Math Geol (2007) 39: 765–780
DOI 10.1007/s11004-007-9123-7

Hydraulic Conductivity Estimation via Fuzzy Analysis of Grain Size Data

James Ross · Metin Ozbek · George F. Pinder

Received: 24 October 2005 / Accepted: 8 February 2007 / Published online: 3 October 2007

Abstract A measure of hydraulic conductivity is arguably the most important variable to practicing hydrogeologists. However, the amount of readily available hydraulic conductivity data at any site is generally small, given the resources required to adequately sample a spatial domain. However, other hydrogeologic data, such as grain size distributions and soil descriptions, are often rather easy to obtain. A fuzzy reasoning algorithm is used to define a relationship between soil grain size and hydraulic conductivity. By introducing soil grain distributions and qualitative borehole log descriptions into this fuzzy inference system, hydraulic conductivity can be estimated. The theory is defined, and an application to data from a Superfund site is provided, where the inference procedure produces accurate hydraulic conductivity estimates.

Keywords Fuzzy sets · Approximate reasoning · Grain size distributions · Borehole logs · Qualitative data

Introduction

Large amounts of information often result from a hydrogeologic investigation, whereas hydraulic conductivity data comprise a relatively small subset because the tests required to determine conductivity are generally both time consuming and expensive. Consequently, regional estimates of hydraulic conductivity are inherently uncertain. Since significant uncertainty in hydraulic conductivity estimates lead to uncertainty in groundwater flow and transport models, less costly means of acquiring hydraulic conductivity data have been explored. Specifically, several attempts

J. Ross (✉) · M. Ozbek · G.F. Pinder
Research Center for Groundwater Remediation Design, Civil and Environmental Engineering Program, College of Engineering and Math, University of Vermont, 33 Colchester Avenue, Burlington, VT 05446, USA
e-mail: jlross@uvm.edu

have been made to relate other more easily measured hydrogeologic parameters to hydraulic conductivity. Porosity, particle sorting, texture and soil type have all been related to hydraulic conductivity. These relationships have intuitive appeal.

The determination of porosity requires little effort and expense when compared to that of hydraulic conductivity, and relationships between porosity and hydraulic conductivity have been modeled (Selley 1985). The degree of sorting can be evaluated from the soil's grain size distribution curve and related to hydraulic conductivity through an empirical equation (Demmico and Klir 2004). Textural characteristics (grain size, ductility) are determined by the examination of a soil sample and have been related to hydraulic conductivity in sandstone through fuzzy relations (Fang and Chen 1997). In a very general sense, an understanding of the ranges of likely hydraulic conductivity values for various soil types has been used to map these two variables (Domenico and Schwartz 1990). Fuzzy sets have been used to characterize an expert's notion of the hydraulic conductivity at unsampled locations (Bardossy et al. 1989), essentially providing opinion-based imprecise values of hydraulic conductivity at different locations in a spatial domain. More appropriate to the theory presented herein, attempts have been made to relate porosity and grain size distributions to hydraulic conductivity through petrophysical relations, the Kozeny-Carman Equation and other pedotransfer functions (Uma et al. 1989; Sen 1992; Alyamani and Sen 1993; Koltermann and Gorelick 1995).

Though the above methods differ in their approaches, they share the common goal to expand available hydraulic conductivity data with data derived from complementary sources. These approaches suffer from various shortcomings. Porosity data, although easily measured in the laboratory is not commonly available in practical applications. The degree of sorting is simply a convenient measure of considerably greater information from a soil grain distribution curve. The relationship of the textural measure to hydraulic conductivity has only been evaluated for sandstone. The soil type-to-hydraulic-conductivity relationship is highly uncertain. Fuzzy set hydraulic conductivities provide little help, as fuzzy kriging algorithms produce fuzzy estimates (Diamond 1989; Bardossy et al. 1989, 1990). Finally, attempts to relate hydraulic conductivity to grain size distributions and porosity have not considered the high degree of uncertainty implicit in such a relationship. So far they have relied upon deterministic equations. Such a relation is generally recognized as fuzzy and should be modeled as such.

In this paper, a method is defined whereby hydraulic conductivity values are estimated from grain size distribution curves and qualitative descriptions of soil samples (for example, describing a soil sample as "a silty fine sand with some clay"). The approach recognizes that the grain size distribution curve implicitly contains some of the measures noted above, such as grain size and sorting.

Soil type is also implicitly defined when common soil classifications are plotted along the horizontal axis of the grain size distribution graph. This crisp representation of soil types, which states that a grain size diameter of 0.002 mm characterizes clay, but a grain size diameter of 0.0021 mm is characteristic of a silt, implies unwarranted accuracy. If the soil grain size axis is partitioned so that the soil classification avoids sharp boundaries (i.e. is fuzzy), and there is a fuzzy rule base that defines an imprecise relationship between soil type to hydraulic conductivity, the grain size distribution curve can be used to estimate a hydraulic conductivity value for the sample.

Table 1 Qualitative field data in the form of Borehole log descriptions of the four soil samples used in the application, below

Sample	Borehole log description
1	Silty Fine Sand
2	Silt to Medium Clay Silt, trace Fine Sand
3	Fine Sand, some Clayey Silt, trace Medium Sand
4	Fine Sand, some Silt, trace Medium Sand

This defined estimate of hydraulic conductivity can be further enhanced using an independent qualitative field description of a soil sample provided by a groundwater professional. Such descriptions by a geologist are done by inspecting soil samples taken at borehole locations. The adjectives used to describe a soil sample in the field, such as sandy or clayey, are inherently imprecise. However, such subjective information reflects valuable insight into the sample's makeup and complements the information obtained by a grain size distribution. This qualitative information, converted into fuzzy sets, will produce results similar to the soil grain distribution data. The information from the grain size distribution curves and field assessments can be combined to produce a knowledge base that facilitates hydraulic conductivity estimates. Such an integration of subjective and objective fuzzy geological data is well founded in mineral potential mapping (Cheng and Agterberg 1999; Brown et al. 2003; Porwal et al. 2003). This method is developed and tested using data from the CIBA-Geigy Superfund site in the Toms River, New Jersey.

In the following sections, the creation of the fuzzy sets is discussed. The introduction and transformations of the qualitative field data (Table 1) and grain size distributions into fuzzy data are explained, as is the juxtaposition of the fuzzy data with fuzzy rules. Finally, an example with results from the application of this reasoning mechanism to Toms River data is provided.

Defining Sets and Distributions

The definition of soil grain fuzzy sets is integral to defining the relationship between soil grain size and hydraulic conductivity. If the soil grain fuzzy sets are poorly defined, then the entire relationship is poorly defined and will produce disappointing estimates. The fuzzy sets defined on the domain soil grain size (mm) will be named after the soil types that are generally found on soil grain distribution curves (Coarse Sand, Medium Sand, Fine Sand, Silt, Clay).

The fuzzy sets that are provided are defined by membership functions and aim to capture the imprecision inherent in soil type names, like fine sand and clay. These functions, typically piecewise linear to create triangular or trapezoidal fuzzy sets, represent, for every soil grain size on the abscissa, the degree to which that soil grain size belongs to a particular fuzzy set. Given that the fuzzy sets represent soil types, the membership functions stipulate the degree to which a given soil grain size belongs to the concepts coarse sand, medium sand, etc.

As a rule of thumb, fuzzy sets should both overlap each other and, when combined, cover the range of reasonable domain values. This serves two purposes. First

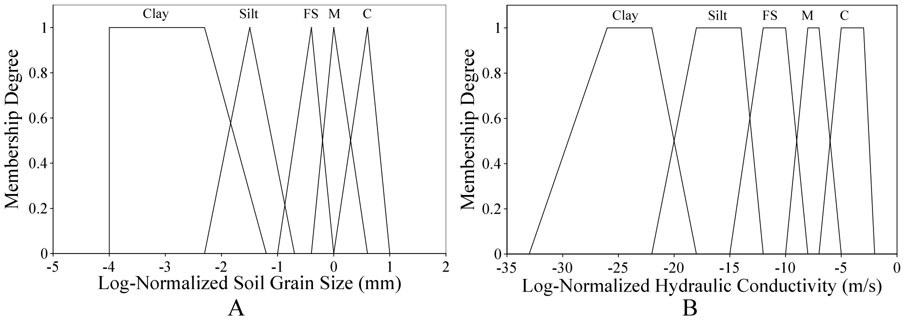


Fig. 1 Fuzzy sets defined by an expert to characterize common soil types (**A**), and their corresponding hydraulic conductivities (**B**). The soil types considered are clay, silt, fine sand (*FS*), medium sand (*M*), and coarse sand (*C*)

Table 2 Parameters for the soil type and hydraulic conductivity fuzzy sets in Fig. 1

Fuzzy set name	Soil type parameters (mm)	Log soil type parameters (log mm)	Log hydraulic conductivity parameters (log m/s)
Clay	[0.0001, 0.0001, 0.005, 0.063]	[−4, −4, −2.3, −1.2]	[−33, −26, −22, −18]
Silt	[0.005, 0.032, 0.2]	[−2.3, −1.5, −0.7]	[−22, −18, −14, −12]
Fine Sand	[0.1, 0.4, 1]	[−1, −0.4, 0]	[−15, −12, −10, −8]
Medium Sand	[0.4, 1, 4]	[−0.4, 0, 0.6]	[−10, −8, −7, −5]
Coarse Sand	[1, 4, 10]	[0, 0.6, 1]	[−7, −5, −3, −2]

is semantic continuity, where a realistic definition of soil types must have gradual and overlapping boundaries. Secondly, overlapping fuzzy sets make for continuous results when used in a rule base. For gradually changing input soil grain size values, the hydraulic conductivity estimates that result from the rule-based reasoning mechanism also gradually change.

The soil type fuzzy sets based on soil grain size are shown in Figure 1A. Notice that they both overlap and overlay the range of reasonable soil grain sizes. These soil type fuzzy sets relied upon traditional crisp delineations of soil types as guidelines. Though there is a precedent for fuzzy soil type definitions (Hsieh et al. 2005), such fuzzy sets did not guide the definitions of the soil types in Figure 1A. The hydraulic conductivity fuzzy sets are defined in a similar manner (Fig. 1B), though the expertise of a hydrogeologist was the only guide in their construction. Information regarding the soil type and hydraulic conductivity fuzzy sets are provided in Table 2. The data in this table describe the parameters of the fuzzy sets. For triangular fuzzy sets the middle value is the median (full membership), while the left and right parameters are the upper and lower bounds of the fuzzy set (zero membership). Trapezoidal fuzzy sets differ by having two median parameters that define an interval of values with full membership.

Fuzzy rules are the means by which an estimate is actually produced for a given input value. Each rule relating a soil type fuzzy set to a hydraulic conductivity fuzzy set defines a vaguely defined "point" that, when plotted along with a number of other

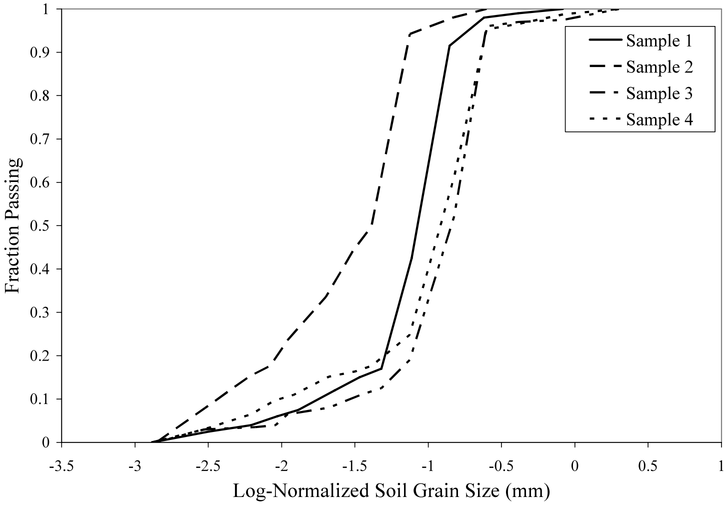


Fig. 2 Grain size distribution curves for the four samples used in the application section, below. The *vertical axis* is analogous to probability

rules relating the same variables, defines an imprecise relationship between these variables. The fuzzy rules provided are very simple in this case. Since the soil grain fuzzy sets share the same names as the hydraulic conductivity fuzzy sets, the rules essentially match like-named fuzzy sets. The rules are structured as follows. If the soil grain size is Coarse Sand then the hydraulic conductivity is that of Coarse Sand. If the soil grain size is Medium Sand then the hydraulic conductivity is that of Medium Sand, etc. The collection of the five rules provides the rule base, and the heart of the fuzzy reasoning process. The reasoning will be discussed later.

The input into each rule is derived from a grain size distribution curve, resulting from sieve and hydrometer analyses. A grain size distribution is a representation of the distribution of soil grain sizes for a given sample of soil (Fig. 2). They are cumulative distribution functions (CDF), which are structurally and theoretically distinct from fuzzy sets. The vertical axis is analogous to a probability value, in that it records what fraction of the soil sample (by weight) is less or equal in size to each grain size value along the horizontal axis. The grain size distribution must be transformed from a probabilistic form to one that is consistent with a fuzzy set representation, without changing or losing the information contained in the original distribution. To achieve this goal, the discrete CDF must first be converted to a discrete probability density function (PDF). The probability values represent the prevalence of each soil grain size.

Using a tool from the Possibility Theory, a probability-possibility transformation (Oussalah 2000) is used to transform the PDF to a possibility distribution (PD), $\pi(x)$, on the variable X (Zadeh 1978). Such a transformation converts uncertain data in 'probabilistic' form to uncertain data in 'possibilistic' (related to fuzzy logic) form. The PD is the sister distribution to the PDF, as it is able to characterize the same data with a different interpretation.

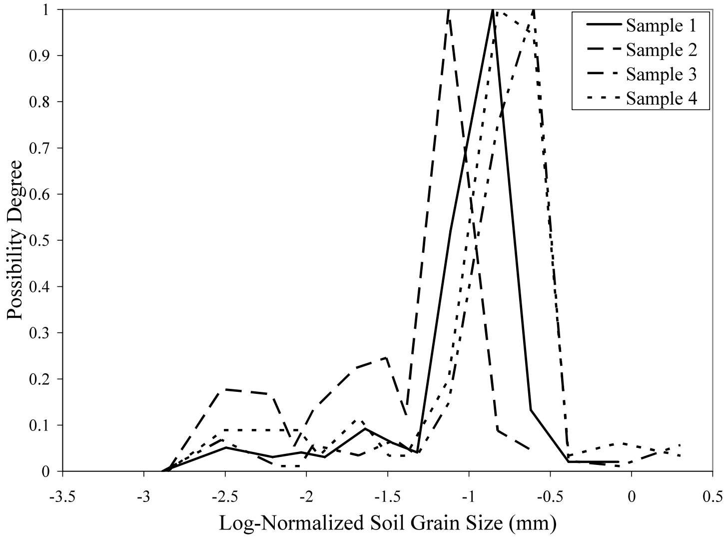


Fig. 3 Possibility distributions of the four samples used in the application, below, transformed from the grain size distribution data

The possibility value for a soil grain size represents the degree to which the soil grain size is possible, given some background information. Possibility distributions are generally used to characterize a person's knowledge regarding a matter about which there is no available evidence from which to make a certain conclusion. Thus, while fuzzy sets characterize certain but imprecise information, possibility distributions define uncertain information. Total uncertainty (or ignorance) would be represented by a possibility distribution where every value along the abscissa has a possibility of one (everything is possible and nothing can be discounted). Interpretation is not of significance in this algorithm, as only the structure of the PD is important here. Fuzzy sets and possibility distributions are generally defined on the same scale, and the two differ only in interpretation. As such, they are defined on the same range, since their scales are consistent.

The transformation used is the most obvious and most common PDF-PD transformation. Each probability value is merely divided by the maximum probability value, which satisfies the constraint that a possibility distribution needs to have at least one domain value that is entirely possible, $\pi(x) = 1$ (Dubois and Prade 1988). The results of the transformation of the soil grain PDFs into soil grain PDs for the four soil samples considered in the Application Section are shown in Fig. 3.

Juxtaposing the PD with the soil grain fuzzy sets, it is evident that the former intersects many of the fuzzy sets. Figure 4 provides insight into how the PD for one sample intersects each of the five soil grain fuzzy sets, which, in turn, reveals how each of the rules, described earlier, will operate when estimating hydraulic conductivity. The maximum degree of intersection with each fuzzy set is important. If one visualizes the intersection of the PD and a certain fuzzy set, the maximum degree

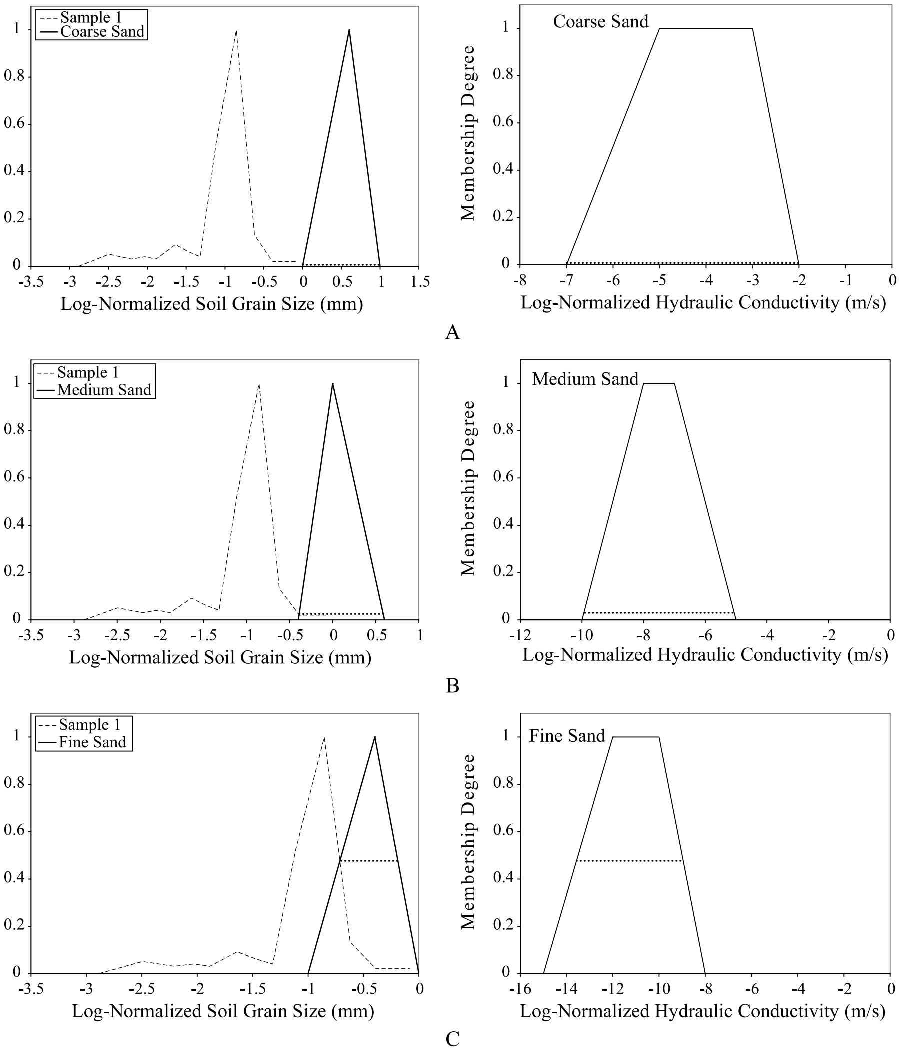
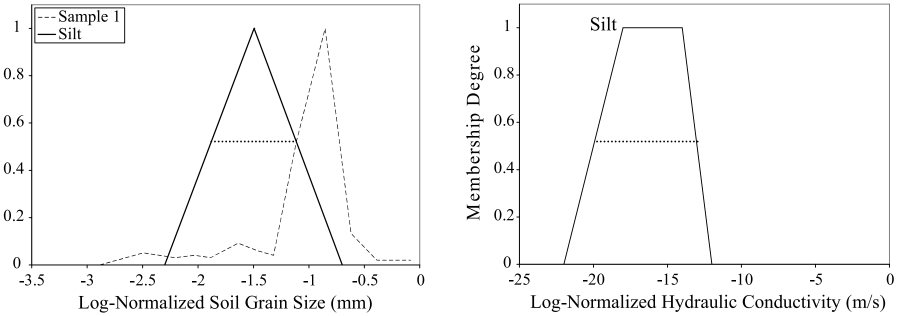


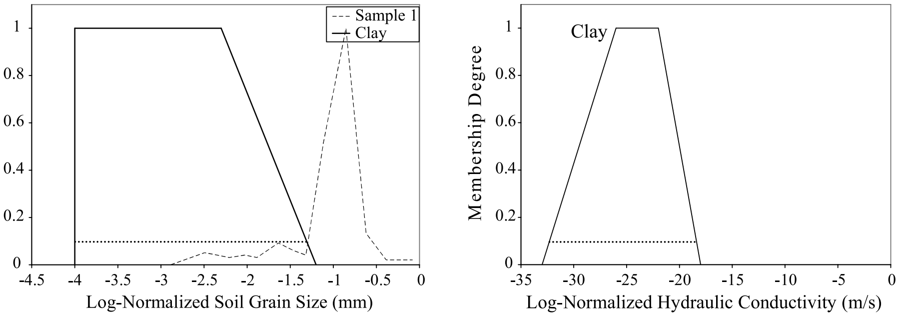
Fig. 4 The possibility distribution representing the grain size distribution data for sample 1 is intersected with the fuzzy sets **A**, coarse sand; **B**, medium sand; **C**, fine sand; **D**, silt; and **E**, clay (*left-most figures*). The maximum degree of intersection between the PD and each fuzzy set defines where each fuzzy set is cut, and as well as the corresponding cut hydraulic conductivity fuzzy sets (areas below the dotted line in right-most figures)

of membership of the intersection is the degree to which the corresponding rule will operate. This is further explained in the following section.

Looking at the left-hand graphs in Fig. 4, it seems rather straightforward to qualitatively describe the soil. For the example given, this soil sample maximally intersects the Fine Sand and Silt fuzzy sets more so than the others. Knowing this, one might describe the soil sample as fine sandy silt, with trace amounts of clay and medium sand. Consider now an alternative strategy to describing this soil. Often soil samples



D



E

Fig. 4 (*Continued*)

are described via visual inspection by the groundwater professional resident during drilling. A soil type identification is recorded in a borehole log. This information is helpful provided that it can be presented in a manageable framework. Fuzzy sets offer a unique approach to exploiting these subjective descriptions, as they generally define linguistic variables. The geologist's description of the soil sample is a linguistic variable of sorts. It may be converted into a fuzzy set by the person preparing the log or an appropriate expert. For instance, if a soil sample is described as being a silty fine sand, a fuzzy set resembling that in Fig. 5 (situated between the Silt and Fine Sand fuzzy sets) may be defined by an expert using a basic understanding of membership functions. In general, one may feel as though the soil sample: (1) is best represented by a single soil grain size (full membership), (2) may be characterized by other grain size values to a lesser degree (membership decreases from unity), and (3) would not be characterized by other soil grain sizes (zero membership). Thus, a geologist, knowing the predefined Fine Sand and Silt fuzzy sets in Fig. 1, could feel comfortable characterizing a soil sample deemed to be a silty fine sand by sketching the triangular membership function labeled Sample 1 (Fig. 5).

Training and experience would be needed to avoid arbitrary definitions. Therefore, experts rather than novices should create these fuzzy sets from qualitative descriptions. In this example, most experts would agree that a sample described as a silty fine sand would be comprised predominantly of fine sand, although the influence of silt should shift the resulting fuzzy set slightly towards the Silt set. It is important to note that an expert-defined fuzzy number must have at least one value with full mem-

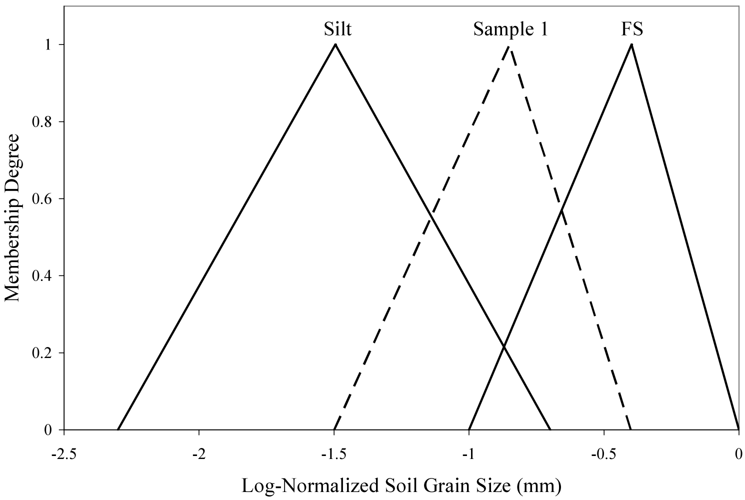


Fig. 5 An expert's interpretation of borehole log data for sample 1 (Table 1), shown in reference to fuzzy sets Silt and Fine Sand (*FS*)

bership and any horizontal cuts through the fuzzy set must create closed and finite intervals along the domain. For example, the Silt fuzzy set in Fig. 1 has at least one log grain size value (-1.5) with full membership, and a horizontal cut at the zero degree membership results in the closed and finite interval $[-2.7, -0.3]$. A horizontal cut through the fuzzy set at any other membership value would produce subsets of this interval. The manner in which this borehole log information is combined with the grain size distribution data is explained in the next section.

Reasoning

The approach taken is most common in interpreting fuzzy IF–THEN rules, using the 'conjunction-based' model. These models are desirable because they treat a rule as defining a fuzzy point in a fuzzy relationship between two or more variables. A collection of these rules becomes a series of fuzzy points that define a complete imprecise relationship, such that any feasible input value will produce a reasonable output value. This is the precise intention of this algorithm, where the two variables are soil grain size and hydraulic conductivity.

Once the possibility distribution is intersected with each fuzzy soil type, the degree to which each rule operates is determined by the maximum point of intersection between the possibility distribution and each predefined fuzzy set. Since the rules relate soil grain size to hydraulic conductivity, the membership of the horizontal 'cuts' of the soil grain fuzzy sets direct the 'cuts' of the hydraulic conductivity fuzzy sets. For example, if the PD intersects the Silt soil fuzzy set to a degree of 0.6, then the Silt hydraulic conductivity fuzzy set is cut at degree 0.6 membership. This membership at which maximum intersection occurs between the PD and a soil type fuzzy set

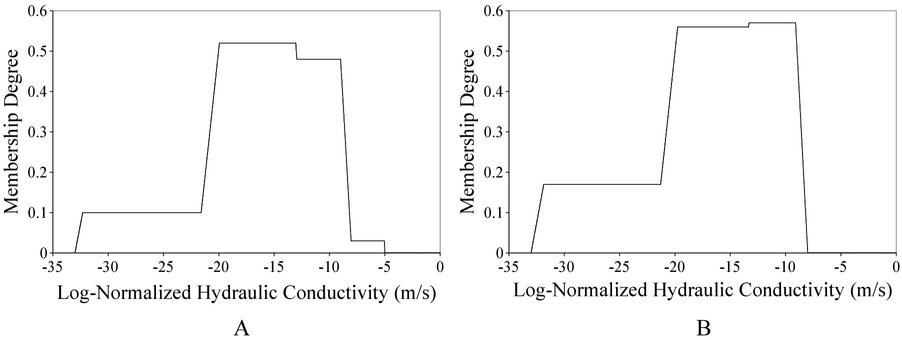


Fig. 6 Fuzzy sets resulting from inputting **A**, the grain size distribution PD, and **B**, the fuzzy characterization of the borehole log data, into each fuzzy rule and aggregating the resulting cut hydraulic conductivity fuzzy sets using the union operator

dictates the degree to which the corresponding rule pertains to the soil sample represented by the PD. The degree of intersection between the PD and soil grain fuzzy sets guides the degree to which the corresponding hydraulic conductivity fuzzy set operates should be intuitive. If one views the intersection as the degree to which the soil type applies to the given soil sample, then the hydraulic conductivity fuzzy set should only apply, or contribute, to the soil sample's true hydraulic conductivity value to the same degree (Fig. 4).

Each rule will operate to a degree between zero and unity, depending upon how strongly each soil grain fuzzy set matches the soil sample's PD. This results in five hydraulic conductivity fuzzy sets, whose peaks are cut off at various degrees (right-hand side of Fig. 4). These abridged fuzzy sets are combined to produce a fuzzy hydraulic conductivity value for the soil sample, using the union operator. The union operator works by assigning to each hydraulic conductivity value along the abscissa the maximum membership degree from the five cut sets for that hydraulic conductivity value (Klir and Yuan 1995). The formula is

$$K_U(k) = \max_i \{K_i(k)\} \tag{1}$$

for all $k, i = 1, \dots, 5$. K_U represents the union of the five cut hydraulic conductivity fuzzy sets, K_i, resulting from the rules, and k are the hydraulic conductivity values along the abscissa over which this operation is performed. For example, given the five rules relating the soil types and hydraulic conductivities in Fig. 1, the union operator in (1) produces the fuzzy hydraulic conductivity value in Fig. 6A based on the cut hydraulic conductivity fuzzy sets on the right-hand side of Fig. 4.

The fuzzy input from the borehole log data provided by the groundwater professional resident at the drilling site will produce another fuzzy hydraulic conductivity result (Fig. 6B). The soil grain fuzzy set is then transposed over the soil type fuzzy sets, as had been done with the PD in Fig. 4. The two hydraulic conductivity fuzzy results need to be combined to produce a single result. This is accomplished through a weighted mean, which aggregates the two fuzzy sets to produce a single representative fuzzy value. The weighted mean equation, acting on fuzzy sets, is a function on

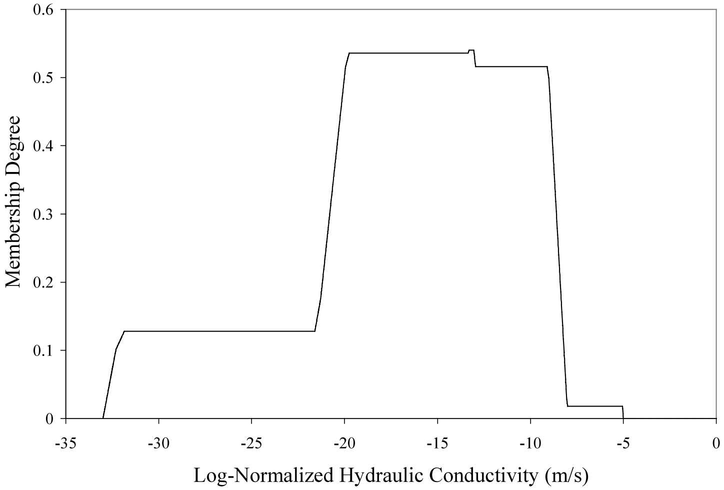


Fig. 7 Aggregate fuzzy hydraulic conductivity estimate for sample 1, from the combination of the individual fuzzy results from grain size data and borehole log data (Fig. 6)

a domain element's (k) membership values ($\mu_K(k)$) from different fuzzy sets (K_i)

$$\mu_K(k) = \sum_i \sigma_i \cdot \mu_{K_i}(k), \tag{2}$$

for all hydraulic conductivity values k along the domain, where σ_i represents the weight given to the ith fuzzy set. This weight measures how influential each information source (borehole log and soil grain distribution) will be to the final answer. These weights must have a value of between zero and unity, and they must sum to one.

Intuitively, the soil grain distribution data, an objective characterization of the soil sample should be weighted more heavily than the subjective characterization from the borehole log data. However, the latter should not be discounted completely, as it is viable data. As such, reasonable weights would be the following

$$\sigma_{\text{Borehole Log}} = 0.4,$$

$$\sigma_{\text{Grain Size Distribution}} = 0.6.$$

Applying these weights and (2), a fuzzy hydraulic conductivity value that represents both sources of hydraulic conductivity values, a weighted average of the fuzzy sets in Fig. 6, is generated (Fig. 7). This fuzzy result in Fig. 7 provides a vague understanding of the estimated hydraulic conductivity value for the given soil sample, and is considered a fuzzy estimate. Its use in geostatistical estimation is not obvious and requires non-traditional estimation procedures (Bardossy et al. 1989, 1990; Diamond 1989).

Because ordinary and simple kriging are considerably more prevalent and popular approaches to estimation than those founded in fuzzy logic, it is more convenient if the hydraulic conductivity value that results from the above reasoning mechanism is a crisp value. The fuzzy estimate must be 'defuzzified' or reduced to a crisp (non-fuzzy) value. The most common defuzzification method is called the Center of Gravity (COG) method. The fuzzy set is considered to be a uniformly thin sheet of metal, and the horizontal component of the center of gravity is the defuzzified value. The finite form of the COG equation is

$$COG = \frac{\sum_j \mu_K(k_j) \cdot k_j}{\sum_j \mu_K(k_j)}, \tag{3}$$

where μ_K is the membership function of the hydraulic conductivity fuzzy set, and k_j are the finite number of domain values used in the calculation. In the case of the fuzzy set in Fig. 7, the defuzzified value, or log hydraulic conductivity estimate for this soil sample is -16.85. The exponential applied to this value ($e^{-16.85}$) gives a hydraulic conductivity estimate of 4.8×10^{-8} m/s. The defuzzified value will change with the discretization of the K-axis ($k_j \in K$) in (3).

Application

The reasoning described above was applied to four soil samples from the CIBA-Geigy Superfund site in Toms River, New Jersey, for which both soil grain distribution data and hydraulic conductivity values were available. At this site, many borehole logs were available, a number of which were accompanied by soil grain distribution curves. Hydraulic conductivity measurements were few compared to the number of soil grain distribution curves. Since there were numerous sample locations with soil grain distribution curves but without hydraulic conductivity values, it was an ideal site for the practical application of the method described.

The particle size distributions of the four soil samples for which hydraulic conductivity measurements were available had been determined by both sieve and hydrometer methods, so that the soil grain distribution curves were complete. In cases where hydrometer analysis has not been performed on a soil sample, a geologist can use expert judgment to complete the curve. The soil type and hydraulic conductivity fuzzy sets have been defined, therefore, the fuzzy relationship is already in place. The input, however, needed to be transformed into the appropriate possibility distributions for the inferential reasoning. Each of the four soil grain distribution curves (Fig. 2) was changed into PDF form and then PD form (Fig. 3) using the aforementioned transformations. Each of these inputs activated the five rules in different ways, resulting in a set of four fuzzy hydraulic conductivity output values (Fig. 8).

Similarly, an expert characterized the borehole log descriptions of these soil samples using fuzzy sets (Table 3), simply by applying his understanding of such descriptions and a basic understanding of triangular membership functions. These fuzzy set versions of borehole log data were also fed through the fuzzy rule base to produce another set of four fuzzy hydraulic conductivity estimates (Fig. 9). Corresponding fuzzy hydraulic conductivity estimates (resulting from the grain size distribution curve and borehole log description of each soil) were combined (Fig. 10) using the weighted mean approach, described in the previous section.

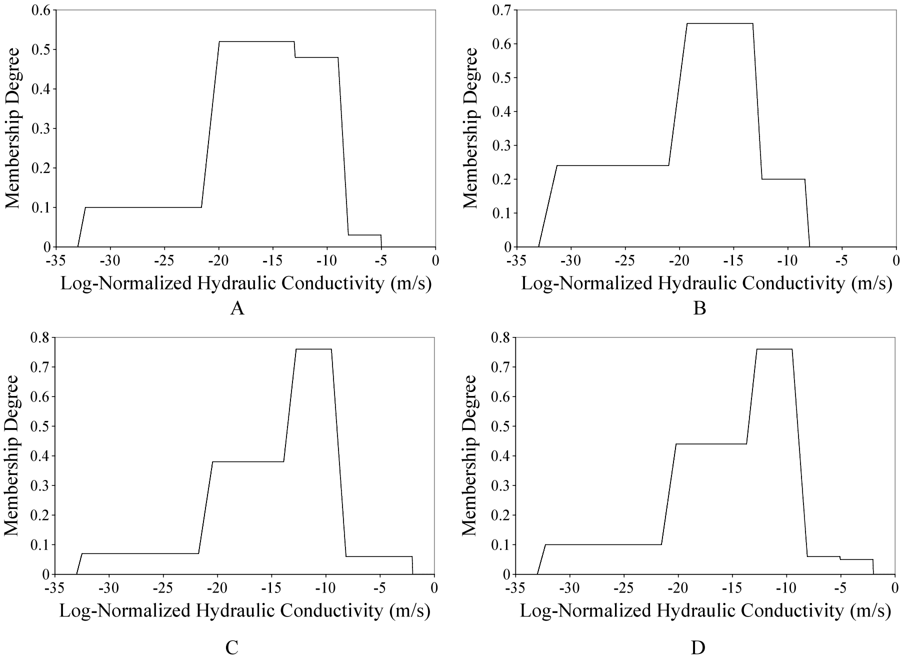


Fig. 8 Fuzzy sets that result from inputting the transformed soil grain distribution from **A**, sample 1; **B**, sample 2; **C**, sample 3; and **D**, sample 4, into the fuzzy rule base

Table 3 Shapes and parameters of expert-provided fuzzy sets interpreted from Borehole log descriptions

Sample	Fuzzy set shape	Parameters (log grain size)
1	Triangular	$[-1.5, -0.85, -0.4]$
2	Trapezoidal	$[-2.8, -1.8, -1.5, -0.7]$
3	Trapezoidal	$[-2.3, -1.1, -0.85, 0.2]$
4	Trapezoidal	$[-1.35, -0.7, -0.45, 0]$

Table 4 Hydraulic conductivity estimates from the combination of two types of soil data and the corresponding measurements from pump tests

Sample	Aggregate estimate (m/s)	Measurement (m/s)
1	4.8×10^{-8}	2.5×10^{-7}
2	2.3×10^{-9}	1.6×10^{-9}
3	7.8×10^{-8}	1.8×10^{-6}
4	3.6×10^{-7}	1.6×10^{-6}

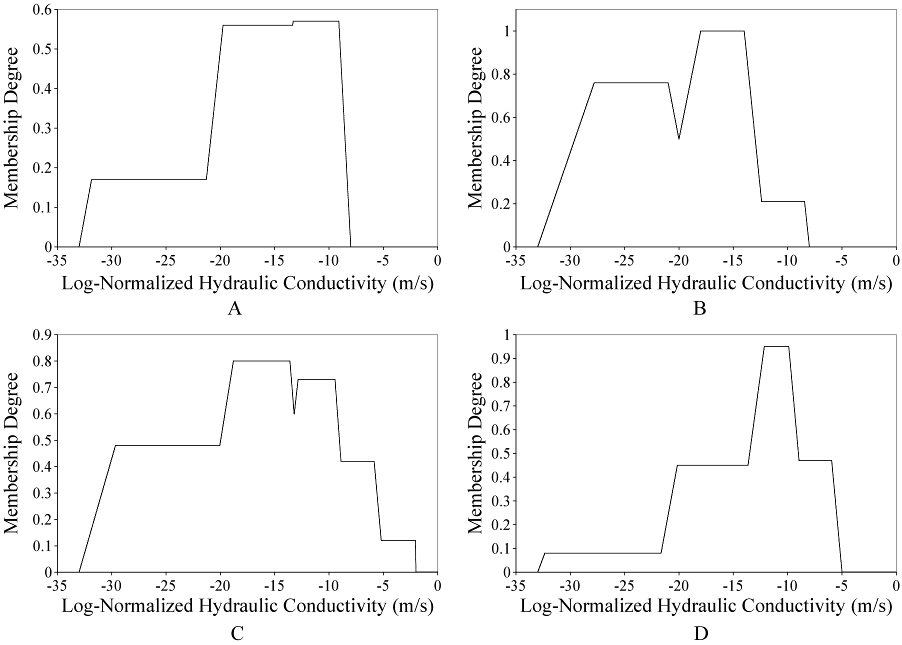


Fig. 9 Fuzzy sets that result from inputting the fuzzy characterizations (Table 3) of the borehole log data from **A**, sample 1; **B**, sample 2; **C**, sample 3; and **D**, sample 4, into the fuzzy rule base

Table 5 Hydraulic conductivity estimates from Borehole log and grain distribution data

Sample	Borehole log	Grain distribution	Measurement
1	3.2×10^{-8}	6.5×10^{-8}	2.5×10^{-7}
2	1.1×10^{-9}	5.2×10^{-9}	1.6×10^{-9}
3	2.6×10^{-8}	3.5×10^{-7}	1.8×10^{-6}
4	7.5×10^{-7}	1.9×10^{-7}	1.6×10^{-6}

Results and Conclusions

Crisp estimates were obtained by defuzzification of the fuzzy estimates (Table 4). The table also gives the measured hydraulic conductivity values for the samples. All estimated values were within two orders of magnitude of the corresponding measured value, which is considered reasonable. Hydraulic conductivity values may vary by up to five orders of magnitude if there are small differences in the percentages of the fine fractions (Koltermann and Gorelick 1995).

Also relevant is the interpretability of the results. With the fuzzy reasoning process, one can understand how the estimates are calculated. For instance, it is evident that the second hydraulic conductivity estimate in Table 5 is lower-valued than the other three estimates because the aggregate estimated hydraulic conductivity fuzzy set in Fig. 10B has high memberships for lower hydraulic conductivity values.

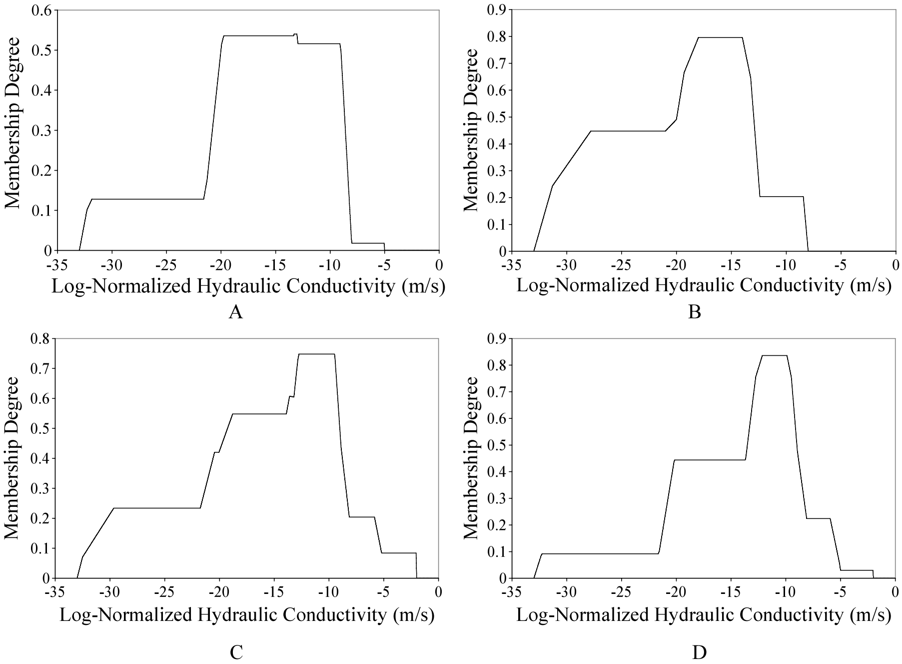


Fig. 10 Fuzzy sets that result from the combination of the fuzzy sets from borehole log data and grain size distribution data for **A**, sample 1; **B**, sample 2; **C**, sample 3; and **D**, sample 4

When this fuzzy set is defuzzified, the lower values of hydraulic conductivity are weighted more heavily than in the other samples' estimated hydraulic conductivity fuzzy sets in Fig. 10. The higher memberships for lower-valued hydraulic conductivity values in this fuzzy set directly result from the significant overlapping of the grain size distribution and borehole log description fuzzy sets with the Clay soil type fuzzy set during inference.

In many applications, borehole log data is not only more prevalent than grain size data, but it is often the only type of data available. Considering this, an investigation was made into the reliability of estimates resulting from just the expert fuzzy characterization of borehole log descriptions of the soil samples. For completeness, the scenario where only grain size distribution data was available was also evaluated. The results of these investigations are summarized in Table 5. In general, for this set of data, it is difficult to distinguish which form of data is more accurate. However, estimates stemming from both forms of data are still within two orders of magnitude of the measurements. Both types of data are not required to produce acceptable estimates. It has been shown that the use of only one data type in the fuzzy reasoning process can still lead to relatively accurate results. Ideally, for the generation of realizations in groundwater modeling, measures of uncertainty should accompany hydraulic conductivity estimates. However, the issue of uncertainty in the results is not addressed, and definitive answers regarding the uncertainty of the fuzzy estimates are unavailable. A thorough description of the uncertainty of fuzzy sets would be requisite and is beyond the scope of this research.

The soundness of this approach to hydraulic conductivity estimation relied greatly upon the definition of the original soil type and hydraulic conductivity fuzzy sets that were used for reasoning. More accurate estimates could be produced had more data from the CIBA-Geigy Superfund site been available. With more data, the soil type and hydraulic conductivity fuzzy sets could have been constructed to respect the site data, thus producing better hydraulic conductivity estimates for that particular site. This would be true for any site with sufficient soil data. Since experts in hydrogeology defined these fuzzy sets, the reasoning process is considered sound and generally applicable to any site with reasonable accuracy.

References

Alyamani MS, Sen Z (1993) Determination of hydraulic conductivity from complete grain-size distribution curves. Ground Water 31(4):551–555
Bardossy A, Bogardi I, Kelly WE (1989) Geostatistics utilizing imprecise (fuzzy) information. Fuzzy Sets Syst 31(3):311–328
Bardossy A, Bogardi I, Kelly WE (1990) Kriging with imprecise (fuzzy) variograms. Math Geol 22(1):63–94
Brown W, Groves D, Gedeon T (2003) Use of fuzzy membership input layers to combine subjective geological knowledge and empirical data in a neural network method for mineral-potential mapping. Nat Resour Res 12(3):183–200
Cheng Q, Agterberg FP (1999) Fuzzy weights of evidence method and its application in mineral potential mapping. Nat Resour Res 8(1):27–35
Demmico R (2004) Fuzzy logic and earth science: an overview. In: Demmico R, Klir G (eds) Fuzzy logic in geology. Elsevier, Amsterdam, pp 63–102
Diamond P (1989) Fuzzy kriging. Fuzzy Sets Syst 33(3):315–332
Domenico PA, Schwartz FW (1990) Physical and chemical hydrogeology. Wiley, New York, 824 p
Dubois D, Prade H (1988) Possibility theory: an approach to computerized processing of uncertainty. Plenum, New York, 263 p
Fang JH, Chen HC (1997) Fuzzy modeling and the prediction of porosity and permeability from the compositional and textural attributes of sandstone. J Petrol Geol 20(2):185–204
Hsieh B, Lewis C, Lin Z (2005) Lithology identification of aquifers from geophysical well logs and fuzzy logic analysis: Shui-Lin area, Taiwan. Comput Geosci 31:263–275
Klir G, Yuan B (1995) Fuzzy sets and fuzzy logic. Prentice-Hall, Upper Saddle River, 574 p
Koltermann CE, Gorelick SM (1995) Fractional packing model for hydraulic conductivity derived from sediment mixtures. Water Resour Res 31(12):3283–3297
Porwal A, Carranza EJM, Hale M (2003) Knowledge-driven and data-driven fuzzy models for predictive mineral potential mapping. Nat Resour Res 12(1):1–25
Oussalah M (2000) On the probability/possibility transformations: a comparative analysis. Int J Gen Syst 29(5):671–718
Selley RC (1985) Elements of petroleum geology. WH Freeman, New York, pp 234–239
Sen Z (1992) New statistical grain-size method for evaluating the hydraulic conductivity of sand aquifers—comment. J Hydrol 130:399–403
Uma KO, Egboka BCE, Onuoha KM (1989) New statistical grain-size method for evaluating the hydraulic conductivity of sandy aquifers. J Hydrol 108:343–366
Zadeh L (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst 1(1):3–28