# ECOSYSTEM RESTORATION COST RISK ASSESSMENT

*Final Report*

by

Charles Yoe, Ph.D.

In association with:

Planning and Management Consultants. Ltd.
6352 South U.S. Highway 51
P.O. Box 1316
Carbondale, IL 62903
(618) 549-2832

A Report submitted to:

U.S. Army Corps of Engineers
Institute for Water Resources
7701 Telegraph Road
Alexandria, VA 22315-3868

Risk Analysis of Water Resources Investments Research Program

IWR Report 02-R-1
June 2001

**Views, opinion, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other official documentation.**

# PREFACE

The work presented in this document was conducted as part of the Risk Analysis of Water Resources Investments Research Program, under the "Identifying and quantifying key sources of risk and uncertainty in production and costs related incremental analysis" work unit. The Program is sponsored by the Headquarters, U. S. Army Corps of Engineers and is assigned to the Institute for Water Resources (IWR). Dr. David Moser is the Program Manager at IWR. Mr. Harry Kitch, Planning and Policy Division, Mr. Jerry Foster and Mr. Earl Eiker (retired), Engineering Division, and Mr. Harold Tohlen (retired), Operations, Construction and Readiness Division, are the Headquarters Program Monitors. Field Review Group Members that provide overall program direction include: Mr. Martin Hudson, Portland District; Mr. S.K. Nanda and Mr. Dale Rossmiller, Rock Island District; Mr. Jerry Smith, Southwest Division; Mr. Gerald Melton, South Atlantic Division; Mr. Paul Wemhoener, Omaha District; and Mr. Franke Walberg, Kansas City District. The report was prepared under the general supervision of Dr. David Moser (Acting) Chief of Decision Methodologies Division, Institute for Water Resources, and Mr. Bob Pietrowski, Director of the Institute for Water Resources. This report was prepared by Dr. Charles Yoe on behalf of Planning and Management Consultants, Ltd (PMCL). Mr. Jack Kiefer of PMCL provided overall contract management. Ms. Joy Muncy of IWR served as the Project Manager.

This document evolved from the contributions of the following people: Ms. Amy Guise (CENAB-PL-P), Ms. Marianne Matheny (CENAB-PL-P), Mr. Cedric Bland (CENAB-EN-C), Ms. Marie DeLaTorre (CENAB-PL-E), Mr. Jon Fripp (CENAB-ENH), Mr. Mark Colosimo (CENAB-PL), Mr. Michael Martyn (CENAB-EN-D), Ms. Meg Jonas (CENAB-EN-G), and Ms. Sue Hughes (NABCE-PL).

Review comments of this document were provided by: Dr. David Moser, Mr. Keith Hofseth, Ms. Lynn Martin and Ms. Joy Muncy of IWR, Mr. Jim Henderson of the Engineer Development and Research Center (formerly known as Waterways Experiment Station), and Ms. Amy Guise and Mr. John Vogel of the Baltimore District.

# TABLE OF CONTENTS

**Preface** ... iii

**List of Tables** ... vii

**List of Figures** ... vii

**I. Introduction** ... 1

**II. Risk Analysis** ... 3

Risk Analysis ... 3
Risk Assessment ... 4
Risk Management ... 5
Risk Communication ... 5

**III. Reasons for Doing Cost Risk Assessment** ... 7

Improved Accuracy ... 7
A Distribution of Costs ... 7
Probability of Costs Exceeding Our Estimate ... 8
Estimate Exposure ... 10
Identify Key Components In Exposure ... 10
Full Knowledge of Estimate ... 11
Improving Decisions with Cost Estimate Risk Analysis ... 12
Contingencies with Confidence Intervals ... 13
Comparing Alternative Designs ... 14
Feasibility ... 15
Arranging Financing ... 15
Construction Profitability ... 16
Aids Cost Management ... 16
Useful Throughout Life Of Project ... 16

**IV. Techniques** ... 19

Sensitivity Analysis ... 19
Range Estimation ... 20
Monte Carlo Simulation ... 21

**V. Case Study Choice** ... 27

**VI. Seeley Creek Cost Risk Assessment** ... 29

Variation in Seeley Creek Costs ... 30
The Model ... 33
Triangular Distribution Parameters ... 40
Interdependence of Variables ... 41
Simulation ... 41

Assessment Results ........ 41
Lessons Learned ........ 47

**VII. Summary and Conclusions ........ 49**

**References ........ 51**

**Appendix A: Describing Uncertainty ........ A-1**

# LIST OF TABLES

Table III-1 Selected Project Costs in Millions ....8
Table III-2 Contingencies with Confidence Levels ....13

Table IV-1 Cost Model ....20

Table VI-1 Project Cost Estimate ....30
Table VI-2 Cost Risk Assessment ....31
Table VI-3 Selected Project Costs in Millions ....40

# LIST OF FIGURES

Figure II-1 Risk Analysis Model ....4

Figure III-1 Empirical Distribution ....9
Figure III-2 Frequency Distribution ....9
Figure III-3 Importance Analysis ....11
Figure III-4 Cost Comparison ....14

Figure IV-1 Monte Carlo Process ....21

Figure VI-1 Empirical Distribution ....27
Figure VI-2 Frequency Distribution ....27
Figure VI-3 Excavation Quantity Input ....32
Figure VI-4 Uniform Distribution ....33
Figure VI-5 Beta Subjective Distribution ....34
Figure VI-6 Empirical Distribution of Seeley Creek Costs ....37
Figure VI-7 Frequency Distribution of Seeley Creek Costs ....38
Figure VI-8 Importance Analysis for Seeley Creek Costs ....39

# I. INTRODUCTION

Cost estimation is a fundamentally uncertain exercise in the best of circumstances. Estimating the future costs of a current decision is inherently uncertain. Cost estimation, like many other professional endeavors, had for some time been loathe to openly admit to that uncertainty. That is no longer the case and has not been for some time. Compelling evidence of this fact can be found throughout the professional literature. One need look no farther than The Association for the Advancement of Cost Engineering International Professional Practice Guide to Risk, a three volume set of 360 articles that address risk and cost estimation. Paul Garvey has recently written *Probability Methods for Cost Uncertainty Analysis, A Systems Engineering Perspective,* a viable textbook resource for a college level course in cost uncertainty. The conclusion is a simple one, cost uncertainty is now a mainstream and important topic.

The U.S. Army Corps of Engineers has not been oblivious to this fact having commissioned the report and case study "Risk Analysis Framework for Cost Estimation" (Yoe, 2000) and a review of ecosystem cost reports in "Analyzing Uncertainty in the Costs of Ecosystem Restoration," (Noble, et al, 2000). These reports build on previous work that is referenced in those documents. The purpose of this report is to draw on the experience and wisdom of this growing literature and summarize the methods that can best be used in analyzing cost uncertainties in ecosystem restoration projects. This has been done by applying risk assessment techniques to the estimation of project costs for a Section 206 Study to restore the in-stream riparian habitat for the brown trout and other species to ecologically sustainable levels along Seeley Creek, an interstate stream in New York and Pennsylvania.

As a relatively new priority output for the Corps' National program, ecosystem restoration projects are challenging for several reasons. First, the projects are often unique. Unlike flood control, where the Corps has many decades of world-class experience, ecosystem restoration projects are often not only one of a kind designs but first of their kind designs. The uniqueness of the measures and the field conditions under which they are constructed contribute significantly to the uncertainty inherent in the estimation of their costs. Second, the study budgets for planning and designing these projects are often limited. Planners must work in data poor environments. This requires cost estimators to make broad assumptions about the details of the plan and the specifics of the design that further contribute to the uncertainty. There is often intense pressure to move forward with ecosystem restoration plans in areas where environmental problems have been exacerbated, in some cases for decades. Ecosystem restoration is popular. Although these, and other, differences in ecosystem restoration projects increase the uncertainty inherent in the estimation of project costs they do not present any unique challenges in terms of the methods, techniques and tools required to address them.

A basic intuitive definition of risk analysis is offered in the next section. The third section identifies some reasons for doing a risk analysis of project costs. The fourth section identifies the most applicable techniques to be used in estimating these cost uncertainties. Section five describes the case study used in this analysis and section six presents the results of the analysis. The report concludes with summary and conclusions. An appendix describes some techniques for quantifying uncertainty that were not used in the case study.

# II. Risk Analysis

Uncertainty is the condition of not being sure. Risk is the chance of a bad thing happening. Analysis is the separation of the whole into its component parts. Risk analysis in the Civil Works Program is a systematic process for describing and quantifying risks associated with processes, actions, or events; taking steps to manage those risks; and communicating about the risks and management actions with all interested parties. Risk analysis, therefore, comprises the three components of risk assessment, risk management and risk communication. A risk analysis of costs necessitates assessing the risks, managing the risks and communicating about those risks. This case study focuses on assessing the risks associated with the estimation of costs for a Section 206 study on Seeley Creek, Bradford County, Pennsylvania. "Costs," as used in this report, refers to the monetary costs of restoration and not the more inclusive definition implied by National Economic Development (NED) costs. The same techniques used here can be readily applied to NED costs.

The language of risk analysis is confusing and messy. Different parties and interests use different definitions to meet their varying needs. This paper offers simple intuitive definitions of the risk analysis components. Although lacking formality they posses a simplicity and a rigor that is consistent with most known, more formal definitions of the terms.

## Risk Analysis

Risk analysis is a decision making tool. It is the cornerstone for decision making under uncertainty. There are many models of risk analysis. The risk analysis model used in the business programs of the Civil Works Program comprises three separate but not always distinct components. They are risk assessment, risk management, and risk communication. For the moment we can think of risk assessment as the technical, analytical work required to describe the major risks and uncertainties of interest in an analysis. Risk management is the process of deciding what to do about the risks that have been assessed. Risk communication is the exchange of information among risk assessors, decision makers, the public and other interested parties throughout the risk analysis.

Conceptually we might represent these components as shown in Figure II-1. The figure indicates the simultaneous and distinct, yet overlapping, nature of the three components of a risk analysis. Although we will present and discuss these components as if they are quite unique, in fact it is often difficult to say where assessment ends and management begins in practice. A risk analysis of an ecosystem restoration cost estimate requires all three of these components.


**Figure II-1. Risk Analysis Model**

## RISK ASSESSMENT

Risk assessment is the component of risk analysis in which analysts describe the risks complete with their associated uncertainties. The product of a risk assessment is (are) the answer(s) to the question(s) asked of the assessment by risk managers. They invariably include a description of what we know about the risks under consideration. Risk assessment is the systematic, scientific characterization of potential adverse effects associated with hazardous substances, processes, actions or events.

At an intuitive level, risk assessment is the work required to adequately answer the following questions of an ecosystem restoration project's cost:

- What can go wrong?
- How can it happen?
- How likely is it?
- How bad can it be?

Ask and answer these generic questions and you have done a risk assessment. Qualitative data and methods lead to qualitative answers and qualitative risk assessment. Quantitative data and methods lead to quantitative risk assessments. The models and methods used to answer these questions are all acceptable, so long as the answers obtained are adequate for decision-making. Risk assessment should include an evaluation of all relevant uncertainties.

## RISK MANAGEMENT

Risk management encompasses the work necessary to adequately answer the following questions of an ecosystem restoration project's cost:

- What specific question(s) do we want the risk assessment to answer?
- What can be done to reduce the impact of the risk described?
- What can be done to reduce the likelihood of the risk described?
- What are the trade-offs of the available options?
- What is the best way to address the described risk?

Risk management is directed at the risks that have been assessed. Risk management does not begin when the assessment ends. It begins when the specific questions to be addressed by a risk assessment are identified. These questions direct the risk assessment. For example, what is the likelihood that our base cost estimate will be exceeded? A good risk assessment directs itself toward answering the questions of concern to decision makers. For our purposes, agency decision makers and risk managers can be thought of as more or less the same. They should get involved from the beginning of a risk analysis by posing the specific questions to be answered by the assessment and then they manage those risks.

## RISK COMMUNICATION

Risk analysis requires a lot of communication. Few cost estimators, for example, consider themselves risk assessors but many of them may eventually be involved in risk assessment. Cost estimators talk to surveys people and geotechnical analysts to decide how best to address uncertainties present in their investigations. Cost estimators talk to economists, their peers and their supervisors. There is a lot of talking that should go on among the study team members to conduct a good risk assessment, to address the uncertainties present, and to manage the risks associated with cost estimating. And then, of course, the results of the risk assessment and the options exercised to manage risks must be explained to the public and others.

Risk communication in general is the work required to answer the following series of questions of an ecosystem restoration project's cost:

- With whom do you communicate?
- How do you get both the information that you need and the information others have?
- How do you convey the information you want to communicate?
- When do you communicate?

# III. Reasons for Doing Cost Risk Assessment

Q: Why do risk assessment of ecosystem restoration cost?
A: To make better decisions.

Traditional, single-point cost estimates are incapable of providing decision makers with such crucial information as:

- The probability of overrunning the cost estimate at all or by some percentage (e.g., the probability of a 20% overrun);
- How much different actual costs can realistically be from the baseline estimate (i.e., exposure to overruns);
- The most important factors contributing to the uncertainty in your ecosystem restoration project costs; and,
- The contingency required to obtain a desired level of confidence in a cost estimate.

For these reasons alone, few people in the construction industry would argue that traditional point estimate cost estimation methods are as reliable for decision making as the probabilistic methods used in risk assessment. The feasibility of projects can be more definitively determined and design alternatives can be more effectively compared, whether it is for value engineering or planning purposes, with cost risk assessment techniques. It is easier to arrange financing and to anticipate budget impacts with full knowledge of the range of potential project costs. Risk assessment of cost estimates enables us to address these and other concerns.

There are two broad categories of reasons for cost risk assessment. They are: (1) improved accuracy of cost estimates; and (2) improved decision-making. Each category is addressed in the paragraphs that follow.

## Improved Accuracy

A point estimate of project costs is very precise. But as long as it is a prediction of a project's true costs we can be virtually assured that it will not be exactly right. If the cost estimators have done their jobs well, the estimate will be close enough to the true cost so as not to cause anyone who uses the point estimate to suffer any extreme consequences. A good risk assessment, however, never fails to encompass the actual costs of a project.

### A Distribution of Costs

In order to understand the points that follow it can help to have some understanding of what a distribution of cost estimates tells us. Imagine preparing a single point estimate of the cost of some project. Because cost estimating is predicting, it is not hard to imagine that if we change one assumption in the cost estimate we might arrive at a somewhat different cost. Imagine all of the different assumptions about quantities and unit costs one could change one at a time and imagine all of the different values one could use for one of those assumptions. Each value produces a different cost estimate. Then imagine all the different combinations of changed assumptions you could make to produce different cost estimates. We would soon have thousands of different cost estimates. Some of them would be more likely than other costs.

Suppose for argument's sake we have 10,000 cost estimates and 1,000 of them are below $6.46 million. Then we could estimate the probability the actual cost of the project will be less than $6.46 million as 0.1 or 10 percent (1,000 chances in 10,000). It may help to think of these 10,000 cost estimates as you consider the points made below. Selected costs from a 10,000 iteration Monte Carlo simulation of project costs for Seeley Creek are presented in Table III-1 below. Costs are no longer point estimates; they are a distribution of many possible costs. Costs are shown in Figures III-1 and III-2.

**TABLE III-1**

**SELECTED PROJECT COSTS IN MILLIONS**

| Item | Cost | Item | Cost | Item | Cost |
|---|---|---|---|---|---|
| Minimum Observed Cost | $5.75 | 30th Percentile | $6.68 | 70th Percentile | $6.99 |
| Maximum Observed Cost | $7.69 | 35th Percentile | $6.73 | 75th Percentile | $7.03 |
| Mean Observed Cost | $6.83 | 40th Percentile | $6.77 | 80th Percentile | $7.08 |
| 5th Percentile | $6.35 | 45th Percentile | $6.81 | 85th Percentile | $7.13 |
| 10th Percentile | $6.46 | 50th Percentile | $6.84 | 90th Percentile | $7.20 |
| 15th Percentile | $6.53 | 55th Percentile | $6.88 | 95th Percentile | $7.30 |
| 20th Percentile | $6.59 | 60th Percentile | $6.92 | | |
| 25th Percentile | $6.64 | 65th Percentile | $6.95 | | |

## Probability of Costs Exceeding Our Estimate

There is no objective way to estimate the probability that the single-point cost estimate prepared via traditional methods will be exceeded. Risk assessment of a cost estimate can produce a distribution of total costs or a distribution for any cost element or subset of total costs. It is simple and straightforward to obtain quantified estimates of the likelihood that a cost estimate will be exceeded when we have a distribution of costs. Not only can we estimate the probability that any particular cost estimate will be exceeded, we can estimate the probability it will be exceeded either by a given percentage, such as 20%, or by a given amount, such as $1 million.

In this example of 10,000 costs the mean is the baseline or best guess cost in this case equal to $6.83 million. In the simulation 5,137 cost estimates exceeded that value so there is a

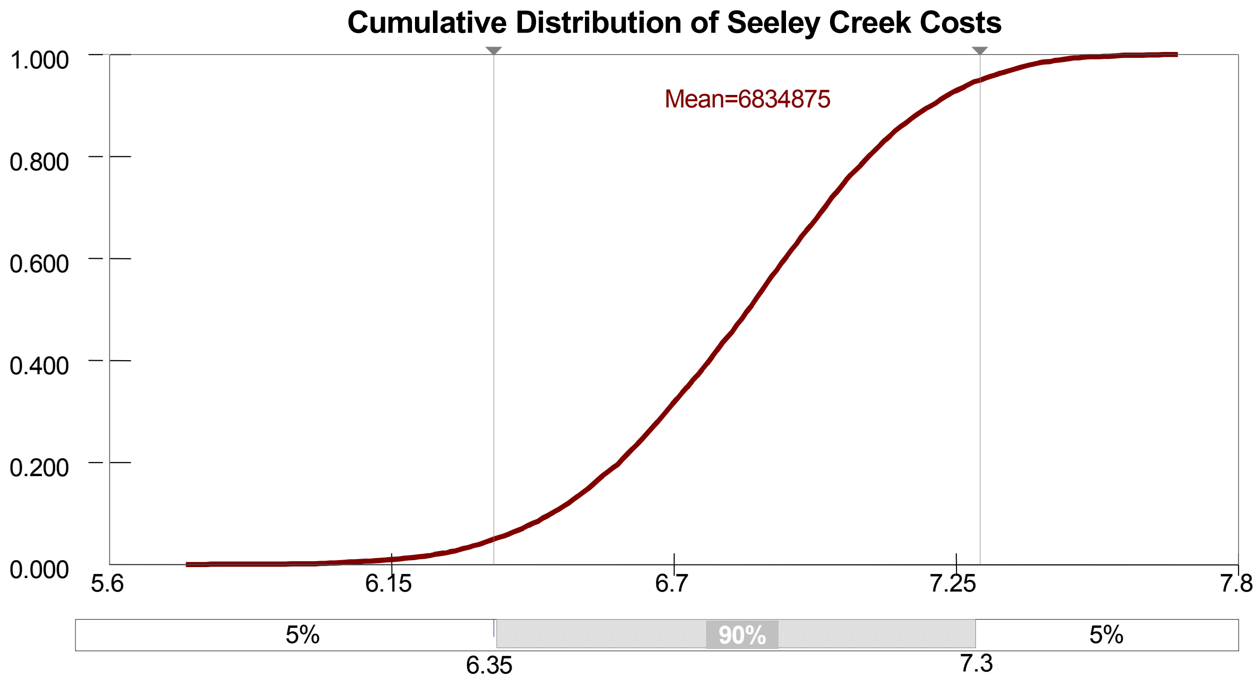



**Figure III-1. Empirical Distribution**

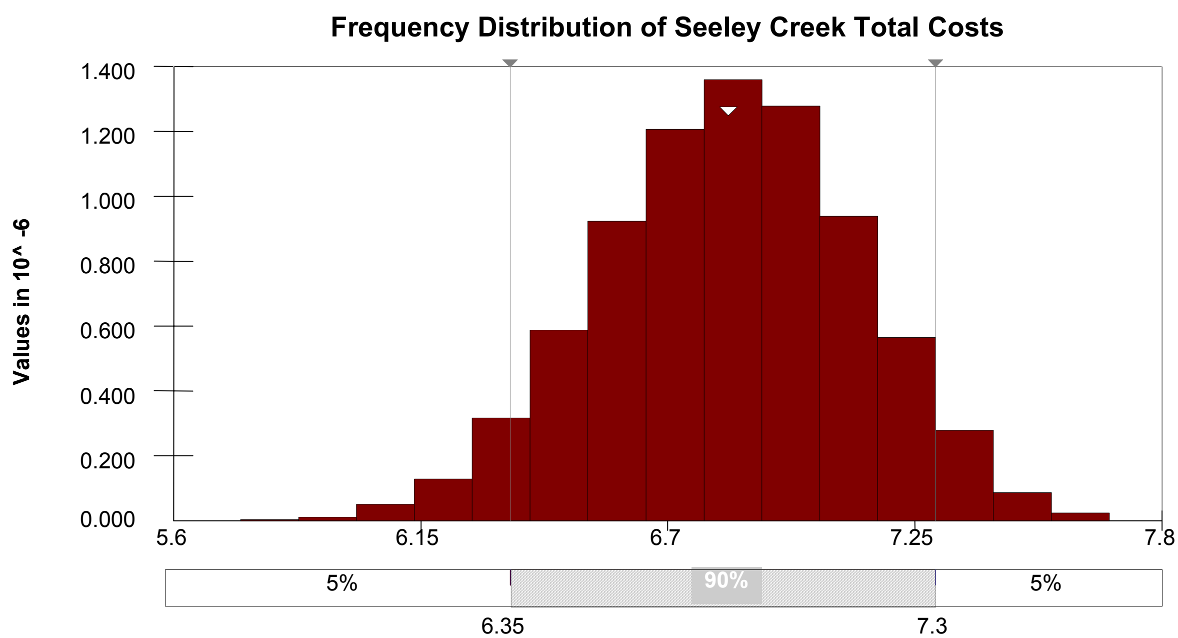



**Figure III-2. Frequency Distribution**

0.51 probability of a cost exceeding the mean estimate. To calculate the probability of a 20% overrun of the mean simply count how many cost estimates exceed $8.20 million (120% of $6.83

million). In this case there were no estimates above this amount. Thus, there is no chance of costs exceeding the mean estimate by 20 percent or more. In a similar fashion we can estimate the probability of a cost of $7 million or more, or any other cost, by counting the values in the range of interest. In this example, there is a 29.07 percent chance (2,907 of the 10,000 estimates actually exceeded this amount) that costs will exceed $7 million.

The percentile values in Table III-1 reveal additional probability information. For example, the 45$^{th}$ percentile is $6.81 million. This means 45 percent of all our costs estimates were $6.81 million or less.

## Estimate Exposure

There is no objective way to estimate one's maximum exposure to cost overruns using the traditional single-point cost estimate. Exposure is defined as the difference between the single-point estimate and the highest realistic estimated cost (Curran and Rowland, 1990). If costs might overrun the estimate it is important to know just how bad the overrun could be. Risk assessment provides a methodology that enables the cost estimator to estimate the Corps' and non-Federal partner's exposure.

In the example of 10,000 cost estimates, look at the maximum cost estimate of $7.68 million. The difference between the best guess cost estimate ($6.83 million) and this maximum value, or $0.85 million, is the maximum exposure to the risk of a cost overrun. Because this maximum cost occurred once in 10,000 estimates the probability of such an extreme exposure is 0.01 percent. More likely overrun risks may be of more interest. The 95 percent exposure to cost overrun risk is $0.47 million, substantially less. There is a 5 percent risk of incurring an overrun of $0.47 million or more.

## Identify Key Components In Exposure

If the probability of any particular overrun is considered too great or if the exposure is unacceptable it is in the decision makers' best interests to know how best to reduce that probability or exposure. That could be readily done if the factors (quantities or unit costs) that contribute most to an overrun or its probability could be identified. Traditional cost estimating techniques provide no systematic means of determining the key cost factors under conditions of uncertainty. Cost estimating models are often too complex to lend themselves readily to such an analysis of key components[1].

Risk assessment lends itself readily to such techniques. The results of an importance analysis for the Seeley Creek project are presented in Figure III-3. The labels on the left of the graph identify cost estimate inputs in order of their contribution to the range in potential total costs by their specific location (i.e., cell address) in the cost estimator's spreadsheet.

[1] In practice, cost estimators are often able to identify the most relevant uncertainties.

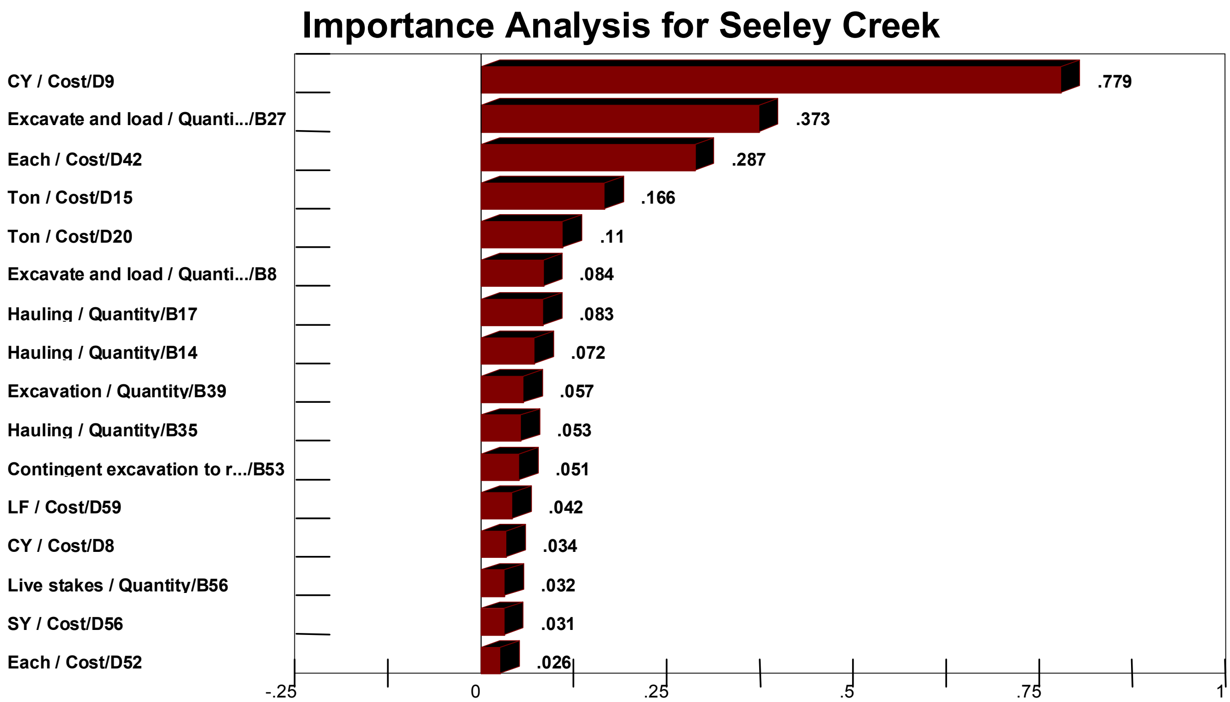



**Figure III-3. Importance Analysis**

This analysis shows that the unit cost of excavating loose rock from the channel is the single most important contributor to the variation in the total costs as shown in Figures III-1 and III-2 above. The quantities of materials required for the stone to revetment is the next greatest contributor to the variation in total costs, followed by armor stone hauling costs and boulder placement costs. This suggests that if we would like to narrow the uncertainty in the final cost estimate as shown in Figures III-1 and III-2 we should reduce the uncertainty in the loose rock excavating costs and the revetment associated quantities first.

## Full Knowledge of Estimate

One of the most enduring and irrefutable points made about single-point cost estimates is that they fail to reveal all that is known about a cost estimate. Curran (1989) says that typically we harness rivers of data, we filter it, we polish it, we reflect upon it, then, finally, we make our selection of "the right number," holding it up for all to see. When that winnowing process is complete the only certainty we can assign to this value is that it is going to be wrong. The actual cost will either be higher or lower than the estimated value.

The point estimate is a single mythical value that masks a great deal of what is known about project costs. We have to filter and polish a great deal of information away in order to get to a single number and no information should be ignored in such an uncertain venture as ecosystem restoration cost estimation. The world is full of probabilities and ranges of possibilities, not single-point numbers waiting to be counted with certitude. Risk assessment of cost estimates can describe the variation in possible cost outcomes. They can be used to answer the questions: What can go wrong? How can it happen? How likely is it? How bad can it get?

Costs could be as low as $5.75 million or as high as $7.68 million, a range of $1.93 million. The range indicates the potential for a wide variety of possible cost estimates. Using the interquartile range between the $25^{th}$ and $75^{th}$ percentiles we see that a full half of all of the cost estimates were within $0.33 million of one another. So although there is considerable overall variability, meaning extreme cost values vary widely, the most likely costs vary much less.

## IMPROVING DECISIONS WITH COST ESTIMATE RISK ANALYSIS

By reducing and addressing the uncertainty about relevant information in the risk assessment step of a risk analysis we can presumably make better decisions in the risk management step of a risk analysis than we would if we ignored that uncertainty or remained unaware of it. An ecosystem restoration cost estimate can be used in a cost-effectiveness analysis to determine whether or not a project is feasible and eligible for Federal support. It can also affect its eligibility for construction under specific Corps programs. Local partners decide whether or not to participate in a project based on its cost. The cost estimate is used as the basis for cost sharing arrangements. Cost estimates form the basis for budget requests. Contractors decide whether or not to bid on construction contracts based on cost estimates. Cost estimates are used as the basis for construction contracts and cost estimates are used to manage project costs.

These activities entail a great many significant decisions. If the cost estimate is inaccurate, mistakes can be made, some of them significant. One of the most important reasons for knowing the accuracy of a cost estimate is the impact this information has on an agency's or company's management as well as their policies and philosophy with regard to cost engineering. People know what to expect when an estimate is presented to them for review if the confidence level in that estimate is known. The problems associated with providing a good estimate, particularly at a concept stage will be better understood, anticipated and appreciated. As the methods for estimating and reporting costs evolve and change, so too can the agency's or company's policies and philosophies change to accommodate the new and improved information.

Risk assessment of costs will provide increased accuracy. Costs can be estimated with ranges, distributions, confidence intervals, and the like. Contingencies can be estimated with greater confidence using risk-based techniques.

## Contingencies with Confidence Intervals

The Corps has long used contingencies in its cost estimates quite successfully. What these traditional methods of contingency estimation did not enable cost estimators or decision makers to do, however, was to understand the confidence associated with that contingency. With cost risk assessment, it is possible to select a contingency so you are 80, 90, 95, 99 or any other percent sure your cost estimate will not be exceeded. Selecting a contingency to acquire a desired level of confidence in a cost estimate is possible with cost risk assessment but not under traditional techniques.

Consider Table III-2 shown below, based on the data presented above. The baseline cost estimate is the expected value, $6.83 million in this case. The median cost is $6.84 million. The actual cost is as likely to be less than that as more than the median cost, it is the 50th percentile. In order to manage the risk associated with cost estimates that underestimate the actual costs, the Corps' cost estimators add a contingency to their cost estimate. Contingencies represent allowances to cover unknowns, uncertainties, and/or unanticipated conditions that are not possible to adequately evaluate from the data on hand at the time the cost estimate is prepared but must be represented by a sufficient cost to cover identified risks (ER 1110-2-1302 12.a.). They are currently determined based on professional judgment. They are sometimes added to individual quantity or unit cost estimates and/or as a lump sum adjustment to total costs.

**TABLE III-2**

**CONTINGENCIES WITH CONFIDENCE LEVELS ($Millions)**

| Desired Confidence Level | Required Contingency % | Contingency Amount | Cost Estimate |
|---|---|---|---|
| 60 | 1.2% | $0.08 | $6.92 |
| 70 | 2.3% | $0.16 | $6.99 |
| 80 | 3.6% | $0.24 | $7.08 |
| 90 | 5.4% | $0.37 | $7.20 |
| 95 | 6.7% | $0.46 | $7.30 |
| 99 | 9.1% | $0.62 | $7.46 |

In this example, the expected value or best estimate of costs is $6.83 million. Eight thousand of our 10,000 cost estimates were $7.08 million or less. Hence we are 80 percent (8,000/10,000) sure the actual cost will be $7.08 million or less. In order to be 80 percent sure costs do not exceed the estimate we would use $7.08 million as the estimate. This requires a contingency of $0.24 million or 3.6 percent of the baseline cost estimate (i.e., the mean) to achieve the desired confidence level of 80 percent. That is considerably less than the 15 percent contingency that is commonly used for costs at this stage of estimation.

## Comparing Alternative Designs

How do you evaluate the costs of two alternative plans or two different designs when one is well known and the other is an experimental design that relies on new technology? Which cost is more uncertain? Suppose the familiar project has a slightly higher single-point cost estimate. Should it be chosen?

Consider the hypothetical data of Figure III-4. Suppose for simplicity that Projects A (the tighter distribution) and B (the wider distribution) will achieve the same outputs and are equal in all significant respects except for costs. Project A has a mean cost of $2 million with a narrow distribution of potential costs because it is a well-known design relying on time-tested technology. Project B, although it has a slightly lower expected cost at $1.9 million, has the potential to cost a great deal more than Project A. The uncertainty in its costs due to the novel design and technologies are much greater. Project B also has a greater potential for lower costs as well. The better choice depends, to some extent, on the managers gambling preferences. If concern for cost overruns is a principle decision factor then it is better to go with Project A. Managers guided by expected values will go for Project B. Those compelled by the possibility of B costing significantly less than A would also prefer B. The choice is largely a function of one's risk preferences.

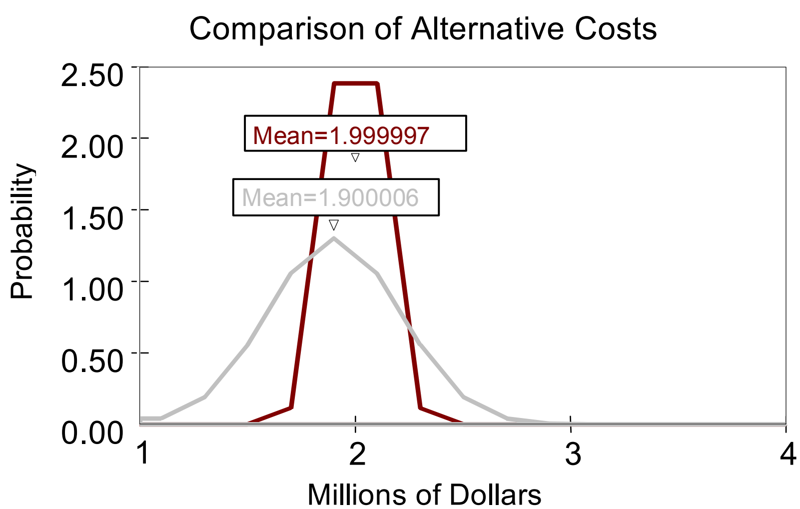



**Figure III-4. Cost Comparison**

A better-informed comparison of project costs requires a distribution of costs. Comparing single-point cost estimates can be misleading during plan formulation and cost-based screenings of alternative courses of action. To avoid disastrous surprises a comparison of risk assessment cost estimate results is preferred. The Seeley Creek case study had only one alternative plan at the time of this analysis, so an actual cost comparison was not possible.

## Feasibility

Traditional Civil Works Program projects rely on a National Economic Development benefit-cost analysis. In these instances the distribution of costs for the alternative plans can be combined with the distribution of benefits to produce a probabilistic estimate of the net economic benefits and benefit-cost ratio. Ecosystem restoration projects do not require an explicit benefit-cost analysis.

The distribution of project costs for an ecosystem restoration project can be used together with a distribution of project outputs to produce a distribution of incremental costs. A cost risk assessment is an essential step toward the Corps desired use of risk-based economic analysis for decision-making. Plans exist to revise IWR-Plan, an incremental cost analysis tool, to include the capability of addressing the uncertainty in ecosystem restoration costs and outputs.

## Arranging Financing

Cost estimates provide the basis for non-Federal partners to decide first, whether or not to participate in a project and second, to arrange their financing of a project if they do participate. A single-point cost estimate gives partners a target level of financing that will ultimately be either too low or too high. An estimate that is too low could jeopardize the partner's ability to support the project. It could also damage the Corps' credibility for future cooperative efforts. A cost estimate that is too high could discourage a partner's participation in a project.

A risk assessment estimate of costs can help partners manage that risk, make better decisions and better arrange for the proper level of funding. The choice of financing vehicle, e.g. tax revenues vs. bonds, may well depend on the actual cost to the partner. With a risk-based estimate of costs, partners can better gage their ultimate share of the costs and the funding vehicle to choose. They can also better anticipate the likelihood and impact of overruns on their budget in the near and long term.

An added advantage is that with more complete information a non-Federal partner can examine the data and apply their own confidence level parameter. For example, if the Corps chose an 80 percent confidence level and their partner prefers a 95 percent level for their financial planning purposes nothing prevents the two partners from using the same information differently. Thus, the Corps might proceed based on a cost estimate of $7.08 million while the partner uses $7.30 million as its financial target. It stands to reason that a small local government might be more risk averse than the Corps of Engineers. Cost risk assessment provides better information for partners.

## Construction Profitability

Although the Federal and non-Federal partners are normally non-profit entities the construction companies that build Corps projects are not. Contractor failures hit a peak in 1975 (Engineering News Record, 1977). In 1976 about 90 percent of construction contracts fell short of their expected profitability (Lewis, 1977). As early as the mid-seventies the inadequacy of the single-point cost estimate was coming under fire. The ENR said the impact of uncertainty was felt nowhere more than in the construction industry. Increasing competition has forced narrower profit margins that have continued to the present. This increases the need for greater bid accuracy. Single-point estimates simply do not provide the accuracy required for construction firms to maintain their profitability.

Suppose, for example, a contractor felt the project would cost him $7 million to build. The cost data suggests the actual cost, hence a government cost estimate, has about a 29.1 percent chance of equaling or exceeding that amount. Hence, the contractor might decide not to bid the project. Contractors are forced to live with risk and uncertainty as a daily way of life. Cost risk assessments provide more and better information to contractors. The probability of overruns and exposure to cost risk can threaten not only profitability on a single project but the very viability of a firm. The Corps and its partners owe the construction industry the best information possible for greater bid accuracy.

## Aids Cost Management

Risk assessment of project costs has the capability of identifying the most critical components of a cost estimate. Through a variety of sensitivity techniques it is possible to determine which cost components have the greatest potential to affect project costs favorably or unfavorably. The importance analysis above identifies those components whose uncertainty should be reduced to provide better cost estimates. An alternative use of that analysis is for risk management. Any contractor who can find an effective way to lower the costs of excavating loose rock from the channel has a greater chance of successfully bidding the job or making a profit on its construction. If the uncertainty cannot be reduced prior to construction, these components are identified as in need of careful management during construction in order to keep costs to a minimum. Thus cost risk assessment aids both the cost estimation and cost management functions of the cost engineer.

## Useful Throughout Life Of Project

Risk assessment of project costs provides information that can be used during the earliest stages of a project's life in plan formulation and in deciding whether to proceed with an alternative or not. As decisions to proceed are made, the same information can be used, during the arrangement of financing, in bid preparation, and in cost management. Risk assessment separates what is known from what is not known. Probabilistic methods are used to express those things that are not known.  Importance analysis identifies the most important of the uncertain factors and this is extremely useful in directing the expenditure of study funds to reduce the uncertainty in the total cost estimates. Potential cost uncertainties can be better investigated during design and specification stages and they can be more carefully monitored and managed during construction to hold down costs.  Thus, risk assessment of costs serves the construction project better than a point estimate from concept through completion.

# IV. TECHNIQUES

The techniques used to identify and describe the uncertainty inherent in a risk assessment of ecosystem restoration costs are the same techniques that would be used for any cost estimation purpose. They are simply adapted as necessary for the unique aspects of ecosystem restoration. The basic techniques that can be applied are sensitivity analysis, Monte Carlo simulation, and range estimation. In all candor, it is a foregone conclusion that Monte Carlo simulation will be used to estimate ecosystem restoration costs in most instances. Indeed, Monte Carlo simulation was used to generate the values presented in the last section. Range estimation, a technique that once garnered much attention in the cost estimation literature has been superseded by the commercially developed software (spreadsheets and Monte Carlo process add-ins) that supports Monte Carlo simulation in most arenas.

It is worth noting that other techniques exist. It has been suggested in discussion of this project with peers that fuzzy sets could be used to a better advantage than Monte Carlo simulation. Another person suggested that Bayesian hierarchical models might be useful. These techniques may ultimately prove to be of great utility but the moment belongs to Monte Carlo simulation. Sensitivity analysis and range estimation are briefly mentioned before the Monte Carlo process is explained.

## SENSITIVITY ANALYSIS

Sensitivity analysis in a risk assessment context is the systematic variation of assumptions, models, model inputs and parameters in order to examine the impact of these changes on the outcome of the risk assessment. It is rather unusual for sensitivity analysis to consider alternative models, especially for cost estimation. Hence, most sensitivity analysis will involve alternative assumptions and alternative input and parameter values. A common form of sensitivity analysis involves the creation of scenarios. When assumptions, parameters and inputs are systematically changed to describe some scenario such as an optimistic or pessimistic scenario, costs are estimated consistent with these assumptions and values. The systematic variation of assumptions and values is repeated for as many scenarios as desired.

One essential caveat of any sensitivity analysis is that each scenario investigated must be possible and realistic. Worst and best case scenarios are sometimes possible but they are so unlikely, so improbable, as to fail the test of realism. And although there is a clear distinction between what is possible and what is probable that distinction is not always or even often recognized by those unfamiliar with risk assessment techniques.

Many Districts already use some sensitivity analysis. It can be considered a minimalist investigation of the uncertainty inherent in the preparation of a cost estimate. It is not a true risk assessment because it does not enable us to estimate the likelihood of these different events' occurrence. Investigation of specific scenarios in the context of a true risk assessment, however, can lend a helpful dimension of information to any cost estimate.

## RANGE ESTIMATION

Range estimation is an alternative to Monte Carlo simulation. It was developed by and for cost estimators. Interest in range estimating arose in the 1980s and seemed, if the literature is a reasonable gage, to have peaked in the early to mid 1990's, as the commercial Monte Carlo software became more user friendly and available. Some people find range estimating more intuitively appealing and consider it easier to develop the input data needed to use it. It requires either the proprietary software of a single firm to run or the user must develop his/her own software. In either case the tools are not as readily available as the Monte Carlo software.

Range estimating is driven in part by the notion that analysts unfamiliar with the sometimes-complex properties of probability distributions could misuse Monte Carlo methods of analyzing costs. Poorly specified uncertainties, for example using an inappropriate distribution to describe the uncertainty in an input (see Appendix), could result in model outputs (i.e., cost estimates) that are misleading. In lay terms, some people are concerned that if the uncertainty in estimates of unit costs and quantities is exaggerated so will be the potential range in project costs. Consequently, some of the risk that appears evident will in fact be iatrogenic risk, i.e., a result of the method used to estimate the risk.

According to a description of range estimating by one of the method's principle proponents, range estimating uses a simple but effective measure of uncertainty: the range. The range is specified with four parameters: the *probability* that the element's actual value will be equal to or less than its target value, a target value, a *lowest* estimate, and a *highest* estimate (Curran, 1989).

Suppose a work element has a target value, i.e. best guess or most likely value, of $10.05. In range estimating the estimator is asked to estimate the probability that the element's actual value will be less than or equal to the target value. So let us estimate that probability as 75 percent[2], the lowest value as $7.80 and the highest value as $14.35. The probability measures the likelihood of an underrun while the lowest and highest values measure the degree of underrun and overrun. Specification of the range is to take all foreseeable circumstances into account. The range is considered far more valuable for decision making than any single number from within it.

Curran describes range estimating as a synergistic combination of Monte Carlo simulation, sensitivity analysis, and heuristics that introduces ranges and other data into a personal computer to obtain the desired results. Although a detailed description of the range estimating algorithm is beyond the scope of this report it yields results conceptually very similar to those produced in a Monte Carlo simulation.

Because range estimating was developed for cost estimating it once had the advantage of offering outputs that Monte Carlo simulation did not. With the advent of commercially available Monte Carlo software it no longer enjoys that advantage. The principle advantage of range

[2] In other words, there is a 25 percent chance the target value will be exceeded.

estimating appears to lie in the belief that estimators will find it easier to estimate the four parameters of a range than the parameters of a distribution as required in Monte Carlo analysis.

## MONTE CARLO SIMULATION

The preferred method of assessing the risks in estimating the costs of an ecosystem restoration project is to calculate the costs for hundreds or thousands of possible scenarios and then to study the results of those many calculations. From the thousands of possible cost estimates we can learn what can go wrong, how it can happen, how likely it is and the consequences as well. What is needed, however, is a reliable and cost effective method for calculating these thousands of estimates. The Monte Carlo process[3] is one such method.

During the development of the atomic bomb it was necessary to simulate a wide variety of circumstances given the theoretical uncertainties of the time. The Monte Carlo process was used and refined to develop values of random variables. It is essentially a sampling process that is a method for generating random values of a random variable based on a probability distribution. It consists of two general steps. First, a random variable value is generated, usually on the interval [0,1]. Second, this value is transformed into a useful value for the problem at hand.

To illustrate the idea, consider the mid-square method[4] of generating a random variable. Suppose we use a seed value of 4745. Square it and take the middle four numbers, 22515025 and divide them by 10,000. Our random value is 0.5150. But in how many problems will that number be relevant? We need to transform it into a useful value.

Suppose we are trying to estimate the number of hours it will take to fill a geotube in the field. Further suppose our best estimate based on limited historical experience indicated that is it will take between 10 and 50 hours. If we are trying to generate a possible time to fill the geotube we need a number between 10 and 50, not a number like 0.5150. Assuming the number of interest has a uniform distribution[5] we can convert our random number using the formula:

$$x = a + (b - a)u \tag{1}$$

where x is a random number between 10 and 50, a is the minimum value (10), b is the maximum value (50), and u is the random number generated over the interval [0,1]. Through simple substitution we get:

$$30.6 = 10 + (50 - 10).5150 \tag{2}$$

[3] This includes the closely allied Latin Hypercube process of sampling.

[4] The mid-square method attributed to John von Neumann was one of the early methods developed to generate random variables. It was soon abandoned because it does not generate true random variables. It is sufficient for our heuristic purposes here, however.

[5] We assume a uniform distribution to keep the arithmetic simple and not because it is the way such a problem should be approached..

Thus, we assume it takes 30.6 hours to fill the tube. The Monte Carlo process is simply a technique for generating random values and transforming them into values of interest. The process continues by squaring 5150 to get 26522500 substituting .5225 into equation (2) and repeating the process as often as desired. The methods of generating random or pseudo random numbers are more sophisticated now and the mathematics of other distributions is more complex, but the process is similar to that in the simple example.

Imagine a cost-estimating model in a spreadsheet software package. Individual numbers in a cell can be replaced by a distribution. For example, if the number of hours required to fill a length of tube were assumed to average 21 hours, 21 would appear in a cell in the model. Now imagine that we replace that single number with a uniform distribution that says the actual average number of hours is unknown but it is believed to be between 10 and 50. The choice of a uniform distribution implies that any number in this range is as likely as any other number. When there are more complex relationships among the unknown values, such as some numbers are very likely and others are extremely unlikely, other kinds of distributions are used.

Imagine the Monte Carlo process generating a number like 30.6 that is used in the cost estimate. Now imagine that a new random number is selected and transformed into a random number between 10 and 50 and costs are calculated with this new number. Let us keep track of all the random numbers of hours it takes to fill the geotube and the resulting costs associated with them. By examining several thousands of these numbers we can learn a great deal about our cost estimate.

Each new calculation of the cost is called an iteration of the model. A simulation is a collection of many iterations. Many simulations employ this Monte Carlo process and they are often called Monte Carlo simulations. Although that is strictly a misnomer (it is a simulation that uses the Monte Carlo process) it is common usage. There are many kinds of simulations that have nothing to do with the Monte Carlo process. The Corps' ship simulators at the Waterways Experiment Station in Vicksburg are but one example.

To develop some intuition for this tool consider a project that requires pouring concrete. There are two input variables, the quantity of concrete and the inclusive costs of placing it. Suppose both the quantity and cost of the concrete are uncertain. Our best guess is that 1,000 cubic yards of concrete will be needed and it will cost $100 per cubic yard. The resulting cost estimate is $100,000. A simple spreadsheet model is shown below in Table IV-1.

**TABLE IV-1**

**COST MODEL**

| Concrete (cy) | Cost per cy | Project Cost |
|---|---|---|
| 1000 | $100 | $100,000 |

Let's introduce a little sensitivity analysis. Suppose we are sure we will need at least 800 cy of concrete and no more than 1,100 cy. Furthermore, we know it will not cost less than

$95/cy but it could cost as much as $200/cy to place it. The best-case possible is small quantities and low costs; the worst case is just the opposite. The best case and worst-case scenarios result in costs of $76,000 and $220,000. Although we have done a decent job of bracketing what costs could be we have no idea how likely either of these extreme scenarios will be.

If we want to incorporate the Monte Carlo process into the model shown in Table IV-1 we must replace one or more of the input variables with distributions. So what distribution will we use? Building on what we have said to this point we have quantities ranging from 800 to 1,100 cubic yards. Do we know anything else about these quantities? Yes, we know that 1,000 cy is the most likely value of all. Minimum, maximum, and most likely values are enough to define a triangular distribution. For simplicity, we'll use that. Likewise we can describe our uncertainty about unit costs with a triangular distribution. Costs are assumed to have a minimum of $95, a most likely value of $100 and a maximum of $200.

Using commercially available software we can replace the point estimates of Table IV-1 with two triangular distributions. A Monte Carlo process takes a random number between 0 and 1 and transforms it into a number from the interval [800,1100] according to the rules of the triangular distribution used for the quantity estimate. It would do similarly for costs. These two randomly selected values are multiplied together and produce one possible cost for this project. This process, repeated 10,000 times, is summarized in the Figure IV-1 below, a graphic representation of a Monte Carlo version of the spreadsheet model above.

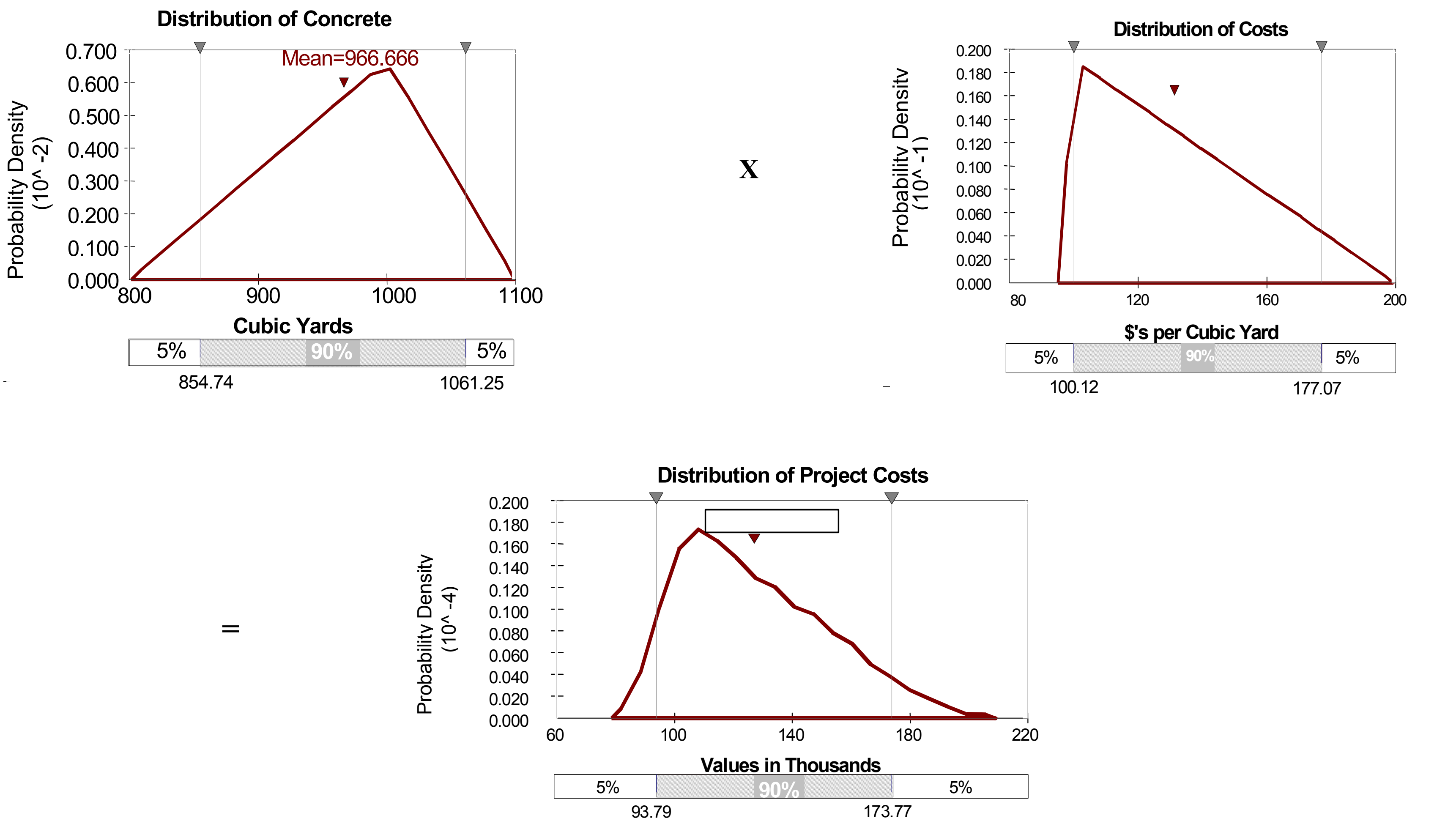



**Figure IV-1. Monte Carlo Process**

A value is randomly selected from the distribution on the left while a second value is independently selected from the distribution on the right[6]. They are multiplied together and the process was repeated 10,000 times to generate the distribution of costs at the bottom of the figure.

This analysis shows costs as low as $78,597 and as high as $209,035; quite a bit different from the best and worst case scenarios. This suggests the chance of either of those extreme scenarios identified in the sensitivity analysis is less than 1-in-10,000. The mean of the 10,000 costs was $127,278, more than our original best estimate. This is because the expected values of the input distributions were different from the best guess point estimates. There is an 11.3 percent chance costs will be less than $100,000, not a very likely outcome. We are 90 percent sure the costs will be between $93,790 and $173,770. The analysis showed the uncertainty in unit cost to be far more significant in determining total costs than the uncertainty in quantities. Thus, if we had resources to refine estimates for only one of these variables it would be better to refine the cost data than the quantity data. That is a simple Monte Carlo simulation.

Ecosystem restoration cost estimates are more complex than this. But virtually all of them can be reproduced in a spreadsheet model. If so, commercial software can be used to produce a risk assessment using the Monte Carlo process. Special software applications can be developed to add Monte Carlo capability to virtually any cost estimating program. Conceptually, there is nothing that would prevent the incorporation of the Monte Carlo process into the Corps' M-CACES, CEDEP or other cost estimating programs.

[6] The values selected from the two distributions can be independent of one another, as was the case for this example, or they can be dependent upon one another in a number of ways.

# V. CASE STUDY CHOICE

A case study for this research was identified by calling several Corps Districts and inquiring about what candidate studies were available and which Districts were interested in participating in this research. The Baltimore District responded quickly and enthusiastically. After a preliminary meeting with District personnel to explain the purpose of the research the District expressed a desire to participate in this effort. Following a second meeting, the District identified a section 206 Study for Seeley Creek, PA as their candidate project.

The Seeley Creek watershed is 134 square miles in Chemung and Steuben Counties, New York, and Tioga and Bradford Counties, Pennsylvania. There are approximately 175 stream miles in the watershed. In Bradford County, Seeley Creek has 29 stream miles and its major tributary, South Creek, has 49 stream miles. The objective of the Corps of Engineers' Aquatic Ecosystem Restoration Project conducted under the continuing authority of Section 206 of Water Resources Development Act of 1996 is to restore the in-stream riparian habitat for the brown trout and other species to ecologically sustainable levels.

Plan formulation has relied upon bioengineering and natural analogy channel design techniques used to restore the stream sites along a two mile stretch of the stream in Bradford County. Design analysis relied extensively upon analogy techniques to mimic the habitat conditions in the non-impacted streams in the immediate vicinity. Comparisons of pre- and post-project environmental outputs and community diversity will be used to assess the projects' success and need for adaptive management and monitoring plan. The project consists of three sites described below in the District's words:

"This project area is located at the mouth of Seeley Creek above the confluence with Hammond Creek in Bradford County, Pennsylvania. While it should be noted that the entire watershed is dynamic to various degrees, the current study focuses on three separate areas that will directly stabilize and restore approximately 4,500 feet of stream habitat. Additional environmental benefits will be accrued through an overall reduction in sediment to the lower watershed from the eroding banks. A 3,500-foot channel realignment and bank protection project has been investigated from just upstream of the State Route 328 bridge at the stream mouth to just below the T-763 bridge. Two additional steep slope bank erosion areas addressed by this project are above Route 549 and above T-763. For descriptive purposes in the current phase of the study, these areas have been designated as Area I (a 550 foot reach above Route 549), Area 2 (a 400 foot section of the T-763 bridge), and Area 3 (the realignment reach). Total distance from the lower end of area 3 to the upper end of Area I is approximately 6,600 feet."

The project consists of grade control weirs and tie-back dikes, stone toe revetments, earthwork to form the channel, gas line relocation, in-stream habitat creation and riparian plantings. This project is currently estimated to cost $8,001,069, based on the Corps' traditional point estimate of project costs.

# VI. SEELEY CREEK COST RISK ASSESSMENT

Q: What can go wrong?
A: There could be an overrun on the costs of the Seeley Creek project.

Q: How can it happen?
A: Actual quantities may exceed the estimated quantities. Unit costs may be higher than estimated. Engineering and design, construction management or price escalation could be more than anticipated.

Q: How likely is it?
A: This is difficult to answer in a generic fashion. It depends on the District's cost engineering philosophy.

Q: How bad could it be?
A: Some overruns are negligible and others are significant. The consequences could be that a project is scrapped and never built. More often it involves embarrassment and difficulty in negotiating the changes. Sometimes, in order to avoid that embarrassment and difficulty costs are "conservatively estimated." That is, costs are estimated on the high side. The downside of this practice is that unrealistically high costs can discourage participation in project construction. When benefit-cost analysis is required the economic feasibility of a project may be jeopardized by such a conservative estimate.

The Seeley Creek cost risk assessment would begin with some specific questions posed by the District's decision makers serving a risk management role. These questions would be specific to the circumstances of the Seeley Creek project. This was a research project, however, and there is no established culture of doing cost risk assessment. Hence, there was no one in a position to articulate the questions that might have guided the assessment of cost risks. But if you return to the intuitive definition of risk management in Section II you will see it begins with the generic question, "What specific questions do we want the risk assessment to answer?"

This is a critical step in the larger risk analysis. Managers have questions they need to have answered so they can properly complete the planning process. These questions need to be specifically articulated for the risk assessors so they can be sure to address them in their assessment. In practice it is likely that a set of routine questions might emerge. These could include:

- What is our best unbiased estimate of project costs?
- What is the maximum likely overrun of our best estimate?
- What are the most uncertain unit costs?
- What are the most uncertain quantities?
- What unit costs contribute the most to the variation in total costs?

- What quantities contribute most to the variation in total costs?
- If we were to do further analysis of the cost estimate on what cost components should we allocate our resources?
- What cost estimate is consistent with a 10 percent or less chance of a cost overrun?

Questions unique to a project design might also arise in some situations. Choices among specific designs or specific technologies might emerge as might questions of timing and so on. Seeley Creek is a relatively simple project. There is only one alternative under consideration. In the present case there were no specific questions articulated prior to the assessment. Hence we revert to the questions posed above as they are likely to always be of interest in ecosystem cost estimation. But we do want to stress that a set of questions such as these should be prepared prior to initiation of the cost risk assessment and communicated to the risk assessors. This is essential to ensure that decision makers get the information they need to make the decisions necessary to execute their mission.

## VARIATION IN SEELEY CREEK COSTS

Let us begin at the end of the cost risk assessment. Figures II-1 and II-2 are reproduced below as VI-1 and VI-2. Bear in mind that the uncertainty in the quantity and unit cost inputs to the cost estimate have been described as distributions, a point explained at length later in this section. As a result these figures present the variation that could exist in the costs of the Seeley Creek project.

Look at Figure VI-2. The figure covers a span of the number line, where we display estimated total costs, between \$5 and 8 million. What does the histogram that shows the results of 10,000 calculations of the Seeley Creek cost estimate tell us? First, we notice that estimates near \$5 or 8 million dollars do not appear. The histogram peaks in the vicinity of \$6.8 million and most of the estimates seem to fall between \$6.6 and 7.0 million. The distribution looks reasonably symmetrical and the likelihood that costs will fall below \$6 million are much less than the chance they'll be more than \$7 million. Figure VI-1 suggests we can be 100 percent sure the cost will be less than about \$7.7 million. The 90 percent confidence interval designated on the bar below the graph suggests we can be 90 "sure" costs will be between \$6.35 and 7.30 million. All of these results, of course, are contingent upon the reasonableness of our model. The model is also discussed at length below.

The big picture conveyed by these two figures is the simple truth that we do not know what the actual costs will be. They are uncertain. We do have an idea what the most likely value is, it is about \$6.8 million dollars. But costs could be anywhere from about \$5.8 to 7.7 million. We also know some cost estimates are more likely than others. The shape of our histogram tells us that.

There are two very useful things we can now do with these graphs. Suppose for argument's sake that we find the spread in possible costs to be too large. Suppose we are uncomfortable proceeding with a project with costs this uncertain. What are our options?

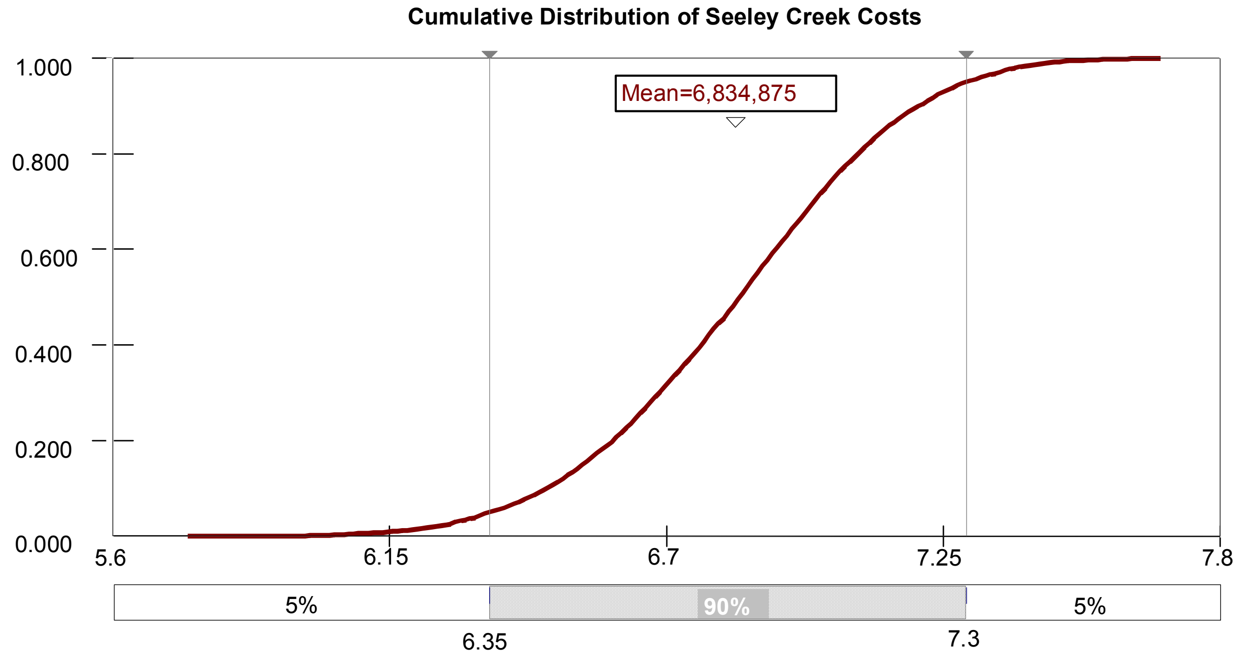



**Figure VI-1. Empirical Distribution**

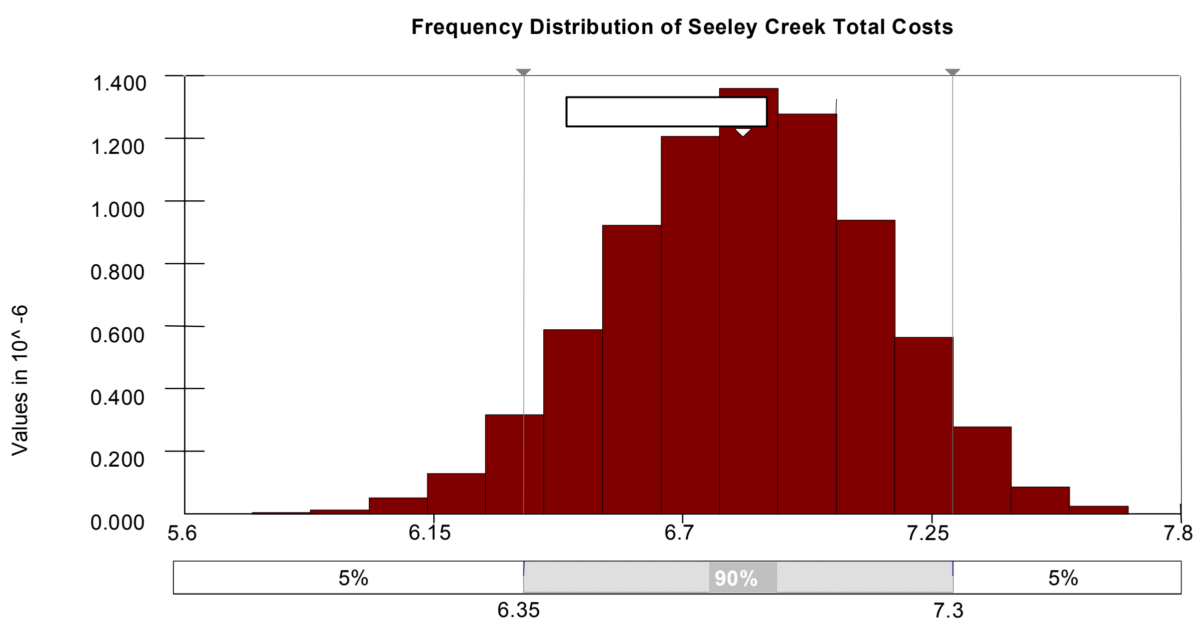



**Figure VI-2. Frequency Distribution**

In order to tighten the distribution of potential total costs we have to address the uncertainty in the inputs. If we can reduce the uncertainty in the inputs we may be able to reduce the uncertainty in the output, i.e. total costs. Fortunately, risk assessment provides a systematic way to investigate the most important input uncertainties through importance analysis, discussed earlier and shown in Figure III-4. By identifying the inputs that contribute most to the spread of costs in the figures below we have a good idea where to concentrate our efforts in order to reduce the uncertainty in our total cost estimate. More will be said on this topic below.

To consider a second major use of this information assume the uncertainty has been reduced as much as is practical for the study and Figures VI-1 and VI-2 represent the final assessment results. The cost engineers task now is to choose a cost estimate to use for this project. That could be any value at all. It might be difficult to justify a cost estimate of say $8 million on any basis other than extreme paranoia about cost overruns. It would be far more realistic to reexamine the uncertainty of one's quantity and unit cost inputs than to leap to such an extremely high estimate of costs.

So what cost should the District use when there are 10,000 candidate costs? The answer lies in the objectives and risk attitudes of the decision makers in their role as risk managers. If the greatest concern of risk managers is to avoid overruns then they will want a cost that is above the unbiased expected value. Bear in mind that cost risk assessment done well identifies the mean of $6.83 million as the single best guess of what the costs will actually be. If anything but that is chosen to represent costs then the risk assessors must have introduced some sort of bias. Perhaps they were conservative and overstated the high side of costs. Or maybe they were naively using outdated data that understates costs. Let's assume the mean is indeed an unbiased estimate of the most likely value.

If the District is risk averse and wants to avoid the problems associated with a potential cost overrun then they will select a cost from the right side, above the mean, of the cost distribution. How far to the right? Well, how important is it to avoid an overrun? Looking at Figure VI-1 we see a cost of $7.03 million provides us with a confidence level of about 80 percent that there will not be a cost overrun. Or in other words, if the District uses a cost estimate of $7.03 million there is a 20 percent chance the costs will eventually exceed that amount. Can you live with a 20 percent chance of an overrun? If so, you have your cost estimate. If not, then select a higher level of confidence. At $7.30 million there is only a five percent chance of an overrun.

On the other hand, risk managers may want to be optimistic or risk seeking, betting that costs will come in under the expected cost. If the project is being constructed under a program that had a $5 million limit for example we could say there is not chance costs will come in under the limit. Thus it is certain the sponsor will have to cover some costs in excess of the $5 million program limit.

The potential to choose any cost estimate from an interval and then to be able to quantitatively estimate the likelihood the actual costs will be above or below that amount is an extremely valuable piece of information. Given that an overrun is what can go wrong, these curves enable us to answer the question, how likely is it. Having previewed some of the uses of the distribution of total costs we now turn to some of the details of obtaining these results.

**Variation in Outputs**

The term *uncertainty* is used to describe a lack of sureness about something. Whenever there is doubtfulness about an event, a piece of information or the outcome of a process, a condition of uncertainty is said to exist. Uncertainty can be attributed to two sources: (1) the anticipated variability of processes ("inherent variability"), or (2) incomplete knowledge ("knowledge uncertainty") .

*Inherent Variability* refers to the ordinary variability in a system. In nature, it refers to the irreducible randomness of natural processes. In man-made systems, it refers to the vagaries of the system, this randomness is irreducible from the perspective of the risk analyst. In the ecosystem restoration context, uncertainties related to inherent variability include things such as stream flow, assumed to be a random process in time, soil properties, assumed to be random in space, or the success rate of propagules purchased to revegetate a project area. Inherent variability is sometimes called aleatory uncertainty.

*Knowledge Uncertainty* deals with a lack of understanding of events and processes, or with a lack of data from which to draw inferences; by assumption, such lack of knowledge is reducible with further information. Knowledge uncertainty is sometimes called epistemic uncertainty.

In the literature of risk analysis, there are a myriad of terms use to describe sources of error, uncertainty and/or risk. All of these definitions can be collapsed into the two above named sources. The taxonomy used to describe the source of uncertainty is not as important as understanding which source the uncertainty comes from.

The analyst, decision maker and stakeholder must understand the source of the uncertainty to properly interpret it. Consider the meaning of a 10% risk that an ecosystem restoration project would fail to satisfy a performance target. If the uncertainty is due to inherent variability, this may mean the ecosystem restoration project would fail to satisfy the performance target 1 year out of 10; however, if knowledge uncertainty is the issue, the a risk of 10 may suggest there is a 10% chance the project will always fail to meet the target (Stedinger 2000). It is critical that this distinction be made, communicated and understood.

Cost estimating is full of examples of uncertainty. The total cost of a project is an estimate, a forecast. Costs are unknown until construction is complete. Given the current state of accounting practices it often remains unknown even after construction is completed. There is variation in the estimate of total costs for two distinct reasons. One of them stems from inherent variability in the factors that cost money, the other is knowledge uncertainty about details of what it will actually take to construct something. For example, the amount of rock in a channel bottom is always going to be uncertain. The number of cubic yards of excavation or loose rock removed from a channel will also be uncertain. In fact most of the variation in construction cost estimates will be due to knowledge uncertainty. Inherent Variability comes into play in defect rates, weather, and other situations where pure chance is a factor. The significance of recognizing the reason for the uncertainty is quite simple. No matter how much money you throw at variation due to chance you cannot reduce it. You might understand it better and describe it more completely but you cannot make it go away. On the other hand, additional resources can often be effectively used to reduce the variation due to knowledge uncertainty. Additional study or investigation, for example more or better cross sections, more foundation exploration, better hydrology, contacting contractors for price information and other techniques can reduce the uncertainty in a cost estimate.

Knowledge uncertainty can be reduced, inherent variability cannot be reduced. Allocate resources to reduce knowledge uncertainty to reduce the uncertainty in your model outputs.

The above definitions are modified from the National Research Council Commission on Geosciences, Environment and Resources Report; *Risk Analysis and Uncertainty in Flood Damage Reduction Studies* (2000). Readers are referred to the NRC discussion of the concepts in the original report.

## The Model

The costs of the Seeley Creek stream restoration are estimated by traditional cost estimating techniques to be $8 million as shown in Table VI-1. Although the project cost exceeds the Section 206 Authority limit of $5 million, the sponsor could choose to pay the addi-

| TABLE VI-1<br><br>PROJECT COST ESTIMATE | | | | |
|---|---|---|---|---|
| **Seeley Creek Stream Restoration 09/01/00** | | | | |
| | **Contract** | **Contingency** | **Escalation** | **Total Cost** |
| 09 Channels and Canals | $5,716,515 | $857,477 | $322,126 | $6,896,118 |
| 30 Engineering and Design | $457,321 | $68,598 | $22,089 | $548,008 |
| 31 Construction Management | $485,904 | $48,590 | $22,449 | $556,943 |
| Total Seeley Creek | $6,659,740 | $974,665 | $366,664 | $8,001,069 |

tional cost. This makes estimating the likelihood that this limit will be exceeded to be even more important in the decision process.

Contingencies for this project are based on fifteen percent of project contract costs. They represent about 12 percent of the total cost. Escalation of prices to the midpoint of construction represents about a 4.9 percent increase in contract plus contingency costs. Total costs of account 09 are $6.90 million. Engineering and design is eight percent of total costs for account 09. Construction management is 8.5 percent of total costs for account 09. E&D and CM account for another $1.1 million in project costs.

The cost risk assessment approaches the notion of contingencies in a different way. Using the detailed cost estimate used to prepare the summary in Table VI-1 cost estimators and design engineers are able to address the uncertainty in individual elements of the cost estimate. By describing these uncertain elements with a probability distribution the expert is able to say which values could occur and which of them are most likely.

A large number (10,000) of possible cost scenarios are investigated using the Monte Carlo process. E&D and CM are estimated as fixed percentages of the account 09 contract cost. Escalation is based on and added to the sum of contract, E&D, and CM costs to obtain the total cost estimate. The results obtained through this process are those described throughout this report.

The Seeley Creek cost risk assessment model is shown in Table VI-2. The values shown represent the expected values of all inputs, one of many possible scenarios for the actual cost. Each iteration of the model selects a new quantity and unit cost value for each of the cells shown according to the rules provided by the District's cost and design experts.

Total channel and cannel costs for the project shown total $4,697,192. Engineering and design, construction management, contractor fees and escalation bring the project cost estimate up to $6.83 million.

| TABLE VI-2<br><br>COST RISK ASSESSMENT | | | | | |
|---|---|---|---|---|---|
| | **Quantity** | **Units** | **Unit Cost** | **Units** | **Total Cost** |
| Mob, demob and preparatory work | 1 | | $16, 747.00 | | $16, 747.00 |
| Grade control weirs and tie-back dikes | | | | | |
| **Excavation** | | | | | |
| Excavation and Load | 33000 | CY | $1.34 | Per CY | $44,115.50 |
| Excavating loose rock | 33000 | CY | $5.50 | Per CY | $181,500.00 |
| Total Excavation | | | | | |
| Total backfill around revetments | 17000 | CY | $1.39 | Per CY | $23,630.00 |
| Stockpile remaining excavated material | 16000 | CY | $5.50 | Per CY | $88,000.00 |
| **Armor stone** | | | | | |
| Hauling | 16167 | Ton | $28.12 | Per Ton | $454,545.98 |
| Placement | 16167 | Ton | $4.31 | Per Ton | $64,980.00 |
| **36" rip rap** | | | | | |
| Hauling | 4500 | Ton | $27.00 | Per Ton | $121,512.00 |
| Placement | 4500 | Ton | $14.44 | Per Ton | $64, 980.00 |
| **Core stone** | | | | | |
| Hauling | 4500 | Ton | $19.27 | Per Ton | $86,736.00 |
| Placement | 4500 | Ton | $3.57 | Per Ton | $16,065.00 |
| **Filter stone** | | | | | |
| Hauling | 7500 | Ton | $20.94 | Per Ton | $157,081.25 |
| Backfill spread | 15000 | SY | $0.23 | Per SY | $3,450.00 |
| Compaction | 5000 | CY | $0.19 | Per SY | $950.00 |
| **Stone toe revetments** | | | | | |
| Excavate and load | 60000 | CY | $1.34 | Per CY | $80,210.00 |
| Excavate loose rock | 60000 | CY | $5.50 | Per CY | $330,000.00 |
| Backfill around revetments | 30000 | CY | $1.39 | Per CY | $41,700.00 |
| Stockpile remaining excavated material | 30000 | CY | $5.50 | Per CY | $165,000.00 |
| **52" rip rap** | | | | | |
| Hauling | 39000 | Ton | $28.12 | Per Ton | $1,096,511.00 |
| Placement | 26000 | CY | $14.44 | Per CY | $375,440.00 |
| **Filter stone (correlate all the stone)** | | | | | |
| Hauling | 16167 | Ton | $24.58 | Per Ton | $397,339.05 |
| Backfill spread | 32329 | SY | $0.23 | Per SY | $7,435.67 |
| Compaction | 10776 | CY | $0.19 | Per CY | $2,047.00 |
| **Earthwork to form channel** | | | | | |
| Excavation | 53667 | CY | $1.34 | Per CY | $71,743.83 |
| Stockpile excavation material | 53667 | CY | $5.50 | Per CY | $295,168.50 |
| **Boulders** | | | | | |
| Hauling | 100 | Each | $29.23 | Each | $2,922.87 |
| Placement | 100 | Each | $41.55 | Each | $4,155.00 |
| Gas line relocation | 160 | LF | 414.92 | Per LF | $2,387.20 |

| TABLE VI-2 (Continued)<br><br>COST RISK ASSESSMENT | | | | | |
|---|---|---|---|---|---|
| | **Quantity** | **Units** | **Unit Cost** | **Units** | **Total Cost** |
| **Temporary work for handling water during construction** | | | | | |
| Excavate trench | 3000 | CY | $2.14 | Per CY | $6,420.00 |
| Sandbags | 3000 | Each | $2.12 | Each | $6,360.00 |
| Piping | 500 | LF | $10.03 | Per LF | $5,015.00 |
| Pump | 1 | Each | $13,981.00 | Each | $13,981.00 |
| Rip rap | 1 | CY | $216.51 | Per CY | $216.51 |
| Geotextile fabric | 4 | SY | $47.77 | Per SY | $191.08 |
| Temporary erosion and sediment control | 1 | Each | $101,892.37 | Each | $101,892.37 |
| Contingent excavation to remove materials | 2500 | CY | $1.34 | Per CY | $3,342.08 |
| Stockpile temporary excavation materials | 2500 | CY | $5.50 | Per CY | $13,750.00 |
| **Plantings** | | | | | |
| Live stakes | 39333 | SY | $2.52 | Per SY | $99,119.16 |
| Live fascine with erosion control | 694 | SY | $34.43 | Per SY | $23,894.42 |
| VRSS | 1125 | SF | $18.00 | Per SF | $20,250.00 |
| Rock | 13425 | LF | $10.00 | Per LF | $134,250.00 |
| Live fascine | 6800 | LF | $1.62 | Per LF | $11,016.00 |
| Joint plant | 106 | SY | $16.75 | Per SY | $1,775.50 |
| Vegetative spurs at grade | 75 | LF | $30.00 | Per LF | $2,250.00 |
| Vegetative spurs above grade | 12 | Each | $3,000.00 | Each | $36,000.00 |
| Rock Toe | 628 | CY | $26.14 | Per CY | $16,415.92 |
| **TOTAL CHANNELS AND CANNELS** | | | | | **$4,697,192.11** |
| Engineering and Design | | | | 8.0% | $375,775.37 |
| Construction Management | | | | 8.5% | $399,261.33 |
| SUBTOTAL | | | | | $5,472,288.81 |
| Prime Contractor's OH, Office, Profit, Bond Escalation | | | | 20% | $1,094,455.76 |
| Escalation | | | | 4.9% | $268,139.21 |
| **TOTAL COST** | | | | | **$6,834,813.78** |

## Input Data

Quantities were estimated using design estimates appropriate to the stage of this analysis. Unit cost data were based on the M-CACES database and the estimator's experience. The uncertainty inherent in both the quantities and unit costs was acknowledged and recognized by District personnel. For reasons discussed at greater length in the lessons learned discussion and the Appendix a relatively straightforward and simple approach was used to describe and quantify the uncertainty.

All inputs were described using either triangular, uniform, or beta subjective distributions. All of these distributions can be used in the absence of extensive databases. They are non-parametric distributions based on expert opinion. As such they represent one of the simplest ways to apply the Monte Carlo process to a cost estimate. There are more sophisticated ways to describe the uncertainty attending a cost estimate as discussed in the Appendix. Thus, the method used here represents a simple application of cost risk assessment.

Probability distributions were used to describe the uncertainty in 28 different quantities shown in Table VI-2. Twenty-seven of these distributions were triangular distributions. Most of them were defined by adding and subtracting a fixed percentage to the design engineers' best estimate of a quantity. Plus or minus five or ten percent were the two most common estimates of the parameters of the triangular distribution.

**Quantifying Uncertainty**

If cost risk assessment is ever to be used regularly by the Corps, quantifying the uncertainty in cost estimate inputs is going to be one of the most important activities. Training in how best to do this will be essential for all Corps personnel. One purpose of this research was to demonstrate the feasibility of applying these techniques to ecosystem restoration projects. The techniques actually used represent a compromise between state-of-the-art uncertainty estimation and a pragmatic approach to elicit the cooperation of otherwise very busy professionals.

In a data-poor environment, such as this one was, the uncertainty in cost inputs is often best described by the design engineers' and cost estimators' experience and best judgment. In this case, the quantity or unit cost value used in the official estimate for Seeley Creek was the starting point. The most likely value used in the risk assessment was to reflect the estimator's unbiased (i.e., the actual value was as likely to be more than as less than this amount) best judgment. Whether this goal was achieved or not is a matter of some speculation. It is difficult for professionals conditioned by many years of doing things in one way to successfully shift their approach after a sixty-minute discussion.

Once the most likely value for a quantity or unit cost was identified the minimum and maximum possible values were identified. The estimators found it easiest to estimate these values using a percentage adjustment to their best estimate, for example the actual unit cost could be 15 percent more or 15 percent less. Reliance on symmetrical adjustments could reflect some lack of experience with the process of quantifying uncertainty.

For another example, District personnel estimated that 33,000 CY of material would have to be excavated and loaded for weir and tie-back dike construction. They judged the actual quantity could be as much as 10 percent less or 10 percent more than this. In other cases, where the data were not as good, the range in percentages may have been greater.

A triangular distribution can be described by estimating the lowest, most likely and highest possible values for a variable. Consider the Figure VI-3 below, which shows the excavation and load quantities for the grade control weirs and tie-back dikes. The quantity cannot be less than 29,700 cubic yards or more than 36,300.

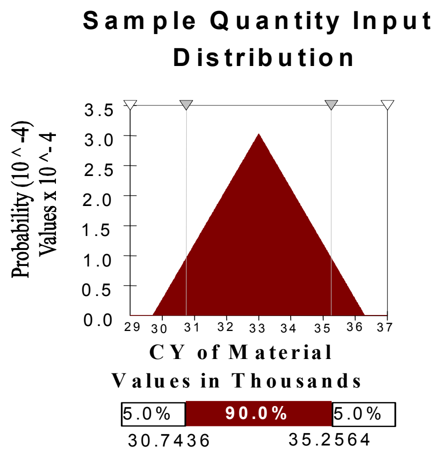



**Figure VI-3. Excavation Quantity Input**

Two of the triangular distribution parameters are the minimum and the maximum values. Together these two parameters place us on the relevant portion of the number line. The next logical question to ask might be, can we say anything else about the excavation quantity? The quantity is somewhere between 29,700 CY and 36,300 CY, but do we know anything else? Are some of the values more likely than others? The distribution shape suggests that is indeed true. And the third parameter for a triangular distribution is the most likely value, or the mode on that line segment. All we need to define a triangular distribution is a minimum, maximum and most likely value.

It is important to remember that the most likely value is the mode, not the mean. The mean of a triangular distribution is obtained by taking the average of the three parameters. In this example the most likely value is 33,000 CY. The mean of this distribution, a number not needed to define the distribution, is the same as the mode because the distribution is perfectly symmetrical. That is, values below the most likely value are as likely as values above it. And the range of values (3,300 CY) below the most likely value is the same as the range above the most likely value.

It may help to think of the distribution in Figure VI-3 as a rule we specify for instructing the computer on which values to choose for this excavation quantity and how often. The choice of a distribution is one of the more difficult things for new risk assessors to understand and master. There are many different “rules” we can specify (see Appendix). Good risk assessment should have good reasons for using the distributions they use. In this example, we used triangular and uniform distributions simply because they were the simplest distributions for District personnel without prior experience with probabilistic scenario analysis to understand and work with.

The second distribution type used for the quantity estimates was the uniform distribution. Figure VI-4 shows the distribution for the material that might be excavated as a result of hydrologic events during construction. We know it will be no less than zero and no more than 5,000 CY.

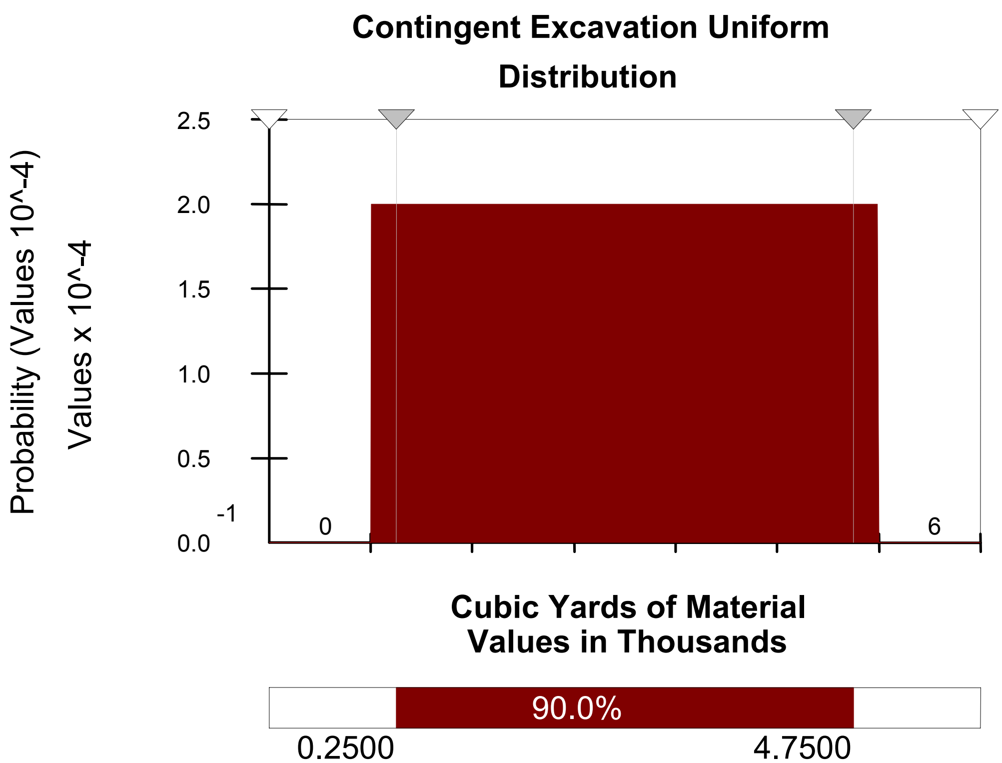



**Figure VI-4. Uniform Distribution**

Other than that, there is nothing else we can say about the most likely value that will occur. The uniform distribution represents a kind of maximum uncertainty situation. After we identify a minimum and a maximum value there is nothing else we can say. This distribution was used once for quantities.

A few quantities are certain. There will be one mobilization and demobilization. One pump will be used. A few other planning quantities were treated as deterministic values.

Uncertainties in unit costs are described with triangular distributions in 32 of 33 cases. The other used a beta subjective distribution. Many of the triangular distributions are based on calculating the interval created by taking plus or minus 15 percent of the cost estimator's best estimate of unit costs. Reliance on this particular percentage reflected the estimator's comfort level with the quality of data he had. Unfamiliarity with the technique may have contributed to some repeated reliance on this percentage once it was used. The distribution of per CY prices for excavating loose rock from the creek channel is shown in Figure VI-5 below. The most likely cost was estimated to be $5.85 but there is a chance the rock could be removed for as little as $2 per CY. This price is not as likely as the other values however and that is reflected by the "rule" embodied in the shape of the distribution.

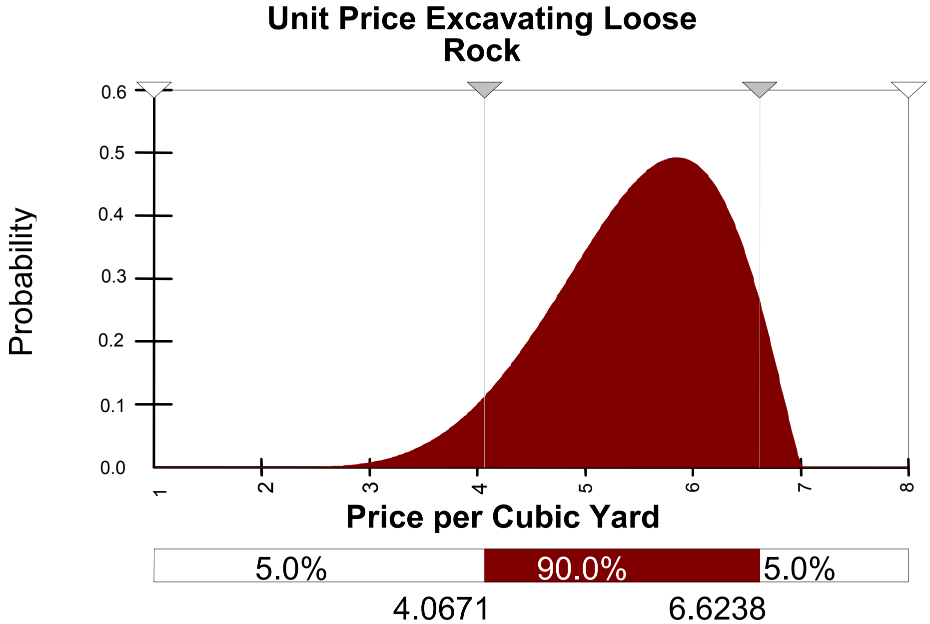



**Figure VI-5. Beta Subjective Distribution**

## Triangular Distribution Parameters

A discussion of the manner in which distributions can be identified is well beyond the scope of this project. However, a few words on how the triangular distributions were identified are in order. The text suggests a minimum and maximum are identified first, then the most likely value is identified. That is a common way of specifying a triangular distribution. In this application a different approach was taken.

The District had prepared a traditional deterministic cost estimate prior to the initiation of the cost risk assessment. In that case the easiest value to begin with was the most likely value, the value that had already been identified. The working assumption is that if a more likely value existed the estimator would have used it instead of what was used. Hence, the point estimate in the District cost estimate was assumed to be the mode. The extreme values were then estimated based on an adjustment to the most likely value. The estimators working on this assessment generally did this by adjusting their best estimate up or down a percentage. In most cases it was a symmetric adjustment, ± some percentage. Percentages in multiples of 5 were most common. Occasionally, asymmetric intervals were specified. In some cases the estimator found it easier to increase or decrease the best estimate by a fixed dollar amount. This technique is an easy and often inaccurate approach to describing uncertainty. Given the realities of involving inexperienced personnel in a relatively sophisticated risk assessment, however, this approach was accepted without much scrutiny.

## Interdependence of Variables

Acknowledging, recognizing and describing the uncertainty is one of the key steps in a cost risk assessment. That is largely accomplished by identifying distributions that describe the uncertainty and variability in key input variables. But it is not the only important consideration.

Equally important is the need to consider how different model inputs may be related to one another. When inputs are independent of one another there is nothing more we need to do to set up our model after we have specified the distribution to use for the input. When they are not independent, more work needs to be done. A number of inputs in the Seeley Creek analysis were dependent upon one another.

Three dependency relationships were identified among the quantities and two were identified among the unit costs. Weir quantities, revetment quantities and planting quantities all tended to move in the same directions within their groupings. For example, weir quantities included armor stone, 36-inch rip rap, core stone and filter stone. When the Monte Carlo process generated a quantity above the most likely value for armor stone the model should show above most likely values for the other three stone quantities as well. This dependency was built into the model. Likewise direction relationships, i.e. positive correlations were used for revetment and plant quantity groups. On the cost side all placement costs were assumed to move together as were all hauling costs. This interdependence of variables is accomplished via a rank correlation coefficient specification that is a feature of the software used to complete the Monte Carlo analysis.

Relationships between quantities and unit costs were explored. The cost estimator felt the potential variation in quantities was not sufficient to affect unit prices directly. Hence, no such interdependencies were used for this model.

## Simulation

The original model was built using the Corps' TRACES software for cost estimating. The detailed report from TRACES was used to build a replica of the model in Excel Office 2000. The Monte Carlo process used @RISK version 4.02. Ten thousand iterations of the model were run in about 30 minutes time. The results have been used throughout the report and are presented in the section that follows.

## Assessment Results

The results of the cost risk assessment have been presented in a series of tables and figures throughout this report. This section summarizes many of those results and offers a few suggestions for presenting the results of a cost risk assessment.

The risk assessment should answer the questions that have been presented to the assessors by the managers who are going to be responsible for making a decision. The identity of these people will vary with the District and the context of the cost estimate. No questions were posed of the risk assessment for the case study. That was primarily due to the fact that there is little to no practical experience with risk cost assessment within the Corps' Civil Works Program culture. Consequently, District personnel have no experience with what sorts of things they make ask or expect of a risk assessment. Earlier it was suggested these questions might include the following:

**What is our best unbiased estimate of project costs?**

The simulation of costs has produced 10,000 estimates for the project cost. To answer this question we need to determine which of those estimates is the best unbiased one. Best and unbiased are not used in their statistical meaning in this context. By best we mean the one that is better than all others for the purposes of the Corps. By unbiased we mean an estimate that is not strategically optimistic or conservative. And so, if the model values for individual model inputs are the best and unbiased the best unbiased estimate of costs is the expected value of our distribution or the mean. In this case the mean is $6,834,875 or $6.83 million. Because the results of a risk assessment are an estimate it is never appropriate to treat all the digits of the mean or any other value as significant.

This is the value that is believed to be the most likely cost of the project. There is a 51 percent chance this cost will be exceeded and a 49 percent chance costs will be this much or less. That places it pretty close to the median cost. All other things equal, this would be the best unbiased estimate of the cost of the Seeley Creek project. But all other things are not equal.

**What is the maximum likely overrun of our best estimate?**

Notice the question. It does not ask the maximum possible overrun. With scenarios of unanticipated hazardous toxic and radioactive wastes, earthquakes, strikes, bad weather, economic upturns and downturns and so on it would not be difficult to imagine virtually any cost for this or any other project. What we seek is the maximum likely overrun. That depends squarely on the District's cost engineering team and their capabilities. In the current context it means based on the assumptions built into the model how high could costs go. That is the difference between the best unbiased estimate and the maximum cost estimated in a simulation of sufficient iterations. The maximum cost in this analysis was $7,682,302 or $7.68 million. The maximum overrun is this less the best estimate or $0.85 million. Because this cost occurred once in 4,863 iterations in excess of the mean there is a $2 \times 10^{-4}$ chance that costs will overrun the best estimate by this much or more.

Risk assessors have done their job to identify this number. Risk managers must now decide if that is an acceptable potential overrun. If the possibility of a $7.68 million project overrunning costs by $0.85 million is unacceptable then managers have several options. One is to try to reduce the uncertainty in the cost estimate. This uncertainty is best demonstrated by the information contained in Figures VI-6 and VI-7. If the spread of output values in Figure VI-6 is to be reduced the uncertainty in the input values will have to be reduced. That leads to the next

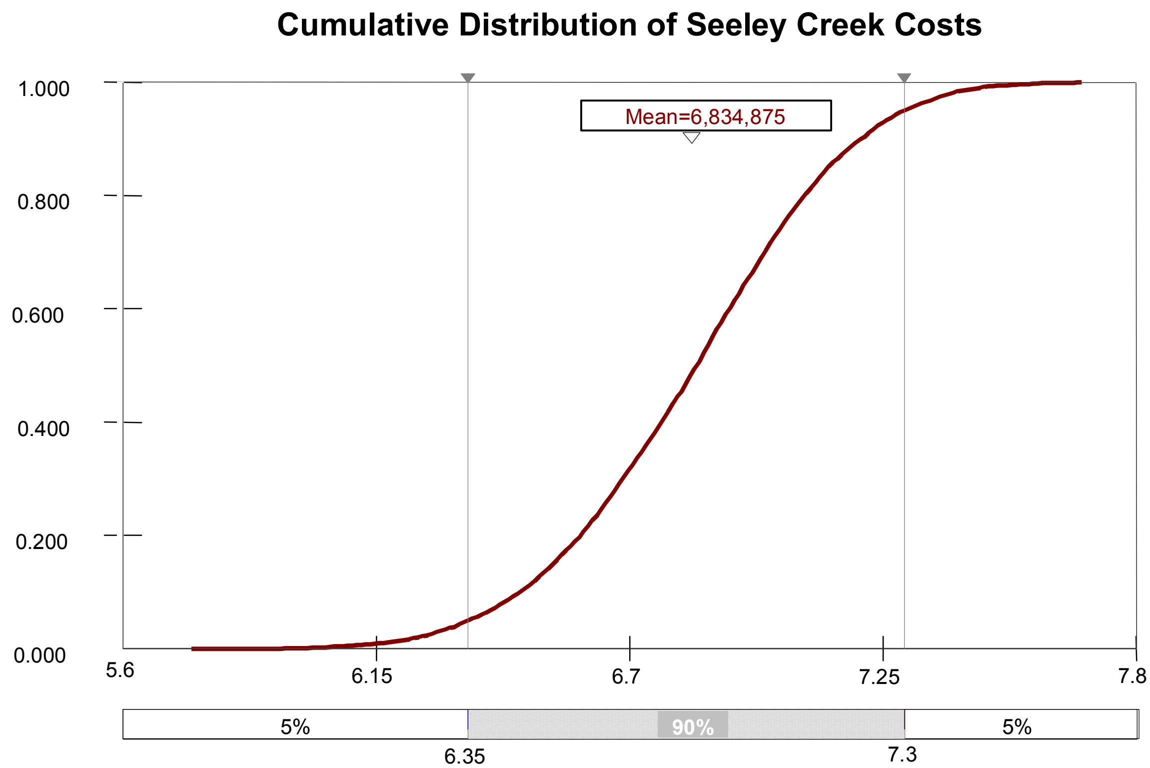



**Figure VI-6: Empirical Distribution of Seeley Creek Costs**

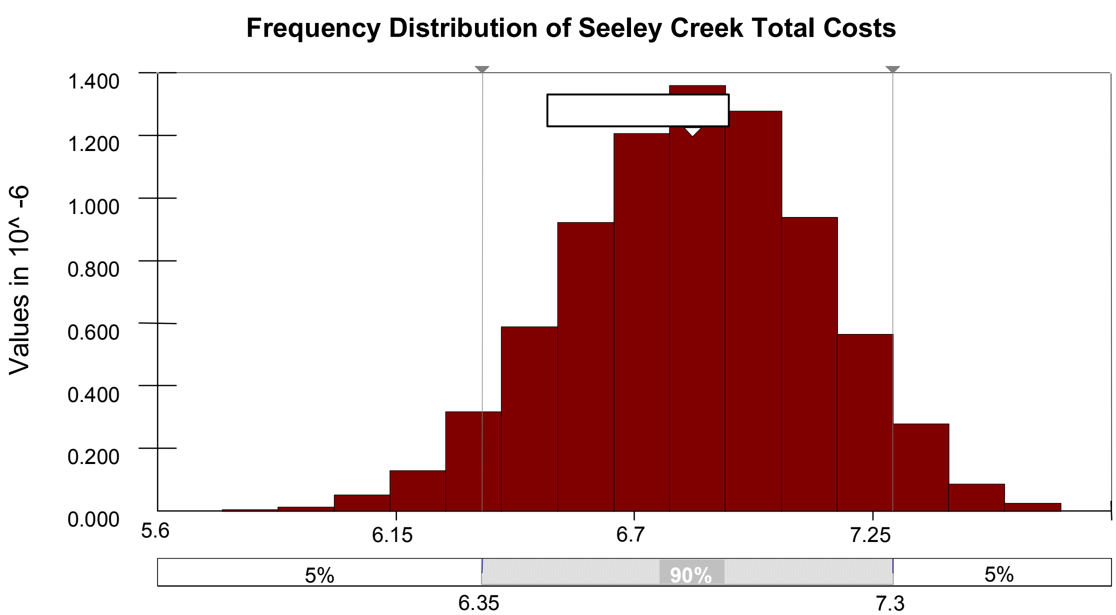



**Figure VI-7. Frequency Distribution of Seeley Creek Costs**

two questions. Another option for minimizing cost overruns is to carefully manage the project so as to minimize costs. This too is facilitated by the answer to the next two questions.

**How Many Iterations?**

How many iterations are enough? The answer depends on what you are interested in knowing. The expected value or best estimate of costs can often be known with a reasonable degree of accuracy after a few hundred iterations. If we are interested in whether we have a symmetric estimate of costs or a rightward skew we need more iterations in order to get a reasonable idea of the shape of the distribution. This is reasonably well ascertained after about a thousand iterations.

With a few thousand iterations we begin to get some idea what the tails of our distribution might look like. The more iterations we do the better defined the tails become. Five to ten thousand iterations will give a reasonable idea of how likely high-end and low-end costs might be. That leaves only one's concern with extreme events to consider. When there is legitimate concern about circumstances that could lead to an unusually high cost it may be wise to do tens of thousands of cost estimates. It would seem rare to ever have to do a simulation of more than 100,000 iterations, but that remains a matter to be determined by the cost estimator.

### What are the most uncertain unit costs?

Identification of the most uncertain costs must be done by the cost estimator. This task is best accomplished before uncertainties are quantified. In most cases quantification of uncertainties for about 20 percent of the cost estimate inputs will be sufficient to capture the bulk of the uncertainty about any given cost estimate. That was not done in this project because of its prototype nature and District personnel's lack of familiarity with the technique. It simply was not realistic to expect very busy volunteers to master the concepts and methods of cost risk assessment. Any estimate is bound to present some unit prices that are harder to estimate than others because there is more uncertainty. This is particularly true with ecosystem restoration costs where components of plans and their work units are less familiar to the estimator. It is important to have the expert's opinion on those prices he considers most uncertain. Here we use uncertainty to include variability as well.

### What are the most uncertain quantities?

In a similar fashion the design engineer should identify those quantities she considers most uncertain. That was not done in this project due to the time constraints of the District.

### What unit costs contribute the most to the variation in total costs?

An importance analysis, often called a sensitivity analysis by software producers, is useful for identifying the most important variables in a probabilistic scenario analysis like this. One importance analysis is reproduced below. Although the cell addresses mean nothing to the reader they can be readily referenced to the model to determine that the three most important unit prices are: the costs of excavating loose rock from the channel, the cost of hauling armor stone, and the price of placing boulders on the job. Figure VI-6 shows the uncertainty in the estimate

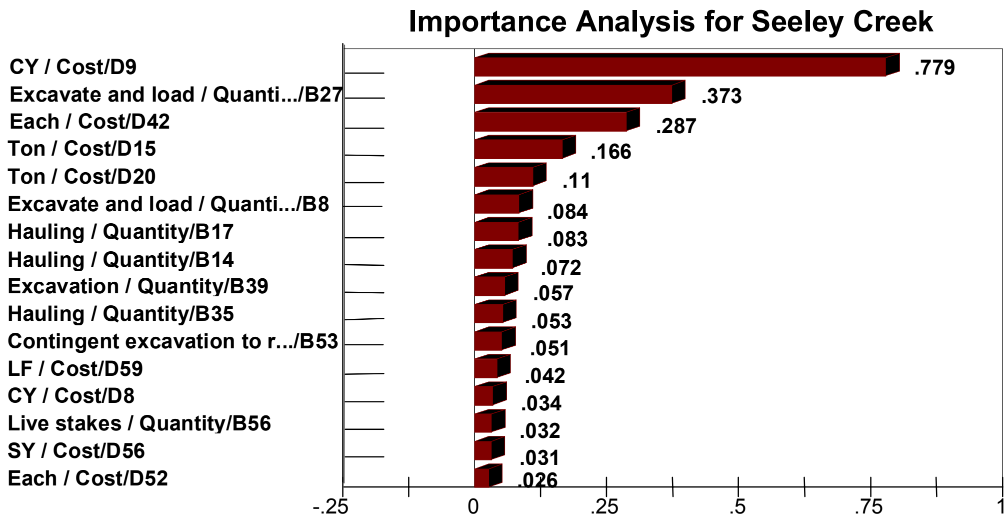



**Figure VI-8. Importance Analysis for Seeley Creek Costs**

of loose rock excavation costs. It suggest that if there is a way to better determine the likelihood of getting rock removed by the local government at cost close to $2 per CY the uncertainty in the overall cost estimate might be reduced.

**What quantities contribute most to the variation in total costs?**

Using the same importance analysis the three most important quantities to investigate to reduce total cost uncertainty for Seeley Creek are: the quantity of stone to revetment excavation, the amount of 36 inch rip rap to haul, and the amount of loose rock to be excavated for the weirs and dikes.

**If we were to do further analysis of the cost estimate on what cost components should we allocate our resources?**

The simple answer to this question would be to examine the intersection of the most uncertain costs with those that contribute most to the total cost variation, likewise for the quantities. In this instance where we lack the analysts' opinions we could look simply to the first three or four items identified in the importance analysis. If more work is to be done to further refine the cost estimate those are the things that can most productively be addressed.

**What cost estimate is consistent with a 10 percent or less chance of a cost overrun?**

Ultimately, the District will have to select an estimate to use for the project. It is unlikely that the best unbiased estimate will be used because there is such a high chance that cost will be overrun. It is a simple fact of life that for the Corps of Engineers cost overruns are more problematic than cost underruns. Normally the District and its partner will be inclined to want to provide protection against a cost overrun. It is worth repeating that cost risk assessment provides an initial line of defense against cost overruns by providing information that enables the Corps to investigate ways to better refine cost estimates in a very focused fashion.

When the investigation of costs has gone as far as desired or possible the kinds of information in Figures VI-5 and VI-6 above aid the choice of a cost contingency in a brand new fashion for the Corps. Table III-1 is reproduced below for your convenience. The best estimate of $6.83 million does not provide sufficient protection against a cost overrun. In this example we have arbitrarily chosen a 10 percent chance of an overrun as a tolerable risk of an overrun. Based on the table below we see that 90 percent of all the cost estimates were $7.20 million or less. In the total cost dataset ten percent of all the values were greater than that value. Hence, we assume there is a ten percent chance that costs will be more than $7.20 million and we choose that as the cost estimate that limits us to a ten percent chance of a cost overrun.

**TABLE VI-3**

**SELECTED PROJECT COSTS IN MILLIONS**

| Item | Cost | Item | Cost | Item | Cost |
|---|---|---|---|---|---|
| Minimum Observed Cost | $5.75 | 30th Percentile | $6.68 | 70th Percentile | $6.99 |
| Maximum Observed Cost | $7.69 | 35th Percentile | $6.73 | 75th Percentile | $7.03 |
| Mean Observed Cost | $6.83 | 40th Percentile | $6.77 | 80th Percentile | $7.08 |
| 5th Percentile | $6.35 | 45th Percentile | $6.81 | 85th Percentile | $7.13 |
| 10th Percentile | $6.46 | 50th Percentile | $6.84 | 90th Percentile | $7.20 |
| 15th Percentile | $6.53 | 55th Percentile | $6.88 | 95th Percentile | $7.30 |
| 20th Percentile | $6.59 | 60th Percentile | $6.92 | | |
| 25th Percentile | $6.64 | 65th Percentile | $6.95 | | |

This cost exceeds the best cost estimate by $0.37 million dollars ($7.20 million - $6.83 million). Thus, starting from our best estimate of costs we add a contingency of $0.37 million to it to obtain the cost estimate that we believe will provide the degree of protection we want from cost overruns. Consequently, the official cost estimate for the project would become $7.20 million. The $0.37 million contingency represents a 5.4 percent increase over the best cost estimate of $6.83 million. This is considerably less than is typically used in a data poor environment. That is due largely to the fact that we have been able to address the uncertainty in the cost estimate on an item-by-item basis and we have been able to choose the degree of overrun protection we want.

Why not choose a higher degree of overrun protection? That is certainly an option that is open to the cost managers. One obvious answer is that the extra cost may make the project a harder sell for the partner or the Corps. It may also be undesirable to focus attention on a high cost that has a relatively small chance of occurring. Remember, when the cost risk assessment is done well there is a 90 percent chance the actual cost will be $7.20 million or less. Those are pretty good odds! At a cost of $7.20 million the maximum exposure to a overrun has been reduced from $0.8 million to $0.5 million.

## Lessons Learned

Few research projects are conducted under ideal circumstances. This one does not reverse that tendency. As a result the opportunity to demonstrate the utility of this method was limited. This merits discussion.

Despite the case study District's generous participation in and support for this project they were unable to find the time to spend on this project that would have yielded the greatest utility for the Seeley Creek project and for this research. Because of the participants' busy schedules it was not possible to get people together to work on this project as often as might have been most fruitful. This included the time required to learn about cost risk assessment, the time required to quantify uncertainty in the most realistic fashion, and the time to review and consider the results of the preliminary analysis. There is currently no culture in the Corps of Engineers Civil Works Program that recognizes or values cost risk assessment. As a result, some of the lessons learned were somewhat different from what was anticipated but they are nonetheless useful.

The District's official estimate of costs is $8 million. The results of the cost risk assessment show the most likely estimate to be $6.83 million with no chance costs will be as high as $8 million. This is certainly an interesting result. It suggests that the cost risk assessment may be overly pessimistic and could warrant reconsideration.

That the cost estimate may be unreasonably high is a piece of information, which, if true, could have important implications for the project's eventual construction. A non-Federal partner expecting a lower cost might summarily dismiss the project based on first costs alone. It would surely seem to be in the District's interests to investigate the possibility that the cost risk assessment has revealed useful things about the project prior to their coordination of costs with the non-Federal partner. The timing was not right for that kind of investigation.

We are left, then, to speculate that cost risk assessment may be very useful in terms of what it might suggest to us about the conservative bias in cost estimates prepared in a data poor environment. The $8 million estimate includes a 15 percent contingency to the overall cost estimate. The risk assessment handled the contingency on each quantity and cost separately. It also took into account the dependence and independence of each of these variables.

The 15 percent contingency in a traditional estimate effectively recognizes that the cost estimate without a contingency could be 15 percent more than was estimated. Cost risk assessment is based on individual descriptions of uncertainty and an acknowledgment that some values are as likely to be less than estimated as more than estimated. In this case, cost risk assessment suggests that a cost as high as $8 million is virtually impossible. Under traditional estimation methods we might consider it high and not likely to be overrun but we might not know that it is virtually impossible to be reached. Nor would we know how much the estimate could be reduced and still have an acceptable chance of being exceeded.

One of the first things a review of this cost risk assessment would do would be to have the design engineers and cost estimator look at each of the uncertainty distributions to examine its adequacy in light of the results obtained. Presumably greater variation in assumptions about "rules" (i.e., choice of distribution or distribution parameters) for selecting values of input variables would result. This might include the use of more asymmetric distributions and or the use of different kinds of distributions. It takes time for personnel with no or limited experience with the Monte Carlo process to become comfortable with what a distribution is saying about the uncertainty in an estimate. The topic is not terribly intuitive and it is often new material. These and other factors can combine to make the informed discussion of probability distributions a time consuming process.

Moving back one more step, experience suggests that analysts are not as likely to address the more arcane quantitative issues of cost risk assessment unless they are motivated to want to do cost risk assessment. Lacking a culture for cost risk assessment the prime motivation for the assessment, i.e., the specific questions the assessment was to answer were missing. As a result, this research was characterized by a dimension of academic curiosity that kept it at arm's length from being seriously considered as a decision making tool or methodology.

And so, if we are to offer a few lessons learned from this experience they would include the following:

(1) Cost risk assessment may well provide District personnel with insights that could materially affect the success of their program, especially when operating in a data poor environment.

(2) The greatest value of a cost risk assessment will be derived when it answers questions that Corps managers and non-Federal partners have posed of it to aid their decision processes.

(3) There is little that is intuitive about using probability distributions to describe uncertainty. If this is to be done as effectively and efficiently as possible Corps analysts are going to need ample support to acquire these skills or the Corps will have to rely on more costly outside experts.

(4) Few analysts will be willing to devote the time and effort to cost risk analysis unless they are properly motivated. Motivation can be top down or bottom up. Top down motivation could come in a requirement to do cost risk assessment and it presupposes recognition of the value of doing cost risk assessment by those higher up in the agency. Bottom up motivation could result from analysts' recognition of the value of cost risk assessment to their own jobs and programs.

(5) Motivation to do cost risk assessment must be accompanied by an agency commitment to cost risk assessment. This must include the development of educational and training materials and opportunities, and an on-going commitment to their delivery to Corps personnel. It should also include the adaptation of current Corps cost estimating tools such as TRACES, PACES, M-CACES and CEDEP to include the Monte Carlo process, the ability to model interdependent relationships and the preparation of meaningful reports.

# VII. SUMMARY AND CONCLUSIONS

The case study presented in this report establishes that it is possible to do simple cost risk assessment for ecosystem restoration projects. The actual time spent on the District's involvement in the cost risk assessment was less than one half a day, not counting coordination meetings. That will rarely represent a hardship to even the smallest budgets. Although this case study could not capitalize on the strengths of risk assessment for reasons beyond the case study itself there is ample reason to suspect the risk assessment results could provide valuable information to the Corps and its partners when planning and designing these unique projects. In short, we believe this research shows cost risk assessment can be done and it offers great promise in the form of new dimensions of information about project costs including maximum exposure to overruns, important input variables, estimated levels of overrun protection and more. The process itself is extremely valuable to those involved in the estimate for what it teaches about what we know and what we do not know in our cost estimates, especially in a data poor environment. In summary, the process works and can offer much.

The experience here and with other Corps Districts leads us to conclude that cost risk assessment has the potential to significantly improve the quality of cost estimation information in the planning stages of a project. Cost risk assessment would, however, represent a significant change in the way cost estimates are prepared by the Corps. The knowledge and skills required to do cost risk assessment are new. A substantial commitment to education, training and tool development must accompany any effort to move the Corps in the direction of using cost risk assessment. Design engineers and cost estimators are going to need motivation, training and support for cost risk assessment to become a reality in the Corps planning process.

Industry already makes effective use of these techniques. There is an extensive construction cost literature on this topic and cost risk assessment is being used more and more. It seems evident, however, that if the Corps is going to develop the in-house capability of doing cost risk assessment, a strategy that strikes us as important if not yet urgent, it must make a substantial commitment to the methodology. Initially that would seem to suggest a top down motivation for doing this sort of analysis, accompanied by substantial support to field elements who must prepare the cost estimates.

The techniques are well known, the Corps' tools are readily adaptable to these techniques, the Corps has a professional staff of cost engineers who are more than capable of learning what can be effectively taught. All that is missing is the organizational encouragement and support that would enable these professionals to begin to apply and use these techniques, not only for ecosystem restoration where they would be especially useful, but throughout the Corps' Civil Works Program.

# REFERENCES

Curran, Michael W. (1998). *AACE International's Professional Practice Guide to Risk.* Morgantown: AACE International.

Curran, M. (1989). "Range estimating." *Cost Engineering*, 31(3):18-26.

Curran and Rowland (1991). "Range estimating in value engineering." *Transactions of the American Association of Cost Engineers,* pp. G.3.1-G.3.5.

Garvey, Paul R. (2000). *Probability Methods for Cost Uncertainty Analysis A Systems Engineering Perspective*. New York: Marcel Dekker, Inc.

Noble, Benjamin D., Ronald M. Thom, Thomas H. Green, and Amy B. Borde (2000). *Analyzing Uncertainty in the Costs of Ecosystem Restoration*. U.S. Army Corps of Engineers, Institute for Water Resources, to be published.

Yoe, Charles (2000). *Risk Analysis Framework for Cost Estimation*. U.S. Army Corps of Engineers, Institute for Water Resources, to be published.

Yoe, Charles and Leigh Skaggs (1997). *Risk and Uncertainty Analysis Procedures for the Evaluation of Environmental Outputs.* U.S. Army Corps of Engineers, Institute for water Resources, IWR Report 97-R-7, August 1997.

# APPENDIX A

# DESCRIBING UNCERTAINTY

# APPENDIX A: DESCRIBING UNCERTAINTY

It is not always easy to find a real case study to apply research techniques to in real time. The Baltimore District and its personnel were most gracious in offering their project for this research. Because it was an ecosystem restoration project, which alone entails substantial uncertainties, and it was being done in a data poor environment, it was not possible to use a full array of techniques in describing the uncertainty encountered in a typical project. For the most part, this research project relied on the simplest means of quantifying uncertainty consistent with the available data and District personnel's available time and interest. The purpose of this Appendix is to illustrate some of the alternative approaches that could have been used to quantify uncertainty.

## THE SETTING

To illustrate alternative techniques the data in the table below will be used. These are real dredging project data from another Corps District. These data are production information for 30 inch pipeline dredging projects. For the purposes of the examples that follow, let us suppose an estimator is trying to quantify the uncertainty in the gross cubic yards per hour on a new project the District is planning. The techniques used in this example have broad carryover value.

| 30-INCH PIPELINE DREDGE PRODUCTION RATES | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Dredge** | **Size Dia. (In.)** | **Max Pipeline (Ft)** | **Avg. Pumping Dist. (Ft)** | **Bank Height (Ft)** | **Net Ewt (%)** | **Gross Cy/Hr** | **Pay Cy/Mo** |
| Alaska | 30 | 22,250 | 11,250 | 6.1 | 54.5 | 1,772 | 50,2965 |
| Illinois | 30 | 18,500 | 11,600 | 70 | 66.8 | 2,363 | 104,0318 |
| Illinois | 30 | 39,500 | 2,700 | 5.3 | 48.3 | 1,752 | 55,9519 |
| Alaska | 30 | 10,000 | 5,000 | | 37.8 | 744 | 14,4808 |
| Bill James | 30 | 20,000 | 11,000 | 3.3 | 29.9 | 929 | 12,5324 |
| R.S. Weeks | 30 | 20,800 | 16,000 | 7.2 | 61.4 | 1,065 | |
| Alaska | 30 | 22,000 | 17,167 | 8 | 27.4 | 2,139 | 36,5618 |
| Meridian | 30 | 26,500 | 19,250 | 4 | 49.1 | 1,786 | 53,3434 |
| R.S. Weeks | 30 | 29,333 | 26,333 | 4.83 | 60.4 | 1,453 | 60,6903 |
| Illinois | 30 | 29,000 | 19,000 | 2.4 | 29.6 | 1,235 | 39,6109 |
| | | **23,788** | **16,360** | **5.3** | **46.5** | **1,524** | **47,5000** |

## QUANTIFYING UNCERTAINTY

The techniques presented here are representative of some of the options that would be available for quantifying uncertainty or describing variability in cost estimate inputs. Variation will be used to include uncertainty and/or variability.

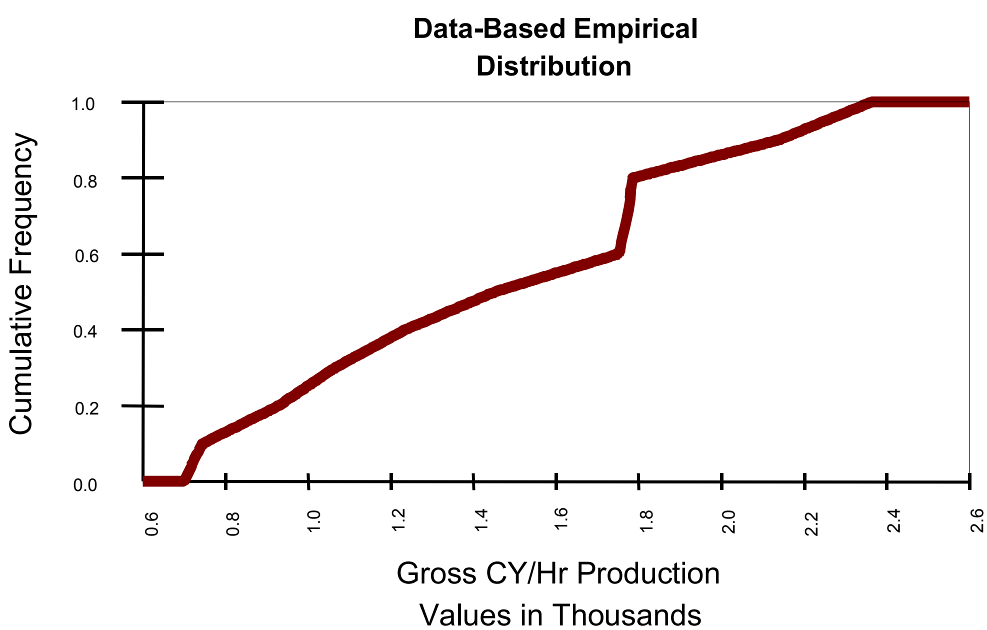
## Use Data

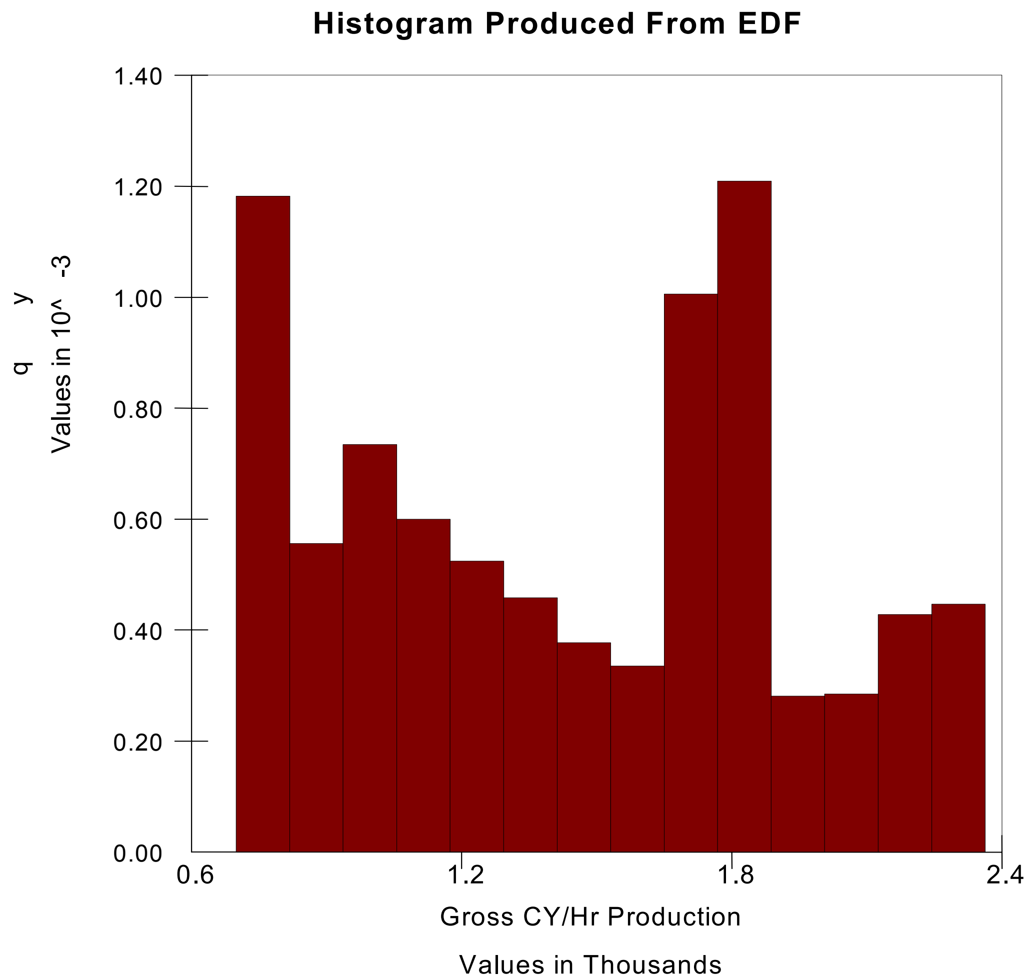
If data are available from other projects or databases, such as M-CACES, it may be advisable to use the data to describe the variation in the input. Using the data directly would be most useful when the data are directly representative of the quantity to be estimated.

The data of interest when sorted yield an empirical distribution as shown below. An empirical distribution simply says, these are the data.

| EMPIRICAL DISTRIBUTION OF PRODUCTION RATES | |
|---|---|
| **Gross CY/HR** | **Cumulative Frequency** |
| 744 | 0.1 |
| 929 | 0.2 |
| 1065 | 0.3 |
| 1235 | 0.4 |
| 1453 | 0.5 |
| 1752 | 0.6 |
| 1772 | 0.7 |
| 1786 | 0.8 |
| 2139 | 0.9 |
| 2363 | 1 |

Graphically this empirical distribution looks like the following. Five thousand iterations of the above distribution yields a histogram like that below. This histogram does not match any known distribution in appearance. It simply shows the data.





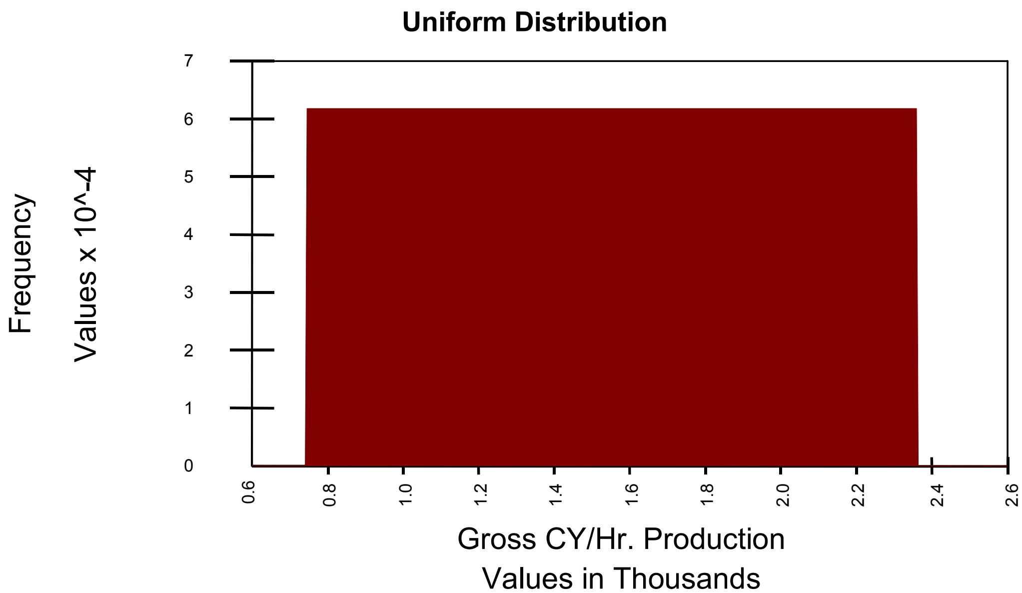

## Non-Parametric Distributions

If you do not use the data directly to describe the variation you will have to resort to some sort of distribution. General distributions that do not require any great knowledge of underlying assumptions can be quite useful. They are often called non-parametric distributions. Some commonly used examples include the general empirical distribution above and the triangular and uniform distributions. Non-parametric distributions can be quite useful in preliminary risk modeling and even advanced risk assessment models.

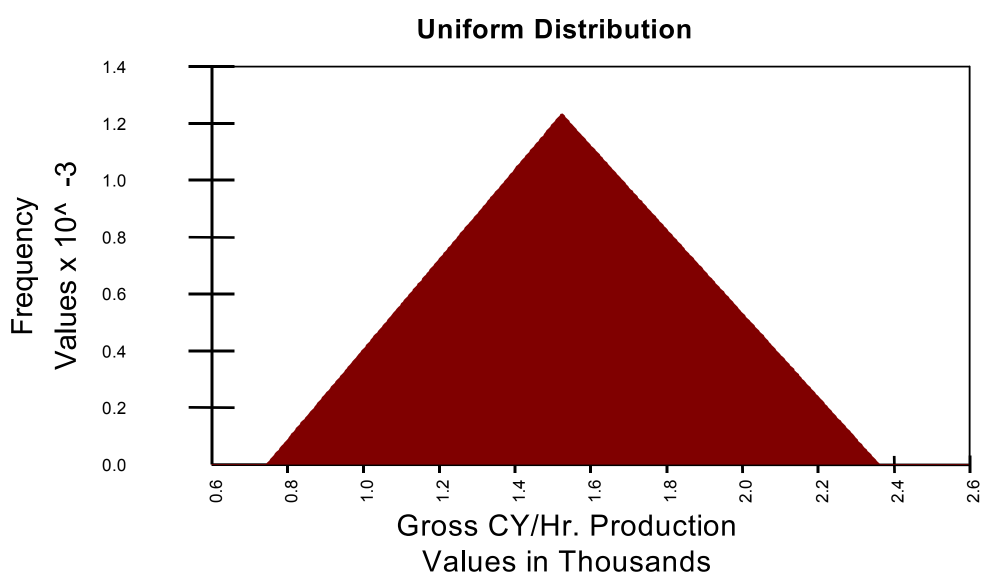
The uniform distribution is a sort of maximum uncertainty distribution. You need only a minimum and a maximum possible value. All values between these two are assumed to be equally likely. It is used in those very rare cases when all we know are the minimum and maximum possible values. It is a rare situation when we do not know at least something more than that. But when we do not, we can use a uniform distribution. The minimum and maximum may be based on data or expert opinion.

Given the data available for dredge production here there would be no reason to use a uniform distribution. However, if we lacked data and knew simply that the rate was 744 CY/Hr.

for one project and 2,363 CY/Hr. for another we would have enough to create a crude distribution. The resulting uniform distribution is shown below.



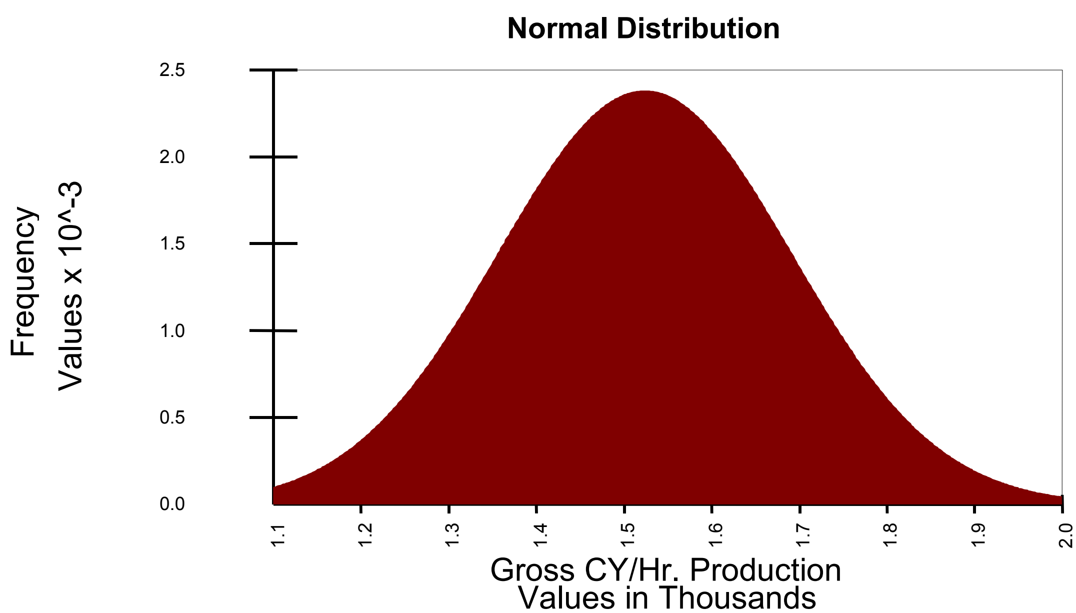

The triangular distribution was discussed in the text. It requires a minimum and maximum but is distinguished from a uniform distribution by identification of a most likely value. Using the data above there are several choices for the most likely value which is usually a mode. Absent a mode, the mean (1524) or median (1603) are reasonable choices. The triangular distribution using the mean is shown below.



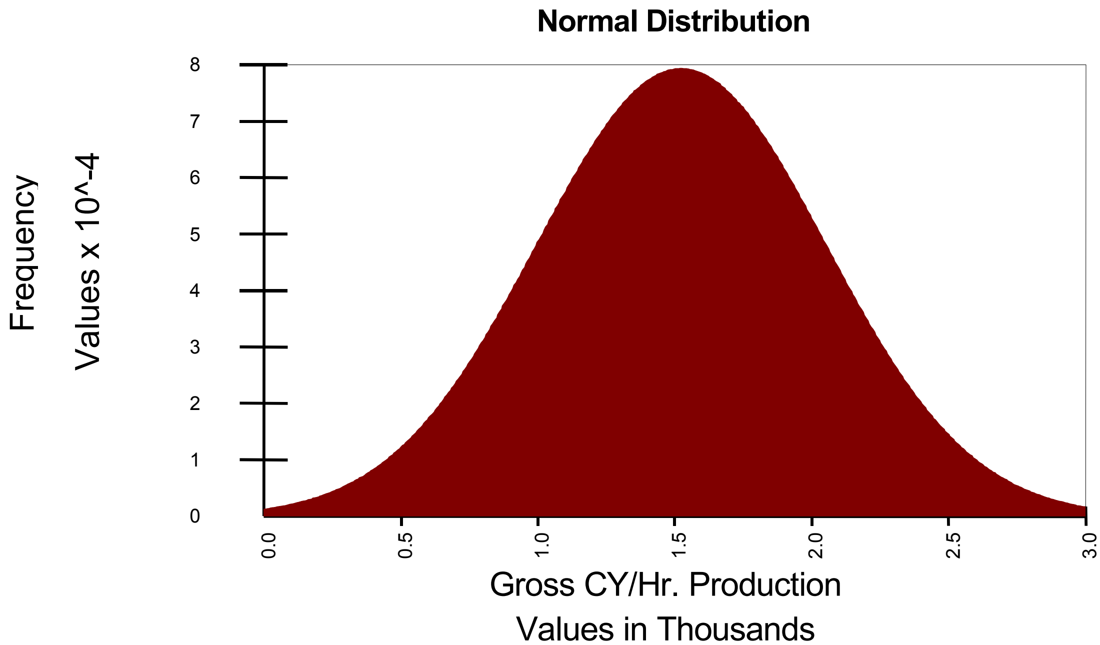
## Parametric Distributions

Theoretical or parametric distributions require more knowledge of probability distributions. They are to be used when there is theory that suggests that a particular distribution should be used. It is also useful when a particular distribution has proven useful in the absence of supporting theory. Parametric distributions often fit expert opinion especially when the required level of accuracy is not great. The normal distribution is a good example of a parametric distribution.

Suppose we are only interested in the expected value of our distribution. If we regard our ten data points as a random sample and calculate the sample mean (1524) and its standard error (168), the Central Limit Theorem suggest the distribution of sample means is itself normally distributed if the sample is large (n>30, not met in this case) or if the population from which our sample is drawn is normal, an assumption made for the convenience of this appendix. The resulting normal distribution of the mean gross cy/hr. production is shown in the sampling distribution below. Note the range of the horizontal axis.

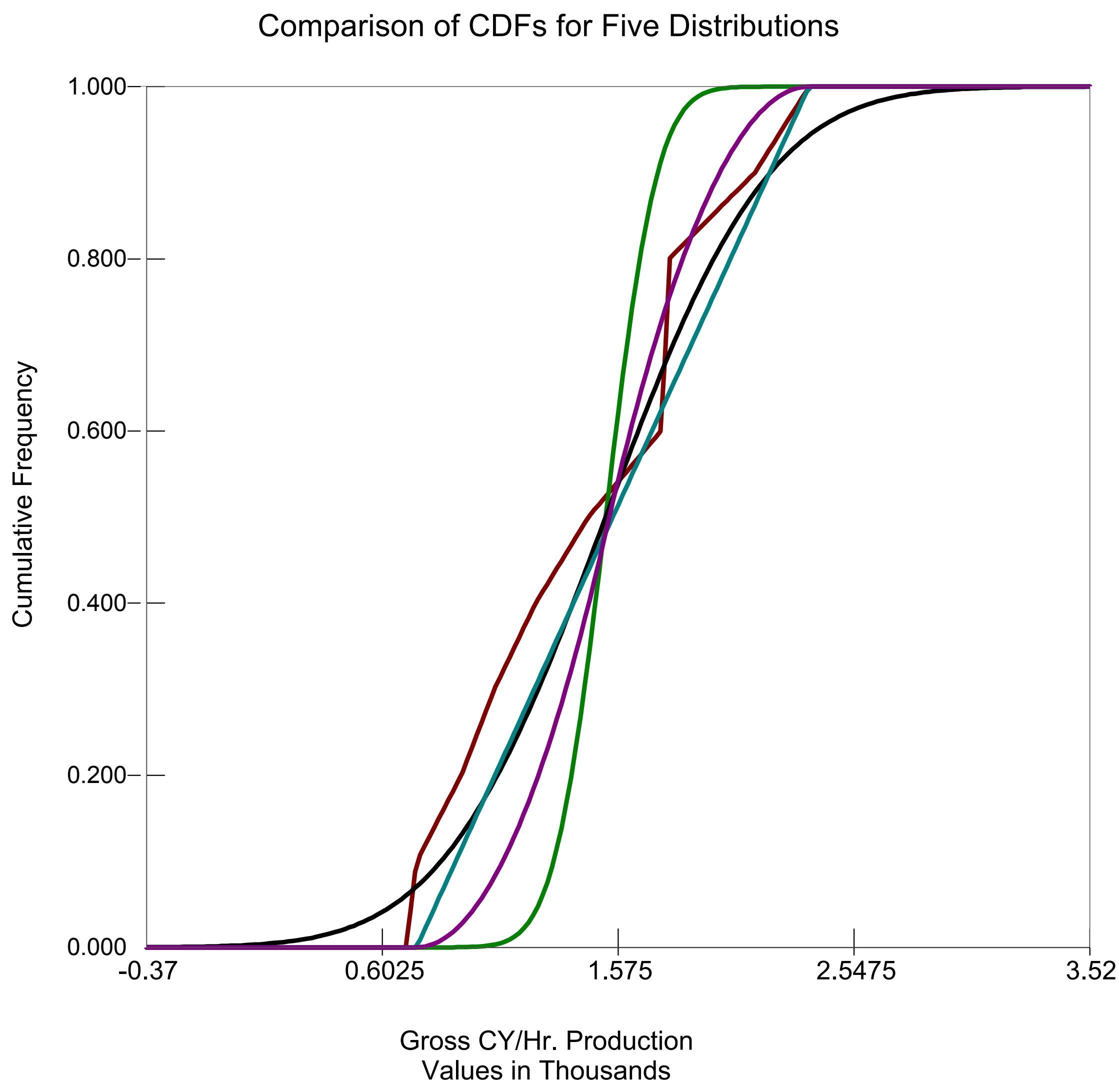



An alternative use of the normal distribution would result if it has been shown that using a normal distribution for production rates has been accurate in the past, another assumption made for the convenience of this appendix. The ten data points are then used to estimate the population mean (1524) and the population standard deviation (530), the two parameters required to define a normal distribution. The resulting normal distribution is shown below. The primary difference between the two distributions is the range of possible values. The former distribution estimates only the mean production rate, this distribution estimates possible individual production rates.




## DOES THE DISTRIBUTION MATTER?

This appendix has suggested empirical, uniform, triangular, sample mean and normal distributions can be estimated from the same ten data points. Does the choice of a distribution matter? It does. A 5000 iteration simulation was run for each of these probability distributions. The results are shown graphically below.




This graph is difficult to read in black and white. The first curve to rise above the axis on the left is the normal distribution. The second one is the empirical distribution. The third, a straight line, is the uniform distribution. The fourth distribution to rise from the axis moving left to right is the triangular distribution. The last distribution to begin its rise is the sample mean distribution.

All of them converge at the 50$^{th}$ percentile indicating similar expected values. These distributions vary in their distribution of values above and below the mean. The steeper the curve the less variation present in the data. In other words, the curve that rises from zero to one on the vertical axis over the least horizontal space has the least variation. In this example the normal distribution shows the greatest variation. The sample mean shows the least variation.

The purpose of this appendix has been to indicate that there are many different ways to characterize the uncertainty about any model input. In any given situation some of these options will be better than others. The resolution of these issues can be daunting for those just learning the techniques. Consequently, this research relied principally on the use of non-parametric distributions.