

Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions

Alex. B. McBratney ^{*}, Inakwu O.A. Odeh

Cooperative Research Centre for Sustainable Cotton Production, Department of Agricultural Chemistry and Soil Science, The University of Sydney, Sydney, NSW 2006, Australia

Received 25 April 1996; accepted 27 February 1997

Abstract

Fuzzy systems, including fuzzy set theory and fuzzy logic, provide a rich and meaningful improvement, or extension of conventional logic. The mathematics generated by this theory is consistent, and fuzzy set theory may be seen as a generalisation of classic set theory. Applications in soil science, which may be generated from, or adapted to fuzzy set theory and fuzzy logic, are wide-ranging: numerical classification of soil and mapping, land evaluation, modelling and simulation of soil physical processes, fuzzy soil geostatistics, soil quality indices and fuzzy measures of imprecisely defined soil phenomena. Many other soil concepts or systems may be modelled, simulated, and even replicated with the help of fuzzy systems, not the least of which is human reasoning itself.

Keywords: fuzzy set theory; fuzzy numbers; fuzzy logic; geostatistics; soil science; geographical information systems

“As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality.” – ALBERT EINSTEIN (as quoted by Black, 1937)

“What makes the society turn is science, and the language of science is math, and the structure of math is logic, and the bedrock of logic is Aristotle, and that’s what goes out with fuzzy.” – BART KOSKO (as quoted by McNeill and Freiberger, 1993)

^{*} Corresponding author. Fax: +61 2 9351-3706. E-mail: alex.mcbratney@cropsci.su.edu.au

1. Introduction

1.1. Historic fuzziness

Fuzzy systems are an alternative to classical notions of set membership and logic that have their origins in ancient Greek philosophy (Brule, 1985). The concept of exact mathematical models owes its success in large part to the efforts of Aristotle and the philosophers who preceded him. These philosophers devised a concise theory of logic, or the so-called *Laws of Thought* (Korner, 1967). Two of the laws are the *Law of the Excluded Middle* which states that every proposition must either be TRUE or FALSE, and the *Law of Extended Contradiction* which states that a proposition cannot be both TRUE and NOT TRUE. Even when Parmenides proposed the first version of these laws (around 400 BC) there were strong and immediate objections: for example, Heraclitus proposed that things could be simultaneously TRUE and NOT TRUE.

The great philosopher Plato laid the initial foundation for what has now become fuzzy logic, indicating that there was a grey area (between TRUE and FALSE) where the antithesis ‘tumbled about.’ The later philosophers echoed these sentiments, notably Hegel, Marx, and Engels, but it was Lukasiewicz who first proposed a systematic alternative to the bi-valued logic of Aristotle (Lejewski, 1967). In the early part of this century, Lukasiewicz described a three-valued logic, along with the mathematics to go with it. The third value he proposed could be ascribed to the term ‘possible’, which he assigned a numeric value between TRUE and FALSE. Eventually, an entire notation and axiomatic system from which he hoped to derive modern mathematics was developed. Later, Lukasiewicz explored four-valued logics, five-valued logics, and then declared that in principle there was nothing to prevent the derivation of an infinite-valued logic. He felt that three- and poly-valued logics were the most intriguing, but ultimately settled on a four-valued logic because it seemed to be the most easily adaptable to Aristotelian logic.

Knuth proposed a three-valued logic similar to that of Lukasiewicz, from which he speculated that mathematics would become even more elegant than in traditional bi-valued logic (Brule, 1985). His foresight, apparently missed by Lukasiewicz, was to use the integral range $[-1, 0 + 1]$ rather than $[0, 1, 2]$. Nonetheless, this alternative failed to gain acceptance, and passed into relative obscurity.

Black (1937) outlined his prototype fuzzy sets which implied that vagueness stems from a continuum. The continuum needs not be continuous, he said, but can be discrete ‘like a dotted line’. He observed that a continuum can be approximated by a discrete dotted line, with each dot allocated a number. The number, he added, can be deduced by the percentage of people who would allocate a given number or usage to an element or object in the continuum. Vagueness, to him, was thus a matter of probability, a chance of allocating, rather than degree of truth. For this he mis-stepped, as he opted for people’s perception of usage of an object instead of the degree of truth or membership grade.

It was not until relatively recently that the notion of an multi-valued logic took hold. That was in the mid-sixties when Zadeh (1965) published his seminal work *Fuzzy Sets* which described the mathematics of fuzzy set theory, and by extension, fuzzy logic. This

theory adduced that by making the membership function (or the values TRUE and FALSE) operate over the range of real numbers $[0, 1]$, new procedures for the calculus of logic could be developed. The theory proved to be a generalisation of classic logic. It is the application of this theory in soil science that is the main topic of discussion in this and subsequent papers in this special issue.

1.2. *Fuzziness in soil science*

Many models in soil studies are interdisciplinary, requiring mathematical models that are built in the hard sciences and which are then linked with connections and subjective rule-based models used in the less exact or soft sciences. The resulting complex models are often difficult to interpret and may not reflect the soil or soil processes of the real world. For example, land evaluation is a decision-making procedure that relies on hard sciences of chemistry and physics but still requires knowledge of social and institutional factors so as to be able to evaluate the consequences of decisions (Waterstone, 1994). Thus models in soil science are characterised by: multiple, usually conflicting attributes; subjective uncertain conception on preferences (of the modeller); and uncertain, imprecise information on data on the models.

The problem of dealing with imprecision and uncertainty is a part of the wider human experience. However, until Zadeh (1965) introduced the theory of *fuzzy sets* there was no mathematically meaningful method of quantifying such imprecision and uncertainty. The theory of fuzzy sets or subsets, as some may prefer, is a generalisation of abstract set theory. The fuzziness stems from imprecision and uncertainty: its chief characteristic is the grouping of individuals into classes that do not have sharply defined boundaries. Fuzzy sets are useful whenever we have to describe ambiguity, vagueness, and ambivalence in conceptual or mathematical models of empirical phenomena (Kandel, 1986). Such conceptual models without fuzzification do not realistically describe physical phenomena as the models are based on a strictly binary approach which implies that an entity either belongs to a class or not at all.

The classical usage of statistics in soil science is based on deterministic analyses involving the imposition of some specific field designs and treatments to minimise the effect of uncertainty (Fisher, 1954). No doubt this approach has been quite successful. However, uncertainty is considered a removable artefact which should disappear with increasing knowledge (Bardossy and Duckstein, 1995). In reality uncertainty, imprecision and ambiguity are inevitable or inherent parts of natural systems. In some cases, it is implicitly assumed that uncertainty is always a result of randomness, and probability theory and statistics are adequate to deal with such uncertainty. Extension of such applications for complex models are often fraught with problems and sometimes difficult, especially when the complexity requires many theoretical assumptions. More often than not the complexity of the models stems from overemphasised precision which does not always mean greater truth. Fuzzy set theory provides the means of dealing with such uncertainty especially due to imprecise boundaries between categories. Even in many cases where the model is precise, fuzziness may be a concomitant of the complexity (Kandel, 1986).

In this paper we review the theory of fuzzy sets and its applications in soil science, focussing on the major topics to be dealt with in greater detail in subsequent papers of

this issue. Specifically, we give a brief introduction to the theory, fuzzy membership functions (MFs) and, fuzzy numbers and operations. Current applications in soil classification and mapping, land evaluation, fuzzy geostatistics and soil physical processes are explored as they are elsewhere in this special issue. A brief sketch of the future applications is given focussing on potential applications in decision-making, fuzzy control and field soil description, with some examples.

2. Theory

2.1. Fuzzy sets

In a formal definition of a fuzzy set, we presuppose that $\mathbf{X} = \{x\}$ is a finite set (or space) of points, which could be elements, objects or properties; a fuzzy subset, \mathbf{A} of \mathbf{X} , is defined by a function, $\mu_{\mathbf{A}}$, in the ordered pairs:

$$\mathbf{A} = \{x, \mu_{\mathbf{A}}(x)\} \quad \text{for each } x \in \mathbf{X} \quad (1)$$

In plain language, a fuzzy subset is defined by the membership function defining the membership grades of fuzzy objects in the ordered pairs consisting of the objects and their membership grades. The relation $\mu_{\mathbf{A}}(x)$, is therefore termed as a membership function (MF) defining the grade of membership x (the object) in \mathbf{A} and $x \in \mathbf{X}$ indicates that x is an object of, or is contained in \mathbf{X} . For all \mathbf{A} , $\mu_{\mathbf{A}}(x)$ takes on the values between and including 0 and 1. In practice, $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ and Eq. (1) is written as:

$$\mathbf{A} = x_1, \mu_{\mathbf{A}}(x_1) + x_2, \mu_{\mathbf{A}}(x_2) + \dots + x_n, \mu_{\mathbf{A}}(x_n),$$

the + is used as defined in the set theoretic sense. If $\mu_{\mathbf{A}}(x) = 0$, then $x, \mu_{\mathbf{A}}(x)$ is omitted.

2.2. Membership function and fuzzy numbers

A fuzzy membership function (FMF) is thus an expression defining the grade of membership of x in \mathbf{A} . In other words, it is a function that maps the fuzzy subset \mathbf{A} to a membership value between and including 0 and 1. In contrast to the characteristic function in conventional set theory which implies that membership of individual objects in a subset as either belonging or not at all, i.e. $\mu_{\mathbf{A}_{nf}}(x) \rightarrow \{0,1\}$ where \mathbf{A}_{nf} is the non-fuzzy equivalent of fuzzy subset \mathbf{A} , the FMF of x in \mathbf{A} is expressed as:

$$\mu_{\mathbf{A}}(x) \rightarrow [0,1]$$

that associates with each element $x \in \mathbf{X}$ its grade of membership $\mu_{\mathbf{A}}(x) \in [0,1]$. Thus $\mu_{\mathbf{A}}(x) = 0$ means that x does not belong to the subset \mathbf{A} , $\mu_{\mathbf{A}}(x) = 1$ indicates that x fully belongs, and $0 < \mu_{\mathbf{A}}(x) < 1$ means that x belongs to some degree; partial membership is therefore possible.

A fuzzy number is a fuzzy subset of real numbers characterised by a possibility distribution (Kandel, 1986). It is a generalisation of the real numbers. Usually, a fuzzy

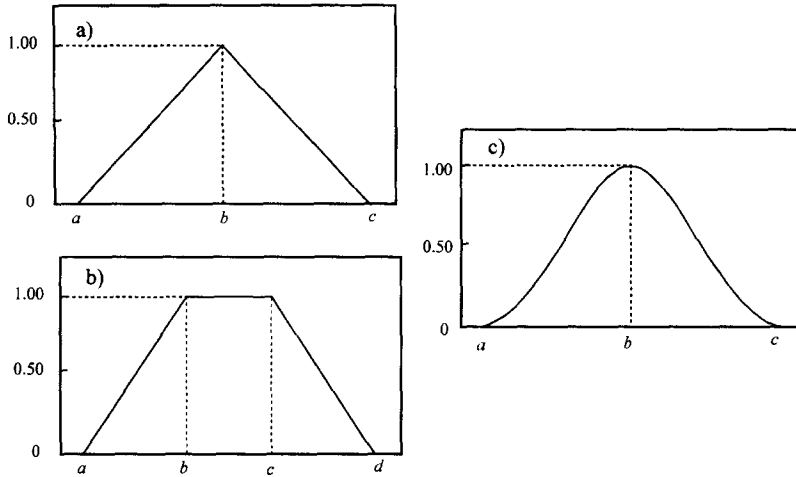


Fig. 1. Examples of fuzzy numbers: (a) triangular, (b) trapezoidal and (c) Gaussian.

number is defined as a normal and convex fuzzy set in the real line $A \subseteq \mathfrak{R}$. By being normal we mean that the maximum value of membership in a fuzzy set in \mathfrak{R} is 1. Convex implies that the fuzzy number consists of an increasing and decreasing part, and sometimes a flat part. Functionally, this can be expressed for each of the real numbers a , b , and c as:

$$\mu_A(b) \geq \min(\mu_A(a), \mu_A(c)), \quad a < b < c \quad (2)$$

The simplest fuzzy number is the so-called triangular fuzzy number (Bardossy and Duckstein, 1995) with its characteristic MF written as:

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x \leq b \\ \frac{c-x}{c-b} & b < x \leq c \\ 0 & c < x \end{cases}$$

Fig. 1(a) illustrates the MF of triangular fuzzy number. Other fuzzy numbers are trapezoidal or Gaussian (Fig. 1b and c).

The fuzzy numbers play an equivalent role in fuzzy models to the classical (real) numbers in non-fuzzy models, hence their universal importance in fuzzy operations.

2.3. Extension principle

The extension principle (Zadeh, 1965) is a general method for extending non-fuzzy mathematical concepts for the purpose of deriving fuzzy quantities (Dubois and Prade, 1980). In other words, it provides the means of extending point-to-point operations to fuzzy sets, a development that makes arithmetic operations on fuzzy numbers possible.

In defining the concept, let us suppose that \mathbf{X} and \mathbf{Y} are two sets, and f the point-to-point mapping from \mathbf{X} to \mathbf{Y} :

$$f: \mathbf{X} \rightarrow \mathbf{Y} \quad \text{for every } x \in \mathbf{X} \quad f(x) = y \in \mathbf{Y} \tag{3}$$

We can then extend f to operate on fuzzy subsets of \mathbf{X} . If \mathbf{A} is a fuzzy subset of \mathbf{X} with membership function $\mu_{\mathbf{A}}$, then the image of \mathbf{A} in \mathbf{Y} is the subset \mathbf{B} with the membership function expressed as (Bardossy and Duckstein, 1995):

$$\mu_{\mathbf{B}}(y) = \begin{cases} \sup\{\mu_{\mathbf{A}}(x); y = f(x), x \in \mathbf{X}\} \\ 0 \text{ if there is no } x \in \mathbf{X} \text{ such that } f(x) = y \end{cases} \tag{4}$$

The extension principle therefore allows for the definition of arithmetic operations on fuzzy numbers in such a way that it preserves the image of the elements of \mathbf{X} . For example, if \mathbf{A} and \mathbf{B} are two fuzzy numbers, their sum is $\mathbf{C} = \mathbf{A} + \mathbf{B}$, where \mathbf{C} is a fuzzy number (fuzzy subset of the real line) with a MF defined as:

$$\mu(y) = \sup\{\max(\mu_{\mathbf{A}}(x_1), \mu_{\mathbf{B}}(x_2)); \text{ such that } y = x_1 + x_2\} \tag{5}$$

When combined with the Cartesian product, the extension principle provides the most useful operations on fuzzy numbers (Bardossy and Duckstein, 1995).

2.4. Basic operations on fuzzy sets

The basic operations on fuzzy subsets are similar, and are a generalisation of classical sets. Readers are referred to a number of authors who have presented details of these operations, e.g., Kaufmann (1975), Kandel (1986) and Burrough (1989). In this paper, we present only a brief summary of each operation as presented in Kaufmann (1975). In defining each of the operations, we assume two fuzzy sets, \mathbf{A} and \mathbf{B} , each belonging to finite sets \mathbf{X} , of real numbers \Re .

Inclusion. \mathbf{A} is said to be included in \mathbf{B} if

$$\mu_{\mathbf{A}}(x) \leq \mu_{\mathbf{B}}(x) \quad x \in \mathbf{X} \tag{6}$$

This can be denoted as:

$$\mathbf{A} \subset \mathbf{B}$$

which obviously is a case of inclusion in the sense of the theory of fuzzy subsets (Kaufmann, 1975).

Equality. \mathbf{A} and \mathbf{B} are equal if and only if

$$\mu_{\mathbf{A}}(x) = \mu_{\mathbf{B}}(x) \quad x \in \mathbf{X},$$

and this can be denoted by:

$$\mathbf{A} = \mathbf{B}.$$

Where at least one $x \in \mathbf{X}$ does not satisfy the equality, $\mu_{\mathbf{A}}(x) = \mu_{\mathbf{B}}(x)$, then \mathbf{A} and \mathbf{B} are not equal as denoted:

$$\mathbf{A} \neq \mathbf{B}$$

Complementation. Fuzzy sets \mathbf{A} and \mathbf{B} are complementary if

$$\mu_{\mathbf{B}}(x) = 1 - \mu_{\mathbf{A}}(x), \quad x \in \mathbf{X} \tag{7}$$

This can be denoted as:

$$\mathbf{B} = \overline{\mathbf{A}} \text{ or } \overline{\overline{\mathbf{A}}} = \mathbf{B}$$

It is obvious that

$$\overline{\overline{\mathbf{A}}} = \mathbf{A}$$

which means that the complement of $\overline{\mathbf{A}}$ is \mathbf{A} . The complement operator corresponds to the set theoretic operator, *NOT*.

Intersection. The intersect of \mathbf{A} and \mathbf{B} is defined as:

$$\mathbf{A} \cap \mathbf{B}$$

meaning that it is the largest fuzzy subset containing elements held in both \mathbf{A} and \mathbf{B} , i.e.:

$$\mu_{\mathbf{A} \cap \mathbf{B}}(x) = (\mu_{\mathbf{A}}(x) \wedge \mu_{\mathbf{B}}(x)) = \min(\mu_{\mathbf{A}}(x), \mu_{\mathbf{B}}(x)) \quad x \in \mathbf{X} \quad (8)$$

This is the *AND* (minimisation) operator.

Union. The union of \mathbf{A} and \mathbf{B} is the smallest fuzzy set that contains all elements held in both \mathbf{A} and \mathbf{B} , i.e.:

$$\mathbf{A} \cup \mathbf{B}$$

Thus the union of the two fuzzy sets is defined as:

$$\mu_{\mathbf{A} \cup \mathbf{B}}(x) = (\mu_{\mathbf{A}}(x) \vee \mu_{\mathbf{B}}(x)) = \max(\mu_{\mathbf{A}}(x), \mu_{\mathbf{B}}(x)) \quad x \in \mathbf{X} \quad (9)$$

This corresponds to fuzzy *OR* (maximisation) operator.

The product of two fuzzy sets \mathbf{A} and \mathbf{B} can be defined as:

$$\mathbf{AB} = \mu_{\mathbf{A}(-)\mathbf{B}} z = \vee(\mu_{\mathbf{A}}(x) \wedge \mu_{\mathbf{B}}(y)) \quad x, y, z \in \mathbf{X}$$

where $z = x - y$. This can be interpreted as the *SOFT AND* operator (Burrough, 1989).

Two other important fuzzy operators are *disjunctive sum* which is defined by:

$$\mathbf{A} \oplus \mathbf{B} = (\mathbf{A} \cap \overline{\mathbf{B}}) \cup (\overline{\mathbf{A}} \cap \mathbf{B}) = \max(\mu_{\mathbf{A} \cap \overline{\mathbf{B}}}, \mu_{\overline{\mathbf{A}} \cap \mathbf{B}}) \quad (10)$$

and the *difference* defined by the relation:

$$\mathbf{A} \ominus \mathbf{B} = \mathbf{A} \cap \overline{\mathbf{B}} = \mu_{\mathbf{A} \cap \overline{\mathbf{B}}} \quad (11)$$

The *disjunctive sum* is the connectives termed as *SOFT OR* (Burrough, 1989).

2.5. Fuzzy logic

Fuzzy logic can be considered as a generalisation of the Boolean logic whereby the latter has been extended to handle the notion of partial-truth: truth values between and including ‘completely true’ and ‘completely false’. In a formal definition, fuzzy logic is a structured model-free estimator that approximates functions through linguistic input/output schemes. Fuzzy logic therefore uses a ‘soft’ linguistic type of variables (e.g. deep, sandy, steep, etc.) which are defined by continuous range of truth values or FMFs in the interval [0,1] instead of the strict binary (TRUE or FALSE) decisions and assignments, as is the case with the Boolean logic. The linguistic input–output associations, when combined with an inference procedure, constitute the fuzzy rule-based systems.

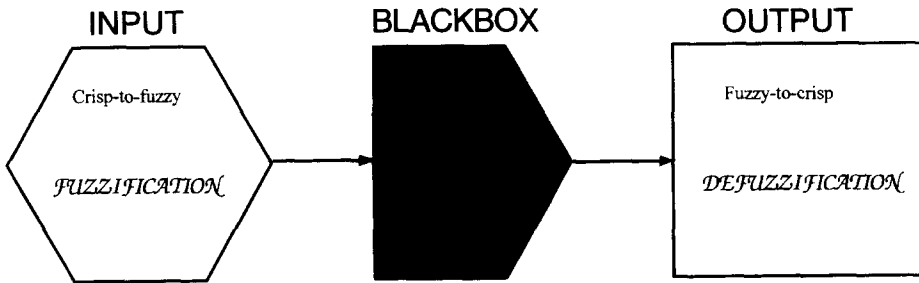


Fig. 2. An input–output map of a typical fuzzy rule-based system.

Fig. 2 shows the input–output map of a typical fuzzy rule-based system. It consists of three components: the input with the input variables fuzzified by the MFs, the rule base or the so-called black box that does the work of inference through the fuzzy operations discussed in Section 2.4, and the output component which may be defuzzified. Many examples of how the fuzzy logic approach has been used in soil studies abound as will be presented in the later sections.

2.6. Hypothetical examples in soil science

We provide examples in soil science to explicitly illustrate the fuzzy set theory. In soil science, we often hear of ‘very deep soil’, ‘deep soil’ or ‘shallow soil’. These combined constitute the universe of ‘soil depths’, taking values between 0 and 200 cm (Burrough, 1989). Soil depth could be regarded as a fuzzy property (Banai, 1993). We assume that **A** is a fuzzy subset of shallow soil, **B** a fuzzy subset of deep soil and **C** a fuzzy subset of very deep soil. For simplicity, we take the subset **B** (deep soil) as represented by soil depths of about 100 cm. The FMF, $\mu_B(x)$, is shown in Fig. 3. Thus 100 cm depth fully belongs to the set of deep soil, and therefore has a membership of 1, whereas depths of $0 \leq x < 100$ or $100 < x \leq 200$ cm have partial memberships. Fig. 3 also shows that the transition between full belongingness and non-belongingness is not clear-cut but gradual.

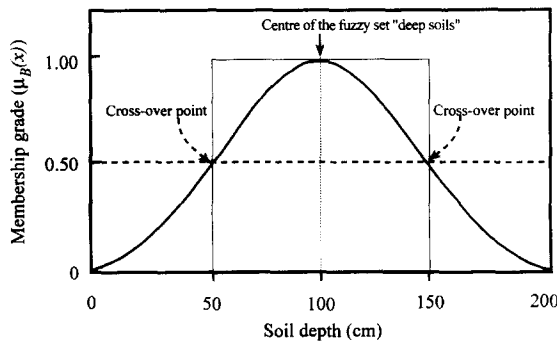


Fig. 3. A Gaussian (fuzzy) membership fitted to a subset of ‘deep soil’.

Assuming that the MF defining the subset of deep soil is an exponential or Gaussian type (Jang and Gulley, 1995) defined by:

$$\mu_B(x) = e^{-\left(\frac{(x-c)}{1.44\sigma}\right)^2} \quad 0 < x \leq 200 \quad (12)$$

where $c = 100$ cm, representing the centroid of the fuzzy subset of deep soil; σ is the lower cross-over point (50 cm). The cross-over point is the x value at which membership is 0.50 (Fig. 3). The higher cross-over point is at $x = 150$ cm depth. The x values 0, 100 and 200 cm depth represent the exponential fuzzy number defined by the MF. This function satisfies the normal and convex assumptions (Eq. (2)) in that maximum membership of $x = 100$ cm depth is 1, while the curve consists of an increasing part and a decreasing part (Bardossy and Duckstein, 1995). The support of the fuzzy number, defined as the set of real numbers of which MF is greater than zero, is $0 < x < 200$ cm. The membership values denote the ‘likeness’ of the occurrence of the real numbers (Kaufmann and Gupta, 1991) within the set interval or the support of the fuzzy number. A somewhat similar proposal is the suggestion by Zadeh (1978) that fuzzy membership is the possibility that the parameter having an imprecise value and denoted by a fuzzy number A would take the value x . It therefore means that $\mu_A(x) = 1$ indicates x that is totally possible, and $\mu_A(x) = 0$ is indicative of x that is impossible.

Fig. 4 shows MFs of more than one fuzzy subset, shown as subsets of shallow, deep and very deep soil. Choosing 0 cm, 100 cm and 200 cm as the standard indices or centroids of the shallow soil, deep soil and very deep soil, respectively. The shallow soil and very deep soil sets have a truncated or asymmetrical FMF. The deep soil FMF is symmetrical, representing a true, normal and convex FMF (also see Fig. 3).

We can also show how several fuzzy subsets can overlap. Fig. 4 clearly indicates the overlapping nature of the three fuzzy subsets within the universe of depths of the soil. It illustrates the fact that classes or sets in the real world do not necessarily have sharply defined boundaries, and that a continuum, such as the soil system that is characterised by multiple attributes, do not always have sets or classes with rigidly defined cut-offs. Individuals in a multivariate system with similar characteristics should have shared memberships to different subsets defining the system.

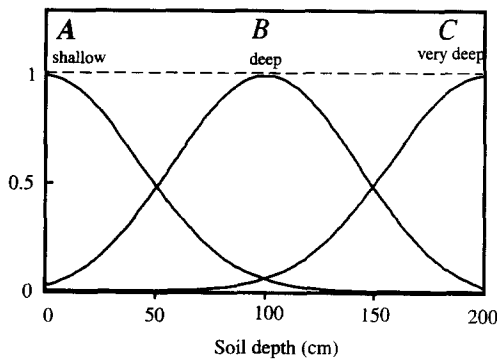


Fig. 4. Gaussian membership functions fitted to fuzzy subsets of ‘shallow soil’ (A) (asymmetrical), ‘deep soil’ (B) (symmetrical) and ‘very deep soil’ (C) (asymmetrical).

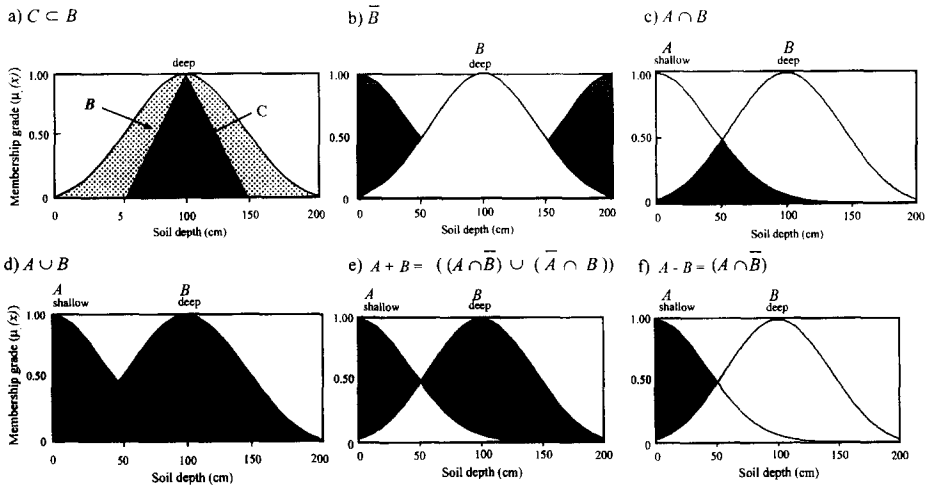


Fig. 5. Extended Venn diagrams illustrating fuzzy sets and fuzzy operations: (a) inclusion, (b) complement, (c) intersection, (d) union, (e) disjunctive sum and (f) difference.

Fig. 5 illustrates graphically the simple operations on fuzzy subsets as defined in Eqs. (6)–(11). In Fig. 5a, a fuzzy set of deep soil (**B**) is fitted by an exponential FMF and a triangular function fitted to the fuzzy subset of the soil types that are about 100 cm deep (**C**). Obviously, **C** is a subset within the set **B**. It means that **C** is less than **B**, and hence **C** belongs to **B** as defined in Eq. (6).

In Fig. 5b, the complement of the subset of deep soil (**B**) is considered. The complement operator is equivalent to not connective, and hence, ‘not a subset of deep soil’ (or not **C**) within the ‘universe of depths of soil’ is represented by the patterned area of Fig. 5b. The latter represents the combined subsets of ‘shallow’ and ‘very deep’ soil. In Fig. 5c, the intersection of the two fuzzy subsets of shallow soil and deep soil, is illustrated by the patterned area. It represents the minimisation of memberships of the two sets corresponding to the values of soil depths within the universe of discourse, i.e., where both subsets have common memberships.

The union of subsets of the shallow and deep soil is shown in Fig. 5d as the patterned area. It is a combination of elements in the two fuzzy subsets (**A** and **B**) and hence, the maximum fuzzy memberships at corresponding values of soil depths, be they within the shallow or deep subset or where both overlap. The two other fuzzy operations illustrated in Fig. 5 are the disjunctive sum and difference. The disjunctive sum is the union of $A \ominus B$ and $B \ominus A$. In other words, it is membership values, up to maximum, for the corresponding soil depths where both **A** and **B** do not intersect (Fig. 5e). As shown in Fig. 5f, the difference between the two fuzzy subsets, expressed as $A - B$, is the intersection of the fuzzy set **A** (shallow soil) and the complement of fuzzy subset **B** (\bar{B}).

3. Fuzzy (continuous) classification and its applications in soil science

In soil science, the fuzzy set theory is principally used for classification. The purpose of classification is to reduce a complex system, represented by some sets of data, into

explicitly defined classes. In using the fuzzy set theory, observations are grouped into *continuous classes* (McBratney et al., 1992; Burrough et al., 1992) in which individuals are assigned *continuous class membership values* instead of classifying the observations into exactly defined (hard) classes. A membership value of 1 is assigned to individuals that exactly matched strictly defined classes; individuals that do not match the strictly defined classes would receive membership values depending on their degree of closeness to the strictly defined classes or the class centroids (or class means). Obviously the manner in which individuals are assigned continuous membership values depends on the MF's used in assigning the membership values and the character space between the individuals and the centroids of the classes.

There are two different but complementary approaches to grouping individuals into fuzzy sets or classes. The first is based on *fuzzy c-means* (FCM, also known as *fuzzy k-means*) partitioning of observations in multivariate space into relatively stable naturally occurring groups. The number of classes are not known a priori. This approach, dealt with in Section 3.1, can be used for exploratory analysis in which case the results may lead to testable hypotheses about the nature of soil and landscape (Burrough et al., 1992).

The second approach is based on what is termed the *Semantic Import model* (SI), where a MF is defined without reference to the data. The class limits are specified based on experience or conventionally imposed definitions before individuals are allocated on the basis of how close they match the requirements of the classes. Very often, the MF's are simple (as indicated in Section 2) and can be seen as simple extensions of Boolean classes. It is for this reason that they were first applied in land suitability classification (as discussed in Section 3.2). When classes are not pre-specified it is reasonable to use the FCM procedure to suggest suitable classes which can then be used as SI classes (Van Gaans and Burrough, 1993).

In either of the two cases the principle of fuzzy logic can be applied to physical 'objects' such as soil profiles, soil polygons or land parcels. Whenever these objects are treated as individuals it is possible to apply the fuzzy logic rules in a soil/geographical information system or GIS (Section 3.3) because the fuzzy logic rules apply only to the attributes. On the other hand, fuzzy membership values for soil profiles or soil layers (point observations) can be interpolated to continuous (discretised) surfaces for further use in GIS if the classes in attribute space are spatially contiguous. This point is taken up by Burrough et al. (1997).

3.1. *Continuous soil classification and mapping using FCM*

Recognising the complexity of the soil system, multivariate techniques or the so-called numerical methods, have been applied to soil studies since the early 'sixties (McBratney, 1992). Most of the techniques attempted to optimise classification, given the available data (Odeh et al., 1992a). However, the resulting classes are discrete, just as with the 'natural' (hard) classification. Where the natural classification is applied, it usually means that the resulting classes are a reflection of the underlying processes. Most soil processes may not produce crisp classes but a continuum (Dale et al., 1989). This is concomitant with the observation by Fridland (1974) that "the soil cover may be

regarded as a discrete–continuous formation that is physically continuous but geographically discrete”, and “classificationally, it (the soil cover) is liable to be either continuous (with gradual transitions between soil (types), though closely related soil forms) or discrete (with sharp transitions between soil (types) and very dissimilar neighbouring soil (types))”. The complexity of soil variation at all scales, the resultant uncertainty in soil mapping and spatial interpolation (Webster, 1985) and our inability to sample and measure physical objects precisely as they exist in the real world (Burrough, 1987), require that a new system of soil classification be developed. Fuzzy set theory has provided the basis for the new and emerging system of classification.

The first well-defined application of the theory of fuzzy sets to data structure identification was made by Ruspini (1969). It was followed by the work of Dunn (1974) and Bezdek (1974) who developed the FCM algorithm that became a landmark in the theory of cluster analysis. McBratney et al. (1992) termed this technique *continuous classification*. Continuous classification is based on the theory of fuzzy sets without the notion of vagueness attached (McBratney, 1992). The continuousness is achieved by the use of fuzzy sets to group individuals into classes with no cut-offs. In order to improve the FCM algorithm for predictive classification De Gruijter and McBratney (1988) modified the FCM objective function by providing for membership to an extragrade class. The modified objective function, defining the within-class sum-of-square errors J_E , is expressed as:

$$J_E(M, c) = \alpha \sum_{i=1}^n \sum_{j=1}^c m_{ij}^{\phi} d_{ij}^2 + (1 - \alpha) \sum_{j=1}^c m_i \cdot \sum_{j=1}^c d_{ij}^{-2}$$

where c is the number of classes, n is the number of individuals or pedons; m_{ij} is the membership of an individual i in class j ; ϕ is the fuzziness exponent ($1 < \phi < \infty$); d_{ij} is the character space between the feature value of an individual, i , and the feature centroidal value for class j ; α is the parameter that determines the mean value of m_i , which is the membership value of an individual, i , in the extragrade class. Minimisation of the objective function is done by heuristic Picard iteration to derive the membership coefficients, the memberships to the extragrade class and the centroids of the soil properties used for classification. Readers are referred to De Gruijter and McBratney (1988), McBratney and De Gruijter (1992) and Odeh et al. (1992a) for detailed treatment of the FCM algorithm.

Since the work of Dunn (1974) and Bezdek (1974), there have been numerous applications of FCM in various other fields (Bezdek, 1981; McBratney and Moore, 1985). Examples in soil science include direct applications to soil classification by Odeh et al. (1990, 1992b), Powell et al. (1991), McBratney and De Gruijter (1992), McBratney et al. (1992) and Odeh et al. (1992a). The examples provided by Powell et al. (1991) and Odeh et al. (1992a) elucidated the soil–landscape inter-relationships as revealed by the resulting ‘natural’ classes. In the examples given by McBratney et al. (1992) and Odeh et al. (1992b), the geostatistical technique of kriging was combined with continuous classification to produce new kinds of maps. Fig. 6 shows an example of composite maps of continuous classes of soil textural profiles (Odeh, 1990), including the extragrade class, resulting from *union* or maximum operation on all the continuous classes generated by the FCM algorithm. The map shows the core areas of continuous

classes and the extragrade with the white areas between the cores as the intragrades, i.e. individuals that have less than 0.5 degree of membership to any of the continuous classes. More examples of continuous classification and mapping using the FCM algorithm are provided by Triantafilis and McBratney (1993), Burrough et al. (1997), De Gruijter et al. (1997), Hendricks Franssen et al. (1997) and Lagacherie et al. (1997).

Continuous (fuzzy) classes can be constructed based on the ‘central concepts’ of classes that are defined a priori, using the conventionally defined classes such as in Soil Taxonomy and other classification systems. In such cases, Picard iteration within the FCM algorithm is not needed to calculate the membership values. More recently, McBratney (1994) developed such an allocation program based on the FCM algorithm without the embedded Picard iteration. A further example of its application was published by Mazaheri et al. (1995) in which the program was applied to the Australian Great Soil Group (GSG) classification scheme. The results show an intuitively plausible allocation system. Further extensions of this approach are presented by Mazaheri et al., 1997. More work needs to be done, especially applications and modification to other classification schemes such as the Key to Soil Taxonomy (Soil Survey Staff, 1994) and the World Soil Reference Base (FAO, 1994).

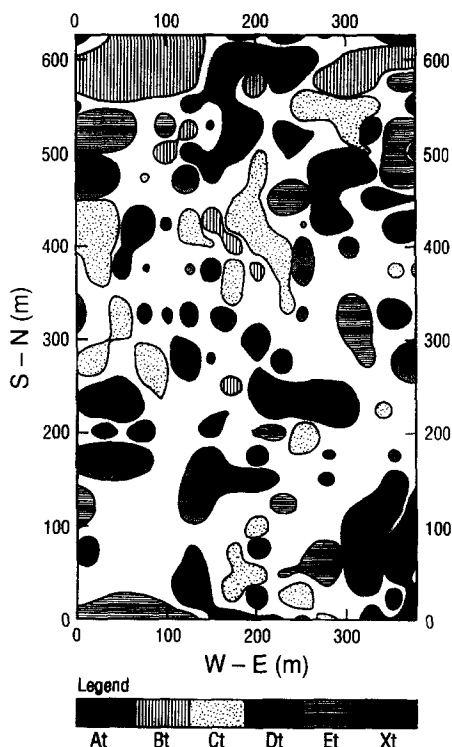


Fig. 6. Composite map showing the core areas of continuous classes of textural profiles: *At* = sandy loam over clay loam; *Bt* = sandy loam over sandy clay loam, most texturally uniform; *Ct* = loam over heavy clay; *Dt* = slightly gravelly sandy loam over heavy clay, relatively coarse-textured; *Et* = moderately gravelly sandy loam over sandy clay loam, also coarse-textured.

3.2. Applications based on semantic import model (SI)

The basic soil information used for land evaluation is mostly described by seemingly vague terms such as ‘poorly drained’, ‘slightly susceptible to soil erosion’, ‘moderate nutrient availability’ etc. (Burrough, 1989). Not even when these terms are defined precisely is the qualitative ambiguity removed. Usually, the land evaluator’s aim is to produce a set of clearly defined classes of land qualities based on specified land use requirements (FAO, 1976). These subsequently provide the means of transferring information about the soil and its use. As land qualities are complex attributes that are derived from land characteristics such as topography, soil, water, or biological and human activity, subsequent Boolean logical operations in the process of land evaluation tend to throw away much useful information (Burrough, 1989; McBratney, 1992). Fuzzy sets theory has been applied to land evaluation in order to deal with such ambiguity and vagueness, and to handle inexactness.

Chang and Burrough (1987) were the first to apply fuzzy sets and logic to land evaluation. Their approach, similar to that of Burrough (1989) and Burrough et al. (1992), was based on the SI in contrast to the FCM approach. The former involves selection of a-priori MFs with which individuals can be assigned membership values. It therefore requires some expert knowledge of data and the structure therein. This can be based on the predetermined FAO framework (FAO, 1976). For example, Burrough (1989), Burrough et al. (1992) and Banai (1993) presented some MF’s suitable for soil and landscape data. The simplest MF is of the form (Burrough, 1989; Burrough et al., 1992):

$$\mu_F(x) = \frac{1}{1 + \{d^{-1}(x - b)\}^2}$$

which depicts a general symmetrical bell-shape as the exponential function given in Eq. (12). The parameter b is the value of attributes x , be it a subset of deep soil as defined above, representing the centroid of the subset. The parameter d , gives the width of the bell-shaped curve at the cross-over points as defined above.

The other variants of FMF are asymmetric types, which define the end member subsets. Taking the subset of say, the ‘shallow soil’, as represented by the asymmetric left curve in Fig. 4, the MF is defined by the exponential function similar to Eq. (12):

$$\begin{aligned} \mu_A(x) &= e^{-\left(\frac{(x)^2}{1.44\sigma^2}\right)} \quad \text{for } x \leq b \\ \mu_A(x) &= 0 \quad \text{for } x > b \end{aligned} \tag{13}$$

where $b = 100$ cm is the maximum value of the subset at which the MF is applicable, σ is the cross-over point at which the membership value equals 0.5. Eq. (13), as applied to a shallow subset within the universe of ‘depths of the soil’, expresses the concept of shallowness rather than defining a soil as shallow by imposing a cut-off at, for example, 50 cm. This introduces us to the concept of *linguistic variable* which is built on top of

the fuzzy set theory. A linguistic variable is a variable representing natural language expressions that refer to some quantity of interest (Schmucker, 1984). The use of such linguistic qualifiers in soil science and land evaluation, such as *fairly*, *slight*, *very* and *extremely*, is an attempt to more precisely model an imprecise phenomenon. Fuzzy sets allow us to fulfil a true representation of imprecise systems as they are in the real world. McBratney (1992) provided some good examples. If we take the term *shallow* as defined by the MF illustrated in Fig. 7. To adequately describe the primary term *shallow*, we need to extend and fine-tune the perception of shallowness by giving an allowance to qualify a qualifier. This is the concept of *hedges*. The hedges are introduced to fuzzy sets by applying the power functions to the MF defining the primary linguistic variables. In Fig. 7, the *very shallow* and *fairly shallow* curves are respectively produced by squaring and taking the square root of the MF (Eq. (13)) defining the *shallow* subset. The *not very shallow* is a negation of *shallow* (also shown by the left curve (subset A) of Fig. 4).

Chang and Burrough (1987) applied fuzzy sets and fuzzy logic to determine suitability for apple growing in an area northeast of China. Burrough (1989) and Burrough et al. (1992) used fuzzy classification to determine land suitability for various purposes. All these studies involved the complex combination operator as defined in fuzzy systems. Heuvelink and Burrough (1993) reported that fuzzy methods were much better in producing suitability classification than Boolean methods. The latter methods were shown to have wrongly rejected a larger number of cells as suitable for research plots than fuzzy methods, the classification of which produced a more continuous variation such as maps showing results of drainage analysis.

A slightly different approach using fuzzy sets was recently developed by Triantafilis

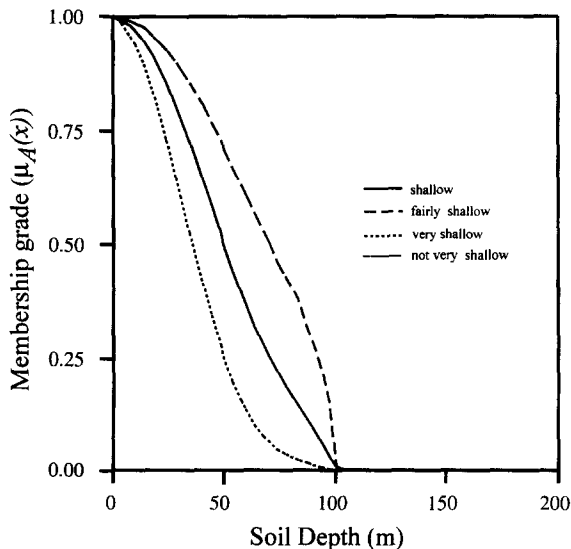


Fig. 7. Membership functions showing linguistic hedges: 'fairly', 'very' and 'not very' shallow soil.

and McBratney (1993) in which a fuzzy derivative of an accumulated suitability score (s) was determined using the expression:

$$\mu_F(s) = e^{-0.1s}$$

The results are presented on a series of maps which show better representation of continuous variation than with conventional methods. The results are also similar to those produced by Burrough et al. (1992). *Land versatility* (Triantafilis and McBratney, 1993), defined as the mean of the suitability (fuzzy) memberships for a range of crops, i.e. wheat, cotton, sunflower, etc., was also derived. It is a quantitative, but continuous substitute for the original notion of land capability. A map of block-kriged land versatility for an area in the northwest of NSW was also produced. This is also an example of complex combination operator (Burrough, 1989). Other important land evaluation work using fuzzy sets application to the FAO framework (FAO, 1976) has been published by Tang et al. (1991), and Tang and Van Ranst (1992). A major problem with this work is the large number of parameters to be modelled. Perhaps, this could be solved by optimisation of the models in order to exclude redundant parameters. Even though there is more scope for the use of methods based on fuzzy sets in environmental management and land evaluation, further work is required for a more objective formulation of MFs and their complex combinations.

3.3. Applications in soil (geographical) information systems

The inadequacy of the traditional Boolean logic for the design of spatial database have been identified since the 'eighties (Robinson and Strahler, 1984; Burrough, 1986, 1989). In his excellent book on Geographic Information Systems (GIS), Burrough (1986) had foreshadowed the potential of fuzzy logic as an alternative to Boolean logical foundation in the GIS design. Robinson and Strahler (1984) had earlier discussed the possibility of directly utilising fuzzy sets theory for the storage of geographic features as linguistic data. More recently, an automatic method for the interactive acquisition of fuzzy information was developed by Robinson (1990). Kollias and Voliotis (1991) also reported a prototype soil information system with fuzzy retrieval capabilities based on fuzzy reasoning. Wang et al. (1990) developed a fuzzy information representative scheme embedded in a conventional GIS package.

Extension and application of fuzzy sets to multi-objective decisions (or multi-attribute decisions) in a spatial context and using fuzzy rules and inference systems, are desirable. Inclusion of other factors, such as socio-economic attributes, that normally would be necessary for a better and more effective land evaluation framework, is possible within a soil information system setup. Sui (1992) demonstrated the combined use of GIS and fuzzy sets for urban land evaluation which was based on a set of physical features and socio-economic criteria. The results provide more details of gradual variation of land use value than the conventional GIS modelling. Recently, Banai (1993) applied Saaty's (Saaty, 1980) Analytical Hierarchy Process (AHP) and fuzzy set theory to deal with factor diversity and complexity in land evaluation involving the location of a public utility. In spite of these isolated successes in applying fuzzy systems, including fuzzy logic and inferences, for spatial analysis, few studies have explicitly described their

incorporation into GIS. Further research is needed to develop a suite of fuzzy membership functions incorporated into spatial decision-making processes.

4. Other (potential) areas of application in soil science

Other areas of application of fuzzy sets to soil studies include modelling of soil physical processes, and fuzzy geostatistics. Potential areas of application could be for field soil description, multiple-object decision-making and fuzzy control. There are many processes and schemes in soil science, particularly edaphological processes, that have imbued complexity associated with not only uncertainty due to measurement errors and imprecise boundaries, but also with, uncertainty due to the choice of a particular analytical model and corresponding model parameters (Bardossy et al., 1990a). The latter is similar to the so-called 'soft' tolerance interval boundaries in tolerance analysis (Dubois and Prade, 1980). Application of fuzzy sets in soil morphological description has not been fully explored, even though many of the terms used in describing soil morphology are vague, and loaded with qualitative ambiguity. This section deals with these and other potential areas of applications of fuzzy sets.

4.1. Application to modelling soil physical processes

Recent concern about quality of the soil and water being affected by agricultural and industrial activities has drawn to our attention the importance of modelling soil processes for better monitoring of the soil and water systems. One of the most important soil physical processes is water movement into and through the unsaturated zone especially through the near-surface layers, the vadose zone. It is the behaviour of the vadose zone that determines aquifer recharge, pollutant transport, salt leaching, and plant growth.

Modelling of soil physical processes generally requires a large number of parameters. More often than not, many of these parameters are determined only at a few locations. Due to the complexity of these processes and the soil system, it is very difficult to estimate these parameters at unsampled locations from the few measured locations. Moreover, several of these parameters influence the models in a highly non-linear manner with consequent results that could be very sensitive to even a small parameter change (Bardossy and Duckstein, 1995). In spite of the contumacious problem of modelling these processes, mathematical models have now become an indispensable tool for quantifying and integrating the important physical, biological and chemical processes operative in the vadose zone. However, whether applications of these models to real-life cases are reliable and precise is questionable. It is this realisation that led Bardossy and Disse (1993) and Bardossy and Duckstein (1995) to fuzzify the commonly used numeric approach to model water movement through the unsaturated zone of the soil. This section gives a brief account of their method and results.

Since the early 'thirties, Richards' equation (Richards, 1931) has been the most basic mathematical expression for describing unsaturated water flow through porous media. A partial differential equation is used to model non-steady flow through a multidimen-

sional anisotropic and non-homogeneous soil matrix. This is achieved by combination of Darcy's law (vertical flow rate) and continuity equation yielding the non-linear Richards' equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[K(\theta) \left(\frac{\partial \psi}{\partial z} - 1 \right) \right]$$

where θ is the moisture content, $K(\theta)$ is the hydraulic conductivity, which is a function of the moisture content, and $\psi(\theta)$ is the matric head (which is also a function of the moisture content). The hydraulic properties of a soil can be described by the curves of $\psi(\theta)$ and $K(\psi)$, respectively (assuming there is a one-to-one correspondence between θ and ψ), using Van Genuchten's equations (Van Genuchten, 1980) expressed as:

$$K_r(\psi) = \frac{\left(1 - \left(\frac{\psi}{h_\alpha} \right)^{n-1} \left[1 + \left(\frac{\psi}{h_\alpha} \right)^n \right]^m \right)^2}{\left[1 + \left(\frac{\psi}{h_\alpha} \right)^n \right]^{\frac{m}{2}}} \quad (14)$$

and

$$\psi(\theta) = h_\alpha \left[\left(\frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^{-\frac{1}{m}} - 1 \right]^{\frac{1}{n}} \quad (15)$$

with $m = 1 - (1/n)$, $n =$ Van Genuchten's shape parameter ($-$), $h_\alpha = 1/\alpha$ (cm), $\alpha =$ Van Genuchten scale parameter (cm^{-1}), $\theta_r =$ residual water content ($\text{cm}^3 \text{cm}^{-3}$), $\theta_s =$ saturated water content ($\text{cm}^3 \text{cm}^{-3}$).

In using Eqs. (14) and (15), five parameters, θ_r , θ_s , K_s , n , and α are required for each spatial unit. All these parameters are difficult to either measure or estimate, which makes it very limiting to apply the models. A more simplified fuzzy approach was developed by Bardossy and Disse (1993) for modelling water flux in unsaturated zone of the soil.

In adopting the fuzzy approach, fuzzy rules are formulated based on the idea that the water movement at given time and location depends on the moisture condition in the immediate neighbourhood of the location. The rule system therefore takes less than the five parameters above as a basis for the rule arguments. In other words, two linear coefficients, K_s (saturated hydraulic conductivity) and θ_s (saturated water content), and all the non-linearity were used to generate the rules for describing vertical movement in the vadose zone. The rule thus formulated is:

If the relative soil moisture in element i is H_{1i} , and the relative moisture content of the adjacent element below is H_{2i} , then the normalised flux between the elements is F_i .

The relative moisture content, θ , is expressed as

$$\theta = \frac{\theta}{\theta_s}$$

The relative moisture content of two adjacent layers, θ_j and θ_{j+1} , are used for calculation of the degree of fulfilment (DOF) grade, v_i , for each rule i . The DOF is the truth value that commensurates with the fulfilment of the conditions of a rule for a given premise (Bardossy and Duckstein, 1995). The DOF is similar to fuzzy sets, as it takes the value [0,1], except that for rules, truth value is formulated to define a more complex statement. The truth value can be calculated using the basic fuzzy operators defined above, including the max–min operator.

The flow between the layers, $Q_{j,j+1}$, is obtained by weighted sum of combination and the mean defuzzification:

$$Q_{j,j+1} = \frac{\sum_i v_i q_i K_s}{\sum_i v_i K_s^*}$$

where q_i is the mean of $F_i K_s^*$ and K_s^* is the estimated saturated hydraulic conductivity. This is followed by conversion of the actual flow to equivalent average moisture content of the layer. The whole process is then repeated for each time step of 0.25, 1.00 and 1.75 h.

Fig. 8a shows the moisture content profiles at different time steps using the finite-element solution of Richards' equation in comparison with the fuzzy rule-based model (Fig. 8b). It is clearly demonstrated that there is close similarity between the two models.

Further experimental work was carried out by Bardossy and Disse (1993) to demonstrate the superiority of the fuzzy rule-based model over Green and Ampt (1911) (AG) model which also requires the five parameters listed above. Comparisons were

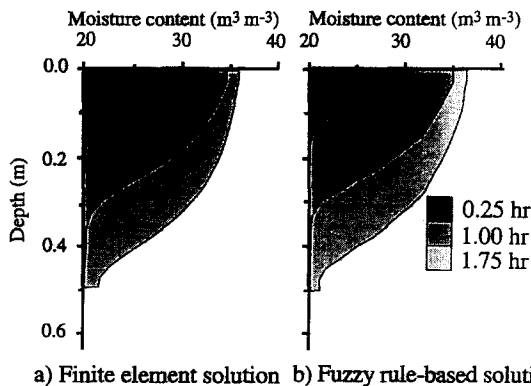


Fig. 8. Soil moisture profiles obtained by (a) finite element solution of Richards' equation, and (b) the fuzzy rule-based solution (after Bardossy and Duckstein, 1995).

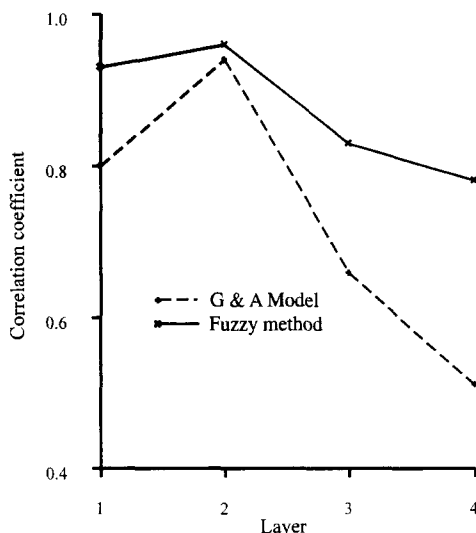


Fig. 9. Correlation coefficients between the measured moisture content and the corresponding simulated moisture content obtained by Green and Ampt's (1911) method and fuzzy rule-based model (data from Bardossy and Duckstein, 1995).

made of moisture content simulated by these two models with measured moisture content. Fig. 9 shows the correlation coefficient between the simulated moisture content by the different models and measured moisture content. The fuzzy Richards' model clearly out-performed the non-fuzzy model. The fuzzy Richards' model can also be extended to three dimensions, which consider flux in lateral and vertical directions. Fig. 10 shows and compares the results of finite-element solution of Richards' equation and that of the fuzzy model. Once again the fuzzy solution is comparable with the conventional finite-element model solution. Furthermore, the model calculations using

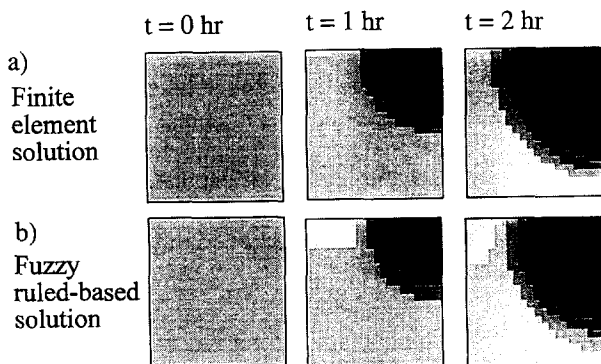


Fig. 10. Infiltration resulting from (a) finite element solution and (b) fuzzy rule-based solution (after Bardossy and Duckstein, 1995).

fuzzy rule-based models require much less time in comparison to the non-fuzzy modelling (Bardossy and Duckstein, 1995).

Research on the use of the fuzzy rule-based approach to model other soil physical and chemical processes would be appropriate. Such models could be incorporated into appropriate complex models, such as LEACHM (Leaching Estimation and Chemistry Model of Wagenet and Huston, 1989), for more accurate simulation of the various physical and chemical processes occurring in the soil.

4.2. *Fuzzy geostatistics in soil spatial modelling*

The problem of applying geostatistics to soil studies is two-fold.

(1) As the soil is multivariate, it is difficult, in practice, to apply geostatistical interpolation methods to model its variation as a whole.

(2) Soil variables may vary continuously and sometimes independently, within the soilscape or the mapping unit, leading to an inherent stochastic–deterministic uncertainty, i.e. there is no certainty in their spatial representations based on stochastic–deterministic models.

The multivariate problem can be resolved severally. There are many attempts at using heterotopic cokriging to simultaneously predict several soil variables onto unsampled locations (e.g., McBratney and Webster, 1983), but this is only readily applicable to numerical interval data that exhibit spatial dependence. Many of soil data are either nominal or ordinal, often lacking spatial autocorrelation, although cokriging can be performed on their indicator transforms. An alternative that takes the uncertainty into account is to use fuzzy sets in combining all types of soil data, yielding membership values as a measure of individual ‘likeness’ to the soil class exemplar. Odeh et al. (1990) used fuzzy *k*-means to produce membership grades that were subsequently used for designing optimal sampling strategy for mapping the soil as a whole. The sampling design was implemented and subsequently followed by kriging of the class membership values which were combined to form a composite map (also see Fig. 6 that illustrates continuousness of the soil system) (Odeh et al., 1992b). McBratney and De Gruijter (1992) transformed the membership values to log-ratios prior to kriging. The ratios were then back-transformed for subsequent mapping of the soil classes. Work is continuing to determine the best kriging technique for continuous class membership values, especially, as the latter are considered as compositional data.

The stochastic–deterministic uncertainty inherent in an intrinsic random function defining a variable has been resolved based on the theory of regionalised variables (Matheron, 1965). But, according to Bardossy et al. (1990a), there are uncertainties due to the choice of a particular analytical model and the corresponding model parameters. The variogram, which summarises the variation of a soil property, plays a crucial role in spatial analysis of soil. Matheron (1965) discussed the fluctuations in the variogram, distinguishing between fluctuations due to different realisations of a generating process and those due to repeated sampling of a single realisation of the process (i.e., sampling error). It is precisely the uncertainties associated with the latter that Bardossy et al. (1988, 1990a,b) applied fuzzy sets to quantify ‘common sense’ and ‘sensible’ variogram determination, and associated fuzzy kriging. Their methods involved the following two main steps.

(1) Determination of fuzzy variogram parameters, a method that relies on the uncertainty caused by the choice of a particular analytical model and uncertainty due to the variogram parameters. This step builds FMFs on both the choice of the model and the variogram parameters.

(2) Calculation of fuzzy kriged estimates and the corresponding estimation variance: using the extension principle (see Eqs. (3)–(5)), the fuzzy kriging operator is developed for which the membership value of any real number resulting from kriging is derived. Thus the MF for the fuzzy kriging estimator integrates the uncertainty of the variogram model (choice) and model parameters as determined in step 1 above. The final kriged estimates are the fuzzy number corresponding to the fuzzy kriging MFs.

The idea of fuzzy kriging is intuitively reasonable, but the validity of the method, especially in soil science, has not been tested. It would require testing of the method against existing methods (ordinary kriging, for example) by external means with a test set. This is a challenge to soil geostatisticians.

4.3. Application to image analysis

The importance of the pore structure in modelling of soil physical processes has been discussed by McBratney and Moran (1994), and Moran and McBratney (1997). A model based on digital images which encompasses many or all scales was developed by Moran et al. (1989). It provides a unique methodology for elucidating the physical behaviour of the soil without the need to measure at all sites of interest. The model is based on two- or three-component digital images obtained by impregnation of the soil with resins. Planar surfaces of the impregnated soil blocks are then video-digitized for further image segregation followed by statistical analysis. In the segmentation process, there is the difficulty associated with fuzzy imprecise boundaries between the solid and the pore space, i.e., the edge boundaries are fuzzy. At certain scales, therefore, it may be difficult to resolve the solid and pore components. McBratney and Moran (1994) developed and applied a model based on the concept of *fuzzy random pseudofractal sets*. In defining structure of a three-dimensional volume of soil, a universal set, Ω , is divided into two subsets as:

$$\Omega = \mathbf{P} \cup \mathbf{S} \text{ or } \mathbf{P} \cup \bar{\mathbf{P}}$$

where \mathbf{P} is the pore or void and \mathbf{S} the soil solid. The is the complement of $\bar{\mathbf{P}}$ (the same as \mathbf{S}).

According to Moran and McBratney (1997), the two subsets lack deterministic geometry as they are considered random. As the subsets are fuzzy, the most rational thing to do is to fit a fuzzy membership function to either of the subsets, preferentially, the pore subset. The membership to the latter subset could be referred to as the soil porosity. The porosity thus is represented by membership values ranging between zero (white) and unity (black). A model based on the *pore subset* was implemented at a coarse scale to simulate using two components: (1) the seed types, and (2) the spatial distribution of the seed centroids. The seed types are of three main shapes: circular, ellipsoidal and linear. The centroids have spatial point distributions which may be either random, clustered or regular. A series of structures was generated varying from circular solid at clustered locations to ellipsoidal pores at regular locations. The model was

extended to a finer scale. It demonstrates how the use of fuzzy sets allow for an extension of binary image analysis to grey-level images. Additionally, Moran and McBratney (1997) defined the concept of a *fuzzy connectivity function* and measure it. On the whole, the model shows promise for development of quantitative soil pore structural studies which could be used to replicate real structure-generating processes, e.g., air entry due to shrinkage and compaction.

4.4. Fuzzy logic and multi-attribute decision-making

As stated above, many soil phenomena are described by multiple attributes which are characterised by vague conception and/or are subjective. Such imprecise information are better analysed and simplified by incorporation of fuzzy sets with fuzzy logic operations that would yield results that are themselves fuzzified, but not as ambiguous. The procedure involves two main steps. The first step is to fit MF's onto the input variables or attributes (note that the input attributes are always measured in discrete numerical values). This requires expert knowledge of the phenomenon in order to discern the MF's appropriate to the degree of which the inputs belong to each of the fuzzy sets. The second step involves the fuzzy rules and logic operations leading to decisions regarding the phenomenon of interest. This is the fuzzy inference system as discussed in Section 2.5. We make rules which are also based on our expert knowledge of the phenomenon. The rules are the *IF ... THEN* type combined with the fuzzy operators, which include the commonly used *AND*, *OR* and *NOT*. The *THEN* in the rules normally refers to the expected outputs. The expected outputs are also fuzzy sets fitted with MF's. Decisions regarding the rules are implied by the outputs.

Application to soil mottle description. The method described above is termed 'a fuzzy multi-attribute decision-making (FMCDM)' approach (Odeh, 1996). An example of its application to description of soil mottle characteristics is given. A mottled soil layer may be indicative of soil drainage condition as affected by the external factors and the inherent soil characteristics. The soil mottle is characterised by several attributes: abundance, size, colour tone and colour contrast, which are based on subjective, ambiguous, vague and conflicting connotations. For this reason, it is difficult to relate them to the soil redox processes. For example, is it reasonable to say that a soil layer with few, coarse and prominent mottles is less mottled than a layer with common, distinct and very coarse mottles? If so what is the precise boundary between the various levels of mottleness? What is the degree of mottleness and how should it be quantified?

In order to answer these questions we formulate decision rules using fuzzy logic operators and the *THEN* implication. Examples of the rules look like this:

IF abundance is many OR contrast is prominent OR size is v.coarse, THEN soil mottleness is v.strong

IF abundance is common OR contrast is distinct OR size is medium, THEN soil mottleness is strong

IF abundance is few OR contrast is faint OR size is fine, THEN soil mottleness is v.slight.

The rules lead to decisions. An example of the first rule above is presented in Fig. 11. A soil horizon with membership value of 0.75 'many' (abundance) and 0.60 'prominent' (contrast) and 0.55 'very coarse' (size), implies that the soil has a 0.75 grade of 'very strongly' mottled. The 'or' operator, as maximisation of the inputs, was used in the logic operations. The implication operator is the minimisation of the output to the degree of the maximised input. Thus the output or decision regarding the degree of mottleness is $\min(0.75, 1) = 0.75$. This quantity is apparently easier to perceive, more readily mappable and amenable to further application in modelling the soil redox process, than the mottle characteristics described using the non-fuzzy (traditional) approach. The FMCDM approach has great potential in its application to the description of other soil morphological properties and more importantly, land evaluation.

Application to soil boundaries. Recently, fuzzy tools have been applied to soil boundaries, both in theory and practice (Burrough, 1996; McBratney and Whelan, 1995). There is also the work by Lagacherie et al. (1997), which tackles the problem of uncertainty of soil boundaries in GIS applications. McBratney and Whelan (1995) developed a method, based on a distance-membership function, for fuzzifying conventional soil maps such that the gradational nature of soil variation and imprecise nature of soil maps between mapping units are accounted for. The result is a continuous soil map that provides enhanced subtlety in depicting changes in soil types. As we move to the next technological stage in site-specific soil management, such a soil map will provide the needed information in a more efficient way.

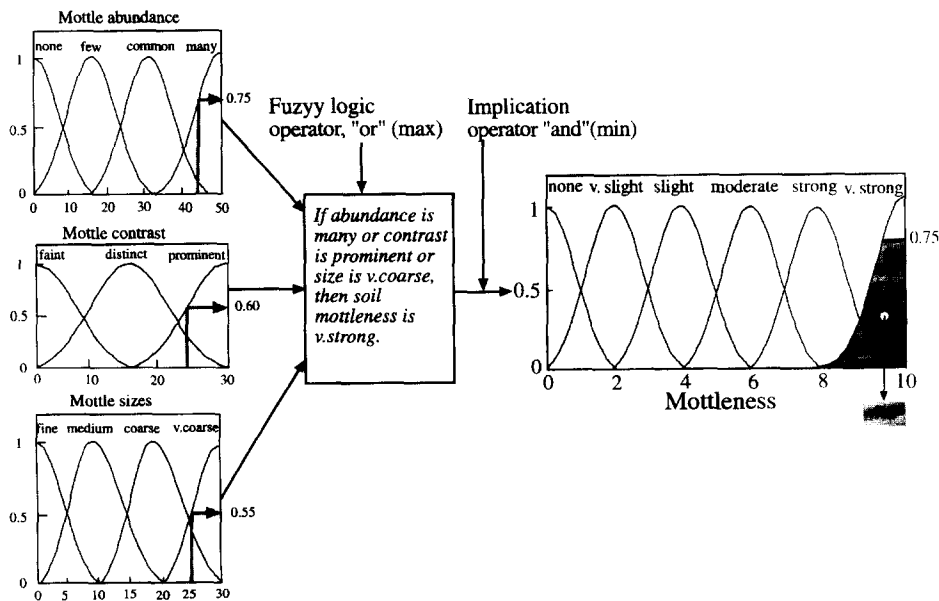


Fig. 11. Fuzzy multi-criteria decision-making on soil mottle description.

Application to development of soil quality indicators. Recently, there has been increased awareness of the importance of *soil quality* for different land utilisation types, including agricultural production. However, like soil morphological properties, the term soil quality, in physical terms, may not connote a single soil attribute, but in most cases is a multi-variable index. More importantly, any method designed to assess such an agglomeratic index as soil quality should include all the crucial attributes or indicators. Smith et al. (1993) proposed a multiple-variable indicator transformed as a technique for integration of specific criteria defining soil quality. Their method does not take into consideration the uncertainty due to ambiguous and vague definition of the term soil quality, and the uncertainty associated with measurement error. The fuzzy set approach offers a variety of techniques that will enable agglomeration of a set of soil criteria/parameters into a single index of soil quality. The use of fuzzy expert knowledge, fuzzy rules and logic, combined with the extension principle (Eqs. (2) and (3)), may achieve such an objective. This needs to be explored.

5. Conclusions

Fuzzy systems, including fuzzy set theory and fuzzy logic, provide a rich and meaningful improvement or extension of conventional logic. The mathematics generated by this theory is consistent, and fuzzy set theory may be a generalisation of classic set theory. The applications in soil science, of which may be generated from or adapted to fuzzy set theory and fuzzy logic, are wide-ranging and include:

(1) *Soil numerical classification and mapping.* The soil is more continuous than it is discrete. The FCM algorithm can classify the soil into continuous classes, therefore incorporating more information onto the resulting maps than is the case with conventional methods. The FCM has been re-programmed to allocate soil individuals into existing systems of classification.

(2) *Land evaluation.* Fuzzy set theory and logic are suitable tools for land evaluation as they take care of the vague terms used in describing the primary landscape variables normally used for land evaluation, and uncertainty associated with measurement errors.

(3) *Modelling and simulation of soil physical processes.* Many models of the soil processes are complex requiring many parameters. Fuzzy rule-based models are capable of optimising conventional models, therefore reducing the number of parameters. Extension to physico-chemical and biological models is desirable.

(4) *Fuzzy variogram and kriging of soil variables.* In addition to stochastic–deterministic uncertainty there are uncertainties due to the choice of the variogram model and the corresponding model parameters. Fuzzy sets allow for the choice of appropriate variogram models from a number of options, and better optimisation of parameter modelling and subsequent fuzzy kriging.

(5) *Grey-scale digital image analysis.* Conventional image analysis relies mostly on binary digital information. Fuzzy sets could extend binary image analysis to grey-level images.

(6) *Fuzzy measures of imprecisely defined soil phenomena.* Many of the morphological and physical soil properties are multi-attributed, with vague terms describing each

property. Fuzzy logical operations in a multi-attribute decision-making algorithm could be used for the derivation of a single quantity for each of the properties.

(7) *Soil quality indices*. Soil quality is a new term requiring proper definition and quantification. It could be quantified by derivation of an index from multiple soil variables. The fuzzy-set approach could provide for agglomeration of a set of crucial soil parameters into a single index of soil quality important for land management.

In conclusion, fuzzy set theory has great potential in soil science. In addition to the applications above, many other soil concepts or systems may be modelled, simulated, and even replicated with the help of fuzzy systems, not the least of which is human reasoning itself.

Acknowledgements

After all that we are a bit fuzzy so we are not sure who to thank. Thanks to the *excluded middle*.

References

- Banai, R., 1993. Fuzziness in geographical information systems: contributions from the analytical hierarchical process. *Int. J. Geogr. Inf. Syst.* 7, 315–329.
- Bardossy, A., Disse, M., 1993. Fuzzy rule-based models for infiltration. *Water Resour. Res.* 29, 373–382.
- Bardossy, A., Duckstein, L., 1995. *Fuzzy Rule-Based Modelling with Applications to Geophysical, Biological and Engineering Systems*. CRC Press, New York, 113 pp.
- Bardossy, A., Kelly, W.E., Bogardi, I., 1988. Imprecise (fuzzy) information in geostatistics. *Math. Geol.* 20, 189–203.
- Bardossy, A., Bogardi, I., Kelly, W.E., 1990a. Kriging with imprecise (fuzzy) variogram, Part I. Theory. *Math. Geol.* 22, 63–79.
- Bardossy, A., Bogardi, I., Kelly, W.E., 1990b. Kriging with imprecise (fuzzy) variogram, Part II. Application. *Math. Geol.* 22, 81–94.
- Bezdek, J.C., 1974. Numerical taxonomy with fuzzy sets. *J. Math. Biol.* 1, 57–71.
- Bezdek, J.C., 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.
- Black, M., 1937. Vagueness: an exercise in logical analysis. *Philos. Sci.* 4, 427–455.
- Brule, F.J., 1985. Fuzzy systems — a tutorial. Source: Internet Newsgroups: comp.ai; <http://www.quadralay.com/www/Fuzzy/tutorial.html>.
- Burrough, P.A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Clarendon Press, Oxford, 194 pp.
- Burrough, P.A., 1987. Mapping and mapping analysis: new tools for land evaluation. *Soil Use Manage.* 3, 20–25.
- Burrough, P.A., 1989. Fuzzy mathematic methods for soil survey and land evaluation. *J. Soil Sci.* 40, 447–492.
- Burrough, P.A., 1996. Natural objects with indeterminate boundaries. In: Burrough, P.A., Frank, A. (Eds.), *Geographic Objects with Indeterminate Boundaries*. Taylor and Francis, London, pp. 3–28.
- Burrough, P.A., MacMillian, R.A., van Deusen, W., 1992. Fuzzy classification methods for determining land suitability from soil profile observations and topography. *J. Soil Sci.* 43, 193–210.
- Burrough, P.A., Van Gaans, P.F.M., Hootsmans, R., 1997. Continuous classification in soil survey: spatial correlation, confusion and boundaries. In: De Gruijter, J.J., McBratney, A.B., McSweeney, K. (Eds.), *Fuzzy Sets in Soil Science*. *Geoderma*, 77: 115–135 (this issue).
- Chang, L., Burrough, P.A., 1987. Fuzzy reasoning: a new quantitative aid for land evaluation. *Soil Surv. Land Eval.* 7, 69–80.

- Dale, M.B., McBratney, A.B., Rusell, J.S., 1989. On the role of expert systems and numerical taxonomy in soil classification. *J. Soil Sci.* 40, 223–234.
- De Gruijter, J.J., McBratney, A.B., 1988. A modified fuzzy k-means method for predictive classification. In: Bock, H.H. (Ed.), *Classification and Related Methods of Data Analysis*. Elsevier, Amsterdam, pp. 97–104.
- De Gruijter, J.J., Walvoer, D.J.J., Van Gaans, P.F.M., 1997. Continuous soil maps: a fuzzy set approach to bridge the gap between aggregation levels of process and distribution models. In: De Gruijter, J.J., McBratney, A.B., McSweeney, K. (Eds.), *Fuzzy Sets in Soil Science*. *Geoderma*, 77: 169–195 (this issue).
- Dubois, D., Prade, H., 1980. *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York, 393 pp.
- Dunn, J.C., 1974. A fuzzy relative of the isodata process and its use in detecting compact, well-separated clusters. *J. Cyber.* 3, 22–57.
- FAO, 1976. *A Framework for Land Evaluation*. Soils Bulletin 32, FAO, Rome, 72 pp.
- FAO, 1994. *Draft World Reference Base for Soil Resources*. FAO, ISSS and ISRIC, Rome–Wageningen.
- Fisher, R.A. 1954. *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh, 354 pp.
- Fridland, V.M., 1974. Structure of the soil mantle. *Geoderma* 12, 35–41.
- Green, W.H., Ampt, G.A., 1911. Studies of soil physics, I. The flow of air and water through soils. *J. Agric. Sci.* 4, 1–24.
- Hendricks Franssen, H.J.W.M., Van Eijnsbergen, A.C., Stein, A., 1997. Use of spatial prediction techniques and fuzzy classification for mapping soil pollutants. In: De Gruijter, J.J., McBratney, A.B., McSweeney, K. (Eds.), *Fuzzy Sets in Soil Science*. *Geoderma*, 77: 243–262 (this issue).
- Heuvelink, G.B.M., Burrough, P.A., 1993. Error propagation in logical cartographic modelling using Boolean methods and continuous classification. *Int. J. Geogr. Inf. Syst.* 7, 321–246.
- Jang, J.-S.R., Gulley, N., 1995. *Fuzzy Logic Toolbox User's Guide*. The MathWorks, Inc., Natick, Mass., 274 pp.
- Kandel, A., 1986. *Fuzzy Mathematical Techniques with Applications*. Addison-Wesley, Mass.
- Kaufmann, A., 1975. *Introduction to the Theory of Fuzzy Sets, 1. Fundamental Theoretical Elements*. Academic Press, New York.
- Kaufmann, A., Gupta, M.M., 1991. *Fuzzy Mathematical Models in Engineering and Management Science*. North-Holland, Amsterdam, 356 pp.
- Kollias, V.J., Voliotis, A., 1991. Fuzzy reasoning in the development of geographical information systems FRIS: a prototype of soil information system with fuzzy retrieval capabilities. *Int. J. Geogr. Inf. Syst.* 5, 209–223.
- Korner, S., 1967. Laws of thought. In: *Encyclopedia of Philosophy*, Vol. 4, MacMillan, New York, pp. 414–417.
- Lagacherie, P., Cazemier, D.R., Van Gaans, P., Burrough, P.A., 1997. Fuzzy k-means clustering of fields in an elementary catchment and extrapolation to a larger area. In: De Gruijter, J.J., McBratney, A.B., McSweeney, K. (Eds.), *Fuzzy Sets in Soil Science*. *Geoderma*, 77: 197–216 (this issue).
- Lejewski, C., 1967 Jan Lukasiewicz. In: *Encyclopedia of Philosophy*, Vol. 5, MacMillan, New York, pp. 104–107.
- Matheron, G., 1965. *Les Variables Regionalisées et leur Estimation*. Masson, Paris, 306 pp.
- Mazaheri, S.A., Koppi, A.J., McBratney, A.B., 1995. A fuzzy allocation scheme for the Australian Great Soil Groups classification system. *Eur. J. Soil Sci.* 46, 601–612.
- Mazaheri, S.A., McBratney, A.B., Koppi, A.J., 1997. Sensitivity of memberships to attribute variation around selected centroids and intergrades in the continuous Australian Great Soil Groups classification system. In: De Gruijter, J.J., McBratney, A.B., McSweeney, K. (Eds.), *Fuzzy Sets in Soil Science*. *Geoderma*, 77: 155–168 (this issue).
- McBratney, A.B., 1992. On variation, uncertainty and informatics in environmental soil management. *Aust. J. Soil Res.* 30, 913–935.
- McBratney, A.B., 1994. Allocation of new individuals to continuous soil classes. *Aust. J. Soil Res.* 32, 623–633.
- McBratney, A.B., De Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. *J. Soil Sci.* 43, 159–175.

- McBratney, A.B., Moore, A.W., 1985. Application of fuzzy sets to climatic classification. *Agric. For. Meteorol.* 35, 165–185.
- McBratney, A.B., Moran, C.J., 1994. Soil pore structure modelling using fuzzy random pseudofractal sets. In: Ringrose-Voase, A.J., Humphreys, G.S. (Eds.), *Soil Morphology: Studies in Management and Genesis*. Proc 9th Int. Working Meet. Soil Micromorphology, Townsville, Australia, pp. 495–506.
- McBratney, A.B., Webster, R., 1983. Coregionalisation and multiple sampling strategy. *J. Soil Sci.* 34, 249–263.
- McBratney, A.B., Whelan, B.M., 1995. Continuous models of soil variation for continuous soil management. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), *Proceedings of Site-specific Management for Agricultural Systems*. ASA, CSSA and SSSA, Madison, pp. 325–338.
- McBratney, A.B., De Grijter, J.J., Brus, D.J., 1992. Spatial prediction and mapping of continuous soil classes. *Geoderma* 54, 39–64.
- McNeill, D., Freiberger, P., 1993. *Fuzzy Logic*. Bookman Press, Melbourne.
- Moran, C.J., McBratney, A.B., 1997. A two-dimensional fuzzy model of soil pore structure. *Math. Geol.*, in press.
- Moran, C.J., McBratney, A.B., Koppi, A.J., 1989. A rapid analysis method for soil macropore structure. I. Specimen preparation and digital binary image production. *Soil Sci. Soc. Am. J.* 51, 921–928.
- Odeh, I.O.A., 1990. *Soil Pattern Recognition in a South Australian Subcatchment*. Ph.D. Thesis, The University of Adelaide, S.A.
- Odeh, I.O.A., 1996. A fuzzy multi-criterial approach to field soil description: an example. Proc. Pedometrics Symposium, Organised by ACLEP and ISSS Working Group on Pedometrics, Melbourne, July 1–4, 1996, pp. 201–202.
- Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1990. Design of optimal sampling spacings for mapping soil using fuzzy-*k*-means and regionalised variable theory. *Geoderma* 47, 92–122.
- Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1992a. Soil pattern recognition with fuzzy-*c*-means: applications to classification and soil–landform interrelationships. *Soil Sci. Soc. Am. J.* 56, 505–516.
- Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1992b. Fuzzy-*c*-means and kriging for mapping soil as a continuous system. *Soil Sci. Soc. Am. J.* 56, 1848–1854.
- Powell, B., McBratney, A.B., MacLeod, D.A., 1991. The application of fuzzy classification to soil pH profiles in the Lockyer Valley, Queensland, Australia. *Catena* 18, 409–420.
- Richards, L.A., 1931. Capillary conduction of liquids through porous media. *Physics* 1, 318–333.
- Robinson, V.B., 1990. Interactive machine acquisition of a fuzzy spatial relation. *Comput. Geosci.* 16, 857–872.
- Robinson, V.B., Strahler, A.H., 1984. Issues in designing geographic information systems under conditions of inexactness. Proc. 10th Int. Symp. Machine Processing of Remotely Sensed Data, pp. 179–188.
- Ruspini, E.H., 1969. A new approach to clustering. *Inf. Control* 15, 22–32.
- Saaty, T.L., 1980. *The Analytic Hierarchy Process*. McGraw-Hill, New York.
- Schmucker, K.J., 1984. *Fuzzy Sets, Natural Language Computations, and Risk Analysis*. Computer Science Press, Rockville, Maryland, 192 pp.
- Smith, J.L., Halvorson, J.J., Papendick, R.I., 1993. Using multiple-variable indicator kriging for evaluating soil quality. *Soil Sci. Soc. Am. J.* 57, 743–749.
- Soil Survey Staff, 1994. *Keys to Soil Taxonomy*, 6th ed. USDA Soil Conservation Services, Washington, D.C.
- Sui, D.Z., 1992. A fuzzy GIS modelling approach for urban land evaluation. *Comp. Environ. Urban Syst.* 16, 101–115.
- Tang, H.J., Van Ranst, E., 1992. Testing of fuzzy set theory in land suitability assessment for rainfed grain maize production. *Pedologie* 42, 129–147.
- Tang, H.J., Debavye, J., Ruan, D., Van Ranst, E., 1991. Land suitability classification based on fuzzy sets theory. *Pedologie* 41, 277–290.
- Triantafyllis, J., McBratney, A.B., 1993. Application of continuous methods of soil classification and land suitability assessment in the lower Namoi valley. CSIRO Australia Division of Soils, Divisional Rep. 121, CSIRO, Melbourne.
- Van Gaans, P.F.M., Burrough, P.A., 1993. The use of fuzzy logic and continuous classification in GIS applications. In: Harts, J.J., Ottens, H.F.L., Scholten, H.J. (Eds.), *Proc. UGIS'93*. Utrecht–Amsterdam, pp. 1025–1034.

- Van Genuchten, M.Th., 1980. A close-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44, 892–898.
- Wagenet, R.J., Huston, J.L., 1989. LEACHM Leaching Estimation and Chemistry Model. A process-based model of water and solute movement, transformations, plant uptake and chemical reactions in the unsaturated zone. Continuum Vol. 2 (Version 2), Water Resource Institute, Centre for Environmental Research, Cornell University.
- Wang, F.-J., Hall, G.B., Subaryono, 1990. Fuzzy information representation and processing in conventional GIS software: database design and application. *Int. J. Geogr. Inf. Syst.* 4, 261–283.
- Waterstone, M., 1994. Institutional analysis and water resources management. In: Duckstein, L., Parent, E. (Eds.), *Natural Resources Management*. Kluwer, Dordrecht.
- Webster, R., 1985. Quantitative spatial analysis of soil in the field. *Adv. Soil Sci.* 3, 1–70.
- Zadeh, L.A., 1965. Fuzzy sets. *Inf. Control* 8, 338–353.
- Zadeh, L.A., 1978. Fuzzy sets as a basis for theory of possibility. *Fuzzy Sets Syst.* 1, 3–28.