ECONOMIC RISK ANALYSIS
Using Analytical and Monte Carlo Techniques

BRENDAN R. O'DONNELL,[1] MICHAEL A. HICKNER,[2] BRUCE A. BARNA
Michigan Technological University • Houghton, MI

Investment decisions are typically based on some form of cash-flow analysis, such as net present value (NPV) or internal rate of return (IRR). The analysis is first performed using predicted performance of the project over the project life as if the predictions were deterministic. The stochastic nature of these predictions can then be handled using a variety of risk analysis techniques, such as: best case/worst case scenarios; single-parameter sensitivity analysis (Strauss plots); analytical error propagation; Monte Carlo simulation; and decision trees. In this paper, we present the development and application of a Microsoft Excel spreadsheet template that facilitates analytical and Monte Carlo risk analysis of investment decisions. We have found the template particularly useful in teaching risk analysis to senior students in the design course.

METHODS FOR ASSESSING RISK

Best/worst case scenarios calculate a return on investment for the most profitable set of investment conditions and another return for the worst possible set of conditions. This approach analyzes both ends of the spectrum in terms of return. Generally, however, the worst case will not exceed the minimum return and the best case will. Because the expected result is somewhere between the two extremes, most evaluations will not be resolved by this method. The method is useful for those few cases where the worst-case scenario is found to be acceptable or the best case is found to be unacceptable.

Single-parameter sensitivity analysis tests the variability of the result with respect to one economic variable. This type of risk analysis is common because the calculations and interpretation are simple. Only one variable is changed at any given time, and the result (which is frequently linear) can be shown graphically on a Strauss plot (NPV versus change in the variable of interest). Because much information can be derived from a small amount of work, some companies mandate that all capital appropriation requests be accompanied by sensitivity tests on key input values such as raw material price, labor, utilities, etc. This technique can show the break-even point for each of the critical economic variables.

Analytical methods use error propagation analysis to evaluate the risk involved. This approach uses statistical identities to relate the variability of each parameter to its distribution. In order to define variability in the desired risk measure, the relationship between the parameters and the desired measure is combined in equation form. This equation combines all facets of variability in the economic input values into a single expression of variability in the desired risk measurement. The inclusion of variability and multiparameter influence upon the result makes the analytical method applicable for examining risk where the variability is well defined and the input parameters are assumed independent.

Monte Carlo simulations have been used to simulate random variation in sets of related variables. First, a statistical distribution is specified for each input. Then the simulation randomly selects one value for every input from the specified distribution for that item. The set of random input values is used to calculate a result. This process is repeated a sufficient number of times so that the distribution of outcomes can be used to reliably predict the variability of the calculated result. The simulation can be run hundreds or even thousands of times to explore every possible combination of variables. Monte Carlo methods have gained increasing attention due to the increased power and decreased cost of desk-

Brendan O'Donnell *received his MS degree in chemical engineering from Michigan Tech in the summer of 2001. His graduate research focused on developing software tools to aid in process design, economic assessment, and process improvement. He holds a BS degree in chemical engineering from Michigan Tech and is currently working as a photolithography engineer for IBM in Burlington, Vermont.*
Bruce Barna *is Professor of Chemical Engineering at Michigan Technological University. He holds BS and MS degrees from Michigan Tech and a PhD from New Mexico State University. He worked as a process engineer for Reynold's Metals and Exxon and as a plant engineer and plant manager for Kalsec, Inc., prior to joining the faculty at MTU.*
Michael Hickner *is currently a chemical engineering PhD candidate at Virginia Tech. His research involves synthesizing and characterizing new proton exchange membranes for fuel cells. He has conducted part of his fuel cell research at Los Alamos National Laboratory. He received his BS in chemical engineering from Michigan Tech in 1999.*

[1] *Currently with IBM, Inc., Burlington, Vermont*
[2] *Currently at Virginia Tech, Blacksburg, Virginia*

top computing. Once the distributions are chosen for each economic input, repeating the calculations becomes trivial.

Decision trees employ a method of weighting an event's economic impact by its probability. The use of decision trees follows a left-to-right progression where each decision builds upon the previous one until a final outcome is reached. Each branch of the decision tree has a probability and an economic value. The expected value of a decision can be calculated by summing the product of the probability and economic outcome to each decision node. Comparing the result of each probability node will result in an ultimate, numerically based decision. The advantage of this method is the ability to incorporate calculated probabilities and economic factors to give a numerical result for a complex decision-making process.[1]

The logical place to teach these risk-analysis techniques in most chemical engineering curricula is in the capstone design course. Best/worse case techniques and single-parameter sensitivity techniques are readily mastered by all students. The analytical, Monte Carlo, and decision-tree techniques can be more of a challenge, depending on the statistical background of the students and the time available for them to write their own simulation routines. Recognition of this was the impetus to develop a spreadsheet-based learning tool that could be used to facilitate risk analysis using both analytical and Monte Carlo methods.

IMPLEMENTING THE ANALYTICAL AND MONTE CARLO METHODS

A cash-flow analysis was used to demonstrate analytical and Monte Carlo techniques. A sample cash-flow table was generated using net present value (NPV) as the result of interest

$$\text{Net Present Value} = \sum_{i=1}^{\text{Nyears}} \frac{\text{Constant Cash Flow}}{(1+\text{MAR})^i} \tag{1}$$

$$\text{Constant Cash Flow}_I = \frac{\text{Cash Flow}}{(1+\text{Inflation Rate})^i} \tag{2}$$

$$\text{Cash Flow} = \text{Income} - \text{Expenses} - \text{Investment} - \text{Working Capital} - \text{Tax} \tag{3}$$

$$\text{Tax} = \text{Profit}_{\text{pre-tax}} * \text{Tax Rate} \tag{4}$$

$$\text{Profit}_{\text{pre-tax}} = \text{Income} - \text{Expenses} - \text{Depreciation} \tag{5}$$

This cash-flow table was incorporated into a spreadsheet template to facilitate analytical and Monte Carlo analysis of the NPV. The template allows the user to tailor the analytical and Monte Carlo analyses to a specific set of economic variables, including the distribution of each variable. The next part of this article will first describe the theoretical basis for each analysis and then its implementation in the actual Excel spreadsheet.

In this paper, we present the development and application of a Microsoft Excel spreadsheet template that facilitates analytical and Monte Carlo risk analysis of investment decisions.

ANALYTICAL METHOD

The cash-flow analysis provides the relationship between our result and the inputs: income, expenses, working capital, investment, and inflation rate. In order to determine the variability in the result, we will use the method of error propagation.[2,3] When a variable, c, is a function of a number of variables, $x_1, x_2, \ldots x_n$, it can be written

$$c = \phi(x_1, x_2, \ldots x_n) \tag{6}$$

It follows that if each x_i is independent and σ_c^2 represents the variance of c, then

$$\sigma_c^2 = \left(\frac{\partial c}{\partial x_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{\partial c}{\partial x_2}\right)^2 \sigma_{x_2}^2 + \ldots + \left(\frac{\partial c}{\partial x_n}\right)^2 \sigma_{x_n}^2 \tag{7}$$

By applying this equation to economic variation, with NPV as c, fixed capital investment (Inv) as x_1, income (Inc) as x_2, expenses (Exp) as x_3, working capital (WC) as x_4, and inflation (Inf) as x_5, we arrive at the following expression for the variance of the net present value:

$$\sigma_{NPV}^2 = \left(\frac{\partial NPV}{\partial Inv}\right)^2 \sigma_{Inv}^2 + \left(\frac{\partial NPV}{\partial Inc}\right)^2 \sigma_{Inc}^2 + \left(\frac{\partial NPV}{\partial Exp}\right)^2 \sigma_{Exp}^2 + \left(\frac{\partial NPV}{\partial WC}\right)^2 \sigma_{WC}^2 + \left(\frac{\partial NPV}{\partial Inf}\right)^2 \sigma_{Inf}^2 \tag{8}$$

Using Eq. (8), if we can define all of the terms on the right side of the equation, we should be able to calculate the variance of the net present value. The problem then becomes one of calculating the components of the right side.

Approximating the partial differentials • Partial differentials represent the slope of the function with respect to the variable of interest at a small increment. If we assume that the function responds nearly linearly due to an incremental change in the variable of interest, then we could approximate the partial differential by changing the variable a small percentage above and below the base case value and calculating the slope from the two resulting points. Figure 1 shows such a change and resultant NPV with a line fit.

We can use this approach to calculate the partial derivatives of NPV with respect to the rest of the variables following a similar procedure. The task then becomes one of estimating the variances for each variable.

Estimating individual variances • Defining the variance requires estimation based on past experience or future prediction. For our purposes, we assumed that each of the five input variables follows one of three distributions: normal, uniform, or modified-beta. Although other distributions could be added, these three distributions can represent most of the types of distributions encountered. Each distribution requires additional inputs to define the variance. The normal distribution requires the standard deviation as an input (the variance is simply the square of the standard deviation). The mean value is assumed to be the base case value, m. For the other two distributions, we need estimates of the maximum, minimum, and the most likely value for the beta distribution. Thus, we define

a = minimum value
b = maximum value
m = most likely value (mode)

The variance for the uniform distribution can be calculated using[4]

$$\sigma^2 = \frac{(b-a)^2}{12} \tag{9}$$

The modified-beta distribution uses these maximum and minimum inputs to calculate variance based on the following PERT (Program Evaluation and Review Technique) simplified formula:[5]

$$\sigma^2 = \left[\frac{(b-a)}{6}\right]^2 \tag{10}$$

Modified-beta distributions can be skewed either positively or negatively. The expected mean is different from the most likely value and is calculated by

$$\mu = \frac{a+4m+c}{6} \tag{11}$$

Implementing the analytical approach on a spreadsheet Once the variance for each variable and the partial differential of NPV with respect to each variable (or input) has been approximated, the overall variance can be calculated. The partial derivative for each variable is estimated by modifying the individual parameter a set percentage (specified by the user) and calculating the slope between the two perturbed points. The variance of each parameter necessary to satisfy Eq. (8) can be correlated from input values of uncertainty. Equation 8 combines the parts to calculate the variance in net present value. We have now quantified the uncertainty in our economic decision variable using the analytical method.

The assumption of independence is one weak aspect of the analytical method. Some of the variables are often interrelated. For example, expenses are often related to investment, working capital is sometimes related to investment, etc. In this respect, the Monte Carlo simulation may be more appropriate where such interrelations exist. The Monte Carlo technique does not explicitly account for interrelations either, but more combinations of variables are explored, as is illustrated in the following paragraphs.

MONTE CARLO METHOD

Monte Carlo simulations can reduce error compared to the analytical approach by performing random walks within specified distributions and determining the results directly from repeated trials. In this case, we are interested in finding the variability in the net present value based on the variability in the five economic parameters mentioned earlier.

Using assumptions for individual variability, we can pick sets of random expenses, incomes, investments, etc., and calculate a result for that set using the cash-flow equations (Eqs. 1-5) to find a net present value for the set. Variance in net present value can then be extracted directly from the distri-

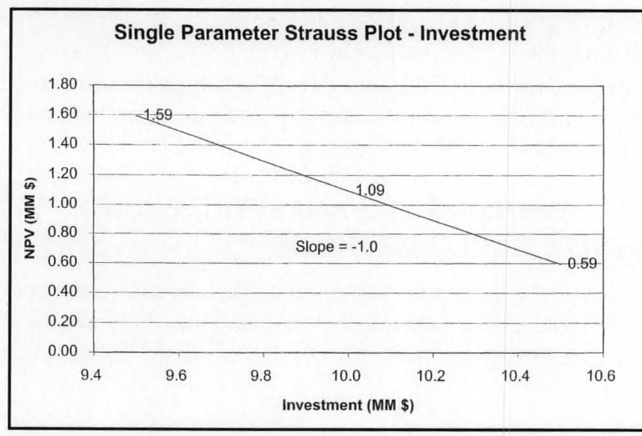

Figure 1. Determining the partial derivative of NPV with respect to investment, using a single-parameter Strauss plot.

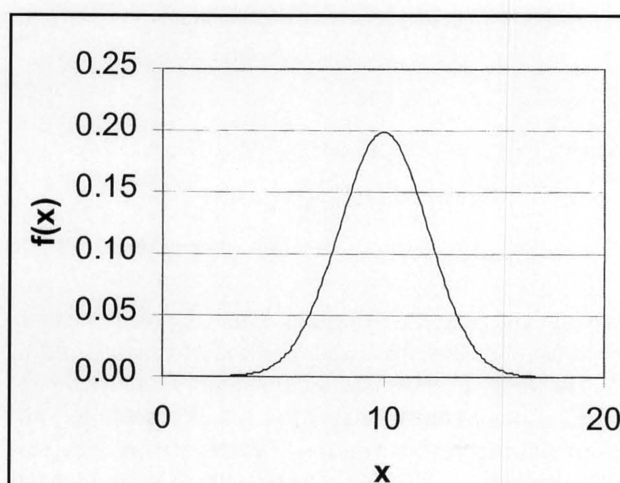

Figure 2. Normal probability plot for mean = 10 and $\sigma^2 = 4$.

bution of net present value results.

Statistical derivation • Normal, uniform, and modified-beta distributions are used for the Monte Carlo simulations as well. The normal distribution represents the standard normal or Gaussian curve. For such a distribution, approximately 67% of all values lie within one standard deviation of the mean, 95% lie within two standard deviations, and 99% lie within three. A normally distributed variable can be characterized by its mean and variance (or standard deviation). Figure 2 illustrates a normal probability curve with mean of ten and variance of four. Of course, the sum of the probabilities for all occurrences is one.

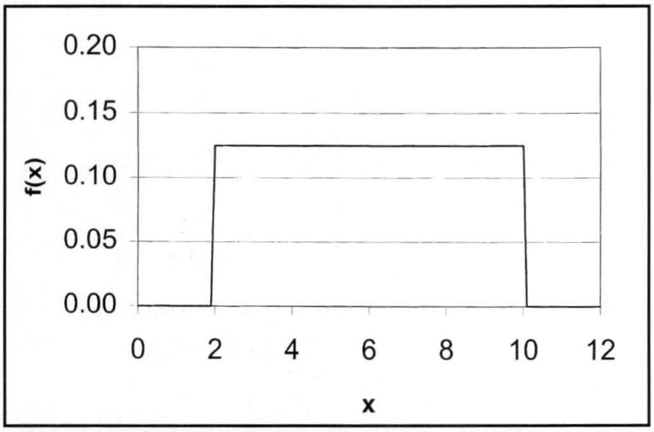

Figure 3. *Uniform probability plot for minimum 2 and maximum 10.*

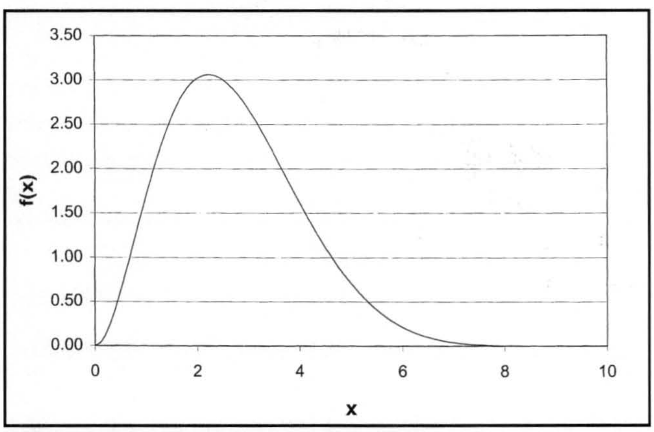

Figure 4. *Modified-beta probability plot for minimum 0, maximum 7, and mode 2.*

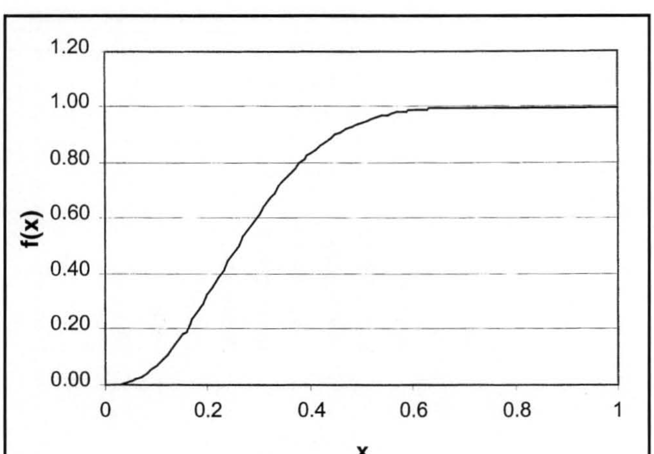

Figure 5. *Cumulative modified-beta probability distribution for minimum 0, maximum 7, and mode 2.*

The uniform distribution gives equal probability for any occurrence between the minimum and maximum endpoints, and is completely characterized by these values. Figure 3 shows a uniform distribution with a minimum and maximum of ten and two, respectively.

The modified-beta distribution can be skewed in either direction from the midpoint. It is characterized by its most likely value (mode) and estimates of low and high values. Figure 4 shows the probability distribution for a sample modified-beta distribution with a mode of two, a low of zero, and a high of seven. In this case, there is a lower probability of values to the left of the mode than to the right.

Using the distributions for a Monte Carlo analysis requires programming our spreadsheet to generate random numbers and then to extract a value from the normal, uniform, or modified-beta distributions established by the input of uncertainty for the variables.

Generating random values within a distribution • The Microsoft Excel spreadsheet has some built-in statistical functions. For instance, given a random number, a mean, a standard deviation, and a standard distribution, the function

NORMINV(RAND(), MEAN, STANDARD-DEVIATION)

will return a random value from that distribution. A similar function exists for the uniform distribution. No such function exists for the modified-beta distribution, however, so it had to be programmed separately. The procedure below outlines the calculation routine for the modified-beta distribution. The same approach could be applied to any desired distribution.

Using the modified-beta probability distribution function (Eq. 12, Figure 4), we integrate to get the cumulative probability distribution (see Figure 5). Then, Excel generates a random number between one and zero corresponding to an f(x) (Figure 5). The x-value is selected based on the random f(x) and scaled to the minimum and maximum range specified. The result is a random value within the distribution

$$f(x)=\frac{(a+b+1)!}{a!*b!}x^a(1-x)^b \tag{12}$$

where

a = minimum value

b = maximum value

Generating the simulations • The randomized value of an input variable (for example, expenses) is then combined with the other randomized values of the variables in a set using Eqs. (1-5) to calculate one NPV. In a Monte Carlo simulation, the calculation of NPV is repeated multiple times, each from a new set of random inputs. Initially, we anticipated that 50 to 100 iterations would produce a stable and reproducible NPV distribution. With this number of iterations we found that the results were very sensitive to the bin size selected for frequency analysis and often had not stabilized. Considering that the computing power is readily available and that a calculation requires only one or two seconds, we increased the iteration count to 500, which proved to be sufficient.

The variability in net present value can be extracted directly from the dataset. The specific use of the spreadsheet is discussed next, followed by a case study.

USING THE SPREADSHEET TOOL

Modern spreadsheets have the usefulness of being programmable, extensible, easy to use, and good teaching tools. Unlike various programming languages that hide the intermediate results, spreadsheets allow the user to see the inputs, the dataset, and any calculations. The risk analysis template developed here provides all these aspects for both the analytical and Monte Carlo analyses.

To use the template, first the user must supply the inputs on the first tab of the spreadsheet. Inputs include base-case (mean or mode) values for investment, expenses, income, inflation rate, and working capital. Cash-flow analyses also rely on pre-set variables such as project life, minimum acceptable return, tax rate, and depreciation schedule. Minimum acceptable return (MAR) and tax rate are the only pre-set variables that can be modified by the user in this tool. The depreciation is fixed at seven-year modified accelerated cost-recovery system (MACRS) and project life is fixed at ten years, all typical values for an industrial project.

Next, the user must select the distributions for each of the variables. There is a pull-down menu for selecting the distribution type: normal, uniform, or modified-beta. Below the pull-downs are cells for specifying the variability of the distribution, selected by

Normal • Requires standard deviation, σ; uses base-case values as mean

Uniform • Requires a (minimum) and b (maximum)

Modified-beta • Requires a (minimum) and b (maximum); uses base-case values as most likely (mode)

These entries link to calculation tabs in the spreadsheet. No further user input is required. The analytical result for variance is shown on the INPUT tab, while the Monte Carlo results are shown on the RESULTS sheet. On the HISTOGRAMS tab are histograms for net present value as well as for each variable. Here the user can assess whether the simulation adequately represents the input distribution. Because the Monte Carlo calculation is dependent on random-number generation, recalculating the spreadsheet can result in slightly different distributions.

Both analytical and Monte Carlo results include a confidence interval for NPV. Sometimes it is more useful to calculate the probability of a net present value greater than zero ($P_{NPV} \geq 0$). Such a calculation would represent the probability of the project meeting or exceeding the minimum acceptable return given the expected variation in the variables. If the distribution result is normally distributed, then a simple technique for calculating p-values in Excel is to use the

NORMDIST (VALUE, MEAN, STANDARD-DEVIATION, TRUE)

function. Setting VALUE=0 and subtracting the result from one will calculate the probability of the project NPV exceeding the minimum acceptable return. If the distribution is not normal, then probabilities must be determined directly from the frequency distribution histogram.

CASE STUDY

A simple case study based on a hypothetical project can help illustrate application of the template. Distributions for the variables have been selected based on historical experience, probable error in cost estimation, etc. The minimum acceptable return (MAR) is set at 20% and tax rate set at the federal corporate level of 34%. Project life is 10 years. Table 1 summarizes the inputs necessary for computing an analytical and Monte Carlo risk analysis.

TABLE 1
Base-Case Parameters for Risk-Analysis Case Study

Variable	*Base Case*	*Distribution*	*Max*	*Min*	*Std. Dev.*
Investment	$10 MM	Normal			2
Working Capital	$1MM	Uniform	2	1	
Expenses	$4 MM	Modified-beta	5	3	
Income	$8 MM	Modified-beta	10	7	
Inflation	3%	Normal			1%

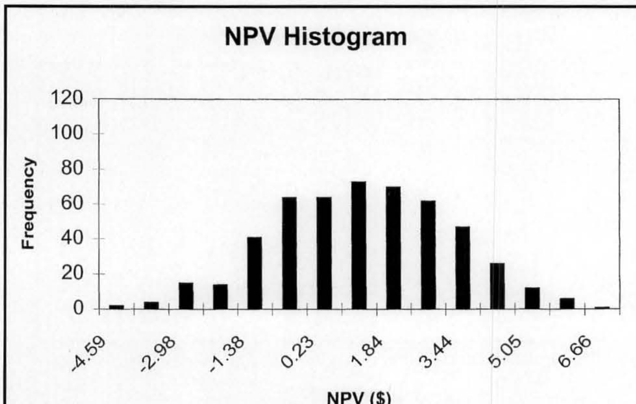

Figure 6. *Histogram of net present value given constraints in Table 1.*

The base-case values seen in Table 1 provide a starting point for the study. The variability information permits study of the variation in the net present value due to the predicted variation of each parameter. With this information, the risk-analysis tool has the necessary inputs to perform both the Monte Carlo and the analytical determination of the variance in net present value.

Using the Excel template, the expected or deterministic NPV for the base case is found to be \$1.18 MM. The template also gives the analytical result for net present value variance as \$4.455 MM^2 with a standard deviation (σ) of \$2.11 MM. Assuming a normal distribution, the 95% confidence interval can be generated by taking the mean NPV \$1.09 MM plus/minus 2 σ, or \$5.28 MM to \$-3.09 MM. The p-value for NPV greater than 0 is 0.7, signifying a 70% chance of the project exceeding the minimum acceptable return.

Monte Carlo results are also generated by the template using 500 iterations. The data set is displayed using histograms. For this case, the mean is found to be \$1.46 MM with a standard deviation of \$2.09 MM. A 95% confidence interval for net present value is \$5.50 MM to \$-2.69 MM. The p-value for NPV greater than zero is 0.77, signifying a 77% chance of the project exceeding the minimum acceptable return. The Monte Carlo results are close to but slightly different than the analytical result.

The histogram in Figure 6 shows a sample of the simulation. Again, because the results are based on random-number generation, each recalculation could have slightly different results. With 500 iterations, the mean and confidence intervals are essentially constant between simulations, but the shape of the histogram varies much more than we had anticipated. The central limit theorem suggests that the sampling distribution of the mean can be approximated by the normal distribution, regardless of the population. Therefore, we would expect the outcome of a calculation involving large numbers of input variables to be normally distributed, regardless of the distribution of the inputs. This does not appear to be the case for our cash-flow analysis and suggests that the Monte Carlo results are probably a better measure of project risk than the analytical results.

USING THE TEMPLATE IN THE CLASSROOM

We have used this template for several years in the capstone design course and have found it to be useful in teaching the concept of risk analysis and analytical estimation of that risk. The students perform a feasibility analysis of a new project or plant in the fall semester. As part of this analysis, they are asked to include an economic analysis and risk analysis of the venture, both in their written report and their oral presentation to the management of Fictitious Chemical Company, their hypothetical employer.

For the economic analysis, students use a spreadsheet template that they have each been asked to generate in a prior homework assignment. For the risk analysis, however, most of the students were historically limited to the single-parameter sensitivity approach since not all of them had the statistical background or time to conduct the more elaborate analyses. Some of the more capable groups are challenged by asking them to perform an analytical or Monte Carlo analysis to illustrate the techniques to the entire class.

With this template now available, we are able to ask all the students to apply the full spectrum of risk analysis techniques to their projects. They are provided with the spreadsheet file and told that they are free to modify it or use it as they see fit. We find that use of the template and this approach allows us to concentrate more on the actual teaching of risk analysis and less on the programming required to do the analysis. All the students are able to successfully apply the software. Those with a strong statistical background tend to do a better job of interpreting the results.

CONCLUSIONS AND FUTURE WORK

Risk analysis is a critical part of any project decision. Increasing the minimum acceptable return or setting higher breakpoints are simple methods of compensating for risk that have been used as shortcuts in the past. The goal of this research was to develop a spreadsheet template for quantifying the risk in the discounted cash-flow measure, net present value. Analytical and Monte Carlo methods were implemented in a Microsoft Excel template for ease of use. Both methods result in a mean and a standard deviation value. The template also calculates confidence intervals based on the results.

The template is a work in progress. We hope, in future versions, to be able to streamline some of the Monte Carlo simulations, to develop macros for group calculations, and to bypass some of the more computationally intensive tasks. We also hope to add the ability to alter project life, iteration number, depreciation schedule, etc.

We have found the template to be quite useful in teaching risk analysis concepts to our seniors in the plant design course. Faculty who would like to try it in their courses may contact Bruce Barna to get the latest version. Be advised that the file is large (approximately 6.5MB). We would appreciate feedback on the experiences of other users.

REFERENCES

1. Blank, L.T., and A.J. Tarquin, *Engineering Economy,* 4th ed., McGraw Hill, New York, NY, pp. 574-580 (1998)
2. Jelen, F.C., and J.H. Black, *Cost and Optimization Engineering,* McGraw Hill, New York, NY, p. 168 (1983)
3. Holland, F.A., F.A. Watson, and J.K. Wilkinson, "Probability Techniques for Estimates of Profitability," *Chem. Eng.,* Jan. 7, pp. 105-110 (1974)
4. Freund, F.E., and I. Miller, *Probability and Statistics for Engineers,* Prentice-Hall, Englewood Cliffs, NJ, p. 114 (1977)
5. Whitehouse, G.E., *Systems Analysis and Design Using Network Techniques,* Prentice Hall, Englewood Cliffs, NJ, p. 42 (1973) ❒