

Incorporating Psychological Influences in Probabilistic Cost Analysis

Edouard Kujawski^{1, *}, Mariana L. Alvaro², and William R. Edwards¹

¹ Engineering Division, Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720

² Department of Statistics, California State University, Hayward, California 94542

ABSTRACT

Today's typical probabilistic cost analysis assumes an "ideal" project that is devoid of the human and organizational considerations that heavily influence the success and cost of real-world projects. In the real world "Money Allocated Is Money Spent" (MAIMS principle); cost underruns are rarely available to protect against cost overruns while task overruns are passed on to the total project cost. Realistic cost estimates therefore require a modified probabilistic cost analysis that simultaneously models the cost management strategy including budget allocation. Psychological influences such as overconfidence in assessing uncertainties and dependencies among cost elements and risks are other important considerations that are generally not addressed. It should then be no surprise that actual project costs often exceed the initial estimates and are delivered late and/or with a reduced scope. This paper presents a practical probabilistic cost analysis model that incorporates recent findings in human behavior and judgment under uncertainty, dependencies among cost elements, the MAIMS principle, and project management practices. Uncertain cost elements are elicited from experts using the direct fractile assessment method and fitted with three-parameter Weibull distributions. The full correlation matrix is specified in terms of two parameters that characterize correlations among cost elements in the same and in different subsystems. The analysis is readily implemented using standard Monte Carlo simulation tools such as @Risk and Crystal Ball®. The analysis of a representative design and engineering project substantiates that today's typical probabilistic cost analysis is likely to severely underestimate project cost for probability of success values of importance to contractors and procuring activities. The proposed approach provides a framework for developing a viable cost management strategy for allocating baseline budgets and contingencies. Given the scope and magnitude of the cost-overrun problem, the benefits are likely to be significant.

Key words: budget allocation, correlation, cost contingency, critical chain project management, engineering/expert judgment, human behavior, probabilistic cost analysis, systems thinking

*Author to whom all correspondence should be addressed (e-mail: e_kujawski@lbl.gov).

1. INTRODUCTION

Real-world experience and intuition both suggest that project costs depend on many factors including technical, organizational, and behavioral considerations. It seems that based on brilliant insight Thucydides got to these as the very root of cost overruns over 2000 years before the formal development of probability and risk analysis when he stated [Augustine, 1997: 255]:

*"Their judgment was based more on wishful thinking than on sound calculation of probabilities."*¹

This observation is still very insightful and applicable to today's cost-overrun problem.

Numerous and significantly different models and techniques have been developed to estimate costs. There is an on-going shift in R&D and complex engineering projects from deterministic to probabilistic approaches. The deterministic approach has major shortcomings such as specifying a single point estimate for each cost element and calculating contingency as a percentage of the total. The probabilistic approach provides the proper framework for handling cost uncertainties based on probability distributions and contingencies based on a desired probability of success. Unfortunately, today's typical Probabilistic Cost Analysis (PCA) seldom addresses the behavioral and organizational considerations that have been documented by Sage [1981] as essential to project success. It therefore generates "mythical project costs" that are based on what Brooks [1995] has termed the "mythical man-month". Monte Carlo simulation being a mathematical tool cannot by itself compensate for "Garbage In...Garbage Out" [Walker II and Cox III, 2003]. It should then be no surprise that actual project costs often exceed the initial estimates and are delivered late and/or with a reduced scope. These include the full spectrum of projects: remodeling one's residence, complex infrastructure construction projects such as Boston's Central Artery/Tunnel Project [Hughes, 1998], R&D oriented projects such as the spectacular Hubble Space Telescope and the canceled Superconducting Super Collider, and multitudes of commercial ventures. Projects that come-in under cost, within schedule, and meet all requirements (including non-functional ones such as quality/reliability) do not necessarily deserve kudos. They may have carried budgets with excessive padding that lead to unnecessarily high costs and misallocation of resources. In today's highly competitive business environment, it is therefore critical to improve the realism of cost estimates and how cost budgets are managed.

In recent years psychologists have studied and quantified human behavior and judgment under uncertainty. Their findings are having a profound influence in many fields including decision-making, management, and economics [Rabin, 1998]. Goldratt [1997], focusing on negative human behavior as major causes of project-scheduling problems, developed the Critical Chain Project Management (CCPM) as a management philosophy and solution that simultaneously reduces project duration and protects against schedule risk. Some of the identified negative behaviors, such as Parkinson's Law - "Work

¹ The ancient Greek word, "eikos", encapsulates ideas of the modern concept of probability [Bernstein, 1996: 16].

expands to fill the time allotted" - and "padding schedule estimates", directly contribute to cost overruns. One of the key principles of CCPM is to aggregate task buffers at the project-level for use where and when needed. The original CCPM also proposed the following guidelines for sizing buffers: (1) Cut task duration estimates in half, and (2) Add approximately 25% of the original estimate to the project buffer. These guidelines appear to be rather arbitrary and many technical managers are uncomfortable with them [Givens Filiatrault and Peterson, 2000]. Another shortcoming of the original method is that it only approximately treats the probabilistic nature of project risks. A number of simple alternatives to estimate and sum buffers have been proposed [Newbold, 1998]. We think that their use is now no longer justifiable because of the availability of MCS tools such as @Risk and Crystal Ball®. Schuyler [2000: 187] proposes integrating the CCPM into a full probabilistic analysis. In a detailed report, Herroelen et al [2002] state: "The critical chain scheduling/buffer management methodology has much to offer if applied wisely and if the practical implications and limitations are well understood." There are numerous real-world applications that substantiate this conclusion [Schragenheim and Dettmer, 2001].

In the 1990's the Lockheed Missiles and Space Co. carried out a study which concluded that the following deficiencies in cost modeling and contingency management have been major contributors to both project high costs and overruns [Gordon, 1997]:

1. Invalid mathematics such as arithmetically summing uncertain cost elements instead of using statistical methods
2. "Money Allocated Is Money Spent" (MAIMS principle)
3. Failure to coordinate cost analysis and cost management
4. Hidden incentives in management styles
5. Hidden incentives in procurement processes.

It is standard practice for project management to allocate definite budgets to cost elements and maintain a budget contingency for dealing with unforeseen in-scope events. The MAIMS principle captures the fact that given this situation, cost under runs are rarely available to protect against cost overruns while task overruns are passed on to the total project cost. Gordon's [1997] numerical studies strongly indicate that a realistic PCA needs to account for the MAIMS principle. Unfortunately today's typical PCA assumes an "ideal" project where cost savings achieved by cost elements which are under cost are passed on to elements requiring additional resources. Given the wide use of this unrealistic assumption, it is not surprising that cost overruns are the rule rather than the exception. The MAIMS principle has important implications for budget allocation and contingency management. To deliver a successful project at an optimal cost, project management needs to allocate "reasonable" budgets to the cost account managers and dynamically manage the contingency funds as a risk portfolio at the project level [Kujawski, 2002a and 2002b].

The premise of this paper is that PCA needs to integrate psychological findings with mathematically valid models and sound management techniques to obtain realistic cost estimates and achieve project success. Building on these concepts, we develop a

practical yet realistic and mathematically valid model that remedies several shortcomings that are prevalent in today's PCAs and adversely impact project management. In Section 2 we review findings on human behavior under uncertainty and organizational considerations that contribute to cost overruns on R&D and complex engineering projects. In Section 3 we examine some mathematical properties of statistical sums and their implications for structuring the cost model and correlations among cost elements. In Section 4 we advocate the use of the Direct Fractile Assessment (DFA) method for subjectively assessing cost elements and fitting them with three-parameter Weibull distributions. In Section 5 we develop a correlation model that accounts for correlations among cost elements at both the subsystem and project levels. In Section 6 we apply the proposed approach to a representative design and engineering project. In Section 7 we integrate the presented concepts into a comprehensive approach to budget management including realistic cost estimates, efficient budget allocation, and dynamic management of contingencies. We conclude in Section 8 with a summary and recommendations for further development.

2. PSYCHOLOGICAL ASPECTS OF COST OVERRUNS

"Seek simplicity, then distrust it."
Alfred North Whitehead

2.1. Causes of Cost Overruns

As discussed in the Introduction, there are numerous elements that contribute to project cost overruns. For the purpose of this paper, we offer the following candidate "Top 10 List" sequentially ordered by the project activities of procurement, cost estimation, and budget management:

- R.1. Ineffective procurement processes that encourage unrealistic low bids
- R.2. Management pressure to produce low, can-do estimates
- R.3. Cost estimates based on poor project definition
- R.4. Overly optimistic assessment of unknowns
- R.5. Use of deterministic rather than probabilistic methods
- R.6. Use of inadequate data elicitation methods
- R.7. Use of improper cost distributions
- R.8. Neglect of interrelationships among cost elements and risks
- R.9. Inadequate allocation of budgets
- R.10. Inadequate management of cost contingencies.

These causes are not simply a lack of technical expertise. They are often driven by psychological, organizational, and institutional considerations. There is no unanimous "Top 10 List" and we apologize for listing these and omitting others.

Causes R.1, R.2, and R.3 are often institutional and organizational problems and they are outside the scope of this paper. These can be addressed through legislation that deals with reforms in procurement or acquisition [Senator William Cohen, 1982].

Cause R.4 is a characteristic of human behavior and judgment under uncertainty. As indicated in the Introduction, psychologists are quantifying these effects and behavioral economists [Rabin, 1998] are building models that remove some of the assumptions of the 100% rational economic person. PCA needs to account for these psychological influences to provide realistic cost estimates and result in successful projects for lower costs.

Causes R.5, R.6, R.7, and R.8 are often justified on the basis of simplicity and expediency. Whether or not they are caused by a lack of technical expertise, they represent over-simplified approaches and invalid assumptions that lead to erroneous cost estimates. We speculate that they may also be symptomatic of analysts and decision makers who have succumbed to Russo and Schoemaker' s [1990: 16] Decision Trap Number 2, Frame Blindness, that leads them to develop only a partial view of the problem.

Causes R.9 and R.10 are often categorized as managerial errors that arise from a lack of coordination between the project cost analysis and management activities. We think that they are also intertwined with a lack of appreciation of the probabilistic concepts and psychological influences. Causes R.9 and R.10 lead to cost overruns on top of high cost estimates. High contingencies by themselves are not sufficient to prevent cost overruns; they also need to be dynamically managed as a portfolio at the project level [Kujawski, 2002a and 2002b].

2.2. Some Heuristics for Cost Analysis

We now present several heuristics that provide useful insights into the psychological considerations of the cost-overrun problem and lay the foundation for the proposed approach.

2.2.1 Systems Thinking

“Dividing an elephant in half does not produce two small elephants.”
Law # 10 of the Fifth Discipline [Senge, 1994: 66]

Similarly, it would be misleading to think that the total cost of R&D and complex engineering projects consists simply of the sum of the individual cost elements. Project characteristics and risk events are likely to affect multiple cost elements in a common way, either negative or positive. The realization of any risk can start a ripple effect, and if uncontrolled it can snowball into catastrophic consequences. The implications are that PCA requires systems thinking to go beyond the individual hardware and software components and that it needs to focus on the overall project structure with emphasis on the interrelationships among project elements. A realistic cost analysis takes a multidisciplinary effort with a systems engineering approach [Garvey, 2000].

2.2.2. Estimating Cost Elements

There are numerous conflicting psychological and organizational influences that motivate technical leads to provide either low or high cost estimates. These need to be considered on a case-by-case basis when calibrating subjectively assessed cost elements. We do not think that there is a one-size-fits-all solution for assessing and calibrating uncertain cost elements. The one rule is to use experienced and knowledgeable analysts.

Underestimates. Overconfidence tends to be a prevalent behavioral trap in human judgment and it can lead to catastrophic consequences [Pious, 1993]. Numerous studies [Alpert and Raiffa, 1982] have found that people, including experts, tend to be overconfident when assessing uncertain quantities and risks. Alpert and Raiffa [1982: 300] strongly advise:

*"For heaven's sake, **Spread Those Fractiles!** Be honest with yourselves! Admit what you don't know!"^{2, 3}*

In R&D and complex engineering projects there are additional considerations that further motivate cost estimates that are biased towards low values. These include (1) R&D folks are intrinsically optimistic regarding new technologies, especially those coming from their own laboratories, (2) psychological influences, such as "impress your boss", (3) organizational influences, such as "hidden incentives", and (4) market forces.

Inflated Estimates. Some project management experts [Goldratt, 1997; Newbold, 1998] believe that technical leads inflate cost and schedule estimates to make it easier to achieve success. This is not our experience on R&D and complex engineering projects, and we surmise that the noted negative behavior depends on the work environment and project type. In organizations where cost overruns are considered unacceptable the staff quickly learns to pad estimates. But the project managers, having risen through the ranks, are familiar with the rules of the game and automatically reduce the estimates often in an ad-hoc manner. PCA cannot work in such an environment; the net result is a lose-lose situation.

2.2.3. The MAIMS Principle

"Money Allocated Is Money Spent."
Gordon [1997]

As discussed in the Introduction, the MAIMS principle is the money-analog of Parkinson's Law and it has been identified as a major cause of significant cost overruns. Unfortunately, today's typical PCA assumes an "ideal" project with a 100% rational economic staff with the consequences that whenever a cost element comes in low the savings are available where they are needed. It is no surprise that actual projects exceed these "ideal" cost estimates.

2.2.4. Assessing Analysis Uncertainties

"It's not what we don't know that gives us trouble, it's what we know that ain't so."

Will Rogers [Russo and Schoemaker, 1990: 95]

PCA is an important and powerful tool when used correctly; but it can also be a minefield for the unwary. The literature abounds with PCAs that contain some or all of the following oversimplifications: use of triangular distributions, omission of interrelationships among cost elements, no treatment of budget allocation and contingency management, no estimates of confidence level or uncertainty for the high percentiles. Such oversimplified PCAs provide erroneous information that can lead cost

² Emphasis in the original text.

³ "Fractile" is a mathematical concept. Details are given in Section 4.1.

account managers and other decision-makers into a false sense of confidence and incorrect decisions [Conrow, 2003].

It is important that the PCA be properly performed and documented in accordance with good practices of probabilistic risk assessment [EPA/630/R-97/001, 1997]. Relying on a poorly documented PCA is well characterized by Russo and Schoemaker's [1990: 95] Decision Trap Number 5, "implicitly trusting the most readily available information or anchoring too much on convenient facts." In order for a PCA to be a viable tool for analyzing and managing costs and cost risks, its documentation should include:

1. The rationale for the selection of the model
2. The methods for evaluating the uncertain cost elements, modeling correlation effects, and performing the statistical sum
3. Graphical representations and statistical parameters of the input and output distributions
4. Sensitivity studies or other appropriate investigations of the data, modeling, and numerical uncertainties
5. Explicit discussion of the range of applicability of the results.

3. STRUCTURING THE PROBABILISTIC COST ANALYSIS

In this section we present a few analytical and numerical results that illustrate the importance of correlations among cost elements and the Cost Work Breakdown Structure (CWBS) to the validity of the PCA. Detailed models and additional analyses are presented in Sections 4, 5, and 6.

3.1. Some Analytical Results

To focus on PCA we consider a generic project with WBS level-3 cost elements [MIL-HDBK-881, 1998]. In general, there are interrelationships among the cost elements because of their dependence on common factors such as state of technology, complexity, criticality, management, staff, and product development process [Browning and Eppinger, 2002]. In anticipation of the treatment of cost correlation presented in Section 5, we denote the cost elements as $C_{i,j}$, where the 1st set of integers refers to the WBS level 2 and the 2nd set of integers refers to the WBS level 3. The total project cost is then a random variable given by

$$C_T = \sum_{m,j} C_{m,j}.$$

The expected value of C_T is simply the sum of the expected values of the individual cost elements,

$$E(C_T) = \sum_{m,j} E(C_{m,j}). \quad (1)$$

In contrast, the variance of C_T depends on the variances of the individual cost elements and the correlation among them [Garvey, 2000],

$$\text{Var}(C_T) = \sum_{m,j} \text{Var}(C_{m,j}) + \sum_{m,j} \sum_{n,l \neq m,j} \text{Corr}(C_{m,j}, C_{n,l}) * [\text{Var}(C_{m,j}) * \text{Var}(C_{n,l})]^{1/2}. \quad (2)$$

The correlation coefficients are dimensionless parameters with values between ± 1.0 for correlated random variables and a value of 0.0 for independent random variables.

Equations (1) and (2) provide much insight into how to structure the cost model and the validity of the PCA. Consider the case of n correlated cost elements $C_{m,j}$. Let σ_M^2 denote the largest $\text{Var}(C_{m,j})$. Since the correlation coefficients range between ± 1.0 , $\text{Var}(C_T)$ satisfies the following inequality,

$$\text{Var}(C_T) \leq n * \sigma_M^2 + n * (n - 1) * \sigma_M^2. \quad (3a)$$

The 2nd term dominates as the number of correlated cost elements increases. The behavior is significantly different if the cost elements are treated as independent random

variables. In the absence of correlations the 2nd term vanishes and the C_T distribution narrows as n increases:

$$\sqrt{\text{Var}(C_T)} / E(C_T) \propto 1/\sqrt{n}. \quad (3b)$$

Equation (3b) clearly flags the trap or fallacy of using low-level cost elements and neglecting their interrelationships. If correlations among cost elements at all levels were negligible, analysts theoretically could arbitrarily reduce the uncertainty in project costs by simply decomposing the total cost into lower and lower level cost elements. Furthermore the Central Limit Theorem [Garvey, 2000] becomes applicable and one would conclude that the total cost of any complex system is adequately characterized by a Gaussian normal distribution. In reality, the costs of R&D and complex engineering projects are best characterized by distributions that are skewed to the right and are relatively broad. In Section 5 we go beyond these mathematical arguments and identify physical causes for the existence of correlations and more complex dependencies among cost elements within a single project.

3.2. Some Quantitative Studies

To illustrate the above observations and gain additional insight into the accuracy of MCS, we present the results for the sum of ten correlated three-parameter Weibull distributions. The calculations were performed using Crystal Ball®. An accurate MCS of the low probability values is challenging especially for open-ended PDFs because even with many trials the number of samples in the tail is small. The reported results with both 10,000 and 100,000 trials are indistinguishable up to and including the 95th percentiles; but, the differences or numerical uncertainties become significant the closer one gets to the extremes even for the closed-ended PDFs.

Figure 1a depicts the impact of different correlations among cost elements. For positively correlated cost elements, a high value of one is likely to be associated with a high value of the other and similarly a low value with a low value. The net effect is to spread the PDF of the total cost. Some of the statistics are summarized in Table I. The PDF for independent cost elements (correlation coefficient $\rho = 0.0$) is approximately given by a Gaussian normal distribution with mean $\mu_T = 10*\mu_i$ and standard deviation $\sigma_T = \sqrt{10} * \sigma_i$, where μ_i and σ_i denote the mean and standard deviation of the individual cost element. This is consistent with the Central Limit Theorem [Garvey, 2000]. For $\rho = 1.0$ the ten identical Weibull distributions are perfectly correlated and their sum is simply the individual component multiplied by 10. This results in a Weibull distribution with $\mu_T = 10*\mu_i$ and $\sigma_T = 10*\sigma_i$. The plots for intermediate values of the correlation coefficients $\rho = 0.5$ and 0.3 are almost indistinguishable. Figure 1b compares the corresponding Cumulative Distribution Functions (CDFs). The impact of correlations among the cost elements appears to be less dramatic, and for greater clarity we explicitly plot the differences in Fig. 1c. It should be noted that we manually set the different 0th and 100th percentile values equal to compensate for numerical uncertainties. The correlation effects

are significant over a wide range of percentiles of interest to PCA such as the 75th to 95th values. The 90th percentile values corresponding to $\rho = 0.0$ and 1.0 differ by $\sim 10\%$. Since the above illustrative examples consist of only ten identical elements, the results should be principally viewed as qualitative. The correlation effects increase with the number of cost elements [Book, 1999 and 2000/2001]. In general, subdividing the project cost into too many bite-size pieces has several negative consequences including an unwarranted level of analysis, an increased likelihood of erroneous results, and a false sense of confidence. We also note that, independent of any model, correlation effects vanish for the mean, 0th and 100th percentile values.

4. SUBJECTIVE ASSESSMENT OF COST ELEMENTS

R&D and complex engineering projects rely heavily on engineering/expert judgment for the assessment of uncertain cost elements. Unfortunately the subjective assessments are often performed in a rather ad-hoc manner, and they have been identified as a critical source of uncertainty in probabilistic risk analyses [Keeney and von Winterfeld, 1991]. In this section we advocate the use of the Direct Fractile Assessment (DFA) method for data elicitation and fitting each cost element with realistic continuous PDFs.

4.1. The Direct Fractile Assessment Method

The DFA method has been investigated in numerous psychological experiments and found to provide one of the most reliable and least bias-prone procedures for eliciting uncertain quantities [Alpert and Raiffa, 1982]. Individuals are asked to assess uncertain quantities by providing several values, referred to as percentiles or fractiles. The k^{th} percentile or fractile x_k is the value that is assessed to have a $k\%$ probability of exceeding the actual value x^* ; mathematically speaking, $P(x^* \leq x_k) = k\%$. We summarize some of Alpert and Raiffa's findings:

1. People have a systematic bias toward overconfidence. The subjective probability distributions tend to be too tight. Typically 33% rather than 50% of the actual values fell within the 0.25 to 0.75 fractiles.
2. The judgment of extreme values is significantly worse. Typically, 20% rather than 2% of the actual values fell outside the 0.01 to 0.99 fractiles.
3. Minimum and maximum values are vague terms. "We really don't know what that means."

The above findings are based largely on laboratory experiments, but they deserve the critical consideration of cost and risk analysts. The DFA has been used in numerous applications including the cost assessment of large technology projects [Dillon et al, 2002]. Based on this experience, we propose the following guidelines for subjectively estimating cost elements of R&D and engineering design projects:

1. Use experts to provide the 10th, 50th, and 90th percentiles of the uncertain cost elements. While other percentiles may be used, these seem to be more easily assessed by experts [Dillon et al, 2002]. For similar reasons, analysts should avoid seeking extreme values, abstract measures such as the mean and the standard deviation, or specific distribution functions. Dillon et al [2002: 56] propose an additional step of "educating the experts about the possibility of bias and discussing what the base cases were and were not."
2. Calibrate each set of percentiles to reflect individual and project specific considerations be it optimism or pessimism. Several different approaches have been proposed, but there is need for further studies [Clemen and Lichtendahl, 2002]. As a default calibration to account for overconfidence and as a defense against overly optimistic estimates, cost analysts might opt to shift the assessed 90th percentile to the

- 80th or 75th percentiles. Because of the MAIMS modification (see Section 4.3) calibration of the low-end percentiles such as 10 vs. 20% is of negligible importance.
3. Explicitly model decision points and low-probability/high-consequence events using scenarios as discussed in Section 4.4.
 4. Regularly update the cost analysis. Cost uncertainties decrease as the design matures and the project evolves.

4.2. Selection of a Probability Distribution Function

There are many good reasons for modeling the cost elements with continuous PDFs:

1. Uncertain cost elements are continuous rather than discrete random variables.
2. There are several well-known analytic PDFs with adequate characteristics.
3. There are several commercial tools such as @Risk and Crystal Ball® that provide extensive choices of PDFs.

To realistically reflect cost uncertainties and readily support the proposed data elicitation method, the selected PDFs should be flexible enough to satisfy the following criteria:

- C1. Capable of fitting three arbitrary percentiles
- C2. A finite lower range
- C3. An infinite upper range with reasonable behavior
- C4. Physically meaningful and easy to estimate parameters.

The need for criteria C1, C2, and C4 is self-evident. Criterion C3 is more subjective, but we consider it to be important because of the difficulty of assessing maximum values and the risks associated with the use of overly restrictive or unreasonably large ranges. We use the above four criteria to evaluate the suitability of the following four PDFs for characterizing uncertain cost elements: triangular, generalized Beta, three-parameter lognormal, and three-parameter Weibull. Analysts and assessors should always verify that they feel comfortable with the shape of the fitted distribution.

4.2.1. The Triangular Distribution

The standard form of the triangular distribution is to specify its minimum, most-likely value or mode, and maximum. Because of its very restrictive shape, it is often impossible to fit a triangular distribution to three arbitrary percentiles and it gets a very low score for C1. It also gets a low score for C3. The triangular distribution has been blamed for both unrealistically low and high estimates. Graves [2001] argues that it leads to underestimates because it has a finite upper bound. Moran [1999] argues that it leads to overestimates because it does not capture the expert's confidence level of achieving the most-likely value and/or knowledge of the shape of the distribution.

4.2.2. The Generalized Beta Distribution

The generalized Beta distribution, Beta (a , b , α , β), is a very flexible PDF with lower range a , upper range b , and two shape parameters α and β . Depending on the values of α

and β , the shape varies from uniform to bell-shaped to U-shaped. Because of this flexibility, it is often used for modeling construction activities [Fente et al, 1999]. Since the generalized Beta distribution has four independent parameters, three percentiles do not uniquely specify it and additional information is required to select a best shape. The generalized Beta distribution scores low on C3 and C4.

4.2.3. The Three-Parameter Lognormal Distribution

The three-parameter lognormal distribution, $\log N(L, \mu, \sigma)$, is an open-ended distribution with location L , mean μ , and standard deviation σ . It arises when many factors interact in a multiplicative way. It is used in diverse situations including reliability analysis [McCormick, 1981], cost analysis [Graves, 2001], and financial analysis [H. Forsey, 2001]. Because of its relatively slow fall-off it has been criticized for providing unreasonably high probabilities at high values. We give it a marginally acceptable score for C3. Since the lognormal distribution is always positively skewed, it scores low on C1.

4.2.4. The Three-Parameter Weibull Distribution

The three-parameter Weibull distribution, $W(L, \alpha, \beta)$, is an open-ended distribution with location L , scale parameter α , and shape parameter β . It is a very flexible PDF that can assume a wide variety of shapes depending on the value of β . It is used in numerous applications including reliability analysis [McCormick, 1981] and cost analysis [Dillon et al, 2002]. It is more flexible than the three-parameter lognormal even though both are characterized by three independent parameters [AbouRizk and Halpin, 1992]. We score it high on all four criteria. We consider the three-parameter Weibull well suited for use in PCA. Furthermore we think that the use of more complex PDFs is unwarranted when the input data is based on subjective assessments.

4.2.5. Some Numerical Comparisons

We use the cost element $C_{1.1}$ in Table IV (see Section 6) to quantitatively investigate the above PDFs and the effects of calibrating percentiles. The three-parameter Weibull, the three-parameter lognormal, and the generalized Beta distributions are flexible enough to provide successful fits in all of the examined cases; the triangular distribution is not. The three-parameter Weibull and generalized Beta PDFs and CDFs are shown Figs. 2a and 2b, respectively. Table II summarizes some key statistical parameters. The differences or errors of the means and standard deviations are minor compared with those of the high-percentile values. For a given set of three percentiles, the CDFs corresponding to the three-parameter Weibull and the generalized Beta distributions differ mainly outside of the specified ranges, where expert opinion is likely to be highly unreliable. The results suggest that while it is important to select suitable PDFs to fit the assessed percentiles, it is unwarranted to further discriminate among these fits because the differences are of secondary importance compared to the accuracy of the assessed percentiles.

4.3. MAIMS-Modified PDFs

The MAIMS principle accounts for the fact that account managers rarely under-run their original allocated budgets. This has important implications for PCA. Once a cost element is allocated a budget x^* it becomes a random variable with minimum value x^* rather than the lower range L of the original PDF. We refer to these PDFs as the MAIMS-modified PDFs. They are proper PDFs with a delta-like function at x^* that accounts for all random values less than or equal to x^* . We stress that the MAIMS-modified PDFs are not the same as the Crystal Ball® and @Risk truncated PDFs. The latter are generated by only considering the trials with random values between the specified ranges.

We use the Weibull (20/50/80) distribution in Table II to illustrate the consequences of the MAIMS principle for the following budget allocations: values of x^* equal to the mean, 50th, 75th, and 85th percentiles. The results are summarized in Table III and Fig. 3. Each MAIMS-modified CDF has a step-function behavior at x^* and is identical to the original CDF above x^* . Applying the MAIMS principle to a PDF increases its mean value and reduces its standard deviation. The impact increases with increasing values of x^* . The results strongly suggest that the MAIMS principle is likely to play a significant role in PCA. We further investigate this in Section 6 for a representative design and engineering project.

4.4. Modeling Specific Risks

The PDFs in Section 4.2 provide a macroscopic rather than a microscopic view of the project cost risk. They effectively model those factors or project characteristics that are ever present and contribute to cost uncertainties. But complex projects often involve a number of key or critical decisions and high-impact risks that need to be explicitly analyzed and risk responses developed. It is tempting to assume or claim that the cost PDFs also quantify these situations. We think that this is an invalid and counterproductive argument because it leads to the loss of valuable information and visibility into the high-consequence events. The analysis of specific risks and risk response actions requires a microscopic view and is best carried out using tools such as decision trees, influence diagrams, or other discrete representations [Kujawski, 2002b]. The microscopic perspective also assists risk analysts and domain experts to think about low-probability/high-consequence events and better deal with overconfidence or optimism biases. But it is too cumbersome and unrealistic to individually analyze every risk and source of cost uncertainties. In this paper we focus on the macroscopic perspective and acknowledge that work is needed to develop an integrated PCA that explicitly models risk reduction activities and provides a better decision-making tool.

5. A TWO-LEVEL CORRELATION MODEL (TLCM)

As discussed in the Introduction and Section 2, psychological influences and project characteristics such as complexity, dependencies, criticality, management, staff, and product development process are likely to simultaneously impact multiple cost elements at the subsystem and system levels. Furthermore, the realization of any one risk is likely to increase the probability and/or consequences of other risks. The cost elements are therefore correlated, and in general given N cost elements there are $N*(N - 1)/2$ correlation coefficients. These cannot assume arbitrary values between ± 1.0 because the correlation matrix must satisfy certain mathematical properties [Embrechts et al, 1999]. Some MCS tools including Crystal Ball® check the correlation coefficients for consistency and offer the option of adjusting them to construct a consistent correlation matrix. But there is danger in relying on an algorithm that may not capture the characteristics of correlation among cost elements [Hulett, 1999]. Clemen and co-workers [1999 and 2000] have done research on using expert judgment to assess correlation coefficients. Dillon and co-workers [Dillon et al, 2002] estimate correlations based on expert judgment about relationships between correlation and shared variance. The assessment of correlation coefficients is a difficult problem, and there is a definite need for a realistic and practical model that accounts for interrelationships among cost elements.

In this section we develop a correlation model that greatly reduces the number of parameters needed to specify a mathematically valid correlation matrix and accounts for the following two types of correlations:

1. Correlations among the cost elements within a given subsystem. These dependencies arise principally from the subsystem characteristics such as complexity and common staff.
2. Correlations among the cost elements in different subsystems. These dependencies arise principally from the organizational and programmatic considerations that are common to all cost elements because they are in the same project.

The model is based on Markowitz's multi-factor model [Markowitz, 1997]. Given its properties, we refer to it as the Two-Level Correlation Model (TLCM).

5.1. Development of the TLCM

Consider the cost elements at the WBS level 3 defined in Section 3. We model them as a sum of two random variables $R_{m,j}$ and F_m :

$$C_{m,j} = R_{m,j} + \alpha_{m,j} * F_m. \quad (4)$$

The $\alpha_{m,j}$ are constants. $R_{m,j}$ and F_m have the following properties:

1. The $R_{m,j}$ are independent random variables. $Cov(R_{m,j}, R_{n,k}) = 0$ for $m,j \neq n,k$.

2. The F_m are correlated random variables. $\text{Cov}(F_m, F_n) \neq 0$ for $m \neq n$.
3. $R_{m,j}$ and F_n are independent. $\text{Cov}(R_{m,j}, F_n) = 0$ for all m and n .

Following some algebraic manipulations, it can be shown that the TLCM correlation coefficients satisfy the following interesting relationship:

$$\text{Corr}(C_{m,j}, C_{n,k}) = \text{Corr}(F_n, F_m) * \alpha_{m,j} * \alpha_{n,k}. \quad (5)$$

The above expression provides a simple model for correlation coefficients that account for interrelationships among cost elements at both the project-wide (WBS level 2) and subsystem (WBS level 3) levels.

5.2. Some Properties of the TLCM

Consider a WBS that consists of M level-2 elements each of which consists of J level-3 cost elements, $C_{m,j}$ where $m = 1, \dots, M$ and $j = 1, \dots, J$, for a total of $M*J$ cost elements. From Eq. (5) it is seen that the TLCM requires $M*(M-1)/2$ values of $\text{Corr}(F_m, F_n)$ and $M*J$ values of $\alpha_{m,j}$. The number of parameters in the TLCM is then $M*(M - 1)/2 + M*J$ versus $M*J*(M*J - 1)/2$ for the general case. For example, for $M = J = 10$ the TLCM requires a total of 145 distinct correlation coefficients, which is a sharp reduction from the mathematically feasible number of 4,950. But even specifying this reduced number of parameters is a sizeable task and is not justified given the lack of relevant data. We therefore propose a simplified version of the TLCM that depends on only two parameters.

5.2.1. A Two-Parameter Model for the Correlation Coefficients

To further simplify the determination of the correlation matrix, we characterize the correlation coefficients with only two parameters, ρ_{int} and ρ_{ext} , that provide measures of the correlation among the cost elements of the same subsystem and different subsystems. The full correlation matrix is then given by

1. $\text{Corr}(C_{m,j}, C_{m,k}) = \rho_{\text{int}}$ for all values of m and $j \neq k$ (i.e. correlation among cost elements in the same subsystem)
2. $\text{Corr}(C_{m,j}, C_{n,k}) = \rho_{\text{ext}}$ for all values of j, k , and $m \neq n$ (i.e. correlation among cost elements in different subsystems)
3. $\rho_{\text{int}} > \rho_{\text{ext}}$.

5.3. Some Limitations of Correlation Coefficients

A correlation coefficient is a scalar quantity that only provides a measure of the linear or monotonic relationship between two random variables. It does not account for any non-linear dependencies. This limited information is somewhat analogous to the situation of different PDFs that have the same mean and variance but different higher-order moments. There is also a lack of statistical data and limited experience with subjectively assessing correlations [Clemen et al, 2000]. But these limitations and difficulties do not warrant

PCAs that assume no correlation among the cost elements. We agree with Book [1999 and 2000/2001] and others [Chapman and Ward, 2000] who advocate the use of reasonable correlation values because it provides more realistic and credible estimates of project costs and cost risks than assuming either no correlation ($\rho = 0.0$) or perfect correlation ($\rho = 1.0$). Studies in various fields, including financial risk management, have shown that correlation coefficients are often a satisfactory measure of association among uncertain elements [Embrechts et al, 1999]. These authors also make the point that rank-order correlation coefficients can assume any value between +/-1.0 and do not depend on any specific PDF. Rank-order correlation is implemented in several commercially available MCS tools including Crystal Ball® and @Risk. But there are other correlation measures, and they are not mathematically equivalent. Pearson's correlation measures linearity rather than monotonicity. Garvey [1999] stresses that the latter should be used for cost risk analysis with a CWBS because the cost variance depends on it. Recent studies on a realistic cost model [Smith and Hu, 2003] indicate that these differences do not have much of an impact on the total cost. We think that at the present time this is more of an academic than a practical issue.

In this paper we deal principally with cost elements that are assessed using expert judgment. But there are numerous projects that use Cost Engineering Relationships (CERs) that model functional relationships and thereby introduce correlations among the cost elements [Garvey, 2000]. For example, the integration and testing costs depends on common factors such as the complexity of the hardware and software, power, weight... There are also functional and statistical dependencies between cost and schedule. Cost-schedule joint probability distributions provide some modeling help, but they also have serious shortcomings. Embrechts et al [1999] stress that "marginal distributions and correlation do not determine the joint distribution." Garvey [2000: 332] states "...*they do not capture causal impacts* that schedule compression or extension has on cost."⁴ The treatment of dependencies in PCA is an active area of research and development. For the general situation, we propose integrating the TLCM with CERs and other functional relationships to account for behavioral and organizational dependencies.

⁴ Italics in the original text.

6. APPLICATION TO A REPRESENTATIVE DESIGN AND ENGINEERING PROJECT

To investigate the concepts and issues discussed in the previous sections we consider the hypothetical project with the Work Breakdown Structure (WBS) in Table IV. It is both rich and simple enough to illustrate the following effects on a PCA: (1) the assessment of the cost elements, (2) the modeling of correlation among them, and (3) the allocation of budgets and management of contingencies.

Given a WBS, the first step of a PCA is to develop an appropriate Cost Work Breakdown Structure (CWBS). Assume that the resulting CWBS is the WBS level 3 in Table IV. The second step is to systematically assess the cost elements using the DFA method. Assume that the 10th, 50th, and 90th percentiles in Table IV specify the assessed cost elements. These values may be further calibrated for biases in the assessments. The third step is to fit realistic PDFs to these percentiles. We use three-parameter Weibull distributions. At this step, the proposed approach further deviates from today's typical PCA and it implements a MCS with the following variations:

1. A baseline or budget is allocated to each cost element. Their sum constitutes the Project Baseline Cost (PBC).
2. Each cost element PDF is modified by setting its lower value to its allocated budget.
3. Statistical interrelationships among the cost elements are modeled using the TLCM.

Figure 4 depicts different budget allocation strategies for a given set of PDFs and correlation matrix. All calculations were performed with Crystal Ball® and 10,000 trials. The “ideal curve” corresponds to the model where each cost account manager rationally spends money only as necessary to satisfy the project requirements. The actual costs may be less than the budgeted costs and the savings are available to support other project elements on an as-needed basis. In the MAIMS_@_X50 and MAIMS_@_X75 curves all cost elements are allocated equal percentiles of 50% and 75%, respectively⁵. In the MAIMS_@_mean curve each cost element is allocated its mean or expected value. Each cost element is then budgeted at a percentile that depends on the PDF. Some interested parties may not consider this to be equitable; but in its defense we note that this approach has mathematical merit and compensates somewhat for high risks [Kindinger, 1999]. The results in Fig. 4 are consistent with those of Section 4.3. The MAIMS effects increase with increased allocated budgets. They depend on and are substantial over a wide range percentile values. We note that the MAIMS principle has little impact at very high confidence levels (CL > 95%) because at these CL values each contributing cost element must be near its maximum or 100th percentile value.

Figure 5 depicts the effects of different correlation values. The behavior is consistent with the results of Section 3.2. The effects of correlation are significant over a wide range of cost percentiles important to PCA. We note that the correlation effects peak

⁵ We use X_n to denote the n^{th} percentile of a cost element to differentiate it from the project cost PoS or percentile.

near the 0th and 100th percentile values and that they vanish at the mean, 0th and 100th percentile values.

Figure 6 depicts the effect of "spreading those fractiles" for a given correlation matrix and two different budget allocations. It has a substantial impact and it increases in importance with high Probability of Success (PoS) values. For example, in the MAIMS_@_mean situation, the effects of calibrating the assessed x_{10} and x_{90} values to the 20th and 80th percentile values dominate the effects of the MAIMS principle for greater than 60% PoS values. For completeness and convenience, some of the data is summarized in Table V.

Based on this analysis, we note that the assessment of the cost elements, the interrelationships among them, the budget allocation and management of contingencies constitute important and confounding factors [Freedman et al, 1991]. The results strongly suggest that realistic cost predictions require PCAs that simulate these effects simultaneously rather than individually.

7. BUDGET ALLOCATION, CONTINGENCY, AND PROJECT COST

In this section we integrate the concepts presented in the previous sections into a sound methodology for determining the Total Estimated Cost (TEC), allocating the baseline budgets, and managing the Management Cost Contingency⁶ (MCC). Our objective is to develop an optimal but realistic TEC for a given PoS that we assume has been specified by either project management, based on the desire to win the project and risk tolerance, or by the procuring activity.

The project manager or a designee, because of either contractual requirement or management expediency, allocates definite budgets that constitute the Project Baseline Cost (PBC) to the cost managers. Typically, he/she also establishes a MCC for management flexibility in executing in-work scope and dealing with unforeseen in-scope events, and allocates available contingency funds on an as-needed basis during the life of the project. Procuring activities in government [NASA, 2002; U.S. Department of Defense, 2002; U.S. Department of Energy, 2000] and commercial organizations [Smith and Merritt, 2002] often establish additional contingencies for work scope changes and other "unknown unknowns". The PBC and MCC represent snapshots in time and they need to be reassessed periodically as the project evolves. Given adequate risk management, the cost uncertainties should decrease and the MCC should track above or at the planned margin [Forsberg et al, 2000]. The MAIMS principle and probabilistic aspects of the project costs complicate the situation. As we have already shown, the TEC depends not only on the desired PoS but also the budget allocation and the management of contingencies. The project cost cannot be estimated until the cost management strategy including budget allocation is specified. We like to think that this contains a flavor of the Heisenberg Uncertainty Principle.

Much has been written on cost contingency; but there is still much confusion [Baccarini, 1999; INCOSE, 2003]. To shed some additional light on this topic, we express the MCC in a form that exhibits its dependence on the PoS and the cost management strategy,

$$\text{MCC}(\text{PoS}, \text{PBC}_1, \dots, \text{PBC}_n) \equiv \text{TEC}(\text{PoS}, \text{PBC}_1, \dots, \text{PBC}_n) - \text{PBC}. \quad (6)$$

PBC_i is the baseline budget for cost element C_i ; PBC is the sum over all cost elements. Equation (6) contrasts with both the deterministic practice that allocates a percentage of the PBC as MCC and today's typical PCA that models an "ideal project" devoid of the MAIMS principle. In contrast to Eq. (6), the MCC based on today's typical PCA is independent of the budget allocation strategy.

Consider the illustrative project in Section 6 and the following budget management strategies: (1) all cost elements are baselined at their mean values; (2) all cost elements are baselined at the 50% CL; and (3) all cost elements are baselined at the 75% CL. Figure 7 depicts the resulting TECs and MCCs and the "ideal" project TEC. The budget management strategy has a significant impact on the project cost and PoS. The cost penalty, either increased budget or reduced PoS, increases with increasing budget

⁶ There is no standard terminology and usage of terms and definitions vary widely with organizations.

allocation and is substantial for all but the very high PoS values. For example, dramatic cost reductions are achieved by allocating budgets to the cost elements at the 50% CL rather than the 75% CL. The standard PCA that assumes an “ideal” project provides a false sense of confidence and it may be a major source of cost overruns even for budgets with high contingencies.

Figure 7 has important implications for both the procuring activity and the contractor. Consider a hypothetical request for proposal for the project in Table IV. To level the playing field, the procuring activity specifies that all bids should provide the 50% CL cost. Contractor A has a certain level of sophistication. He prepares a PCA with every bid; he systematically assesses the cost elements including uncertainties; he baselines and allocates budgets to the cost elements at their mean values; management establishes and controls a contingency that equals the difference between the bid and the mean TEC. But Contractor A is not cognizant of the MAIMS principle and other psychological considerations that impact project cost. He performs the standard PCA and obtains the CDF in Fig. 7 labeled “TEC Ideal” and a P50 TEC of 7,348 K\$. The P50 value or median is 317 K\$ less than the mean value of 7,655 K\$ because the cost elements are given by positively skewed PDFs. To increase the chance of winning the project, Contractor A submits a bid of 7,348 K\$ and rationalizes that his practices are too conservative given that this value is 30% above the low estimate of 5,633 K\$. But because of the MAIMS principle Contractor A's risks are significantly greater than he thinks. Given that the cost elements are budgeted at their mean values, the TEC is really given by the CDF in Fig. 7 labeled “PEC MAIMS_@_mean”, the P50 TEC is 8,071 K\$, and the PBC of 7,665 K\$ is the lowest achievable cost. Based on our analysis we conclude that there is a negligible likelihood that Contractor A given his practices can deliver the project for the submitted bid of 7,348 K\$. The criticality of the situation is further aggravated by the fact that Contractor A has stumbled onto Russo and Schoemaker's [1990: 95] Decision Trap Number 5 [see Section 2.2.4]; he is implicitly relying on misleading information. Table VI summarizes this and several other scenarios.

8. SUMMARY AND RECOMMENDATIONS

This paper begins by exploring some of the limitations of today's typical probabilistic cost analysis that assumes an “ideal” world devoid of the psychological considerations and other complexities that are so critical to project success. It then proceeds to develop a practical and theoretically valid framework for quantifying the influences of human behavior, judgment under uncertainty, management practices and interrelationships among cost elements presents. The key elements include:

1. The use of an appropriate cost work breakdown structure. Subdividing the project costs into too many bite-size pieces is likely lead to erroneous results and false sense of confidence. Analysts should be wary of the pitfalls of performing a probabilistic cost analysis that consists of hundreds of cost elements that are subordinate to level 3.
2. The assessment of uncertain cost elements using the direct fractile assessment method and findings about human judgment. We recommend that analysts elicit uncertain cost elements at the 10th, 50th, and 90th percentiles and, if judged appropriate, calibrate them to account for human and organizational behaviors, be it overconfidence or cost/schedule padding.
3. The selection of flexible and realistic probability distribution functions. We favor the three-parameter Weibull distribution for fitting uncertain cost elements; but there are many other suitable distributions. The values of the input percentiles are much more likely to have a significant impact on the probabilistic cost analysis than the choice of the fitted distributions.
4. A realistic and practical treatment of correlation among cost elements. We propose a model for the full correlation matrix that depends on two parameters: ρ_{in} for correlation among cost elements in the same subsystem, and ρ_{ext} for correlation among cost elements in different subsystems. Reasonable correlation values in the range 0.3 to 0.6 should lead to more realistic cost estimates than the overly optimistic values assuming independence or the overly pessimistic values assuming perfect correlation
5. Incorporation of the “Money Allocated Is Money Spent” (MAIMS principle) and budget management practices. In the Monte Carlo simulation, all sampled values less than their associated cost baseline budgets are set equal to the latter.
6. The explicit treatment of confounding effects. The assessment of the cost elements, correlation effects, budget allocation, human behavior, and organizational considerations items all influence each other and have a significant impact on the project cost and/or probability of success. For enhanced credibility and realism, we recommend that the probabilistic cost analyses consider these influences simultaneously rather than individually.
7. The use of commercially available and widely used software. The proposed analysis is readily implemented using standard Monte Carlo simulation tools such as Crystal Ball® and @Risk.

We think that the proposed approach provides a framework for obtaining more accurate predictions than those provided by today's typical probabilistic cost analysis. With more accurate predictions and realistic expectations project managers can develop more viable

plans and make better decisions. The results are projects that are delivered for a lower cost and higher probability of success. We acknowledge that it takes effort to develop these more realistic models and that all models are only approximations to reality. But given the magnitude of the cost overrun problem, there is no excuse for accepting the status quo; the benefits are likely to be significant.

We, however, do not claim that the proposed approach is the silver bullet that will slay the cost overrun monster. We have focused only on cost and the macroscopic perspective. Cost is but one element of the performance-cost-schedule triad. By their very nature R&D and complex engineering projects are susceptible to high-consequence risks that are better modeled with decision trees, influence diagrams, and other decision tools. The ever present and more generic cost uncertainties are better and more conveniently modeled with continuous probability distribution functions. In the near future, probabilistic cost analysis should integrate the microscopic and macroscopic approaches to ensure that they properly address all risks and cost uncertainties and that they adequately support risk reduction activities. Other areas that need additional research and development include (1) eliciting and integrating data from multiple experts [Clemen and Winkler, 1999]; (2) budgeting and managing contingencies for multiple projects [Dillon and Paté-Cornell, 2001]; and (3) quantifying human and organizational behaviors of R&D and complex engineering projects. Our experience is that the single greatest challenge to the development and use of improved probabilistic cost analysis is the implementation of systems thinking [Richmond, 2000] at the personnel, organizational, and institutional levels. One of the anonymous referees has suggested the development of a tracking system that identifies the assumptions for the high, medium, and low percentiles and tracks their evolution. This would provide management with a powerful forecasting tool and the ability to dynamically adjust the budget of each cost element. He states " ...in the future -accurately watching and forecasting - can grow into the aspects of controlling and suppressing adverse behaviors." We agree and think that this paper offers a springboard to develop and implement more refined cost models and improved budget management practices.

ACRONYMS

CCPM	Critical Chain Project Management
CDF	Cumulative Distribution Function
CER	Cost Engineering Relationship
CL	Confidence Level
CWBS	Cost Work Breakdown Structure
DFA	Direct Fractile Assessment
MAIMS	Money Allocated Is Money Spent
MCC	Management Cost Contingency
MCS	Monte Carlo Simulation
PBC	Project Baseline Cost
PCA	Probabilistic Cost Analysis
PDF	Probability Distribution Function
PoS	Probability of Success
TEC	Total Estimated Cost
TLCM	Two-Level Correlation Model
WBS	Work Breakdown Structure

ACKNOWLEDGMENTS

The authors thank Paul Garvey (The MITRE Corporation) for his valuable comments and suggestions on the treatment of correlations among cost elements. They acknowledge the contributions of several anonymous referees to the improvement of this paper. Edouard Kujawski extends his special thanks to Dr. Richard Bloom for providing his professional expertise on human behavior.

REFERENCES

- S.M. AbouRizk and D.W. Halpin, Statistical Properties of Construction Duration Data, *Journal of Construction Engineering and Management* 118 (3) (1992), 525-544.
- M. Alpert and H. Raiffa, "A progress report on the training of probability assessors," in *Judgment under uncertainty: Heuristics and biases*, D. Kahneman, P. Slovic, and A. Tversky (Editors), Cambridge University Press, 1982, pp. 294-305.
- N. R. Augustine, *Augustine's Laws*, American Institute of Aeronautics and Astronautics, Reston, 1997.
- D.J. Baccarini, Cost Contingencies - A Review, *Australian Institute of Quantity Surveyors Refereed Journal* 2 (1) (1998), 8-15.
- P.L. Bernstein, *Against the Gods: The Remarkable Story of Risk*, Wiley, New York, 1996.
- S.A. Book, Why Correlations Matter in Cost Estimating, 32nd Annual DoD Cost Analysis Symposium, Williamsburg, 1999.
- S.A. Book, Estimating Probable System Cost, *Crosslink* 2 (1) (Winter 2000/2001), 12-21.
- F.B. Brooks, Jr., *The Mythical Man-Month*, Addison-Wesley, Reading, 1995.
- T.R. Browning and S.D. Eppinger, Modeling Impacts of Process Architecture on Cost and Schedule Risk in Product Development, *IEEE Transactions on Engineering Management* 49 (4) (2002), 428-442.
- C. Chapman and S. Ward, Estimation and Evaluation of Uncertainty: A Minimalist First Pass Approach, *International Journal of Project Management* 18 (2000), 369-383.
- R.T. Clemen and T. Reilly, Correlations and Copulas for Decision and Risk Analysis, *Management Science* 45 (2) (1999), 208-224.
- R.T. Clemen and R.L. Winkler, Combining Probability Distributions From Experts in Risk Analysis, *Risk Analysis* 19 (2), 1999.
- R.T. Clemen, G.W. Fischer, and R.L. Winkler, Assessing Dependence: Some Experimental Results, *Management Science* 46 (8) (2000), 1100-1115.
- R.T. Clemen and K.C. Lichtendahl, Jr., Debiasing Expert Overconfidence: A Bayesian Calibration Model, *PSAM6*, June 2002.
- Senator William Cohen, *Federal Agency Acquisition Management*, Hearing before the Senate Committee on Government Affairs 97, April 1982, p. 3.
- R.L. Coleman, J.R. Summerville, and M.E. Dameron, The Relationship Between Cost Growth and Schedule Growth, *Acquisition Review Quarterly* (Spring 2003), 117-122.
- E.H. Conrow, *Effective Risk Management: Some Keys to Success*, American Institute of Aeronautics and Astronautics, Reston, 2003.
- R.L. Dillon, R. John, and D. von Winterfeld, Assessment of Cost Uncertainties for Large Technology Projects: A Methodology and an Application, *Interfaces* 32 (4) (2002), 52-66.
- R.L. Dillon and M.E. Paté-Cornell, APRAM: An Advanced Programmatic Risk Analysis Method, *International Journal Technology Policy Management* 1 (1) (2001), 47-65.
- P. Embrechts, A. McNeil, and D. Straumann, Correlations: Pitfalls and Alternatives, *Risk* (May 1999), 67-71. <http://www.defaultrisk.com/pp%5Fcorr%5F08.htm>.
- EPA/630/R-97/001, *Guiding Principles for Monte Carlo Analysis*, Washington, DC, 1997.

J. Fente, K. Knutson, and C. Schexnayder, Defining a Beta Distribution for Construction Simulation, Proceedings of the 1999 Winter Simulation Conference, pp. 1010-1015.

K. Forsberg, H. Mooz, and H. Cotterman, Visualizing Project Management: a Model for Business and Technical Success, John Wiley & Sons, New York, 2000.

H. Forsey, "The mathematician's view: modeling uncertainty with the three-parameter lognormal," in Managing Downside Risk in Financial Markets, F.A. Sortino (Editor), Butterworth-Heinemann, 2001, pp. 51-58.

D. Freedman, R. Pisani, R. Purves, and A. Adhikari, Statistics, Norton & Company, New York, 1991.

P.R. Garvey, Do Not Use RANK CORRELATION in Cost Risk Analysis, 32nd Annual Department of Defense Cost Analysis Symposium, February 1999.

P.R. Garvey, Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective, Marcel Dekker, New York, 2000.

C.L. Givens Filiatrault and C.D. Peterson, Five Behaviors That Can Reduce Schedule Risk, Proceedings of the Project Management Institute Annual Seminars & Symposium, Houston, September 2000.

E.M. Goldratt, Critical Chain, The North River Press, Great Barrington, 1997.

C. Gordon, Risk Analysis and Cost and Cost Management (RACM): A Cost/Schedule Management Approach using Statistical Cost Control (SCC), 1997.
<http://www.acq.osd.mil/pm/paperpres/1097conf/gord/1097gord.htm>.

R. Graves, Open and Closed: The Monte Carlo Model, PM Network (December 2001), 48-53.

T.P. Hughes, Rescuing Prometheus, Random House, New York, 1998.

D.T. Hulett, Project Cost Risk Analysis Using Crystal Ball®, 1999.
<http://www.decisioneering.com/articles/hulett.html>

W. Herroelen, R. Leus, and E. Demeulemeester, Critical Chain Project Scheduling: Do Not Oversimplify, Project Management Journal 3 (4) (2002), 48-60.

INCOSE, Figuring Out Contingency Budgets, Internet thread, Risk Management Working Group, January 2003, <http://www.incose.org/rmwg/RMWG>.

R.L. Keeney and D. von Winterfeld, Eliciting Probabilities from Experts in Complex Technical Problems, IEEE Transactions on Engineering Management 38 (3) (1991), 191-201.

J.P. Kindinger, Use of probabilistic cost and schedule analysis results for project budgeting and contingency analysis at Los Alamos National Laboratory, Proceedings of the 30th Annual Project Management Institute 1999 Seminars & Symposium.

E. Kujawski, Why Projects Often Fail Even with High Cost-Contingencies, Systems Engineering 5 (2) (2002a), 151-155.

E. Kujawski, Selection of Technical Risk Responses for Efficient Contingencies, Systems Engineering 5 (3) (2002b), 194-212.

H.M. Markowitz, Portfolio Selection: Efficient Diversification of Investments, Blackwell, Cambridge, 1997.

N.J. McCormick, Reliability and Risk Analysis, Academic Press, New York, 1981.

MIL-HDBK-881, Handbook Work Breakdown Structure, US Department of Defense, 1998.

S.R. Moran, Cost-Risk Analysis - The Devil Is In The Details, Space Systems Cost Analysis Group, Los Angeles, January 1999.

NASA, Program and Project Management Processes and Requirements NPG: 7120.5B, 2002.

R. Newbold, Project Management in the Fast Lane: Applying the Theory of Constraints, St. Lucie Press, Boca Raton, 1998.

S. Pious, The Psychology of Judgment and Decision Making, McGraw-Hill, New York, 1993.

M. Rabin, Psychology and Economics, Journal of Economic Literature XXXVI (1998), 11-46.

B. Richmond, The "Thinking" in Systems Thinking: Seven Essential Skills, Pegasus Communications, Waltham, 2000.

J.E. Russo and P.J.H. Schoemaker, Decision Traps: The Ten Barriers to Brilliant Decision-Making and How to Overcome Them, Simon & Schuster, New York, 1990.

P. Sage, "Behavioral and Organizational Considerations in the Design of Information Systems and processes for Planning and Decision Support, IEEE Transactions on Systems, Man, and Cybernetics SMC-11 (9) (1981), 640-678.

E. Schragenheim and H.W. Dettmer, Does Your Internal Management Meet Expectations, CrossTalk 14 (4) (2001), 19-25.

J. R. Schuyler, Risk and Decision Analysis in Projects, Project Management Institute, Newton Square, 2001.

P.M. Senge, The Fifth Discipline: The Art & Practice of The Learning Organization, Doubleday, New York, 1994.

U.S. Department of Defense, Interim Defense Acquisition Guidebook (Formerly 5000.2-R), 2002.

U.S. Department of Energy, Project Management Practices, Practices 8 and 12, Draft, October 2000.

A. Smith and S-P. Hu, Impact of Correlating CER Risk Distributions On A "Realistic" Cost Model, ISPA/SCEA Conference, June 2003.

P.G. Smith and G.M. Merritt, Proactive Risk Management: Controlling Uncertainty in Product Development, Productivity Press, New York, 2002.

E.D. Walker II and J. F. Cox III, An Analysis of the Theory and Practice of Project Management in Order to Identify the Causes of Failures of Projects, Submitted to the International Journal of Operations and Production Management (2003).

Table I. Some Summary Statistics for the PDFs in Figure 1a

Statistics	$\rho = 1.0$	$\rho = 0.8$	$\rho = 0.5$	$\rho = 0.3$	$\rho = 0.1$	$\rho = 0.0$
Mean, K\$	4,326	4,326	4,326	4,326	4,326	4,326
Median, K\$	4,210	4,230	4,266	4,263	4,307	4,318
Standard Deviation, K\$	481	433	350	350	206	153
Skewness	1.31	1.22	1.00	1.03	0.56	0.38
Kurtosis	5.13	4.85	4.27	4.45	3.30	3.06

Table II. Parameters and some statistics for different PDFs fitted to three percentiles

Distribution (fractiles*)	Fitted Parameters ⁺				Mean K\$	SD [^] K\$
Weibull(10/50/90)	L = 370 K\$	$\alpha = 67$ K\$	$\beta = 1.29$		432	49
Weibull(20/50/80)	L = 364 K\$	$\alpha = 83$ K\$	$\beta = 0.97$		448	85
Beta(10/50/90)	a = 375 K\$	b = 607 K\$	$\alpha = 0.94$	$\beta = 2.91$	432	45
Beta(20/50/80)	a = 375 K\$	b = 584 K\$	$\alpha = 0.49$	$\beta = 1.10$	439	60

* 10/50/90: $x_{10} = 382$ K\$, $x_{50} = 421$ K\$, $x_{90} = 499$ K\$

20/50/80: $x_{20} = 382$ K\$, $x_{50} = 421$ K\$, $x_{80} = 499$ K\$.

⁺ Parameters defined in text.

[^] SD: Standard deviation.

Table III. Means and standard deviations for an illustrative cost element with and without the MAIMS principle

Probability Distribution [^]	Mean K\$	Percentile of mean	SD K\$
W($x_{20} = 382, x_{50} = 421, x_{80} = 499$)	448	63	85
WM($x_{20} = 382, x_{50} = 421, x_{80} = 499; x^* = \text{mean} = 448$)	479	75	66
WM($x_{20} = 382, x_{50} = 421, x_{80} = 499; x^* = x_{50} = 422$)	463	72	73
WM($x_{20} = 382, x_{50} = 421, x_{80} = 499; x^* = x_{75} = 482$)	502	81	55
WM($x_{20} = 382, x_{50} = 421, x_{80} = 499; x^* = x_{85} = 521$)	535	87	48

[^] W () is the three-parameter Weibull distribution in Table II. The WMs are the corresponding MAIMS-modified distributions with the specified values of x^* . All percentiles are in units of K\$.

Table IV. Sample project WBS and assessed cost elements

WBS Cost Elements ^	Estimated Percentiles K\$		
	X ₁₀	X ₅₀	X ₉₀
1.0 Total project/system, C_T			
1.1 Project/system-level, C₁			
1.1.1 Project management, C _{1,1}	382	421	499
1.1.2 Systems engineering, C _{1,2}	220	232	257
1.1.3 Integration & test, C _{1,3}	887	1,010	1,256
1.2 Subsystem X, C₂			
1.2.1 Mechanical components, C _{2,1}	970	1088	1,323
1.2.2 Electrical components, C _{2,2}	742	846	1,054
1.2.3 Integration & test, C _{2,3}	596	724	980
1.3 Subsystem Y, C₃			
1.3.1 Software development, C _{3,1}	1,069	1,282	1,708
1.3.2 Firmware, C _{3,2}	634	743	961
1.3.3 Integration & test, C _{3,3}	541	656	886

^ C_{m,j} is defined in Section 3.1.

Table V. Some summary data of the PCA for the project in Table IV

Model^		Estimated Project Cost*, K\$					
$\rho_{int}/\rho_{ext}^{\wedge}$	Budget Allocation ⁺	Mean	SD	P0 Min value	P50	P80	P95
0.0/0.0	Ideal	7,676	785	5,633	7,576	8,295	9,114
0.6/0.4	Ideal	7,661	1,367	5,633	7,348	8,626	10,301
0.6/0.4	X50	8,059	1,145	7,002	7,692	8,771	10,427
0.6/0.4	Mean	8,438	1,010	7,665	8,071	8,987	10,494

^ Cost elements with 10th and 20th percentile values calibrated to the 20th and 80th percentiles, respectively.

ρ_{int}/ρ_{ext} : TLCM coefficients.

Ideal: Today's typical analysis; no MAIMS principle.

X50: MAIMS principle with cost elements allocated budgets at the 50th percentile.

Mean: MAIMS principle with cost elements allocated budgets at the mean.

* SD: Standard deviation.

Pn: Value with n% probability of success or confidence level.

Table VI. Some summary data of the different cost management strategies depicted in Figure 7

Management		Strategy		MAIMS-Modified PCA				Typical PCA			
Budget Allocation	Desired PoS	TEC \$K	MCC \$K	MCC %	Real. PoS	TEC \$K	MCC \$K	MCC %	Real. PoS		
Mean	20%	7,673	0	0%	20%	6,445	-1,220	-16%	0%		
	50%	8,071	406	5%	50%	7,348	-317	-4%	0%		
	80%	8,987	1,322	17%	80%	8,626	961	13%	73%		
50% CL	20%	7,111	0	0%	20%	6,445	-557	-8%	0%		
	50%	7,692	690	10%	50%	7,348	346	5%	37%		
	80%	8,771	1,769	25%	80%	8,626	1,624	23%	77%		
75% CL	20%	8,466	0	0%	20%	6,445	-2,021	-24%	0%		
	50%	8,613	147	2%	50%	7,348	-1,118	-13%	0%		
	80%	9,330	864	10%	80%	8,626	160	2%	52%		

LIST OF FIGURE CAPTIONS

Figure 1a. Illustrative example of the impact of correlations on the probability distribution of the sum of ten cost elements. Each cost element is given by a three-parameter Weibull distribution with location $L = 370$ K\$, scale parameter $\alpha = 67.5$ K\$, and shape parameter $\beta = 1.3$. The distribution and notation are defined in Section 4.2. The units of K\$ are introduced for additional relevance to PCA.

Figure 1b. Comparison of the impact of correlations on the cumulative distribution functions corresponding to the data in Figure 1a.

Figure 1c. Quantitative comparison of the impact of correlations for the illustrative example in Figure 1b.

Figure 2a. Illustrative example of the differences between the fits provided by the three-parameter Weibull and generalized Beta distributions to the two sets of three percentiles in Table II. W10/50/90: Weibull(10/50/90), B10/50/90: Beta(10/50/90), W20/50/80: Weibull(20/50/80), B20/50/80: Beta(20/50/80).

Figure 2b. Comparison of the cumulative distribution functions corresponding to the probability distributions in Figure 2a.

Figure 3. Comparison of CDFs with and without the MAIMS principle for the PDFs in Table III. W: original PDF; WM_@_mean: allocated budget $x^* = \text{mean}$; WM_@_X50: $x^* = x_{50}$; WM_@_X75: $x^* = x_{75}$; WM_@_X85: $x^* = x_{85}$.

Figure 4. Impact of different budget allocation strategies on the PCA for the project in Table IV. Cost elements with Weibull distributions fitted to the 10th, 50th, and 90th fractiles; TLCM parameter values of 0.6 and 0.4.

Figure 5. Impact of different correlation models on the PCA for the project in Table IV. Cost elements with Weibull distributions fitted to the 10th, 50th, and 90th fractiles.

Figure 6. Impact of calibrating cost elements on the PCA for the project in Table IV. The 10th and 90th fractiles are equated to the 20th and 80th fractiles and fitted with Weibull distributions; TLCM parameter values of 0.6 and 0.4.

Figure 7. Impact of different cost management strategies on the cost and contingency for the project in Table IV. This is a different representation of the data in Figure 4.

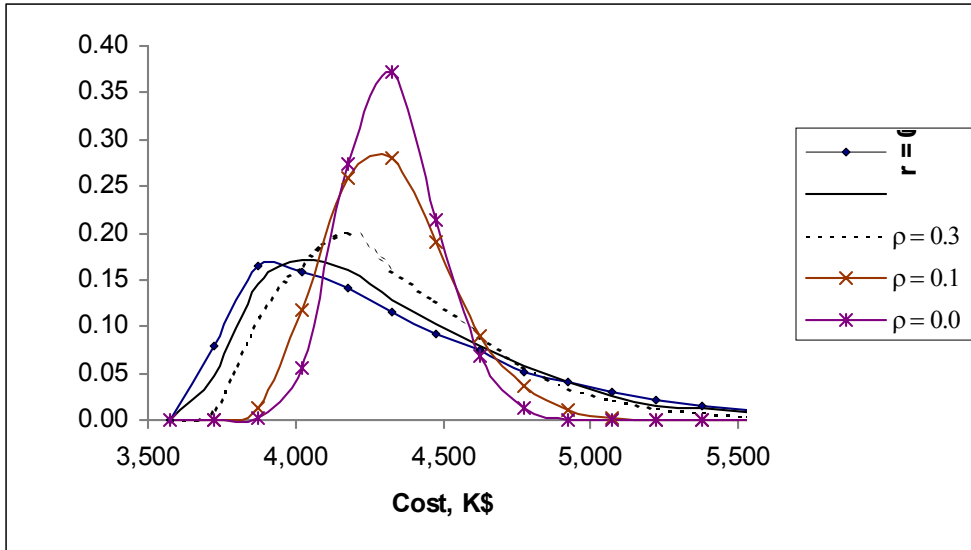


Figure 1a. Illustrative example of the impact of correlations on the probability distribution of the sum of ten cost elements. Each cost element is given by a three-parameter Weibull distribution with location $L = 370$ K\$, scale parameter $\alpha = 67.5$ K\$, and shape parameter $\beta = 1.3$.

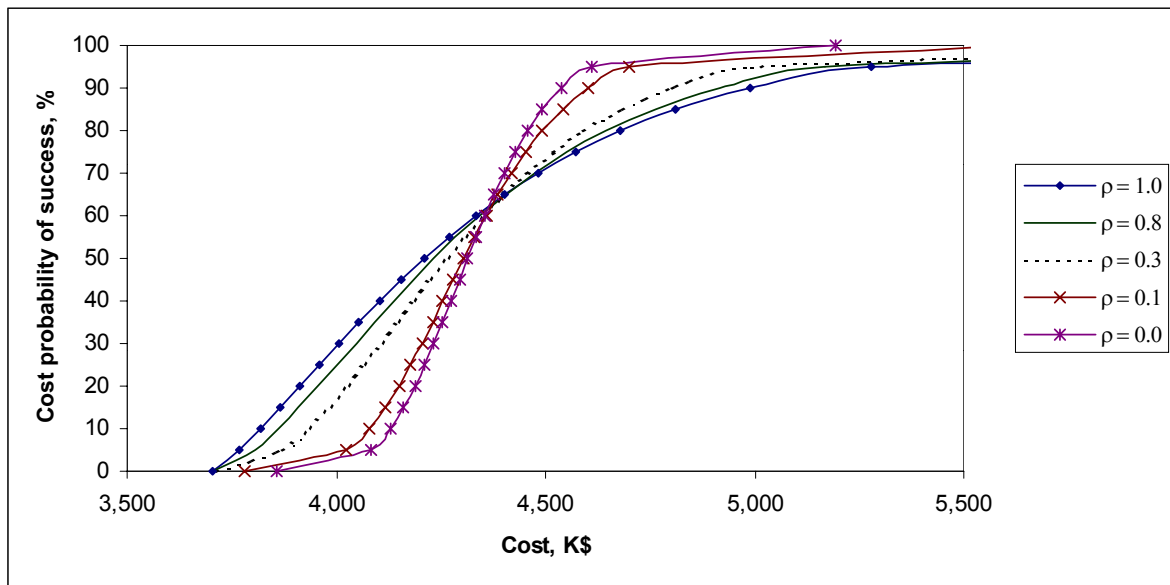


Figure 1b. Comparison of the impact of correlations on the cumulative distribution functions corresponding to the data in Figure 1a.

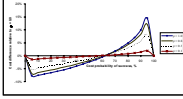


Figure 1c. Quantitative comparison of the impact of correlations for the illustrative example in Figure 1b.

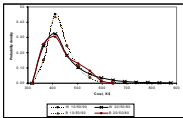


Figure 2a. Illustrative example of the differences between the fits provided by the three-parameter Weibull and generalized Beta distributions to the two sets of three percentiles in Table II. W10/50/90: Weibull(10/50/90), B10/50/90: Beta(10/50/90), W20/50/80: Weibull(20/50/80), B20/50/80: Beta(20/50/80).

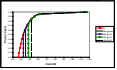


Figure 2b. Comparison of the cumulative distribution functions corresponding to the probability distributions in Figure 2a.

Figure 3. Comparison of CDFs with and without the MAIMS principle for the PDFs in Table III. W: original PDF; WM_@_mean: allocated budget $x^* = \text{mean}$; WM_@_X50: $x^* = x_{50}$; WM_@_X75: $x^* = x_{75}$; WM_@_X85: $x^* = x_{85}$.

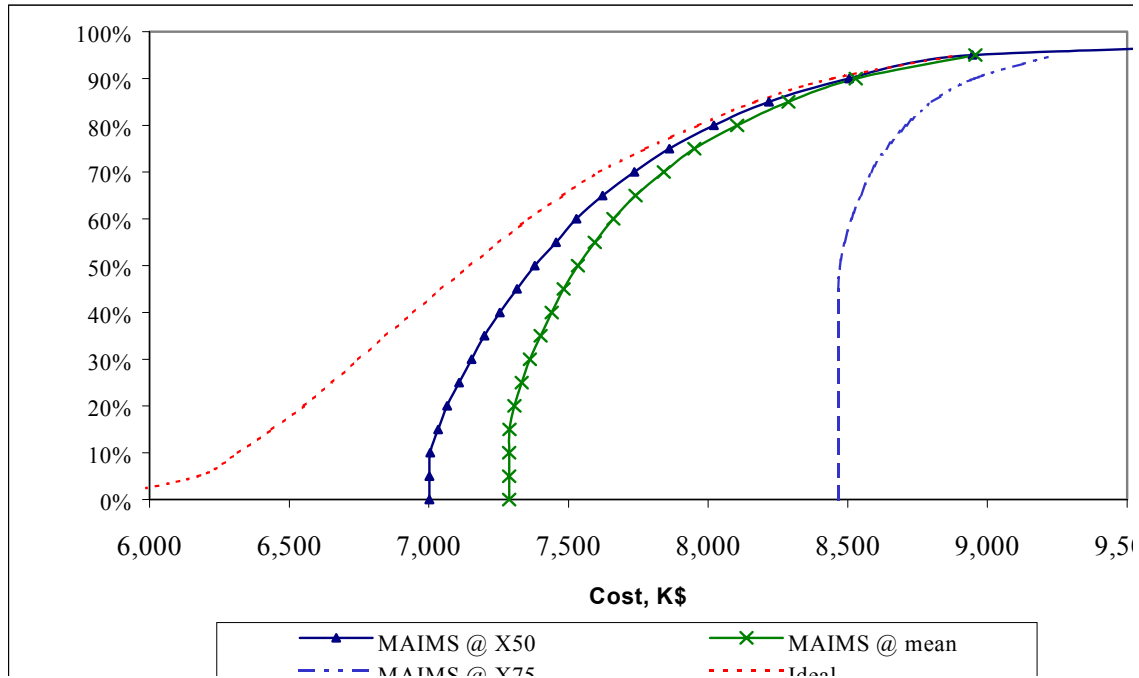


Figure 4. Impact of different budget allocation strategies on the PCA for the project in Table IV. Cost elements with Weibull distributions fitted to the 10th, 50th, and 90th fractiles; TLM parameter values of 0.6 and 0.4.

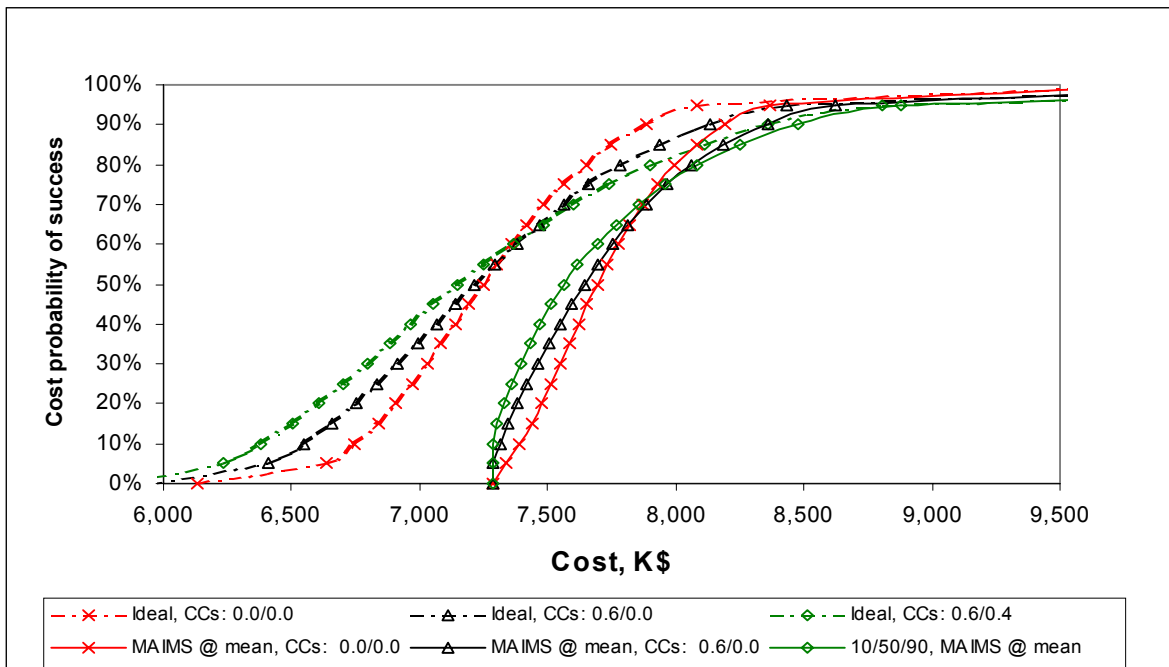


Figure 5. Impact of different correlation models on the PCA for the project in Table IV. Cost elements with Weibull distributions fitted to the 10th, 50th, and 90th fractiles.

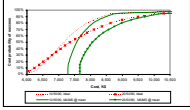


Figure 6. Impact of calibrating cost elements on the PCA for the project in Table IV. The 10th and 90th fractiles are equated to the 20th and 80th fractiles and fitted with Weibull distributions; TLM parameter values of 0.6 and 0.4.

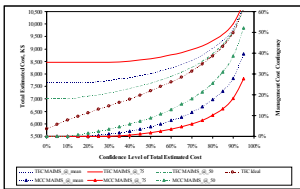


Figure 7. Impact of different cost management strategies on the cost and contingency for the project in Table IV. This is a different representation of the data in Figure 4.