Journal of Hydrology (2008) **354**, 149–159

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jhydrol

Kalman filter updating of possibilistic hydraulic conductivity

James Ross [a,*], Metin Ozbek [b], George F. Pinder [a]

[a] *Research Center for Groundwater Remediation Design, Environmental Engineering Program, College of Engineering and Mathematical Sciences, University of Vermont, 33 Colchester Avenue, Burlington, VT 05405, USA*
[b] *Environ International Corporation, Princeton, New Jersey, USA*

Received 16 July 2007; received in revised form 14 February 2008; accepted 5 March 2008

KEYWORDS
Fuzzy sets;
Hydrogeology;
Kalman filter;
Geostatistics;
Boreholes;
Hydraulic conductivity

Summary Preferential sampling regimes too often leave hydrogeological site characterization in need of hydraulic conductivity data. Furthermore, faith in a kriged hydraulic conductivity field (*K*-field) wanes when viable data are few. A straightforward solution to this problem is the application of expert knowledge to the identification of hydraulic conductivity at specified locations. However, the nature of this supplementary data suggests that its use should be limited to the initial estimation of a *K*-field. Available hydraulic conductivity measurements are used to update the initial estimate of the *K*-field in a fuzzy Kalman filtering procedure. The results are fuzzy and can be defuzzified into traditional crisp estimates. An application of these protocols to the CIBA-Geigy site in New Jersey is provided to illustrate the ease of the procedure and the impact of viable expert input.

Introduction

The characterization and propagation of uncertainty in hydrogeological applications is important in groundwater flow system modeling. Though inappropriate groundwater model selection is a significant source of model uncertainty, so to is the uncertainty surrounding hydraulic conductivity (*K*) that is introduced primarily via measurement and interpolation.

Traditionally, uncertainty in hydrogeology has been framed in probabilistic terms. Nevertheless, fuzzy set and possibility theories, as well as fuzzy logic, have recently provided more intuitive means of uncertainty characterization (Zadeh, 1965, 1978). In fact, the introduction of such frameworks has revealed new sources and enhanced upon traditional sources from which viable hydrogeological data may be extracted, including the facilitation of direct expert knowledge extraction (Bardossy et al., 1989; Piotrowski, 1996). In this latter approach, an expert opines upon the value of hydraulic conductivity at specific locations, utilizing only an implicit understanding of the particular site's characteristics drawn from available field information,

* Corresponding author. Tel.: +1 802 656 7789; fax: +1 802 656 8446.
E-mail address: jlross@uvm.edu (J. Ross).

0022-1694/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2008.03.005

education and experience, information which cannot otherwise be explicitly and mathematically defined. The use of possibility theory to accomplish this task is considerably more flexible and intuitive than permitting the petitioned expert to frame his/her knowledge in a probabilistic framework.

The hydrological site characterization process using non-traditional data forms has two distinct issues. Hydraulic conductivity estimates determined through the site characterization process using non-traditional data forms can incorporate soil physical measurements as well as expert knowledge. These different data sources possess different degrees of integrity. While the physically-based methods are reliant upon relatively precise hard data, expert knowledge is a very imprecise data source, perhaps implying a lower degree of quality.

This paper suggests an approach to hydrogeological site characterization that not only accommodates the use of imprecise data, but also sets guidelines governing how data from physical measurements and expert knowledge can be properly utilized. Specifically, the authors propose that (1) hydraulic conductivity measurements resulting from direct expert input be interpolated using fuzzy kriging (Bardossy et al., 1989; Bardossy et al., 1990a,b) to produce an initial estimate of the *K*-field; (2) a static fuzzy Kalman filter (Ross et al., 2006) be used to update this first estimate with available measurement data to produce the final *K*-field for a site. This fuzzy Kalman filtering approach is considered to be an alternative to Bayesian fuzzy kriging (Bandemer and Gebhardt, 2000). In the latter, experts are required to frame their knowledge as expectation values and covariance matrices of trend parameters, rather than membership functions and possibility distributions representing uncertain hydraulic conductivity values (Fig. 1).

In the next section, the application of possibility theory to hydrogeological site characterization is discussed. Section 'Interpolating fuzzy measurements' highlights the major implications of fuzzy kriging. The fuzzy Kalman filter is introduced in Section 'Fuzzy Kalman filtering', followed by a field application of the algorithm presented herein, in Section 'Field application'. Introductions to pertinent tools of fuzzy set and possibility theories are provided in Appendix.

While fuzzy sets and fuzzy numbers are inherent in the groundwater flow modeling process, there has been no discussion, as far as the authors are aware, regarding the best

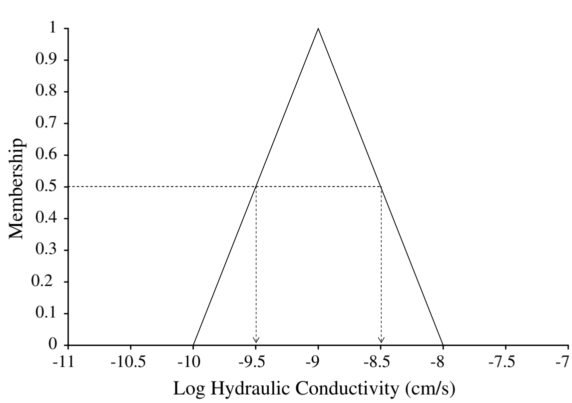

Figure 1 This fuzzy set is a possible characterization of the expert's notion of *low* log hydraulic conductivity (cm/s). The 0.5 level α-cut produces the interval $[-9.5, -8.5]$.

use and propagation of fuzzy data relative to the data source. Such a discussion follows.

Expert judgement

As noted, data sources in hydrogeological site characterization can be partitioned into two groups: physical measurement and expert knowledge. The former data source produces hydraulic conductivity values based upon direct measurements of that variable or through translations of secondary variable measurements, while the latter dictates that an expert armed with geological knowledge of a given site, can identify, via his/her understanding, the hydraulic conductivity values at various spatial locations. Hydraulic conductivity data derived from expert knowledge may be considered somewhat less reliable than physics-based data and perhaps should be treated as such.

Traditionally, hydraulic conductivity measurements are made via indirect methods. Indirect methods include both inverse approaches, where head measurements lead to hydraulic conductivity values, and functional transformations of secondary data, most often derived from soil sample analyses (Alyamani and Sen, 1993; Uma et al., 1989; Sen, 1992; Masch and Denny, 1966). Estimated *K*-fields rely upon the number and location of hydraulic conductivity measurements made throughout the site under study. However, the spatial distribution of measurement locations is often not conducive to accurate spatial estimation regionally, and measurements are often too few for the accurate estimation of a *K*-field (Piotrowski, 1996).

Numerous approaches have been devised to correct some of the deficiencies resulting from sparse sampling (Gotway and Hartford, 1996; Odeh et al., 1999; Asli and Marcotte, 1995). Of these, cokriging (Issaks and Srivastava, 1989) is most often encountered in practice. Though the use of secondary data in this context improves an understanding of the spatial correlation of the site, it requires the modeling of multiple semivariograms and may not positively impact the quality of the values being interpolated. Moreover, in cases where the primary variable is hydraulic conductivity (as it is in this research) and secondary data results from soil analysis, certain information must either be discarded or simplified. For instance, grain size distributions may need to be reduced to a single representative value and qualitative data resulting from borehole investigations must be transformed into crisp soil classes. As such, the quality of primary variable estimates is arguably more positively impacted by the availability of more primary data rather than secondary data.

A much more straightforward solution to the problem of too few measurement data is the application of expert intuition to the identification of hydraulic conductivity values at various locations throughout the spatial domain. In this work, an expert is any agent familiar with the hydrogeological aspects of the site of interest. It is assumed here that such an appropriate expert could roughly estimate, without sampling, the approximate hydraulic conductivity value or representative soil grain size at any number of locations throughout the spatial domain. Expert knowledge applied toward this end is arguably quite valuable. With an assumed minimal loss of accuracy, a well-informed expert can