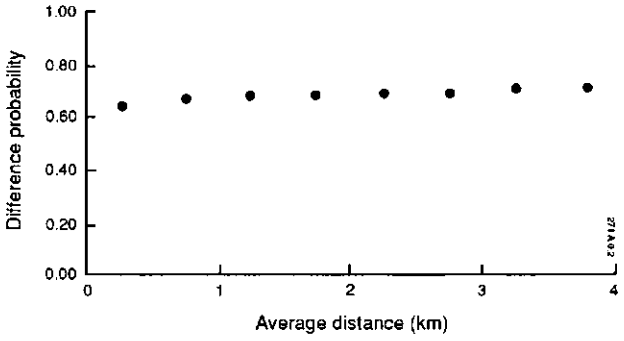


Fig. 1 Spatial-difference probability graph for suitability classes based on all borings.

In Fig. 2 the s-d-p graph for some soil and land characteristics are presented which are used in the land evaluation procedure. All the variables show a slight increase in difference probability at greater distance, indicating that observations farther away differ slightly more than nearby observations. The short-range variation of all the variables is quite large. For example, the probability of finding a different texture class at 15 cm depth at a distance of 250 m is more than 50%.

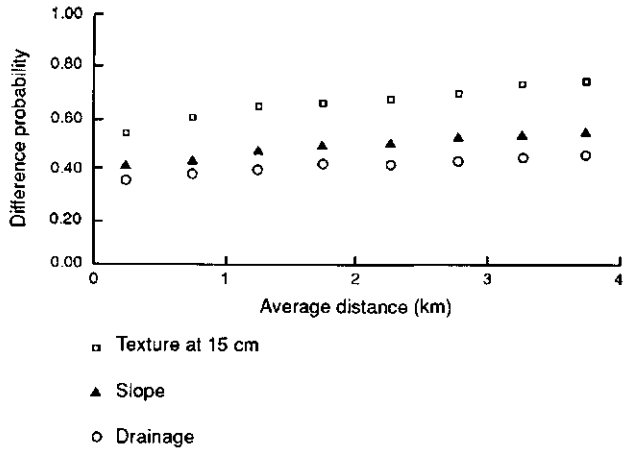


Fig. 2 Spatial-difference probability graph for some soil and land characteristics based on all borings.

Suitability maps

The three procedures resulted in quite different maps. Map I, based on the soil map 1 : 200 000, is dominated strongly by suitability class 3, which occupies about 60% of the map. It contains 5 different mapping units, and 17 mapping areas. Map II, based on the soil map 1 : 50 000, contains 5 different mapping units and 56 mapping areas. As a result of the Thiessen interpolation of the borings, map III has quite a "blocky structure". It contains 5 different mapping units and 189 mapping areas. Fragments of the three maps produced are presented in Fig. 3. In Table 4 some summary statistics for each mapping unit of the three maps are presented. These statistics were calculated by using the test data within the mapping units. For all the classes of map II the mode class, based on the test borings, does not correspond with the suitability class of the map. For map I this is the case for mapping units 2 and 5, and for map III for the units 1 and 5. In these situations

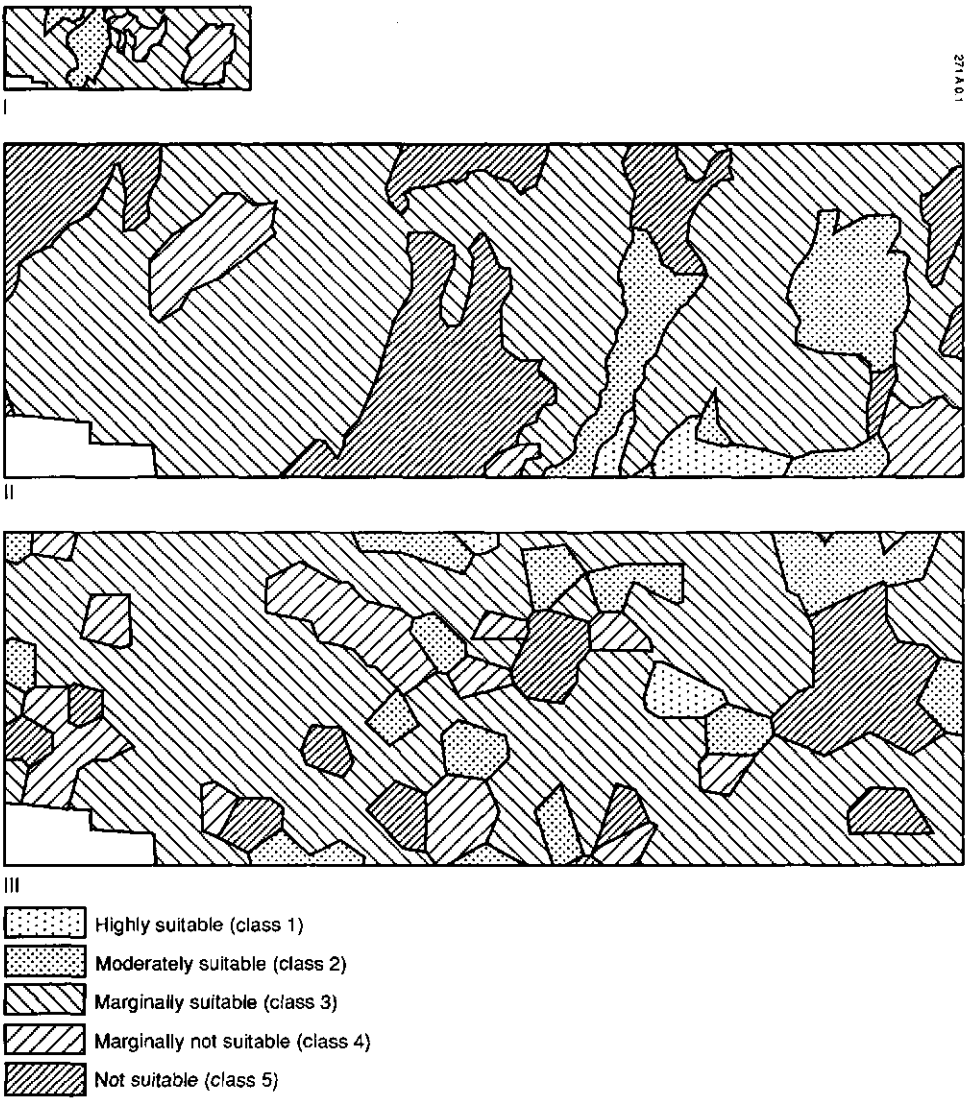


Fig. 3 Fragments of the three different suitability maps for bananas.

I) Suitability map derived from soil map 1 : 200 000

II) Suitability map derived from soil map 1 : 50 000

III) Suitability map produced by interpolation of point data, schale 1 : 50 000

the suitability class of the mapping unit is not supported by the borings within the mapping unit. For almost all the units of the three maps the mode and median class are the same.

Table 4. Summary statistics for mapping units of the three suitability maps based on suitabilities for 98 test borings.

Summary statistics	Land unit map (map I)*					Soil map (map II)*					Point observations (map III)*				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Number of borings	1	8	59	18	12	4	7	42	11	34	3	7	50	23	15
Mode	1	4	3	4	4	3	4	4	3	3	2	2	3	4	4
Median	1	3	3	4	4	3	4	4	3	3	2	3	3	4	4
Range	1	2-4	1-5	3-5	4-5	1-4	3-5	2-5	1-4	2-5	2-3	2-5	1-5	1-5	2-5

*) Classes 1: highly suitable; 2: moderately suitable; 3: marginally suitable; 4: marginally not suitable; 5: not suitable

Comparing the quality of the three maps

The **reliability** of information on all the maps was characterized by the purity. In Table 5 both the purity for the individual mapping units and for the whole map are presented. The purities of all the mapping units are quite low. Extremely low purities were found for the mapping units 2, 4 and 5 of map II, and unit 1 of map III. The overall purity of map I has the largest value (49%). An extremely low overall purity (15%) was found for map II, the map based on the soil map 1 : 50 000. An explanation for the low purities can be found in the high short-range variability as indicated by the s-d-p function. Based on these results we conclude that the information on map I, the map based on the soil map 1 : 200 000, is more reliable than the information on maps II and III. This is surprising. A possible explanation might be the absence of a clear spatial structure in the data, as indicated by the s-d-p function, which made mapping of the data difficult. The homogeneity of the mapping units, indicated by the range (Table 4), does not show a clear difference between the three maps; all units are quite heterogeneous.

Table 5. Purity of individual mapping units for map I, II and III and overall purity for maps I, II and III.

Purity (%)																		
Mapping units map I*						Mapping units map II*						Mapping units map III*						Map I	Map II	Map III
1	2	3	4	5		1	2	3	4	5		1	2	3	4	5				
100	13	51	72	25		25	0	26	0	9		0	42	48	39	20		49	15	40

*) Classes 1: highly suitable; 2: moderately suitable; 3: marginally suitable; 4: marginally not suitable; 5: not suitable

The **relevance** of information on a map was characterized by the magnitude of areas incorrectly classified as "suitable" and "not suitable". The results are presented in Table 6. From Table 6 we see that the percentages "suitable" on the map and "not suitable" in reality, and "not suitable" on the map and "suitable" in reality are the smallest for map III, 5% and 10% respectively. This means that map III, produced by interpolating point data, is the most appropriate one for selecting potential sites for banana plantations. For all three maps the percentage of suitable areas not identified by the map is large compared with the area correctly identified as suitable.

Table 6. Identification of comparison between suitability on map and suitability in reality.

Map	Identification (%)			
	Suitable on map & suitable in reality	Suitable on map & not suitable in reality	Not suitable on map & suitable in reality	Not suitable on map & not suitable in reality
I	2	7	13	78
II	1	10	14	75
III	5	5	10	80

The **presentation** of information was evaluated by looking only at the boundary index of the maps. Another important factor in the quality of presentation of information, namely the cartographic methods used, did not differ between the maps. The boundary index is 1.28 cm^{-1} (cm/cm^2) for map I, 0.75 cm^{-1} for map II, and 1.12 cm^{-1} for map III. With respect to the presentation of information, map II, the one based on the soil map 1 : 50 000, is the best.

CONCLUSIONS AND RECOMMENDATIONS

Many soil variables are on an ordinal measurement scale. Commonly used statistics, such as mean, standard deviation and semivariance, and most spatial interpolation techniques are not useful for this type of data. When discussing future developments in the field of soil data processing, it is useful to focus not only on which data could or should be available, but also on which data are available. Quite often due to limited resources only ordinal soil data can be collected. Developing useful processing techniques for ordinal soil data should be encouraged. For example, in the area of spatial interpolation the usefulness of indicator kriging (Journel, 1983) needs to be investigated. Soil scientists have to realize, on the other hand, that most data processing techniques are available or are being developed for quantitative data. Within a standard soil survey, more quantitative data should be collected as is done at this moment.

In the case study presented, the quality of three different suitability maps based on ordinal data was evaluated. The suitabilty map based on the soil map 1 : 200 000 was the most reliable one. The suitability map produced by interpolating point data contained the most relevant information. The suitabilty map based on the soil map 1 : 50 000, however, was the most readable one. The choice to be made for a particular map depends on the relative importance which is assigned to the various quality criteria by a particular user.

ACKNOWLEDGEMENT

Soil survey data for this study were obtained in Costa Rica by the second author, working with E. Spaans and H. van Leeuwen under the guidance of

W.G. Wielemaker. The soil survey was part of a joint research programme between CATIE (Centro Agronómico Tropical de Investigación y Enseñanza), MAG (Ministerio de Agricultura y Ganadería) and the Agricultural University Wageningen, the Netherlands.

REFERENCES

Agbu, P.A., D.J. Fehrenbacher & I.J. Jansen, 1990. Soil property relationships with SPOT satellite digital data in East Central Illinois. Soil Sci. Soc. Am. J. 54: 807-812.

Beek, K.J., P.A. Burrough & D.E. McCormack (Ed.), 1987 Quantified land evaluation procedures. ITC Publ. 6. ITC, Enschede.

Beckett, P.H.T. & R. Webster, 1971. Soil variability: a review. Soils and fertilizers 34: 1-15.

Bie, S.W. & P.H.T. Beckett, 1973. Comparison of four independent soil surveys by airphoto interpretation, Paphos Area (Cyprus). Photogrammetria 29: 189-202.

Bouma, J. & A.K. Bregt (Eds.), 1989. Land qualities in space and time. Proc. ISSS sympos., Wageningen, the Netherlands, 22-26 August 1988. Pudoc, Wageningen.

Bregt, A.K. & M.C.S. Wopereis, 1990. Comparison of complexity measures for choropleth maps. Cartogr. J. 27: 85-91.

Bregt, A.K., J. Bouma & M. Jellinek, 1987. Comparison of thematic maps derived from a soil map and from kriging of point data. Geoderma 39: 281-291.

Bregt, A.K., L.A. Ankum, J. Denneboom & Y. van Randen, 1989. Gridded soil map of the Netherlands: gridding error. p. 110-117. In R.J.A. Jones and B. Biagi (Eds), Computerisation of land use data. European Communities, Luxembourg.

Burrough, P.A., 1986. Principles of geographical information systems for land resources assessment. Clarendon Press, Oxford.

Davis, J.C., 1986. Statistics and data analysis in geology. John Wiley & Sons, New York.

Dent, B.D., 1985. Principles of thematic map design. Addison-Wesley, Reading.

Edmonds, W.J., S.S. Iyengar, L.W. Zelanzy, M. Lentner & C.D. Peacock, 1982. Variability in family differentia of soils in a second-order soil survey mapping unit. Soil Sci. Soc. Am. J. 46: 88-93.

Edmonds, W.J., J.C. Baker & T.W. Simpson, 1985. Variance and scale influences on classifying and intrepreting soil map units. Soil Sci. Soc. Am. J. 49: 957-961.

FAO, 1976. A framework for land evaluation. Soils Bull. 32. FAO, Rome.

Gajem, Y.M., A.W. Warrick & D.E. Myers, 1981. Spatial dependence of physical properties of a typic torrifluvent soil. Soil Sci. Soc. Am. J. 45: 709-715.

Hodgson, J.M., 1978. Soil sampling and soil description. Clarendon Press, Oxford.

Journel, A.G., 1983. Nonparametric estimation of Spatial distributions. Mathematical Geology 15: 445-468.

Lam, N.S., 1983. Spatial interpolation methods: a review. The Am. Cartographer 10: 129-149.

Marsman, B.A. & J.J. de Gruijter, 1986. Quality of soil maps: A comparison of survey methods in a sandy area. Soil survey papers No. 15. Stiboka, Wageningen.

Monmonier, M.S., 1974. Measures of pattern complexity for choroplethic maps. The Am. Cartographer 9: 159-169

Ogunkunle, A.O., 1986. Spatial variability of some chemical properties in two ultisol mapping units in southern Nigeria. Soil survey and land evaluation 6: 26-32.

Oliver, M.A. & R. Webster, 1987. The elucidation of soil pattern in the Wyre Forest of the West Midlands, England. I. Multivariate distribution. J. of Soil Sci. 38: 279-291.

Shaw, G. & D. Wheeler, 1985. Statistical techniques in geographical analysis. John Wiley & Sons, Chichester.

Soil Survey Staff, 1951. Soil survey manual. USDA Agric. Handb. 18. Washington, DC.

Soil Survey Staff, 1975. Soil Taxonomy: a basic system of soil classification for making and interpreting soil surveys. USDA-SCS Agric. Handb. 436. U.S. Gov. Print Office, Washington, DC.

Soto, B.M., 1985. Bananas, cultivo y comercialización. Litografía e Imprenta LIL, S.A. San José, Costa Rica.

Stevens, S.S., 1946. On the theory of scales of measurement. Science 103: 677-680.

Wang, C., 1982. Variability of soil properties in relation to size of map unit delineation. Can. J. Soil Sci. 62: 657-662.

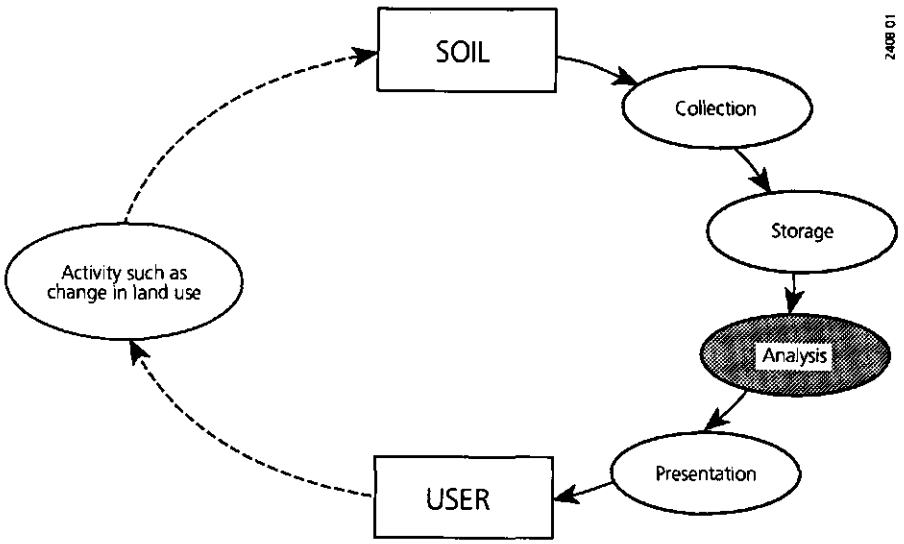
Webster, R., 1985. Quantitative spatial analysis of soil in the field. p. 1-70. In B.A. Stewart (Ed.) Advances in soil science, Vol. 3. Springer-Verslag, New York.

Western, S., 1978. Soil survey contracts and quality control. Clarendon Press, Oxford.

Wilding, L.P. & L.R. Drees, 1978. Spatial variability: a pedologist's viewpoint. p. 1-12. In Diversity of soils in the tropics. ASA Spec. Publ. 34 ASA and SSSA, Madison, WI.

5 Accuracy in predicting moisture deficits and changes in yield from soil maps

Published in Geoderma, 43: 301-310 (1989)



Accuracy in Predicting Moisture Deficits and Changes in Yield from Soil Maps

A.K. BREGT and J.G.R. BEEMSTER

Soil Survey Institute, P.O. Box 98, 6700 AB Wageningen (The Netherlands)

(Received March 14, 1988; accepted after revision September 1, 1988)

ABSTRACT

Bregt, A.K. and Beemster, J.G.R., 1989. Accuracy in predicting moisture deficits and changes in yield from soil maps. Geoderma, 43: 301–310.

In a survey area of 1435 ha, water-tables have been lowered as a result of water extraction for drinking water supplies. We simulated the change in grass yield in this area by comparing the former and present hydrological situations. A soil map on a scale of 1:10,000 of the area was used to derive soil maps on scales of 1:25,000 and 1:50,000. Representative profile descriptions from the mapping units of the three soil maps were physically interpreted and were used thereafter as a basis for simulating changes in yield.

We wanted to know whether the quality of representative profile descriptions would permit correct predictions for the whole area and for specific locations within the area. Therefore, we compared the simulated values of the representative profile descriptions with simulated values of individual borings. Two attributes were investigated: the simulated change in yield due to extraction of water and the simulated average moisture deficit in a 30-year period under the present hydrological situation (after water extraction).

For the whole area, the mean error was used as a quality measure. No differences in quality were found among representative profile descriptions for the three map scales. As quality measures for point predictions, the mean absolute error and the mean square error were used. The calculated errors are quite large. For instance, the error in the estimation of the moisture deficit is 50% or more for all three map scales. Spatial variability within a mapping unit forms a large proportion of this error. The best predictions at point locations were obtained with the representative profile descriptions for the 1:10,000 map.

INTRODUCTION

Soil maps are used and interpreted for a variety of purposes (Western, 1978). Soil surveys also form an important data source for land evaluation (Young, 1980; Bouma et al., 1986; Beek et al., 1987). Traditionally, soil survey data are mainly interpreted for agricultural purposes. They are, however, being used increasingly for environmental and other interpretations as well. Breeuwsma et al. (1986) present examples of the use of soil maps for the estimation of

travel times of water in the unsaturated zone and of phosphate sorption capacity. Other examples of soil map interpretations in The Netherlands were reported by Wösten et al. (1987).

Interpretations of soil maps are generally based on descriptions of representative profiles made by soil surveyors for mapping units. The quality of these interpretations depends entirely on the quality of the representative profile descriptions. A faulty representative profile description will automatically lead to wrong interpretations. Therefore, knowledge of the quality of these descriptions is important. In this study, we investigated the quality of representative profile descriptions to see if they permitted correct predictions for an entire area or for specific locations within that area. The influence of map scale on the quality of the predictions was studied as well.

The soil map interpretations were focused on predicting the sensitivity of soils to lowering of the water-table as a result of the extraction of water for domestic use.

A previous study in the same area was carried out by Wösten et al. (1987). They investigated the influence of map scale on predictions of the changes in yield for the entire area and for individual fields. They concluded that for the area as a whole the average change in yield could be predicted correctly with the 1:50,000 map. Changes in yield for specific fields, however, could be predicted most accurately from maps with scales of 1:10,000. In their study, the representative profile descriptions for the 1:10,000 map were used as references for the less detailed scales. Thus, no information about the 'real' quality of representative profile descriptions was obtained. In our study, the interpretations of representative profile descriptions were compared with the interpretations of data from borings. Thus, a more realistic estimate of the quality of the representative profile descriptions for interpretations was obtained.

MATERIALS AND METHODS

Study area

Known as Mander, the study area of 1435 ha lies near Tubbergen in the eastern part of The Netherlands. A sample area of 404 ha from this study area was considered in our investigation. The principal soils of the sample area are classified according to Soil Taxonomy (Soil Survey Staff, 1975) as Typic Haplaquods, Typic Humaquepts and Plaggepts. These soils have been formed in a regolith consisting of aeolian sands and glacial drift, which rests in some parts on Tertiary clay. Depth to the Tertiary clay is 150 cm in the eastern part of the sample area.

TABLE I

Number of mapping units used in the analysis and average number of borings per mapping unit for the three scales

Scale	Number of borings	Number of mapping units	Average area of mapping unit (ha)	Average number of borings per mapping unit
1: 10,000	410	83	3.3	4.9
1: 25,000	410	70	4.0	5.9
1: 50,000	410	41	6.8	10.0

Soil survey

The soils were mapped at a scale of 1:10,000, with an average observation density of 1.5 borings per hectare (Stoffelsen and Van Holst, 1985). In the whole area 2150 borings were made. The locations of the borings were selected by a soil surveyor. Borings extended to the upper surface of the Tertiary clay or to the mean lowest water-table (MLW), with a maximum depth of 3.2 m.

For each mapping unit at the scale of 1:10,000 the soil surveyor prepared a representative profile description (RPD). This description contains the sequence of horizons in the mapping unit and the average value and range of properties within each horizon. The description is presented in the same form as for a profile. The RPD is an average characterization of the mapping unit and not of an existing profile. It is compiled by the soil surveyor on the basis of his experience and on data for borings in the mapping unit.

From the soil map at a scale of 1:10,000 and its associated RPD's, a map at a scale of 1:25,000 with RPD's was derived through generalization without new fieldwork. In the same way, a soil map at a scale of 1:50,000 with RPD's was derived from the soil map at a scale of 1:25,000. Only one experienced soil surveyor was involved in delineating map units and in preparing representative profile descriptions. Thus, errors due to differences among surveyors did not occur.

For statistical analysis of the data, only units of the 1:10,000 map with at least three borings were used. This meant that the effective study area was reduced from 404 ha to 278 ha. To permit comparison of the three maps, the same effective study area (278 ha) was used for the 1:25,000 and the 1:50,000 map (Table I).

Interpretations

The soil map interpretations in this study were focused on the sensitivity of soils to drawdown of the water-table. The sensitivity was calculated with a simulation model developed by De Laat (1980).

As a first step, physical properties were assigned to individual horizons recorded in the representative profile descriptions (RPD's). In order to do so use was made of a translation set derived for soils in The Netherlands, which relates horizons to soil-physical properties (Wösten et al., 1987). Soil-physical properties in this context are water retention and hydraulic conductivity characteristics. If adjacent horizons had the same physical properties, they were combined into one physical unit layer. These physical unit layers were used later in calculating water balances of the soils.

To test the quality of each RPD for interpretations, the procedure for establishing physical unit layers was applied to all of the borings made within a map unit.

The next step was to calculate with the simulation model the water balance of RPD's and borings for successive 10-day periods. The calculations were made for the period of 1956–1985. We assumed that the soils were under grass vegetation for a growing season of 180 days, viz., from April 1 to October 1. Moisture deficits were calculated for two situations, one before and one after water extraction. These deficits were then transformed into relative yields on the basis of the relation between yield reduction per mm moisture deficit and potential production, as reported by Van Bohemen (1981). The difference between relative yields in the former and present hydrological situations forms the yield reduction due to increased moisture deficit. Water extraction may, however, also lead to increases in yield. This occurs in soils that were initially too wet. Such changes were also calculated with standard tables relating the increase in yield to the water-table fluctuations.

Final results are presented as an average, net change in yield per hectare, expressed in terms of a yearly percentage, due to drawdown of the water-table. Negative percentages indicate a net decrease in yield due to a lower water-table during the growing season. Positive percentages indicate a net increase in yield due to a lower water-table.

In this study, two simulated attributes were selected for further investigation, namely the change in yield (% ha^{-1} yr^{-1}) due to extraction and the average moisture deficit (mm) in a 30-year period under the present hydrological situation, which we refer to as the change in yield and the moisture deficit.

Using the same model, other authors have validated the procedure described for a clay area (Bouma and De Laat, 1981) and for a sandy area (Wösten et al., 1985).

STATISTICAL ANALYSIS

The quality of a representative profile description (RPD) was estimated by comparing its value for a mapping unit with the values for the individual borings in that mapping unit.

As a quality measure of the RPD for estimating an average value for the

total study area the mean error (ME) was used. The ME is of particular interest for agencies interested in results for the total study area, e.g. water extraction companies and provincial water boards. They want to know the quality of the map for predicting a correct average value for the entire area. The ME was calculated according to:

$$ME = \frac{1}{n} \sum_{u=1}^{m} \sum_{i=1}^{k} (rpd_u - b_{ui}) \quad (1)$$

where: ME = mean error; rpd_u = value of representative profile of uth mapping unit; b_{ui} = value of ith boring in the uth mapping unit; n = total number of borings; m = number of mapping units; k = number of borings in the uth mapping unit.

The ME can be zero, positive or negative. In the case of a zero ME the average prediction of the variable by the RPD for the entire area is perfect. A positive ME means an overestimation of the variable by the RPD. A negative ME means an underestimation of the variable by the RPD.

As a quality measure of the RPD for estimating values at specific locations within the area, the mean absolute error (MAE) and the mean square error (MSE) were used. The value of MSE is more influenced by outliers in the data set. The MAE and MSE are of particular interest to the local user, e.g. the farmer. The MAE was calculated according to:

$$MAE = \frac{1}{n} \sum_{u=1}^{m} \sum_{i=1}^{k} |rpd_u - b_{ui}| \quad (2)$$

where MAE = mean absolute error and the other symbols are the same as in eq. 1.

The MSE was calculated according to:

$$MSE = \frac{1}{n} \sum_{u=1}^{m} \sum_{i=1}^{k} (rpd_u - b_{ui})^2 \quad (3)$$

where MSE = mean square error and the other symbols are the same as in eq. 1.

Either large variations within mapping units or high errors in the representative profile descriptions will lead to high MAE and MSE values (Fig. 1). Separation of these two effects is meaningful. This can be done by splitting the MSE into two parts (Bunke and Droge, 1984; Kempthorne and Allmaras, 1986):

$$MSE = MSE_s + MSE_r \quad (4)$$

where: MSE_s = contribution in mean square error due to systematic errors; MSE_r = contribution in mean square error due to random errors.

The systematic error (MSE_s) is caused by the difference between the value of the RPD and the mean value of data from borings in the mapping unit and

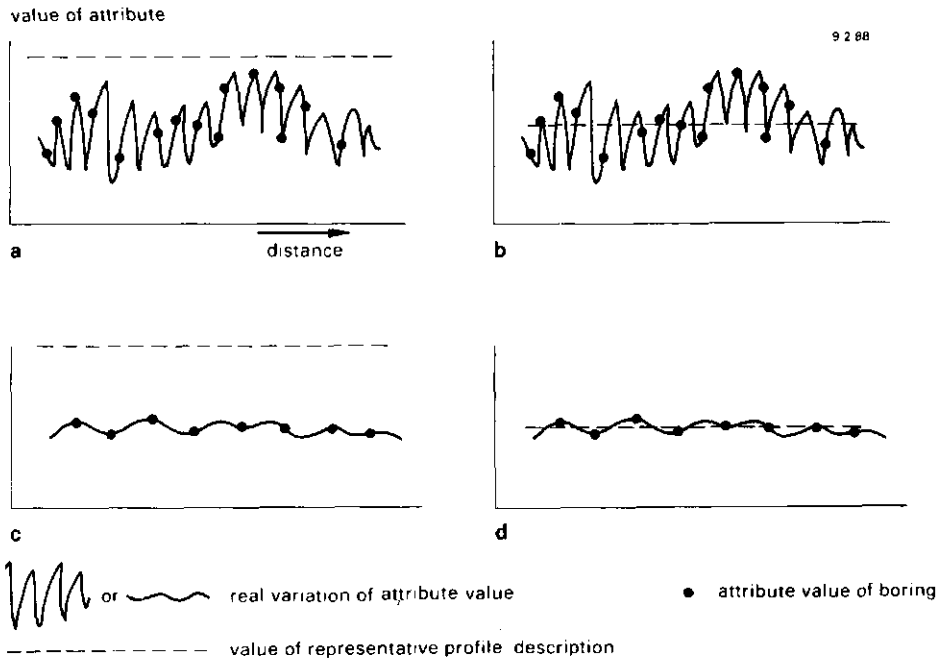


Fig. 1. Conceptual examples of attribute variations in mapping units and values of associated representative profile descriptions (RPD): a) large variation and large error in RPD results in high MAE and MSE values; b) large variation and small error in RPD results in high MAE and MSE values; c) small variation and large error in RPD results in high MAE and MSE values; d) small variation and small error in RPD results in low MAE and MSE values.

a systematic component in the measurement error. The random error (MSE_r) is caused by variation within the mapping unit and a random component in the measurement error. The MSE_r was estimated according to:

$$MSE_r = \frac{1}{n-m} \sum_{u=1}^{m} \sum_{i=1}^{k} (\bar{b}_u - b_{ui})^2 \quad (5)$$

where $\bar{b}_u$ = mean value of the borings in the uth mapping unit and the other symbols are the same as in eq. 1.

The MSE_s was estimated according to:

$$MSE_s = MSE - MSE_r$$

RESULTS AND DISCUSSION

In Table II the ME, MAE and MSE values are given for the moisture deficit calculated for the maps at scales of 1:10,000, 1:25,000 and 1:50,000. The estimated mean value of the moisture deficit (MV) for the study area by averaging

TABLE II

Mean error (ME), mean absolute error (MAE) and mean square error (MSE) values for the average moisture deficit (mm) in a 30-year period under the present hydrological situation for three map scales (the 95% confidence intervals for ME, MAE and MSE are given between brackets)

Scale 1:...	Mean value (MV) (mm)	Mean error (ME) (mm)	Mean absolute error (MAE) (mm)	Mean square error				
				total (MSE) (mm^2)	systematic component		random component	
					(MSE_s) (mm^2)	% of MSE	(MSE_r) (mm^2)	% of MSE
10,000	11.4	2.8 (1.9-3.8)	5.5 (4.7-6.4)	103 (74-132)	13	13	90	87
25,000	11.4	2.9 (1.9-3.9)	6.1 (5.2-7.0)	120 (89-151)	13	11	107	89
50,000	11.4	1.6 (0.4-2.8)	7.2 (6.2-8.2)	155 (116-194)	65	42	90	58

all the borings is 11.4 mm. Because the borings are more or less evenly distributed over the area, this value is considered to be the best possible prediction of the average moisture deficit in the area. The mean error (ME) gives an indication of the quality of prediction of this value from the RPD. The ME ranges from 2.8 mm for the 1:10,000 map to 1.6 mm for the 1:50,000 map, indicating an overall overestimation of the average moisture deficit by the RPD from maps at all three scales. The scale of the map hardly influences the prediction of the average moisture deficit of the whole area. This means that someone who is only interested in an average value for the whole area can use the 1:50,000 soil map. This conclusion corresponds with the one reached by Wösten et al. (1987).

The MAE gives an indication of the quality of the RPD for making predictions at specific locations. The MAE values range from 5.5 mm for the 1:10,000 map to 7.2 mm for the 1:50,000 map. This means that from the 1:10,000 map predictions for a particular spot in the area based on representative profiles are on average 5.5 mm too low or too high, which is an error of 48%. With increasing map scale the MAE decreases. From the scale of 1:50,000 to 1:25,000 the MAE decreases by 1.1 mm. From the scale of 1:25,000 to 1:10,000 the decrease in MAE is smaller. The transition from a scale of 1:25,000 to 1:10,000 yields only a small benefit in decrease of the mean absolute error. The MSE has a pattern similar to the MAE. The transition from the scale of 1:25,000 to 1:10,000 also yields only a small benefit in terms of decrease of the mean square error. An explanation for this is the small reduction in the number of mapping units, between the 1:10,000 and 1:25,000 maps (Table I). The random component (MSE_r) in the MSE is quite large for maps at all three scales. For the

TABLE III

Mean error (ME), mean absolute error (MAE) and mean square error (MSE) values for the change in yield (% ha^{-1} yr^{-1}) due to water extraction for three map scales (the 95% confidence intervals for ME, MAE and MSE are given between brackets)

Scale 1:...	Mean value (MV) (% ha^{-1} yr^{-1})	Mean error (ME) (% ha^{-1} yr^{-1})	Mean absolute error (MAE) (% ha^{-1} yr^{-1})	Mean square error				
				total (MSE) (% ha^{-1} yr^{-1})^2	systematic component		random component	
					(MSE_s) (% ha^{-1} yr^{-1})^2	% of MSE	(MSE_r) (% ha^{-1} yr^{-1})^2	% of MSE
10,000	4.6	−1.4 (−1.9 to −1.0)	3.3 (3.0-3.6)	21 (17-25)	7	33	14	67
25,000	4.6	−1.8 (−2.3 to −1.4)	3.6 (3.3-3.9)	23 (20-26)	7	30	16	70
50,000	4.6	−1.7 (−2.3 to −1.2)	4.5 (4.2-4.9)	35 (30-40)	21	60	14	40

1:50,000 map 58% of the MSE is caused by this error, which increases to about 90% for the scales of 1:10,000 and 1:25,000. As shown by the MSE_s, the RPD performs better at a larger scale.

In Table III the calculated ME, MAE and MSE values are given for the changes in yields for scales of 1:10,000, 1:25,000 and 1:50,000. The estimated mean value for the study area is 4.6% ha^{-1} yr^{-1}. This means that drawdown of the water-table has a net positive effect on the yields in this area. The ME ranges from −1.4 for the 1:10,000 to −1.7 mm for the 1:50,000 soil map. A negative ME means an underestimation of the average change in yield on the basis of the RPD's. Again, differences in quality between the three scales are small.

The total mean absolute error (MAE) ranges from 3.3 for the 1:10,000 map to 4.5 mm for the 1:50,000 map. The values for the 1:10,000 and the 1:25,000 maps differ little. The MSE has a pattern similar to the MAE. For the 1:50,000 map, 40% of the MSE is caused by the random error which increases to about 70% for scales of 1:10,000 and 1:25,000.

The MSE_s can be reduced by improving the quality of the RPD, e.g. by using better estimation procedures. With an improved RPD it is, for instance, possible to reduce the MAE_s on a map with a scale of 1:10,000 for the change of yield by 33% (Table III). The remaining 67% is a result of random errors and cannot be resolved by improvement of the RPD. For the scale of 1:50,000 improvement of the RPD would result in a possible reduction of the MSE by 60% (Table III). This means that improvement of the RPD is more effective at a 1:50,000 scale. Also for a moisture deficit, improvement of the RPD will be most effective on the scale of 1:50,000 (Table II).

Comparing Tables II and III we found that for all scales the contribution of the random component to the MSE is quite large. Due to this random component, studies with very detailed simulations that are based on data for one point and that are extrapolated to a large area of land, have to be considered with great scepticism. Very large errors can be made.

The individual farmer is less interested in the change in yield at a particular spot than for a field. In this study, we have no information about the 'real' change in yield values for fields to check the quality of the RPD on this point. We expect, however, that the error made in predicting the value for a field will be less than the error in predicting for points, because: (i) on all the scales a large proportion of the MSE is caused by random errors, of which soil variability forms a large component; for a field, the differences in change in yield due to spatial variability are likely to be averaged; (ii) the growth of crops need not reflect all the spatial variability observed in soils; the extent of the rooting system has the effect of averaging variability.

CONCLUSIONS

Representative profile descriptions (RPD) gave reasonable results in this study for predicting the average moisture deficit or the change in yield for the

total area. A map scale of 1:50,000 is sufficient for this purpose. When information has to be provided for point locations within the area, however, large prediction errors are made. These errors decrease when maps with larger scales are used. The best prediction is obtained with a 1:10,000 map, but the error at this scale is still quite large. Only small reductions of this error are possible by improving the RPD because of the large random component in the error.

ACKNOWLEDGEMENTS

The authors are grateful to Mr. J. Bouma, Mr. J.J. de Gruijter, Mr. A.B. McBratney and Mr. J.H.M. Wösten for their helpful comments.

REFERENCES

Beek, K.J., Burrough, P.A. and McCormack, D.E. (Editors), 1987. Proceedings of the International Workshop on Quantified Land Evaluation Procedures Held in Washington, D.C., 27 April-2 May 1986. ITC Bull., 6, 165 pp.

Bouma, J. and De Laat, P.J.M., 1981. Estimation of the moisture supply capacity of some swelling clay soils in the Netherlands. J. Hydrol., 49: 247-259.

Bouma, J., van Lanen, H.A.J., Breeuwsma, A., Wösten, J.H.M. and Kooistra, M.J., 1986. Soil survey data needs when studying modern land use problems. Soil Use Manag., 2: 125-130.

Breewsma, A., Wösten, J.H.M., Vleeshouwer, J.J., van Slobbe, A.M. and Bouma, J., 1986. Derivation of land qualities to assess environmental problems from soil survey. Soil Sci. Soc. Am. J., 50: 186-190.

Bunke, O. and Droge, B., 1984. Estimators of the mean squared error of prediction in linear regression. Technometrics, 26: 145-155.

De Laat, P.J.M., 1980. Model for unsaturated flow above a shallow water table, applied to a regional subsurface flow problem. Agric. Res. Rept., 895. Pudoc, Wageningen, 126 pp.

Kempthorne, O. and Allmaras, R.R., 1986. Errors and variability of observations. In: A. Klute (Editor), Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods. American Society of Agronomy, Madison, Wisconsin, 2nd ed., pp. 1-31.

Soil Survey Staff, 1975. Soil Taxonomy. A Basic System of Soil Classification for Making an Interpreting Soil Surveys. USDA-SCS Agric. Handb. 436, Washington, D.C.

Stoffelsen, G.H. and van Holst, A.F., 1985. Bodemkundig-Hydrologisch Onderzoek in het Waterwingebied Mander. Fase II: De Bodemgesteldheid, de Voormalige en Huidige Hydrologische Situatie en de Grondwaterstandsverlaging. Stiboka, Wageningen, 94 pp.

Van Bohemen, P.J.M., 1981. Toename van de Produktie van Grasland bij Verbetering van de Watervoorziening. ICW, Wageningen, 68 pp.

Western, S., 1978. Soil Survey Contracts and Quality Control. Clarendon, Oxford, 284 pp.

Wösten, J.H.M., Bannink, M.H. and Bouma, J., 1987. Land evaluation at different scales: you pay for what you get! Soil Survey and Land Evaluation, 7: 13-24.

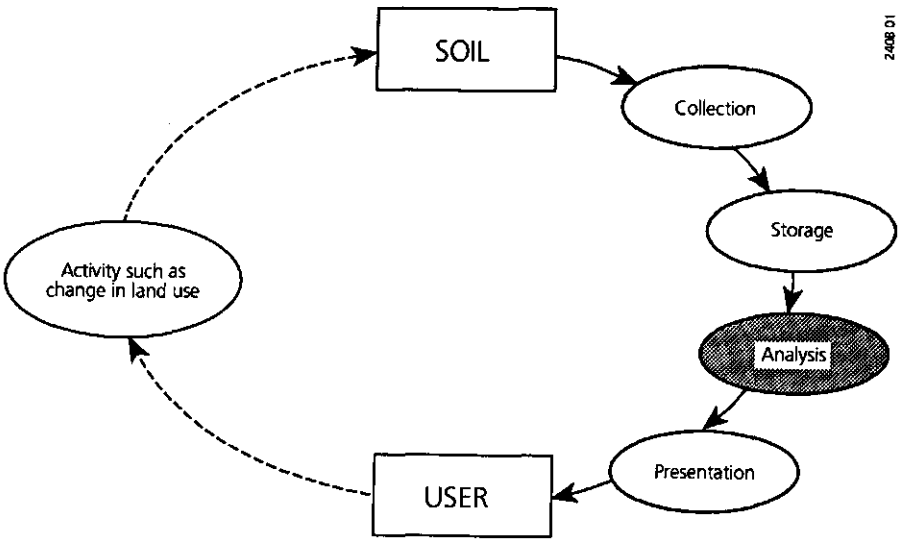
Wösten, J.H.M., Bouma, J. and Stoffelsen, G.H., 1985. Use of soil survey data for regional soil water simulation models. Soil Sci. Soc. Am. J., 49: 1238-1244.

Wösten, J.H.M., Bannink, M.H. and Beuving, J., 1987. Water Retention and Hydraulic Conductivity Properties of Top- and Subsoils in The Netherlands: The Staring Series. Stiboka, Wageningen, 75 pp.

Young, A., 1980. Tropical Soils and Soil Survey. Cambridge University Press, Cambridge, 278 pp.

6 Comparison of thematic maps derived from a soil map and from kriging of point data

Published in Geoderma, 39:281-291 (1987)



Comparison of Thematic Maps Derived from a Soil Map and from Kriging of Point Data

A.K. BREGT, J. BOUMA and M. JELLINEK

Netherlands Soil Survey Institute, P.O. Box 98, 6700 AB Wageningen (The Netherlands)

(Received May 6, 1986; accepted after revision October 31, 1986)

ABSTRACT

Bregt, A.K., Bouma, J. and Jellinek, M., 1987. Comparison of thematic maps derived from a soil map and from kriging of point data. Geoderma, 39: 281–291.

A general-purpose soil survey was carried out in the eastern part of The Netherlands. A thematic map was derived from the produced soil map, containing basic soil data that are important for soil-hydrological behaviour. A comparable map was obtained by interpolation of data obtained from the individual borings. The quality of both maps in terms of purity was tested by means of 60 independent test borings. The purity of each of the maps was 77%, indicating no significant difference in quality.

SCOPE OF THE STUDY

Soil differs from place to place, both laterally and vertically. In normal soil survey the surveyor describes the vertical variation of the soil by distinguishing different soil horizons. The lateral variation is described by delineating boundaries of soil units in the field. The surveyor analyses borings and interprets landscape features such as relief and vegetation patterns. He classifies the soils according to a classification system, which uses a variety of basic soil properties. Such general-purpose classifications are most useful for resource inventories and surveys for various types of land use. Various thematic and interpretative maps can be derived from a general-purpose soil map, e.g. maps showing the starting-depth of sand or boulder clay or maps showing suitabilities for grazing or forestry.

Soil variability can also be described by interpreting boring data by statistical methods. A review of soil-variability studies following classical procedures was presented by Beckett and Webster (1971). The use of classical statistics in describing soil variability neglects, however, the spatial dependence between neighbouring observation points. In the last ten years, geostatistical techniques were developed to define this spatial dependence using the

$$K(h)_1 = \mathrm{Co} + a_1 |h| + a_3 |h|^3$$

$$K(h)_2 = \mathrm{Co} + a_1 |h| + a_3 |h|^3 + a_5 |h|^5 \qquad (6)$$

respectively, where $K(h)$ = generalized covariance; Co = nugget variance; a_0, a_1, a_3, a_5 = coefficients to be determined.

Kriging in the presence of drift is called universal kriging. More information about the kriging procedure and theory is given by Olea (1975), David (1977), Journel and Huijbregts (1978), and Kafritsas and Bras (1981).

MATERIALS AND METHODS

Study area and soil map

The study was carried out on soil data collected in the survey area "Hupselse Beek", which covers an area of 650 ha. It is located in the eastern part of The Netherlands, near Groenlo. From this area, a sample area of 125 ha was considered for this study. The area consists of cover sand overlying miocene clay and boulder clay. The boulder clay (glacial till) was deposited in an Early Pleistocene period. A detailed soil map was made at a scale of 1:5000. The soils were classified according to Soil Taxonomy (Soil Survey Staff, 1975) mainly as Typic Haplaquods, Typic Haplaquepts, Plaggeptic Haplaquods and Typic Udipsamments (see also Wösten et al., 1985). Mapping units of the soil map were delineated in the field by using soil borings and by interpreting landscape features. For the soil map, 198 soil borings to a depth of 2 m were made. Profile data are field estimates of basic soil properties.

The Dutch soil-classification system, normally used for detailed soil surveys, mainly reflects properties of the horizons near the soil surface. When the soil map has to be used for hydrological interpretations, it should, however, include additional data of subsurface soil horizons that are expected to be important for soil-physical properties. Therefore, additional observations were made of textures and structures of all horizons. Also, the starting-depths of gravel and boulder clay or miocene clay were recorded. The depth classifications for gravel and clay were adopted from existing mapping criteria, and they were presented as overlays on the soil map.

In the sandy surface sediments various types of soils were formed, which differ mainly in thickness of the A horizon. These soils are underlain by a sequence of gravel and clay sediments starting at various depths.

Thematic map derived from soil map

From the general-purpose soil map, as discussed above, a thematic map was derived, containing soil data especially important for soil-hydrological behav-

iour. Wösten et al. (1985) concluded that in this area differences in the thickness of the A horizon and the starting-depths of gravel and boulder clay were mainly associated with differences in soil-hydrological characteristics. This was concluded on the basis of multiple measurements of hydraulic-conductivity and moisture-retention curves in the different soil horizons.

In the context of this study, the legend of the original thematic map, as presented by Wösten et al. (1985), was somewhat simplified. The starting-depths of boulder clay and miocene clay were combined into one diagnostic criterion. Instead of the thickness of the root zone, the thickness of the A horizon was used here.

Thematic map obtained by kriging of point data

In this study, a thematic map was also obtained by kriging of soil data from individual borings, as described below. The values for the thickness of the A horizon and for the starting-depths of gravel and boulder clay were selected from the individual boring descriptions. The values for each soil property were interpolated to a regular grid of 20 by 25 points. Interpolation was carried out using the universal kriging program AKRIP (Kafritsas and Bras, 1981). For each property, the kriging procedure involved the following four steps: (1) identification of the order (k) of the intrinsic random function, which indicates the order of the trend in the data; (2) determination of the coefficients of the generalized covariance models appropriate for the pre-determined order of the intrinsic random function; (3) selection of the best generalized covariance model; (4) application of point kriging.

This procedure yielded an estimated value for the three soil properties for every grid point. Next, the estimated values of properties at each grid point were classified and combined with the same criteria as used for the thematic map derived from soil map. The resulting grid map was transformed into a polygon map using a grid-polygon conversion algorithm.

Validation procedure

The quality of both maps was tested by calculating the purity of these maps. The purity, in terms of percentages, indicates agreement between test-boring data and data according to the legends of both maps. Sixty test borings were made for calculating the purity. The locations of these test borings were chosen according to a procedure described by De Gruijter and Marsman (1985). By this procedure, the study area was divided into 20 blocks of 250 m by 250 m. Twelve blocks were selected randomly with replacement. In each selected block, a transect was located by determining a random starting-point. The direction of the transect was selected at random from two alternatives: east–west and north–south. The distance between the test points within the transect was

50 m. From the sixty test locations selected this way, one was located on a road and was dropped.

For both maps, two types of purity measures were calculated, namely partial purity and average purity. Partial purity is defined as the percentage agreement between one classification criterion of the map and the corresponding property of the test borings. Average purity is defined as the arithmetic mean of all the partial purities (Marsman and De Gruijter, 1986). Three partial purities and one average purity were calculated for each map.

Calculations of map purity can be performed quickly and easily. A disadvantage is that all deviations from the legend are equally weighted as they are independent of the magnitude of the deviation.

RESULTS AND DISCUSSION

Thematic map derived from soil map

The thematic map derived from the soil map (map 1) is shown in Fig. 1 and the corresponding legend is presented in Table I. In total, 15 different mapping units were distinguished. The mapping units 15 and 16 covered large areas on this map. Map 1 contains 41 delineated areas compared with 110 on the detailed soil map. Various areas on the soil map were combined because they did not differ from a soil-physical point of view (Wösten et al., 1985). When a thematic map is made from a general-purpose soil map, the number of delineated areas can often be reduced.

Thematic map obtained by kriging of point data

For every selected soil property, calculations were made for the order of the intrinsic random function (IRF) and the generalized covariance that best describes the data (Table II). The thickness of the A horizon and the starting depth of gravel show no trend in the data ($k=0$). The generalized covariances ($K(h)$) of these properties are constant, which implies the occurrence of a 100% nugget variance. This, in return, means that no spatial structure in the data is shown. Lack of spatial structure does not imply that kriging should not be used. The AKRIP program will calculate unweighed average values based on a limited number of surrounding points. This represents an optimal interpolation procedure when no spatial structure is present. Nugget effects are caused by measuring errors and variation of a soil property within the shortest sampling interval. In this study we are dealing with estimated soil properties, and measuring errors are expected to be considerable. These errors are probably the main cause for the observed 100% nugget variance. Similar high nugget variances are reported by Burrough (1983) and Bregt (1985) when describing the spatial variability of soil properties.

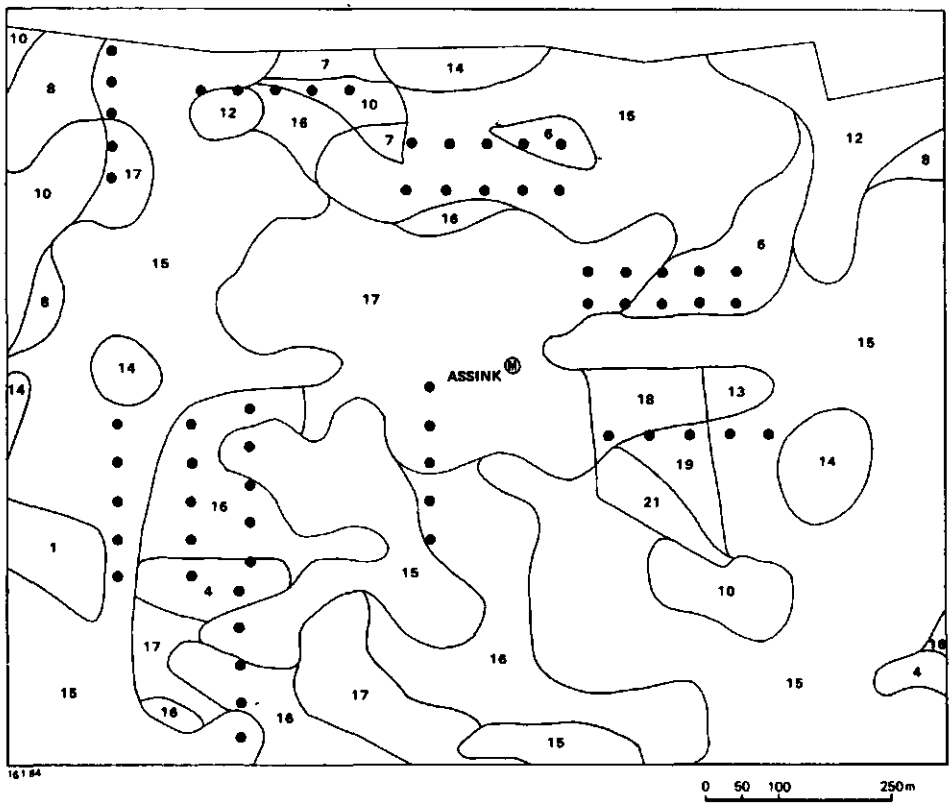


Fig. 1. Thematic map derived from the soil map (map 1) of the study area of 125 ha (legend in Table I); the locations of the 60 test borings are indicated.

The single soil properties were interpolated to a regular grid. Through the use of an allocation algorithm, results were combined into a grid map using classification criteria identical to the ones used for map 1. The resulting grid map was transformed into a polygon map. The latter map is shown in Fig. 2 and the corresponding legend is presented in Table I. Henceforth, this map will be referred to as map 2. Map 2 contains 63 delineated areas, compared with 41 on map 1. The difference is caused by small areas, covering a single grid cell.

Comparing both maps

The partial purity and the average purity were calculated for maps 1 and 2 (Table III). Consideration of all properties together resulted in 77% average purity for both maps. These results show that the quality of map 1 does not differ significantly from map 2.

The partial purity of the starting-depth of boulder clay was significantly

TABLE I

Characteristics of 21 units distinguished on one of the two or on both thematic maps

Unit of thematic map		Thickness of A horizon (cm)	Starting-depth (cm below surface)	
derived from soil map	kriged		gravel	boulder/miocene clay
1	1	<25	*	40-120
-	2	<25	*	120-200
-	3	<25	*	*
4	-	<25	120-200	*
-	5	25-35	< 40	< 40
6	6	25-35	< 40	40-120
7	7	25-35	< 40	*
8	8	25-35	40- 80	40-120
-	9	25-35	40- 80	120-200
10	10	25-35	40- 80	*
-	11	25-35	80-120	40-120
12	12	25-35	80-120	120-200
13	13	25-35	80-120	*
14	14	25-35	*	< 40
15	15	25-35	*	40-120
16	16	25-35	*	120-200
17	17	25-35	*	*
18	-	>35	80-120	*
19	19	>35	*	40-120
-	20	>35	*	120-200
21	21	>35	*	*

* = > 200 or absent; - = unit not present on map.

lower on both maps than the partial purities of the thickness of the A horizon and the starting-depth of gravel. For map 1 this can be explained by local soil conditions. The topography of the surface of the boulder clay is independent

TABLE II

Selected intrinsic random functions (IRF) for the three properties

Property	Order of IRF	Generalized covariance
Thickness A horizon	0	$K(h)_0 = 63.4$
Starting-depth gravel	0	$K(h)_0 = 9$
Starting-depth boulder clay	1	$K(h)_1 = 0.1 + 45\|h\|^3$

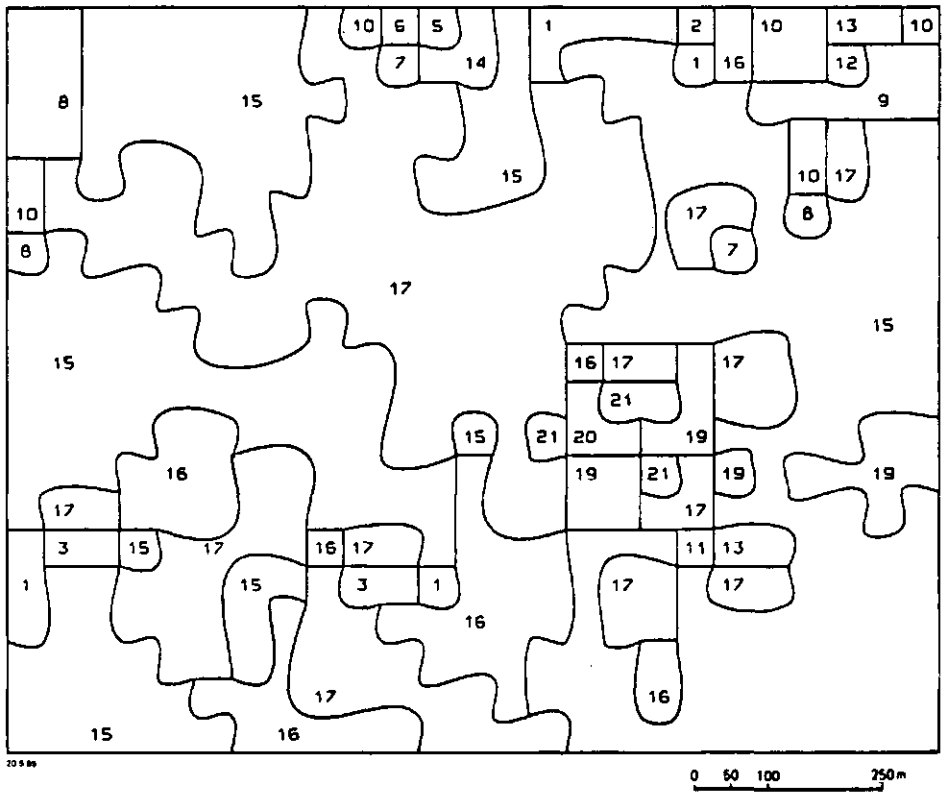


Fig. 2. Kriged thematic map (map 2) of the study area of 125 ha (legend in Table I).

TABLE III

Validation of the thematic map derived from the soil map and the kriged thematic map, based on comparison between data from 60 independent borings (locations indicated in Fig. 1) and the corresponding units of the two maps (the 90% confidence interval is given in brackets)

Purity measure	Thematic map	
	derived from soil map	kriged
Partial purity: thickness A horizon (%)	80 (68-92)	83 (71-95)
Partial purity: starting-depth of gravel (%)	83 (77-89)	86 (75-97)
Partial purity: starting-depth of boulder clay (%)	68 (57-79)	61 (51-71)
Average purity: all data (%)	77 (72-82)	77 (69-85)

of patterns formed by various soils in the sandy surface layers and is therefore difficult to predict from landscape patterns. The soil surveyor has to rely mainly on his limited number of borings when delineating boundaries of the boulder-clay surface. This may be expected to result in a lower purity for this property. The purity of the starting-depth of boulder clay is, however, even lower for map 2. This is surprising because the intrinsic random function shows a clear spatial structure, and one would therefore expect a better prediction by the kriging method.

The results were based on the use of an empirically determined classification of the thickness of each horizon in each mapping unit. In retrospect, the obtained purity of approximately 80% is indicative for what turned out to be a satisfactory legend, because this purity is commonly used as a guideline. For example, soil survey in The Netherlands attempts to achieve a purity of at least 70% (Buringh et al., 1962). Comparable USDA values range from 80 to 90% (Soil Survey Staff, 1951).

CONCLUSIONS

No significant difference in quality in terms of purity was found between a large-scale thematic map derived from a general-purpose soil map as compared with a similar map obtained by kriging of point data.

Good results obtained in this study by interpreting soil maps are encouraging for their future use, because they form a widely available soil-data base.

ACKNOWLEDGEMENT

The authors are grateful to Dr. J.J. de Gruijter for his helpful comments.

REFERENCES

Beckett, P.H.T. and Webster, R., 1971. Soil variability: a review. Soils Fertilizers, 34: 1-15.

Bregt, A.K., 1985. Geostatistical techniques and spatial variability of soil physical properties. In: Soil Physics and Rice. International Rice Research Institute, Los Banos, Philippines, pp. 85-98.

Burgess, T.M. and Webster, R., 1980a. Optimal interpolation and isarithmic mapping of soil properties, I. The semi-variogram and punctual kriging. J. Soil Sci., 31: 315-331.

Burgess, T.M. and Webster, R., 1980b. Optimal interpolation and isarithmic mapping of soil properties, II. Block kriging. J. Soil Sci., 31: 333-341.

Buringh, P., Steur, G.G.L. and Vink, A.P.A., 1962. Some techniques and methods of soil survey in the Netherlands. Neth. J. Agric. Sci., 10: 157-172.

Burrough, P.A., 1983. Multiscale sources of spatial variation in soil, I. The application of fractal concepts to nested levels of soil variation. J. Soil Sci., 34: 577-597.

David, M., 1977. Geostatistical Ore Reserve Estimation. Elsevier, Amsterdam, 364 pp.

Davis, J.C., 1973. Statistics and Data Analysis in Geology. Wiley, New York, N.Y., 550 pp.

De Gruijter, J.J. and Marsman, B.A., 1985. Transect sampling for reliable information on mapping

units. In: D.R. Nielsen and J. Bouma (Editors), Soil Spatial Variability. Proceedings of a Workshop of the ISSS and the SSSA, Las Vegas, 30 November–1 December 1984. Pudoc, Wageningen, pp. 150–163.

Delfiner, J.P., 1976. Linear estimation of non-stationary spatial phenomena. In: M. Guarascio et al., Advanced Geostatistics in the Mining Industry. NATO Advanced Study Institute Series. Reidel, Boston.

Dubrule, O., 1984. Comparing splines and kriging. Comput. Geosci., 10: 327–338.

Gajem, Y.M., Warrick, A.W. and Mijers, D.E., 1981. Spatial structure of physical properties of a typic Torrifluvent soil. Soil Sci. Soc. Am. J., 45: 709–715.

Giltrap, D.J., 1983. Computer production of soil maps, I. Production of grid maps by interpolation. Geoderma, 10: 327–338.

Journel, A.J. and Huijbregts, Ch.J., 1978. Mining Geostatistics. Academic Press, London, 600 pp.

Kafritsas, J. and Bras, R.L., 1981. The Practice of Kriging. Report No. 263. Mass. Inst. Technol., Cambridge, Mass.

Marsman, B.A. and De Gruijter, J.J., 1986. Quality of Soil Maps: A Comparison of Survey Methods in a Sandy Area. Soil Survey Papers No. 15. Stiboka, Wageningen.

Matheron, G., 1971. The theory of regionalized variables and its applications. Cah, Centre Morphol. Mathém. Fontainebleau, No. 5.

McBratney, A.B. et al., 1982. Regional variation of extractable copper and cobalt in the topsoil of southeast Scotland. Agronomie, 2: 969–982.

Olea, R.A., 1975. Optimal mapping techniques using regionalized variable theory. Series on Spatial Analysis, 2. Kansas Geol. Surv., Lawrence, Kans.

Soil Survey Staff, 1951. Soil Survey Manual. US Dept. Agric. Handb. 18, Washington, D.C.

Soil Survey Staff, 1975. Soil Taxonomy. A Basic System of Soil Classification for Making an Interpreting Soil Surveys. USDA-SCS Agric. Handb. 436, Washington, D.C.

Ten Berge, H.F.M. et al., 1983. Spatial variability of physical soil properties influencing the temperature of the soil surface. Agric. Water Manage., 6: 213–226.

Van Kuilenburg, J., De Gruijter, J.J., Marsman, B.A. and Bouma, J., 1982. Accuracy of spatial interpolation between point data on soil moisture supply capacity, compared with estimates from mapping units. Geoderma, 27: 311–325.

Vieira, S.R., Nielsen, D.R. and Biggar, J.W., 1981. Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J., 45: 1040–1048.

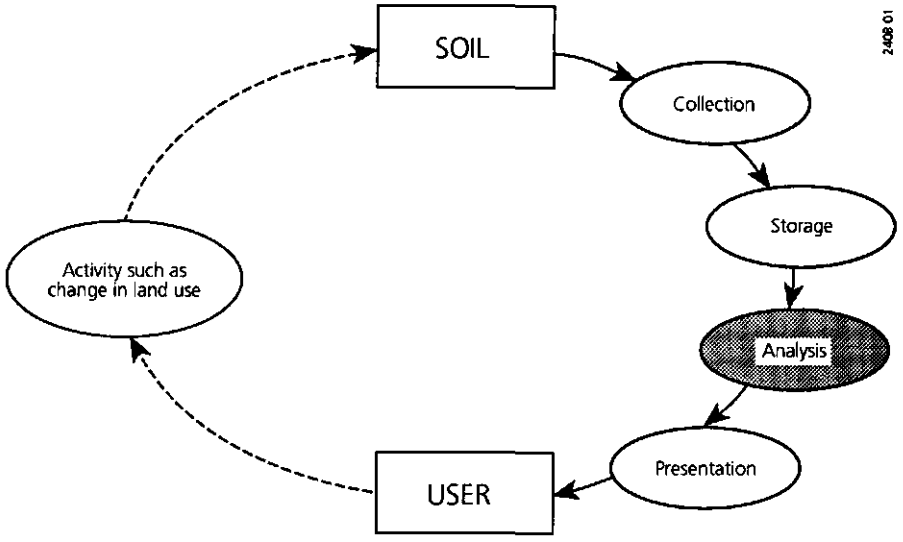
Webster, R. and Burgess, T.M., 1980. Optimal interpolation and isarithmic mapping, III. Changing drift and universal kriging. J. Soil Sci., 31: 505–524.

Webster, R. and Nortcliff, S., 1984. Improved estimates of micro nutrients in hectare plots of the Sonning Series. J. Soil Sci., 35: 667–672.

Wösten, J.H.M., Bouma, J. and Stoffelsen, G.H., 1985. Use of soil survey data for regional soil water simulation models. Soil Sci. Soc. Am. J., 49: 1238–1244.

7 Determination of rasterizing error: a case study with the soil map of the Netherlands

Published in International Journal of Geographical Information Systems, 5:361-367 (1991)



Determination of rasterizing error: a case study with the soil map of The Netherlands

A. K. BREGT, J. DENNEBOOM, H. J. GESINK
and Y. VAN RANDEN

DLO-The Winand Staring Centre for Integrated Land, Soil and Water Research, P.O. Box 125, 6700 AC Wageningen, The Netherlands

Abstract. The digital soil map of The Netherlands (scale 1 : 50 000) in vector form was rasterized using three sizes of raster cell and two different rasterizing methods. The rasterizing errors were estimated for several map sheets of different complexity using the double-conversion method, the Switzer method and the Goodchild method. The relationship between the complexity of the maps, indicated by the boundary index, and the rasterizing error is presented as a regression equation. The double-conversion method provided a better estimation of the rasterizing error than the other two methods. Differences of less than 1 per cent were found between the rasterizing methods used.

1. Introduction

In vector to raster conversion an error is introduced, the rasterizing error. The influence of rasterizing method, size of raster cell and map complexity on this error have been studied, and the purpose of this paper is to report the results of a case study of the soil map of The Netherlands.

At the Winand Staring Centre a geographical information system with mainly soil data has been established during the past ten years. One part of the system consists of the soil map of The Netherlands at a scale of 1 : 50 000, digitization of which in vector form began in 1976. All the map sheets (75) have now been digitized and were recently rasterized, mainly to combine soil data with other geographical data. For example, in the landscape ecological mapping project of The Netherlands (Canters *et al.* 1991), soil data were linked with flora and fauna data in raster form. Organizations in The Netherlands also require a rasterized version of the soil map to be used within their own systems.

2. Map rasterization and rasterizing error

When rasterizing a soil map it is necessary to select the orientation of the raster, the size of raster cell and the rasterizing method. In this study the orientation of the raster was chosen to be parallel to the co-ordinate lines of the topographical map. Both size of cell and method of rasterizing are important factors influencing the rasterizing error, which is also affected by the complexity of the map pattern.

2.1. *Rasterizing methods and rasterizing error*

Raster to vector conversion and the resulting errors arising from these processes are illustrated in figure 1. In figure 1a a fragment of the soil map in vector form is presented with two units A and B on which a raster with 16 cells has been superimposed (figure 1b). For rasterizing the vector map two methods were used: central point and dominant unit rasterization (Wehde *et al.* 1980). With central point rasterization the

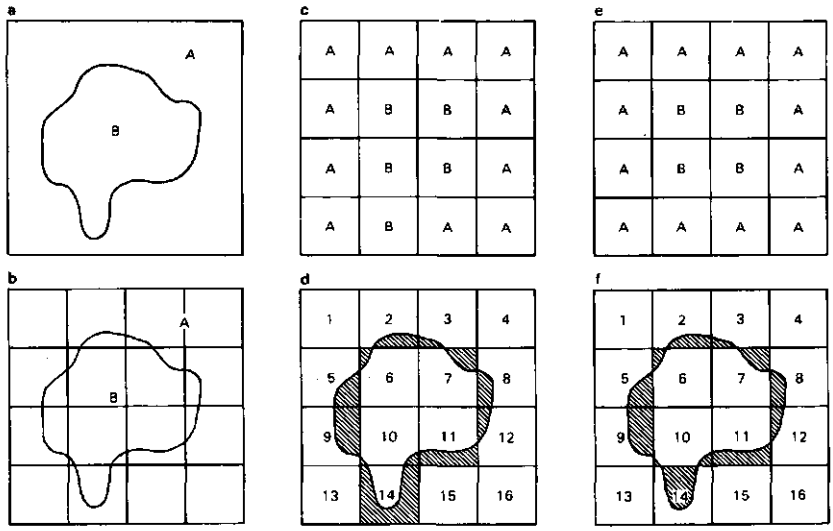


Figure 1. Fragment of soil map. A and B are arbitrary soil units. (a) Soil map in vector form; (b) soil map in vector form with raster superimposed; (c) soil map in raster form by central point rasterizing of the original vector map; (d) rasterizing error by central point rasterizing (shaded) (1....16 are cell numbers); (e) soil map in raster form by dominant unit rasterizing of the original vector map; and (f) rasterizing error by dominant unit rasterizing (shaded) (1....16 are cell numbers).

soil unit on the vector map which is located at the central point of each raster cell is considered to be representative of the whole cell (figure 1c). With dominant unit rasterization it is the soil unit occupying the largest area of each raster cell which is considered to be representative (figure 1e). When more than one soil unit occurs in a raster cell, an error appears. For example, in cell 2 (figure 1d), the shaded area is incorrectly considered to be part of unit A, as it is in cells 3, 5, 8, 9 and 12. There are also cells (6, 7, 10, 11 and 14) where parts are incorrectly considered to be part of unit B. The difference between cell 14 in figure 1d and that in figure 1f is caused by the difference in rasterizing method. The shaded area in figure 1d and 1f is the total area incorrectly classified by the rasterizing process. This area (the mis-match area), expressed as a percentage of the total area, forms the rasterizing error, which will be referred to in this paper as 'the error'.

2.2. *Raster cell size*

The relationship between size of raster cell and error has been discussed by Müller (1977) and Burrough (1986), who state that smaller raster cell sizes will result in smaller errors. There are therefore advantages in making the raster cell as small as possible, on the other hand, with smaller cells, the conversion time and the necessary storage capacity increase rapidly. A size of raster cell should therefore be chosen where both the conversion time and the error are acceptable. In this study the error was evaluated for three sizes of cell, 1×1 mm, 2×2 mm and 4×4 mm.

2.3. *Complexity of the map pattern*

According to Switzer (1975) a relationship exists between the error and complexity of the map and the boundary index (BI) was used in this study as a measure of that complexity (Bregt and Wopereis 1990). The BI is defined as the boundary length in centimetres per square centimetre of the map and is calculated by dividing the total length of the polygon boundaries by the surface of the map sheet.

The error was calculated for 11 of the 75 soil map sheets of The Netherlands, each of which has a surface area of 2000 cm^2, except sheet B (which covers 2750 cm^2). The complexity of the map sheets chosen varies considerably, and the BI ranges from 0·2 for sheet A to 2·2 for sheet K. The boundary indices of the remaining soil maps of The Netherlands fall within this range.

2.4. *Measurements of the error*

For the determination of the error, three different methods were used, two of which were adapted from the literature and the third was developed by the authors.

The first method is described by Switzer (1975). According to Switzer (1975), the error between two map units i and j can be estimated by

$$F_{ij} = 0{\cdot}76P1_{ij}/N1 - 0{\cdot}19P2_{ij}/N2 \quad (1)$$

where: F_{ij} = that part of the map that belongs to map unit i but is represented as map unit j; $P1_{ij}$ = the number of cell pairs (horizontal and vertical) that lie in the different map units i and j at a distance of one cell width; $N1$ = the total number of cell pairs at a distance of one cell width: $N1 = 4(XY) - 2(X + Y)$, where X = number of columns, Y = number of rows; $P2_{ij}$ = the number of cell pairs (horizontal and vertical) that lie in the different map units i and j at a distance of two cell widths; $N2$ = the total number of cell pairs at a distance of two cell widths; $N2 = 4(XY) - 4(X + Y)$, where X = number of columns, Y = number of rows.

The total error, F_i, of map unit i is given by

$$F_i = \sum_{i \neq j} F_{ij} \text{ (i fixed)} \quad (2)$$

The total error, F, for the entire map is given by

$$F = \sum_{i=j}^{k} F_i \quad (3)$$

where k = number of map units.

The second method is that described by Goodchild (1980), based on the concept of fractals. Goodchild (1980) changed the parameters in equation (1) to

$$F_{ij} = 0{\cdot}60P1_{ij}/N1 - 0{\cdot}11P2_{ij}/N2 \quad (4)$$

The next steps are similar to those described by the Switzer [equations (2) and (3)].

The third method is the double-conversion method, according to which a map sheet is rasterized twice. First, a normal vector to raster conversion is undertaken (1 × 1 mm, 2 × 2 mm, or 4 × 4 mm, see figure 2b), the product of which is called the base raster. The map sheet is then rasterized again, but with a much smaller raster cell size (e.g. 0·1 × 0·1 mm) to produce the fine raster (figure 2d). Subsequently, the map unit of the fine raster is compared with the cell of the base raster. If they differ, that part of the base raster occupied by a cell of the fine raster has been incorrectly coded. The total error is

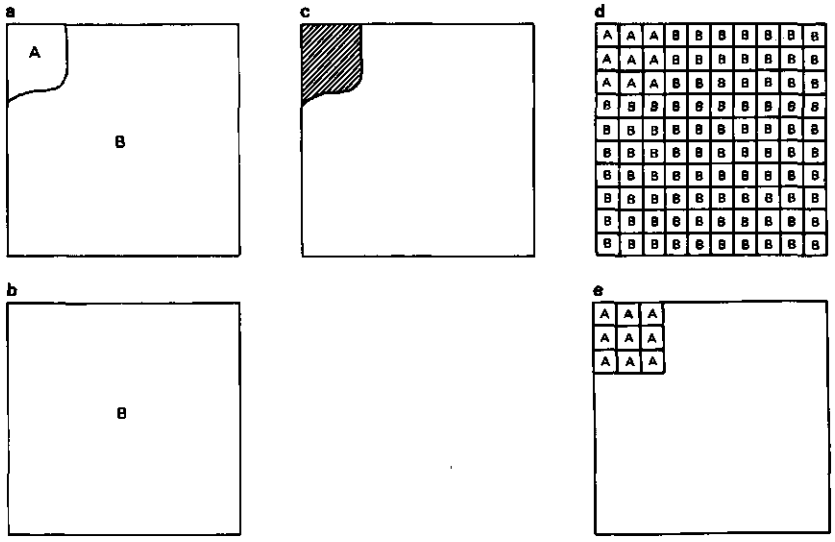


Figure 2. Calculation of the error using the double-conversion method. (a) Fragment of soil map with units A and B; (b) rasterizing to 'base raster'; (c) rasterizing error (shaded); (d) rasterizing to 'fine raster'; and (e) estimation of the error by comparing the fine raster with the base raster.

the sum of all cells in the fine raster with a code that differs from the base raster (figure 2e).

Using the double-conversion method it is possible to calculate the error with any desired accuracy. It was found that a fine raster of 0·02 × 0·02 mm, 0·1 × 0·1 mm, or 0·2 × 0·2 mm influences the error only in the second decimal place, and to limit processing time, a fine raster of 0·1 × 0·1 mm was chosen for further studies.

3. Results and discussion

The errors for the 11 map sheets using the double-conversion method are given in table 1. The error for a raster cell of 4 × 4 mm increased to 20 per cent for a complex map (K), twice that for a raster cell size of 2 × 2 mm (which in turn was almost twice the error for a raster cell size of 1 × 1 mm).

From table 1 it can be seen that a relationship exists between map complexity and error, the latter increasing with an increasing value of BI. The relationship between the BI and the error is presented in figure 3, which shows a linear relationship between the two values. The calculated regression equations are given in table 2. The percentage variance explained by the regression equations was large, indicating a good fit, so that when the boundary index is known for a map sheet, the error can be predicted using these equations.

Table 1 shows that for raster cells of sizes 1 × 1 mm and 2 × 2 mm the difference in error between central point and dominant unit rasterization is very small. For a cell size of 4 × 4 mm, the difference was approximately 0·5 per cent (with one exception for map sheet J, where the difference was 0·9 per cent). For larger cell sizes dominant unit rasterization gave a smaller error than central point rasterization. The difference in

Table 1. Rasterizing error of 11 map sheets for three raster cell sizes and two rasterizing methods, calculated using the double-conversion method.

Map sheet	BI (cm^{-1})	Error for raster cell sizes (%)								
		1 mm × 1 mm			2 mm × 2 mm			4 mm × 4 mm		
		Cent.	Dom.	Diff.	Cent.	Dom.	Diff.	Cent.	Dom.	Diff.
A	0·20	0·5	0·5	0·0	0·9	0·9	0·0	1·7	1·7	0·0
B	0·91	2·2	2·2	0·0	4·3	4·3	0·0	8·3	8·0	0·3
C	1·11	2·6	2·6	0·0	5·2	5·1	0·1	9·7	9·2	0·5
D	1·31	3·1	3·1	0·0	6·2	6·1	0·1	11·6	11·1	0·5
E	1·42	3·4	3·4	0·0	6·7	6·6	0·1	12·8	12·2	0·6
F	1·47	3·5	3·5	0·0	6·9	6·8	0·1	13·6	13·3	0·3
G	1·85	4·3	4·3	0·0	8·5	8·3	0·2	16·0	15·6	0·4
H	1·94	4·6	4·6	0·0	9·2	9·1	0·1	17·8	17·4	0·4
I	1·97	4·7	4·7	0·0	9·3	9·2	0·1	18·0	17·4	0·6
J	2·01	4·8	4·7	0·1	9·4	9·2	0·2	17·4	16·5	0·9
K	2·17	5·3	5·3	0·0	10·5	10·3	0·2	20·2	19·7	0·5

BI = Boundary index; Cent. = central point rasterizing; Dom. = dominant unit rasterizing; Diff. = differences in error between Cent. and Dom.

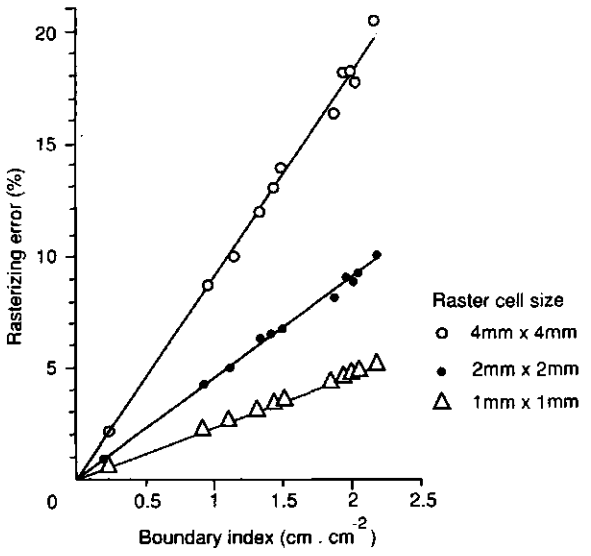


Figure 3. Relationship between rasterizing error (central point rasterizing) and BI for 11 map sheets three raster cell sizes.

Table 2. Relationship between rasterizing error, calculated using the double-conversion method, and the BI for three raster cell sizes and two rasterizing methods for the 11 map sheets.

Rasterizing method	Raster cell size (mm)	Regression equation	Explained variance (%)
Central	1 × 1	E = 2·4BI	99·8
Dominant	1 × 1	E = 2·4BI	99·8
Central	2 × 2	E = 4·7BI	99·8
Dominant	2 × 2	E = 4·6BI	99·8
Central	4 × 4	E = 9·0BI	99·4
Dominant	4 × 4	E = 8·7BI	99·0

E = Rasterizing error (%); BI = boundary index ($cm/cm^2 = cm^{-1}$).

Table 3. *Comparison between rasterizing error (central point rasterizing), using the double-conversion method (DCM), and the calculated error using the Switzer (1975) and Goodchild (1980) equations for a raster cell size of 2 × 2 mm.*

Map sheet	Error DCM (%)	Switzer (1975)		Goodchild (1980)	
		Error (%)	Difference with DCM	Error (%)	Difference with DCM
A	0·9	1·0	0·1	1·0	0·1
B	4·3	5·0	0·7	4·6	0·3
C	5·2	5·6	0·4	5·2	0·0
D	6·2	6·9	0·7	6·4	0·2
E	6·7	7·9	1·2	7·2	0·5
F	6·9	7·7	0·8	7·3	0·4
G	8·5	9·4	0·9	8·8	0·3
H	9·2	10·8	1·6	9·9	0·7
I	9·3	10·5	1·2	9·8	0·5
J	9·4	11·7	2·3	10·3	0·9
K	10·5	12·3	1·8	11·3	0·8

error between the two methods did not increase with an increase in the boundary index; map complexity seems to influence the error equally for both methods.

In the results presented so far the double-conversion method was used for calculating the error. For a raster cell of 2 × 2 mm a comparison was made between this method and the methods adapted from the literature (Switzer 1975, Goodchild 1980). These results are presented in table 3. In all situations the latter methods overestimate the error, although Goodchild's equation gives a closer estimate than that of Switzer.

Although the soil map of The Netherlands at a scale of 1 : 50 000 was used in this study for the estimation of the regression equations, the use of the equations obtained is not restricted to this particular map. The BI is presented in map co-ordinates, which makes the equations independent of map scale, so that they can be applied to other vectors maps if raster cells of 1 × 1 mm, 2 × 2 mm, or 4 × 4 mm are used.

The advantages of using these regression equations are that they provide better estimate of the rasterizing error than do the Switzer (1975) and Goodchild (1980) methods; they can easily be applied, since only calculation of the BI is required; and the rasterizing error can be estimated before rasterizing the map. The disadvantage is that their scope is limited in possible sizes of raster cells and BI interval, whereas the Switzer (1975) and Goodchild (1980) methods can be applied in all possible situations.

4. Conclusions

When a vector map has to be converted into a raster map with cells of 1×1 mm, 2×2 mm or 4×4 mm, the rasterizing error can be estimated simply by calculating the BI and using the regression equations calculated in this study (Table 2).

The difference in rasterizing error between central point and dominant unit rasterization is very small, especially for raster cells, of 1×1 mm and 2×2 mm. For a cell of 4×4 mm dominant unit rasterization gives a smaller rasterizing error than central point rasterization. Map complexity influences the rasterizing error equally for both methods. The double-conversion method provides a better estimate of the rasterizing error than the equations proposed by Switzer (1975) and Goodchild (1980).

References

BREGT, A. K., and WOPEREIS, M. C. S., 1990, Comparison of complexity measures for choropleth maps. *The Cartographic Journal*, **27**, 85.

BURROUGH, P. A., 1986, *Principles of Geographical Information Systems for Land Resources Assessment* (Oxford: Clarendon Press).

CANTERS, K. Y., DEN HERDER, C. P., DE VEER, A. A., VEELENTURF, P. W. M., and DE WAAL, R. W., 1991, Landscape—ecological mapping of The Netherlands. *Landscape Ecology*, **5**, 145.

GOODCHILD, M. F., 1980, Fractals and the accuracy of geograpical measures. *Mathematical Geology*, **12**, 85.

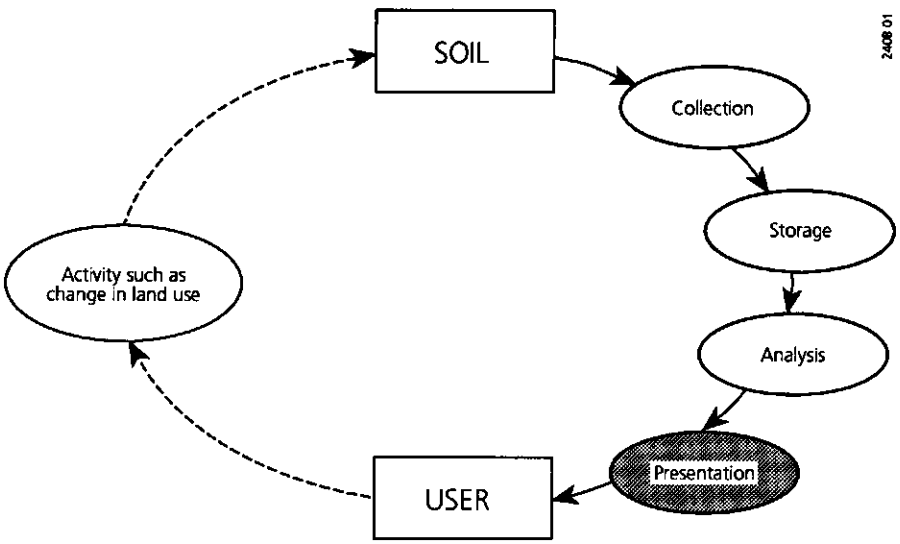
MÜLLER, J. C., 1977, Map gridding and cartographic errors: a recurrent argument. *The Canadian Geographer*, **2**, 152.

SWITZER, P., 1975, Estimation of the accuracy of qualitative maps. *Display and Analysis of Spatial Data*, edited by J. C. Davis and M. J. McCullagh (New York: Wiley), pp. 1–13.

WEHDE, M. E., DALSTED, K. J., and WORCESTER, B. K., 1980, Resource applications of computerized data processing: the areas example. *Journal of Soil and Water Conservation*, **1**, 36.

8 Construction of isolinear maps of soil attributes with empirical confidence limits

Published in Soil Science Society of America Journal, 55:14-19 (1991)



Construction of Isolinear Maps of Soil Attributes with Empirical Confidence Limits

A. K. Bregt,* A. B. McBratney, and M. C. S. Wopereis

ABSTRACT

Moisture-deficit data were obtained at 398 point locations in the Mander area in the Netherlands by simulation using soil-profile attributes. From this data set, a test subset of 75 points was selected at random. The test set was used to select the best of four prediction procedures: ordinary and universal kriging with the original data set, and ordinary and universal kriging with a log-transformed data set. The mean square error of prediction (MSEP) was used to evaluate prediction quality. The lowest MSEP of prediction was obtained using ordinary kriging of the untransformed data. Ordinary kriging was then used to predict moisture deficits and estimate their associated kriging variances at points on a 50 by 50 m grid. The average ratio of actual square errors of prediction to the estimated kriging variances at the 75 test locations was used to adjust the kriging variance estimates on the regular grid to get more realistic estimates. These empirically derived, more realistic estimated kriging variances were then used to construct an isolinear map of moisture deficit with confidence limits. The resulting map showed that, in most of the region under study, the 90% confidence interval for moisture deficit contains 15 mm. This type of map allows the user to obtain confidence limits as well as the predicted value for any point on the map.

AN ISOLINEAR OR ISARITHMIC MAP is a planimetric graphic representation of a three-dimensional volume (Dent, 1985). The lines on an isolinear map connect points with equal value. There are a large number of named isolines, e.g., isohypse or contour for elevation, isobar for atmospheric pressure, isotherm for temperature, and isobront for occurrence of thunderstorms (Thrower, 1972).

In soil science, isolinear maps have been used by various authors for the presentation of continuous soil attributes; they have not yet obtained wide recognition, however, and are not used routinely. These maps are usually constructed from soil punctual data in two steps. In the first step, a grid is produced from point observations by interpolation. This can be termed *primary interpolation*. For this primary interpolation, many different techniques may be used, e.g., inverse squared distance interpolation, Akima's interpolation, Laplacian smoothing splines, and kriging (Laslett et al., 1987). The next or *secondary interpolation* stage positions the isolines with respect to the grid. Generally, linear interpolation is used, as it is assumed that the grid in the primary interpolation stage is fine enough for this to be adequate. Kriging has become popular in soil science as the primary interpolation technique. This popularity is partly due to the fact that, besides a prediction grid, a measure of the predictor's precision (estimated kriging variance) is also obtained for every grid point.

A.K. Bregt and M.C.S. Wopereis, Winand Staring Centre for Integrated Land, Soil, and Water Research, P.O. Box 125, 6700 AC Wageningen, the Netherlands; A.B. McBratney, School of Crop Sciences, Univ. of Sydney, New South Wales 2006, Australia. Contribution from the Winand Staring Centre for Integrated Land, Soil, and Water Research. Received 25 Aug. 1989. *Corresponding author.

Published in Soil Sci. Soc. Am. J. 55:14–19 (1991).

Isolinear maps of soil attributes produced by kriging and associated estimated kriging variance maps have been published by, among others, Burgess and Webster (1980a,b), Byers and Stephens (1983), Dubrule (1984), McBratney et al. (1982), Royle et al. (1981), Webster and McBratney (1987), Webster and Burgess (1980), and Yost et al. (1982). They all present the estimated kriging variances on a separate map. These maps can be used to *identify* sparsely sampled areas.

Some authors (Philip and Watson, 1986; Srivastava, 1986; Henley, 1987) doubt the value of the estimated kriging variance. Laslett et al. (1987) showed that the sum of the estimated kriging variances at 64 test sites for soil pH (H_2O and $CaCl_2$) was considerably lower than the sum of squared differences between observed values at these 64 test sites and their values predicted from grid points.

The purpose of this study was to develop procedures for obtaining an empirically derived, more realistic estimation of the kriging variance and for the construction of isolinear maps with confidence limits. The kriging technique that was used was selected by comparing the influence of stationary and nonstationary conditions on the kriged predictions.

THEORY

Kriging was used in this study as the primary interpolation or spatial-prediction technique. The technique is based on the theory of regionalized variables (Matheron, 1965, 1971, 1973; Journel and Huijbregts, 1978). Any variable, distributed in space, is by definition regionalized. Examples are geological, hydrological, ecological, and pedological data.

Kriging is carried out in two steps. The first step involves modeling the spatial structure of the regionalized variable. The spatial structure can be described by the semivariogram, in the case of stationary conditions, and by the order of the drift and the generalized covariance function, in the case of nonstationary conditions. In the second step of the kriging procedure, the selected model for the spatial structure is applied to the data set to predict values at desired and (usually) unmeasured locations.

The value $\hat{s}(x)$ at a point x is predicted by a linear combination of the values of n surrounding data points:

$$\hat{s}(x_0) = \sum_{i=1}^{n} \lambda_i s(x_i) \qquad [1]$$

where λ_i is the weight of the ith neighboring value, $\hat{s}(x_0)$ is the predicted value, and $s(x_i)$ is an observed value. Kriging is optimal in the sense that it is the best linear unbiased estimator (BLUE) of $s(x_0)$:

$$E[\hat{s}(x_0) - s(x_0)] = 0 \qquad [2]$$

$$\text{var}\,[\hat{s}(x_0) - s(x_0)] \text{ is a minimum} \qquad [3]$$

An interesting by-product of kriging is the estimated kriging variance. This variance is generally used as a measure of the goodness of prediction. It depends on the data configuration and the model of the spatial structure of the data and is not related to the value of the data points directly (Journel, 1986).

MATERIALS AND METHODS

The Data

The data used in this study were obtained from the Mander area in the eastern part of the Netherlands. In this area, water tables are lowered as a result of groundwater extraction for the drinking-water supply. The necessary data for predicting the influence of groundwater extraction on the production of grassland were obtained by a detailed soil survey in 1985 (Bregt and Beemster, 1989; Wösten et al., 1987). In the Mander area of 404 ha, 528 soil borings were made. The separate soil horizons described from each boring were allocated to various soil physical horizon classes with known physical attributes (Wösten et al., 1985). Using these attributes as input data, the sensitivity of the soil to lowering of the water table was estimated by the dynamic simulation model proposed by De Laat (1980). A water balance was estimated for each boring for successive 10-d periods for a 30-yr period. Moisture deficits were calculated for two hydrological situations, one before water extraction and one after water extraction. For this study, attention was focused on the moisture deficit in millimeters for the 30-yr period after water extraction, which, for the sake of brevity, we refer to here as the *moisture deficit*.

A rectangular subarea of 270 ha was selected out of the irregularly shaped 404 ha. This subarea contained 330 borings but, to limit the boundary effect in the kriging procedure, an additional 68 points were used outside the subarea for kriging (Fig. 1). For the evaluation of the quality of kriging predictions, 75 borings were selected at random from the data set within the subarea to serve as a test set (Fig. 1). The moisture-deficit data without the test points have a mean of 10.5 mm and a standard deviation of 15.5 mm. A stem-and-leaf diagram of the moisture-deficit data without the test points is presented in Figure 2.

Statistical Procedures

Recently, Hamlett et al. (1986) stated that nonstationarity should always be considered when analyzing the spatial variability of soil properties. In our case, there was indeed some reason to suspect a trend in the data (Stoffelsen and van Holst, 1985). Moreover, because of the skewness of the original data set, the possibility of transformation to approximate normality also had to be considered. Bearing these two factors in mind, we carried out four sets of structural analysis: two for the untransformed dat and two for log-transformed data. Data were transformed by $z(x) = \log[s(x) + 0.05]$. The small constant 0.05 was added because the minimum of the data set was 0. For each of the untransformed and log-transformed data sets (323 observations), structural analyses were performed by (i) estimating the semivariogram using the method described by McBratney and Webster (1986) (assuming stationarity), and (ii) estimating the order of the drift, k, and the generalized covariance function by the jackknifing procedure of Delfiner (1976) (assuming nonstationarity).

After having found the four possible structural models, spatial prediction at each of the 75 test locations was performed by ordinary kriging in the case of a semivariogram and by universal kriging in the case of a drift and a generalized covariance function. Ordinary kriging was carried out by a kriging program developed by McBratney (1984); universal kriging was carried out by AKRIP (Kafritsas and Bras, 1981).

A consequence of kriging on log-transformed data (Journel and Huijbregts, 1978) is the need for back-transformation of the resulting predictions. Journel and Huijbregts (1978, p 572) give the following solution:

$$\hat{s}(x) = 10^{\hat{z}(x) + s_k^2/2} - 0.05 \qquad [4]$$

where $\hat{z}(x)$ is the predicted value and s_k^2 is the estimated kriging variance of the log-transformed data.

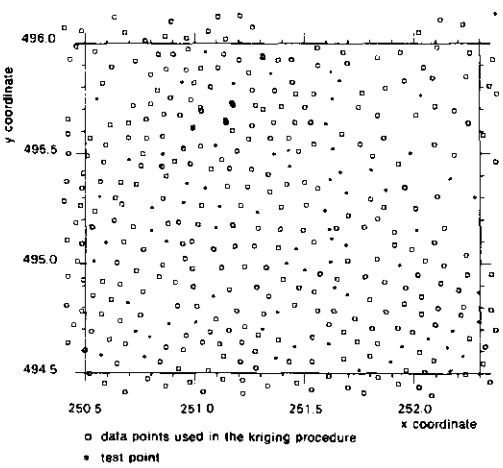


Fig. 1. Map with locations of borings.

Decimal point is at the colon

0: 0000111112223333444555666666777889
1: 000011222222223333333344455556666666666788999999
2: 00001222222233344444444455566777888888
3: 00111112234444444555666666677888899999
4: 011234445556667777778888 9
5: 00112222233444567
6: 12344444688
7: 0001113469
8: 111367779
9: 699
10: 26
11: 0223346678
12: 1222449
13: 11667
14: 234788
15:
16: 0026
17: 1338
18: 002
19: 06
20: 7
21: 6999
22: 9
23: 55
24: 67
25:

High: 27.4 29.2 30.7 30.9 33.4 33.9 34.2 34.2 35 35.1
High: 35.3 35.7 36.8 37.3 37.6 37.9 39.3 39.9 40.5 41.4
High: 41.6 42.7 43.9 44 44.2 45.3 45.4 45.5 48.9 53.7
High: 54.8 57.2 60.6 65.4 67.8 79.1 83.9 90.2 99.1

Fig. 2. Stem-and-leaf diagram of the moisture-deficit data (mm) used in the kriging procedure.

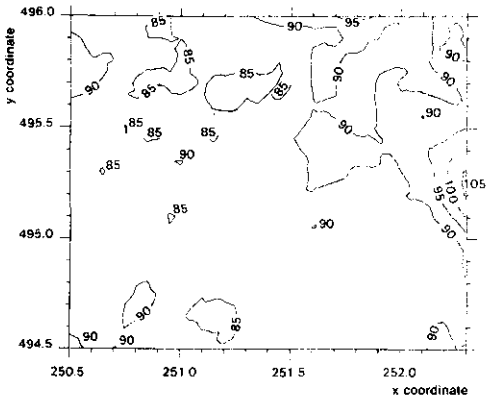


Fig. 4. Isolinear map of kriging variance (mm²) estimated on a 50 by 50 m grid by ordinary point kriging.

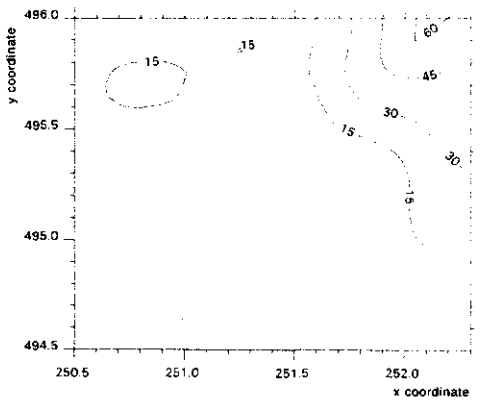


Fig. 5. Isolinear map of moisture deficit (mm) estimated on a 50 by 50 m grid by ordinary point kriging.

underestimation probably arises because kriging assumes the semivariogram or generalized covariance function to be known without error. Clearly, this is not the case and the lack of uncertainty in the semivariogram leads to a marked deflation of the estimated kriging variance relative to the true variance.

An isolinear map of moisture deficit is presented in Fig. 5. The variable shows very little variation over the area studied, except in the upper right corner. In this part of the region, the highest moisture deficits are reached. This is also evident from the perspective view given in Fig. 6. Referring back to Fig. 1, we see that there were only a few test points in this area; this might explain the good performance of ordinary, relative to universal, kriging. In Fig. 7, empirical confidence limits were added to the map as short dashed lines for the lower confidence limits corresponding to the 90% confidence interval, and unevenly dashed lines for the upper confidence limits corresponding to

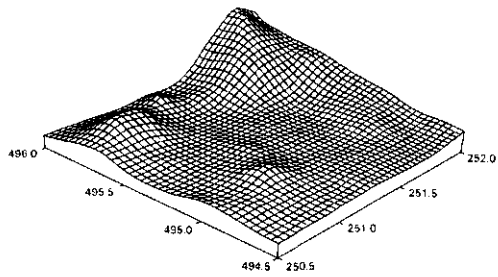


Fig. 6. Perspective view of kriged moisture deficits.

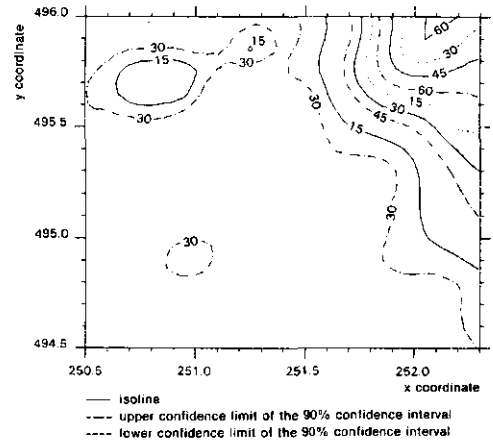


Fig. 7. Isolinear map of moisture deficit (mm) with empirical confidence limits.

the 90% confidence interval. This type of map allows the user to obtain confidence limits, as well as the predicted value, for any point on the map. For example, in Fig. 7, if we look at the point $x = 252.0$ and $y = 495.6$, we can, by visual linear interpolation, obtain a predicted value of 37 by concentrating on the solid isolines only, and values of 14 and 57 for the lower and upper 90% confidence limits by referring to the short and unevenly dashed isolines, respectively. As a check, we see that the confidence limits are approximately symmetric about the predicted value, which they should be using the normal distribution. It can be seen from this map that the 90% confidence interval for moisture-deficit predictions contains 15 mm for almost the complete region under study. This is illustrated in Fig. 8, where this area is shaded. Figure 9 shows the area where the 90% confidence interval for moisture-deficit predictions contains 30 mm.

These results have a great impact on the financial consequences of water extraction. The resulting moisture deficits are translated into yield changes and finally, in practice, into a financial cost to farmers (Wösten et al., 1987). The results show that an accurate estimation of costs is not easily made.

In order to determine the best of the four prediction procedures in practice, we used the statistical relation (Kempthorne and Allmaras, 1986):

$$\text{MSEP} = \text{variance} + (\text{bias})^2 \qquad [5]$$

The MSEP measure was estimated by

$$\text{MSEP} = \frac{1}{75}\sum_{i=1}^{75}[s(x_i) - \hat{s}(x_i)]^2 \qquad [6]$$

where $s(x_i)$ and $\hat{s}(x_i)$ represent actual and predicted values, respectively, at the 75 test locations.

The kriging technique and data set that gave the lowest MSEP value was selected. With the selected technique, moisture deficits were predicted and associated kriging variances were estimated at points on a 50 by 50 m grid.

We calculated the ratio of actual MSEP to average estimated kriging variance at the 75 test locations. The estimated kriging variances on the 50 by 50 m grid were multiplied by this ratio to give more realistic estimates.

The adjusted kriging-variance estimates were used to determine confidence intervals for each point predictor on the 50 by 50 m grid. For the calculation of a confidence interval, a distribution function was also needed. Journel and Huijbregts (1978, p. 15) advised use of the standard Gaussian distribution. In our study, 90% Gaussian confidence intervals ($\pm 1.65 s_k$, where s_k is the realistic estimated kriging standard deviation) were calculated.

Mapping

Once predictions on a regular grid are obtained (at the primary interpolation stage), the position of isolines on the map can be determined (at the secondary interpolation stage). In our study, the location of the isolines was calculated by linear interpolation of adjacent grid points. Reemphasizing the point made above, liner interpolation was only used in the secondary interpolation stage because the grid from the primary interpolation was fine enough to define a prediction surface for which linear interpolation is adequate. Separate maps for predicted moisture deficits, upper confidence limits, and lower confidence limits were plotted. By overlaying and omitting upper- and lower-confidence isolines outside the range of the predicted values, an isolinear map of moisture deficit with empirical confidence limits was obtained.

RESULTS AND DISCUSSION

The structural models for the four prediction procedures are presented in Table 1. Log transformation had little effect on the form of the models. Linear semivariograms with large nugget effects were selected for ordinary kriging for both data sets. For nonstationary kriging, drifts of Order 1 were found with associated generalized covariance functions showing pure nugget effects. Both approaches, i.e., semivariogram and generalized covariance function, suggest a similar model for the data: a gently sloping planar trend across the study area with a large-amplitude short-distance random component.

Estimates of the MSEP of the four kriging procedures are presented in Table 2. Ordinary kriging on the original data gives the lowest MSEP value, followed by universal kriging on the original data.

Log transformation of the data produced higher MSEP values, which perhaps agrees with Puente and Bras' (1986) findings concerning nonlinear estimators. Ordinary kriging performed quite well, although there

Table 1. Selected models for original and log-transformed data using ordinary and universal kriging. The variable h is the distance in km between two points. [$\delta(h) = 1$, if $|h| > 0$, otherwise, $\delta(h) = 0$]

Data transformation	Order of trend	Form of kriging	Fitted model for generalized covariance function $K(h)$ or semivariogram $\gamma(h)$
None	0	Ordinary	$\gamma(h) = 71\ \delta(h) + 138.2h$
None	1	Universal	$K(h) = 81.6\ \delta(h)$
Log	0	Ordinary	$\gamma(h) = 0.172\ \delta(h) + 0.296h$
Log	1	Universal	$K(h) = 0.203\ \delta(h)$

Table 2. Mean square error of prediction (MSEP) of the four kriging procedures.

Form of kriging	Data transformation	MSEP
		mm²
Ordinary	None	151
Universal	None	156
Ordinary	Log	176
Universal	Log	196

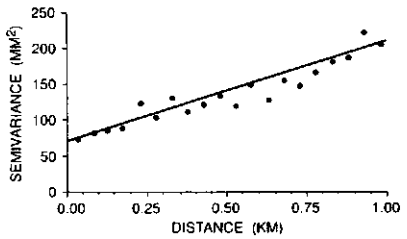


Fig. 3. Semivariogram of moisture deficit used in the selected kriging procedure.

was a significant trend in the data. Similar results were reported by Yost et al. (1982), who stated that ordinary kriging seems to be quite robust to certain degrees of nonstationarity. The predictions on a 50 by 50 m grid were, therefore, made using the simplest method, i.e., ordinary kriging with the original data set. The semivariogram used in the kriging procedure is presented in Fig. 3. A map of the resulting estimated kriging variances is presented in Fig. 4. The variance estimates range from 83 to 106 mm². The largest values are found on the right of the map, which is the result of the lower density of data points in this area (Fig. 1).

The mean estimated kriging variance (MKV) at the 75 test locations using ordinary kriging was 89 mm². This value clearly underestimates the MSEP (Table 2): 151 mm². Laslett et al. (1987) calculated the percentage underestimation (U) of the estimated kriging variances at 64 test sites predicting soil pH (H_2O and $CaCl_2$) using different kriging techniques. The values for U ranged from 22 to 77%. In this study, U has a value of 70%. We decided to adjust the variances on the 50 by 50 m grid by a factor MSEP/MKV = 151/89 to get more realistic kriging-variance estimates. By doing this, we assumed that the kriging predictions were unbiased (bias = 0 in Eq. [5]).

The underestimation of the true kriging variance may arise for several reasons, one of which is the misspecification of the model. We believe, however, the

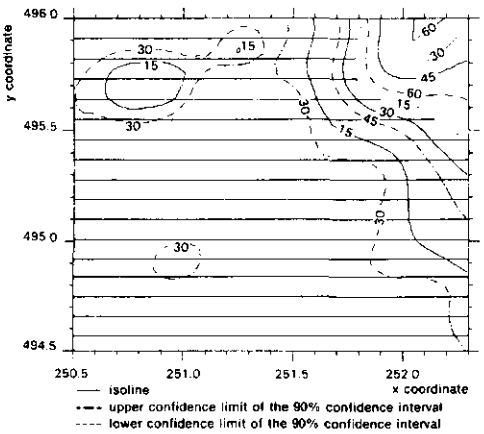


Fig. 8. Map showing the set of all the locations x (shaded) for which the 90% confidence interval for moisture deficit contains 15 mm.

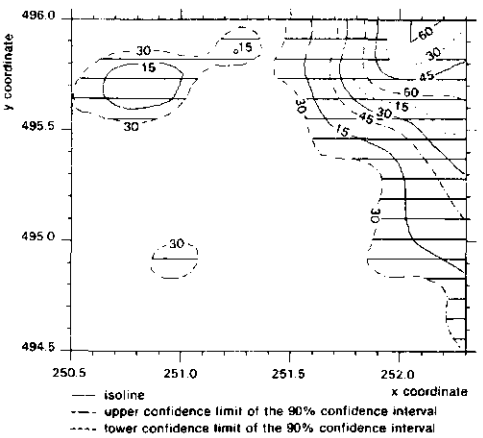


Fig. 9. Map showing the set of all the locations x (shaded) for which the 90% confidence interval for moisture deficit contains 30 mm.

CONCLUSIONS

Moisture deficits of the Mander area in the Netherlands were predicted on a 50 by 50 m grid using ordinary kriging. Universal kriging or log-transformation of the data did not reduce the mean square error of prediction obtained by ordinary kriging at test locations, in spite of a significant trend in the data and the skewness of the distribution. It is salutary to note that the simplest model gave the best predictions, suggesting we should apply Occam's razor even when more complex methods seem appropriate. The ratio of mean square error of prediction to the average estimated kriging variance at the test locations was used to upgrade the kriging variances on the predicted grid to obtain more realistic estimates. We expect that more realistic kriging-variance estimates can be obtained in the future by an iterative procedure in which our modification of using a ratio of square errors is only the first step. Where estimates of kriging variances are required to obtain realistic uncertainty information on predictions, we suggest following a procedure such as the one presented here, at least until further theoretical advances can be made. The realistic kriging-variance estimates were used to construct an isolinear map with 90% confidence limits, assuming a normal distribution of the experimental errors. This type of map has great value in showing the uncertainty in the mapped attribute, allowing users to obtain the predicted moisture deficit, along with its 90% confidence limits, at any point in the region. In our example, the resulting map shows that the 90% confidence interval for moisture deficit contains 15 mm for almost the complete region under study. To obtain a more realistic picture of any mapped attribute, construction of isolinear maps including confidence limits is strongly recommended.

ACKNOWLEDGMENTS

A.B. McBratney would like to thank the Netherlands Soil Survey Institute (STIBOKA) and the Department of Soil Science and Geology of the Agricultural University, Wageningen, for financial support, and D.E. Smiles, CSIRO Division of Soils, Australia, for permission and encouragement to take up the study leave during which this work evolved. The authors kindly acknowledge helpful comments by J. Bouma, J.J. de Gruijter, and A. Stein.

REFERENCES

Bregt, A.K., and J.G.R. Beemster. 1989. Accuracy in predicting moisture deficits and changes in yield from soil maps. Geoderma 43:301–310.

Burgess, T.M., and R. Webster. 1980a. Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. J. Soil Sci. 31:315–331.

Burgess, T.M., and R. Webster. 1980b. Optimal interpolation and isarithmic mapping of soil properties. II. Block kriging. J. Soil Sci. 31:333–341.

Byers, E., and D.B. Stephens. 1983. Statistical and stochastic analyses of hydraulic conductivity and particle size in a fluvial sand. Soil Sci. Soc. Am. J. 47:1072–1081.

De Laat, P.J.M. 1980. Model for unsaturated flow above a shallow water table, applied to a regional sub-surface flow problem. Agric. Res. Rep. 895. Pudoc, Wageningen, the Netherlands.

Delfiner, J.P. 1976. Linear estimation of non-stationary spatial phenomena. p. 49–68. *In* M. Guarascio et al. (ed.) Advanced geostatistics in the mining industry. NATO Advanced Study Inst. Ser. D. Reidel Publ., Boston.

Dent, B.D. 1985. Principles of thematic map design. Addison-Wesley Publ. Co., Reading, MA.

Dubrule, O. 1984. Comparing splines and kriging. Comput. Geosci. 10:327–338.

Hamlett, J.M., R. Horton, and N.A.C. Cressie. 1986. Resistant and exploratory techniques for use in semivariogram analyses. Soil Sci. Soc. Am. J. 50:868–875.

Henley, S. 1987. Letter to the editor: Kriging—blue or pink? Math. Geol. 19:155–158.

Journel, A.G. 1986. Geostatistics: Models and tools for the earth sciences. Math. Geol. 18:119–140.

Journel, A.G., and C.J. Huijbregts. 1978. Mining geostatistics. Academic Press, London.

Kafritsas, J., and R.L. Bras. 1981. The practice of kriging. Rep. no. 263. Mass. Inst. Technol., Cambridge.

Kempthorne, O., and R.R. Allmaras. 1986. Errors and variability of observations. p. 1–31. *In* A. Klute (ed.) Methods of soil analysis. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.

Laslett, G.M., A.B. McBratney, P.J. Pahl, and M.F. Hutchinson. 1987. Comparison of several spatial prediction methods for soil pH. J. Soil Sci. 38:325–341.

Matheron, G. 1965. Les variables regionalisées et leur estimation. Masson, Paris.

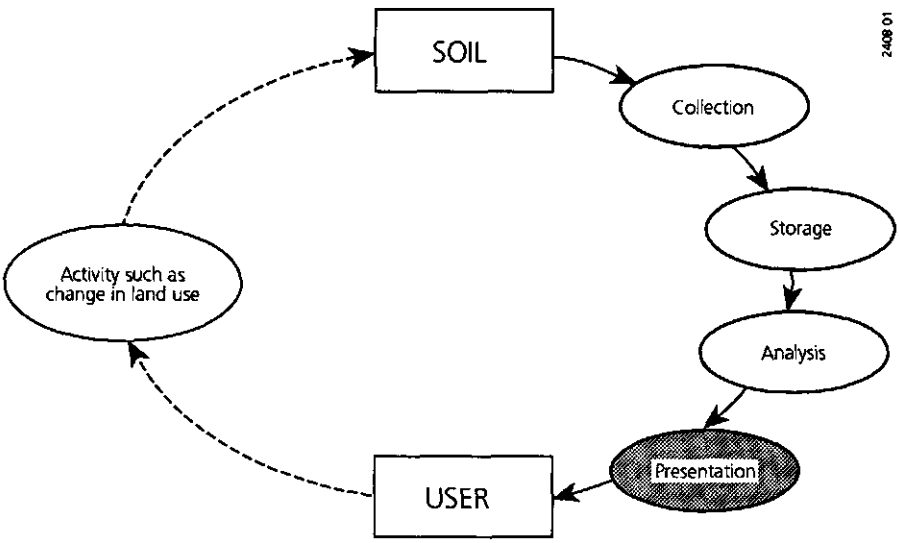
Matheron, G. 1971. The theory of regionalized variables and its applications. Cah. Cent. Morphol. Math. Fontainebleau, No. 5. Ecole Nationale Superieure des Mines de Paris.

Matheron, G. 1973. The intrinsic random functions and their applications. Adv. Appl. Prob. 5:439–468.

McBratney, A.B. 1984. Geostatistical soil survey. Ph.D. diss. Univ. of Aberdeen, Scotland.
McBratney, A.B., and R. Webster. 1986. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. J. Soil Sci. 37:617–639.
McBratney, A.B., R. Webster, and R.G. McLaren, and R.B. Spiers. 1982. Regional variation of extractable copper and cobalt in the topsoil of south-east Scotland. Agronomie (Paris) 2:969–982.
Philip, G.M., and G.S. Watson. 1986. Matheronian geostatistics—quo vadis? Math. Geol. 18:93–117.
Puente, C.E., and R.L. Bras. 1986. Disjunctive kriging, universal kriging, or no kriging: Small sample results with simulated fields. Math. Geol. 18:287–305.
Royle, A.G., F.L. Clausen, and P. Frederiksen. 1981. Practical universal kriging and automatic contouring. Geo-Processing 1:377–394.
Srivastava, R.M. 1986. Philip and Watson—quo vadunt? Math. Geol. 18:141–146.
Stoffelsen, G.H., and A.F. van Holst. 1985. Bodemkundig-hydrologisch onderzoek in het waterwingebied Mander, fase II: De bodemgesteldheid, de voormalige en huidige hydrologische situatie en de grondwaterstands-verlaging. Rapport 1807. Netherlands Soil Surv. Inst., Wageningen.
Thrower, N.J.W. 1972. Maps and man. Prentice-Hall, New York.
Webster, R., and T.M. Burgess. 1980. Optimal interpolation and isarithmic mapping of soil properties. III. Changing drift and universal kriging. J. Soil Sci. 31:505–524.
Webster, R., and A.B. McBratney. 1987. Mapping soil fertility at Broom's Barn by simple kriging. J. Sci. Food Agric. 38:97–115.
Wösten, J.H.M., M.H. Bannink, and J. Bouma. 1987. Land evaluation at different scales: You pay for what you get! Soil Surv. Land Eval. 7:13–24.
Wösten, J.H.M., J. Bouma, and G.H. Stoffelsen. 1985. Use of soil survey data for regional soil water simulation models. Soil Sci. Soc. Am. J. 49:1238–1244.
Yost, R.S., G. Uehara, and R.L. Fox. 1982. Geostatistical analysis of soil chemical properties of large land areas. II. Kriging. Soil Sci. Soc. Am. J. 46:1033–1037.

9 Mapping the conditional probability of soil variables

Accepted for publication in Geoderma



Mapping the conditional probability of soil variables

A.K. BREGT[1], H.J. GESINK[1] and ALKASUMA[2]

[1] *DLO The Winand Staring Centre for Integrated Land, Soil and Water Research (SC-DLO), P.O.Box 125, 6700 AC Wageningen, The Netherlands.*

[2] *Centre for Soil and Agroclimate Research, Jl. Juanda 98, Bogor 16123, Indonesia.*

ABSTRACT

In a study area of 410 ha, located in a region with acid sulphate soils in Indonesia, 790 soil observations were made at a regular grid of 75 x 75 m. Additionally, 74 test observations were made, randomly distributed over the area. The variable "depth to the pyritic layer" was selected for this study.

Two different methods were presented for the production of conditional probability maps of this variable. The first method is based on the use of kriging. The second is based on the use of the inverse distance spatial prediction technique and test points. Both techniques were used to predict values of the variable for a grid of 25 x 25 m. Besides predicted values for every grid point, an estimation of the prediction error, expressed as the standard deviation, was also made. Kriging produces this standard deviation automatically. With the inverse distance technique the standard deviation was estimated using test borings. The standard deviations obtained were used to calculate confidence intervals for each grid point.

Results are presented in maps showing the probability that the depth to the pyritic layer exceeds 50 cm. Only small differences were found between the probability maps produced using the two procedures.

INTRODUCTION

In soil science, maps are generally used to present the spatial distribution of soil variables. These maps are produced either by the soil surveyor delineating

areas in the field or by spatial prediction from point observations, using a mathematical prediction technique (Bregt et al., 1987; Laslett et al., 1987; Bouma, 1989). Choropleth and isolinear maps are quite popular as map forms. On choropleth maps, areas with equal values for a quantitative variable are separated by boundaries. On isolinear maps points of equal value are connected by lines (Dent, 1985).

When using spatial prediction to produce a map from point observations prediction errors occur. These errors, however, are not presented on the map. The maps produced suggest that the actual situation is being presented, which could be misleading for a map user. The uncertainty in the information should, therefore, be made visible. This would enable land-use planners to take into account the uncertainty of the provided information in the planning process.

Recently, Bregt et al. (1991) and Webster and Oliver (1989) presented methods to incorporate the prediction uncertainty in the map itself. Bregt et al. (1991) constructed isolinear maps with confidence limits using kriging. One disadvantage of their method is that the readability of the map decreases due to the increasing number of lines. Webster and Oliver (1989) use disjunctive kriging to delineate areas where predicted values exceed a threshold and represent this probability on the same map. Using the disjunctive kriging technique the original data is transformed.

In this paper two procedures for producing probability maps are presented. The first uses the simple inverse distance spatial prediction method, however, any other spatial prediction technique could have been used. The second uses the conceptional, more complicated kriging technique without data transformation. The results of both procedures are compared.

The data used in this study were obtained from a survey in an area with acid sulphate soils in Indonesia (South Kalimantan). The variable depth to the pyritic layer was used. The presented study focused on the presentation of this variable in the form of conditional probabilities. Research on the optimum observation density for mapping depth to the pyritic layer is reported by Bregt et al. (1990).

MATERIALS AND METHODS

Study area

The study was made in the Pulau Petak region in Southern Kalimantan, Indonesia. Pulau Petak is located west of the town of Banjarmasin, between the rivers Barito and Kapuas Murung. The Pulau Petak region is about 210 000 ha; 90 km long and 20-30 km wide. The area is part of a coastal plain in which physiographic units such as alluvio-marine plains, levees, coastal ridges, old river beds and peat domes can be distinguished (Janssen et al., 1990).

The area is situated in a wet tropical climate, with an average annual temperature of 27 °C and an average annual precipitation of 2200 mm. The climax vegetation is fresh water swamp and peat forest, with riverine forest on the levees and salt and brackish water vegetation along the coast. Due to human activities most of the natural vegetation has disappeared. In the present situation, reclaimed areas with rice fields, abandoned rice fields and secondary, acid tolerant vegetation can be found.

The conditions along the coast, with mangrove vegetation and continuous marine sedimentation are typical for the formation of pyrite (Pons and Van Breemen, 1982; Dent, 1986). As a result, the soils in the Pulau Petak region are dominated by potential and, after drainage, actual acid sulphate soils.

Soil survey

Within the Pulau Petak region a detailed soil survey was made in a study area of 410 ha (the Belawang area) located in an alluvio-marine plain. In the study area, soil patterns cannot be revealed from landscape features, so the soil surveyor must rely mainly on his soil profile observations. Auger borings were made in a regular grid of 75 x 75 m. To get an impression of the short range variation, additional observations near some of the grid points were made at distances of 1, 5 and 25 m. In total 790 borings were made, which in this article are referred to as "observation points" (193 observations/km^2). Besides the observation points, 74 extra borings were made to serve as a test set (test points). The locations of the test points were selected by a random procedure.

At the observation and test points several soil characteristics were recorded in the field and soil samples were taken for analysis in the laboratory. Potential

and actual acidity were determined for each sample, using the method described by Konsten et al. (1988).

In this study the depth from the surface to the pyritic layer is the main variable of interest. This variable determines to a large extent the agricultural possibilities of the soil. The presence of pyritic material in the rooting zone may, after drainage, cause severe acidification which restricts the development of plant roots.

The following operational definition for the variable depth to the pyritic layer was used:

> "If in a soil horizon the potential acidity exceeds 32 mmol H^+ per 100 g soil, the depth from the surface to the top of this horizon is considered to be the depth to the pyritic layer."

The value of 32 mmol H^+ per 100 g soil is the equivalent of 0.75% total S, used as a criterion for pyritic materials in Soil Taxonomy (Soil Survey Staff, 1975).

Spatial prediction and prediction error

The values of the variable at the observation points were used to predict the value of the variable at unknown points in a grid (25 x 25 m), using kriging and the inverse distance method. Spatial prediction can be described as:

$$\hat{Z}_p = \sum_{i=1}^{n} W_i \cdot Z_i \qquad (1)$$

Where: $\hat{Z}p$ = Predicted value of variable Z at point p;
W_i = Weighting factor assigned to point i;
Z_i = Measured value of variable Z at observation point i;
n = Number of observation points used in the prediction.

The predicted value $\hat{Z}$ at point p is a weighted average of the values of variable Z at n surrounding observation points. In both prediction methods, the 24 closest observation points were used to predict points. The two methods differ in the way the weighting factors are determined.

The predicted value at a point will differ from the true value: the prediction error. The prediction is unbiased if the average prediction error is zero. However, predictions may be scattered widely around the true values. This scatter can be expressed as the prediction error variance S_E^2 (Davis 1986):

$$S_E^2 = \frac{1}{n} \sum_{p=1}^{n} (Z_p - \hat{Z}_p)^2 \qquad (2)$$

Where: Z_p = Measured value of variable Z at test point p.

This prediction error variance is unknown and must be estimated. This was done differently for the two spatial prediction methods.

Inverse distance

The inverse distance method is based on the assumption that the value Z, at a point p that is to be predicted, is more similar to nearby points. The calculation of the weights is based on the distance between the points to be predicted and the observation points (Ripley, 1981).

The inverse distance method provides no information about the prediction error. We used the test points to estimate the prediction error variance. Inverse distance prediction to the test points was made. For each test point the difference between the measured value and the predicted value was calculated to give the prediction error. The distribution of prediction errors of the test points was used to estimate prediction error variance, assumed to be valid for the whole area. This assumption is defensible in a situation with a regular observation grid together with only slight differences in soil variation in the area. Both requirements were met in this study.

The described procedure for obtaining uncertainty estimates is not restricted to the inverse distance method any other spatial prediction technique can be used.

Kriging

In this study ordinary kriging was used. In the kriging procedure the weights to be used in the prediction (see Eq. (1)) are calculated using semivariances (Journel and Huijbregts, 1978; Webster, 1985; Cressie, 1989). The semivariance is a measure of the degree of spatial dependence between observation points at a certain distance h. The sample semivariogram was calculated using the observation points.

In a sample semivariogram, semivariance is known only at discrete points at distance intervals h. The kriging procedure requires semivariances for any distance. For this reason the discrete sample semivariogram must be modelled by a

continuous function. Journel and Huijbregts (1978), Cressie (1985) and McBratney and Webster (1986) describe several possible semivariogram models.

In Eq. (1) an infinite number of possible combinations of weights exist, each of which will give a different prediction and a different prediction error. However, only one combination will give a minimum prediction error. This unique combination of weights is calculated in the kriging procedure. If the model assumptions are correct, kriging produces predictions that have the smallest possible error. At the same time, kriging produces an explicit statement of the magnitude of this error in the form of estimated kriging standard deviation (KSD) for every predicted location.

The semivariogram is of major importance in the kriging procedure. Errors in the semivariogram model not only influence the predicted value itself but also the estimated KSD.

Prediction intervals and mapping

Using the distribution of the prediction error, a confidence interval around the predicted value can be determined for each grid point, within which the true value will fall with a pre-specified probability. The probability of the true value of a grid point falling within a pre-specified interval can also be calculated.

Assuming the prediction error is normally distributed, the standard deviation of the prediction error can be used to calculate prediction intervals (Journel & Huijbregts, 1978; Davis, 1986; Bregt, 1991):

$$P\left(\hat{Z}_i - S_i * z_{\alpha/2} < Z_i < \hat{Z}_i + S_i * z_{\alpha/2}\right) = 1 - \alpha \qquad (3)$$

Where: P = The probability of the true value Z_i falling between the stated limits;
Z_i = The true but unknown value of variable Z at point i;
$\hat{Z}_i$ = The predicted value at point i;
S_i = The standard deviation of the prediction error at point i;
$z_{\alpha/2}$ = The z-score derived from the standard normal distribution, corresponding to significance level $\alpha/2$.

Or, if one-sided probabilities need to be calculated:

$$P\left(Z_i > \hat{Z}_i - S_i * z_{\alpha}\right) = 1 - \alpha \qquad (4)$$

and

$$P\left(Z_i < \hat{Z}_i + S_i * z_\alpha\right) = 1 - \alpha \qquad (5)$$

Now, for each grid point the condition of whether the true value of variable Z - with a certain value of the significance level α - will exceed a certain threshold value or not ("true or false") can be tested. The desired probability, as well as the threshold value, can be varied (Fig. 1).

If in Figure 1, for a given value of α, the minimum value of the prediction of the prediction interval $(\hat{Z}_i - S_i * z_\alpha)$ exceeds the threshold value T, the true value of variable Z at point i will - with a probability of $(1-\alpha) * 100\%$ - exceed the threshold value. In Figure 1 this condition is true for grid point B and false for grid point A.

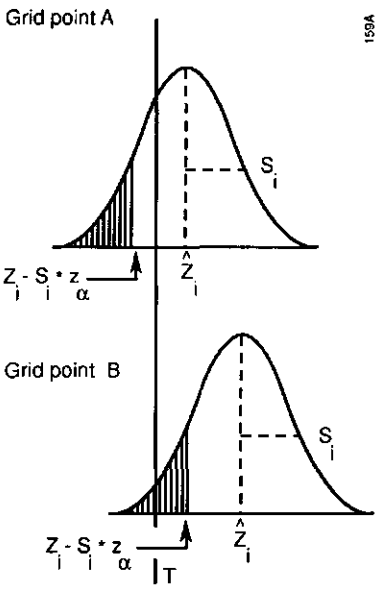


Fig. 1. Testing the condition that the true value (Z_i) of a variable will exceed a threshold value (T).

$\hat{Z}_i$ = Predicted value of variable Z at point i;
S_i = Standard deviation of the prediction error at point i;
z_α = z-score derived from standard normal distribution for significance level α;
T = Threshold value.

When presenting information one must choose between varying the probability (threshold value fixed) and varying the threshold (probability fixed). We carried out the procedure for the first possibility: the probability (1-α) was varied, the threshold value was kept constant. This resulted in a map with delineated areas where, with increasing probability, the true value of the variable exceeded the threshold value. As a threshold value for depth to the pyritic layer 50 cm was chosen. For soil classification purposes this value is relevant (Soil Survey Staff, 1975).

Using the results of both prediction methods, the grid points were mapped for five different values of α. The differences between both methods were evaluated by calculating the number of grid points differently classified in probability classes. The operations needed to derive the maps were made using with the GIS-package ARC/INFO (ESRI, 1989).

RESULTS AND DISCUSSION

Spatial variability of depth to the pyritic layer

The semivariogram of the variable depth to the pyritic layer is presented in Figure 2. A linear model with parameters C_0 (nugget variance) and C/a (slope) was fitted

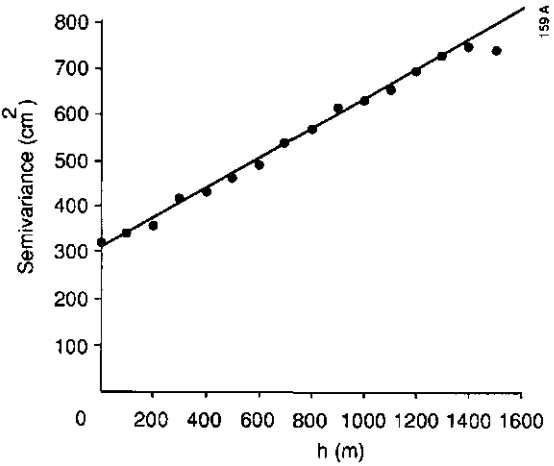


Fig. 2. Semivariogram of depth to the pyritic layer in the study area. Nugget variance (C_0) = 311 cm^2; slope (C/a) = 325.

through the sample semivariances. Several interval sizes and directions were examined. In all cases the linear model showed a good fit for distances of less than 500 m. The large nugget variance indicates measurement errors and/or variation of the variable within the shortest observation distance. The magnitude of the short range variation is illustrated in Figure 3. Differences of more than 50 cm in the depth to the pyritic layer can be found within 25 m. Such high short-range variability is often found in acid sulphate soils due to the genesis of such soils (accumulation of secondary pyrite along root channels) and/or partial oxidation of pyrite along the root system or structure elements (Dent, 1986).

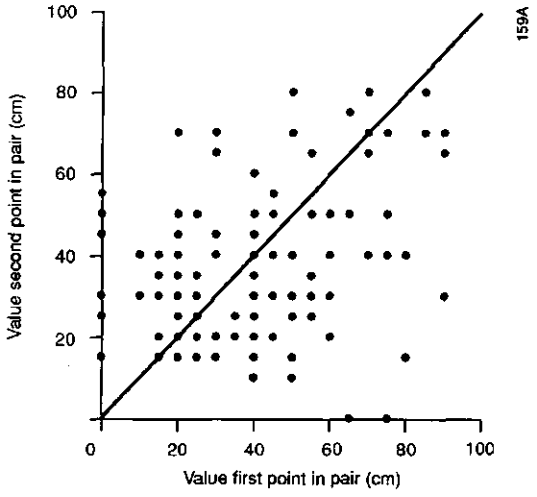


Fig. 3. Scatter diagram of depth to the pyritic layer in the study area for pairs of observation points with a mutual distance of less than 25 m.

Predicted values for grid points

Predictions were made for grid points using kriging and the inverse distance method. Distribution of the differences in predicted values produced by the two prediction procedures is presented in Figure 4. The differences prove to be rather small: in only 3% of the points did the difference exceed 10 cm, and 86% had a difference of less than 2.5 cm. Taking soil description accuracy into account these differences can be neglected. The map with the values obtained by kriging and classified in intervals of 10 cm is presented in Figure 5. A choropleth map

is normally used to present quantitative variables and provides no information about the reliability of the presented data.

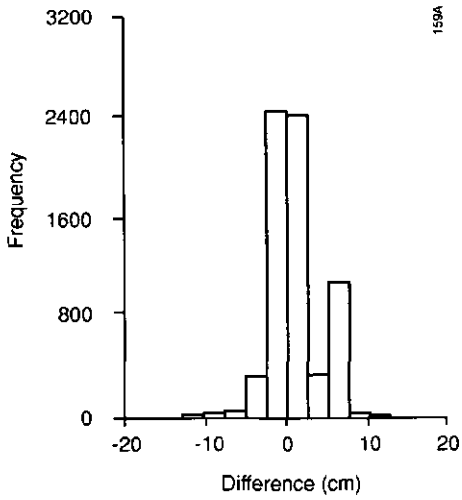


Fig. 4. Distribution of the differences in predicted values of starting depth of the pyritic layer for grid points obtained with kriging and the inverse distance method.

Prediction error

The kriging procedure produces a kriging standard deviation (KSD) for every predicted grid point. The KSD varies with observation density. For the Belawang area with its regular pattern of observation points it was not surprising that the KSD shows little variation, as can be seen in Figure 6. The KSD itself had a standard deviation of a few millimetres. The mean KSD was slightly less than 19 cm, a rather high value.

The results of inverse distance prediction to test points are presented in Figure 7. As mentioned in the previous section it must be assumed that this distribution is valid for the whole area. For a number of reasons this assumption seems acceptable for this data set. First, the mean error of the inverse distance prediction to test points was about zero (Fig. 7). Second, the mean KSD almost equalled the standard deviation of the prediction error resulting from the inverse distance prediction to test points. Third, the KSD hardly varied over the area.

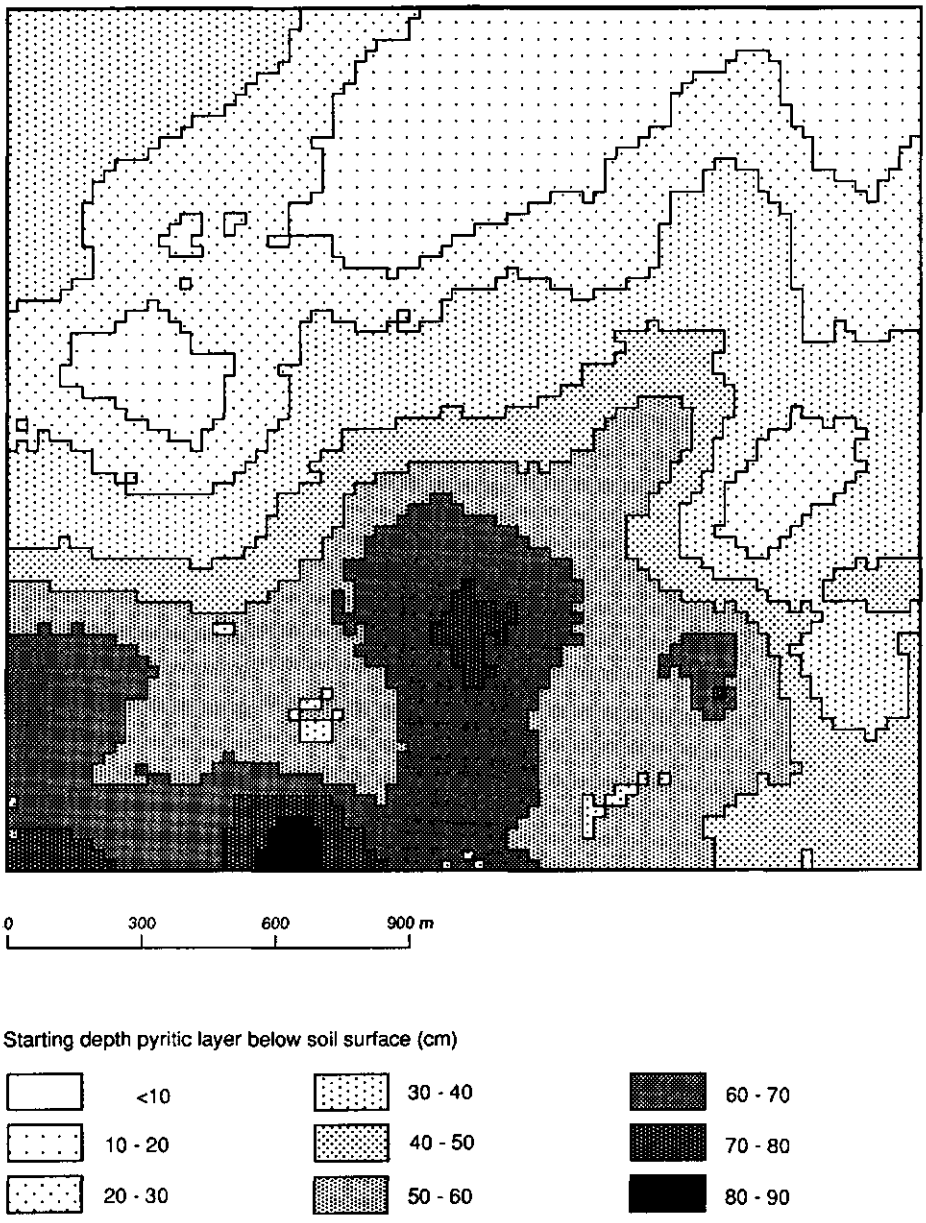


Fig. 5. Map showing the depth to the pyritic layer in the study area, predicted by kriging.

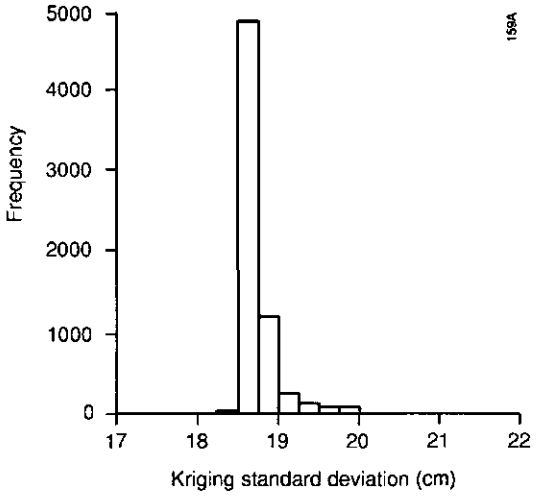


Fig. 6. Distribution of the estimated kriging standard deviation (KSD) for the study area.

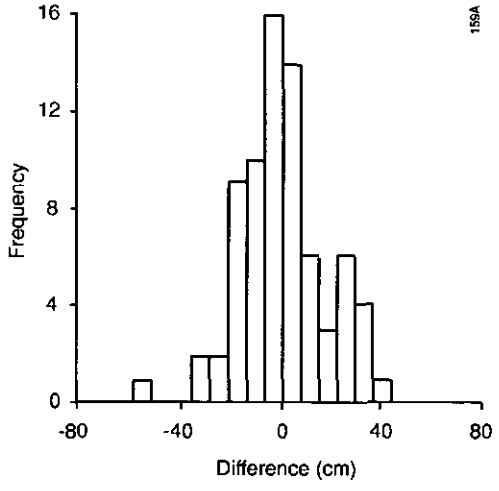


Fig. 7. Distribution of the prediction errors, calculated with inverse distance prediction to measured test points. The distribution has a mean of 0.06 cm and a standard deviation of 18.06 cm.

In Figure 8, the condition of whether or not the true value of the variable with a probability of 80% ($\alpha = 0.2$) will exceed 50 cm is evaluated for two grid points. If 80% of the area under the curve lies to the right of the threshold value it is true (grid point B), otherwise the condition is false (grid point A).

This condition tested for each grid point and the procedure repeated for several values of α, leads, after classification, to the maps shown in Figures 9 and 10. The kriging procedure produced a more smoothed map. In general, both maps give similar information about the variable.

The area on the maps in Figures 9 and 10 where the actual depth to the pyritic layer with a probability of 80 % will exceed 50 cm, is much smaller than the area

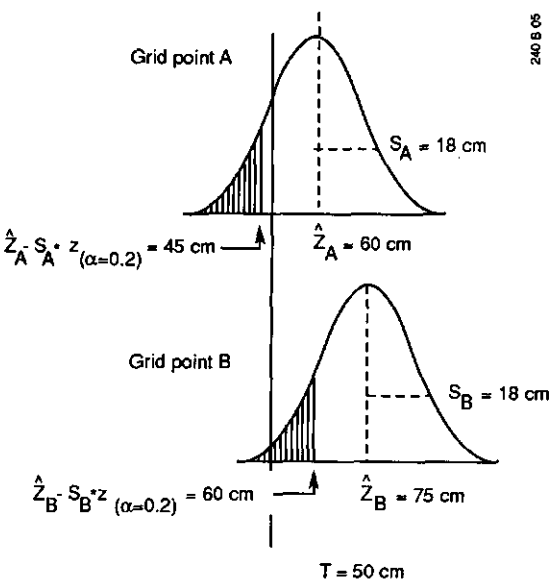


Fig. 8. Testing the condition of whether or not the true value of the variable at a grid point with a probability of 80% will exceed the threshold value of 50 cm.
$\hat{Z}_A$ and $\hat{Z}_B$: Predicted value for grid points A and B;
S_A and S_B : Standard deviation of the prediction error for grid points A and B;
$z_{(\alpha=0.2)}$: z-score of the standard normal distribution for a significance level $\alpha = 0.2$.

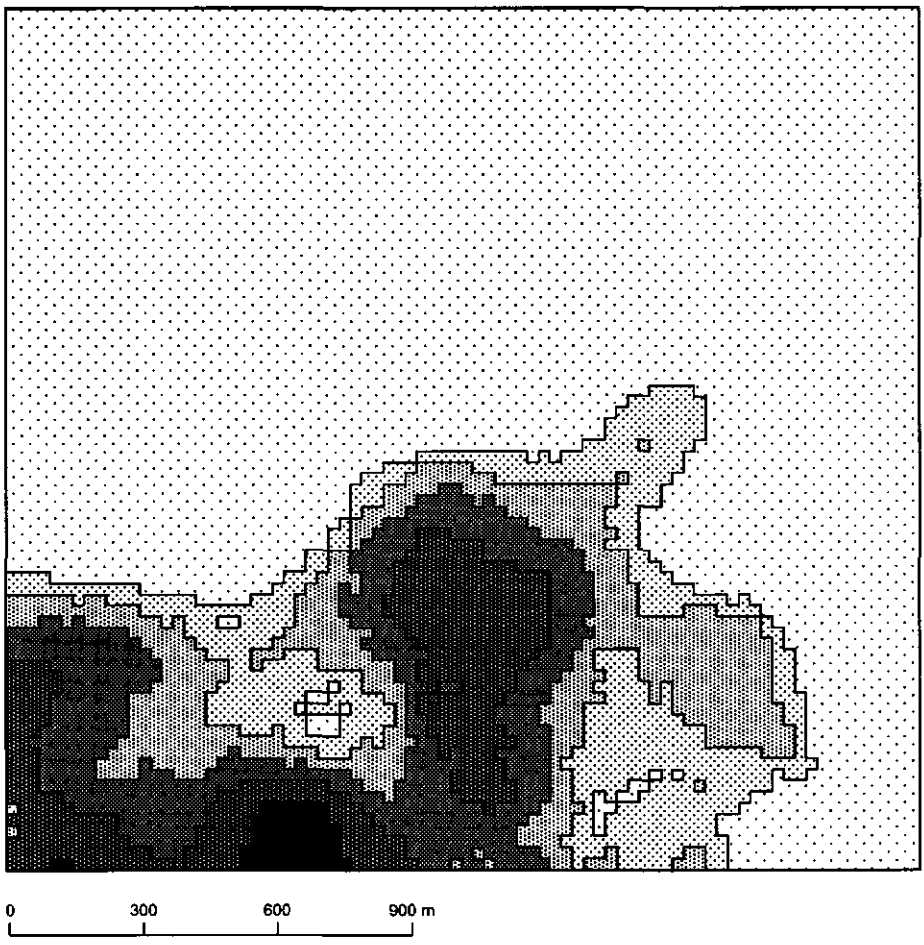


Fig. 9. Map showing probability on depth to the pyritic layer > 50 cm, based on predictions by kriging and the estimated Kriging Standard Deviation (KSD).

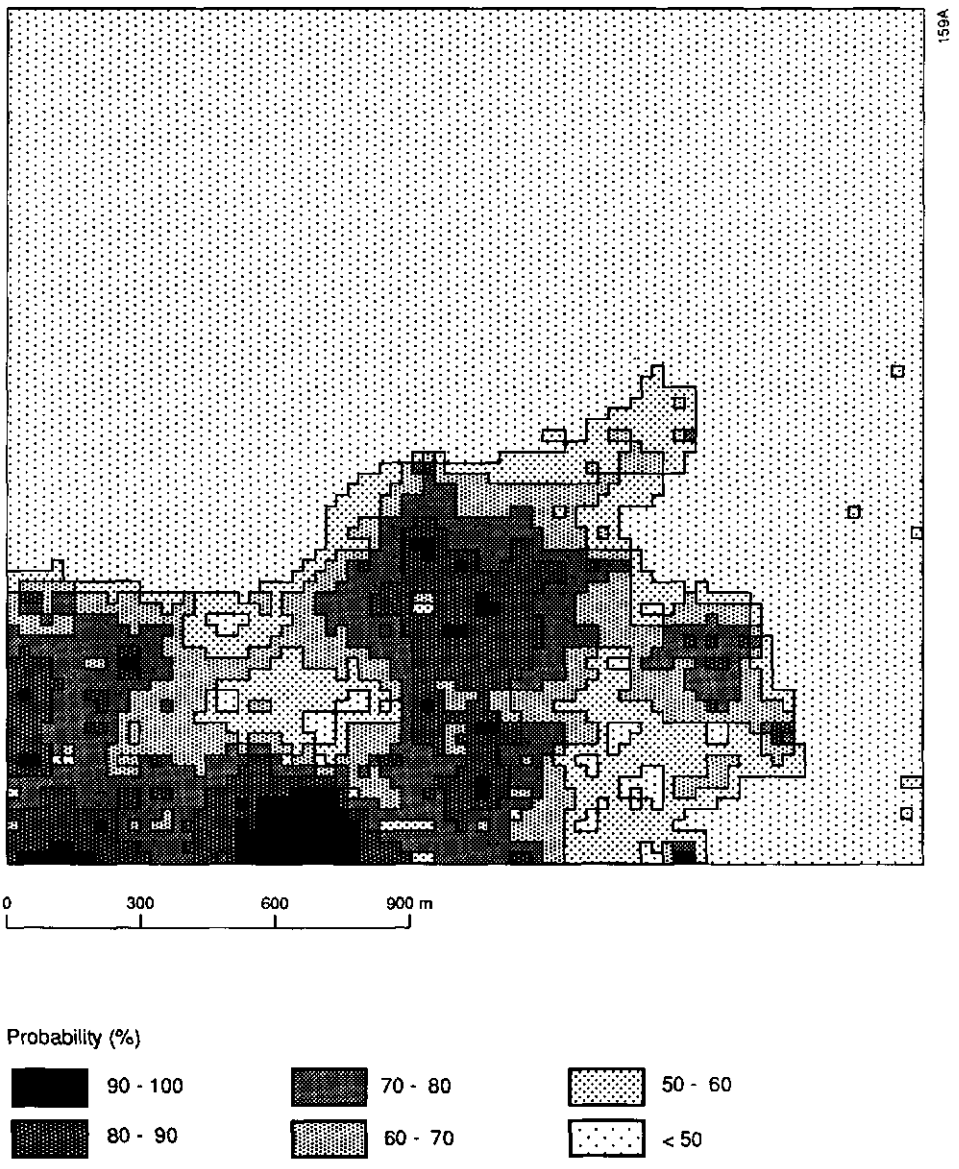


Fig. 10. Map showing probability on depth to the pyritic layer > 50 cm, based on predictions using the inverse distance method and the standard deviation of the prediction errors, calculated using the test points.

mapped on the map in Figure 5 having a depth greater than 50 cm. The conclusion must be that the boundaries on a map based only on predicted values (Fig. 5) are not very reliable, especially when the prediction error is high. This could lead to wrong interpretations by map users.

A possible difficulty in interpreting the maps in Figures 9 and 10 should be mentioned. For instance, in the 80-90% probability class on those maps, the probability of the true value of the variable exceeding the threshold is 80-90% for each separate point. When joining points to areas, the expected value of the area fraction where the true value exceeded the threshold value was also 80-90%. In practice severe deviations from the expected value may occur. A measure of this deviation is the variance of the area fraction exceeding a threshold value. When the individual grid points are dependent, the variance is greater than in the situation where the grid points are independent. The dependent situation occurred in our study.

CONCLUSIONS

The variable depth to the pyritic layer shows a strong variation in space. The boundaries on a map based only on predicted values are not very reliable, especially when the prediction error is high. Mapping this variable in the form of conditional probabilities gives a better picture of the real nature of this variable, which is important for a proper use of the soil data.

Only slight differences for mapping uncertainty were found between kriging and the procedure based on the use of test points and the inverse distance method. In the case of many, regularly distributed, observation points and a rather constant spatial variation of the variable of interest, the proposed procedure can be used as it is cheap and efficient. In other cases, however, the more sophisticated kriging procedure is preferable.

ACKNOWLEDGEMENTS

The authors are grateful to W. Andriesse, J. Bouma, J.J. de Gruijter and J.A.M. Janssen for their helpful comments on an earlier draft of this paper.

REFERENCES

Bregt, A.K., Bouma, J. & Jellinek, M., 1987. Comparison of thematic maps derived from a soil map and from kriging of point data. Geoderma, 39: 281-291.

Bregt, A.K., McBratney, A.B. & Wopereis, M.C.S., 1991. Construction of isolinear maps of soil attributes with empirical confidence limits. Soil Sci. Soc. Am. J., 55: 14-19.

Bregt, A.K., Alkasuma, Janssen, J.A.M., Van de Griendt, J.S. & Andriesse, W., 1990. Optimum observation density for mapping the starting depth of the pyrite layer in Conoco. In: Papers workshop on acid sulphate soils in the humid tropics. AARD & LAWOO, Bogor, p. 62-70.

Bouma, J., 1989. Land qualities in space and time. In: J. Bouma and A.K. Bregt (Editors), Land qualities in space and time. Proceedings of an ISSS symposium, Wageningen, 22-26 August 1988. Pudoc, Wageningen, p. 3-13.

Cressie, N., 1985. Fitting variogram models by weighted least squares. Mathematical Geology, 17: 563-586.

Cressie, N., 1989. Geostatistics. The American Statistician, 43: 197-202.

Davis, J.C., 1986. Statistics and data analysis in geology. John Wiley & Sons, New York, 646 pp.

Dent, B.D., 1985. Principles of the thematic map design. Addison-Wesley, Reading, 398 pp.

Dent, D.L., 1986. Acid sulphate soils: A baseline for research and development. Publication 39, ILRI, Wageningen, 200 pp.

ESRI, 1989. ARC/INFO Users Guide. ESRI, Redlands.

Janssen, J.A.M., Hendro Praseyto and Alkasuma, 1990. Landscape genesis and physiography of Pulau Petak as basis for soil mapping. LAWOO/AARD Scientific Report no. 18, ILRI, Wageningen, 38 pp.

Journel, A.J. & Huijbregts, CH.J., 1978. Mining Geostatistics. Academic Press, Londen, 600 pp.

Konsten, C.J.M., Andriesse, W. & Brinkman, R., 1988. A field laboratory method to determine total potential and actual acidity in acid sulphate soils. In: H. Dost (Editor), Selected papers of the Dakar symposium on acid sulphate soils. Publication 44, ILRI, Wageningen, p. 106-134.

Laslett, G.M., McBratney, A.B., Pahl, P.J. & Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. J. Soil Sci., 38: 325-341.

McBratney, A.B. & Webster, R., 1986. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. J. Soil Sci., 37: 617-639.

Pons, L.J. & Van Breemen, N., 1982. Factors influencing the formation of potential acidity in tidal swamps. In: H. Dost and N. van Breemen (Editors), Proceedings of the Bangkok symposium on acid sulphate soils. Publication 31, ILRI, Wageningen, p. 37-51.

Ripley, B.D., 1981. Spatial Statistics. John Wiley, New York, 252 pp.

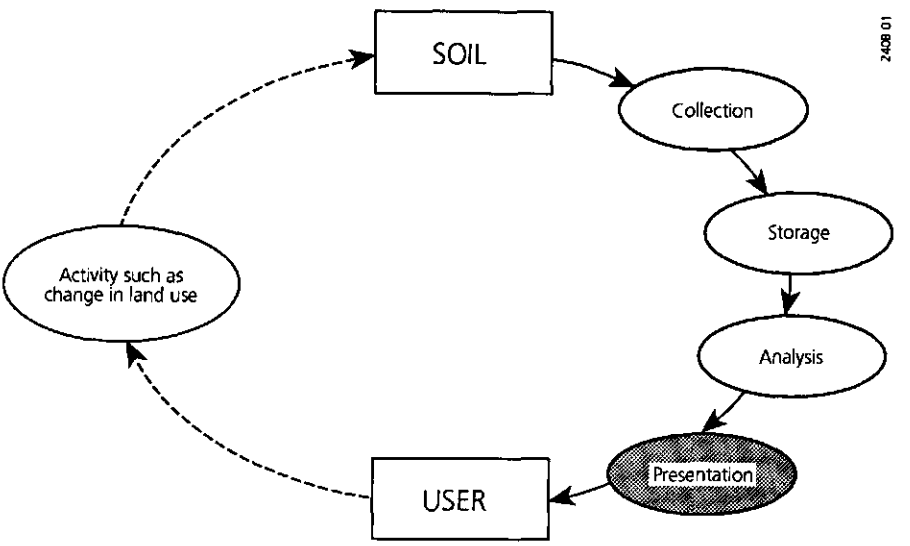
Soil Survey Staff, 1975. Soil Taxonomy. A Basic System of Soil Classification for Making and Interpreting Soil Surveys. USDA-SCS Agric. Handb. 436, Washington, D.C., 754 pp.

Webster, R., 1985. Quantitative spatial analysis of soil in the field. In: B.A. Stewart (Editor), Advances in soil science, Vol. 3. Springer-Verlag, New York, p. 1-70.

Webster, R. & Oliver, M.A., 1989. Optimal interpolation and isarithmic mapping of soil properties. VI. Disjunctive kriging and mapping the conditional probability. J. Soil Sci., 40: 497-512.

10 Comparison of complexity measures for choropleth maps

Published in The Cartographic Journal, 27:85-91 (1990)



Comparison of complexity measures for choropleth maps

A.K. BREGT and M.C.S. WOPEREIS

The Winand Staring Centre for Integrated Land, Soil and Water Research, P.O. Box 125, 6700 AC Wageningen, The Netherlands

ABSTRACT

The complexity of a choropleth map has an important influence on its readability. Six map complexity measures (aggregation index, compactness index, boundary contrast index, fragmentation index, boundary index, size disparity index) described in literature are presented in this study. Correlation coefficients between these indices are calculated, using 90 different maps. It is shown that most of the measures are highly correlated and therefore redundant.

Visual judgement of 13 maps by 28 subjects shows that four indices (aggregation index, compactness index, fragmentation index, boundary index) can serve as reliable indicators of map complexity. As a result, each of these measures can be used to compare pattern complexity among different maps. The fragmentation index seems to be the best choice in view of its simplicity and calculation ease.

INTRODUCTION

Choropleth maps are often used in cartography for presenting spatial information. For the production of a choropleth map, grouping of data into classes is often necessary, although some authors (Tobler, 1973; Peterson, 1979) have proposed the use of classless choropleth maps. Classification has a great impact on the ultimate map. Even with the same number of classes two different classification methods can yield completely different maps. Various methods for data classification in choropleth mapping are presented by e.g. Evans, 1977; Jenks, 1977 and Stevanovic and Vries-Baayens, 1984. It is important to choose a method which

yields on the one hand a map containing relevant information for the user and on the other hand a map which is easy readable. The readability of a map is a function of:

- the presentation quality;
- the map reader's experience with cartographic materials and;
- the complexity of the map pattern (Monmonier, 1974; Dent, 1985).

This study is focused on the complexity of the map pattern. The readability of a map decreases with increasing complexity of the map pattern. Complexity is related to the number, size and distribution of delineated areas on the map. In literature, various measures for map complexity are described. In the study presented here some of these measures are compared in order to select a reliable measure for map complexity.

MEASURES OF MAP COMPLEXITY

Map complexity has been studied by several cartographic researchers (Olson, 1972; Monmonier, 1974; Müller, 1975 and 1976; Chang, 1978; MacEachren, 1982) and various measures of map complexity have been defined, six of which were selected in this study. These six complexity measures are described in detail below. All measures can be used if the individual base enumeration units that compose the choropleth map are known. For raster maps these units are rectangular and of equal size. Only two of the six complexity measures (boundary index and size disparity index) can be used without knowing the base enumeration units of a choropleth map.

We designed six map fragments of 4 by 5 grid cells (MF1 to MF6) to compare the impact of different map patterns on the selected measures (Figure 1).

Cells with the same class that touch each other in a corner point are considered to represent different polygons, unless a mutually adjacent cell also falls in that class. For each map fragment, values of the six complexity measures are calculated (Figure 2).

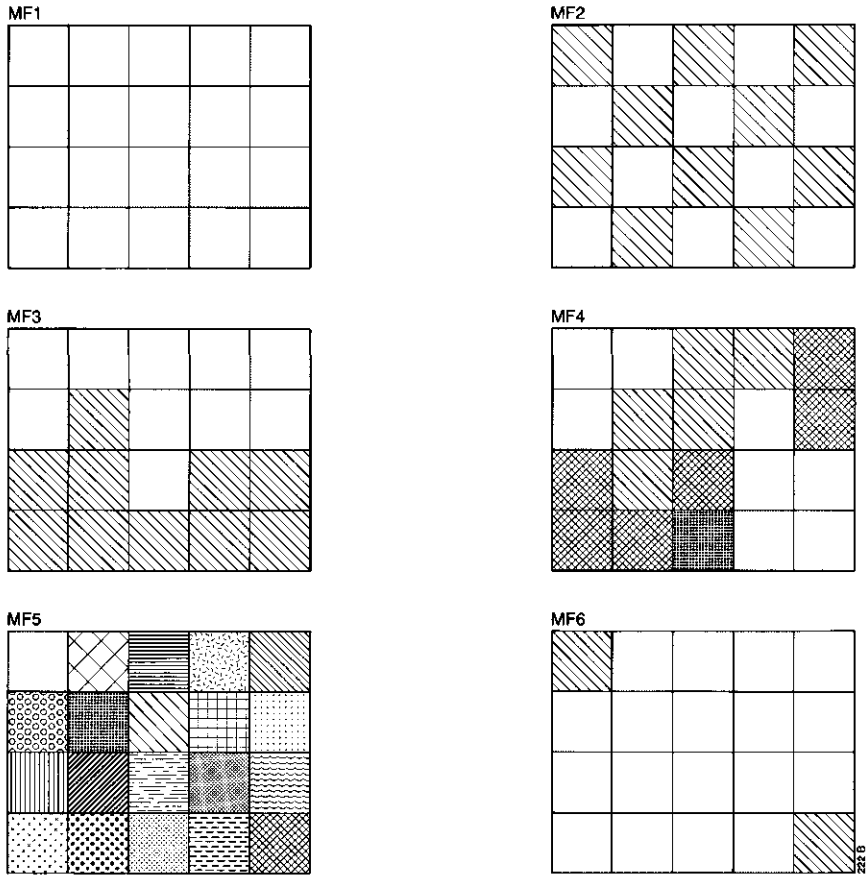


Fig. 1 Six map fragments (MF1 to MF6) of twenty grid cells with different complexities.

Aggregation index

A map pattern looks 'aggregated' if a group of adjacent enumeration units belongs to the same class. Müller (1975) defined the aggregation index (*AG*) as:

$$AG = \sum_{i=1}^{n} \sum_{j=1}^{n} ab_{ij} \Big/ \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \qquad (1)$$

where n = number of enumeration units; $a_{ij} = 1$ if cells i and j are adjacent, $a_{ij} = 0$ when otherwise and when i = j; $ab_{ij} = 1$ if cells i and j are adjacent and belong to the same class; $ab_{ij} = 0$ otherwise and when i = j.

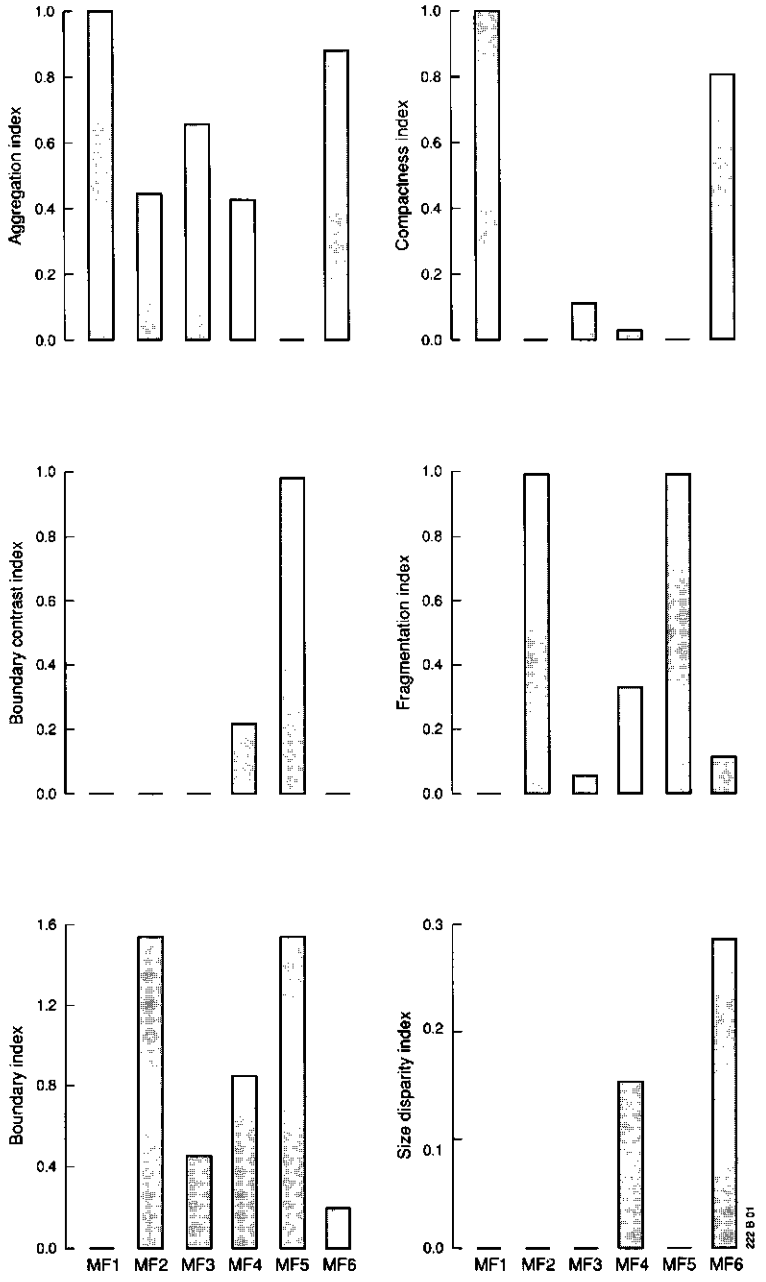


Fig. 2 Histograms of the six complexity measures. Each histogram shows values for one complexity index, calculated for the six map fragments of Figure 1.

The aggregation index ranges between 0 and 1. High values of this index indicate a low complexity of the map pattern.

For MF1 in Figure 2 the maximum value of the aggregation index is reached because all grid cells belong to the same class. Most cells of MF6 fall in one class, which results in a high index value. All cells in MF5 are different and the aggregation index is therefore 0. The indices of MF2 and MF3 are not the same, in spite of the equal distribution of grid cells over the two classes. The aggregation index of the 'chess board pattern' MF2 is lower than the index of MF3 because of the higher amount of contiguous grid cells.

Compactness index

A map can be considered compact if the map pattern consists of clustered regions of three or more enumeration units belonging to the same class. Compactness is a special case of aggregation and is determined by the number and size of clustered regions. A group of i cells belonging to the same class is compact if i - 1 cells are all adjacent to another similar cell and at least i - 2 pairs of those i - 1 surrounding cells are adjacent to each other (Müller, 1975).

We computed the total number of compact regions of three up to and including nine enumeration units. The total compactness (*TCP*) can then be calculated as:

$$TCP = \sum_{i=3}^{g} W(i)*E(i) \qquad (2)$$

where $W(i)$ = total number of cells in compact region and
$E(i)$ = total number of compact regions of i cells.

In this study, a *TCP*-value is divided by the maximum value of *TCP* to derive the compactness index (CP), which ranges between 0 and 1. High values of this index indicate a low complexity of the map pattern.

All the 20 grid cells of MF1 belong to the same class: the compactness index is maximal. MF2 and MF5 both have no group of at least three compact cells and the index is therefore 0. MF6 is very compact, which results in a high compactness index.

Boundary contrast index

Müller (1975) defined map complexity as the total amount of three mutually adjacent enumeration units belonging to three different classes. In this study, this value is divided by its maximum value for a map to derive the boundary contrast index (BC). This index ranges between 0 and 1. High values indicate a high complexity of the map pattern.

MF1, MF2, MF3 and MF6 have only one or two classes and the index value is therefore 0. The maximum value is reached for MF5 because all the 20 grid cells belong to different classes.

Fragmentation index

A map pattern looks fragmented if it consists of a large number of small polygons. The fragmentation index (*FI*) is a standardized measure of this number of polygons and is described by Monmonier (1974):

$$FI = (M - 1)/(N - 1) \quad (3)$$

where M = number of polygons and N = number of enumeration units.

FI varies from 0 to 1. High values indicate a high complexity of the map pattern.

The fragmentation index of MF2 and MF5 is maximal because in these map fragments each grid cell forms one small polygon.

MF1 consists of only one polygon; the index value is therefore 0. The increasing amounts of polygons in MF3, MF6 and MF4 result in increasing values of the fragmentation index.

Boundary index

Bregt et al. (1988) defined the boundary index (BI) as the total length of boundary lines between polygons divided by the total area of the map. The index has a minimum of 0 and no specific maximum value. High index values indicate a high complexity of the map pattern.

MF1 has no boundaries and the index is therefore 0. The highest boundary indices are found for MF2 and MF5, which both consist of 20 individual polygons.

Size disparity index

The disparity index, also described by Monmonier (1974), is a measure of areal inequality of polygons, based on the Lorenz curve (Yeates, 1968, pp. 90-92). First, the area of each polygon is determined. These areas are divided by the total area of the map, so that their sum is 1.0. Next, these proportionate areas are rank-ordered from lowest to highest. A vector of cumulative proportionate areas is obtained. These cumulative proportionate areas are plotted against the ranks. If all the polygons are equal in size, a straight line (diagional) will be the result. If this is not the case, the line will be a curve under the diagonal. The area between diagonal and line is the size disparity index (SD). The index ranges from 0 for polygons of uniform size to almost 0.5 in case of maximum areal inequality of polygons. High values usually indicate a low complexity of the map pattern.

MF1, MF2, MF3 and MF5 all have polygons of the same size and their size disparity indices are therefore 0. The differences in size are especially pronounced in case of MF6, which is indicated by its high index value.

EXPERIMENTAL DESIGN

To compare the different complexity measures, data from a study area in the eastern part of The Netherlands were used. In this area water-tables have been lowered as a result of water extraction for drinking water supplies. For 410 different locations the average moisture deficit for grass under the present hydrological situation (variable A) and the change in grass yield due to water extraction (variable B) were determined (Bregt and Beemster, 1989). With the interpolation method kriging predictions were made on a raster of 31 by 37 raster cells of 50 m x 50 m for both variables. The resulting data were grouped into a maximum of ten classes using different classification techniques. A total of 90 classifications were carried out, resulting in 90 different maps. For each map, the values of the selected six complexity measures were computed. Finally, correlation coefficients between these indices were calculated.

An experiment was conducted to select the best indicator of map complexity. Twenty-eight subjects (most of them without significant knowledge of cartography) were asked to participate in a test on complexity of map patterns. Six four-class

maps of data set A (test set A) and seven four-class maps of data set B (test set B) were used. Both test sets were presented to each subject, one at a time. Each subject was first informed of the nature of the experiment and was then asked the following question: "Please arrange these two sets of four class maps from low to high complexity. A map is considered to be complex if the pattern looks intricate or involved. You have approximately five minutes for this task. Maps that in your opinion do not differ in complexity may have the same order."

Finally, the six measures of pattern complexity were computed and rank-ordered for each test set. The ranked data given by the subjects were compared with the rank orders of each complexity index to select the best indicator of pattern complexity.

STATISTICAL COMPARISON

The correlation coefficients between the six complexity measures of the 90 maps are presented in Table 1. High absolute values indicate a strong correlation. All

Table 1. Correlation coefficients between the various complexity indices. AG: aggregation index; CP: compactness index; BC: boundary contrast index; FI: fragmentation index; BI: boundary index; SD: size disparity index.

	AG	CP	BC	FI	BI	SD
AG	1.00	0.98	-0.82	-0.90	-1.00	0.07
CP	0.98	1.00	-0.69	-0.79	-0.97	-0.04
BC	-0.82	-0.69	1.00	0.98	0.83	-0.29
FI	-0.90	-0.79	0.98	1.00	0.91	-0.24
BI	-1.00	-0.97	0.83	0.91	1.00	-0.08
SD	0.07	-0.04	-0.29	-0.24	-0.08	1.00

the indices, except the size disparity index, are highly correlated. Simultaneous calculation of all of these indices as measures for map complexity is not very reasonable.

The size disparity index clearly represents another aspect of the map pattern than the other selected measures.

RELATIONS BETWEEN VISUAL JUDGEMENT AND COMPLEXITY MEASURES

The results of the complexity judgement of the two test sets by 28 subjects are presented in Tables 2 and 3. The maps in these tables are ordered according to their average ranks (*C*m). This value is calculated according to:

$$Cm = 1/28 \sum_{1=1}^{n} Ci*Pi \qquad (4)$$

where n = number of maps in test set; *C*i = place in complexity order and *P*i = number of subjects assigning a particular map to place *C*i.

Table 2. Number of subjects assigning a place in the complexity order (1 to 6) to a particular map for test set A; *C*m = mean complexity order (see text).

Map	1	2	3	4	5	6	*C*m
A1	28	0	0	0	0	0	1.0
A2	6	22	0	0	0	0	1.8
A3	0	0	27	1	0	0	3.0
A4	0	0	1	15	9	3	4.5
A5	0	0	0	12	14	2	4.6
A6	0	0	0	0	5	23	5.8

A1 and B2 are the least complex maps (lowest *C*m) and A6 and B7 are the most complex maps (highest *C*m). In test set A, map A1 was clearly the least complex one. All the subjects agreed on this, although six persons did not see any

difference with map A2. The differences between maps A4 and A5, and between B4 and B5 are small. Most subjects found ordering of test set B more difficult than test set A: this can also be concluded from the larger spread of the subjects over the complexity classes. As an illustration two maps of test set A are presented in Figure 3.

Table 3. Number of subjects assigning a place in the complexity order (1 to 7) to a particular map for test set B; *Cm* = mean complexity order (see text).

Map	1	2	3	4	5	6	7	*Cm*
B1	22	6	0	0	0	0	0	1.2
B2	7	21	0	0	0	0	0	1.8
B3	0	0	26	1	0	0	1	3.2
B4	0	0	6	10	7	5	0	4.4
B5	0	0	5	11	6	6	0	4.5
B6	0	0	4	8	5	8	3	4.9
B7	0	0	3	2	7	6	10	5.6

In Table 4 the values of the complexity indices for the 6 maps of test set A are given. The maps in this table are ordered according to increasing complexity, as experienced by the subjects. A1 is the least complex map and A6 the most complex. The ordering of the calculated complexity indices according to increasing complexity is presented in brackets. The lowest complexity is indicated by 1. Similar ordering of calculated measures and maps indicates a correct prediction of the experienced complexity by the complexity measures. This is the case with the aggregation index, compactness index, fragmentation index and the boundary index. The boundary contrast index has a value of 0 for four of the six maps. Therefore, this measure is not very useful, especially not for the least complex maps. The size disparity index does not show a clear relationship with the map sequence.

In Table 5 the values of the complexity indices for the seven maps of test set B are given. The complexities of the maps B1, B2 and B3 are very well estimated by the aggregation index, compactness index, fragmentation index and the boundary

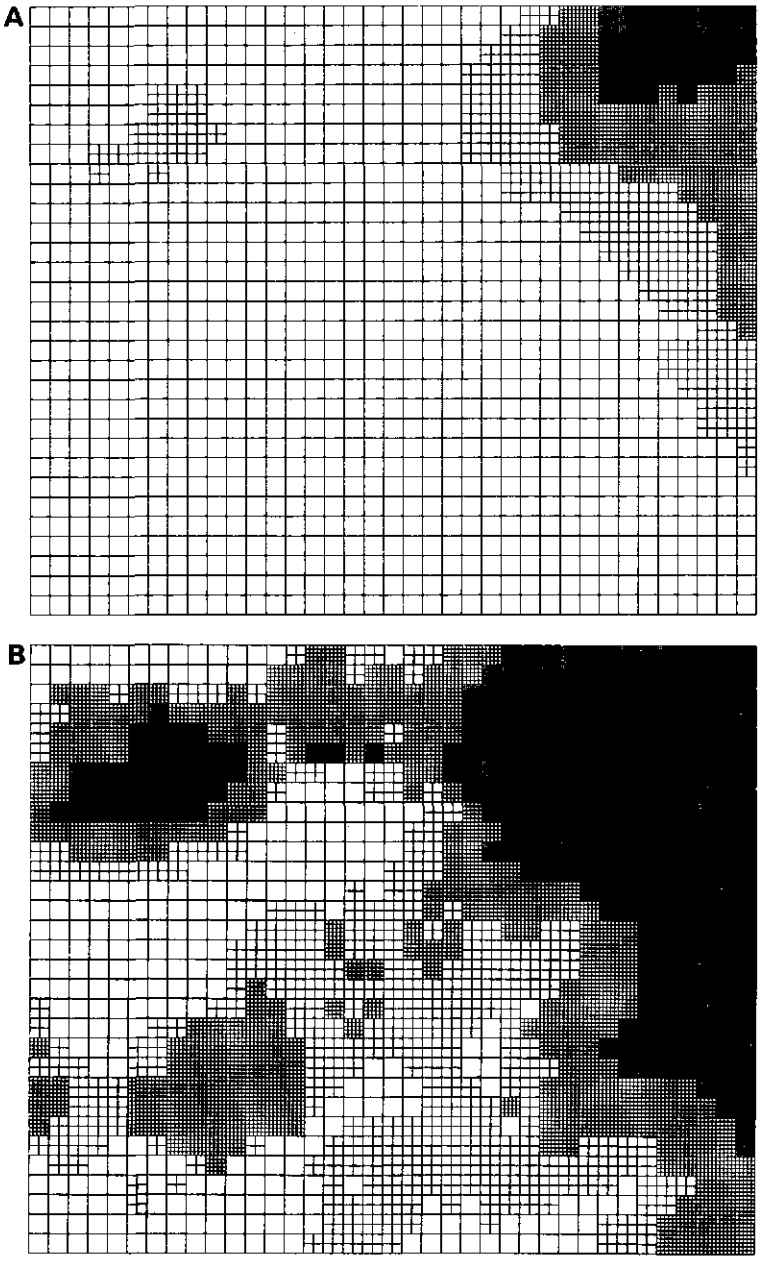


Fig. 3 a) Map of test set A (A1): this map is considered to be least complex by the subjects.
b) Map of test set A (A6): this map is considered to be most complex by the subjects.

index. The complexities of B4 and B5, however, are confused. This is understandable because the average ranks (*Cm*) of B4 and B5 are almost equal. The two most complex maps are ordered correctly by the fragmentation index and boundary contrast index only. The complexity sequence given by the size disparity index again differs clearly from that given by the subjects.

Table 4. Complexity indices computed for the test maps A1 to A6. The subjects considered A1 as the least, and A6 as most the complex map. Each index is ordered according to increasing computed complexity. For each index the place of each map in this order is given in brackets. AG: aggregation index; CP: compactness index; BC: boundary contrast index; FI: fragmentation index; BI: boundary index; SD: size disparity index.

Map	AG	CP	BC	FI	BI	SD
A1	0.937 (1)	0.842 (1)	0 (1)	0.003 (1)	2.061 (1)	0.325 (6)
A2	0.928 (2)	0.815 (2)	0 (1)	0.004 (2)	2.342 (2)	0.342 (5)
A3	0.868 (3)	0.684 (3)	0 (1)	0.012 (3)	4.309 (3)	0.404 (3)
A4	0.852 (4)	0.639 (4)	0 (1)	0.019 (4)	5.089 (4)	0.423 (1)
A5	0.802 (5)	0.529 (5)	0.004 (2)	0.020 (5)	6.681 (5)	0.380 (4)
A6	0.751 (6)	0.451 (6)	0.020 (3)	0.043 (6)	8.511 (6)	0.406 (2)

Table 5. Complexity indices computed for the test maps B1 to B7. The subjects considered B1 as the least, and B7 as most the complex map. Each index is ordered according to increasing computed complexity. For each index the place of each map in this order is given in brackets. AG: aggregation index; CP: compactness index; BC: boundary contrast index; FI: fragmentation index; BI: boundary index; SD: size disparity index.

Map	AG	CP	BC	FI	BI	SD
B1	0.832 (1)	0.591 (1)	0 (1)	0.018 (1)	5.771 (1)	0.412 (1)
B2	0.827 (2)	0.569 (2)	0 (1)	0.019 (2)	5.836 (2)	0.411 (2)
B3	0.779 (3)	0.477 (3)	0.003 (2)	0.024 (3)	7.464 (3)	0.384 (5)
B4	0.749 (5)	0.444 (4)	0.012 (4)	0.034 (5)	8.612 (5)	0.383 (6)
B5	0.753 (4)	0.444 (4)	0.008 (3)	0.031 (4)	8.487 (4)	0.373 (7)
B6	0.732 (7)	0.425 (6)	0.015 (5)	0.042 (6)	9.261 (6)	0.393 (3)
B7	0.735 (6)	0.438 (5)	0.020 (6)	0.052 (7)	9.261 (6)	0.390 (4)

Combining the results of Tables 4 and 5, it is clear that aggregation index, compactness index, fragmentation index and boundary index are good indicators for the experienced complexity by the subjects. Between these four measures also a strong correlation exists, as is indicated by the calculated correlation coefficients (Table 1). As a result of this strong correlation, calculation of one measure is enough for characterizing the complexity of the map pattern. The fragmentation index appears to be the best choice in view of its simplicity and ease of calculation.

Sometimes the total number of base enumeration units in a choropleth map is unknown. In this case, neither the fragmentation index nor the aggregation index nor the compactness index can be calculated. Then, the complexity of the map pattern can be described by the boundary index, because this measure is not based on the original individual enumeration units.

CONCLUSIONS

When producing a choropleth map it is important to pay attention to its readability. One way of increasing the readability of the map is to limit the complexity of the map pattern, which can be quantified by a complexity measure. The fragmentation index is the most suitable measure.

One has to be aware that a too strong simplification of the map pattern might result in the loss of valuable information for the map reader. The challenge is to produce a readable map, which contains the desired information. Evaluation of a complexity measure during the process of map design is strongly recommended.

REFERENCES

Bregt, A.K., L.A. Ankum, J. Denneboom & Y. van Randen, 1988. Gridded soil map of the Netherlands: gridding error. In R.J.A. Jones and B. Biagi (Eds) Proceedings of a workshop on Computerisation of land use Data. 1987. Commission of the European Communities, Pisa, Italy, p. 110-117.

Bregt, A.K. & J.G.R. Beemster, 1989. Accuracy in predicting moisture deficit and changes in yield from soil maps. Geoderma, 43: 301-310.

Chang, K., 1978. Visual Aspects of Class Intervals in Choroplethic Mapping. The Cartographic Journal, 15: 42-48.

Dent, B.D., 1985. Principles of Thematic Map Design. Addison-Wesley, Reading.

Evans, I.S., 1977. The selection of class intervals. Transaction Institute of Britisch Geographers, 2: 98-124.

Jenks, G.F., 1977. Optimal data classification for choropleth maps. Occasional Paper No. 2, Dept. of Geography, University of Kansas.

MacEachren, A.M., 1982. Map complexity: Comparison and Measurement. The American Cartographer, 9: 31-46.

Monmonier, M.S., 1974. Measures of Pattern Complexity for Choroplethic Maps. The American Cartographer, 1: 159-169.

Müller, J.C., 1975. Definition, measurement, and comparison of map attributes in choroplethic mapping. In Proceedings of the Association of American Geographers, 7: 160-164.

Müller, J.C., 1976. Objective and Subjective Comparison in Choroplethic Mapping. The Cartographic Journal, 13: 156-166.

Olson, J., 1972. Autocorrelation as a measure of map complexity. In Proceedings of the American Congress on Surveying and Mapping, p. 111-119.

Peterson, M.P., 1979. An evaluation of unclassed crossed-line choropleth mapping. The American Cartographer, 6: 21-37.

Stefanovic, P. & A. Vries-Baayens, 1984. Classification systems, choropleth maps and the computer. ITC-journal, 1: 52-57.

Tobler, E.R., 1973. Choropleth maps without class intervals. Geographical analysis, 3: 262-265.

Yeates, M.H., 1968. An introduction to quantitative analysis in economic geography. McGraw-Hill, New York.

PART III

11 Concluding remarks

The reason of existence of soil survey is its recognized ability to provide relevant soil information to users. As the questions change soil survey must also react accordingly to provide the right answers to the questions raised. The present changes in demand can be summarized in the following three points:
- new information;
- rapid production of information;
- quality of information.

New information

The traditional soil survey products in The Netherlands mainly focus on answering agricultural questions. Most soil classification systems are designed for this purpose. Although the agricultural questions remain relevant, new questions are raised as a result of environmental problems. For instance, questions about the behavior of pollutants in the soil and the behavior of soil attributes (e.g. organic matter, soil structure) in time in relation to sustainable land use. Soil survey data is also increasingly being used as input data for simulation models. These models generally require quantitative data and information on variability. To sum up, we can say that there is a demand for information about new attributes, attributes measured at a quantitative scale, variability of attributes and behavior of attributes in time.

For the derivation of new attributes many pedotransfer functions have been established (e.g. Breeuwsma et al., 1986; Wösten, 1990). Most attributes collected in soil surveys in the world are recorded on a qualitative measurement level (see Chapter 3). Commonly used statistics, such as mean, standard deviation and semivariance, and most spatial interpolation techniques are not permissible for this data type. In soil survey it must be realized that increasingly more questions raised require answers in quantitative terms, and also that for quantitative data, processing techniques are available or are being developed. Within a standard soil survey, more quantitative data should be collected as is done at this moment. Although it is useful to focus not only on which data should be available, but also

on which data are available. The development of useful processing techniques of qualitative data should be encouraged. Or to cite Hillel (1991):

> *"The utility of soil physics in the solution of many environmental problems is thwarted by the absence of a comprehensive and readily accessible data bank in which all sound measurements of soil properties and processes are recorded and correlated. To be useful, such data must be quantitative and fully specified. The acquisition of such data should be an integral activity of national and of international soil surveys that at present are still too qualitative and in some cases subjective. Rather than merely criticize such surveys, soil physicists should seek to participate in their work and contribute to their improvement."*

The variability of soil attributes within mapping units or small areas has been studied in many case studies (e.g. see Chapter 5; Marsman & De Gruijter, 1984), but has not yet yielded an approach which is incorporated in standard soil surveys. The most common model for describing soil variation is still the discrete spatial model without variability within the mapping units.

The behavior of soil attributes in time did not receive much attention in soil survey. Up to now most recorded soil attributes are considered to be constant in time. We know that this is not true when longer time intervals are involved. In The Netherlands we have, for instance, a decrease in the thickness of peat layers and changes in density and organic matter contents of the soil. The dimension of time can no longer be ignored in soil survey. We must consider soil survey more as a "monitoring" activity, with monitoring intervals of 10-20 years. As present technology (soil information systems) makes this new attitude possible, it is time to change our mental concepts about soil survey.

Rapid production of information

When soil survey data are needed for policy support studies the time available for answering questions is often limited. For instance, there is no time to spend three years on data collection to get an answer. We must be prepared to rapidly analyze effects of different scenarios. In order to produce information fast, computer storage, analysis and presentation is necessary. In most countries soil- and geographical information systems have been implemented, which produce

information rapidly. The implementation of automated analysis techniques is less far developed than the storage of data alone.

Quality of information

Product specifications and usually provided with industrial products such as chips, electronic components and cars. These specifications also include an indication of the quality of the product. The user decides which product suits his requirements, based on the quality and the price. A similar tendency can be found in the use of soil data. Providing soil information alone is not enough; the user also wants information about the quality of the information. Based on the specifications, he can decide if the information provided meets his needs. It might be possible that less, or more information is needed. To satisfy the demands of the users we must provide them with the necessary parameters by which they can make their judgements.

For a long time the only statement about quality of soil information has been the purity of the soil map, which was assumed to be constant for all units. In a standard soil survey no check on the validity of the assumed purity was done. It is obvious that this is not very reliable as a measure of quality. Providing information about the means and variances of attributes in mapping units as demonstrated by Visschers (1992) are essential parameters in the future. Furthermore, the quality of data should be made visible in our presentations (see e.g. Chapter 8 and 9).

We are facing an interesting period in soil survey. A period in which, as a result of the changing demand, traditional procedures are no longer able to answer all questions raised. It is a challenge for soil survey to develop and adopt new strategies in the years to come.

REFERENCES

Breeuwsma, A., J.H.M. Wösten, J.J. Vleeshouwer, A.M. van Slobbe & J. Bouma, 1986. Derivation of land qualities to assess environmental problems from soil surveys. Soil Science Society American Journal 50: 186-190.

Hillel, D., 1991. Research in soil physics: a re-view. Soil Science 151: 30-34.

Marsman, B.A. & J.J. de Gruijter, 1984. Dutch soil survey goes into quality control. In: P.A. Burrough and S.W. Bie (Eds.), Soil information systems technology. Pudoc, Wageningen, p. 127-134.

Visschers, R., 1992. Upgrading van de bodemkaart van Nederland door steekproeven in kaarteenheden. Rapport 186, DLO The Winand Staring Centre, Wageningen.

Wösten, J.H.M., 1990. Use of soil survey data to improve simulation of water movement in soils. PhD Thesis, Agricultural University, Wageningen.

12 Samenvatting

Verwerking van gekarteerde bodemgegevens

De bodem speelt een belangrijke rol in de landbouw, de landgebruiksplanning en de milieu- en natuurbescherming. Al deze activiteiten vragen informatie over de bodemsamenstelling en over bodemprocessen. Er is behoefte aan een grote verscheidenheid aan informatie, die tot gevolg heeft dat de bodem vanuit diverse invalshoeken bestudeerd wordt, onder andere de bodemchemie, de bodemfysica, de bodembiologie en de bodemkartering. Bodemkartering is de studie van de ruimtelijke verbreiding en de eigenschappen van bodems.

In het omzetten van de verzamelde basisgegevens naar voor een gebruiker relevante informatie kunnen de stappen verzamelen, opslaan, analyseren en presenteren worden onderscheiden (bodemkundige gegevensverwerking). Hoe het proces van gegevens verzamelen tot en met het presenteren van informatie verloopt hangt af van de vraag. De vraag bepaalt uiteindelijk welke gegevens er verzameld en hoe ze verwerkt moeten worden. Aan de vraagkant naar bodemkundige informatie zijn drie ontwikkelingen te noemen:

Ten eerste is er vraag naar nieuwe informatie. Dit is een gevolg van de sterk toegenomen aandacht voor milieu en duurzaam landgebruik. In bodemkundige termen vertaald betekent dit een vraag naar andere bodemeigenschappen en het gedrag hiervan in de tijd.

Ten tweede is een snelle levering van informatie ook gewenst. Bepaalde beleidsvragen staan niet toe dat er bijvoorbeeld eerst twee jaar onderzoek nodig is om een antwoord te kunnen geven. Het gebruik van informatiesystemen en geautomatiseerde analysemogelijkheden speelt bij het snel beantwoorden van een vraag een belangrijke rol.

Ten derde neemt de vraag naar de kwaliteit van de geleverde informatie toe. De bovengenoemde ontwikkelingen aan de vraagkant zijn niet van de ene op de andere dag ontstaan, maar voltrekken zich langzaam. Dit proefschrift levert een bijdrage aan procedures om met bodemkartering beter op de veranderende vraag van gebruikers te kunnen inspelen.

Hoofdstuk 2 gaat dieper in op de verschillende stappen die nodig zijn om bodemkundige basisgegevens om te zetten in informatie die voor een gebruiker relevant is. Allereerst ga ik in op modellen voor het beschrijven van de bodem. Hierbij is onderscheid gemaakt tussen ruimtelijke modellen, modellen in de tijd en ruimte en tijdmodellen. In dit proefschrift ligt het accent op het beschrijven van de bodem met ruimtelijke modellen.

De ruimtelijke modellen zijn ingedeeld in discrete en continue modellen. Een beschrijving van de bodem in de vorm van een bodemkaart is een voorbeeld van een discreet ruimtelijk model, geostatistische technieken daarentegen leveren een continue beschrijving van de bodem op. Om de bodem te beschrijven in de vorm van het discrete ruimtelijk model wordt de physiografische en de vrije kartering gebruikt. Grenzen tussen bodemeenheden zijn tijdens de kartering bepaald. Het continue model heeft waarnemingen uitgevoerd op puntlokaties nodig. De keuze voor de ligging van de puntlokaties kan zowel select als aselect zijn.

De laatste 20 jaar is opslag van verzamelde gegevens in informatiesystemen sterk in opkomst om snel analyses van de gegevens te kunnen uitvoeren. We kunnen acht categorieën van mogelijke analyses onderscheiden. Na een analyse volgt de presentatie van de resultaten. Dit kan zowel in tekst als in beeld. De meest gebruikelijke beeldvorm is de kaart: chorochromatische (= vlakkenkaart voor het weergeven van kwalitatieve verschillen) en choroplethkaarten (= vlakkenkaart voor het weergeven van kwantitatieve verschillen) voor discrete ruimtelijke gegevens en isolijnkaarten (= kaart met lijnen die punten met gelijke waarden van een continu veranderend verschijnsel met elkaar verbindt) voor continue ruimtelijke gegevens. De laatste tijd zien we steeds vaker animaties die veranderingen van bodemgegevens in ruimte en tijd presenteren.

In het tweede gedeelte van dit proefschrift, de hoofdstukken 3 t/m 10, worden onderdelen van het proces van bodemkundige gegevensverwerking beschreven.

Hoofdstuk 3 gaat in op het effect van de waarnemingsdichtheid op de nauwkeurigheid van ruimtelijke voorspellingen. In een studiegebied in Indonesië zijn voor verscheidene waarnemingsdichtheden voorspellingen uitgevoerd naar testpunten met vier verschillende voorspellingstechnieken (kriging, inverse afstand, lokaal gemiddelde en gebiedsgemiddelde). Uit de berekende fouten blijkt er geen verschil te zijn in nauwkeurigheid tussen de kriging, inverse afstand en het lokale gemiddelde als voorspeltechniek. Het gebiedsgemiddelde als ruimtelijke voorspeller

presteert duidelijk minder. De nauwkeurigheid van de voorspellingen en de kosten per waarnemingsdichtheid zijn gebruikt om een optimale waarnemingsdichtheid vast te stellen.

Hoofdstuk 4 beschrijft de aard van bodemgegevens verzameld in het veld. De opname van de meeste bodemvariabelen blijkt op een ordinale meetschaal te gebeuren. Bij ordinale variabelen is het niet mogelijk beschrijvende statistische grootheden zoals gemiddelde, standaard afwijking en semivariantie en de meeste ruimtelijke voorspeltechnieken te gebruiken. Aan de hand van gegevens uit een kartering in Costa Rica zijn verwerkingsmogelijkheden van ordinale gegevens behandeld. De voorgestelde ruimtelijke verschilkansfunctie is gebruikt om de ruimtelijke variabiliteit van ordinale variabelen te karakteriseren.

Hoofdstuk 5 handelt over de bruikbaarheid van bodemkundige gegevens verzameld op verschillende schalen voor het voorspellen van vochttekorten en opbrengstveranderingen. In een studiegebied van 1435 ha zijn de grondwaterstanden verlaagd als gevolg van wateronttrekking voor drinkwaterwinning. Om het effect van deze onttrekking op de grasproduktie te berekenen, is de huidige toestand (na onttrekking) vergeleken met de voormalige toestand (voor onttrekking). Als basisgegevens hiervoor zijn een bodemkaart op schaal 1 : 10 000, 1 : 25 000 en 1 : 50 000 gebruikt. De kaarteenheden van de verschillende bodemkaarten zijn bodemfysisch geïnterpreteerd. Met een simulatiemodel zijn vochttekorten en opbrengstveranderingen berekend. Om de kwaliteit van de drie bodemkaarten voor deze toepassingen te kunnen vaststellen, zijn er ook simulatieberekeningen voor een groot aantal puntlokaties uitgevoerd. De drie bodemkaarten verschillen niet in kwaliteit als het gaat om het vaststellen van gebiedsgemiddelden voor het vochttekort en verandering in opbrengst. De bodemkaart schaal 1 : 10 000 levert de beste voorspellingen van puntlokaties op.

In hoofdstuk 6 wordt een vergelijking gemaakt tussen een kaart afgeleid van een bodemkaart en een kaart vervaardigd via ruimtelijke voorspelling vanuit puntgegevens. De gebruikte ruimtelijke voorspellingsmethode is kriging. De geproduceerde kaarten beschrijven bodemvariabelen die voor het transport van water in de bodem van belang zijn. De zuiverheid van beide kaarten is bepaald met 60 onafhankelijke testboringen. Tussen beide kaarten blijkt geen verschil in zuiverheid te bestaan.

Hoofdstuk 7 gaat in op de fout die optreedt bij het omzetten van een vectorkaart in rastervorm. Elf vectorkaartbladen van de 'Bodemkaart van Nederland

schaal 1 : 50 000' zijn omgezet in rastervorm voor rastercel grootten van 1 mm x 1 mm, 2 mm x 2 mm en 4 mm x 4 mm en twee verschillende verrastertechnieken (centraal punt en dominante eenheid). De verrasterfout is bepaald met de Switzermethode, Goodchildmethode en de zelf ontwikkelde dubbele-conversiemethode. Uit het onderzoek bleek dat er een sterk verband bestaat tussen de complexiteit van de kaart en de verrasterfout. Voor dit verband zijn regressievergelijkingen opgesteld. Deze regressievergelijkingen kunnen gebruikt worden voor het voorspellen van de verrasterfout van een nieuw te verrasteren kaart voordat tot verrasteren wordt overgegaan. Uit het onderzoek bleek ook dat de dubbele-conversiemethode de verrasterfout weer beter voorspelde dan de Goodchild- en Switzermethode. De Goodchildmethode voorspelde de verrasterfout beter dan de Switzermethode. De gehanteerde verrastertechniek bleek slechts een gering effect op de optredende fout te hebben.

Hoofdstuk 8 beschrijft een procedure om isolijnkaarten met betrouwbaarheidsintervallen te vervaardigen. De procedure is gebaseerd op het gebruik van de ruimtelijke voorspellingsmethode kriging. Kriging levert naast een voorspelling van de waarde op een punt ook informatie over de nauwkeurigheid van de voorspelling in de vorm van de kriging-variantie. Kriging-variantie en voorspelde waarde zijn gebruikt om per punt de 90% betrouwbaarheidsboven- en -ondergrens te berekenen. Deze berekeningen zijn uitgevoerd voor punten op regelmatig raster. Vervolgens zijn voor bepaalde waarden de betrouwbaarheidsboven- en betrouwbaarheidsondergrens met lijnen aan elkaar verbonden. Deze lijnen gecombineerd met de isolijnkaart met voorspelde waarden leverde de isolijnkaart met betrouwbaarheidsintervallen. De kaart geeft de gebruiker een veel beter inzicht in de onzekerheid van de gepresenteerde gegevens. Ook de juistheid van de berekende kriging-varianties is nagegaan. Het blijkt dat deze de werkelijkheid aanzienlijk onderschatte. In de procedure voor het vervaardigen van de isolijnkaart zijn hiervoor correcties uitgevoerd.

Ook hoofdstuk 9 gaat in op het presenteren van onzekerheid. Twee procedures worden behandeld voor het vervaardigen van kaarten die de kans op overschrijding van een bepaalde grenswaarde aangeven. De procedures zijn vergeleken aan de hand van gegevens uit een bodemkartering in Indonesië. De geproduceerde kaarten geven de kans aan dat de pyriethoudende laag dieper zit dan 50 cm beneden het bodemoppervlak. Deze presentatievorm stelt de gebruiker in staat om bijvoorbeeld zoals in Indonesië bij het inrichten van een gebied voor rijstteelt rekening te hou-

den met onzekerheid in de gegevens.

Hoofdstuk 10 ten slotte gaat in op de complexiteit van kaarten. De complexiteit van kaarten blijkt een grote invloed te hebben op de leesbaarheid. Voor het evalueren van de complexiteit bestaan complexiteitsmaten. Zes van deze maten zijn met elkaar vergeleken. De fragmentatie-index is het meest geschikt voor het evalueren van de kaartcomplexiteit.

13 Curriculum Vitae

Arnoldus Koenraad Bregt werd op 24 juli 1959 geboren in Varik. In 1977 behaalde hij zijn Atheneum diploma aan de Rijksscholengemeenschap te Tiel. In datzelfde jaar begon hij met de studie bodemkunde en bemestingsleer aan de Landbouwhogeschool te Wageningen. De praktijktijd werd doorgebracht in Roemenië, waar hij bij het 'Research Institute for Soil Science and Agrochemistry (ICPA)' bodemkundig onderzoek verrichtte in de Donaudelta. In 1983 studeerde hij met lof af als landbouwkundig ingenieur met als hoofdvak regionale bodemkunde en met als bijvakken geomorfologie, bodemnatuurkunde en informatica.

In 1983 trad hij in dienst van de Stichting voor Bodemkartering (STIBOKA) als hoofd van de afdeling Toegepaste Informatica en Statistiek. In deze functie is hij o.a. intensief betrokken geweest hij het opzetten van het bodemkundig informatiesysteem voor Nederland (BIS) en de introductie van geografische informatiesystemen (GIS) binnen de Dienst Landbouwkundig Onderzoek (DLO). Bij het opgaan in 1989 van de STIBOKA in het DLO-Staring Centrum (SC-DLO) werd hij hoofd van de afdeling Kwantitatieve Methoden, GIS en Informatica. Het in de STIBOKA- en SC-DLO-periode uitgevoerde onderzoek naar verwerking van bodemkundige gegevens vormt het onderwerp van dit proefschrift.

In de periode 1988-1989 was hij secretaris van het organisatiecomité voor het internationale symposium 'Land qualities in space and time'. Hij fungeerde tevens als editor van de proceedings van dit symposium. In 1989 trad hij ook op als editor van de proceedings van een EG-workshop in Wageningen over 'Application of computerized EC soil map and climate data'.

In 1983 trouwde hij met Nelleke Schuurmans. Zij hebben twee kinderen: Annemoon en Tijmen.