LA-UR-95- 95-1117
TSA-11-95-R114

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

TITLE: APPROACHES TO IMPLEMENTING DETERMINISTIC MODELS IN A PROBABILISTIC FRAMEWORK

Conf-950740--21

AUTHOR(S): Dale V. Talbott

SUBMITTED TO: 1995 ASME/JSME Pressure Vessels and Piping Conference
July 23-27, 1995
Honolulu, Hawaii

Los Alamos

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

DISCLAIMER

APPROACHES TO IMPLEMENTING DETERMINISTIC MODELS IN A PROBABILISTIC FRAMEWORK

Dale V. Talbott
Los Alamos National Laboratory
Los Alamos, New Mexico

ABSTRACT

The increasing use of results from probabilistic risk assessments in the decision-making process makes it ever more important to eliminate simplifications in probabilistic models that might lead to conservative results. One area in which conservative simplifications are often made is modeling the physical interactions that occur during the progression of an accident sequence.

This paper demonstrates and compares different approaches for incorporating deterministic models of physical parameters into probabilistic models: parameter range binning, response curves, and integral deterministic models. An example that combines all three approaches in a probabilistic model for the handling of an energetic material (i.e. high explosive, rocket propellant,...) is then presented using a directed graph model.

NOMENCLATURE

$\Phi(x)$ Heaviside step function. Equal to 0 if x is negative, and 1 otherwise.

μ Mean

σ Standard Deviation

$P(x)$ Probability of x.

$P(x|y)$ Probability of x given y.

rnd A random number between 0 and 1.

v_{lower} Velocity at a lower bin boundary

v_{upper} Velocity at an upper bin boundary

INTRODUCTION

Deterministic models describe the behavior of a system with certainty. For a given set of inputs a deterministic model will produce a single answer. Examples of deterministic models are finite element or finite difference models that always give the same answer each time the model is run. In the real world, the outcome of a system is generally not so well defined. Inherent randomness in material properties, geometry, initial conditions, and other parameters lead to randomness in the outcome. Probabilistic models account for this natural randomness as well as modeling uncertainties that arise as a result of simplifying assumptions and incomplete knowledge.

Probabilistic models are used to help understand real world systems. However, just as a scale model of an aircraft that is designed for use in a wind tunnel does not have all of the details such as engines, control systems, and radios of an actual aircraft, probabilistic models cannot contain all of the details of the physical interactions that may occur in the real world. Probabilistic models should contain enough detail to answer the question being asked, just as the wind tunnel model of the aircraft contains adequate detail to answer specific questions about the aircraft design.

This paper takes a physical problem, modeling the probability of an energetic material reacting given an impact, and compares various approaches for modeling it probabilistically. This is done to give the reader some insight into the strengths and weaknesses of each solution technique.

SIMPLE PROBLEM

A simple problem is used to illustrate a case where physical responses need to be modeled probabilistically. The problem is: given an energetic material such as a high explosive or missile propellant impacting a surface, what is the probability of an energetic reaction of the material. An event tree depicting the problem is shown in figure 1.

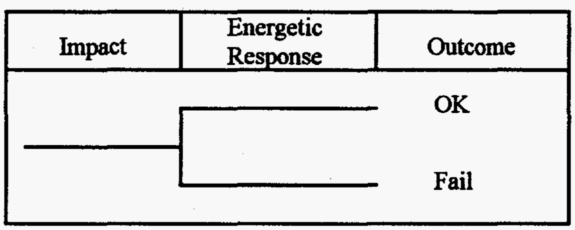

Figure 1. Event Tree for Simple Problem

Even for this simple problem there are many physical parameters describing the impact that are important in determining if the material will react. Some of these parameters are: material configuration, impact surface roughness, impact surface hardness, orientation, mass, velocity, impact area, electrostatic charge and temperature. Modeling the influence of each parameter, accounting for all of the inherent randomness in the materials and impact conditions as well as accounting for other modeling uncertainties can quickly become an overwhelming task.

For this simple example problem, the probability of a reaction is assumed to depend only on the impact velocity. The initial conditions of this problem are modeled probabilistically because of variation in the impact velocities due to natural randomness. There may also be uncertainty as to where the velocity distribution lies due to incomplete knowledge. The family of velocity distributions shown in figure 2 includes distributions for the lower bound, expected value, and upper bound showing both the natural randomness and the uncertainty in the location of the velocity distribution.

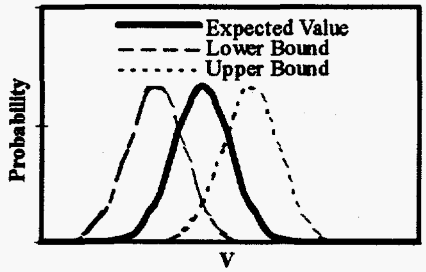

Figure 2. Impact Velocity Distributions

The model must predict if a reaction occurs given an impact velocity. The response model should reflect the inherent randomness in the material as well as any additional uncertainties due to incomplete knowledge. The probability of the material responding at a given impact velocity is represented by a family of curves as shown in figure 3. The shape of the curve represents the random variation within the energetic material and the expected value, lower bound, and upper bound curves represent the uncertainty in the actual location of the response curve.

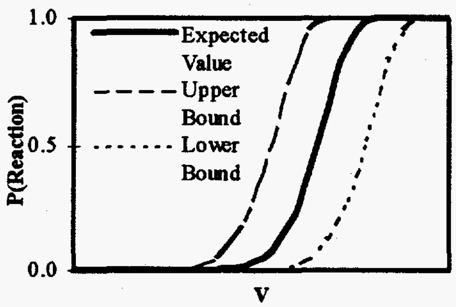

Figure 3. Response Curves

To facilitate an analytic solution of the simple problem, all of the uncertainty is assumed to be captured in a single velocity distribution and a single response curve. This assumes that the curves are well defined and no uncertainty as to the location of the curve exists. This allows the results generated by parameter range binning and by a Monte Carlo simulation (Sobol', 1994) with response curves to be compared directly to an analytic solution. For the simple problem the impact velocities are assumed to be distributed normally with a mean of 170 and a standard deviation of 30. The material is assumed to respond 5% of the time at a velocity of 200 and 95% of the time at a velocity of 300. A response curve is approximated using a Weibull function.

Analytic Solution

To compute an analytic solution, functions describing the velocity distribution as well as the response curve must be developed. The velocity distribution, a normal distribution with a mean of 170 and standard deviation of 30, is calculated using equation 1.

$$P(V) = \frac{1}{\sigma\sqrt{2A\pi}} e^{-\left(\frac{(v-\mu)^2}{2\sigma^2}\right)} \tag{1}$$

A plot of this function is shown in figure 4.

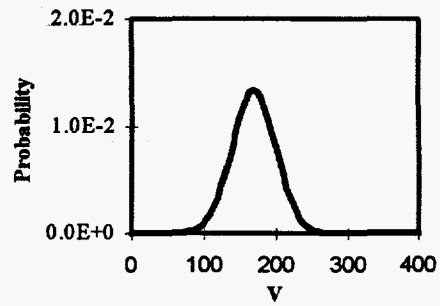

Figure 4.

A Weibull curve is fit to the known response values using equation 2.

$$P(Reaction|V) = 1 - e^{-\left(\frac{v}{\alpha}\right)^{\beta}} \tag{2}$$

where,

$$\beta = \frac{\ln\left(\frac{\ln(1-0.05)}{\ln(1-0.95)}\right)}{\ln\left(\frac{200}{300}\right)} \quad \text{and} \quad \alpha = \frac{200}{\left(-\ln(1-0.05)\right)^{\left(\frac{1}{\beta}\right)}}$$

This response curve is shown in figure 5.

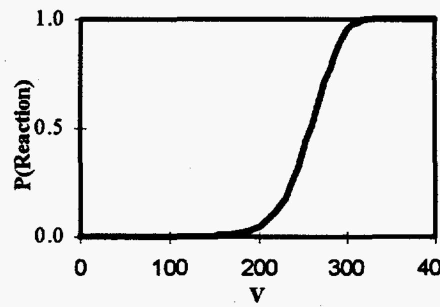

Figure 5. Response Curve

Given the distribution of impact velocities and the response curve, the expected value of the probability of a reaction can be calculated analytically using equation 3.

$$P(Reaction) = \int_0^{\infty} \left(P(Reaction|V) \times P(V) \right) dv \tag{3}$$
$$= 0.029$$

A standard deviation is calculated for the associated probability density function using equation 4.

$$\sigma = \sqrt{\int_0^{\infty} \left(P^2(Reaction|V) \times P(V) \right) dv - P^2(Reaction)} \tag{4}$$
$$= 0.055$$

The standard deviation is large (relative to the mean) due to a highly skewed probability distribution.

The probability that a given velocity contributed to the reaction is a measure of the importance of a given velocity. This is an indication of what velocity range contributes most to the risk of a reaction. A density function for the probability of a velocity given a reaction can be calculated by applying generalized Bayes theorem (Sveshnikov, 1978) as given in equation 5.

$$P(V|Reaction) = \frac{P(Reaction|V) \times P(V)}{\int_0^{\infty} \left(P(Reaction|V) \times P(V) \right) dv} \tag{5}$$

A plot of this function is shown in figure 6.

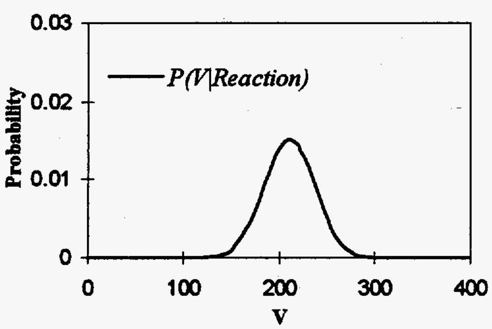

Figure 6. Velocity Importance Distribution

Another measure of interest is the cumulative probability of a reaction given a velocity. This is used to determine the risk reduction obtained if all velocities higher than a given velocity are eliminated. This is calculated by integrating the probability of a reaction from 0 to the velocity.

$$CDF(Reaction|V) = \int_0^{v} \left(P(Reaction|V) \times P(V) \right) dv \tag{6}$$

A plot of this function is shown in figure 7.

It can be seen from figure 7 that most of the probability of a reaction is a result of velocities above 150. It also shows that eliminating impacts with a velocity above 200 (≅ 15% of the impacts) reduces the risk of a reaction by about 65%.

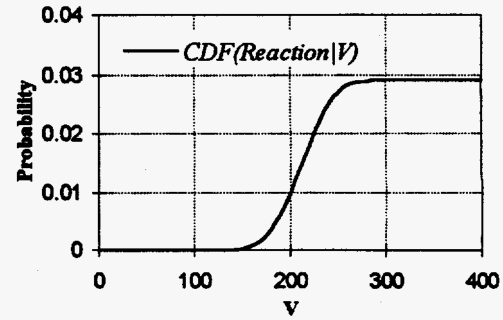

Figure 7. Reaction CDF

Parameter Range Binning

The next approach used to solve the simple problem is parameter-range binning. This approach is similar to damage-state binning which is a technique that has been used successfully in nuclear reactor PRAs for many years. Damage state binning defines accident bins and groups accident sequence with similar endstates into a bin. An advantage of binning is that a probabilistic model can be developed that expresses the physical parameters as probability distributions assigned to each bin.

Parameter-range binning is similar to damage-state binning except that it breaks distributions of physical parameters into bins and treats all values within the parameter range of a bin the same. This makes the models compatible with most common PRA techniques such as fault tree linking, event tree linking or directed graph models without the PRA software being required to track physical parameters through a sequence. It also allows parameter importance to be calculated based on the bin ranges using standard event tree importance measures.

This technique is applied to the simple problem by breaking the velocity distribution into discrete ranges and assuming that the probability of a response is the same for any velocity within a range. The event tree in figure 7 demonstrates the approach.

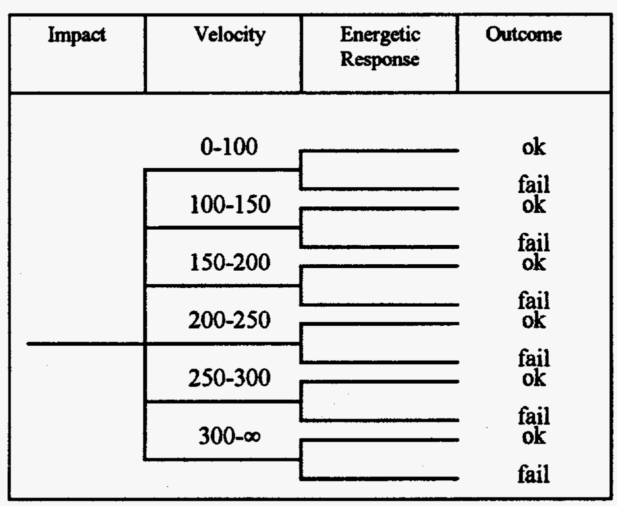

Figure 8. Parameter Range Binning

For the simple problem, the probability of being in each velocity range is determined based on the velocity distribution shown in figure 4. The probability of impacting at a velocity in a given bin can be determined either by integration

$$P(Bin) = \int_{v_{lower}}^{v_{upper}} P(V)dv \tag{7}$$

or by Monte Carlo simulation

$$P(Bin) = \frac{\sum_{i=1}^{n}\Phi(v_{upper} - v_i) - \sum_{i=1}^{n}\Phi(v_{lower} - v_i)}{n} \tag{8}$$

where v_i are random samples of the velocity distribution. As expected, for large sample sizes the total probability for any velocity bin is very similar using either approach. Table 1 lists the bin probabilities calculated using each method.

Table 1. Bin Probabilities for n=500

Bin	Integration	Monte Carlo
P_{0-100}	0.010	0.010
$P_{100-150}$	0.243	0.224
$P_{150-200}$	0.589	0.600
$P_{200-250}$	0.155	0.164
$P_{250-300}$	0.004	0.002
$P_{300-\infty}$	0.000	0.000

The next step is to determine the probability of a reaction given a bin. Six different approaches for determining the probability of a response within a bin are compared. These are calculating the response at the upper bin boundary, at the center of the bin, calculating a mean response for the bin, uniform distribution across the bin, and calculating the total probability for a reaction in each bin based on the velocity distribution and the response curve. Comments on the applicability of each approach are made assuming that the probability of a reaction increases from the lower bin boundary to the upper bin boundary as it does in the simple problem.

Response At Upper Bin Boundary

This is the easiest and also the most conservative approach to using bins for this type of problem. It assumes that the response of every sample in a bin is equal to the response that would be seen at the upper (most conservative) boundary of the bin. As shown in equation 9, the probability of a reaction for each sample placed in the bin is assumed to be the same as that of a sample at the upper bin boundary.

$$P(Reaction|Bin) = P(Reaction|v_{upper}) \tag{9}$$

This type of bin response may be appropriate in cases where detailed deterministic calculations or tests can be run to determine the response at the bin boundaries but no inference can be made to the shape of the response curve between bin boundaries. This approach is simple to apply but can lead to conservative results. Care should be taken to estimate the level of conservatism.

Response At Center Of Bin

This approach applies the response of a sample at the center of a bin (i.e. 225 for a bin from 200 to 250) to the entire bin. This is less conservative than taking the response at the upper bin boundary. Equation 10 is used to calculate the probability of a reaction for a bin.

$$P(Reaction|Bin) = P\left(Reaction\left|\frac{v_{lower} + v_{upper}}{2}\right.\right) \tag{10}$$

Using equation 10 to generate bin probabilities produces a good estimate for many problems but it should be noted that in certain circumstances it can produce a nonconservative result.

Mean Response For Bin Boundaries

Taking the average of the response at the upper bin boundary and the response at the lower bin boundary is another way of approximating the probability of a response given a bin. This estimate is also less conservative than taking the response at the upper bin boundary. Equation 11 is used to estimate the probability of a response given a bin.

$$P(Reaction|Bin) = \frac{P(Reaction|v_{lower}) + P(Reaction|v_{upper})}{2} \tag{11}$$

This approach also produces a good estimate for many problems. It should be noted however that like taking the response at the center of the bin, this approach can produce a nonconservative result in some cases.

Uniform Distribution Across The Bin

This method uses the response curve to develop a probability distribution for each bin assuming that the velocity samples are evenly distributed across a bin. This technique may be appropriate when the response curve is known but the velocity distribution is not. Equation 12 calculates the bin probability by integrating the response curve across the bin.

$$P(Reaction|Bin) = \int_{v_{lower}}^{v_{upper}} \frac{P(Reaction|V)}{v_{upper} - v_{lower}}dv \tag{12}$$

Equation 13 calculates the bin probability using Monte Carlo sampling.

$$P(Reaction|Bin) = \frac{\sum_{i=1}^{n} P(Reaction|v_{lower} + \mathrm{rnd}(v_{upper} - v_{lower}))}{n} \tag{13}$$

Equations 12 and 13 give equivalent results for large n. This approach is less conservative than choosing the response at the upper bin boundary but can be nonconservative for some velocity distributions.

Velocity Distribution And Response Curve Integration

This approach takes all of the information about the velocity distribution and the response curve and calculates the probability of a response given the bin. This requires detailed information about the velocity distribution as well as the response curve. The probability of a bin is calculated with equation 14.

$$P(Reaction|Bin) = \frac{\int_{v_{lower}}^{v_{upper}} P(Reaction|V) \cdot P(V)dv}{\int_{v_{lower}}^{v_{upper}} P(V)dv} \quad (14)$$

As expected, this produces a very good answer but requires information about both the velocity distribution and the response curve.

Binning Results

Results for the simple problem generated using each of these approaches and the bins shown in table 1 are listed in table 2.

Table 2. Simple Problem Results Using Binning

Response Determined At	Reaction Probability
Upper Bin Boundary	.095
Center of Bin	.035
Mean Response for Bin	.053
Uniform Distribution Across Bin	.041
Integrated Curves	.029

The results range from the guaranteed conservative results obtained by taking the response at the upper bin boundary to the very accurate integration of the velocity distribution and the response curve across the bin. It should also be remembered that if conservative estimates are made in successive events in an accident sequence, the conservatism is multiplicative and can easily generate very conservative results.

Monte Carlo Simulation Without Binning

Another approach to solving the simple problem is by direct application of the Monte Carlo method (Sobol', 1994) without binning. Using the Monte Carlo method, the problem is solved by sampling the velocity distribution and calculating the probability of a response for each sample. The probability of a reaction for any given velocity is defined by the Weibull function $P(Reaction|V)$ and velocities are sampled from the function $P(V)$. Equation 15 is used to calculate the expected value for the probability of a reaction.

$$\mu = \frac{\sum_{i=1}^{n} P(Reaction|v_i)}{n} \text{ where,} \quad (15)$$

v_i are random samples of the velocity distribution. For this exercise, 500 samples of the velocity distribution were used to develop a probability distribution for the response. An expected value, μ, of 0.029 and a standard deviation, σ, of 0.0501 were calculated. These agree well with the results of the analytic solution for μ of 0.029 and σ of 0.054. Figure 9 shows a histogram of the probability density function for a reaction generated by the Monte Carlo simulation.

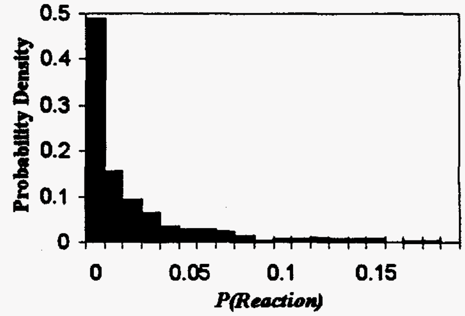

Figure 9.

The highly skewed curve causes the standard deviation to be high relative to the expected value. Figure 10 shows the velocity importance in the form of a histogram based on the samples from the simulation.

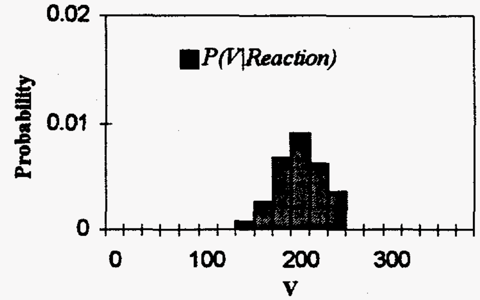

Figure 10

This compares well with the plot shown in figure 6 developed in the analytic solution. The difference in the probability between figure 6 and figure 10 is an artifact of the ranges selected for the histogram and doesn't reflect on the accuracy of the solution.

The last measure that will be compared to the analytic solution is the cumulative probability of a reaction at a velocity. Figure 11 shows the cumulative probability of a reaction given a velocity calculated based on the results of the Monte Carlo simulation. This plot also compares very will with the corresponding plot from the analytic solution presented in figure 7.

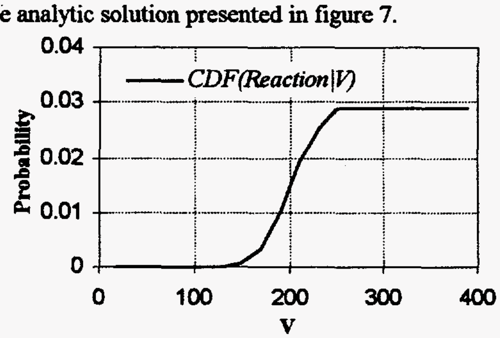

Figure 11

INTEGRATED DETERMINISTIC MODELS

Many problems of interest are much more complex than the simple example problem and are often dependent upon multiple physical parameters. For these cases, simple response curves with respect to one variable, like the one used in the example problem, are not adequate. Instead, fast running deterministic models that can take a set of initial conditions and model the outcome are needed. By varying model parameters, these fast running models can be used in a Monte Carlo simulation to determine probability distributions for the outcomes.

A number of techniques exist for developing fast running models for use in Monte Carlo simulations. These techniques fall into two categories; theoretical models based on physical laws and empirical models based on experimental data or the results of more detailed theoretical models such as finite element or finite difference models.

Fast running theoretical models are developed using engineering tools like spring mass models for impact forcing functions or lumped capacitance models for thermal transfer. The amount of simplification that must be employed in a theoretical model is dictated by the detail with which the system is known, the number of runs required, and the speed with which the model runs. As the processing speed of computers increases, the level of detail that can be employed directly in fast running theoretical models will also increase.

Empirical models use curve fitting or look up tables to determine the response to the input parameters. Curve fitting techniques develop multidimensional response surfaces using data points obtained experimentally or by running detailed deterministic models. The equations defining the response surfaces are then used to determine the probability of a response. Look up tables use the data points directly and either pick the most appropriate data point from the look up table or they interpolate between points to determine a response.

A fast running theoretical model of tool drops in a silo was developed by Applied Research Associates (ARA) as part of the Minuteman III weapon system safety assessment (Sues, 1994) performed by the Defense Nuclear Agency. This model contained a full three dimensional model of the important silo geometry and tracked the path of a tool drop to determine impact points for tools dropped in a silo. The model takes the geometry and physical characteristics of the tools and parameter distributions describing the kinematics of the tool as it is dropped. Based on this information and the silo geometry, the model calculates all of the impact points in the silo and the impulse imparted as a result of each impact. This model runs very quickly, allowing hundreds of tool drops to be simulated in minutes on a personal computer. The model was incorporated into a Monte Carlo simulation using software developed by ARA.

Without this model, conservative assumptions would have been made in developing the risk model. These conservative assumptions would have given an answer that could have misled analysts and potentially caused them to give poor advice to decision makers. Identifying the possible impact of the conservatism and developing a fast running, detailed deterministic model resulted in a realistic and defensible result.

RESPONSE CURVES WITH UNCERTAINTY

The family of curves in figure 3 show the uncertainty of the response due to the natural randomness in the material as well as the uncertainty as to where the response curve actually lies. The top curve shows the upper bound of possibilities, the middle curve the best estimate, and the bottom curve a lower bound on where the response curve could be. One approach to handling this is to develop a response distribution that has a lower bound at the lower curve and an upper bound at the top curve for each value of the parameter. This distribution can be sampled to determine the probability of a reaction given a value of the parameter. The simplest distributions to apply are a uniform distribution between the upper and lower bounds or a triangular distribution with the point on the expected value as the mode.

Figure 12 shows a family of response curves. For each sample of V, a probability density function is developed and then sampled.

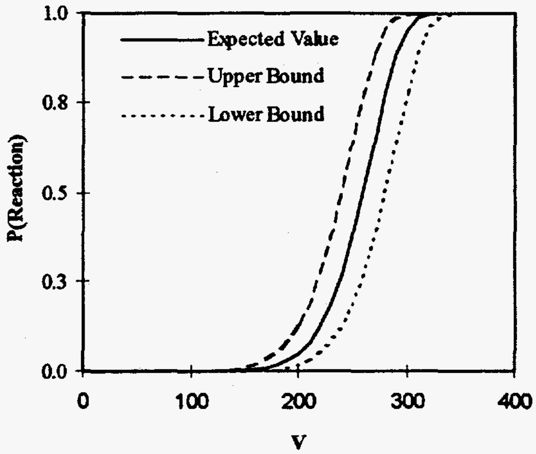

Figure 12.

If a triangular distribution is assumed between the curves the resulting probability distribution for V equal to 200 is shown in figure 13. Sampling the distribution in figure 13 accounts for the uncertainty in the location of the response curve.

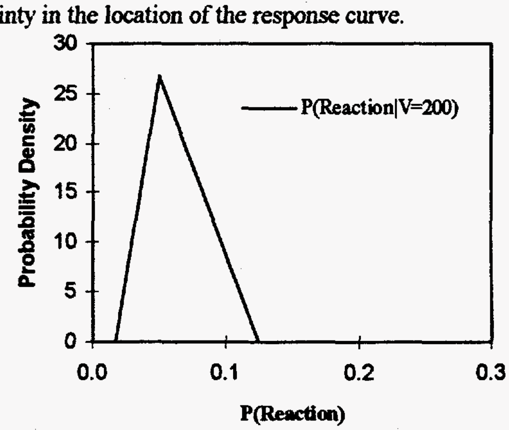

Figure 13.

This technique has been used to develop probabilistic models of the response of energetic materials to impacts. It allows the probabilistic analyst to capture the randomness present in experimental data on energetic materials as well as incorporate an estimate of the uncertainty about the applicability of the data to the problem being modeled.

COMBINED EXAMPLE

The techniques presented thus far represent a set of tools that can be combined to develop probabilistic models of complex accident sequences. One example problem that combines the techniques is a tool drop that could bounce off of a handling fixture and impact an energetic material. Binning is used to differentiate between different tool types, integrated deterministic model track the motion of the tool being dropped, and response curves with uncertainty determine the response of the energetic material upon impact. This model is represented by the directed graph in figure 14.

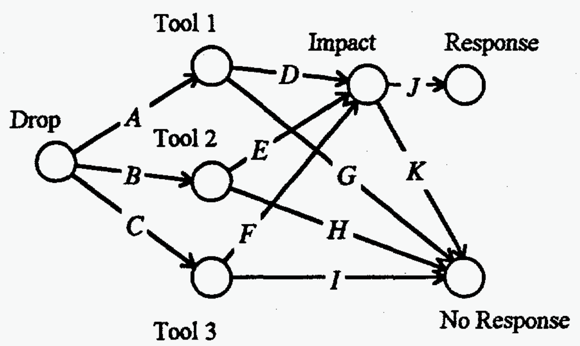

Figure 14. Directed Graph of Combined Example

The model is quantified by developing the minimal path sets through the graph in terms of arcs and performing a Monte Carlo simulation using the path sets. The path sets of interest are those going from the node labeled "Drop" to the node labeled "Response." Equation 16 is a Boolean representation of the path sets in terms of arcs.

$$\textit{Response Path Sets} = A \cdot D \cdot J + B \cdot E \cdot J + C \cdot F \cdot J \qquad (16)$$

The arcs A, B, and C represent bins of tool types. Each bin is assigned a probability distribution that represents the fraction of the time that a particular tool is dropped given a tool drop. Integrated deterministic models are required to determine the probability distributions for D, E, and F. These models take distributions of the initial kinematic parameters associated with the tool drop bin and determine if the tools trajectory causes an impact with the energetic material. They also determine the appropriate parameters to evaluate the response given an impact. These parameters might be the impulse imparted to the material and the effective impact area. Response curves with uncertainty then use the parameters generated by the integrated deterministic models to develop probabilities of a reaction. These probabilities are used to quantify J. The arcs G, H, I, and K are quantified by taking the complement of the probabilities assigned to D, E, F, and J respectively.

LANSE

LANSE is a program developed at Los Alamos that performs Monte Carlo simulations using a directed graph methodology that tracks physical parameters through an accident sequence, and allows deterministic models and response curves to be integrated directly into probabilistic models. LANSE has the capability to evaluate the combined example as presented.

LANSE models the combined example by first evaluating the minimal path sets through the graph and then performing a Monte Carlo simulation using the path sets. Linking rules are used to assign the proper models and parameters to each sequence. A dynamic parameter spreadsheet is updated, as each sample is taken in a Monte Carlo simulation, to track physical parameters through a sequence. Integrated deterministic models, developed as dynamic link libraries, pass parameters back and forth from the parameter spreadsheet for each sample in the simulation. LANSE also has the capability to develop response curves with uncertainty by simply entering known values along the curves. It uses a parameter in the spreadsheet to develop a probability density function based on the response curves. This probability density function is sampled to evaluate the probability of a response given an impact.

SUMMARY AND CONCLUSIONS

Probabilistic risk assessment analysts should give decision makers the best possible information on risk and on ways to reduce risk. Several techniques are presented in this paper to refine the results of a risk assessment by including more detail about the physical processes at work in an accident sequence. Employing techniques like those presented can produce more realistic results and will serve to increase the credibility of the PRA process among decision makers.

ACKNOWLEDGMENTS

The author wishes to thank Jane Booker for assistance with some of the statistics; Marty Fuentes for discussions on binning; Greg Buntain and Bill Tao for information on energetic materials; David O'Brien, Tom Bement and Paul Kuam for ideas on response curves with uncertainty; Michael Caffery for help in incorporating response curves into LANSE; and finally Desmond Stack and Sharif Heger for encouragement and help on the application of directed graphs.

REFERENCES

Sobol', Ilya M., 1994, "A Primer for the Monte Carlo Method," CRC Press, Boca Raton, FL.

Sveshnikov, A. A., 1978, "Problems In Probability Theory, Mathematical Statistics And Theory Of Random Functions," Dover Publications, Inc., New York, NY.

Sues, Robert H, et al., 1994, "Minuteman III Weapon System Safety Assessment: Weapon States 8 and 9," Pre-release draft, Applied Research Associates, Raleigh NC.