ANS Reprint

RADIOACTIVE WASTE MANAGEMENT AND DISPOSAL

KEYWORDS: *approximate reasoning, nuclear waste, safety analysis*

AN APPROXIMATE REASONING-BASED METHOD FOR SCREENING HIGH-LEVEL-WASTE TANKS FOR FLAMMABLE GAS

STEPHEN W. EISENHAWER,* TERRY F. BOTT, and RONALD E. SMITH
Los Alamos National Laboratory, MS K557, Los Alamos, New Mexico 87545

Received January 20, 1999
Accepted for Publication December 28, 1999

The in situ retention of flammable gas produced by radiolysis and thermal decomposition in high-level waste can pose a safety problem if the gases are released episodically into the dome space of a storage tank. Screening efforts at the Hanford site have been directed at identifying tanks in which this situation could exist. Problems encountered in screening motivated an effort to develop an improved screening methodology. Approximate reasoning (AR) is a formalism designed to emulate the kinds of complex judgments made by subject matter experts. It uses inductive logic structures to build a sequence of forward-chaining inferences about a subject. Approximate-reasoning models incorporate natural language expressions known as linguistic variables to represent evidence. The use of fuzzy sets to represent these variables mathematically makes it practical to evaluate quantitative and qualitative information consistently. In a pilot study to investigate the utility of AR for flammable gas screening, the effort to implement such a model was found to be acceptable, and computational requirements were found to be reasonable. The preliminary results showed that important judgments about the validity of observational data and the predictive power of models could be made. These results give new insights into the problems observed in previous screening efforts.

I. INTRODUCTION

High-level waste (HLW) produces flammable gases as a result of radiolysis and thermal decomposition of organics. The gases of concern include hydrogen, ammonia, and methane. Under certain conditions, it is possible for these gases to be retained within the waste for extended periods and then to be released quickly into the dome space of the storage tank. This behavior has been observed in a number of tanks at the Hanford Site. It is known that in at least one such tank (241-SY-101 before it was mitigated), the concentration of flammable gases has been at or above the lower flammability limit (LFL) for short periods. As part of the effort to reduce the safety concerns associated with flammable gas in HLW tanks at Hanford, a flammable gas watch list (FGWL) has been established. Inclusion on the FGWL is based on criteria intended to measure the risk associated with the presence of flammable gas. It is important that all high-risk tanks be identified with high confidence so that they may be controlled. Conversely, to minimize operational complexity, the number of tanks on the watch list should be reduced as near to the "true" number of flammable risk tanks as the current state of knowledge will support.

There are several functional steps in the FGWL screening process. The first is to determine if the available information is sufficient to allow a meaningful evaluation of the tank. If so, the evaluation is performed. The result of the evaluation is a recommendation on whether the tank should be on the watch list. In this latter step, some statement of the confidence associated with the recommendation is required. The actual process of going from some universe of information for a tank to a clear recommendation on tank classification is a complex, frequently implicit combination of inferences about flammable gas phenomenology. These inferences about gas generation, composition, retention and release characteristics for a tank are drawn from a large, diverse, uncertain, and often contradictory universe of information. This universe includes

1. observations associated with gas release events
2. measurements and associated models for predicting volumes of retained gas
3. waste properties associated with empirical models to roughly estimate the potential for gas generation, retention, and release.

*E-mail: seisenhawer@lanl.gov

It is quite common for a conclusion drawn from one set of data and models to be diametrically opposed by some other set of data and models. Data vary in terms of quality and the degree of associated uncertainty, and models have varying powers of prediction. As a result, the evaluation must contend with an entire series of qualitative judgments about what inferences regarding flammable gas phenomenology are possible and how to resolve discrepancies among them. Compounding this problem is the fact that perhaps as many as 177 tanks, with widely varying waste characteristics, must undergo the screening, and some demonstration of consistent evaluation is needed. This is difficult to achieve because of the wide variations in waste types and the large differences in installed instrumentation and the historical database.

In this paper, we present an alternative to existing approaches for FGWL screening based on the theory of approximate reasoning[1] (AR). Our AR-based model emulates the inference process used by an expert when asked to make an evaluation. We start with a brief history of the FGWL and the current approach to screening. This is followed by an overview of AR methodology. We briefly describe the overall logic structure of the pilot model developed for flammable gas screening. The evaluation process for an AR model is examined using one small segment of the complete model, and we show how a likelihood statement about retained gas is obtained. The computer implementation for the FGWL AR model is explained, and insights into the screening problem are presented.

II. FGWL HISTORY

A useful starting point in the description of screening methodology is to examine the evolution of the FGWL since 1990. Tanks on the original FGWL were determined using a "slurry growth" criterion based on a cursory review of tank records and supporting documents. The original FGWL had a total of 20 tanks.[2] Later, a ranking system was developed to evaluate all waste tanks in terms of the potential for flammable gas production and the potential for gas retention and release at concentrations above the LFL. By 1994, there were 25 tanks on the FGWL.

The ranking system was an initial, informal attempt to relate FGWL membership to the three requirements for a hazard to exist. These are

1. one or more processes to generate flammable gases
2. the existence of a mechanism to store these gases in the waste for an extended interval
3. the possibility of releasing these gases over a short period of time so that a gas concentration above the LFL could exist for an extended interval.

As noted earlier, flammable gas generation in HLW is to be expected as a result of radiolysis and a number of chemical decomposition reactions known to be important in the range of storage temperatures. Extensive research has lead to a reasonable understanding of how gas composition and generation rates vary with waste type and temperature. The ability to retain gas was originally related to a specific type of waste, similar to the slurry in 241-SY-101, where the increase in waste level was known to be related to gas retention, hence the term "slurry growth." Here again, research has helped to illuminate how retention mechanisms act to store flammable species as gas or in solution and to explain how these mechanisms are related to the basic chemical and physical characteristics of the waste. If the flammable gas is released slowly, then natural or forced ventilation will almost certainly keep the concentration in the dome space below the LFL. However, early observations of tank 241-SY-101 clearly indicated that large amounts of gas could be released in a short period of time. The release mechanism in this tank has been identified as being related to the Rayleigh-Taylor hydrodynamic instability. For other tanks, gas release either by this mechanism or as a result of operations that disturb the waste are considered credible.

Inadequacies in the FGWL criteria were well known, and new criteria were proposed in 1994 by the then-primary contractor, Westinghouse Hanford Company.[3] The criteria are based on the concentration of flammable gas needed to support combustion and the subsequent overpressure that could be produced in a flammable gas accident. A safety factor of 4 is used. In practice, this means that to avoid watch list membership, the flammable gas concentration in the dome space of a tank must not exceed 25% of the LFL. The factor of 4 is based on common standards for fire safety and pressure vessel design.

An effort has been under way since 1995 to evaluate all of the waste tanks against the 1994 criteria using a methodology proposed by Hopkins[4] and implemented by Hodgson et al.[5] Several problems have been encountered. First, the calculated gas concentration C_{fg} varies widely depending on which sensor data and models are used. This is to be expected. The instrumentation types and the associated data quality vary considerably from tank to tank. Models for gas generation and retention are approximate in nature and depend in any case on a reasonable understanding of what types of waste exist in a tank. This is often very uncertain because of gaps in the waste transfer records and incomplete understanding of the long-term tank chemistry. For comparison to the threshold of 0.25 LFL, Hodgson et al. use $\max[C_{fg}(i)]$, where i denotes the concentration values calculated using discrete data/model sets. However, this preliminary comparison to the LFL is not the final screening result. An additional evaluation referred to here as the ad hoc judgment follows and is made for two good reasons:

1. Calculation of $\max[C_{fg}(i)]$ does not use all of the information available about gas generation, retention, and release phenomenology in a tank.

2. Practically, too many tanks have a value of $\max[C_{fg}(i)]$ above the threshold criterion.

In the ad hoc step, other information or expert judgment is incorporated, and a final evaluation judgment is made. Thus, much of the decision process is done "offline" but is nevertheless an essential element in the screen being used. Note that while uncertainty in observational or parametric data can be propagated when determining $\max[C_{fg}(i)]$, there is no consistent method applied to represent uncertainty in expert judgment. However, the relative quality of sensors or calculational models represent the major sources of uncertainty that must be evaluated by the experts during the screening process. This makes it extremely difficult to provide best-estimate/degree-of-conservatism comparisons. Finally, with the structure discussed here, it is difficult to ensure that the screening is consistent, and this poses problems during review. The foregoing considerations suggest that the design and implementation of a new screening method should be based on a formalism that is robust and adaptable and in which all of the necessary judgments are defined explicitly. The theory of approximate reasoning provides such a formalism.

III. OVERVIEW OF AR METHODOLOGY

The general structure underlying the AR method described in this paper is shown in Fig. 1. We begin with some universe of information about a tank to be screened. This universe consists of both qualitative and quantitative data. This information is not necessarily in a form that it is directly useful. Therefore, some processing of the data is required. We denote this processed data as a body of evidence, and only elements within it will be considered in the screening process. Elements of evidence must be related to each other in some meaningful manner. This is done through formal structures with logical operations relating the evidence to produce a series of forward-chaining inferences. The output from the logic structure is a description of the system called a state vector. The state vector here is a concise description of the aspects of the tank that affect flammable gas screening. The elements of a state vector are always assumed to include some component of uncertainty that reflects imprecision or ambiguity in the knowledge of the system state. Finally, the system state vector is used in a decision model in which some definite statement about the system is made. Note that the level of abstraction increases as we move through the process. In this paper, we concentrate on how evidence is incorporated into a logic structure and how a useful state vector may be obtained.

In an AR model, the elements of evidence are handled as linguistic variables; that is, natural language descriptors are used. For example, we can characterize the temperature in a room as "too cold," "comfortable," or "too hot" without actually measuring the temperature. The descriptors are used to define sets in which the variable of interest, in this case the temperature in the room, may belong. We say that the universe of discourse for the room temperature T is $T \in$ {{Too Cold}, {Comfortable}, {Too Hot}}. The sets used in AR are fuzzy. That is, a variable

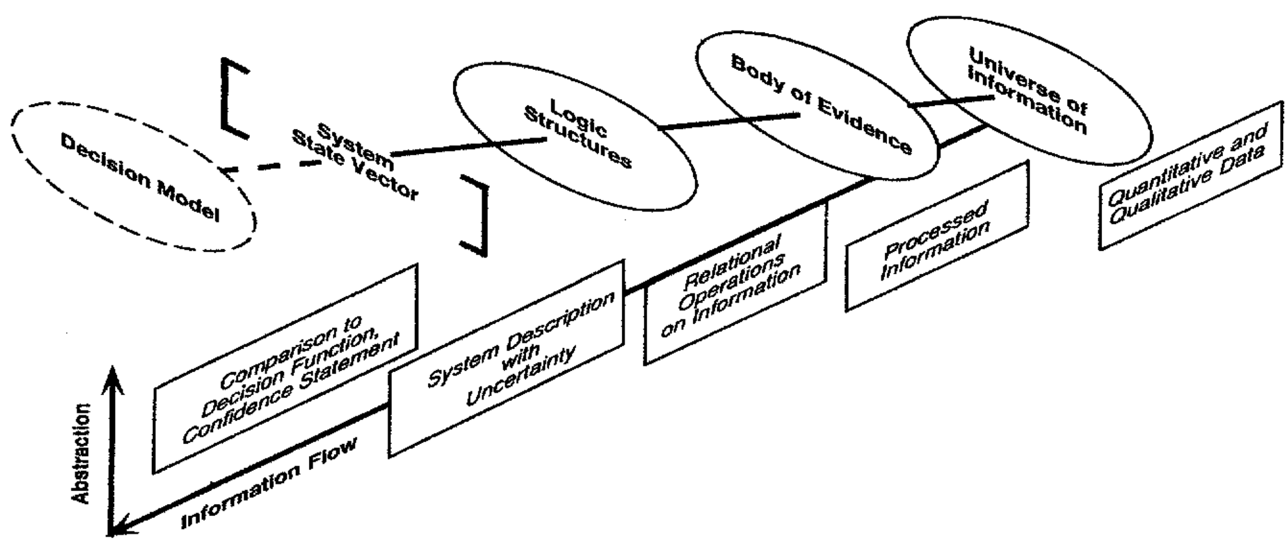

Fig. 1. Overall structure for an approximate reasoning model.

may belong to sets that traditionally might be considered to be mutually exclusive. For example, a temperature of 70°F could belong to both the {Comfortable} and {Too Hot} sets.

An element of information can be either quantitative or qualitative, but it is important to note that in either case, it is almost inevitably uncertain. If an element is defined numerically, it is treated as a classic random variable characterized by a probability density function (pdf). Defining the parameters in the density function then characterizes the uncertainty. Measured room temperature could be defined in this way. A qualitative element can also be a random variable. The total uncertainty associated with an element of evidence is composed of two components: aleatory and epistemic. The aleatory component represents the inherent variability in a parameter. Processes such as radioactive decay and turbulence exhibit aleatory uncertainty. The epistemic component represents state-of-knowledge uncertainty. For example, the assumptions and approximations made in a model induce epistemic uncertainty in the results. That is, there is some doubt about how well the model represents physical reality. It is important to note that in many problems, epistemic uncertainty is greater than the aleatory component.

The logic structure defines a set of relationships between the elements of evidence. The nature of the individual branch junctions depends upon the particular type of relation used. A relation is a type of function that maps multiple inputs into a single output. In this paper, we consider only relations with two inputs and one output. Many different types of relations, both numerical and logical, are possible. However, in our AR model, the only relation used is formal logical implication. The implications are of the form "If A and B then C," or "A and B implies C," written symbolically as $(A \wedge B) \rightarrow C$. Here, A and B are called the antecedents and C is the consequence. If the universe of discourse for A has i elements and B has j elements then we need $i \times j$ different implications about C to cover all the possible combinations of the two antecedents. We refer to this set of implications as a rule base. The complete form of the inference rule is

$$\text{(A is } A_i \text{ and B is } B_j\text{) and (If } A_i \text{ and } B_j \text{ imply } C_k\text{) then } C_k\text{, where } k \leq i \times j \ , \tag{1a}$$

or

$$[(A_i \wedge B_j) \wedge ((A_i \wedge B_j) \rightarrow C_k)], C_k \quad k \leq i \times j \ . \tag{1b}$$

This statement is a special logical construct known as the *modus ponens* tautology and is the basic form of rule base used in all AR models. Later, we discuss how such rule bases are evaluated.

IV. DESIGN OF THE FGWL AR MODEL

An important consideration in the initial development of an AR model is the scope of the model. This determines the size of the required logic structure and is a major determinant in the amount of work required to complete the model. To illustrate the AR approach to FGWL screening, we chose to restrict ourselves to an evaluation of retained gas. This is smaller in scope than the analysis undertaken by Hodgkins et al. There are several reasons for this. First, the body of evidence concerning gas generation and retention appears to be generally more mature than that associated with other aspects of flammable gas phenomenology. Consideration of gas release processes did not appear to provide any additional insight into the issue of implementing an AR model. Second, the body of evidence for retention provides a diverse set of data and models that is sufficient to illustrate the ability of an AR model to combine quantitative and qualitative information and make sophisticated judgments about model validity and the resolution of conflicting results. Application of the model in this form to a large number of tanks also provides an understanding of the effect of data uncertainty on the screening evaluation.

The parameters used in the algorithm are grouped into three general classes: predictors, enablers, and indicators. These three classes correspond to the primary modules in the forward-chaining logic structure shown in Fig. 2. Each class of parameters provides a distinct judgment concerning the likelihood of a significant quantity of retained gas. We denote these judgments using the linguistic variable of likelihood as L_P, L_E, and L_I, where the subscripts denote the predictor, enabler, and indicator parameter groups. Each linguistic variable represents an independent evaluation of the likelihood for a significant quantity of retained gas based on a particular combination of logical inferences. Predictor parameters act as antecedents for inferences about the volume of gas that may be retained. All of the submodules used to infer L_P involve one or more measurements of waste level and an associated model that combines this level data with other information to infer gas volume. Examples of measurements used in the predictor module include long-term changes in waste level and correlations between barometric pressure and level fluctuations. We

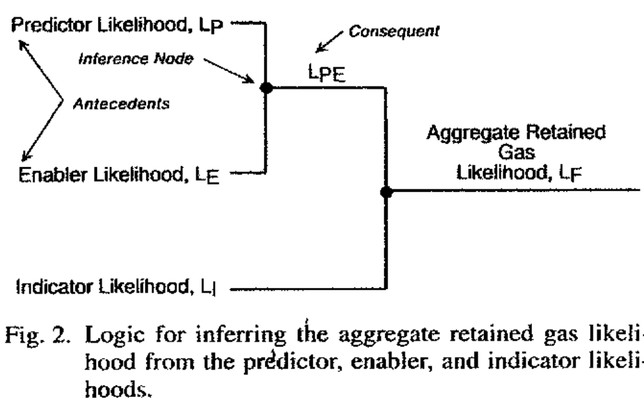

Fig. 2. Logic for inferring the aggregate retained gas likelihood from the predictor, enabler, and indicator likelihoods.

discuss both of these predictor models later. Enablers are sets of parameters that, when properly combined, provide a basis for estimating the gas generation rate and the gas retention effectiveness for a tank. Elements of evidence used in the enabler module include such waste characteristics as organics concentration, waste temperature, and volumetric heat generation rates. Conceptually, the inference structures associated with the evidence in this module are similar in principle to those used in the early FGWL screening process. Gas indicators are parameters that can be used to infer the existence of a gas release event. Positive indicators are direct measurements of an unambiguous nature, such as a dome space flammable gas concentration measurement. The absence of such positive indicators does not prove that a tank is a nonflammable gas tank. Similarly, a negative indicator has a threshold value that indicates conclusively that significant gas retention is not possible in the tank because of some distinct combination of physical characteristics of the waste and tank.

Another important consideration in designing an AR model is determining the form in which the final output is to be expressed. This is equivalent to specifying the format in which a subject matter expert is expected to state his or her conclusions. Ideally, the natural language expressions associated with the output of the model are developed in conjunction with the experts used in building it. The linguistic variable chosen for the final output in the pilot model was "likelihood of a significant quantity of retained gas." The adjective "significant" means that there is sufficient gas retention so that safety or regulatory concerns exist. We express the output likelihood, referred to as the aggregate likelihood, L_F with the following universe of discourse:

$$L_F \in \left\{\begin{array}{l}\{\text{Extremely Unlikely}\},\{\text{Very Unlikely}\},\\ \{\text{Quite Unlikely}\},\{\text{Unresolved}\},\{\text{Quite Likely}\},\\ \{\text{Very Likely}\},\{\text{Extremely Likely}\}\end{array}\right\} . \tag{2}$$

We soon show how these descriptors are used directly in approximate reasoning. The hedges "extremely," "very," and "quite" are intended to provide sufficient resolution to allow meaningful distinctions to be made and are characteristically used by subject matter experts. The set {Unresolved} is used for evaluations in which the results are inconclusive. This is equivalent to an expert saying "I do not know" or "The data are inconclusive or contradictory." We use the expression "likelihood" in the sense that it "supplies a natural order of preference among the possibilities under consideration."[6] That is, something that is said to be "very likely" is understood to have a more realistic chance of happening or to occur more frequently than something that is "quite likely." However, it must be emphasized that the likelihood linguistic variable is not to be confused with quantitative probability nor do we intend our use of likelihood to be associated directly with the likelihood function of probability theory.

V. ILLUSTRATION OF AR-BASED FGWL SCREENING

To illustrate the operation of an AR model, we will use a short excerpt from the complete FGWL screening algorithm. This short segment is one of the logic submodules used to infer the predictor likelihood L_P. Hopkins[4] first recognized that under the correct circumstances, the absolute level of the waste in a tank could provide information on the amount of retained gas. A substantial difference between the measured waste level and the waste level predicted by the fill/transfer history of the tank, when corrected for evaporation, can be evidence of gas retention in the waste. The greater the unexplained level change Δh, the greater is the potential volume of trapped gas. Hopkins defines the effective long-term level change to be attributed to retained gas as[a]

$$\Delta h = h' - h_{81} + \Delta h_{81} + \Delta h_E = \Delta h_M + \Delta h_{81} + \Delta h_E \ , \tag{3}$$

where

h' = recently measured level corrected for transfers since 1981

h_{81} = level measured in 1981 (used as a datum)

Δh_{81} = estimated gas retention level change prior to the 1981 measurement

Δh_E = correction to the level to account for evaporation after 1981.

This model is conceptually simple; however, as Hopkins recognized, its application can be difficult. All waste transfers and other losses from the tank including evaporation must be accounted for. Given the state of the historical records, the large uncertainty in level measurements for some sensors, and the possibility of slow leaks or intrusions, calculating accurate level changes can be complex, and in some cases, the results to be inferred from it are problematic.

The difference between the first two terms in Eq. (3) is the measured level change denoted by Δh_M. The value of Δh as a predictor of retained gas depends to a large degree on how large the correction terms Δh_E and Δh_{81} are relative to Δh_M. The relative importance of Δh_E and Δh_{81} in determining Δh is represented by the parameters M_{81} and M_E, which are defined as

$$M_{81} = |\Delta h_{81} / \Delta h_M|$$

and

$$M_E = |\Delta h_E / \Delta h_M| \ .$$

The larger the absolute value of these ratios, the larger is the influence of the poorly known parameters Δh_E and

[a]Other factors that may affect Δh have been neglected here.

Δh_{81}. If these correction terms are large, then it is reasonable to discount the importance of this model prediction. This is exactly the type of expert judgment that an AR model is designed to emulate.

We wish to draw an inference in this example about the likelihood $L_{\Delta h}$ of a significant quantity of retained gas. To determine this likelihood, both the unexplained level change Δh and the quality of the data used to calculate this parameter should be evaluated. The logic structure for this evaluation is shown in Fig. 3. The three inputs are the long-term level change and the two parameters M_{81} and M_E used to measure the effect of correction terms on the estimate for Δh. These two parameters act as antecedents to allow us to infer a quality parameter Q that represents an expert judgment about how much confidence we should have in applying the model for a particular set of inputs. We can then use this judgment and the level change Δh to infer a likelihood of a significant quantity of retained gas $L_{\Delta h}$. The logic structure in Fig. 3 is a two-step inference chain. Large inferential structures can be built by the simple process of connecting many small chains together in the proper sequence. This forward chaining of inferences is characteristic of AR models.

To proceed further, one must first define the fuzzy sets in which the two correction terms may belong. In this case, we assert that both terms have membership in the same three sets: $(M_E, M_{81}) \in \{\{\text{Small}\}, \{\text{Medium}\}, \{\text{Large}\}\}$. In practice, the number of sets and the actual set linguistics are developed in conjunction with subject matter experts. The two correction terms initially will be expressed numerically, so we need a way to assign the degree of membership in each set for a particular value of the correction term. This is done using membership functions such as those shown in Fig. 4.

If $M_E = 3$ (that is, the evaporation correction is three times larger than the measured level change), then M_E is said to have degrees of membership, γ_E in the three fuzzy sets of $\gamma_E = [0, 0.5, 0.5]$. That is, the degree of membership in {Medium} is $\gamma(M_E, \text{Medium}) = 0.5$. Similarly a

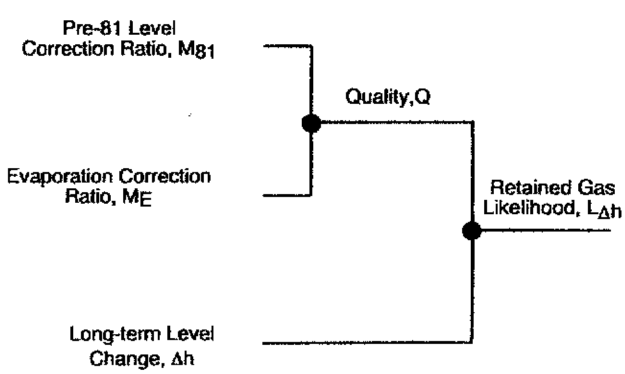

Fig. 3. Logic structure for the determination of the likelihood of retained gas using the long-term level change evidence.

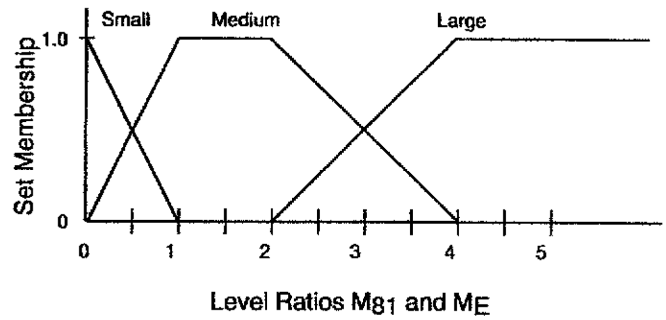

Fig. 4. Fuzzy set membership functions for the level change correction magnitude variables.

value of $M_E = 0.25$ would imply degrees of membership of $\gamma_E = [0.75, 0.25, 0]$. In natural language, this might be expressed as "the evaporation correction ratio is fairly small."

Given the antecedents M_E and M_{81} and the fuzzy sets to which they belong, we are now prepared to define a set of expert judgments that relate them to the quality Q of the long-term level change prediction. We chose to use the fuzzy sets {{Poor}, {Fair}, and {Good}} to describe Q, $Q \in$ {{Poor}, {Fair}, and {Good}}. It is not necessary to define membership functions for Q because it is not itself an element of evidence and exists only as an internal linguistic variable. There are nine rules in the rule base for Q; they are shown in Table I. The shaded box corresponds to the rule:

> "IF the evaporation correction ratio is *medium* AND the Pre-1981 level change correction ratio is *medium* THEN the quality of the unexplained level change model is *fair*."

Referring to Fig. 3, the next inference is made about $L_{\Delta h}$ using Q and the value of Δh itself. Again, we define the fuzzy sets in which Δh may have membership: $\Delta h \in$ {{Very Small}, {Quite Small}, {Moderate}, {Quite Large}, {Very Large}}, and the associated membership functions

TABLE I

Rule Base for Inferring the Measurement Quality Q from the Correction Ratios M_{81} and M_E

M_{81}	Large	Fair	Poor	Poor
	Medium	Fair	Fair	Poor
	Small	Good	Fair	Poor
		Small	Medium	Large
			M_E	

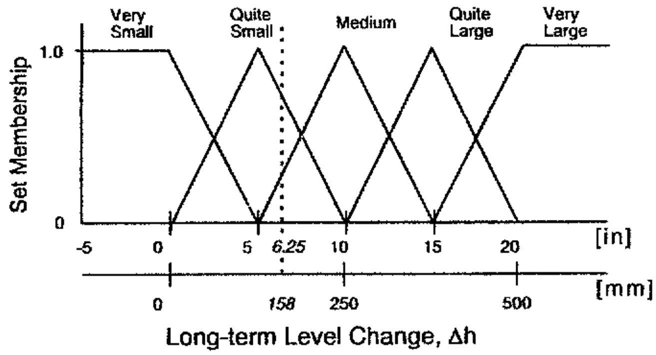

Fig. 5. Fuzzy set membership functions for the long-term level change.

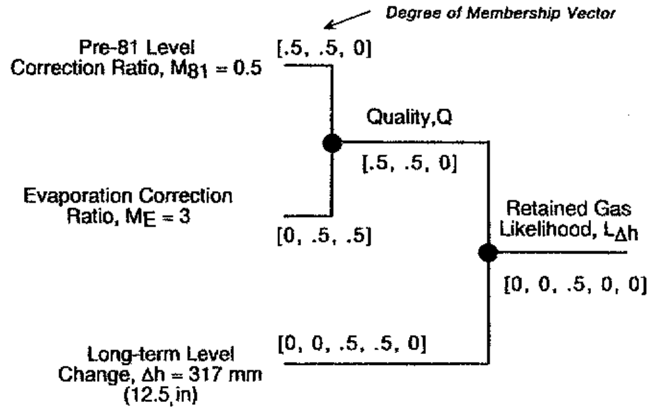

Fig. 6. Numerical example of inferring retained gas likelihood for the logic structure in Fig. 3.

shown in Fig. 5. The likelihood of retained gas $L_{\Delta h}$ is a linguistic variable that we choose to characterize by its membership in a series of sets that describe degree of likelihood:

$$L_{\Delta h} \in \left\{ \begin{matrix} \{\text{Very Unlikely}\},\{\text{Quite Unlikely}\}, \\ \{\text{Unresolved}\},\{\text{Quite Likely}\},\{\text{Very Likely}\} \end{matrix} \right\} .$$

Note that this is a subset of the universe of discourse used for the aggregate likelihood L_F. It should be reemphasized that "Unresolved" does not mean "Equally Likely" but rather "No definite statement can (or should) be made." The rule base for inferring $L_{\Delta h}$ is given in Table II. Note in particular the bottom row in the rule base. If the quality is poor, then $L_{\Delta h}$ always evaluates to "Unresolved." This row of the rule base deals with the situation in which the quality of the data does not allow a strong conclusion to be reached with this model.

Consider now a numerical example using this set of linguistic variables, membership functions, and rule bases. For example, assume that the following differential level data are available:

$$\Delta h_M = 71 \text{ mm (2.8 in.)} ;$$

$$\Delta h_{81} = 35 \text{ mm (1.4 in.)} ;$$

$$\Delta h_E = 211 \text{ mm (8.3 in.)} .$$

These correspond to $M_{81} = 0.5$, $M_E = 3$, and $\Delta h =$ 317 mm (12.5 in.). The degrees of membership for M_{81} and M_E from Fig. 4 are $\gamma_{81} = [0.5, 0.5, 0]$ and $\gamma_E = [0, 0.5, 0.5]$, respectively. These membership vectors are shown as inputs to the logic submodule in Fig. 6. This particular combination of degrees of membership means that four of the rules in Table I will be operative—the lower two rows by the two rightmost columns. We say that these four rules "fire."[b] The mathematical operations associated with the evaluation of the activated rules are discussed in Ref. 7. For Q, evaluation of the rule base yields $\gamma_Q = [0.5, 0.5, 0]$; that is, the quality has equal membership in the {Poor} and {Fair} sets. We could express this as "the quality is poor to fair," which reflects the judgment incorporated into the rule base that if either ratio is large, then the quality cannot be good. The degrees of membership for $\Delta h = 317$ mm (12.5 in.) are $\gamma_{\Delta h} = [0, 0, 0.5, 0.5, 0]$. That is, the level change has nonzero membership only in {Moderate} and {Quite Large}. Evaluation of the rule base for $L_{\Delta h}$ using Q and

[b]The firing of *modus ponens* rule bases with fuzzy antecedents is determined using the max-min rule. The details of the operation of this rule are discussed in Ref. 7.

TABLE II

Rule Base for Inferring the Gas Retention Likelihood from the Measurement Quality Q and the Waste Level Change Δ_h

Q		Very Small	Quite Small	Medium	Quite Large	Very Large
	Good	Very Unlikely	Quite Unlikely	Unresolved	Quite Likely	Very Likely
	Fair	Quite Unlikely	Unresolved	Unresolved	Unresolved	Quite Likely
	Poor	Unresolved	Unresolved	Unresolved	Unresolved	Unresolved

Δh

Δh as antecedents yields $\gamma_{L\Delta h} = [0, 0, 0.5, 0, 0]$. Figure 6 shows how degrees of membership are propagated for this example as the rule bases act on the elements of evidence.[c]

We recognize $\gamma_{L\Delta h}$ as a simple state vector. With this set of data, there is only nonzero membership for $L_{\Delta h}$ in the likelihood fuzzy set {Unresolved}. The only possible conclusion is that the likelihood of a significant quantity of retained gas using the long-term level change is "unresolved." This agrees with the premise stated earlier that if correction terms are large, then the inference based on the long-term level change must be weak.

VI. EXPRESSING THE EVALUATION RESULT

In our example for $L_{\Delta h}$, the interpretation of the membership vector $\gamma_{L\Delta h} = [0, 0, 0.5, 0, 0]$ is straightforward because there is a nonzero membership only in {Unresolved}. In general, there will be nonzero memberships in a number of the likelihood fuzzy sets. This is true not only for $L_{\Delta h}$ but also for the aggregate retained gas likelihood L_F as well. In the latter case, the vector describes the memberships in the sets defined in Eq. (2): {{Extremely Unlikely}, {Very Unlikely}, {Quite Unlikely}, {Unresolved}, {Quite Likely}, {Very Likely}, and {Extremely Likely}}. We will denote this vector as

$$L_F = \gamma(EU, VU, QU, U, QL, VL, EL)$$
$$= [\gamma_i(S_j)\ j = 1{,}7] \ , \tag{4}$$

where $\gamma(S_j)$ is the degree of membership in set j. Consider the vector

$$L_F = \gamma[0, 0.1, 0.3, 0.4, 0.8, 0.4, 0] \ .$$

What does this mean in terms of the likelihood of retained gas and how can this be compared with a screening criterion that is used to classify a tank? The answer to both of these questions requires that we convert the fuzzy set membership vector to a single measure. This operation is referred to as "defuzzification." There are a number of defuzzification operators.[8] We present here a very simple operator that illustrates the general principle.

The output from the defuzzification process is defined using a simple max operator:

$$D(L_F) = \max[\gamma] \ . \tag{5}$$

For our sample vector, $D(L_F) = 0.8$, and this value is associated with the set {Quite Likely}, $S(D(L_F)) \rightarrow$ {Quite Likely}. This association of γ_{LF} with a single set is the defuzzification. The output from the inference model is the specific statement "the likelihood of retained gas is quite likely."

The final step in the evaluation process is to compare this natural language expression to some criterion that classifies the tank. That is, we infer the classification of the tank based on the result from the inductive logic structure. This step is an example of a very simple decision model. Logically, if $S(D(L_F)) \rightarrow$ "quite," "very," or "extremely" {Likely}, then the conclusion of the AR model is that the tank fails the screening process. This is consistent with the design of the logic structure and the definition of the output form discussed earlier. Similarly for the "unlikely" expressions, we conclude that the tank passes and that if $S(D(L_F)) \rightarrow$ {Unresolved}, then the tank requires further study. These statements are simple implications with $S(D(L_F))$ as the antecedent and can be summarized as follows:

1. If $S(D(L_F))$ is associated with the sets {Extremely Likely}, {Very Likely}, or {Quite Likely}, the tank fails the screening and is classified as an FGWL tank.

2. If $S(D(L_F))$ is associated with the sets {Extremely Unlikely}, {Very Unlikely}, or {Quite Unlikely}, the tank passes the screen and is classified as a non-FGWL tank.

3. If $S(D(L_F))$ is associated with the set {Unresolved}, there is insufficient information to assign the tank to the pass or fail sets, and it is classified as an unresolved tank.

A tank with the membership vector of the example would fail the screen and be classified as an FGWL tank.[d] In practice, one might prefer to use different rules for classification based on degree of conservatism considerations. However, note that the classification is based on degree of conservatism considerations, and note that the classification decision rules are independent of the evaluation logic structure.

The aggregate likelihood L_F is a random variable. This is true because most of the inputs to the inductive logic structure, such as the level measurements discussed earlier, are themselves random. The inescapable uncertainty in L_F means that any useful statement about the aggregate likelihood will be statistical in nature. Each of the inputs to the model that is a random variable is represented by a pdf. For example, the pdf for the current waste level is a normal distribution that expresses the uncertainty in the level measurement.[e]

There is no practical way to determine the pdf for L_F directly from the input parameters' pdf's because the total number of inputs is large and because of the nonlinear evaluation operations performed for each implication

[c] An interesting characteristic of the forward-chaining evaluation operation is that the only numerical values for set membership that can exist in an AR model are associated with the input variables. This makes it straightforward to trace the influence of a particular input in the inference model.

[d] Not all vectors are so well-behaved as the examples here, and other defuzzification operators may be required.

[e] If the uncertainty associated with a particular input is epistemic rather than aleatory, then a pdf may be used to represent it.

rule base. Therefore, the statistics for L_F must be obtained from Monte Carlo sampling. The Monte Carlo simulation consists of N trials in which for each trial, all of the input parameters are sampled from their defining pdf's and a complete evaluation is performed with these sample inputs. The immediate output from each trial is an estimate for L_F that is a degree of membership vector γ_{LF}.

At the conclusion of the Monte Carlo simulation, there are seven distinct pdf estimates associated with L_F:

$$\mathrm{pdf}(L_F) = [\,\mathrm{pdf}(\gamma(S_j)\, j = 1,7)]\ , \tag{6}$$

where $\gamma(S_j)$ is the degree of membership in set j. We can derive the required statistics from this vector. If we ask about the value of L_F at some quantile q_i associated with $\mathrm{pdf}(L_F)$, we use the vector

$$q_i^* = [q_i(\gamma(S_j)\, j = 1,7)]\ . \tag{7}$$

This vector contains the degrees of membership at the q_i quantile of the cumulative distribution function (cdf) for each set in the universe of discourse for L_F. For example, at the median ($q_i = 0.5$), assume that $q_i^* = [0.05, 0.08, 0.32, 0.65, 0.41, 0.39, 0.09]$. Each element in this vector is the median degree of membership for the associated fuzzy set in the universe of discourse for L_F. However, note that the vector q^* is not itself the q_i quantile for L_F, $q_i(L_F)$. We must specify how to process the vector to compute $q_i(L_F)$. A natural approach is to define $q_i(L_F)$ using the defuzzification operator of Eq. (5):

$$q_i(L_F) = \max[q_i(\gamma(S_j)\, j = 1,7)]\ . \tag{8}$$

For our example vector, this is the value of 0.65 for {Unresolved}, and the associated statistical statements now associated with the inference model are "The median likelihood for retained gas is unresolved" and "At the median the tank FGWL membership is unresolved."

VII. IMPLEMENTATION OF THE ALGORITHM

The complete AR model for flammable gas screening has 41 primary inputs with an average of three membership functions per input. This requires a total of 40 rule bases with approximately 12 separate rules per rule base. Every rule must be evaluated for each of the Monte Carlo trials. Based on these computational considerations, we chose to implement the screening model as a computer program written in the C programming language. The fuzzy rules are evaluated using a modified version of the commercial software package FuzzyCLIPS.[f] The basic structure of the computer implementation is given as follows:

[f]The FuzzyCLIPS software package is by Togai InfraLogic, Inc.

1. Read in data describing the inputs in the algorithm.
2. Read in the fuzzy rule bases.
3. For each trial in the Monte Carlo simulation,
   (a) select each input from the appropriate distribution
   (b) propagate the membership values through the logic structure using the implication rule bases
   (c) defuzzify the membership values for the aggregate FGWL likelihood L_F
   (d) place L_F in the appropriate bin
   (e) write all values selected from the distributions, intermediate membership values, and crisp value of L_F to a file
4. Create and store the pdf and cdf from the stored values of L_F.
5. Perform postprocessing with a Microsoft Excel spreadsheet to generate quantile statistics and plots of the pdf and cdf.

The AR program was originally run on an IBM 486-66–MHz personal computer (PC) with 16 Mbytes of RAM. Running the entire algorithm for a tank required ~6 h of computing time for 2000 Monte Carlo trials. Running just the barometric pressure logic submodule required between 30 and 80 min/tank depending on the number of level sensors (one to four) that were available for each tank. On a PentiumPro 200-MHz PC with 64 Mbytes of RAM, the run times were reduced by a factor of ~3.

VIII. MODEL TESTING

The FGWL model described here was developed to explore the issues associated with using an AR screening algorithm. Two of the problems considered were the following:

1. A complete tank evaluation where the entire algorithm is used. This was done for two tanks, U-106 and AW-104. U-106 is a single-shell tank with large sludge and salt cake layers. AW-104 is a double-shell tank with more than 3.785 megalitres (1 Mgal) of supernate. Both of these tanks had failed the screening performed by Hodgson et al.
2. partial evaluations using a submodule for the predictor likelihood for all of the tanks on the FGWL that had been flagged previously by Whitney.[9]

The first problem provides insight into the effort required to assemble the input data, the computation time to carry out the Monte Carlo trials needed to generate useful statistics, and the amount of effort necessary to interpret the results. Because the AR model is quite different from the approach used by Hodgson et al., detailed comparisons of the two methods were not considered practical. For the second problem, the submodule in the AR model could be compared to the Whitney model from which it evolved. A detailed discussion of the test results is beyond the scope of this paper, and only some general observations can be given here. For further detail, the reader is referred to Ref. 10.

One significant difference between the AR model results and those of Hodgson et al. for the complete tank evaluations involved the interpretation of the long-term level data. Both of the tanks failed the Hodgson screen based on this model for retained gas volume. However, the quality associated with the data was inferred to be "poor" in the AR model because of the large influence of the correction terms. Thus, as was the case in the example evaluation discussed earlier, the likelihood of a significant quantity of retained gas $L_{\Delta h}$ was inferred to be "unresolved."

The correlation between barometric pressure and waste level was examined in the analysis by Whitney. If a large amount of gas is present, then with some important qualifications that must be neglected here, there should be a strong, negative linear correlation between pressure and level. Whitney examined a large number of tanks and found 37 tanks in which the correlation was found to be strong for at least one level sensor. These tanks were also examined using the barometric pressure submodule in the AR model for level-pressure correlation. In addition to the threshold criterion used by Whitney, additional statistical measures and judgments about data quality were used. Rule bases were also added to take into account relative instrument quality and to resolve differences between the inferences drawn for individual sensors. At the 0.95 quantile, the AR model classified 11 of these tanks as having a strong correlation and inferred that the correlation was weak for two of the tanks. Of the remaining 24 tanks that were classified as unresolved, an additional classification could be made. By observing the cdf for the output likelihood, it was possible to differentiate between tanks in which the data were of reasonable quality but contradictory and tanks in which the data were judged too poor to allow a definitive judgment to be made. Eleven tanks fell in this later category. The capability to make these types of judgments explicit is an important attribute of AR.

IX. CONCLUSIONS

Screening waste tanks for flammable gas is a difficult undertaking. The difficulty arises because of the incomplete understanding of the relevant phenomenology and the need to use partial and apparently contradictory data in models that are themselves incomplete. Our pilot study of the application of the AR methodology to this problem is encouraging. The inductive logic structure and the associated series of implication rule bases make possible a realistic representation of the current state of knowledge. The use of linguistic variables and fuzzy sets provides a way to combine qualitative and quantitative data in a consistent way. The combination of fuzzy and probabilistic approaches in the same model allows for a natural treatment of both uncertainty and ambiguity.

The pilot model showed that the effort required to build an AR model for a relatively complex problem is reasonable and that computational requirements are acceptable. Preliminary analyses with the model clearly demonstrated the value of incorporating qualitative judgments about data and models directly into the screening logic. Differences between the results obtained with the AR model and those obtained previously could often be explained as a consequence of the more detailed inferences about model and data validity included in the rule bases. We conclude that AR is a promising tool for this type of screening problem, as well as for similar assessment tasks in which the data sets and models are incomplete and uncertain, and that further development in this area would be useful.

REFERENCES

1. L. ZADEH, "A Theory of Approximate Reasoning," *Machine Intelligence*, J. HAYES, D. MICHIE, and L. MIKULICH, Eds., Halstead Press, New York (1976).

2. "Operating Specifications for Watch List Tanks," OSD-T-151-00030, Westinghouse Hanford Corporation (1990).

3. J. D. HOPKINS, "Criteria for Flammable Gas Watch List Tanks," WHC-SD-WM-TI-724, Westinghouse Hanford Corporation (1995).

4. J. D. HOPKINS, "Methodology for Flammable Gas Evaluations," WHC-EP-0702, Westinghouse Hanford Corporation (1994).

5. K. M. HODGSON, R. P. ANANTATMULA, S. A. BARKER, K. D. FOWLER, J. D. HOPKINS, J. A. LECHELT, and D. A. REYNOLDS, "Evaluation of Hanford Tanks for Trapped Gas," WHC-SD-WM-ER-526, Rev. 0, Westinghouse Hanford Corporation (1995).

6. S. F. THOMAS, *Fuzziness and Probability*, ACG Press, Wichita, Kansas (1995).

7. T. J. ROSS, *Fuzzy Logic with Engineering Applications*, McGraw-Hill Book Company, New York (1995).

8. H. HELLENDOORN and C. THOMAS, "Defuzzification in Fuzzy Controllers," *Intelligent and Fuzzy Systems*, Vol. 1, p. 109 (1993).

9. P. WHITNEY, "Screening the Hanford Tanks for Trapped Gas," PNL-10821, Pacific Northwest Laboratory (1995).

10. S. W. EISENHAWER, T. F. BOTT, and R. E. SMITH, "An Approximate Reasoning-Based Method for Screening Flammable Gas Tanks," LA-UR-98-1207, Los Alamos National Laboratory (1998).

Stephen W. Eisenhawer (BS, mechanical engineering, Seattle University, 1972; MS, 1974, and PhD, 1977, nuclear engineering, University of Washington) is a technical staff member in the Probabilistic Risk and Hazard Assessment Group at Los Alamos National Laboratory (LANL). His background includes experimental studies in gas/liquid systems, design of high-pressure combustion systems, and the development of methods to represent expert knowledge using formal logical models.

Terry F. Bott (BS, physics, University of Utah, 1970; PhD, chemical engineering, Brigham Young University, 1977) is a technical staff member in the Probabilistic Risk and Hazard Assessment Group at LANL. His background includes nuclear reactor thermal-hydraulics code development, development of mathematical tools for risk and reliability analysis, and the development of methods to represent expert knowledge using formal logical models.

Ronald E. Smith (BS, chemical engineering, Lehigh University, 1986; PhD, chemical engineering, University of Wisconsin–Madison, 1994) is a technical staff member in the Energy and Process Engineering Group at LANL. His background includes control system design, fuzzy logic, neural networks, and reliability studies.