CREATING A BALANCED RESPONSE: AN APPLICATION OF DATA MINING TO DETERMINE WHEN TO SHELTER OR EVACUATE

Michael Boechler[1]
Innovative Emergency Management, Inc.
Baton Rouge, Louisiana

INTRODUCTION

Technological hazards are a fact of modern life. Chemicals and elements that, when properly harnessed, are of enormous benefit to humans, can also be among the deadliest agents when control over them is lost. In these circumstances, emergency managers must attempt to ensure that the public is safe during the response and recovery phases of an accident. Emergency planners can greatly aid in this process. Well-designed plans, carried out in a well-rehearsed manner, can be the difference between public well being and grievous injury.

Following the release of a toxic compound into an inhabited environment, emergency response personnel can employ three strategies. The first is to shelter the population downwind of the hazard in place; the second is to evacuate the potentially affected public (Glickman and Ujihara, 1990, Rogers, et al., 1990). Another option is to employ a combination of shelter-in-place and evacuation. That is, a balanced response can be used to protect the public.

This paper briefly describes how a guidebook for such a balanced response was created for use in an area surrounding a US Army chemical weapons storage depot (IEM, 2000). This was accomplished through the use of a suite of modeling and simulation tools, including a technique that is employed in the growing area of data mining.

Data mining techniques were developed in order to help decision-makers optimize choices based on the analysis of very large data sets (Foss, 1997). In the project described here, a data mining tool was used to characterize a very large set of chemical weapons hazards in a manner that will allow emergency managers, planners and responders to formulate a set of optimal protective actions for several emergency planning zones downwind of a chemical agent release.

BACKGROUND

The Chemical Stockpile Emergency Preparedness Program (CSEPP) is jointly administered by the Federal Emergency Management Agency and the Department of the Army. It was created in order to ensure that the public receives maximum protection (FEMA, 1991) in the case of an accidental release of chemical weapons as they are being destroyed, as called for under the provisions of the Chemical Weapons Convention treaty.

One of the CSEPP sites is located in Anniston, Alabama. It is the area surrounding Anniston Army Depot, which stores 2254 tons of weapons containing chemical agents,

including mustard and the nerve agents VX and GB, also known as Sarin (ANCDF, 1998). Planners and policy makers in the Anniston area have designated 104 emergency planning zones, in which resides a population of approximately 350,000 around the depot. These zones form the basis for planning and response in the event of an accidental release of agent while the chemical weapons are being stored or destroyed.

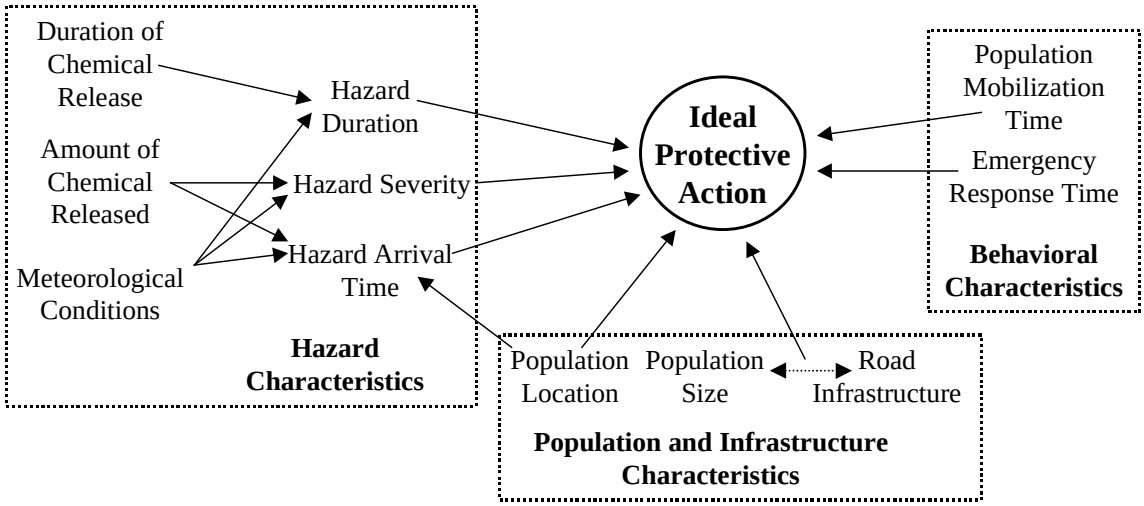
PROBLEM

In order to maximize public protection following the release of a chemical agent, emergency management officials sought to tailor protective action decisions to the characteristics of the release, the meteorological conditions, and the population and infrastructure characteristics of the Anniston area. In accomplishing this, a large number of factors were to be considered. An additional requirement was that the decision support tool be created in hardcopy form so that emergency management staff would not be reliant upon automation, which might be unavailable in the event of a loss of electricity concomitant with the accidental release of the chemical agent (e.g., in the event of an earthquake).

APPROACH

Many factors interact to influence an optimal protective action in response to a toxic vapor release. Three main groups of factors include hazard characteristics, population and infrastructure characteristics, and behavioral characteristics. Some of the ways in which these factors causally interact are shown in Figure 1. For example, the time at which a hazard arrives at a given population center depends on the amount of chemical that is released into the environment, meteorological conditions such as wind speed and direction, season, and atmospheric turbulence, and the location of the population center relative to the site of the release.

Figure 1



In the case of a short duration, quickly moving release, it will be desirable to shelter the downwind population in place, as an evacuation cannot be mobilized in time for the population to avoid the effects of the release. On the other hand, if the release is of long duration and moves slowly over the downwind area, evacuation is the best strategy, as long as the population can be removed from the downwind area prior to receiving a dosage of agent that impairs their ability to complete the evacuation. This will be influenced by the size of the evacuating population relative to the capacity of the road network.

The speed of the emergency response can also affect the decision to shelter or evacuate a given downwind area. A slowly moving, short duration release may require that affected zones be sheltered if there is no time to mount a successful evacuation due to an unavoidable delay in emergency response.

Finally, there are cases in which no single, simple protective action strategy minimizes exposure to the downwind population. For example, populations located close to the source may have sufficient time to evacuate if the road network is relatively clear, but not if the network is congested. In such a case, if the populations located further downwind of the release location can be successfully temporarily sheltered, the strategy that minimizes exposure to the downwind population as a whole may be to temporarily shelter some of the outer zones and evacuate the inner zones.

The number of important factors that determine the best protective action strategy makes planning and decision making during a response impossible to carry out on the basis of ‘back of the envelope’ calculations. What is needed, particularly during an emergency response, is a method of determining the protective action quickly. One approach is to employ a rough rule of thumb or default protective action. However, as discussed above, a pure strategy of shelter-in-place nor evacuation may not provide minimal exposure to the downwind population.

SOLUTION

One way this problem can be addressed is through comprehensive analysis and planning. By combining research on emergency management processes, evacuation time estimates, toxic vapor dispersion and infiltration into structures, and population distribution, computer models can be constructed which reveal combinations of sheltering and evacuation that minimize risk under a wide range of hazard profiles.

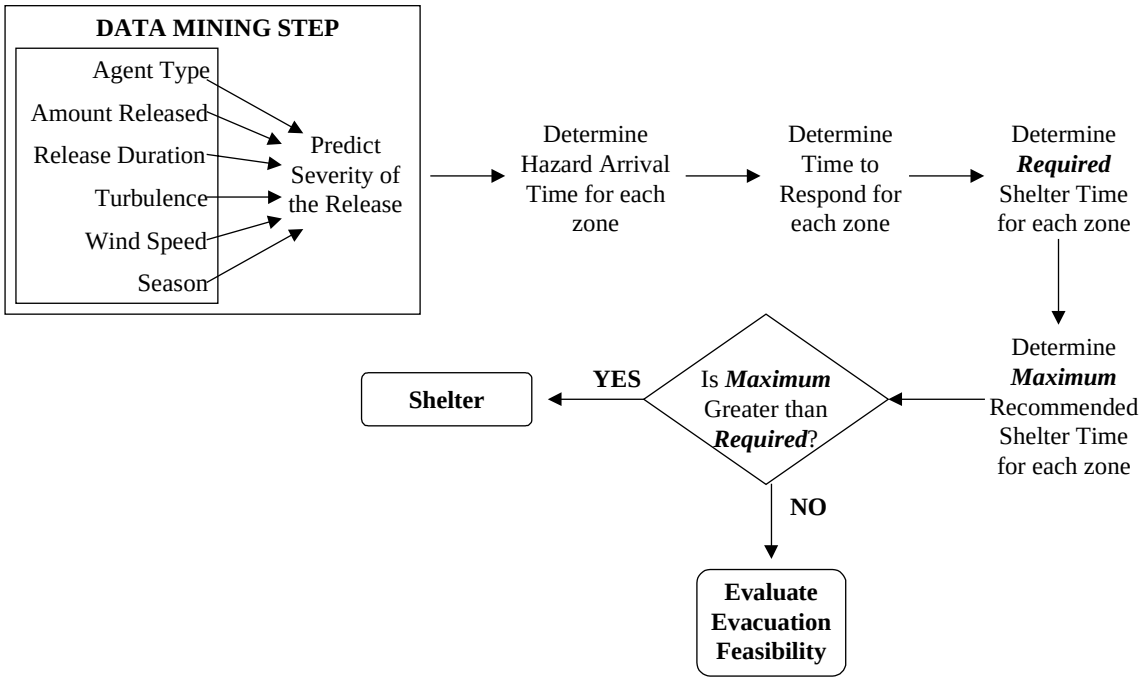
Previous modeling revealed that an evacuation rule of thumb was particularly problematic for the Anniston community. Therefore, the Alabama CSEPP Guidebook was designed around a concept of sheltering populations that could be protected with that protective action strategy in order to ensure that evacuating populations could remove themselves from the downwind hazard area prior to receiving adverse consequences.

A high level diagram of this decision logic concept is shown in Figure 2. A detailed description of this process is available (Wilson, et al., 2000), so will be described only briefly here.

The first step in creating a protective action decision for a given zone downwind of the hazard is to predict the severity of the hazard from the perspective of a downwind population that shelters. This prediction is made with a set of decision rules that were derived using a data mining technique, a topic that will be discussed below. The user then determines the hazard arrival time for each zone. This would be done using a gas dispersion model such as D2PC (IEM, 1993).

Next, the user calculates the time to respond for each zone. This is based on the time that has elapsed since the onset of the accident and the hazard arrival time. Following this, the required shelter time is calculated. This is based on the calculated duration of the hazard once it reaches the zone. The results of the first step are then combined with information on the timeliness of the emergency management response to determine a maximum recommended shelter time. If this time is greater than then required time, the zone is sheltered. If it is not, the final step is to use a lookup table to determine whether evacuation is feasible for the zone.

Figure 2



Data and Modeling

In order to produce this decision support tool, a number of data sources and modeling techniques were employed. The goal of the analysis was to roll up the results of millions of simulations involving the simulated hazards into a set of lookup tables and decision trees that can be used, in combination with some simple arithmetic calculations, to make

a balanced protective action decision. The following is an abbreviated discussion of some of the models that were used and the data they incorporated.

For the first step of the decision logic, a means of predicting the amount of time that a sheltered public could remain in their structures and avoid fatalities was created using information about some of the most important factors that determine the characteristics of a chemical agent plume. These include the type of chemical agent, the amount released, and the duration of the release. This information would be available to the emergency response community from the Army depot in the case of a CSEPP emergency and is available through the use of a maximal credible event, which is a chemical agent source term that has been prepared by the Army in advance of an emergency.

Comprehensive quantitative risk assessment (QRA) studies of potential release sources have been published for each of the CSEPP sites. The QRA for the Anniston site (SAIC, 1997) was used to determine the range of release amounts and times for each type of chemical agent stored at the site. An accident set was then created which included a total of 1728 different possible source terms. These were then combined with 172 different meteorological combinations, each consisting of a wind speed, Pasquill stability class (a measure of atmospheric turbulence), and season (which affects the height of the mixing layer in the atmosphere).

Each source term and meteorological combination was then modeled with a D2PC-compliant gas dispersion model, and the time to exceed a No Deaths dosage of agent was computed for a structure with an air change rate of 1.94 air changes per hour 3 km from the source. This distance corresponds to the populated area nearest the location where the weapons are stored. The value for the air change rate was chosen based on previous work (Karl, et al., 1999, Murray and Burmaster, 1995) that indicated that 9 out of 10 structures in the Anniston area have lower air exchange rates. Thus, this is a fairly conservative estimate of the actual air exchange rate of structures in the area.

The resulting data set of predictor variables and sheltered hazard arrival times was then mined and a set of decision trees was created.

Data Mining

Data mining methodologies have become quite popular in recent years as computing speed and storage capabilities have increased. There are a number of classification and prediction algorithms that fall under the rubric of data mining. These include neural networks, clustering programs, association rules, and classification and regression trees (Aggarwal and Yu, 1999). The typical reason for application of a data mining technique is knowledge discovery within a large database. This usually involves the prediction of some criterion variable value from a number of predictor variable values.

Data mining has successfully been employed in a wide variety of applications, including market segmentation models, credit risk analysis, and medical diagnosis. The classification and regression tree (CART) approach used in this project was developed within the context of medical decision-making (Breiman, et al., 1984) in order to provide

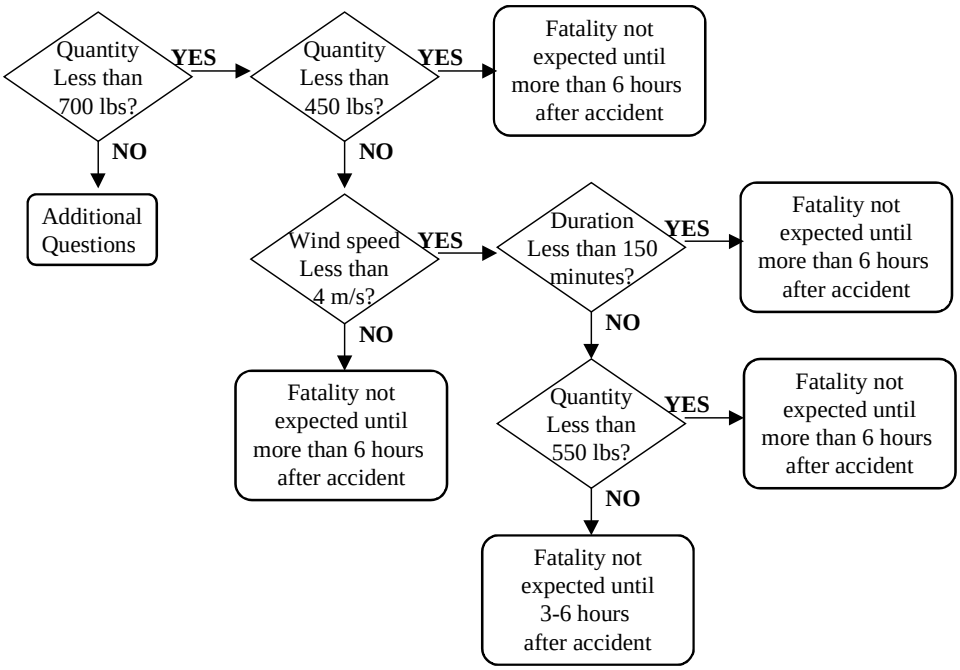
medical practitioners with an easily used method of predicting the outcomes of various treatment options.

Perhaps the most common forms of statistical prediction involve calculating a regression equation, which is then used to predict an outcome of interest. One problem with this approach is that making a prediction involves computing a series of multiplications and additions, which can become time-consuming and prone to error if they must be done without the aid of a computer. This makes the regression approach difficult to employ in a hardcopy format, particularly under the stressful conditions present during an emergency response. The CART approach, in contrast, produces a decision tree that can form the basis of simple if-then rules that can be used to predict an outcome of interest.

For the Guidebook project, the variables shown in the box in Figure 2 formed the basis of the CART analysis, which was carried out using a commercially available statistical package (Mathsoft, 1999). The resulting decision rules allow emergency managers and planners to predict the amount of time that persons could stay sheltered and avoid fatality on the basis of information available from the Army depot emergency operations center very shortly after the start of an accident.

A notional example of a decision tree produced by the CART process is shown below in Figure 3. For ease of use, these diagrams were converted into a series of tables. Each entry in the tables is a question. The user then moves through the tables, answering the questions until a prediction is reached regarding the expected outcome if the downwind population shelters.

Figure 3



DISCUSSION

Making the correct protective action decision in response to a release of hazardous materials involves accounting for a large number of important factors. The number of interactions among these factors can quickly lead to millions of plausible accident scenarios. Planning based on such an astronomical number of possible scenarios is prohibitive. However, if these factors and their interactions are not taken into account in some manner, the emergency manager runs the risk of making a decision that is less than optimal. Some middle ground between simple, sub-optimal rules of thumb and paralysis by analysis must be found in order to make the best decisions possible under stressful conditions. The project described in this paper was conducted with the goal of finding that middle ground.

The approach taken was to use a set of models to simulate toxic vapor dispersion and infiltration into shelters, emergency response system behavior, evacuation network characteristics and traffic flow, and population distribution and mobilization in response to a chemical event. These models produced quantities of data that are not directly useable by emergency planners and managers. Therefore, a decision structure was created that would allow decision-makers to take advantage of the simulation results.

The decision structure underlying the Guidebook consists of lookup tables and some simple calculations—nothing more difficult than the subtraction of 3-digit numbers and the comparison of pairs of numbers. Training in the use of the Guidebook showed that, after a 30-minute training session, planners were able to work through it and arrive at a correct answer in less than 5 minutes. In the future, this structure will be able to provide shelter-in-place and evacuation decisions for the Alabama CSEPP emergency planning zones even more rapidly. Work is currently being undertaken to create an automated version of this methodology to complement and extend the capabilities of the paperback version of the Guidebook.

There are many other aspects of the Guidebook project that were not within the scope of this paper, which focused on the application of a data mining methodology to one aspect of the overall problem. The CART methodology yielded a set of decision trees that can be used without any requirement for calculating additional numbers in order to make a prediction regarding the time that downwind populations can shelter before receiving a lethal dose of chemical agent. The prediction is made using six pieces of information that are available to emergency personnel within minutes of an accident and include some of the most important determinants of a plume’s behavior. This prediction is then used as a basis for deciding if it is possible to shelter the downwind area, thus clearing the road network for other populations that may not be able to shelter in-place and avoid lethality.

Although the Guidebook was developed specifically for the Alabama CSEPP community, the process used to create it is general in nature. It could be used for almost any kind of acute toxic hazard. It requires an adequate plume model for the type of chemical release, information regarding the traffic network around the source of the release, and the possible meteorological conditions for the area. In addition, information regarding the

distribution of the surrounding population and characteristics of warning and response systems is required.

It may appear that such data requirements would prohibit the development of such a system for many jurisdictions. However, much of the required data can be obtained cost-effectively through public sources and through focused tabletop exercises—and the benefits of a balanced response can be very great. A calculation of the risk reduction due to employing the balanced response in the Anniston area showed that the reduction was most commonly 75% with respect to shelter-in-place for the emergency planning zones (Lemcke, et al., 2000).

REFERENCES

Aggarwal C C and Yu P S. "Data Mining Techniques for Associations, Clustering and Classification." In Zhong, N and Zhou L eds. *Methodologies for Knowledge Discovery and Data Mining. Lecture Notes in Artificial Intelligence 1574.* Berlin: Springer, 1999.

Anniston Chemical Agent Disposal Facility, Presentation for 1998 Joint World Wide Chemical Conference XVI & NBC Operations Symposium, Ft. McLellan, Alabama. <http://www.dtic.mil/ndia/chemical/demilitarize.pdf>, 1998.

Breiman L; Friedman J; Olshen R; and Stone C. *Classification and Regression Trees*. New York: Chapman and Hall, 1984.

Glickman, T S and Ujihara, A M "Deciding between in-place sheltering and evacuation in toxic vapor cloud emergencies." *Journal of Hazardous Materials* 23 (1990), pp. 57-72.

Federal Emergency Management Agency, *CSEPP Policy Paper 1: Definition of Maximum Protection*, <http://www.csepp.apgea.army.mil/biblio/policypaper/paper1.pdf>, 1991.

Foss, B A "Knowledge Discovery – A Control Theory Perspective." In Komorowski, J and Zytkow, J eds. *Principles of Data Mining and Knowledge Discovery. Lecture Notes in Artificial Intelligence 1263.* Berlin: Springer, 1997.

Innovative Emergency Management, Inc., *Reference Manual: D2PC and Hazard Analysis*, Baton Rouge, LA: Innovative Emergency Management, Inc., 1993.

Innovative Emergency Management. *Alabama CSEPP Protective Action Recommendation Guidebook.* Baton Rouge, LA: IEM, Inc., 2000.

Karl B, Lemcke T, Martin S A, Morgan L E, Reichardt T, Wilson K., *Public Protection Criterion: a Publication Review for Alabama CSEPP*, Baton Rouge, LA: Innovative Emergency Management, Inc., IEM/TEC99-041, 1999.

Lemcke T; Martin S A; Wilson K. *Scenario and Critical Evaluation Element Report*, Baton Rouge, LA: Innovative Emergency Management, Inc., IEM/TEC00-024, 2000.

MathSoft, *S-Plus Guide to Statistics*, Vol. 1. Seattle: MathSoft Data Analysis Products Division, 1999.

Murray, D M and Burmaster D E, "Residential Air Exchange Rates in the United States: Empirical and Estimated Parametric Distributions by Season and Climatic Region." *Risk Analysis*, 15 (1995), pp. 459-465.

Rogers, G O; Watson A; Sorensen J; Sharp, R; and Carnes S, Evaluating Protective Actions for Chemical Agent Emergencies, for Program Manager for Chemical Demilitarization, and Federal Emergency Management Agency, ORNL-6615, 1990.

Science Applications International Corporation, *Anniston Chemical Agent Disposal Facility Phase 1, Quantitative Risk Assessment*, Revision 1, Abingdon, MD: SAIC, 1997.

Wilson J K; Boechler M; Boyle B; Cochran B; Coles C; Martin S A; Misra R; Roussel L, *Technical Addendum to Alabama CSEPP Protective Action Recommendation Guidebook*. Baton Rouge, LA: IEM, Inc., IEM/TEC00-023, 2000.

---

[1] Michael Boechler, PhD, is a Senior Systems Analysis with Innovative Emergency Management. After receiving his PhD in biological psychology from the State University of New York – Albany in 1992, he worked in the field of healthcare as a statistician and as a Director of Information Systems. For further information on this paper contact: Michael Boechler, PhD, IEM, Inc., 8555 United Plaza Blvd. Ste. 100, Baton Rouge, LA 70809. Phone: 225.952.8274; Fax: 225.952.8122; E-mail: <michael.boechler@ieminc.com>.

Author wishes to thank J. Krause Wilson and Summer Doucet for helpful comments on a draft of this paper.