Cost Modeling for Architectural Comm Trade Studies

Bruce E. MacNeal[1] and Kar-Ming Cheung[2]
Jet Propulsion Laboratory[3], California Institute of Technology, Pasadena, California, 91109

Architecture trade studies require that multiple cost models from various options be compiled and compared to determine relative cost advantages. The time-evolution of full life-cycle costs through implementation and maintenance/operations phases must be modeled. The effect of risk on cost must also be taken into account. Finally, the statistical significance of cost differences must be assessed. This paper describes the methods used for a recent study of future architectures for NASA's communications network. Typical results are presented.

Time-evolution is implemented by creating a separate WBS for each time period in the study, and then inserting/removing cost elements at appropriate times. Cost roll-ups at any level can then be shown as a function of time. Costs for different options are created by selecting and scaling cost elements from a common library. Cost elements can have up to four different types of cost: implementation, maintenance, sustainment and operations. Cost types are treated differently in time according to user defined models. Each cost in each element is assigned a mean and variance. Random analysis is then used to determine cost distributions, while the Kolmogorov-Smirnov Test is used to determine whether similar distributions for different options are statistically distinct. Risk is taken into account by defining risk events applied to individual cost elements or to WBS elements at particular times. Depending on the likelihood and severity of the risks, and where they occur in the WBS, cost distributions can be far from Gaussian. Models can be complex, so model debugging tools are also supplied. The implementation in object-oriented visual basic for MS EXCEL is entirely data-driven and is easily portable.

Nomenclature

p	= risk element probability of occurrence
k	= risk element cost increase
f	= fraction risk element cost increase
δ	= sampling function (Dirac Delta Function)
σ	= cost variance
m	= mean cost

I. Introduction

Architectural studies impose unique requirements on cost modeling. Normal costing (for proposal purposes, for example) must produce a total cost for a single point design with as little cost uncertainty as possible. Great care is taken to provide the best available estimates through extensive "bottom-up" exercises and cost modeling, while using institutionally-approved rates and factors. Such costing exercises can be expensive and time consuming. Cost schedules are important, but full life-cycle costs are often omitted. The concern is to commit the organization to perform work at the right price so these costing exercises can be expensive and time consuming.

---

[1] Program System Engineer, Telecommunications Architectures, 238-411.
[2] Group Supervisor, Telecommunications Architectures, 238-411.
[3] This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

In contrast, *architectural cost modeling* must address multiple, often immature designs with as quickly and inexpensively as possible. Since architectural *options* represent a process as well as a result, cost evolution across the entire life-cycle is important. As an aid to decision-making, it is also important to understand cost uncertainty, the effects of risk, and the statistical significance of cost differences between architectural options.
Costs associated with *risk* can be decisive in selecting options, so the cost impact of risk must be assessed quantitatively[4]. Normal statistical treatments of risk deals with expanded cost (and perhaps schedule) distributions for cost items[5]. We treat discrete risk events applied either to cost items or groups of cost in a manner similar to Reference [3]. The probability distribution function for a risk event is derived and shown to be consistent with random analysis results. In a practical example taken from a recent NASA study on communications architectures, it is shown that risk can have strong and unexpected effects on cost distributions.

II. Implementation

Architectural cost modeling was implemented using the Visual Basic programming language available in every installation of Microsoft EXCEL®. This software application is widely available and provides a powerful, object-oriented implementation of Visual Basic that is tightly connected with native EXCEL data objects (sheets, charts, etc.). User-defined data objects were used heavily, e.g., to implement WBS node objects, cost element objects and risk event objects. This approach is more efficient that using spreadsheet functions, and provides structured, self-documenting software in which the order of execution can be tightly controlled. Being native to EXCEL, the software is highly portable.

A. Data Structure

For architectural studies, it is import that the program be flexible and reusable in order to treat a wide and evolving range of cost options without change the underlying software (see Figure 1). The normal *Work Breakdown Structure* (WBS) is adopted as a scheme for categorizing cost, a method for accumulating and reporting costs, and a method for generating new costs based on accumulated totals. The structure is implemented (a separate WBS for each year of the study[6]) as interconnected instantiations of a single WBS node object. Each node object contains a descriptive title and up to six levels of indices used to specify a connected tree structure. Methods implement roll-up and random analysis functions.

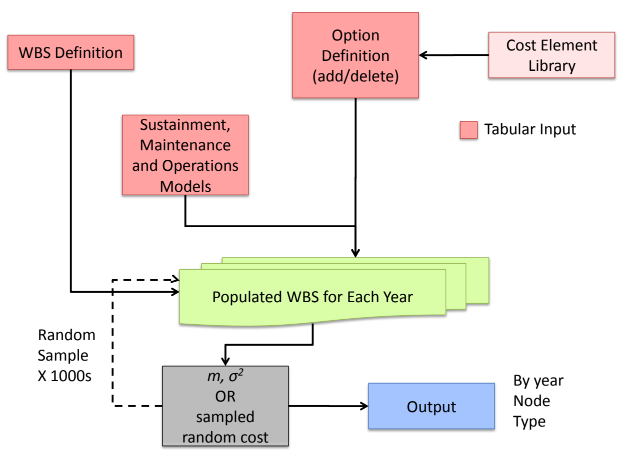

Figure 1: Schematic diagram of data flow.

Costs elements represent individual sources of cost at the lowest level of granularity throughout the entire cost life-cycle. All cost elements for the entire study are kept in a single *cost element library* (a single input table). Each element may have up to four cost types: implementation, sustainment, maintenance and operations (see Figure 2). Implementation costs represent the procurement or labor costs needed to, e.g., acquire equipment, develop software or manage projects. These costs are normally spread evenly across the implementation phase of the option[7]. Sustainment costs represent the cost of labor and parts needed to keep equipment and software running properly. Custom sustainment models are applied over the practical life of the equipment or software. Maintenance refers to the cost of replacing the equipment or software once it has exceeded its useful life. These costs are incurred at the end of a sustainment cycle.

[4] See Reference [1]
[5] See Reference [2]
[6] Annual WBS structures are normally used, but monthly or even daily WBS's are also possible, depending on the time scale of interest.
[7] In principal, implementation costs may be spread unevenly according to some desired time distribution, but such detailed information is usually not available in architectural studies.

After maintenance, a new sustainment cycle begins. Finally, operations costs are the labor costs associated with operating equipment or software. A frequently, options are designed to reduce operations costs by adopting new equipment, software and processes. Each type of cost may have its own cost uncertainty and an implementation risk event (see below).

The biases and uncertainties of cost estimation are a subject for study in themselves[8]. However, architectural cost studies seldom have the time or resources for systematic estimation of uncertainty. Instead, a blanket 30% uncertainty was assumed for the study reported here. The much more significant uncertainties due to identifiable risk are handled separately (see below).

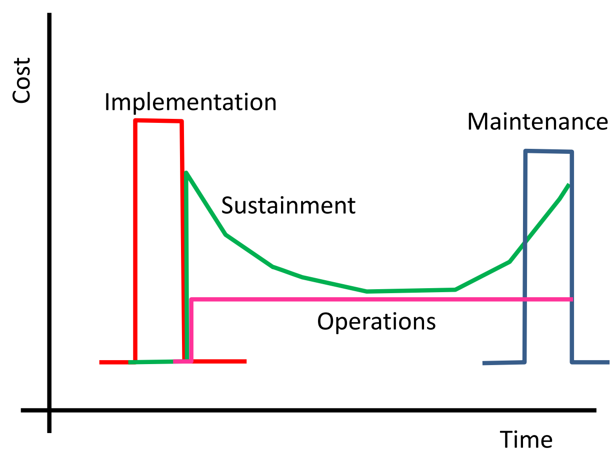

Figure 2: Schematic diagram of cost models for the four different cost types.

Architectural options are created by inserting cost elements from the library into the separate annual WBS structures. Distinct options are created by adding and/or deleting different numbers of cost elements from the library into nodes of the WBS structures for different years of the study (see Figure 1). The year an element is inserted defines the beginning of the implementation cost for the element. Further implementation, sustainment, maintenance and operations costs are automatically defined in subsequent WBS's. Removing an element ends all costs in the given year. Each element may also be giving independent data *types* in four separate categories (independent of the WBS node). These are used, e.g., to distinguish the source of the cost data, the physical location of components, or the managing organization.

Nodal Cost Elements are a special element used to create costs based on the accumulated costs at another node. Nodal elements represent an implementation costs[9] inserted at a particular node, calculated as a specified fraction of the accumulated implementation costs in some other node in the same year. Nodal costs last for a specified number of years. To avoid an infinite loop a nodal cost element may not be inserted in the structure if it is based on costs from a node above it in the same branch of the WBS. Nodal elements are useful for generating management costs or system engineering costs as a fixed fraction of implementation costs.

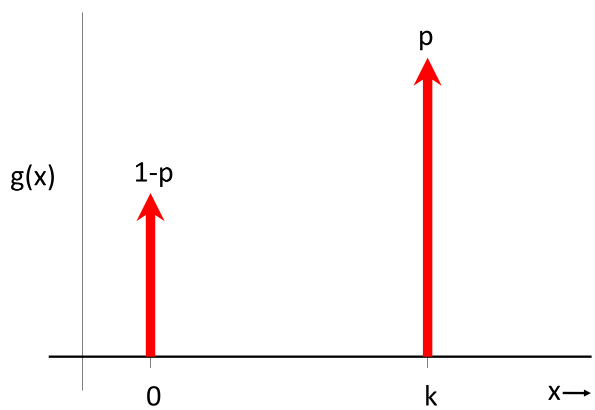

Figure 3: The cost distribution function for a risk event

B. Output and Debugging

The usual cost roll-up functions are provided. Mean cost *m* and variance σ^2 may be interrogated at any node of the WBS in any year. Cost and variance may also be computed for any node as a function of time. Cost roll-ups filtered by data type are a useful debugging tool, as are cost element counts across all options. Complete listings of WBS content by year are also useful, but voluminous.

The accumulated mean cost and variance are a true representation of the cost distribution at any node, regardless of the shape of the cost distribution. But, it is usually necessary to see the actual distribution in order to assess differences between options. The distribution is constructed using the same roll-up functions used for mean cost and

8 See Reference [4]

9 This concept could be extended to include sustainment, maintenance and operations costs when the need occurs.

variance. A Gaussian-distributed[10] random cost is generated based on the m and σ^2 for each element. The roll-up is then repeated thousands of time to generate a histogram of cost at any node, in any year.

C. Risk Events

A risk event is some potential occurrence outside of normal processes that could affect the cost of an option. The effect of risk events is captured in element or nodal *risk elements*. Risk elements may describe a perceived risk in acquiring a single component, or the risk in generating a particular piece of software. They may also be used to describe the programmatic risk associated with a block of costs at a WBS node. At the moment, risk elements are only applied to implementation costs, but extension to sustainment, maintenance and operations is planned. The effect of risk on schedule is not treated here.

Assume that a risk event has a probability of occurrence p. If the event occurs, the cost of the element or node will increase by an amount k. The normalized probability function for the risk event is given by:

$$g(x) = (1 - p)\,\delta(x) + p\delta(x - k) \tag{1}$$

where x is the cost and $\delta(x)$, the impulse function, is the first derivative of the unit step function (or Heaviside Step Function). The schematic diagram of g(x) shown in Figure 3 consists of two infinite "spikes", one at $x = 0$ and the other at x=k. The area under the first spike is 1-p, while the area under the second is p.

The mean cost of this risk element is:

$$\langle \mathrm{x} \rangle = \int_{-\infty}^{\infty} dx\, x\, g(x) = kp \tag{2}$$

The variance of the risk element about the mean is:

$$\langle (x - kp)^2 \rangle = \int_{-\infty}^{\infty} dx\, (x - kp)^2\, g(x) = k^2 p(1 - p) \tag{3}$$

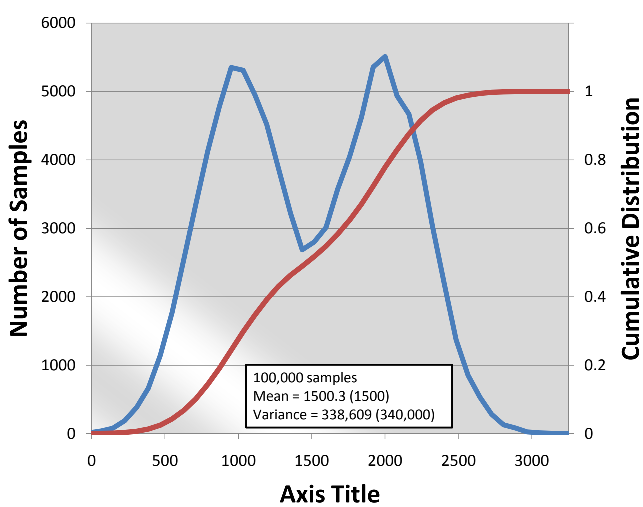

Figure 4: Cost distribution and cumulative distribution for a single cost element with a risk event

If the cost of the risk element is added to the cost of the associated component (or node) then the cost distribution is the convolution of the distributions for the two separate costs. The mean cost of the combination is $m + pk$ and the variance is $\sigma^2 + k^2 p(1 - p)$. It is important to note that the distribution of the combined costs may not be Gaussian even in the limit of adding together a large number of costs. The result depends on the size of the risk k and its location in the WBS structure. The cost produced by a risk operating on a single element or small amount of accumulated cost by appear Gaussian. The cost produced by a severe risk operating on a large amount of accumulated cost will tend to be bimodal.

Figure 4 shows an example of cost for single cost element with a single risk event. The cost of the element in this illustration is 1000 with an uncertainty of 30%. The risk event has a 50% probability of increasing the cost by 100%. The distribution function shown is the convolution of a Gaussian distribution with a mean of 1000 and variance of 90,000 with the distribution function of Figure 4 for k = 1 and p = 0.5. The mean and

[10] If more information on uncertainty is available, triangle or more complex distributions may be used.

variance of the sampled distribution for 100,000 samples are both within 0.5% of the theoretical values (in parentheses).

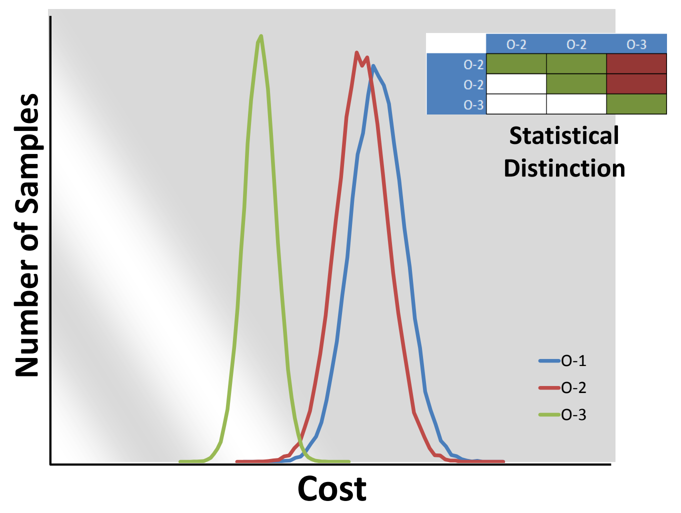

Figure 5: Cost distribution for architectural options with risk events

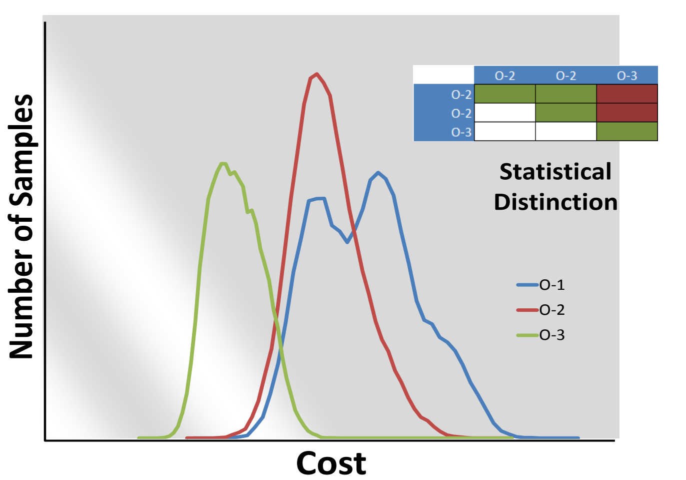

Figure 6: Cost distributions for options with risk events applied

In NASA, risk events are usually assigned a position on a 5x5 matrix[11]. The aces, labeled *likelihood* and *consequence*, are each assigned an integer value (1 through 5), which places the event in one cell of the matrix. For quantitative work, each likelihood integer is assigned a specific probability *p*, and each consequence integer is assigned a specific fractional cost increase *f*, $k = (\text{cost})\ x\ f$. Alternatively, the values of *p* and *f* can be specified directly for each event. For risk assessment/management of mature systems, *p* is 10%-30% and *f* is a few percent. For architectural studies, where large risks have not yet been mitigated, *p* can be over 50% and *f* can be well over 100%.

In order to make better decisions based on cost, it is important to know whether cost distributions for different options are really statistically distinct. The *Kolmogorov-Smirnov Test*[12] [6] is used here to assess the statistical distinctness of empirical cost distributions. A rigorous discussion of this test is beyond the scope of this paper and will be published at some later time. The result of the test is a quantitative score that is <u>independent</u> of the shape of the two empirical distributions being compared. Assuming that the two distributions are drawn from the same (unknown) cost distribution, the known theoretical distributions of scores is used to determine the fraction of comparisons that would result in higher scores. To the extent that this fraction is small, the hypothesis (that they are drawn from the same underlying distribution) may be rejected.

III. Example: Communications Architecture Cost Study

The methods described were applied in a recent study of communication architectures for NASA's Space Communications and Navigation directorate (SCAN). The study involved three architectural options, designated here as O-1, 2 and 3.

Figure 5 shows the distribution of cost for the three options before risk events are applied. The costs for O-1 and O-2 are nearly the same while the cost for O-3 is considerably lower. The colored inset shows the results of the KS test applied to the three options. Green cells indicate that the two options identified in the row and column should be considered to represent the same cost; they are not statistically distinct. The red cells indicate that the cost for O-3 is statistically distinct from the other two[13].

---

[11] See Reference [5]

[12] See Reference [6]

[13] The green cells show that the two empirical distributions produced a KS score that is exceeded by more than 5% of empirical distributions so the hypothesis should not be rejected. The red cells show that just 1% or less of empirical distributions would produce a larger KS score. Most likely they do not come from the same distribution, and should be considered distinct.

Risks were identified for each option separately, and applied to appropriate nodes in the WBS. Figure 6 shows the distribution of cost for the three options with risk included. The effect on distribution shape is significant, particularly for O-1, where the distribution is essentially bimodal. This results from a significant risk event being applied high in the WBS, operating on a significant fraction of the total option cost. The other two options are Gaussian-like, but skewed toward higher costs; the effect of risks located farther down in the WBS. The inset shows that even with the greater overlap, the cost for O-3 should be treated as statistically distinct from the other two.

IV. Conclusion

We have demonstrated a practical analysis tool that meets the unique requirements of architectural cost modeling. Early studies of immature design options require a simple and flexible tool to accommodate rapid design changes. The tool is implemented using the objected-oriented Visual Basic programming environment provided in EXCEL®. Cost models for multiple options are constructed using cost elements from a common library. Mean costs and variances are accumulated according to a user-defined WBS, and may be filtered by node, data type or year. Random methods are used to determine actual cost distributions. An empirical test for the distinctness of cost distributions acts as a guide to decision-making. The use of risk elements in a study of communications architectures reveals their potentially dominant effect on cost distributions.

Appendix B
Glossary

Term	Description
Architectural Option	One of several proposed methods for reaching a goal. Costs for each option are to be compared.
Consequence	The fractional increase in cost associated with the occurrence of a risk event
Cost Elements	A structure used to generate a cost at the lowest level of granularity.
Kolmogorov-Smirnov Test	A statistical test for the distinctness of two empirical distributions
Likelihood	The probability that a risk event will occur
Risk Element	A structure used to generate a cost associated with a risk event, based either on a single cost element, or on the costs associated with a WBS accumulation node.
SCAN	NASA's Space Communications and Navigation directorate
Visual Basic	Programming language supplied in Microsoft EXCEL® for custom applications
Work Breakdown Structure	A tree structure defining the relationship between cost components. A template for cost accumulation.

Acknowledgments

We thank those involved in the cost study for their important contributions to this work: Wallace S. Tai (JPL), Nathaniel Wright (NASA), Richard R. Kunath(GRC), Michael L. Anderson(GSC), Steve Kowal (Honeywell), Michael Prior (ITT), and Peter M. Shames(JPL)

References

[1]Galway, Lionel, “Quantitative Risk Analysis for Project Management”, RAND Corporation, http://sbir.nasa.gov/SBIR/successes/ss/10-018text.html

[2]Thompson, C. M., Smith, Alfred, Cyrulik, Melissa A., “How to Capture Discrete Cost Risks in Your Project Cost Model,” *ACEIT User’s Conference*, Santa Barbara, California, 2009. https://www.sceaonline.org/awards/papers/2008_risk.pdf

[3] Monica, Doyle, Feingold, Harvey, Hayes, Shawn and Jacobs, Mark, “Integrated Cost and Schedule Risk Analysis”, *NASA Cost Symposium*, Kennedy Space Center, April 30, 2009.

[4]Galway, Lionel, “Subjective Probability Distribution Elicitation in Cost Risk Analysis: A Review”, RAND Corporation, http://www.rand.org/content/dam/rand/pubs/technical_reports/2007/RAND_TR410.pdf.

[5]*NASA Risk Management Handbook*, National Aeronautics and Space Administration, NASA Headquarters, Washington, D.C. 20546, November 2011, p 145. http://www.hq.nasa.gov/office/codeq/doctree/NHBK_2011_3422.pdf

[6]Simard, Richard and L’Ecuyer, Pierre, “Computing the Two-Sided Kolmogorov-Smirnov Distribution”, *Journal of Statistical Software*, Vol. 39, No. 11, March, 2011.