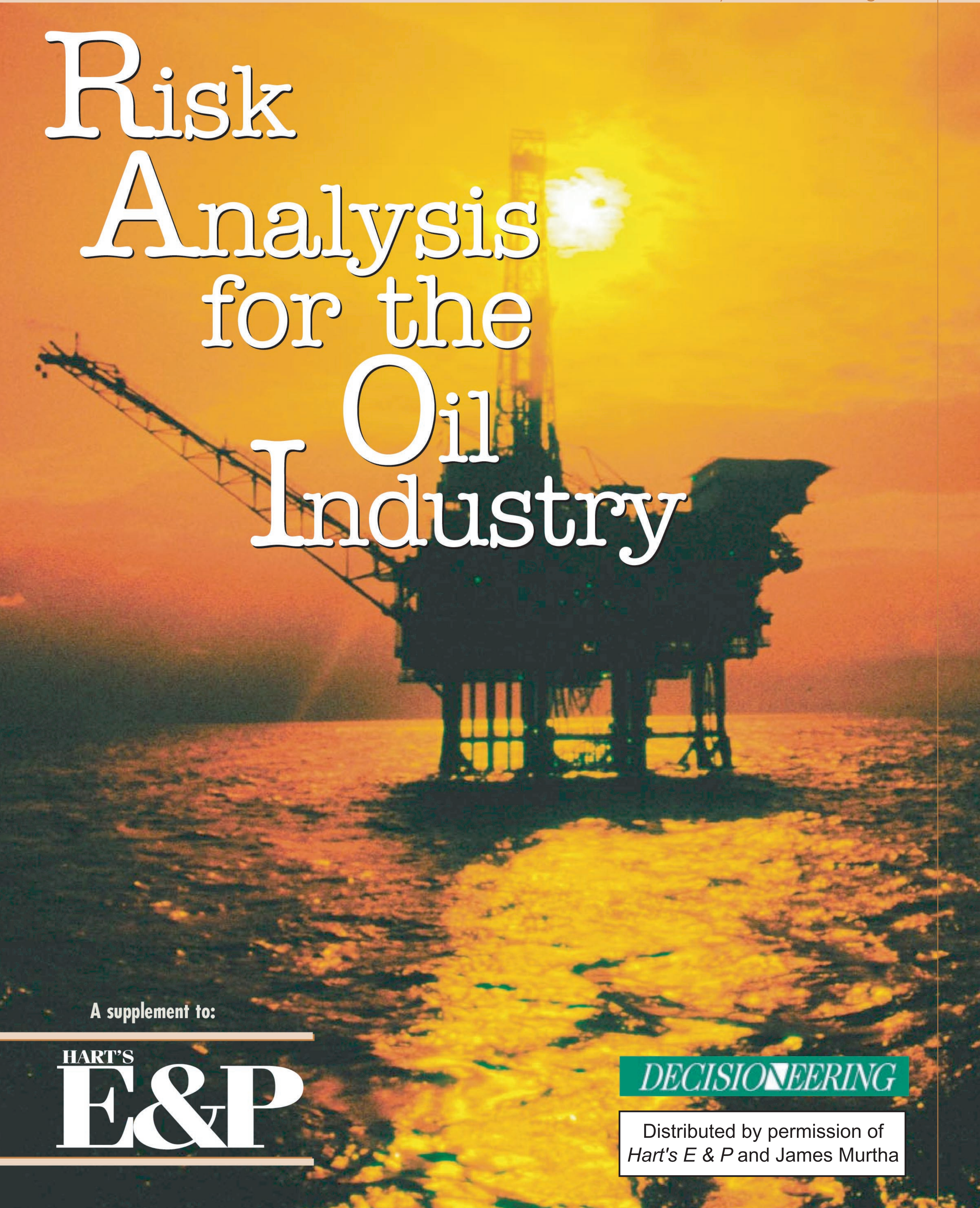
A Comprehensive
Study of Risk Management

Risk Analysis for the Oil Industry



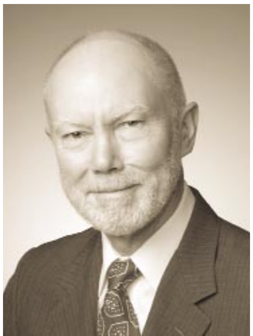
A supplement to:

HART'S E&P

DECISIONEERING

Distributed by permission of
Hart's E & P and James Murtha

Biography



Jim Murtha, a registered petroleum engineer, presents seminars and training courses and advises clients in building probabilistic models in risk analysis and decision making. He was elected to Distinguished Membership in SPE in 1999, received the 1998 SPE Award in Economics and Evaluation, and was 1996-97 SPE Distinguished Lecturer in Risk and Decision Analysis. Since 1992, more than 2,500 professionals have taken his classes. He has published Decisions Involving Uncertainty – An @RISK Tutorial for the Petroleum Industry. In 25 years of academic experience, he chaired a math department, taught petroleum engineering, served as academic dean, and co-authored two texts in mathematics and statistics. Jim has a Ph.D. in mathematics from the University of Wisconsin, an MS in petroleum and natural gas engineering from Penn State and a BS in mathematics from Marietta College. ◆

Acknowledgements

When I was a struggling assistant professor of mathematics, I yearned for more ideas, for we were expected to write technical papers and suggest wonderful projects to graduate students. Now I have no students and no one is counting my publications. But, the ideas have been coming. Indeed, I find myself, like anyone who teaches classes to professionals, constantly stumbling on notions worth exploring.

The articles herein were generated during a few years and written mostly in about 6 months. A couple of related papers found their way into SPE meetings this year.

I thank the hundreds of people who listened and challenged and suggested during classes.

I owe a lot to Susan Peterson, John Trahan and Red White, friends with whom I argue and bounce ideas around from time to time.

Most of all, these articles benefited by the careful reading of one person, Wilton Adams, who has often assisted Susan and me in risk analysis classes. During the past year, he has been especially helpful in reviewing every word of the papers I wrote for SPE and for this publication. Among his talents are a well tuned ear and high standards for clarity. I wish to thank him for his generosity.

He also plays a mean keyboard, sings a good song and is a collaborator in a certain periodic culinary activity.

You should be so lucky. ◆

Table of Contents

Decision Trees vs. Monte Carlo Simulation

Decision trees and Monte Carlo simulation are the two principal tools of risk analysis. Sometimes users apply one tool when the other would be more helpful. Sometimes it makes sense to invoke both tools. After a brief review of their objectives, methods and outputs, we illustrate both proper and improper applications of these well-tested procedures.

Decision trees: Objectives, methodology, results

Decision trees select between competing alternative choices, finding the choice – and subsequent choices along a path – which either maximizes value or minimizes cost.

The output of a tree is a combination of the optimal path and the expected value of that path. Thus, decision trees yield one number. The rules of solving the tree force acceptance of the path with the highest expected value, regardless of its uncertainty.

In theory, preference functions can be used in place of monetary values, but in practice, users seldom go to this level.

Decision trees allow limited sensitivity analysis, in which the user perturbs either some of the assigned values or some of the assigned probabilities, while monitoring the overall tree value. Typically, the user varies one or two parameters simultaneously and illustrates the results with graphs in the plane (how the tree value changes when one assigned value is varied) or in space (how the tree value varies when two assigned values are varied together). The traditional tornado chart also is used to show how each perturbed variable affects the tree value when all other values are held fixed. This chart takes its name from the shape it assumes when the influences of the perturbed variables are stacked as lines or bars, with the largest on top.

Trees, along with their cousins influence diagrams, are particularly popular for framing problems and reaching consensus. For small to moderate size problems, the picture of the tree is an effective means of communication.

One of the most important problem types solvable by trees is assessing the value of information. In this case, one possible choice is to buy additional information (seismic interpretation, well test, logs, pilot floods). Solving the tree with and without this added-information branch and taking the difference between the two expected values yields the value added by the information. If the information can be bought for less than its imputed value, it is a good deal.

Monte Carlo simulation: Objectives, methodology, results

Monte Carlo models focus on one or more objective functions or outputs. Favorite outputs include reserves, total cost, total time and net present value (NPV). Their respective inputs include areal extent, net pay and porosity; line item costs; activity times; and production

Trees, along with their cousins influence diagrams, are particularly popular for framing problems and reaching consensus.

	P0	P5	Mode	P95	P100
Area	976.3	1,300	2,000	2,700	3,023.7
Pay	10.75	20	40	60	69.25
Recovery	76.9	100	150	200	223.1

Table 1. Defining parameters for inputs (Area, Pay, Recovery) to Monte Carlo model

forecasts, price forecasts, capital forecasts and operating expense forecasts.

A Monte Carlo (MC) simulation is the process of creating a few thousand realizations of the model by simultaneously sampling values from the input distributions. The results of such an MC simulation typically include three items: a distribution for each designated output, a sensitivity chart listing the key variables ranked by their correlation with a targeted output, and various graphs and statistical summaries featuring the outputs.

Unlike decision trees, MC simulations do not explicitly recommend a course of action or make a decision. Sometimes, however, when there are competing alternatives, an overlay chart is used to display the corresponding cumulative distributions, in order to compare their levels of uncertainty and their various percentiles.

Example: A tree that can be converted to simulation model

Sometimes a decision tree can be reconstructed as a Monte Carlo simulation. This is especially true of trees with one root-decision node. The simulation would present the result as a distribution, whereas the tree solution would only give the mean value of the distribution. Take, for example, the tree in Figure 1, where the object is to estimate reserves. Note the pattern of a sequence of chance nodes without interspersed choices. This is a giveaway for conversion. Incidentally, a + sign signifies where a node and its subsequent branches and nodes have been collapsed. In this case, imagine copying the node at "big area" and pasting it at the "med area" and "small area" nodes. We shall convert this tree to a simulation and compare the nature and extent of information that these two models provide us.

In fact, the tree of Figure 1 is not a standard decision tree. Instead of the convention of adding the values as one moves out along a path, this tree multiplies them together to get barrels. The output is simply the expected value obtained by considering the possible combinations of area, pay and recovery

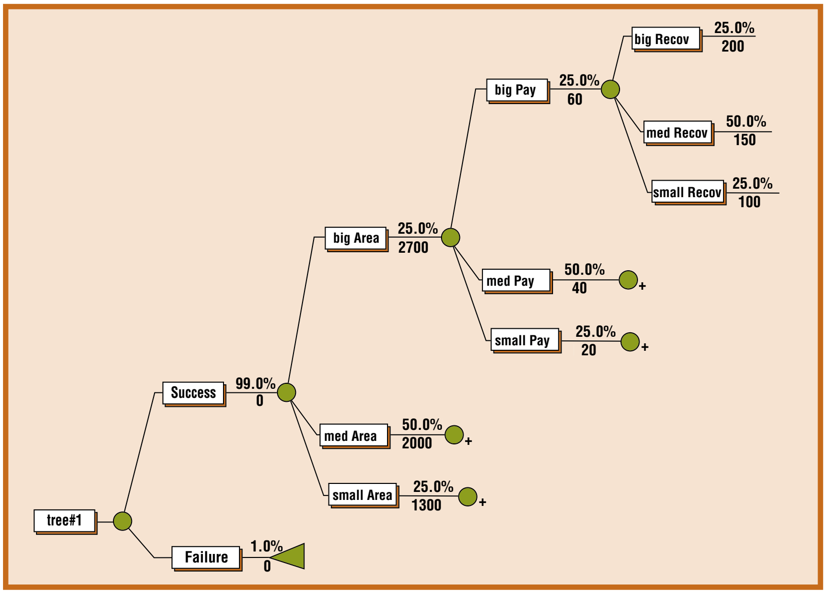


Figure1. A decision tree begging to become a Monte Carlo model.

obtained by following some path, multiplying each of them by the corresponding probability, which is obtained by taking the product of the branch probabilities, and summing these weighted values.

In the tree, each parameter is discretized to only three representative values, signifying small, medium, and large. Thus, area can take on the values 1,300, 2,000 or 2,700 acres. Of course in reality, area would be a continuous variable, taking on any value in between these three numbers. In fact, most people would argue 1,300 is not an absolute minimum, and 2,700 is not an absolute maximum. They might be more like P10 and P90 or P5 and P95 estimates. We can think of a small area being is some range, say from 1,000 to 1,500 with 1,300 being a suitable representative of that class. Similarly, 2,700 might represent the class from 2,500 to 3,000 acres. Each of these subranges of the entire range carries its own probability of occurrence.

For simplicity, we have made all the triples of values symmetric (for example, 1,300, 2,000 and 2,700 are equally spaced), but they could be anything. For instance, area could have the values 1,300, 2,000 and 3,500 for small, medium and large. Likewise, we have assigned equal weights to the small and large representatives, again for simplicity and ease of comparison.

We have made another simplification: we assume all possible combinations are realizable. Sometimes, the large value of area would be paired with three

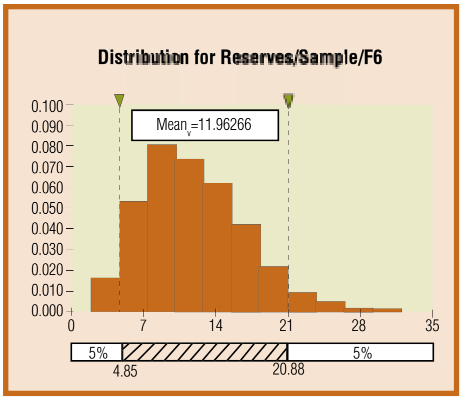


Figure 2. Histogram from simulation

Percentage	MMSTB
5%	4.9
10%	6.0
15%	6.8
20%	7.5
25%	8.2
30%	8.9
35%	9.5
40%	10.1
45%	10.8
50%	11.4
55%	11.9
60%	12.6
65%	13.4
70%	14.2
75%	15.0
80%	16.0
85%	17.3
90%	19.0
95%	21.2

Table 2. Percentiles for output (reserves, MMSTB) from simulation

relatively large values of pay, in the belief these two parameters are dependent. This is simple to accommodate in a tree.

Building the corresponding Monte Carlo simulation model

Converting to an appropriate Monte Carlo model requires finding a suitable distribution for each of the three inputs – area, pay and recovery. In light of the discussion above, we took the small values to be P5 and the big values to be P95. We also selected triangular distributions is each case (which is how we obtained our P0 and P100 values). The resulting distributions are shown in Table 1.

From the tree analysis, we can calculate the expected value as well as finding the two extreme values (smallest and largest) and their respective probabilities:

Expected value 12MMSTB
Minimum value 2.6MMSTB, P(min) = 1/64
Maximum value 32.4 MMSTB, P(max) = 1/64

What we cannot determine from the tree analysis is how likely the reserves would exceed 5 MMSTB, how likely they would be between 5 MMSTB and 15 MMSTB, how likely they would be less than 12 MMSTB, and so on.

The histogram from the Monte Carlo analysis is shown in Figure 2, and its corresponding percentiles in Table 2. Thus, while the mean value coincides with the mean from the tree analysis (in part because of the symmetry of the inputs and lack of correlation), we learn much more about the range of possibilities:

- 90% of the values lie between 4.9 and 20.9 MMSTB;
- there is about a 56% chance of finding less than 12 MMSTB (it is close to the median);
- there is only about a 20% chance of exceeding 16 MMSTB; and

Modifying the tree to account for dependency among inputs

Consider the modified tree branches shown in Figure 3, where the user believes larger areas correspond to larger net pays. Without going into detail, the corresponding Monte Carlo simulation would handle this relationship by specifying a number between -1 and +1 to indicate the degree of correlation between the inputs. If the correlation were between 0 and +1, for example, then during the simulation, realizations sampling larger values of area would tend to be matched with larger values of pay. This technique is routine, simple to implement and can use historical data when it is available to estimate the correlation coefficients. The resulting simulation in this case would feature more extreme values (both large and small) resulting in a larger standard deviation and wider 90% confidence interval than the uncorrelated case. Somewhat surprisingly, the mean value will also increase, but not as much as the range.

Introducing correlation and examining its effect is a standard exercise in Monte Carlo classes. If anything, the simulation handles these paired relationships more easily than the tree, where the dependency is more hard-wired.

Example: Comparing alternative mud systems

Sometimes a problem can be described with a tree and enhanced with a simulation. A case in point would be a drilling cost estimate where there is a choice of mud systems. Consider the following problem.

For 4 years, you have been drilling in a field where there is a danger of differential sticking. Of the 20 wells drilled to date, six of them encountered stuck pipe. In five of those wells, the drillers fixed the problem, spending from 4 to 18 days of rig time with an average of 9 days. In one well, they eventually had to sidetrack at a marginal cost U.S. $700,000 for materials and 12 days of rig time beyond the attempt to free the stuck pipe, which had been 18 days. Average time to drill the wells is approaching 45 days. The rig rate is $60,000/day.

One option for the next well is to use oil based mud, which would greatly reduce the chance of stuck pipe (to 1/10) and the time to fix it (to 6 days), but cost $70,000/day,

Ordinarily, one would look to the decision tree to compare these two alternatives. First, we calculate the expected value of drilling with the conventional water based mud system for the three cases: no problem, stuck but fixed and sidetrack.

No problem cost: (45 days) * ($60,000) = $2,700,000
Stuck and fixed cost: (45 + 9 days)*($60,000) = $3,240,000
Sidetrack cost: (45 + 30)*($60,000) + $700,000 = $5,200,000

The respective probabilities we assign are 14/20, 5/20 and 1/20.

Next, we estimate the costs and probabilities for the oil based mud system.

No problem cost: (45 days)*($70,000) = $ 3,150,000
Cost when stuck pipe: (45+6)*($70,000) = $3,570,000

The respective probabilities are 9/10 and 1/10.

The resulting tree is shown in Figure 3, indicating the correct decision would be to use the water based mud with an expected value of $2,960,000 rather than the oil-based alternative with an expected value of $ 3,171,000.

The Monte Carlo approach

The Monte Carlo model shown in Table 3 captures the same information as the tree model, but uses it differently. Specifically, the simulation model allows the user to estimate ranges for the various activities and costs. For comparison purposes, the historical minimum and maximum values are treated as P5 and P95 to build input distributions, with the option for the user to set these percentages. Most people estimating ranges of this sort, especially cost estimators, tend toward right-skewed distributions.

P(StuckWater)	30.0%			
P(SideTrack)	5%	LowP	5	
P(StuckOil)	5%	HighP	95	
	Sample	**Plow**	**Mode**	**Phigh**
DayRateWater	$ 60	60	60	60
DayRateOil	$ 70	70	70	70
DaysWater_NoStuck	45.0	40.5	45	49.5
DaysOil_NoStuck	45.0	40.5	45	49.5
DaysSideTrack	12.0	10.8	12	13.2
StuckWater?	0	1 = stuck		
SideTrack?	0	1=sidetrack		
StuckOil?	0	1 = stuck		
DayWater_Stuck	10.4	4	8	18
DaysOil_Stuck	7.2	3	6	12
ExtraCost_ST	700.0			
CostWater	$ 2,700			
CostOil	$ 3,150			
TimeWater	45.0			
TimeOil	45.0			

Table 3. Monte Carlo model to compare oil based and water based mud systems

	Decision Trees	Monte Carlo simulation
Objectives	make decisions	quantify uncertainty
Inputs	discrete scenario	distributions
Solution	driven by EV	run many cases
Outputs	choice and EV	distributions
Dependence	limited treatment	rank correlation

Table 4. Comparison between Decision trees and Monte Carlo simulation

Certainly, in this model, we should care about more than just the expected value. For instance, we may want to know the probability of exceeding a certain value, such as $4,000,000. We may also want to question other inputs (see discussion below).

Figure 5 shows the comparison with the assumptions taken from the problem statement. Figure 6 shows a variation where the oil-based mud takes only 40 days on average to drill and the chance of sidetrack with water mud is 10%, not 5%. The difference between the two cases is clear: as the probability of sidetracking increases, the oil-based mud presents a less risky alternative (the density function of outcomes is much narrower). Similar analyses can be made easily in the Monte Carlo model.

Which model is better in this instance?

Both the tree and the simulation offer information to the user and to the larger audience of the presentations that can be prepared from them. The tree emphasizes the two contrasting choices and identifies the extreme outcomes and their probabilities of occurrence. It is both simple and effective. Yet, the tree does not explicitly acknowledge the underlying uncertainty in each of the contributory estimates (such as days for problem free drilling, cost of mud system, days for trying to free the stuck pipe, days for sidetracking). The user can handle these uncertainties – one or two at a time – by using the tree framework to do a trial-and-error sensitivity analysis, and any user would be remiss in avoiding this. But the overall uncertainty is made more explicit in the Monte Carlo simulation, where the

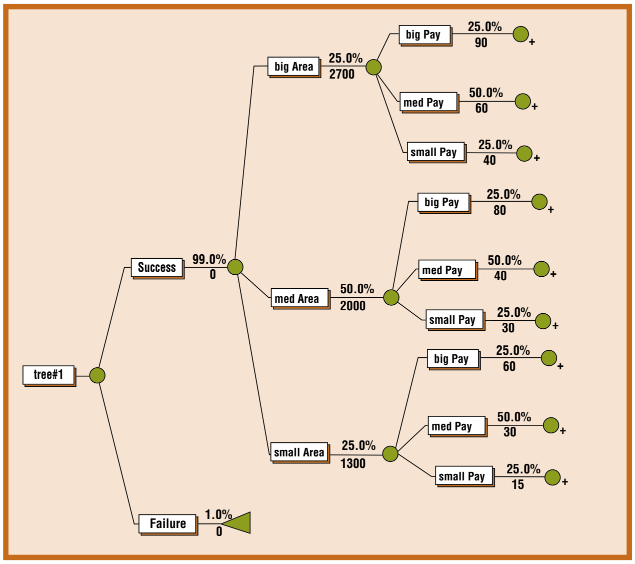


Figure 3. Tree incorporating dependence between area and pay

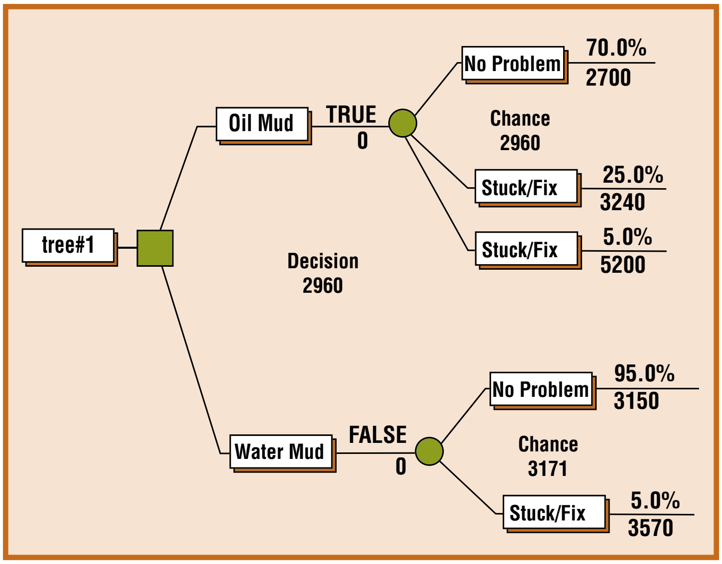


Figure 4. Tree to compare oil based vs. water based mud systems

very nature of the model begs the user to specify the range of uncertainty.

On balance, if I had to pick one model, I would pick the simulation, in part because I have had far better results analyzing uncertainty and presenting the details to management when I use simulation. But whatever your preference in tools, for this problem the combination of both a tree and the simulation seems to be the most useful.

Which model is better in general?

When is Monte Carlo more appropriate? In general, Monte Carlo seems to have more varied applications. Any (deterministic) model you can build in a spreadsheet can be converted to a Monte Carlo model by replacing some of the input values with probability distributions. Thus, in addition to the popular resource and reserve volumetric product models, people build Monte Carlo models for costs of drilling and facilities, time for projects (up to a point, when project schedule software is more applicable), production forecasts, and all sorts of cashflows, including those with fiscal terms. Decision trees must come down to something that compares alternative choices under uncertainty.

Combining simulation and trees

It is now possible to build a decision tree where some or all of the branches from chance nodes contain probability distributions rather than values. Then one can run a Monte Carlo simulation, where each iteration is a new solution of the tree. This technology is relatively new and relatively untested. Although it holds promise, this combination requires the user to be expert at both decision trees and simulation. Some people raise questions of logic when we solve a tree populated with values that are not mean estimates, but rather extreme possibilities of the inputs. In time, when technical papers appear, many of these issues will be resolved.

Summary

Table 2 summarizes some of the features of the two methods, illustrating their principle differences. Don't be surprised to find someone using a tree solution to a problem you elect to solve with simulation or vice versa. Do try to use common sense and do as much sensitivity analysis as you can regardless of your choice. ◆

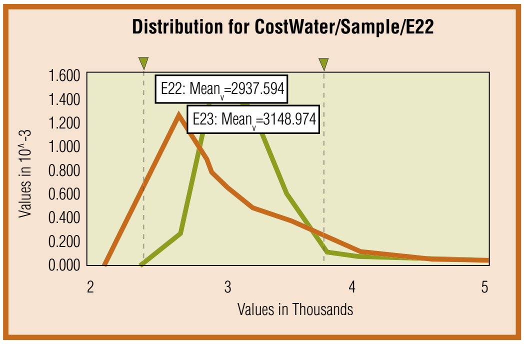


Figure 5. Comparing oil and water based mud systems

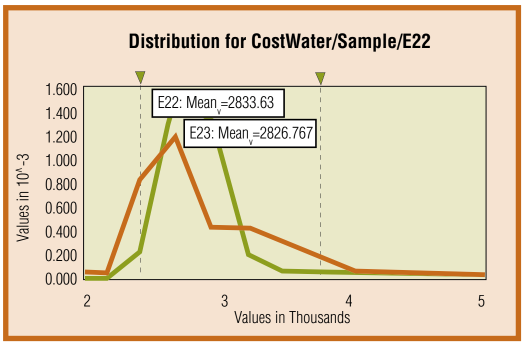


Figure 6. New cost comparison when oil based mud results in faster drilling.