ELSEVIER

Journal of Hazardous Materials 110 (2004) 151–160

Journal of Hazardous Materials

www.elsevier.com/locate/jhazmat

Embedding expert knowledge in a decision model: evaluating natural attenuation at TCE sites

Neil A. Stiber[a], Marina Pantazidou[b,*], Mitchell J. Small[c]

[a] USEPA Office of Research and Development, Office of Science Policy, Washington, DC 20460, USA
[b] Faculty of Civil Engineering, National Technical University of Athens, Iroon Polytechniou 9, Zografou 15780, Greece
[c] Departments of Civil and Environmental Engineering and Engineering and Public Policy,
Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

This paper describes a generalized methodology that enables the translation of expert knowledge about any complex process involved in a remedial decision into easy-to-use decision tools. The methodology is applied to evaluate reductive dechlorination as a remedial possibility at sites contaminated with trichloroethene (TCE), building on an existing protocol/scoring system put forth by the US Air Force and the US EPA. An alternate scoring system is proposed, which has two major advantages, namely that it: (i) attributes relative weights to findings based on expert beliefs; and (ii) systematically includes negative weights for negative findings. The ability of the proposed scoring system to assess the bioattenuation potential of TCE is demonstrated using data from extensively studied sites.

Keywords: Environmental decision making; Expert knowledge decision models; Remedial technology screening tools; TCE site remediation; TCE biotransformation

1. Introduction

The evaluation of remedial options at hazardous waste sites requires the quantification of several uncertain physical, chemical, and biological processes. To reduce the inherent complexity, remedial decisions must use models to account for these processes simplistically or ignore them altogether. The goal of this paper is to restore some balance between the abstraction of a model and the complexity of the real physical system. A methodology is presented for the development of decision tools that embody the knowledge base of expert systems. This methodology is demonstrated by evaluating natural attenuation as a remedial option at sites where ground water is contaminated with trichloroethene (TCE).

A widely used solvent, TCE is the most frequently detected organic compound at hazardous waste sites in the USA [1] and a recalcitrant contaminant in ground water [2,3]. Compared to other remedial alternatives, when successful, natural attenuation achieves significant cost savings, requiring only long-term monitoring beyond site characterization. Natural attenuation of ground-water contaminants encompasses processes that dilute them, remove them from ground water by transferring them to other media (i.e., soil gas, soil particles), or transform them to simpler compounds. Among these processes, chemical or biological transformation is the most protective of the environment, since it reduces not only concentrations but also the total contaminant mass present in the subsurface.

For certain organic contaminants, such as fuel constituents, observations from most sites have confirmed that biological transformation mediated by naturally occurring microorganisms proceeds at adequately fast reaction rates. In contrast, the biotransformation of TCE is currently at a state of emerging scientific understanding. To make up for gaps in scientific knowledge, a significant site characterization and data analysis effort may be required to substantiate the effectiveness of natural attenuation for a TCE site [4]. The proposed methodology and the corresponding scoring system are intended to support these efforts with expert knowledge.

Due to its low cost, natural attenuation is being selected as a remedy at hazardous waste sites with increasing frequency.

---

[*] Corresponding author. Tel.: +30-210-772-3745; fax: +30-210-772-3428.
E-mail address: mpanta@central.ntua.gr (M. Pantazidou).

0304-3894/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhazmat.2004.02.048

Its popularity prompted the codification of recommended procedures for site evaluation in several natural attenuation protocols [4]. Most relevant to the present work is the widely used protocol developed by the US Air Force and the US EPA [5,6] for evaluating natural attenuation of chlorinated solvents. Meanwhile, the increasing numbers of the natural attenuation candidate sites brought about questions concerning the limits of this approach [7,8]. To address these questions, a committee of the National Research Council (NRC) was called to provide a critical review of the scientific basis of natural attenuation and of the criteria used to evaluate its remedial potential. Among its salient recommendations is that protocols should be peer-reviewed. Moreover, the committee recommends that sites should be evaluated by relying on conceptual models and the "footprints" of natural attenuation, i.e., the consequences of contaminant transformation (e.g., presence of daughter products) and not on scoring systems alone [4]. The methodology for developing expert and scoring systems as proposed by the authors embodies several of the main technical recommendations of the NRC report [4], as discussed in more detail in the following sections.

2. Developing an expert system

This section discusses the steps involved in building an expert system that uses a Bayesian Belief Network [9] to conceptualize an event of interest and estimate its probability of occurrence. A Bayesian Belief Network is a causative model capable of describing probabilistic relationships between the main factors influencing the outcome of interest and its consequences. The probabilities describing these relationships were obtained for the model developed herein through structured expert elicitations. The section also describes, in tandem, selected features of the developed expert system in order to: (i) exemplify the kinds of choices facing the expert system developer; and (ii) provide the necessary context for the subsequent discussion of the corresponding scoring system. A detailed discussion of the expert system has been given elsewhere [10]. It should be stressed that the step-wise presentation does not imply that the process of model development is linear. On the contrary, it is highly iterative and incorporates feedback on the model structure from the experts.

The proposed methodology for the development of a decision tool is summarized schematically in Fig. 1. It starts with stating the event of interest and identifying its key influencing factors and major consequences; these are numbered 1–3 in figure. It should be noted that Fig. 1 explicitly differentiates between the physical system and its simplification, i.e., the causative model. The model must achieve a balance between being comprehensive, in order to be as faithful as possible to the known science, and being simple, in order to facilitate the probabilistic assessment of its causal relationships. The causative model by necessity ignores relevant factors and consequences that have yet to be identified; these unknown influencing factors and consequences are placed in Fig. 1 outside the box that circumscribes the causative model. The model also does not account for, by choice, relevant influences and consequences that are of mi-

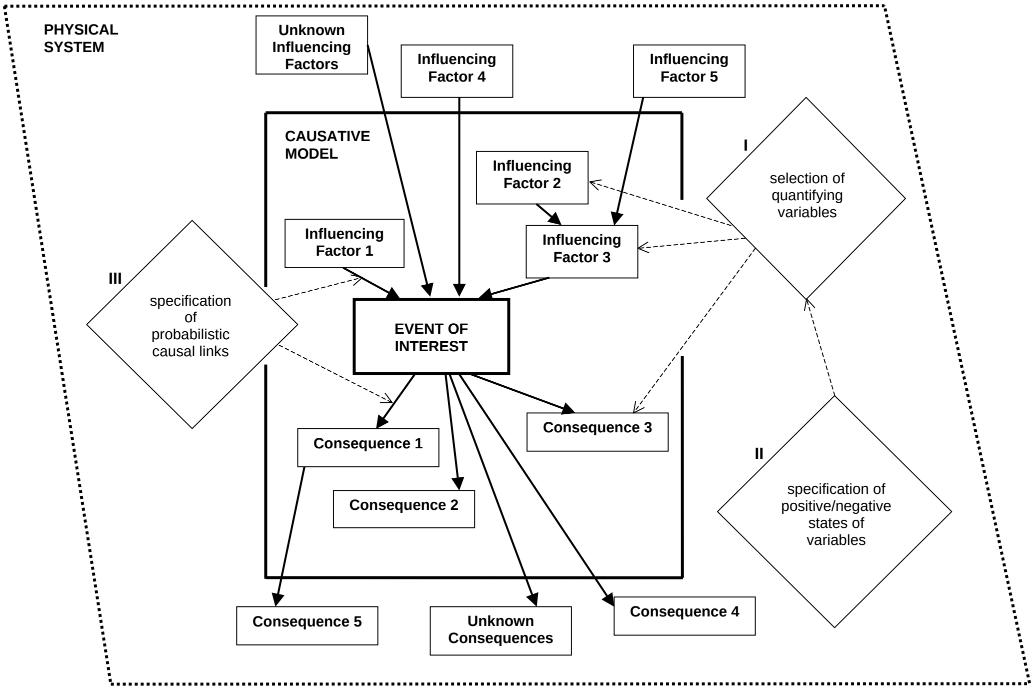

Fig. 1. Schematic of decision tool development. A causative model simplifies the real physical system. The diamonds describe the information needed to fully quantify the model.

nor or uncertain importance, or would significantly complicate the model; these are represented schematically in Fig. 1 by influences and consequences 4 and 5.

For the expert system developed, the simplicity requirement dictated a narrowing of focus on biotransformation, the most protective of the natural attenuation processes. The biotransformation of TCE has yet, as already mentioned, to be completely understood. Nevertheless, it is well established that reductive dechlorination is the most important process for the natural attenuation of the more highly chlorinated solvents [4,6]. It should be noted that biotransformation of TCE is possible under aerobic conditions as well, but at significantly slower pace than in anaerobic environments. Hence, the event of interest is stated as “anaerobic degradation by reductive dechlorination is occurring.” The chosen event statement renders “reducing conditions” an obvious choice for an influencing set of factors. The availability of “electron donors,” necessary for reduction reactions, and a few key ground-water quality indicators, grouped for simplicity under “environmental conditions,” complete the search for the main influencing factors for the event of interest. The daughter and end products of the transformation of TCE, as well as of CO_2 and other carbon sources, represent varying indicators confirming that reductive dechlorination is under way. In other words, these compounds are the major consequences in the model. At the end of this first step, the coarse features of the causative model that describes the event of interest are in place.

Next, the model needs to be finalized with the choice of the specific variables that will quantify influences and consequences. When developing a decision model for a remediation project, these variables become the key types of evidence that need to be measured at the site. Hence, both science-based and practice-related criteria affect the model variable choices at this second step. The Air Force/EPA protocol [5,6] for evaluating natural attenuation at chlorinated solvent sites provided the basic guidelines for choosing the types of corroborating evidence for the reductive dechlorination of TCE. Several modifications were made in order to reduce the number of model variables, by omitting some and aggregating others, and accommodate practical realities (e.g., favoring one measurement over another less frequently collected during site investigations).

Fig. 2 provides an overview of the developed model. Arrows represent causative links between the nodes they connect. The central node is occupied by the event of interest, “anaerobic degradation by reductive dechlorination.” The three major sets of influences for the event of interest, which appear in the nodes “reducing conditions,” “electron donors,” and “environmental conditions,” are quantified by the variables in the nine nodes near the top end of figure. These variables will be referred to herein as pre-conditions. They include terminal electron accepting process (TEAP), hydrogen (H_2), and oxidation reduction potential (ORP), as alternate indicators of a reducing environment. Dissolved organic carbon (DOC) was selected as a “catch-all” indicator of the availability of electron donors. The addition of more specific pre-conditions was also deemed appropriate, since frequently detected TCE co-contaminants are known to serve as electron donor sources. Hence, total petroleum hydrocarbons (TPH) and benzene, toluene, ethylbenzene and xylenes (BTEX) were included in the model, anticipating existing measurements at many mixed-waste sites. Finally, three common ground-water quality measurements, namely, temperature, pH, and oxygen, form the environmental conditions set of influences.

It must be mentioned that the root nodes of a causative model are assumed to be independent. Hence, the chosen structure of the model implies that all nine pre-conditions except hydrogen (which is not a root node since it depends

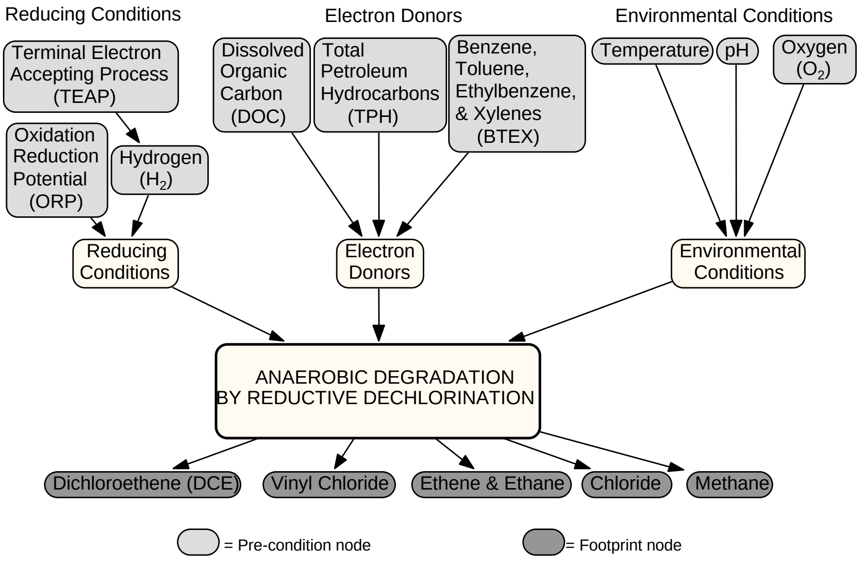

Fig. 2. Causative model for the reductive dechlorination of TCE in ground water.

on TEAP) are independent of each other. The obvious fact that this is not valid for all root nodes is another example of the model-building necessary compromises. Accounting for all the interdependences between pre-conditions would make the model too difficult to grasp and quantify. Leaving only the truly independent variables would exclude valuable corroborating evidence. In the case of TEAP and H_2, the interdependence was deemed too strong, so only TEAP features as an independent root variable, being connected with a causative link to H_2. In contrast, interdependence in the electron donor group is left unaccounted for. This model simplification was deemed acceptable, considering that in most cases it will result in a conservative estimate, i.e., lower probability, for the event of interest.

Fig. 2 also shows the five footprint nodes representing the consequences of anticipated reductive reactions. They include five possible products of the reductive dechlorination of TCE: dichloroethene (DCE), vinyl chloride, ethene and its possible product ethane (grouped in one node for simplicity), and chloride. Methane is also included, being a product of the transformation of carbon sources, which serve as electron donors, in highly reduced environments. In summary, a total of 14 variables, or 14 types of evidence, are needed to fully quantify the causal model for the reductive dechlorination of TCE. At this point, the “selection of quantifying variables” step of model development (decision in diamond I of Fig. 1) is complete.

The next task of model development requires for each model variable the specification of its positive or negative states with reference to the event of interest (decision in diamond II of Fig. 1). Extensive applied experience is required for these decisions. The Air Force/EPA protocol once more provided the basis for the specification of the variables of the developed expert system, again with the necessary modifications and additions. To the extent possible, binary specifications were selected for simplicity. When this was not feasible, as in the case of TEAP, a descriptive specification was chosen to offset the difficulty of handling a quaternary variable: TEAP is denitrification, or iron reduction, or sulfate reduction, or methanogenesis (the most positive state). (This choice in essence leaves further specification of each of these four ground-water conditions up to the model user.) Dissolved oxygen has the only ternary specification: $O_2 \leq 0.5\,\text{mg/L}$ (the most positive state), or $0.5\,\text{mg/L} < O_2 < 1\,\text{mg/L}$, or $1\,\text{mg/L} \leq O_2$. All other variables are made binary. One specification is expressed as absence or presence of a constituent: vinyl chloride is detected (the positive state), or not detected. Another combines a descriptive and a numeral part: *cis*-DCE >80% total DCE (the positive state), or not. The remaining specifications compare the variable measurements to cutoff values, e.g., BTEX >0.1 mg/L (the positive state), or not, $5 < \text{pH} < 9$ (the positive state), or not.

In the final step of model development, the causative structure of the model is specified probabilistically (decision in diamond III of Fig. 1). When the strength of the causative links of the model cannot be determined with the known science, it may be necessary to rely on expert opinions to complete this step. For this study, elicitations of 22 experts practicing in the USA were conducted to provide the conditional probabilities for the causative relationships indicated by the arrows in Fig. 2. The experts were also asked to estimate the prior probabilities of the root nodes, for the case of a “generic” TCE site. For example, in the case of the root node for pH, the experts were ask to give a probability for $5 < \text{pH} < 9$ when there is no information about the site’s previous uses or waste management history, i.e., considering a site representative of the total population of all the TCE-contaminated sites in the USA. As indicated earlier, expert knowledge was also used to refine the model structure in an iterative fashion. The probabilities provided by each of the 22 experts were averaged to construct an average expert system. (Evaluating alternative schemes for aggregating expert beliefs is addressed in a forthcoming publication [11].) The complete expert system consists of the causative model and the averaged expert probabilities.

The Netica^{TM} software [12] was used to develop the expert system and evaluate its outcome. This software is affordable (US$ 585 commercial or US$ 285 educational) and readily available (http://www.norsys.com). The outcome of the expert system gives the probability that anaerobic degradation by reductive dechlorination is occurring, expressed symbolically as Pr(ADbRD is occurring). Once the causal relationships of the model are specified, this probability can be computed for any amount of available data, or even with no site-specific data at all. Using only the expert conditional probabilities for the model links and the expert prior probabilities for the root nodes, Netica^{TM} assigns probabilities to the states of each variable and computes the probability that anaerobic degradation by reductive dechlorination is occurring when no site-specific data are available, expressed as Pr(ADbRD is occurring | no evidence). This value is equal to 0.355 and reflects the average expert probability for the aforementioned “generic” TCE site. This low probability is consistent with the NRC estimate that reductive dechlorination will be protective in <25% of the TCE sites [4], in the absence of site-specific data that may increase, or further decrease, the likelihood of success. Once a site-specific measurement becomes available for a variable, its state is considered to be known with certainty. As new measurements of variables become available, they can be incorporated in the corresponding nodes during subsequent model analysis. Bayesian updating will then permit the influence of this new information to flow both with and against the arrows associated with the updated nodes. However, despite its ease of updating, the full expert system is rarely appropriate as a screening tool or as a periodic long-term monitoring tool. Such considerations provided the impetus to develop a scoring system based on expert knowledge.

3. Developing a scoring system

The development of a scoring system requires that the key pieces of evidence be identified and their decision worth be quantified. The former task has been accomplished during expert system development: the key pieces of evidence are the variables of the causative model. The latter task can be achieved again with the aid of the expert system, by performing a sensitivity analysis for all the model variables. In this section, alternative measures of sensitivity are evaluated in order to develop a point scoring system. The proposed scoring system is accompanied with an interpretation screening guide, consisting of threshold scores and corresponding recommendations. The rationale for the screening guide is supported with systematic comparisons of the probability resulting from the expert system to the number of points produced by the scoring system for the same data. The goal for developing a new scoring system was to add to the simplicity afforded by the Air Force/EPA protocol the science base of the expert system. First, by giving each finding the weight attributed to it by experts, the scoring system clearly indicates the individual pieces of information that are highly compelling or precluding. Secondly, by including negative weights for negative findings, the scoring system makes it easier to conclude that a site is inappropriate candidate for reductive dechlorination.

The first step in developing the point scoring system was to assess the importance of each of the 14 variables for the expert system outcome. This was achieved by studying the sensitivity of the system outcome to each type of evidence, as will be explained with the aid of Table 1. Table 1 lists the 14 variables involved in the causative model of the expert system and in the corresponding scoring system. The output of the expert system, the probability that anaerobic degradation by reductive dechlorination is occurring given some state of evidence, will be referred to as: Pr(ADbRD is occurring | evidence). The first two columns in Table 1 give the possible variation of this probability. The table lists in the first row the value of Pr(ADbRD is occurring | no evidence) = 0.355. The remaining rows of the first two columns show the sensitivity of Pr(ADbRD is occurring | evidence) to the potential states of each variable. These maximum and minimum values were obtained by changing the state of the variable of interest, while assuming no site-specific evidence for the other 13 types of evidence and observing the effect on Pr(ADbRD is occurring). For example, the maximum value of Pr(ADbRD is occurring | TEAP measurement) is obtained for the most positive state of TEAP (methanogenesis), while the minimum value corresponds to its most negative state (denitrification). The possible outcomes of the expert system show that, when the only site-specific evidence available is TEAP, Pr(ADbRD is occurring) does not increase significantly compared to the no-evidence case, even when the reducing conditions are most favorable for reductive dechlorination. It is instructive to note the influence of the negative findings, i.e., variables in negative states, especially for some key pieces of evidence, such as dichloroethene: the corresponding minimum value, 0.074, is significantly smaller than that of the no-evidence case, 0.355. From the table, it is also clear that Pr(ADbRD is occurring) is less sensitive to the pre-conditions (the smallest range, less than 0.01, being observed for temperature) compared to the end and daughter products (the widest range, 0.67, corresponding to vinyl chloride).

Table 1
Three measures of importance of site-specific evidence for the outcome of the expert system that evaluates reductive dechlorination at a TCE site

Type of evidence (model variable)	Probability (anaerobic degradation by reductive dechlorination is occurring | evidence)		ΔProbability due to evidence (Eq. (1))		ΔLOR due to evidence (Eq. (2))	
	Maximum	Minimum	Positive finding	Negative finding	Positive finding	Negative finding
None	0.355					
TEAP	0.399	0.298	0.044	−0.057	0.081	−0.112
Hydrogen	0.414	0.293	0.060	−0.061	0.110	−0.122
ORP	0.404	0.294	0.049	−0.060	0.091	−0.120
DOC	0.403	0.318	0.049	−0.037	0.090	−0.072
TPH	0.394	0.328	0.039	−0.026	0.073	−0.051
BTEX	0.397	0.323	0.042	−0.032	0.078	−0.061
Temperature	0.358	0.349	0.004	−0.006	0.007	−0.011
pH	0.361	0.315	0.006	−0.039	0.012	−0.076
Oxygen	0.385	0.298	0.030	−0.057	0.056	−0.112
Dichloroethene	0.704	0.074	0.350	−0.280	0.637	−0.836
Vinyl chloride	0.840	0.171	0.486	−0.184	0.980	−0.425
Ethene and ethane	0.765	0.220	0.410	−0.135	0.772	−0.290
Chloride	0.647	0.216	0.292	−0.138	0.522	−0.299
Methane	0.599	0.194	0.245	−0.161	0.435	−0.359

Of the 14 model variables, the nine pre-conditions are listed first, followed by the five footprints. *Note*: TEAP: terminal electron accepting process; ORP: oxidation reduction potential; DOC: dissolved organic carbon; TPH: total petroleum hydrocarbons; BTEX: benzene, toluene, ethylbenzene, and xylenes; LOR: Logarithm-of-the-Odds-Ratio.

The last four columns of Table 1 provide two alternate measures of importance for each variable. The middle two columns give change in probability (ΔProbability) due to evidence, expressed as:

$$\Delta\text{Probability} = \begin{cases} \Pr(\text{ADbRD is occurring}|\text{evidence}) \\ \quad - \Pr(\text{ADbRD is occurring}|\text{no evidence}) \\ \Pr(\text{ADbRD is occurring}|\text{evidence}) - 0.355 \end{cases} \quad (1)$$

Finally, in the last two columns of Table 1, the importance of each variable is measured by the change in the Logarithm-of-the-Odds-Ratio (ΔLOR) due to evidence, given by:

$$\begin{aligned} \Delta\text{LOR} &= \log\left[\frac{\Pr(\text{ADbRD is occurring}|\text{evidence})}{\Pr(\text{ADbRD is not occurring}|\text{evidence})}\right] \\ &\quad -\log\left[\frac{\Pr(\text{ADbRD is occurring}|\text{no evidence})}{\Pr(\text{ADbRD is not occurring}|\text{no evidence})}\right] \\ &= \log\left[\frac{\Pr(\text{ADbRD is occurring}|\text{evidence})}{\Pr(\text{ADbRD is not occurring}|\text{evidence})}\right] \\ &\quad - \log\left[\frac{0.355}{1-0.355}\right] \end{aligned} \quad (2)$$

Table 1 shows that Eqs. (1) and (2) provide measures of importance that are similar in scale, with the values for ΔLOR being approximately twice the magnitude of ΔProbability. These measures of importance give significantly more weight to the five footprints of reductive dechlorination compared to the nine pre-conditions. This is because experts find the production of end and daughter products to be compelling evidence; but, the presence of favorable pre-conditions has lower decision value.

3.1. The proposed scoring system and screening guide

The development of the proposed point scoring system was based primarily on the importance of each variable. Additional factors considered were: (i) the input required for building a comprehensive conceptual model for the evolution of natural attenuation processes at a site; and (ii) the realities of a long-term site remediation program. The magnitudes of importance for the different types of evidence in Table 1 can be translated into candidate scoring systems as shown in Table 2. As mentioned previously, ΔProbability and ΔLOR are approximately proportional; and hence, the corresponding scoring systems are very similar. The values of the first two scoring systems in Table 2 are directly proportional to the numbers in the last four columns of Table 1. These values, which have been rounded off for simplicity, were used with consideration of other factors important to a sampling program, discussed in detail below, to develop the proposed scoring system highlighted (in bold font) in Table 2.

The proposed scoring system assigns the same points for the nine pre-conditions as the scoring systems directly based

Table 2
Alternative scoring systems evaluating reductive dechlorination at TCE sites

Type of evidence (model variable)	Proportional to ΔProbability		Proportional to ΔLOR		Proposed scoring system		Air Force/EPA protocol	
	Positive finding	Negative finding	Positive finding	Negative finding	Positive finding	Negative finding	Positive finding	Negative finding
TEAP	1	−1	1	−1	**1**	**−1**	10[a]	0
Hydrogen	1	−1	1	−1	**1**	**−1**	3	0
ORP	1	−1	1	−1	**1**	**−1**	1 or 2	0
DOC	1	−1	1	−1	**1**	**−1**	2	0
TPH	1	−1	1	−1	**1**	**−1**	–	–
BTEX	1	−1	1	−1	**1**	**−1**	2	0
Temperature	0	0	0	0	**0**	**0**	1	0
pH	0	−1	0	−1	**0**	**−1**	0	−2
Oxygen	1	−1	1	−1	**1**	**−1**	3	−3
Dichloroethene	7	−6	6	−8	**3**	**−4**	2	0
Vinyl chloride	10	−4	10	−4	**5**	**−2**	2	0
Ethene and ethane	8	−3	8	−3	**3**	**−2**	2 or 3	0
Chloride	6	−3	5	−3	**3**	**−2**	2	0
Methane	5	−3	4	−4	**3**	**−2**	3	0
Volatile fatty acids	–	–	–	–	–	–	2	0
Carbon dioxide	–	–	–	–	–	–	1	0
Alkalinity	–	–	–	–	–	–	1	0
Chloroethane	–	–	–	–	–	–	2	0
Total	43	−27	40	−30	**24**	**−20**	41	−5

Note: TEAP: terminal electron accepting process; ORP: oxidation reduction potential; DOC: dissolved organic carbon; TPH: total petroleum hydrocarbons; BTEX: benzene, toluene, ethylbenzene, and xylenes; LOR: Logarithm-of-the-Odds-Ratio.

[a] This includes four measures: nitrate (2 points), iron(II) (3 points), sulfate (2 points) and sulfide (3 points).

on the ΔProbability and ΔLOR values. In contrast, it awards about half the points of the ΔProbability- and ΔLOR-based systems for the five footprints. In addition, some of the values are simplified so that positive evidence receive possible scores of 0, 1, 3, and 5, while negative evidence may be assigned scores of 0, −1, −2, and −4. The proposed scoring system awards no points for any findings of temperature or for positive findings of pH ($5 < \mathrm{pH} < 9$), because these were not found to be important to most of the experts, for the majority of the TCE-contaminated sites.

For the proposed scoring system, the influence of the end and daughter products was slightly diminished relative to the pre-conditions. Even though the detection of footprints is of high decision value in this tool, it is important not to neglect measuring pre-conditions, in order to account for all major factors needed to construct a site-specific conceptual model for the process of reductive dechlorination. In particular, when the scoring system is used to complement long-term monitoring, the independent assessment of the presence of electron donors will provide an indication of the sustainability of the process [4].

The Air Force/EPA protocol scoring system is included in Table 2 for easy reference. In addition to the 14 variables of the proposed expert system, the Air Force/EPA protocol also measures volatile fatty acids (another source of energy for the microorganisms), carbon dioxide (the ultimate daughter product of aerobic degradation, also produced during dechlorination reactions), alkalinity (high alkalinity is a consequence of carbon dioxide interacting with aquifer minerals, but also of H^+ consumption involved in some reductive dechlorination reactions), and chloroethane (a potential daughter product of vinyl chloride and/or dichloroethene) [4,5,13]. The Air Force/EPA protocol mainly awards points for positive evidence with only two exceptions: points are subtracted for negative findings of pH and O_2. Consequently, the more data collected, the greater the chance that some will contribute positive points toward the necessary threshold value. At least 20 points are required to conclude that there is "strong evidence" for reductive dechlorination. The points listed in Table 2 show that this rating can be achieved with measurements of pre-conditions alone. The proposed scoring system avoids this over-emphasis on reducing conditions, which was noted with concern by the NRC committee [4].

With the above considerations in mind and the data analysis discussed below, the proposed scoring system is complemented with the following screening guide:

- 12 or more points → excellent candidate for reductive dechlorination;
- 10 or 11 points → strong candidate for reductive dechlorination;
- 5–9 points → moderate candidate for reductive dechlorination;
- 0–4 points → improbable candidate for reductive dechlorination;
- <0 points → reductive dechlorination is probably not appropriate.

This screening guide was developed by analyzing over 100 cases of real and hypothetical data. The scoring system was compared with the probabilistic output of the expert system. As there is not a unique correspondence between the probabilistic output of the expert system and a score from the proposed scoring system, two cases that receive the same

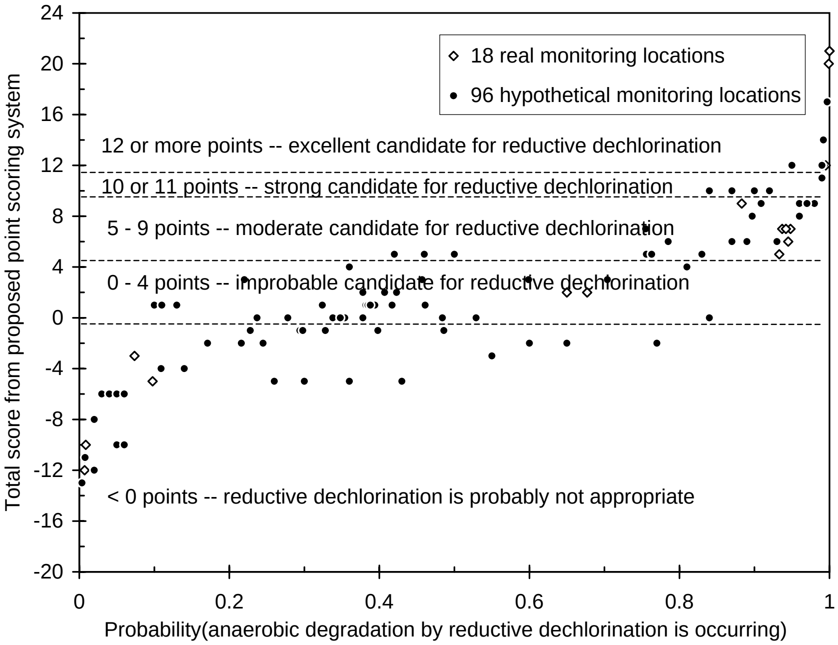

Fig. 3. Relationship between the probability that anaerobic degradation by reductive dechlorination is occurring, computed with the Stiber et al. [10] expert system, and the total score from the proposed scoring system.

score may have different values for the Pr(ADbRD is occurring). The relationship between probabilistic output of the expert system and scores from the proposed scoring system is shown in Fig. 3, with the screening guide superposed. As an example, this figure shows that a site with a score of four can have a probability ranging from 0.36 [when TEAP, hydrogen, ORP, DOC, TPH, and BTEX (all pre-conditions) are positive and vinyl chloride is negative] to 0.81 (when hydrogen and ethene-ethane are both positive). It is also observed that scores of 0–4 points may be associated with a large range of probabilities in the original expert system. These low scores may result from conflicting (positive and negative) evidence or from only a small amount of data being available.

As mentioned earlier, because there are prior beliefs embedded in the expert system, a site with no evidence, and 0 points, has a probability of 0.355. As a rule of thumb, a score of 12 points guarantees a probability of at least 0.9 and a score of 6 points guarantees a probability of at least 0.4. This guide requires at least four types of evidence in order to achieve the threshold of 12 points for an interpretation of "excellent" and cannot give this highest rating in the absence of positive evidence for the critical footprints.

4. Performance of proposed scoring system and screening guide

For purposes of illustration, the proposed scoring system was used for a TCE site in Niagara Falls, New York, described in the literature by Yager et al. [14]. Table 3 compares the probabilities produced by the expert system and the scores of the proposed scoring system, for three monitoring wells at the Niagara Falls site. The probabilities and the scores were calculated taking into account measurements of nine variables. These include five pre-conditions, namely, TEAP, hydrogen, DOC, pH, and oxygen, and four footprints: vinyl chloride, ethene, chloride, and methane. Although wells 87-20 and 89-02 within the plume were assigned high probabilities (0.95 and 0.94, respectively) by the expert system, they each scored 7 points (designating them as "moderate" candidates for reductive dechlorination) under the proposed scoring system. Despite the fact that vinyl chloride was detected in these wells, they had poor evidence for the pre-conditions TEAP and hydrogen. The proposed scoring system gives more weight to these pre-conditions and thus scored the site less favorably than did the expert system. This example and Fig. 3 show that while the scoring system does not capture every subtlety of the expert system, the two forecasts are highly correlated. Hence, the proposed scoring system is still able to differentiate much better than the Air Force/EPA protocol between the two wells in the plume (wells 87-20 and 89-02) and the downgradient well (well 89-06), as indicated by the respective scores in Table 3. Nyer et al. [15] report a related experience: the Air Force/EPA protocol indicated "limited evidence" for reductive dechlorination at a site, while further data showed that there was in fact significant biodegradation of chlorinated solvents taking place.

The performance of the proposed scoring system was further assessed with published data from the St. Joseph, Michigan site, directly west of Lake Michigan [16]. Several studies have confirmed the extensive natural attenuation of TCE at this site [4]. However, detailed multilevel sampling suggests that infiltration of oxygenated lake water is associated with the slower rates of reductive dechlorination at

Table 3
Comparison of the expert system and the proposed scoring system for a TCE site in western New York [14]

	Probability from the expert system	Score from the proposed scoring system	Score from the Air Force/ EPA protocol
No evidence	0.36	0	0
Well 87-20 (in plume)	0.95	7 (moderate candidate)	13 (limited evidence)
Well 89-02 (in plume)	0.94	7 (moderate candidate)	10 (limited evidence)
Well 89-06 (downgradient)	0.10	−5 (inappropriate candidate)	9 (limited evidence)

Table 4
Application of the proposed scoring system at the St. Joseph, Michigan, site [16]

Sampling location	Points awarded for each type of evidence by the proposed scoring system							Total points	Proposed interpretation
	TEAP	H_2	ORP	O_2	VC	Ethene and ethane	Methane		
1	−1	1	−1	−1	5	−2	3	4	Improbable candidate
2	−1	1	−1	−1	5	−2	3	4	Improbable candidate
3	−1	NA	1	1	5	−2	3	7	Moderate candidate
4	−1	−1	1	1	5	3	3	11	Strong candidate
5	−1	1	1	1	5	3	3	13	Excellent candidate
6	−1	1	1	1	5	3	3	13	Excellent candidate
7	−1	1	1	1	5	3	3	13	Excellent candidate
8	−1	1	1	−1	5	3	3	11	Strong candidate

Ground-water samples were collected at the same location at eight different elevations below the water table (sampling location 1 is the closest to the ground surface). *Note*: TEAP: terminal electron accepting process; ORP: oxidation reduction potential; VC: vinyl chloride; NA: not available.

shallow depths, close to the interface between the ground water and the surface water of the lake [16]. These multilevel sampling data were used to further test the ability of the proposed scoring system to distinguish among varying reducing environments. Table 4 gives results for eight sampling depths at the same location. Ground-water samples were collected from the same boring at increasing depths below the water table (sampling location 1 is closest to the surface). The samples spanned approximately a 6 m depth. Three or four samples were obtained from each location within a 6-month period. To assign a score, a variable was considered to be positive if its state was positive at least twice during the sampling rounds. The consistent presence of vinyl chloride indicates that reductive dechlorination of TCE takes place at all depths. However, the total points awarded by the proposed scoring system reveal a gradual overall increase of the effectiveness of reductive dechlorination with depth.

It should be stressed that the proposed scoring system is intended to be a tool for structuring data when the natural attenuation of TCE is being considered as a remedial alternative. Its intended use is as a first-level screening tool to identify those sites with the greatest probability of adequate reductive dechlorination and to complement long-term monitoring. Consequently, it would be inappropriate to rely solely on this scoring system for final decisions on remedial strategy and compliance, as was also cautioned against by the NRC committee [4]. It should further be made clear that the scoring system only evaluates the technical adequacy of the biodegradation processes at a site. In addition, there are many other factors that merit consideration. Among these are regulator approval, community concerns, time and cost constraints, and potential human health and ecological impacts. Decisions for the selection of *enhanced* as opposed to *intrinsic* monitored remediation must balance the risks resulting from reliance on natural attenuation processes versus the cost savings sought from this approach.

5. Summary and conclusions

This paper described a generalized methodology for the development of decision tools to support decisions involving complex processes, and demonstrated its use for the evaluation of natural attenuation at TCE sites. The methodology consists of building a causative model for a complex event of interest and specifying probabilistically the causal structure of the model through expert elicitations. The essence of the expert system can then be embedded in a scoring system by determining the decision worth of the expert system variables through a sensitivity analysis of the model. A scoring system thus developed is simple to use yet transparent. This transparency permits not only peer reviewing but also subsequent refinements as the science base evolves. Such a scoring system attributes relative weights to findings based on expert beliefs and includes negative points for negative findings. Awarding negative points makes possible sharper distinctions between promising and problematic sites. Hence, cost savings are achieved not only when natural attenuation solutions are pursued at suitable sites, but also when the site characterization efforts required for such solutions are aborted earlier at inappropriate candidates. Example applications of the proposed scoring system and the accompanying screening guide confirmed that the methodology results in decision tools that are consistent with recent consensus recommendations for evaluating the remedial potential of natural attenuation by reductive dechlorination for TCE.

Acknowledgements

The authors thank Richard Luthy and 22 expert elicitees for their contributions to this research. Financial support for this work was provided partly by NSF CAREER Grant CMS-9502546. The research presented herein was conducted while Dr. Stiber was a graduate student at Carnegie Mellon University. This paper presents the personal views of its authors and is not intended to portray policies of the U.S. EPA.

References

[1] US EPA, Microbial Processes Affecting Monitored Natural Attenuation of Contaminants in the Subsurface, EPA/540/S-99/001, US EPA Office of Research and Development, Washington, DC, 1999.

[2] NRC, In Situ Bioremediation, National Academy Press, Washington, DC, 1993, p. 207.

[3] NRC, Innovations in Ground Water and Soil Cleanup, National Academy Press, Washington, DC, 1997, p. 310.

[4] NRC, Natural Attenuation for Groundwater Remediation, National Academy Press, Washington, DC, 2000, p. 274.

[5] T.H. Wiedemeier, M.A. Swanson, D.E. Moutoux, E.K. Gordon, J.T. Wilson, B.H. Wilson, D.H. Kampbell, J.E. Hansen, P. Haas, F.H. Chapelle, Technical Protocol for Evaluating Natural Attenuation of Chlorinated Solvents in Groundwater, US Air Force Center for Environmental Excellence, San Antonio, TX, 1996.

[6] US EPA, Technical Protocol for Evaluating Natural Attenuation of Chlorinated Solvents in Groundwater, EPA/600/R-98/128, US EPA Office of Research and Development, Washington, DC, 1998.

[7] S. Hart, In situ bioremediation: defining the limits, Environ. Sci. Technol. 30 (9) (1996) 398A–401A.

[8] M.B. Powers, D.K. Rubin, Doing what comes naturally—nothing, Eng. News Record 237 (16) (1996) 26–28.

[9] J. Pearl, Probabilistic Reasoning in Intelligent Systems, Morgan Kaufman, San Mateo, CA, 1991, p. 552.

[10] N.A. Stiber, M. Pantazidou, M.J. Small, Expert system methodology for evaluating reductive dechlorination at TCE sites, Environ. Sci. Technol. 33 (17) (1999) 3012–3020.

[11] N.A. Stiber, M.J. Small, M. Pantazidou, Site-specific updating and aggregation of Bayesian Belief Network models for multiple experts, Risk Anal., submitted for publication.

[12] Norsys, Netica^{TM} Application for Belief Networks and Influence Diagrams: User's Guide, version 1.03 for Macintosh and version 1.05 for Windows, Norsys Software Corporation, Vancouver, Canada, 1998.

[13] E.K. Nyer, M.E. Duffin, The State of the Art of Bioremediation, Groundwater Monitoring and Remediation, Spring, 1997, pp. 64–69.

[14] R.M. Yager, S.E. Bilotta, C.L. Mann, E.L. Madsen, Metabolic adaptation and in situ attenuation of chlorinated ethenes by naturally occurring microorganisms in a fractured dolomite aquifer near Niagara Falls, New York, Environ. Sci. Technol. 31 (11) (1997) 3138–3147.

[15] E.K. Nyer, P. Mayfield, J. Hughes, Beyond the AFCEE Protocol for Natural Attenuation, Groundwater Monitoring and Remediation, Summer, 1998, pp. 70–77.

[16] J.M. Lendvay, S.M. Dean, P. Adriaens, Temporal and spatial trends in biogeochemical conditions at a groundwater—surface water interface: implications for natural bioattenuation, Environ. Sci. Technol. 32 (22) (1998) 3472–3478.