REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 1995	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
A Fuzzy Approach to Quantification of Technical and Schedule Uncertainty In DoD Weapon System Cost Estimates

5. FUNDING NUMBERS

6. AUTHOR(S)
H. Samuel Cooke, D.Sc.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

19980805 051

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED

12b. DISTRIBUTION CODE

13 ABSTRACT *(Maximum 200 words)*

This research describes a methodology and provides an analysis for quantifying technical and schedule uncertainty utilizing elements of fuzzy logic. Application of this methodology is context specific, related to a particular fielded missile system. Data is characterized using triangular and trapezoidal functions clearly depicting overlapping ranges of categories into classes of fuzzy membership sets. Methodology presented in the research is sufficiently robust to accommodate either linguistic or quantitative input, process the input through a fuzzy algorithm and defuzzify output of the algorithm into a "crisp" solution. A rule base, in the form of a look-up table matrix, is presented for combining technical and schedule uncertainty.

The author may be contacted for further information/application: P.O. Box 22616 Huntsville, Alabama 35814, 256-895-8161

14. SUBJECT TERMS	15. NUMBER OF PAGES 141
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)

A FUZZY APPROACH TO QUANTIFICATION OF TECHNICAL AND SCHEDULE UNCERTAINTY IN DOD WEAPON SYSTEMS COST ESTIMATES

by

Henry Samuel Cooke, Jr.

A Dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Science

Southeastern Institute of Technology

November 1995

APPROVALS:

Research Committee:

Julian S. Kobler, D.Sc.

Charles A. Cockrell, D.P.A.

Raymond C. Watson, Jr., Ph.D.

President:

Raymond C. Watson, Jr.

ABSTRACT

Cost estimating, within the Department of Defense for acquisition of weapon systems, is a target of criticism. In particular, quantification of uncertainty in cost estimates is a subject of concern at the highest levels of review. Often, due to time constraints associated with developing an estimate to accompany a procurement package through the acquisition process, the cost analysis community reverts to giving less than adequate attention to quantification of technical and schedule uncertainty.

This research describes a methodology and provides an analysis for the quantification of technical and schedule uncertainty using two elements of fuzzy logic, linguistic variables and fuzzy membership sets. Application of the methodology presented is context specific, related to a particular fielded weapon system. Foundation for the methodology rests on a survey of technical and schedule uncertainty that requires input, in the form of opinions of percent ranges of uncertainty, for five categories; very low, low, medium, high and very high. Thirty professional personnel supporting the particular weapon system office completed all sections of the survey for each of the five linguistic categories for both technical and schedule uncertainty. Survey input is characterized using triangular and trapezoidal functions clearly depicting overlapping ranges of each category into classes of fuzzy membership sets. A copy of the survey instrument is provided for completeness.

A rule base, in the form of a look-up table matrix, is presented for combining technical and schedule uncertainty. Methodology presented in the research is robust enough to accommodate either linguistic or quantitative input, process the input through a fuzzy algorithm and defuzzify output of the algorithm to a "crisp" solution. QuickBASIC computer programs are documented for insight into precisely how the degrees of membership for each fuzzy set are generated. Pattern search algorithms for specifying triangular and trapezoidal parameters for each fuzzy membership set are presented.

Basic descriptive statistics, distribution functions, analysis of variance and paired t tests are presented to support testing of several hypotheses related to actual cost estimate and contract price data. The conclusion of the research is that a fuzzy approach to quantification of technical and schedule uncertainty can improve a cost estimate even if applied just at a top level.

ACKNOWLEDGMENTS

This Dissertation is the result of seven and one half years of academic effort. Other critically formative years however provided a foundation that enabled the courage and discipline for initiation and completion of this work. That foundation was laid by people who cared about me and who continue to have a most positive and pervasive influence on my life: Honey and Cookie, Jay and Dale Thacker, and Ronald Arthur Edwards.

Honey and Cookie had a dream; a college education for their son. Somehow in the challenges of life they steadfastly maintained a focus on that dream and successfully transferred it into a goal that I ultimately embraced. In retrospect, there are at least two elements of their dream to be greatly appreciated; a deep abiding love and an unfaltering faith in what they knew would benefit the object of their love. I have that dream also.

Jay and Dale Thacker, life long neighbors, demonstrated care for the children in their community. They were uncompromising in setting a high standard for academic accomplishment in their children, Cheryl and Shawn. Although observing from a "distance," I was impressed by, and later deeply respected, their continuous dedication to the pursuit of educational accomplishment as a worthy goal. I understand the benefit of the discipline they taught and have experienced the exhilaration of achievement possible when one extends themselves toward a worthy goal. Jay and Dale your teaching lives on.

Ronald Arthur Edwards, my life long friend whom I appreciate more as the years pass, represents the very best example of a marvelous combination of intellect and undaunted courage. Early on he taught me concerning athletic and academic competition, areas in which he excelled. I observed a wholesome attitude in Ronnie of confident expectation that anything he attempted he could accomplish. On several occasions during this research my version of courage gave way to a vision of discouragement and failure. During these periods I "remembered Ronnie" and was always encouraged to persevere. His influence still makes a difference in my life.

It is not by chance I am blessed with these particular parents, friends, and counselors. In His manifold grace, the only true and living God, the One in whom there is no shadow of turning, no evidence of variation, enriches my life through these precious people. Thank you, Heavenly Father, because with You chance plays no part.

TABLE OF CONTENTS

		Page
List of Tables		v
List of Figures		vii
Chapter		
1.	Introduction	1
	The Context of Uncertainty in Weapon System Acquisition	1
	Background of Uncertainty Quantification	2
	Uncertainty and Technical Requirements	3
	Statement of the Problem	7
2.	Literature Review	12
	Set Theory Background Summary	12
	Set Theory Fundamentals Applicable to Fuzzy Logic	15
	Concept of Fuzzy Logic	17
	Concept Development and Applications	20
	A Fuzzy Approach to Uncertainty	23
3.	Methodology	25
	Survey Instrument and Respondent Profile	25
	Context Specific Linguistic Variables	26
	Characterization of Fuzzy Membership Functions	27
	Simulation Approach to Fuzzification	38
	Uncertainty Rule Base	42
	Defuzzification	44
	Earned Value Concepts	47
	Cost and Price Data	48
4.	Analysis	50
	Context Specific Survey Instrument	50
	Characterization of Fuzzy Membership Functions	51
	Defuzzification	71
	Cost and Price Data	79
	Hypothesis Testing	91

TABLE OF CONTENTS
(Continued)

Page

5. Summary and Conclusions 94
   Discussion of Results 94
   Limitations 96
   Future Research 97

References 100

Appendices

A 103
B 107
C 113
D 116
E 125
F 131
G 137

TABLES

Table Page

1. Technical Uncertainty Fuzzy Membership Sets........................ 28

2. Schedule Uncertainty Fuzzy Membership Sets........................ 29

3. Category Task Type Functions... 40

4. Fuzzy Rule Strengths for Min Function................................ 43

5. Fuzzy Rule Strengths for Max Function................................44

6. Parameters for Best-Fit Triangular Functions for Technical and Schedule Uncertainty Data................................ 52

7. Results of Best-Fit Triangular Functions for Schedule Data.......... 56

8. Best-Fit Trapezoidal Functions: Technical and Schedule 61

9. Technical Uncertainty Fuzzy Set Triangular Functions................ 73

10. Modified Fuzzy Rule Strengths for Min Function.................... 78

11. Modified Fuzzy Rule Strengths for Max Function.................... 78

12. Adjusted Data and Descriptive Statistics for IGE Uncertainty Ratios.. 80

13. Adjusted Data and Descriptive Statistics for Negotiated Price Uncertainty Ratios...81

14. IGE Uncertainty Frequency Distribution.............................. 83

15. Negotiated Price Uncertainty Frequency Distribution................ 83

16. ANOVA for IGE and Negotiated Price................................ 84

17. Comparative Descriptive Statistics for 14 Point IGE "Re-quantified" Uncertainty and IGE Historical Uncertainty.................... 85

18. Historical IGE Uncertainty Frequency Distribution for 14 Points.... 86

19. Re-quantified IGE Uncertainty Frequency Distribution for 14 Points.. 86

20. ANOVA for IGE Uncertainty and IGE Uncertainty Re-quantified 14 Point.. 87

21. Technical Uncertainty Fuzzy Algorithm for 14 Points 88

22. Schedule Uncertainty Fuzzy Algorithm for 14 Points 89

23. Paired t Test of Historical and Re-quantified IGE Uncertainty....... 90

24. Comparison of IGE Historical Top Level Estimate with
Historical Negotiated Contract Price............................ 94

25. Paired t Test of Negotiated Price Percent Uncertainty Vs
Historical IGE Percent Uncertainty............................. 95

26. Paired t Test of Negotiated Price Percent Uncertainty Vs
Re-quantified IGE Percent Uncertainty......................... 95

27. Comparison of Re-quantified IGE Top Level Estimate with
Historical Negotiated Contract Price............................ 96

FIGURES

Figure Page

1. System Effectiveness Concepts.. 5

2. Degree of Membership in Very Low Technical Uncertainty........... 31

3. Degree of Membership for Each of Five Technical Uncertainty Fuzzy Sets.. 33

4. Degree of Membership for Each of Five Schedule Uncertainty Fuzzy Sets.. 33

5. Degree of Membership for the Medium Technical Triangular Function Uncertainty Fuzzy Set................................ 34

6. Degree of Membership for the Medium Technical Trapezoidal Function Uncertainty Fuzzy Set................................ 37

7. Triangular Function Variables Associated with Defuzzification...... 45

8. Triangular Function Discrete Variables of Overlapping Technical Uncertainty Fuzzy Sets.............................. 46

9. Degree of Membership for "Very Low" Technical Uncertainty Fuzzy Set.. 53

10. Degree of Membership for "Low" Technical Uncertainty Fuzzy Set.. 53

11. Degree of Membership for "Medium" Technical Uncertainty Fuzzy Set.. 54

12. Degree of Membership for "High" Technical Uncertainty Fuzzy Set.. 54

13. Degree of Membership for "Very High" Technical Uncertainty Fuzzy Set, with Triangular Fit #1............... 55

14. Degree of Membership for "Very High" Technical Uncertainty Fuzzy Set, with Triangular Fit #2............... 55

15. Degree of Membership for "Very High" Technical Uncertainty Fuzzy Set, with Triangular Fit #3............... 56

FIGURES
(Continued)

Figure Page

16. Degree of Membership for "Very Low" Schedule Uncertainty Fuzzy Set... 57

17. Degree of Membership for "Low" Schedule Uncertainty Fuzzy Set... 58

18. Degree of Membership for "Medium" Schedule Uncertainty Fuzzy Set... 58

19. Degree of Membership for "High" Schedule Uncertainty Fuzzy Set... 59

20. Degree of Membership for "Very High" Schedule Uncertainty Fuzzy Set, with Triangular Fit #1................ 59

21. Degree of Membership for "Very High" Schedule Uncertainty Fuzzy Set, with Triangular Fit #2................ 60

22. Degree of Membership for "Very High" Schedule Uncertainty Fuzzy Set, with Triangular Fit #3................ 60

23. Degree of Membership for "Very Low" Technical Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 63

24. Degree of Membership for "Low" Technical Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 64

25. Degree of Membership for "Medium" Technical Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 65

26. Degree of Membership for "High" Technical Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 66

27. Degree of Membership for "Very Low" Schedule Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 67

28. Degree of Membership for "Low" Schedule Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 68

FIGURES
(Continued)

Figure Page

29. Degree of Membership for "Medium" Schedule Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 69

30. Degree of Membership for "High" Schedule Uncertainty Fuzzy Set Triangular and Trapezoidal Comparison........... 70

31. Technical Uncertainty Fuzzy Membership Sets Overlapping Regions... 72

1. INTRODUCTION

Military weapon system development is characterized by increasing complexity of hardware and software. Improvements required to counter evolving threats have driven complexity to the point where hardware and software design emphasis is to take the "man out of the loop" due to constraints of reaction time and volume of data to be analyzed prior to an engagement decision. Changes in the balance of power among threat nations have also created significant impact on military force structure. Draw down of active duty units in all services has been implemented with corresponding impacts on Reserve and National Guard units. Along with threat changes, collective concern for national debt reduction, a balanced budget, renewed interested in national infrastructure, and the perceived opportunity to realize the "peace dividend" combine to constrain the Department of Defense (DOD) budget for the foreseeable future (Dorr, 1990:5-8).

A constrained budget environment encourages identification of opportunities to emphasize cautious stewardship of reduced program funding. Within the U. S. Army Air Defense Weapon System Community program focus includes control of system design, maintenance and reliability concepts to assure the best equipment is available to U.S. military personnel and that equipment is affordable. On par with concern for hardware and software design is the need for identification of potential cost savings or "cost reduction." DOD is concerned with all aspects of weapon system affordability including the capability to develop credible cost estimates. The DOD Cost Analysis Community, including government and contractor personnel, is searching for approaches to improve the cost estimating process. Part of this concern is the recognized need to improve the quantification of uncertainty associated with the cost estimating process.

1.1 The Context of Uncertainty in Weapon System Acquisition

A cost estimate, which accompanies a contract through the acquisition process, should quantify technical and schedule uncertainty associated with the system being estimated and express this uncertainty in terms of funding requirements. A source of uncertainty is the inexact nature of the technical development process in general. Additionally, imprecision is introduced into the cost elements of an estimate particularly elements in which human judgment enters the process. Introduction of uncertainty, inherent in the analyses, arises within a development program from unplanned system changes, technical problems, schedule shifts, withdrawal of program advocacy etc.

A contract for weapon system development typically contains explicit reference to hardware and software specifications and a Statement of Work (SOW) depicting the technical overview of system functional capabilities and requirements. In advance of contract negotiations cost estimates are generated to quantify contractor cost for the phases of a system life cycle: (1) Development, (2) Production, (3) Military Construction, (4) Fielding, (5) Sustainment, and (6) Defense Business Operating Fund (DA Cost Analysis Manual, 1992:49-50). A cost estimate encompassing all phases is termed a life cycle cost estimate (LCCE). Often cost estimates consider only one or two phases, perhaps Development and Production, or are for certain aspects of a phase. Within Development several well defined phases, Concept Development, Demonstration and Validation, and Engineering and Manufacturing Development, are frequently individually of interest for development of a cost estimate. A cost estimate may encompass only contractor cost and be silent regarding government cost. This research develops a new methodology, not currently in use within the DOD Cost Analysis Community, for the quantification of technical and schedule uncertainty within the Development Phase of weapon system acquisition activities.

1.2 Background of Uncertainty Quantification

Weapon system baseline cost estimates have a credibility gap due to an often repeated problem of inaccurately quantifying expected cost of prospective systems. Numerous examples exist of cost estimates significantly understating contractor cost compared to subsequently negotiated contracts or extensions to existing contracts.

Uncertainty analysis and quantification, an element of cost estimates, has been a target of criticism at the highest levels within the DOD. The following statement documented in the Department of Defense Manual, DOD 5000.2-M, highlights the importance and the problem associated with uncertainty analysis within military cost estimates: "The purpose of cost uncertainty analysis is to bound an estimate...using an arbitrary plus or minus percentage to denote a range is not uncertainty analysis."

Numerous studies document methodologies for cost uncertainty while stressing the need for additional research to quantify uncertainty inherent in estimates. Generally, cost uncertainty analyses have been described in terms of statistical constructs, probability density functions, accompanied by rationale for describing subjective estimators of the shape of a particular distribution (McNichols, 1984:149). From a viewpoint of production cost Vinod and Basu (1990:4-6) discuss the boundary of a feasible set of opportunities in which technological uncertainty is treated in terms of rational expectations about a producer's objective probability distribution. In a general approach to cost uncertainty analysis Garvey (1992:163-167) describes the Analytic Cost Probability (ACOP) model which focuses on statistical technical issues such as correlation and covariance among work breakdown structure (WBS) elements. Accentuating technical statistical aspects of estimating, the ACOP model provides a probability distribution around an estimated cost which quantifies technical and acquisition risks. These documented studies while providing sound rationale for their methodologies have been statistical in concept.

1.3 Uncertainty and Technical Requirements

Considering the range of complexity, weapon system design, test and production is the most complicated of technical processes (Department of the Navy, Best Practices, 1986:Introduction). Specifying weapon system technical requirements, a major aspect of this complexity, is heavily dependent on human judgment which introduces the initial source of vagueness and imprecision in the requirements development process. Fuzzy logic and its subsets effectively address vagueness through its capability of modeling non statistical imprecision (Valluru and Hayagriva, 1993:26; Nikhil and Bezdek, 1994:107).

The relationship of uncertainty to technical requirements will be mentioned in the context of reliability engineering relating to hardware system effectiveness and design specification. In each of these reliability areas the introduction of imprecision and vagueness sets the stage for use of a methodology to quantify uncertainty for inclusion in the cost estimating process.

1.3.1 Hardware System Effectiveness

System effectiveness is the term which describes the capability of any system to accomplish the mission for which it was designed. The degree to which a system achieves the function for which it was designed determines its effectiveness. Of the several major attributes determining system effectiveness, reliability has received a very thorough and systematic study in recent decades as a result of increased reliability problems of complex weapon systems particularly since World War II (Ahmed, 1990:290-295; Bai et al., 1989:528-532; Bowles, 1992:1-12). Influence on system effectiveness comes from two notable sources e.g. (1) newness of equipment design and (2) interrelationships among system priorities. A major influence is hardware design involving either new equipment with predominantly new design or new equipment with standard design and previously tested parts.

Experience demonstrates that development of new equipment with a new design generally suffers from a lack of attention to reliability engineering in the early design phase. This results in inadequacy of design which translates to imprecision or vagueness of design criteria related to reliability. Traditional reasons for inattention to reliability have centered around two primary program constraints: (1) minimizing cost and (2) accelerating development schedules. This combination of constraints has operated to significantly reduce time allocated for proving the feasibility of new design. Not surprisingly the resulting hardware prototypes are overweight, oversized, not designed for production, not designed for ease of maintenance, and characterized by frequent failure. Neglect of reliability in early design activities inevitably requires extraordinary redesign effort during a later development (or procurement) phase to realize the degree of system effectiveness originally envisioned. Lack of focus regarding reliability in new equipment design reduces system effectiveness.

Effectiveness depends on performance capabilities and reliability and also on other system properties: operational performance, availability, maintainability, and reparability. These interrelated concepts should be viewed collectively within the system structure. Each of the properties are inherently vague. Interrelationships among system properties implies that maximizing all properties simultaneously would probably not be desirable. Even casual consideration of the properties gives rise to the notion of "trade-off" relationships.

It is not difficult to imagine trade-offs between reliability and cost, reliability and maintainability, and maintainability and cost. Optimization of system effectiveness is a complex task due to the high degree of interaction among the properties and the inherent vagueness. Figure 1.3.1-1 displays an overview of system effectiveness concepts and associated time related measurements. Review of the figure suggests trade-off situations each of which would generate unique technical uncertainties needful of a methodology addressing quantification. Imprecision in development of operational requirements during design assures unresolved reliability parameter problems produce risk situations. When risk candidates (uncertainty) are not identified and managed, the opportunity to positively influence system operational performance is diminished (Transition from Development to Production, 1986:1-3).

Current Project Management practice places emphasis on managing hardware and software delivery schedules. This occurs to the extent that flow down of total design

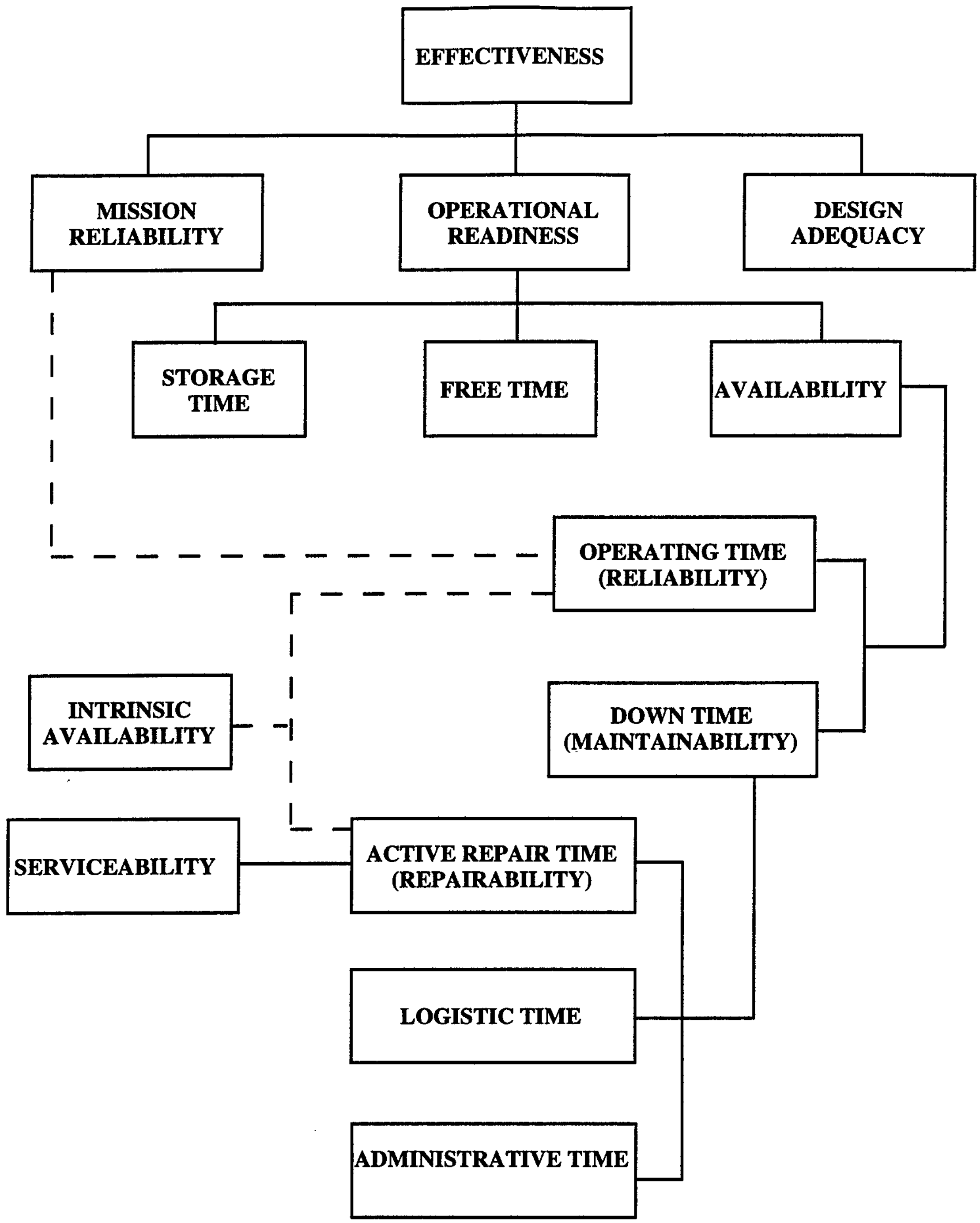


FIGURE 1.3.1-1.
SYSTEM EFFECTIVENESS CONCEPTS
SOURCE: RELIABILITY ENGINEERING, ARINC RESEARCH CORPORATION, 1965: 14.

requirements, including operational design requirements, may not be complete thus producing a technical risk (uncertainty) situation. Wehrle (1990:20-21) agrees with this assessment when he states that "in today's acquisition environment, decisions on how to spend those funds will favor performance considerations over supportability." Given this preference, Wehrle states that logistics and the "ilities": reliability, availability and maintainability, historically receive minor attention during the early design phase and thus operational considerations further compound the problems of lack of early design influence. The practice of neglecting to fully define the system early in the design phase has been described by Wiskerchen and Pittman (1989:29) as a linear orientation to system development. Linear orientation focuses on physical design parameters and ignores environmental parameters. Therefore this practice leads to design habits which optimize technical parameters with minimal emphasis on optimizing operational parameters. A suggested resolution to this situation is to restructure the procurement system to force recognition of full system requirements including operational maintainability and availability issues (Wiskerchen and Pittman, 1989:31).

1.3.2 Design Specification

At the design specification level problems are observed when a linear orientation is followed in early design efforts. System design requirements are flowed-down from operational requirements. Department of the Navy Best Practices (1986:4-3) states that operational requirements have "traditionally been tactically oriented" emphasizing technical performance parameters such as range, speed, etc. System specifications should define not only threat, theater environments, and performance parameters, but the total envelope of external environments. The linear approach requires the hardware developer to augment, derive, specifications to define system internal environmental conditions. External and internal environmental conditions become the design criteria for component parts of a system. Operational requirements alone, from which design requirements are derived, are of limited value to system designers. Vague operational requirements may be translated into design requirements that are not inherently measurable by the design process and therefore must be exhaustively and formally tested. Specifying vague operational requirements, from a design viewpoint, leads to assumptions which may or may not be in agreement with the reference mission profile. Soft requirements lead to unnecessary formal testing to develop measurable requirements. Deriving requirements with an imprecise or inaccurate mission profile ultimately leads to increased reliability problems (Department of the Navy, 1986:4-4 - 4-11).

1.4 Statement of the Problem

DOD Weapon system acquisition is a complex business which spans many years from concept development through introduction into inventory. Cost estimating, a component of the acquisition process, is critically reviewed by congressional members and DOD component agencies many of which are searching for methodologies for improvement of cost estimates. The focus of this research is to adapt two aspects of fuzzy logic, linguistic variables and membership sets, for the quantifying of technical and schedule uncertainty associated with the cost estimating process. Various elements of the Theory of Fuzzy Logic, developed by professor Lofti Zadeh, have been used in numerous applications but to date there is no evidence of elements of fuzzy logic applied to the quantification of technical and schedule uncertainty associated with the world's most complicated process: DOD weapon system acquisition.

Assessment of technical and schedule uncertainty is currently accomplished in a group environment where knowledgeable personnel make informed judgments relative to specific aspects of system development. The group environment generally represents skills in engineering, program management, product assurance, reliability, logistics, configuration management, production and software. Often a form of the Delphi Technique is employed to gather expert opinion via a questionnaire to quantitatively assess technical and programmatic uncertainty. Monte Carlo simulations are conducted to determine the impact of uncertainties in specified technical and schedule areas. Technical uncertainty, quantified by expert opinion, is translated into schedule and cost uncertainty to provide an analysis of the likelihood of meeting a particular set of technical goals within a specified time. Questionnaires solicit uncertainty (risk) ratings using specified criteria related to consequences of occurrence (catastrophic, critical, marginal, negligible) and probabilities of occurrence (frequent, probable, improbable). Aggregated technical ratings are translated into linguistic descriptions such as "high, medium and low" in accord with definitions provided in a survey questionnaire. Probabilities of occurrence are stated as "crisp" numbers such as >75%, < 25% etc. Membership in these descriptive sets follow the Aristotelian logic of either 0 or 1.

Cost data, used as a basis for quantifying potential cost growth associated with these ratings, is derived from a program specific baseline cost estimate (BCE) which accompanies each acquisition phase of weapon system development. Cost data within a BCE can be categorized fixed or variable. Fixed costs are associated with sunk costs, fixed component or other government agency costs. Variable costs, time dependent, are represented by level of effort (labor) and are assumed to behave linearly with changes in schedule duration. Program BCEs address cost at a relatively top level, rather than at work package or below level, and therefore cost uncertainty can be analyzed at a top level.

Technical and schedule uncertainty quantification associated with cost estimating has been a constant subject of concern in the cost estimating community. Methodologies for quantifying uncertainty are numerous. The objective of this research is to document a new methodology for quantification of uncertainty which utilizes elements of fuzzy logic.

1.4.1 Methodologies to Deal with Uncertainty

Quantification of uncertainty is a recognized cost element within accepted cost estimating techniques associated with weapon system BCEs. The terms uncertainty and risk, occasionally used interchangeably in connection with cost estimating, are distinctive. Where risk deals with measurable probabilities, uncertainty is evidenced by the lack of probabilities associated with an event. Uncertainty is concerned with possibilities where sufficient knowledge is available to make subjective judgments rather than probability statements (D.A. Cost Analysis Manual, 1992:35). Quantification of uncertainty associated with cost has not progressed to the point of acceptability to DOD agencies responsible for review and approval of cost estimates. The need for application of relevant theories addressing uncertainty to the cost estimating challenges of complex weapon systems is evidenced by guidelines promulgated by DOD: "The purpose of cost uncertainty analysis is to bound an estimate...using an arbitrary plus or minus percentage to denote a range is not uncertainty analysis" (DOD 5000.2-M, 1991:8-8 - 8-9). Further the DOD Manual states, cost uncertainty arises from unplanned system changes, technical problems, schedule shifts and in early development from key performance relationships. From this perspective a credible cost estimate must account for uncertainty while not being arbitrary. Additional research is needed in the quantification of uncertainty to address concerns about the arbitrary nature of current methodologies.

Certain elements of fuzzy logic are demonstrated in this research as a technique for handling the imprecision, vagueness, that inherently accompanies judgmental inputs such as in the quantification of technical and schedule uncertainty. The application of elements, linguistic variables and membership sets, of the Theory of Fuzzy Logic to cost uncertainty provides a new application of this theory.

1.4.2 Research Hypothesis

Information used in this research consists of two specific sets of data: (1) cost estimates which are a required procurement related item in the weapon system acquisition management process and (2) negotiated contract prices which result from a procurement action for which a cost estimate is generated. Cost estimates are presented at a total Research, Development, Test and Evaluation (RDT&E) level with no visibility into the CES within the estimates. The total cost estimate is composed of two data elements, the portion of the cost estimate which is "uncertainty free" and the portion which represents the quantified uncertainty. When summed these two data elements are the total RDT&E cost estimate. The actual individual top level cost estimates are compared to the actual individually negotiated contract prices and a variance is calculated.

The hypothesis for this research is: through the application of linguistic variables and fuzzy membership sets technical and schedule uncertainty can be quantified and added to the "uncertainty free" portion of the actual cost estimates such that the variance between the recalculated total cost estimates and the negotiated contract prices will be reduced.

The null and alternate hypotheses are:

> H_0: There is no difference in the mean of the top level historical IGEs and the mean of the top level IGEs calculated using the "re-quantified uncertainty."
>
> H_1: There is a difference in the means of the top level historical IGEs and the means of the top level IGEs calculated using the "re quantified uncertainty."

The hypotheses are expanded to a four step process that includes testing two sets of data: (1) a complete forty six pair set, (2) a subset of the forty six pair set, the fourteen pair set, (3) a paired t test on the fourteen pair data set prior to application of the "crisp" output of the fuzzy algorithm and (4) a paired t test on the fourteen pair data set after applying the "crisp" output of the fuzzy algorithm.

1.4.3 Research Framework

The objective of this research is to demonstrate a methodology for improving the quantification of uncertainty related to cost estimates for a specific weapon system. This will be accomplished through use of the fuzzy logic "degree of truth" concept which extends traditional logic by labeling sets qualitatively using linguistic variables and quantifies the truth of output solutions to a strength reflecting the truth value which is called "degree of membership." Components of the research include: (1) survey data, (2) characterization of membership functions, (3) fuzzifying data inputs, (4) fuzzy associative memory rule base, (5) defuzzifying outputs for "crisp" solution variables.

This research is not focused on the mathematics of fuzzy set theory but on the practical implementation level. In the perspective of Tyler Sperry "fuzzy logic is not the same thing as the mathematics of fuzzy set theory" (Sperry, 1993:33). Research presented here specifically develops linguistic variables and fuzzy membership sets for the purpose of practical application in the quantification of technical and schedule uncertainty within the context of a major U.S. Army missile system. Use of any elements of the Theory of Fuzzy Logic to uncertainty related to weapon system acquisition is a new application of this theory.

Weapon system cost estimates must conform to a cost element structure (CES) specified by the service proponent requiring the estimate. U.S. Army weapon systems CES is specified in the Department of the Army Cost Analysis Manual and defined by element i.e. RDT&E Funded Elements, Procurement Funded Elements, Military Construction Funded Elements, Military Personnel Direct Funded Elements, Operations and Maintenance Funded Elements and Defense Business Operations Fund Elements. The CES provides system-specific, appropriation discrete, uniform cost structures with standardized cost elements and definitions. Further, the CES accommodates recent changes in the defense acquisition management processe including defense management review decisions and program budget decisions. Tailoring of the CES is accomplished for each specific cost estimate and includes cost elements for hardware, software, and services. Quantification of cost related to technical and schedule uncertainty within a system specific CES is accommodated in the elements Other RDT&E, Other Procurement, Other Military Construction, and Other Operations and Maintenance. This study demonstrates a fuzzy logic methodology for the quantification of uncertainty applied to top level estimated costs.

Data was generated specifically for this study which required development of a survey questionnaire. Respondents to the survey were full time professional or managerial employees supporting a U.S. Army Missile System Project Office. Other data used in this study was obtained by permission from a U. S. Army Missile Project Office and is considered "proprietary." Due to the proprietary nature of this data, actual values relating to historical and present cost have been adjusted. Specific adjustments were made through use of a common indexing routine applied to the actual values. The indexing has no significance to the outcome of this research and will not be presented in this study in order to preserve the proprietary nature of the data.

2. LITERATURE REVIEW

2.1 Set Theory-Background Summary

Set theory, originated by Georg Cantor (1845-1918), forms an extension for the mathematical basis for fuzzy logic. Review of a few early contributions to the mathematics of set theory will provide appreciation of the struggle by Cantor to establish its foundation. Development of set theory provided fundamental insights into the study of virtually every other branch of mathematics. It was a launching point for advances in abstract mathematics, such as consideration of infinity, a curiosity to both mathematicians and philosophers. One point of connectivity between mathematics and philosophy is evident from the early contributions to the concept of infinity provided by Galileo Galilei. In his Discorsi (including discussions with Salviati, Sagredo and Simplicio) Galilei investigated the notion of continuous infinity. Considering the "number of numbers", sets of natural numbers and squares of natural numbers were determined to be equipollent giving rise to the concept that attributes such as equal, larger, and smaller have no application relative to infinite quantities but only to finite quantities. This notion of a part not being smaller than the whole is referred to as the "paradox of Galilei." In an attempt to make the concept of the infinite less suspect Hilbert popularized the paradox of Galilei through his example of a queue of arriving guests at a hotel with a "countable" number of rooms yet always having sufficient vacancies to accommodate the guests (van Dalen, 1972:3-4).

Contributing to the concept of sets and infinity by borrowing examples from the finite, Leibniz and Johann Bernoulli made analogy to the concept of a line segment. They argued the implication that given a line segment where all members of a series exist (1/2, 1/4, 1/8...) an infinitely small member exists of every definable fraction. Therefore, a member of the line segment exists which can be called the infiniteth (infinitesimal) member. Bernard Bolzano, one of the first to analyze the concept of the infinite, defined a collection as a whole containing certain "parts" where a group of parts, i.e. set, could be considered distinct. His example of a "whole" drinking glass versus a glass broken into several pieces focused attention to the concept of ordering as the key to distinction. To this notion Menge added that if a collection is defined whereby ordering is not relevant then this collection is a set. This gave rise to the idea that a set was a natural extension of the concept of property. Definition of the concept of sequence by Bolonzo included each element (part) having a successor and predecessor element and introduced the concept of countable and finite sets and integers. This work gave birth to the idea that with the property of addition a successor is generated from a predecessor. This concept of sequence moved mathematics toward acceptance of the infinite set (van Dalen, 1972:6).

In 1888 Richard Dedekind, in his proposal of a set based on the concept of a number, made a philosophical contribution to number theory and algebra by arguing that "numbers are free creations of human mind." In this work he promulgated the notion of a number on the idea of a set separating arithmetic from intuitive observation and linking it with consequences of the laws of thought. Cantor's reasoning regarding sets held similar philosophical connotations according to his definition of a set as "a collection into a whole, of definite, well distinguished objects of our perception or of our thought" (Kamke 1950:1). In Dedekind's monograph he expounds thinking on "systems" made up of objects in which objects are collected that in turn make up a set. He grants the existence of a one object set while not recognizing the existence of an "empty system" in which there were no objects. Dedekind provided early definition of the concepts of union, intersection and subsystem through his definition of mapping: "a law assigning objects to all elements of S." He is famous for his definition of finite and infinite using the concepts of system and mappings without reference to natural numbers: "a system S is called infinite if it can be mapped by a one-one mapping onto a proper subsystem of itself otherwise it is finite" (van Dalen 8).

During the period 1870-1872 collaboration between Dedekind and Cantor resulted in the establishment of the countability of real algebraic numbers and the uncountability of the set of reals. Set theory was formally documented soon afterward in Cantor's 1874 paper, a milestone in the history of mathematics, in which he provided a proof of the existence of transcendent real numbers. The theorem of unaccountability of the set of reals which Cantor proved was: let a countable set S of reals be given, then we can find in every interval (a,b) a real number not belonging to S. The proof follows.

> Consider the set S with elements $s_1, s_2, s_3 ...$, let a and b be given, such that $a < s_j < b$ and let $s_i < s_j$ and put $a_1 = s_i$, $b_1 = s_j$. Suppose that $a_1, ... a_n$, $b_1, ... b_n$ are defined. Locate the first s_i and s_j such that $a_n < s_i$, $s_j < b_n$, if there are such elements in S. If not, the proof is finished because the interval (a_n, b_n) contains at most one point of S. Now suppose an infinite sequence of intervals (a_n, b_n). There are two possibilities:
>
> (i) $\lim a_n = \lim b_n =$ r, then: $\lim a_n = \lim b_n = r \in S$ because otherwise $r = a_k .. or .. r = b_k$ for some k, which is impossible as $a_k < r < b_k$ for all k..

(ii) $a = \lim a_n < \lim b_n = b$.. In that case any point between a and b satisfies the requirement.

Dedekind provided insight to Cantor in the establishment of a one to one correspondence of *n-dimensional* Euclidean space R^n and *1-dimensional* space R (van Dalen, 1972: 10-11).

Other contributions of Cantor include: topological notions such as well-ordering, ordinal, cardinal and diagonal procedure. Well-ordering was characterized by Cantor as sets, non-empty sets, which have a first element and each element, except the last, has a successor, and if a subset has an upper bound in a set, it has a least upper bound. The "paradox of the set of all ordinals," absurd contradictions inherent in sets, was known to Cantor as he discussed in the "system" Ω of all ordinals: "If Ω is a set it would have an ordinal δ which would be greater than all ordinals α , in particular δ would be greater than δ." This paradox caused Cantor to segregate "multitudes" into separate categories: those which could be collected into a whole or completed thing, and those that could not be considered completed without contradiction (later Von Neumann called the multitudes sets and classes) (van Dalen, 1972:21-22).

Cantor defined cardinal number M "the general concept, which by means of our active thinking-faculty results from the set M by abstracting from the nature of the elements and the order in which they are given." At the time (1895-1897) Cantor did not have the concept of equivalence classes to provide a mathematical definition suitable for dealing with this double abstraction of the cardinal number M within set theory. Cantor used the diagonal procedure for proofs of countable sequences for both two elements m and w and for characteristic functions. His used set theoretical notions, mappings on mappings, to argue the cardinality of the set of subsets L is greater than L (van Dalen, 1972:17-18).

Cantor's development of set theory was accompanied by detractions from mathematical, philosophical and theological scholars. Notable mathematical opponents included his former professor Leopold Kronecker a well known number theorist and algebraist. In his efforts to "purge" mathematics, Kronecker directed some of his attacks toward his former student's work in transfinite mathematics holding that "only integers are eligible for meaningful mathematics" (van Dalen, 1972: 15). Regardless, Cantor is recognized as the originator of the greater part of set theory.

2.2 Set Theory Fundamentals Applicable to Fuzzy Logic

The Theory of Fuzzy logic derives a strong theoretical background from fundamental set theory. Lofti Zadeh, known for his formulation of the theory, extended traditional set theory for application to control and decision making problems by incorporating a "degree of truth" concept rather than the "all or nothing" classifications of Aristotelian logic. This concept allows sets to be characterized qualitatively through use of linguistic variables and provides measurement techniques of the strength of the degree of truth (Voit, 1993: 26-27).

Set theory investigates the properties of sets from a general point of view independent of consideration of the nature of elements comprising a set. Basic operations on sets: union, intersection and complement are fundamental to the extension of set theory to "fuzzy sets." Referred to as the algebra of sets, union, intersection and complement are studied in relation to pairs of sets i.e. sets A and B on which these operations are performed. The following definitions are universally accepted (Kuratowski, 1972: 27-33):

- The union of two sets A and B is the set of all elements of set A plus all elements of set B. The operation of union of sets is denoted by $A \cup B$.

- The intersection of two sets A and B are the elements which are common, belong simultaneously, to A and B. The operation of intersection of sets is denoted by $A \cap B$.

- The difference of two sets A and B is the set consisting of elements belonging to A but not to B.

- The complement of a set A is the set of elements not belonging to A. The operation of complement of a set is denoted by $1 - A$. In the theory of sets all sets are considered as subsets of some fixed set called space. The set of real numbers form a space. Theorems of the algebra of sets are closely aligned with propositional calculus.

Propositional calculus applies logical propositions, derived from Aristotelian logic, which have only one of two possible values, either 0 (false) or 1 (true). Basic operations on propositions are disjunction, conjunction, negation and implication. When presented two propositions α *and* β the disjunction (the sum) is denoted by $\alpha \vee \beta$ and the conjunction, the product, is denoted by $\alpha \wedge \beta$. It is noted that the disjunction proposition is true if at lease one of the sets (components) are true. Conjunction propositions are true if both components are true. Negation of a proposition, denoted α', is 0 for a true proposition (negation of a true position is a false proposition) and is 1 for a false proposition. Applying Aristotelian logic negation is written;

$$1' \equiv 0$$
$$\text{and conversely } 0 \equiv 1'.$$

The law of double negation is derived from this logic:

$$\alpha'' \equiv \alpha.$$

From disjunction, conjunction and negation propositions Aristotelian logic derives the fundamental theorems of the law of the excluded middle (principium tertii exclusi) and the law of contradiction fundamental to classical logic states that; "from two contradictory propositions, one is true; no proposition can be true simultaneously with its negation." From this logic Demorgan's law states;

$$(\alpha \vee \beta)' \equiv (\alpha' \wedge \beta'),$$
$$(\alpha \wedge \beta)' \equiv (\alpha' \vee \beta')$$

In terms of implication a proposition α implies the proposition β. Stated in terms of "if, or, then" terminology;

$$\text{if } \alpha \text{ then } \beta.$$

The implication is written;

$$\alpha \Rightarrow \beta.$$

and the following is deducted;

$$\textit{if } \alpha \Rightarrow \beta \;\; \textit{and} \;\; \beta \Rightarrow \alpha \;\; \textit{then}\, \alpha \equiv \beta$$

Further, the law of syllogism (transitivity of implication) and the law of contraposition yield the following;

$$if\ \alpha \Rightarrow \beta\ and\ \beta \Rightarrow \gamma\ then\ \alpha \Rightarrow \gamma$$

and;

$$if\ \beta' \Rightarrow \alpha'\ then\ \alpha \Rightarrow \beta \text{ (Kuratowski, 1972:23-26).}$$

This information provides background for Zadeh's discussion regarding propositional calculus of "if-then" rules.

2.3 Concept of Fuzzy Logic

Foundations for the Theory of Fuzzy Logic, formalized by Lofti Zadeh professor of electrical engineering at the University of California at Berkeley, were published in 1965 under the title "Fuzzy Sets." In this work, Zadeh expounds the merits of a concept which "may be of use in dealing with imprecisely defined classes (real numbers, beautiful women, tall men etc.) which do not constitute classes or sets in the usual mathematical sense of these terms." He notes these "classes" are important in human thinking particularly related to the domain of pattern recognition and also important to the communication of information and other abstract notions. The phenomenon of decision making utilizing imprecise qualitative (non-numerical) information occurs daily. Zadeh argues the ability to make decisions with vague information as the key to why humans are extremely capable in situations requiring "control" judgments. Continuing the reasoning, he believes performance of control systems would be improved with the capability to use imprecise input. In his 1965 paper, Zadeh uses the terminology "fuzzy set" as a concept for effectively dealing with imprecise or vague input data. He defines fuzzy set as a class with a continuum of grades of membership. He envisioned a conceptual framework which drew on the existing body of knowledge of set theory and which could be applied (in a more general way) to multiple fields of study in which the absence of sharply defined criteria for class membership introduced vagueness or imprecision (Zadeh, 1965: 338). Zimmerman (1987:10) points out that imprecision is used in the sense of vagueness rather than to imply a lack of knowledge. Further he states the generalization of the classical notion of a set aids in the accommodation of uncertainty in a non-stochastic sense. Relating his comment to descriptive decision theory Zimmerman states that "vagueness only enters the picture when considering decisions under risk or uncertainty and then this uncertainty concerns the happening of a state or event and not the event itself." Further he adds that precision, "crispness", is not assumed in the presence of ambiguity and vagueness but is rather verbally modeled which normally does not permit full use of powerful mathematical approaches for analysis (Zimmerman, 1989:10-11). This concept of a class of objects with continuum of membership (e.g. from 0 to 1) accommodates modeling ambiguity and vagueness and represents a departure from the Aristotelian logic of "truth values" which have one of two possible values, either 0 (false) or 1 (true).

Entrenched in set theory, Zadeh's definition of fuzzy set uses the classical elements of set theory, null set (not considered part of set theory by Cantor), union, intersection, complement of a fuzzy set, algebraic operations, and convex and non-convex fuzzy sets in Euclidean space. He argues the appropriateness of the fuzzy set concept in that it effectively addresses the imprecision observed throughout the practical world. A fuzzy set resembles a probability function except it is completely non-statistical. Defining a fuzzy set in terms of a continuum of membership (i.e. from 0 to 1) provides a contrast to traditional set theory, as well as Aristotelian logic, in which set membership is considered only one of two possible values. The membership function is characterized by a range in which closer to unity represents a greater degree of membership of an element within a set. Each element of a set is associated with a real number in the continuum. This degree of membership, indicating an element can be a member of multiple sets simultaneously, is the distinguishing departure of Zadeh's theory of fuzzy logic from traditional set theory (Zadeh, 1965:339).

Basic notions of set theory and certain properties of the notions described by Zadeh in the 1965 paper provide a strong theoretical foundation for application of fuzzy sets to problems involving vagueness. Those notions include the characteristic function, empty set, union, intersection and complement. Notations vary but the concepts are extended from set theory discussed above;

The Characteristic (membership) Function: $f_a(x)$

associates for every point in X a real number in the interval [0,1] with a value of $f_a(x)$ at x which determines the membership of x in A. A fuzzy set is termed empty *iff* its characteristic (membership) function at X is zero. The notion of containment is integral to the related notions of union and intersection. Set A is contained in set B *iff* $f_a \le f_b$, *less than or equal to* written;

$$A \subset B \Leftrightarrow f_a \le f_b,$$

A union of two fuzzy sets *A and B* with membership functions;

$f_a(x)$ *and* $f_b(x)$ is a fuzzy set C written

$C = A \cup B.$

The membership function is related to *A and B* by;

$$f_c(x) = Max\ [f_a(x),\ f_b(x)],\ x \in X$$

in abbreviated form;

$f_c = f_a \vee f_b$, where $\vee$ indicates maximum.

Union has the property of association;

$$(A \cup B) \cup C = A \cup (B \cup C)$$

An intersection of two fuzzy sets *A and B* with membership functions;

$f_a(x)$ *and* $f_b(x)$ is a fuzzy set C written

$C = A \cap B$

The membership function is related to *A and B* by;

$$f_c(x) = Min\ [f_a(x),\ f_b(x)],\ x \in X$$

in abbreviated form;

$f_c = f_a \wedge f_b$ where $\wedge$ indicates minimum.

Intersection also has the associative property.

The complement of a fuzzy set A denoted A' is written;

$$f_{a'} = 1 - f_a \text{ (Zadeh, 1965:340-341).}$$

Basic algebraic operations on fuzzy sets include product, sum and absolute difference denoted below;

Product; $f_{ab} = f_a f_b.$

$and\ AB \subset A \cap B$

Sum; $f_{a+b} = f_a + f_b$

Absolute Difference; $|A - B| = f_{|a-b|} = |f_a - f_b|$

Zadeh (1965:345) defines a fuzzy relation in the context of fuzzy sets as a generalization of a function in the product space $X * X$. He reasons the utility of fuzzy relations due to pervasiveness of fuzziness in decision processes, thinking, which is not characterized by the traditional two-valued logic. Further, Zadeh sees the "fuzzy logic" in thinking as key to the ability of humans to summarize information. From a system analysis perspective, as complexity increases our ability to make precise, "crisp", statements (quantitative analysis) reduces to a point where precision and relevance diverge to become mutually exclusive characteristics. Traditional system analysis techniques therefore are considered intrinsically unsuited for effectively quantifying "humanistic systems" (Zadeh, 1973:28). In their paper concerning decision analysis Watson argues the need for a methodology to handle the imprecision generated by expert opinion inherent in the decision analysis process. They conclude that fuzzy set theory effectively models imprecision resulting from human judgment by allowing verbal inputs rather than numerically quantified inputs (Watson et al., 1979:1-8).

2.4 Concept Development and Applications

In engineering fuzzy logic has been extensively applied to control theory, particularly process control systems. Representative of process control systems is the steam turbine and vehicle speed controllers discussed by Cox (1992:59-60). These examples contrast the design of proportional-integral-derivative (PID) mathematical model systems with fuzzy logic designed systems in which one fuzzy rule replaces several rules of a conventional system. In a PID system, based on precise modeling of a process, model development rests on a set of equations describing a control surface in which coefficients are developed for the proportional, integral and derivative aspects of the system. A sensor provides a precise, "crisp", input value, the system applies the mathematical model, and a precise output value is calculated which translates into an appropriate action for the system to adjust its physical state to the environment and the process continues iteratively. A fuzzy logic process control system utilizes subjective input values derived from control statements represented by "natural language" terms. Characterization of input variables in a natural language is less precise than assignment of a numerical value for each expected state of a PID system. While less precise, this characterization is technically expressive allowing design of process control even when a rigorous mathematical understanding of a system has not been achieved (Cox, 1992:58-60).

Initial application of fuzzy logic as a practical tool occurred in conventional control systems technology. Japan leads in application and holds nearly 80 percent of the world market in fuzzy logic. Looking to the future the Japanese Science and Technology Association, equivalent of the U.S. National Science Foundation, has numerous projects planned concerning research in fuzzy logic development. The Fuzzy Logic Systems Institute in Iizuka, Japan subsidizes studies for forty full time researchers. The Japanese Ministry of International Trade and Industry (MITI) pursues fuzzy logic development in decision support, robotics, natural language and image understanding and fuzzy computing including fuzzy associative memories. A private institution in Kyoto specializing in fuzzy systems research and development investigates tracking problems, tuning, human factors, interpolation and classification problems including handwriting. This same organization is pursuing fuzzy applications in anti-lock breaks, automatic transmissions, impact warning and monitoring, windshield washers, and light dimmers. Developed at the Toyko Institute of Technology, a fully automated pilot-less helicopter responds to voice commands via a fuzzy logic controller. Accelerating, breaking and stopping for the Japanese Sendia Metro, which opened in 1987, is governed by a fuzzy logic controller. Fuzzy logic control system technology has been applied by the city of Toyko to its subway system. The classic control problem of the inverted pendulum was demonstrated at the Second Congress of International Fuzzy Systems Association in Tokyo by Takeshi Yamakawa. For non-linear control, Yamakawa added a live mouse on a platform atop the inverted pendulum experiment and demonstrated fuzzy logic compensation for the movement of the mouse. Other examples of Japanese application of fuzzy control include: auto focus and image stabilization mechanisms for cameras, automatic adjustment of washing cycle for load size, type and amount of dirt and fabric type for washing machines, vacuum cleaners, air-conditioners, electric fans and hot plates (Swartz and Klir, 1992:32-35).

Applications of Fuzzy Logic, initially expected to be in pattern recognition, communication of information, and abstraction, have become extensive since the mid sixties. Control system design continues to be a successful and rapidly expanding field for application of fuzzy logic. In information technology fuzzy logic applied in a decision support role provides additional reasoning capabilities requiring a minimum of "rules" for execution within expert systems. Rule reduction is possible because imprecision is tolerated due to variables having a relatively wide range of states. Control statements are written in terms allowing utilization of linguistic expression to aid in quantifying states of a variable. Using a process control example, linguistic expression of input variables can take the form of "this is cold", "this is warm", "this is very warm", "this is hot", "this is very hot" rather than having a rule for every degree of temperature in the expected range. "Fuzzified" expressions execute rules in the knowledge base and output values are "defuzzified" to provide a crisp solution for the control process.

Adaptive control in a fuzzy system allows control systems to adjust to environmental changes (Cox, 1993:28). This adjustment, necessary because physical systems are subject to permanent alterations over time due to wear, is particularly interesting because it represents the ability to become self-organizing. The concept of self organization, a major topic in the science of complexity, is currently under study at the Santa Fe Institute. In a paper on classifier systems, John Holland of the University of Michigan an active participant in complexity studies at the Santa Fe Institute, notes that systems of all kinds are intrinsically dynamic and are continually adapting to environments, often with accompanying improvements in performance. He labels these systems adaptive nonlinear networks (ANNs). His work provides insights into the question "how does an ANN perpetually adapt to new environments that continually offer opportunities for improvement" (Holland, 1989: 463-464). Adaptive Fuzzy Controllers are a technology example of self organization.

The assurance sciences are a subject area for application of a fuzzy approach due to the inherent degree of imprecision (uncertainty). Reliability design analysis and evaluation activities associated with the development and/or upgrade of weapon systems is known for uncertainty due to imprecision. Stringent operational requirements combined with the trend toward increasing technological complexity provides a range of opportunities for the application of fuzzy logic. One requirement of concern is the availability cycle for weapon systems (Verma, 1994:436-439).

Much of the uncertainty associated with system operational availability stems from the nature of reliability which considers inherent, environmental and operational aspects. Stochastic in nature, this approach to uncertainty has been thoroughly addressed through techniques of probability theory. Uncertainty associated with reliability analysis of weapon systems is also inherent in human thinking, reasoning, cognition and perception processes. The origin of this uncertainty is what Zadeh describes as humanistic systems: those strongly influenced by human judgment, perception, or even emotions. From a reliability viewpoint, this uncertainty is initially reflected in the imprecision associated with the development of a system level reliability requirement. For example, just how reliable must the system be in terms of mean time between failure (MTBF)? Zadeh contends that it is just this type of imprecision, invoked through human judgment, where the application of fuzzy logic proves useful.

2.5 A Fuzzy Approach to Uncertainty

The purpose of quantifying uncertainty is to provide an "upper bound" on estimate. Bounding can be accomplished objectively, by statistical analysis, or subjectively, through the use of expert opinion (Defense Acquisition Management Documentation and Reports, DOD 5000.2-M, 1991: 8-9). This research will demonstrate a methodology for improvement of cost estimating by developing a "fuzzy approach" for quantification of technical and schedule uncertainty. The methodology will be accomplished by developing fuzzy sets, "membership functions," grouped by "linguistic variables." The terms membership functions and linguistic variables are well documented subsets of fuzzy logic.

These subsets of fuzzy logic, membership functions and linguistic variables, will be used in conjunction with contract specific data, negotiated prices and independent cost estimates (IGEs) developed by the government for the specific contracts. A methodology for a new application for these subsets of fuzzy logic will result in quantifying technical and schedule uncertainty associated with research and development effort of a weapon system program.

Uncertainty increases as a product of the number of uncertain inputs into a study. Uncertainty analysis, criticized by DOD for using an arbitrary percent to bound a cost estimate, is a specific area in which expert opinion, leads to a quantification of a cost element within a cost estimate. To date no DOD cost estimating methodology has been documented to deal with the vagueness, or imprecision, associated with the expert opinion which forms the basis of quantification of cost uncertainty analysis. This particular aspect of the field of cost estimating for weapon systems is ripe for a new application of a subset (membership functions and linguistic variables) of an important theory (Fuzzy Logic). These subsets of the theory of Fuzzy Logic facilitate quantification of uncertainty, within a specific context, without use of statistical approaches. The importance of these subsets of fuzzy logic were pointed out by Hisdale (1994:22) in her paper concerning Interpretative Versus Prescriptive Fuzzy Set Theory: "I believe that the fundamentally great achievements of fuzzy set theory are: (1) the introduction of linguistic values of variables, e.g. "tall" into a theory of mathematical logic, and the attempt to explain how human communication works in everyday life when such labels are used instead of numerical values; and (2) the introduction of the possibility of a partial grade of membership value of an object in a class instead of having a choice solely between the membership values 0 and 1."

Use of linguistic variables is context dependent (Zadeh, 1984:29). The context in which linguistic variables are used in this research is specifically U.S. Army Missile System cost estimating dependent. A survey instrument was utilized to gather context specific data associated with the linguistic variables in two categories; technical and schedule uncertainty. Participants in the survey were experienced with U.S. Army missile systems and represented a range of skills: engineering, logistics, product assurance and program management. Their collective DOD missile weapon system peculiar experience provides a robust heritage of specific context dependent knowledge. Data gleaned from this survey provided the basic information for the research methodology and its application to the theory of fuzzy logic.

3. METHODOLOGY

This research was conducted within the context of the acquisition management processes of a U.S. Army Missile System and uses elements of the fuzzy logic for quantification of technical and schedule uncertainty associated with cost estimating. Major elements of the research include: (1) development of a context specific survey instrument, (2) context specific linguistic variable definitions, (3) characterization of fuzzy membership functions, (4) simulation approach to fuzzification, (5) development of a technical and schedule uncertainty rule base, (6) defuzzifying outputs for "crisp" solution variables, (7) applying concepts of earned value management in a feedback role to provide a starting point for updating the uncertainty quantification process and, (8) cost and price data. This research formalizes and documents a procedure which can be iterated to update uncertainty cost estimating predictions for a specific system.

3.1 Survey Instrument and Respondent Profile

A questionnaire, provided at Appendix B, was developed to provide input data for formalizing this methodology. A sample size of thirty personnel completed both the technical and schedule sections of the questionnaire with specific weapon system professional experience ranging from a minimum of four to a maximum of twenty three years. Fourteen of the respondents had more than ten years continuous experience with the specific weapon system to which this data applied. Areas of functional experience represented in the sample included: test and evaluation, program management, hardware, software, system engineering, product assurance, safety, logistics, cost estimating, and performance simulation.

The questionnaire clearly stated the purpose of the survey: to determine the potential for improving quantification of uncertainty in the cost estimating process. Two categories of uncertainty, technical and schedule, were defined. For each of these categories of uncertainty linguistic variables were defined; very low, low, medium, high and very high. Respondents were tasked to provide a percent range of uncertainty for each of the linguistic variables in the form of From____% To ____%. To focus the intent of "percent range" in the mind of each respondent a concise statement for technical and schedule percent range was delineated that featured increasing levels of difficulty.

Each linguistic variable was defined with one major emphasis reiterated throughout the definitions. Technical uncertainty definitions focused on requirements.

An increase in technical difficulty was imposed by defining an increasing level of constraints for each linguistic variable from very low to very high. Schedule uncertainty definitions focused on the major emphasis of contract duration, number of months of a program. An increase in schedule difficulty was imposed by defining an increasing level of constraints for each linguistic variable from very low to very high. The questionnaire recognized, by a specific statement, that schedule uncertainty is not independent of technical uncertainty. To address the imprecision associated with attempting to differentiate schedule from technical uncertainty the questionnaire included a listing of activities or events whose occurrence would indicate "relatively more" schedule impact and "relatively less" technical impact. The usefulness of fuzzy logic in a methodology to quantify uncertainty is that it effectively address the inherent vagueness and imprecision in human thinking. It was observation of this fact, that humans make decisions using vague, imprecise, non-numerical information, that prompted development of the theory of fuzzy logic (Zadeh, 1965:338; Zadeh, 1973:28; Cox, 1992:58; Cox, 1993:27, Schwartz and Klir, 1992:32; Self, 1990:42). Inspection of the survey data quickly reveals the need for a methodology that can be adapted to handle the inherent vagueness.

The range of data points, from the completed survey questionnaires, were associated, through specific definitions, with linguistic variables to characterize technical and schedule fuzzy membership sets. Use of the data as input variables to a computer program provided graphic evidence of fuzzy sets as "classes of objects in which the transition from membership to non-membership is gradual rather than abrupt" (Zadeh, 1973:28). A simulation and a curve fitting routine was developed that identified each data point with a fuzzy membership set. The simulation and curve fitting routine provided quantification and a graphical depiction of the concept of a fuzzy set as a "class" representing a continuum of grades of membership (Zadeh, 1965:339; 1984:26).

3.2 Context Specific Linguistic Variables

Linguistic variables used as labels for fuzzy membership sets are related in this methodology to the context in which they are used. The idea of context led Pal and Bezdek (1994:108) to define the concept of relativistic fuzzy sets to describe the fuzziness with respect to a particular observer. A sharpened version of this concept is provided by the example of the linguistic variable TALL. Defined in the context of U.S. standards the variable TALL would differ from the same variable in the context of Asian standards. Zadeh (1984:29) describes context specific fuzzy terms relating to the control of a process controller for a kiln operation. He specifically states the fuzzy terms "OK," "normal," "low," and "high" represent definitions based on measured quantities of input variables to the process.

This methodology uses five context specific linguistic variables: very low, low, medium, high and very high for both technical and schedule uncertainty definitions. Cox (1992:61) recommends an odd number of variables between 5 and 9 with overlapping regions between ten to fifty percent of the adjacent membership function regions. These variables are context specific in respect to industry (DOD weapon system acquisition), proponent service (U.S. Army), type of system (weapon), type of weapon system (missile), professional skills (all weapon system acquisition related) and system specific personnel experience (with multiple years experience on a particular missile system). Further delineations are possible but would provide no value added. In this methodology each of the linguistic variables are the label of a fuzzy membership function which semantically characterizes the concept of set.

Context specificity extends to individual interpretation of the definitions used in the questionnaire. Respondent opinion was expressed based on experience on similar RDT&E and Procurement programs of the missile system. Ambiguity in interpretation and application of linguistic variable definitions was reduced, but not eliminated, due to the homogeneity inherent within the context specific areas enumerated above.

3.3 Characterization of Fuzzy Membership Functions

The essence of the methodology for establishment of membership functions is contained within the concept of set membership. Within this research the five context specific linguistic variables, labels, describe the states, or regions, of the fuzzy sets for technical and schedule uncertainty. These labels were chosen because technical personnel in the field of weapon system acquisition are accustomed to thinking of uncertainty in the form of a LOW-MEDIUM-HIGH framework. Similarity between human thinking and application of linguistic labels facilitated the development of the methodology and illustrates one the strengths of fuzzy logic: it simulates human thinking patterns. Information provided by each respondent to the questionnaire makes up the data of each set.

The fuzzy sets are composed of unsorted data with each row representing the opinions of a respondent for each linguistic variable label. These fuzzy sets allow partial membership states which are referred to in this research as degrees of membership. The methodology allows overlap with the space of adjacent fuzzy sets. Overlap provides a smooth transition from one fuzzy membership set to another. Raw data from the completed technical and schedule uncertainty questionnaires are presented in Table 3.3-1 and 3.3-2.

Table 3.3-1

Technical Uncertainty Fuzzy Membership Sets

Technical Input Data By Respondent

Linguistic Variables	Very Low	Low	Medium	High	Very High
1	.10-.20	.15-.40	.35-.60	.55-.80	.75-1.00
2	.00-.03	.03-.25	.25-.35	.35-.60	.60-1.00
3	.05-.20	.20-.35	.35-.65	.65-.85	.85-1.00
4	.00-.15	.10-.20	.15-.80	.80-.95	.90-1.00
5	.00-.10	.05-.45	.40-.70	.70-.90	.85-1.00
6	.05-.10	.10-.25	.25-.50	.50-.75	.75-1.00
7	.01-.07	.05-.15	.15-.30	.25-.45	.45-1.00
8	.00-.10	.10-.20	.18-.50	.50-.90	.75-1.00
9	.00-.10	.10-.33	.33-.66	.66-.90	.90-1.00
10	.00-.10	.10-.30	.30-.50	.50-.75	.75-1.00
11	.00-.10	.11-.25	.26-.50	.51-.75	.76-1.00
12	.00-.20	.20-30	.30-.45	.40-.65	.65-1.00
13	.00-.15	.15-.35	.35-.55	.55-.85	.85-1.00
14	.00-.20	.20-.40	.40-.60	.60-.80	.80-1.00
15	.20-.30	.40-.50	.50-.65	.65-.80	.80-1.00
16	.00-.10	.11-.35	.36-.79	.80-.94	.95-1.00
17	.05-.10	.10-.25	.25-.40	.35-.65	.60-1.00
18	.05-.15	.16-.25	.26-.75	.76-.85	.86-1.00
19	.00-.05	.05-.15	.15-.70	.70-.85	.85-1.00
20	.05-.10	.10-.25	.20-.35	.35-.55	.50-1.00
21	.05-.15	.16-.35	.36-.65	.66-.80	.81-1.00
22	.05-.15	.15-.25	.25-.40	.40-.55	.50-1.00
23	.10-.25	.20-.40	.40-.65	.65-.80	.75-1.00
24	.00-.10	.10-.20	.20-.35	.35-.65	.65-1.00
25	.00-.14	.11-.25	.19-.45	.40-.70	.65-1.00
26	.00-.15	.16-.29	.30-.70	.71-.90	.91-1.00
27	.00-.10	.05-.20	.20-.40	.40-.60	.50-1.00
28	.00-.06	.05-.27	.20-.50	.50-.80	.80-1.00
29	.10-.18	.15-.22	.18-.30	.30-.50	.50-1.00
30	.00-.05	.06-.15	.16-.70	.71-.75	.76-1.00

The data represents the percent range of uncertainty associated with each linguistic label. From the respondent viewpoint the percent range of uncertainty represents their opinion of the increase in technical difficulty imposed by the definitions of the labels. Technical Uncertainty definitions for each linguistic label are found in Appendix B. From a context specific viewpoint the methodology takes advantage of the experience of each respondent in regard to their opinion of "increasing technical difficulty" relative to a weapon system program where virtually no technical difficulty existed.

Table 3.3-2

Schedule Uncertainty Fuzzy Membership Sets

Schedule Input Data By Respondent

Linguistic Variables	Very Low	Low	Medium	High	Very High
1	.05-.10	.10-.25	.20-.55	.50-.85	.80-1.00
2	.03-.15	.25-.45	.45-.70	.65-.85	.85-1.00
3	.05-.20	.20-.40	.40-.60	.60-.80	.80-1.00
4	.00-.10	.05-.20	.15-.50	.70-.95	.85-1.00
5	.00-.05	.05-.20	.45-.65	.65-.85	.80-1.00
6	.05-.12	.10-.25	.20-.60	.50-.75	.75-1.00
7	.03-.10	.05-.18	.15-.30	.30-.50	.50-1.00
8	.00-.03	.02-.08	.08-.30	.25-.50	.50-1.00
9	.00-.10	.10-.40	.40-.60	.60-.80	.80-1.00
10	.00-.15	.15-.35	.35-.60	.60-.80	.80-1.00
11	.00-.20	.21-.40	.41-.60	.61-.80	.81-1.00
12	.00-.05	.05-.15	.15-.25	.30-.50	.55-1.00
13	.00-.15	.10-.35	.35-.60	.60-.80	.80-1.00
14	.00-.25	.25-.50	.50-.75	.75-.90	.90-1.00
15	.10-.20	.15-.40	.40-.60	.60-.85	.85-1.00
16	.00-.10	.11-.35	.36-.79	.80-.94	.95-1.00
17	.05-.15	.10-.30	.25-.45	.40-.60	.55-1.00
18	.05-.15	.16-.25	.26-.75	.76-.85	.86-1.00
19	.00-.05	.05-.15	.15-.60	.60-.80	.80-1.00
20	.05-.10	.10-.25	.25-.35	.35-.55	.50-1.00
21	.05-.15	.16-.40	.41-.70	.71-.85	.86-1.00
22	.05-.10	.08-.15	.15-.25	.20-.35	.35-1.00
23	.05-.15	.10-.30	.30-.65	.65-.85	.85-1.00
24	.00-.10	.10-.20	.20-.35	.35-.65	.65-1.00
25	.05-.15	.12-.25	.20-.50	.45-.70	.65-1.00
26	.00-.10	.11-.30	.31-.60	.61-.79	.80-1.00
27	.00-.10	.05-.20	.20-.40	.40-.60	.50-1.00
28	.00-.03	.03-.10	.10-.35	.40-.60	.70-1.00
29	.01-.10	.10-.19	.15-.30	.30-.60	.55-1.00
30	.00-.10	.11-.20	.21-.30	.31-.40	.41-1.00

The data represents the percent range of schedule uncertainty associated with each linguistic label. From the respondent viewpoint the percent range of uncertainty represents their opinion of the increase in schedule difficulty imposed by the definitions of each label. Schedule Uncertainty definitions for each linguistic label are found in Appendix B. From a viewpoint of being context specific the methodology takes advantage of the experience of each respondent in regard to their opinion of "increase in contract time/duration" relative to a weapon system program where virtually no schedule difficulty existed.

Development of the methodology required computation of the degree of membership in each of the five fuzzy sets. Review of the data in table 3.3-1 and 3.3-2 indicates each fuzzy set, five fuzzy sets for technical and five for schedule uncertainty, had thirty input values. The possible uncertainty percentages of the thirty input values ranged from zero to 100 inclusive. A program was written in QuickBASIC to account for the membership in each fuzzy set for every percent from zero to 100. Appendix D contains code for the program DEGMEMB.BAS which generated the tabular output and the number of occurrences of each percent uncertainty in each of the five fuzzy sets for technical and schedule uncertainty. Appendix E contains search pattern algorithms used to generate triangular and trapezoidal least square error fit for technical and schedule fuzzy membership sets. Although the literature thoroughly documents use of triangular and trapezoidal functions for fuzzification of data, this methodology investigates the usefulness of these functions in the specific context of a "range of survey data" related to the weapon system acquisition process.

Plots of the degree of membership in each set as a function of X, f(x), was generated by developing order pairs of integers representing every percent interval from 0.00 to 1.00. The result was a "stair-step" graph, with a zero slope at all points except where the value increased or decreased at certain multiples of 0.01. A set of approximately 200 ordered pairs were required to enable a line graph to produce the desired result. The ordered pair format was based on one percent intervals containing the number of survey responses for each percent interval ;

(0.00, 18), (0.01, 18)
(0.01, 19), (0.02, 19)
(0.02, 19), (0.03, 19)
(0.03, 18), (0.04,18)

(0.04, 18), (0.05,18)
(0.05, 23), (0.06,23)

Appendix D contains the QuickBASIC program which generated the entire set of ordered pairs for five membership sets for each of technical and schedule uncertainty. The plot constructed by generation of the ordered pairs is presented in Figure 3.3-3.

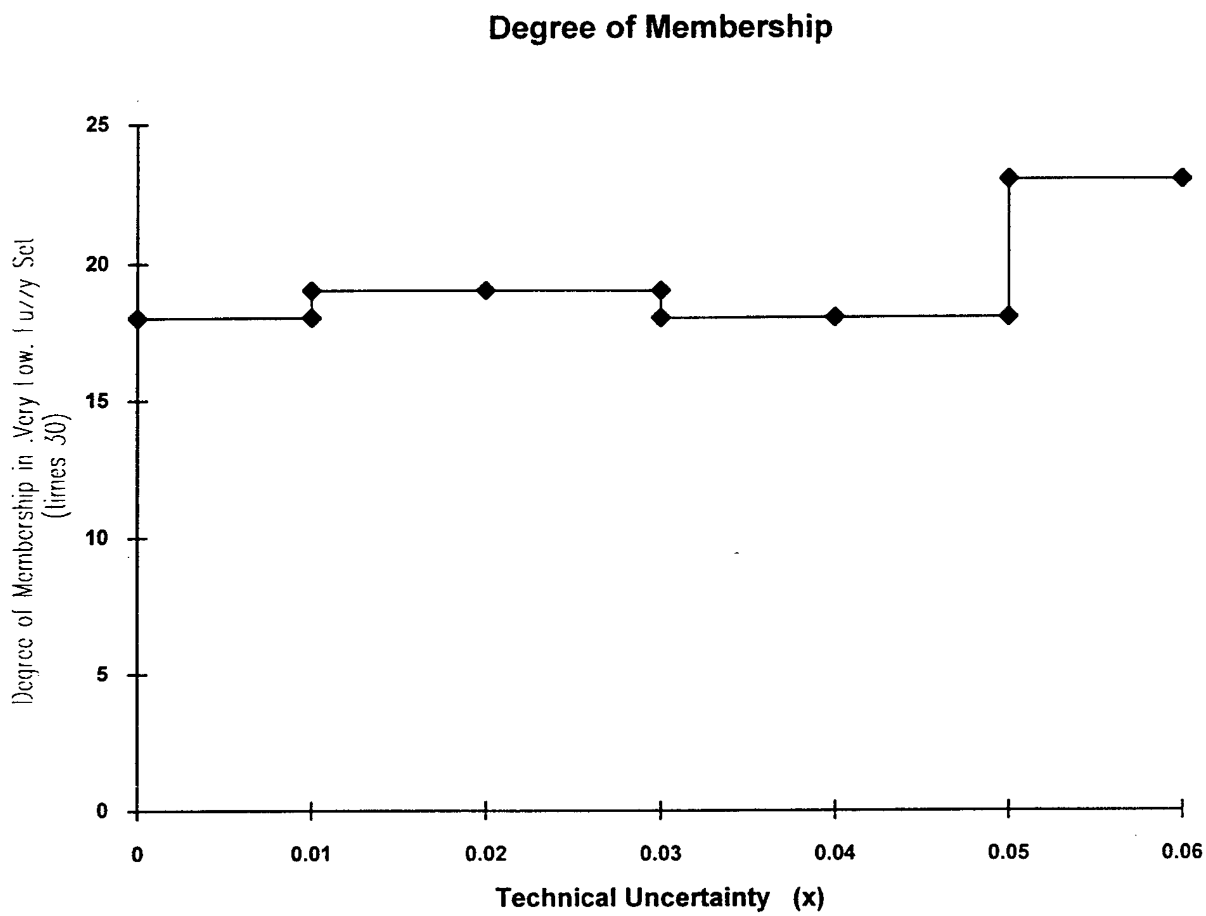


Figure 3.3-3 Degree of Membership in Very Low Technical Uncertainty
Generated by Plots of Order Pairs

Graphical output of the program clearly defined a degree of membership histogram depicting overlap of each fuzzy set with the membership areas of adjacent fuzzy sets for technical and schedule uncertainty. The histograms in Figures 3.3-4 and 3.3-5, suggest the possibility of triangular and/or trapezoidal functions. According to Schwartz and Klir (1992:33) trapezoidal shapes are common in most current industrial applications. Voit (1993:28) mentions "often simple shapes such as trapezoids and triangles ...define membership within fuzzy sets...triangles and trapezoids are the most popular and have proven to be effective and efficient." McCauley (1994:49) and Lindh (1993:36) agree that although many shapes are used for membership functions, Gaussian or trapezoid forms are common.

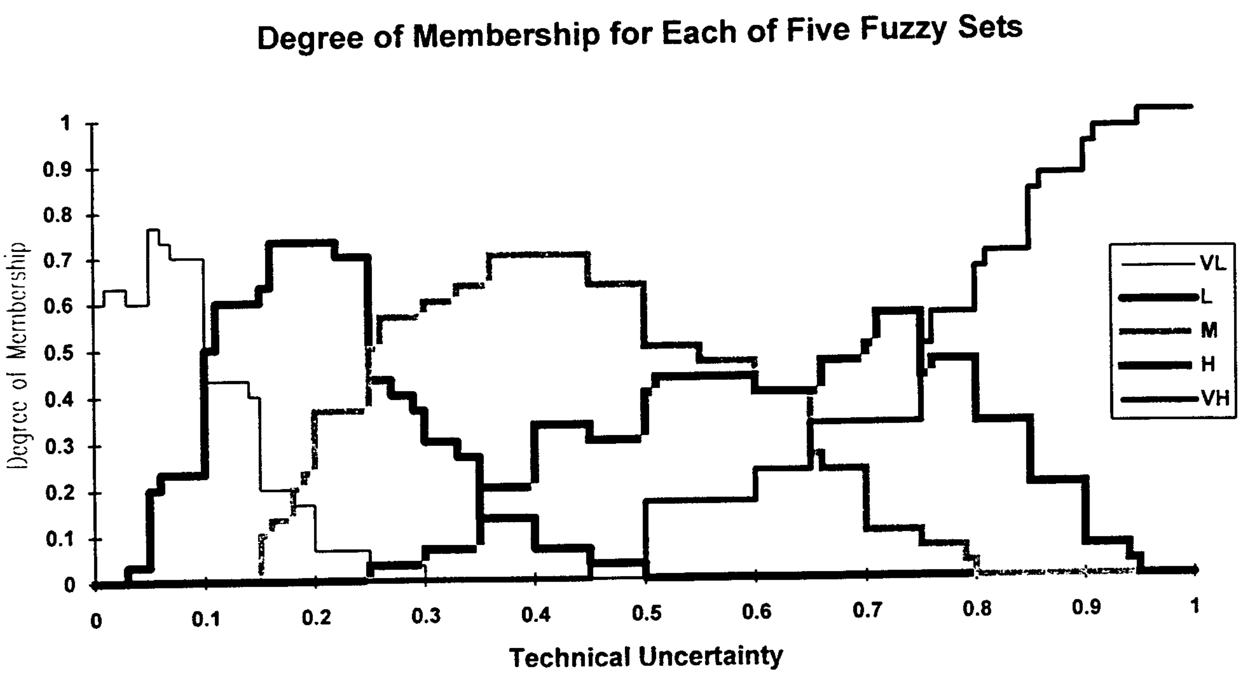


Figure 3.3-4
Degree of Membership for Each of Five Fuzzy Sets for Technical Uncertainty

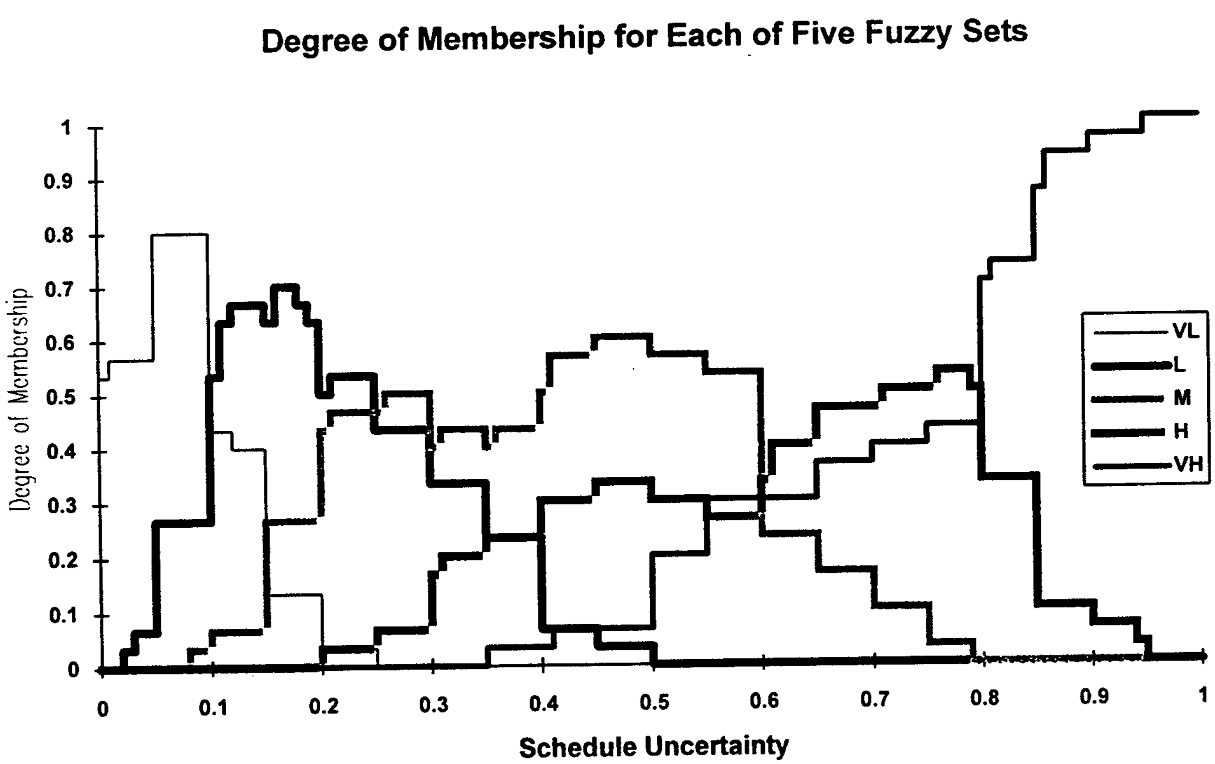


Figure 3.3-5
Degree of Membership for Each of Five Fuzzy Sets for Schedule Uncertainty

Histogram plots for each of the five fuzzy memberships in both technical and schedule uncertainty were characterized using both triangular and trapezoidal functions. Analysis of each functin was necessary because no research was found which specifically addressed data ranges for linguistic variables generated from survey questionnaire input. To assist in the characterizing of the parameters for each function a pattern search program was written to derive values of the specified parameters which would yield a minimum total squared error. Appendix E contains the pattern search algorithm for triangular functions and Appendix F contains the pattern search algorithm for trapezoidal functions.

Each triangular form was defined by specifying four values:

x_0: the x (uncertainty) value at left end of triangle base
x_1: the x value for the apex of the triangle
x_2: the x value for the right end of the triangle base
y_1: the y (degree of membership) value for the apex

Figure 3.3-6 depicts the degree of membership and a first guess fit of the curve to the data for the Medium Technical Uncertainty Fuzzy Set.

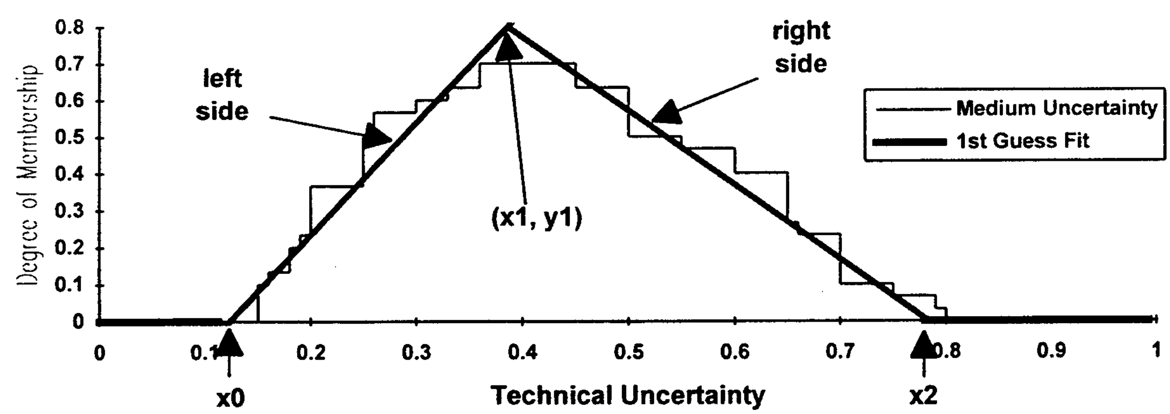


Figure 3.3-6 Degree of Membership for the Medium Technical Triangular Function Uncertainty Fuzzy Set

Two approaches were considered for fitting a triangular function: (1) fit to the line segments which make up the horizontal slope of the curve or (2) fit to the midpoints of the intervals. Given that 100 intervals were available from the data generated via the ordered pairs a fit to the midpoints of the intervals was determined suitable. Fitting to the midpoints of the intervals required ordered pairs;

(.005, 18/30), (.015, 19/30), (.025, 19/30), (.035, 18/30), etc. for all pairs.

The methodology requires a function whose graph consists of four line segments: (1) where Y (the membership function) equals zero for X values less than that of the left end of the base of the triangle, (2) where Y increases linearly from zero to a maximum value (the left side of the triangle), (3) where Y decreases linearly from the maximum value to zero (the right side of the triangle, and (4) where Y equals zero for X values greater than that of the right end of the base of the triangle. The "very low" and "very high" fuzzy sets would begin at zero and end at maximum values respectively. The horizontal line segments to the left and to the right of the base of the triangle lie along the X-axis where;

$$Y = 0$$

The left side of the triangle lies along the line segment that passes through (x_0, 0) and (x_1, y_1) where;

$$Y = \frac{y_1}{x_1 - x_0}x - \frac{x_0 y_1}{x_1 - x_0}$$

The right side of the triangle lies along the line that passes through (x_1, y_1) and (x_2, 0) where;

$$Y = -\frac{y_1}{x_2 - x_1}x + \frac{x_2 y_1}{x_2 - x_1}$$

A QuickBASIC program was developed to allow a trial-and-error search for values of the four parameters x0, x1, x2, and y1. The objective was to obtain the minimum total squared error. Given a set of values for the parameters the program computes;

$$Q = \sum_{i=1}^{n} (y_{a_i} - y_{p_i})^2$$

Where;

Q = the sum of the squared errors, residual variability

n = the number of points to use in the curve fit (i.e. 100)

y_{a_i} = the actual y value (degree of membership for ith point)

y_{p_i} = the predicted y value for the ith point

As a measure of "goodness of fit" the sum of the squares of the differences between the various y-values and their mean was calculated;

$$S = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

A comparison of the "goodness of fit" was made for each type function for every fuzzy membership set for both technical and schedule uncertainty. This data analysis sufficiently justified use of triangular (or trapezoidal) functions as a basis for characterizing fuzzy membership functions for technical and schedule uncertainty.

Note that Q, as defined above, is a measure of the residual variability in the data after application of the model. As a measure of overall variability of the data the S is calculated as the squared sum of the differences between the observed values and their means. The difference S-Q is a measure of the variability accounted for by the model and Q is a measure of variability not accounted for by the model therefore the closer ρ is to unity the better the fit.

$$\rho = \frac{S-Q}{S}$$

A triangular fit pattern search program TRIFIT.BAS yields a minimum value of Q given a first-guess value for each of the parameters. TRIFIT.BAS appears in Appendix E. The search algorithm finds a local optimum which is estimated to be acceptable given the amount of error (Q) and the "goodness of fit" parameter used. Results are depicted in Chapter 4.

The pattern search program was also used to find the apparent best-fit trapezoidal functions for comparison with results of the triangular functions. The trapezoid functions were fit to interval midpoints just as was accomplished with the triangular functions. Each trapezoidal function was defined by specifying five values;

x_0: the x (uncertainty) value at left end of the trapezoid base

x_1: the x value for the left end of the top of the trapezoid

x_2: the x value for the right end of the top of the trapezoid

x_3; the x value for the base of the trapezoid

y_1: the y (degree of membership) value for top of the trapezoid

Figure 3.3-7 depicts the degree of membership for Medium technical uncertainty fuzzy set for the trapezoid function. The methodology requires a function whose graph consists of five line segments: (1) where Y (the membership function) equals zero for values less than that of the left end of the base of the trapezoid, (2) where Y increases linearly from zero to a maximum value (the left side of the trapezoid), (3) where Y remains constant at the maximum value (top of the trapezoid), (4) where Y decreases linearly from the maximum value back to zero (right side of the trapezoid) and (5) where Y equals zero for values greater than that of the right end of the base of the trapezoid.

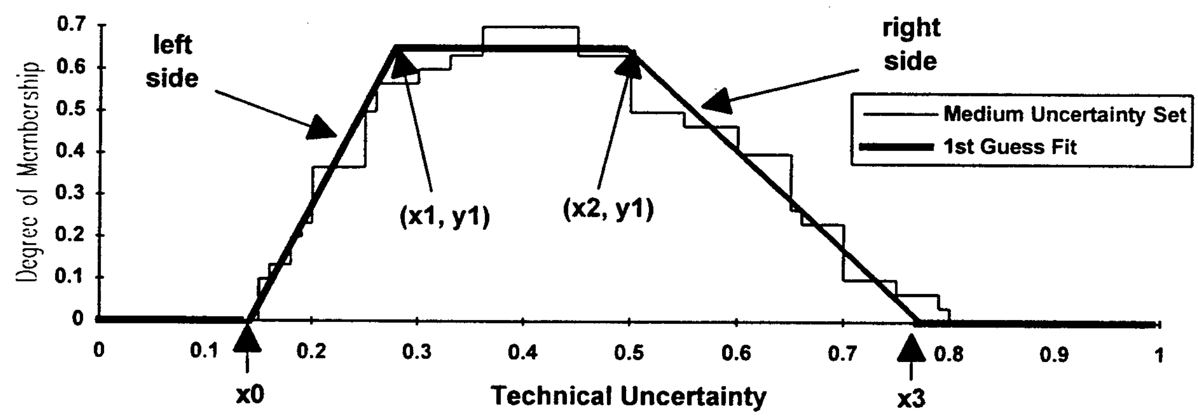


Figure 3.3-7 Degree of Membership for Medium Technical Trapezoidal Uncertainty Fuzzy Set

Trapezoid functions were not fit for the "very high" uncertainty category because the data sets with triangular functions had included functions that were virtually trapezoidal and reached the maximum value at the right end and did not decrease.

The horizontal line segments to the left and to the right of the base of the trapezoid lie along the x-axis where;

$$Y = 0$$

The left side of the trapezoid lies along the line that passes through (x_0, 0) and (x_1, y_1) where;

$$Y = \frac{y_1}{x_1 - x_0} x - \frac{x_0 y_1}{x_1 - x_0}$$

The top of the trapezoid lies along the line;

$$Y = Y_1$$

The right side of the trapezoid lies along the line that passes thorough (x_2, y_1) and (x_3, 0) where;

$$Y = -\frac{y_1}{x_3 - x_2} x + \frac{x_3 y_1}{x_3 - x_2}$$

A trapezoidal least square error fit pattern search program TRIFIT1.BAS yields a minimum value of Q given a first-guess value for each of the parameters. TRIFIT1.BAS appears in Appendix F. The search algorithm finds a local optimum which is estimated to be acceptable given the amount of error (Q) and the "goodness of fit" parameter used. Results are depicted in Chapter 4.

3.4 Simulation Approach to Fuzzification

A thorough review of available literature did not yield a case in which simulation was employed in substitute of a mathematical approach to fuzzification. The methodology in this section of the research applies simulation to fuzzify membership sets characterized by triangular or trapezoidal functions. The simulation runs on a Sun Workstation using a multi-tasking GNUAda translator (GNAT) and an Ada compiler that allows a source program to run on multi-platforms.

Using a simulation approach the data used could be the predicted cost displayed in the CES of a weapon system specific cost estimate. Research presented here utilizes top level RDT&E and Procurement cost but in an actual acquisition environment uncertainty would be applied to many combinations of detailed level work breakdown structure (WBS) over-runs and under-runs which eventually represent a projects' final cost. Instead of using cost, the methodology presented here uses the percent ranges of uncertainty and processes the data using a fuzzy logic algorithm.

For purposes of exactness a distinction between WBS and CES must be made. A WBS is a product-oriented family tree depicting interrelated work elements (work packages) for a total prime mission system. The CES, aligned closely with the new defense acquisition management process, is an appropriation oriented (RDT&E, Procurement, Military Construction, Military Personnel, Operations and Maintenance, and Defense Business Operations Fund) structure that supports the integration of cost analysis with the Planning, Programming, Budget Execution System (PPBES). The WBS may be composed of elements very similar to elements within the CES.

A simulation representing the CES can be run many times by varying the "chance" of a cost over-run for each cost element. To introduce a degree of realism into the simulation this methodology uses input from subject matter experts which was gleaned via the survey questionnaire. The questionnaire solicited expert opinion to approximate an unknown: the possibility of increases in technical and schedule uncertainty related to a specific weapon system acquisition program. In effect this methodology used aspects of the modified Delphi Technique. Quantification of uncertainty into percent ranges was required for all linguistic variables: very low, low, medium, high and very high. Inherent in the methodology was the assumption that if every aspect of an acquisition program went well there would be very low uncertainty and if major problems developed there would be the possibility of very high uncertainty. The "very low " category included the possibility of zero uncertainty while the "very high" category included the possibility of each CES doubling in cost when a "crisp" solution was derived during the defuzzification process. When all inputs were collected via the questionnaire the relationship between uncertainty and impact on predicted cost was "acted out" through the simulation. The ultimate objective is to apply a percent increase to top level predicted costs based on the fuzzy inputs developed through the simulation. Appendix G contains the Fuzzify Simulation Program and test case data.

Primary aspects of the fuzzy logic algorithm are the category package and the logic to manipulate the category. For the category, a Category Name Type is specified as a string of 30 characters representing an English name description. Each category is specified by an upper and lower boundary generated through the characterization process and each data point within the category represents a degree of membership in that specific category.

A set of functions and procedures are defined into one task called Category Task Type. Basic functions performed are: Set_Name, Get_Name, Set_Range, Get_Low_Value, Get_Mid_Value, Get_High_Value. Table 3.3-1 summarizes the functions performed by the Category Task Type of the fuzzy logic algorithm.

Set_Name	Allows the category user to give the new name a category name if needed.
Get_Name	Allows the user to request the category name from the category task, helpful when numerous categories exist.
Set_Range	Allows the user to define upper and lower boundaries.
Get_Low Value Get_Mid_Value Get_High_Value	Allows the category user to retrieve the upper, lower, or mid points of a category range.
Get_Share	Presents a solution to the question "where does the data point fall within this category?"

Table 3.4-1 Category Task Type Functions

Performance of the Category Task Type functions are code in the Category Package Body that contains program implementation details. The Get_Share function is calculated using the triangular method. Incoming data values are tested to determine inclusion or exclusion in membership sets. When data values are included in multiple fuzzy membership sets a routine in the fuzzy program selects a category using a random process. Input/output routines provide the capability of manipulating English names and data values. The methodology for text processing requires creation of predefined templates for definitions of: How_Many_IO, Range_IO, Number_of_Categories, Number_of_Data_Values and Test_Data. The random number generator utilizes templates for subtype and package procedures for selecting membership categories which fit defined criteria. A "begin" routine resets the random number generator, opens test data files and processes data.

The PROCESS DATA section dynamically brings each category into existence and assigns names and boundaries for each category based on descriptions found in the Test_Data file. This file identifies number of categories to be fuzzified and how many data values are to be processed in each category via subsequent "loops" of the simulation. Data, in the form of a "crisp" value, is read in and calculations initiated to determine where the value falls within each defined category. Given the overlapping boundaries of the membership sets it is likely the value will fall within two or perhaps three membership sets. Later in this research additional methodology will describe how this input value will be obtained from "earned value" data generated by schedule performance measurement defined in the DOD Cost/Schedule Control Systems Criteria (CS/CSC). A "running total" of distance along the X axis of each membership set is tabulated by spreading the bases of the fuzzy membership sets defined by the triangular or trapezoidal function. The spreading of the bases along the X axis in effect takes out the "overlap" of the sets and generates a cumulative total value. The "running total" is then multiplied by a random number between zero and one. The sum of the spread values, possibly greater than one, represent the individual contribution of each category's base length defined by triangular or trapezoidal curve fit methodology for characterization of fuzzy membership sets. This routine prepares for the final output of PROCESS DATA section: a single fuzzy membership set. As the simulation loops through each category's contribution to the total range, the Random_Number_In_Total ultimately falls into only one of the categories.

The last step in the simulation is to display the weight of each category's contribution. At this point all simulation loops are finished and the Test_Data file is closed. The combined total of all weights will equal unity by definition. The result of the simulation is a selection of category and a display of the degree of membership of each value in that particular fuzzy membership category. When dealing exclusively with technical or schedule uncertainty but not both simultaneously this simulation provides a fuzzy output category which can be defuzzified. If however the problem is one of technical and schedule impacts we must have a methodology to stipulate an ultimate fuzzy output if the simulation for technical indicates one category and the simulation for schedule represents a different category. This is a problem of combining fuzzy membership sets which requires a rule base.

3.5 Uncertainty Rule Base

An uncertainty rule base provides a mechanism for combining conflicting fuzzy outputs (Sperry, 1993: 33-34) generated from the simulation. A technical and schedule uncertainty rule base was constructed in the form of a look up table matrix that allows combining fuzzy membership sets and assigning output rule strengths based on a minimum or maximum (min-max) function decision routine. Inputs to the matrix represent the outputs of the simulation which in this specific case are the linguistic labels for technical and schedule uncertainty. Output of the rule base is governed by a decision routine which also assigns a linguistic label. Rules were developed in the form of "if"-"and"-"then" and "if"-"or"-"then" in which a rule executes to a "truth value" represented by either a min or a max function.

The "if" portion of a rule has a condition called an "antecedent" and the "then" portion of a rule has a condition called a "consequence" both of which correspond to fuzzy membership sets. A min function decision routine, represented by an "and" operation, is defined such that rule evaluation output is assigned the value of its least true antecedent. A max function decision routine, represented by an "or" operation, is defined such that rule evaluation output is assigned the value of the most true, or strongest, antecedent.

Described in terms of implication a proposition, antecedent and consequence, is symbolically stated in "if"-"or"-"then" terminology as $\alpha \Rightarrow \beta$. Applying set theory terminology the methodology of the rule base uses the concepts of containment, intersection and union, of fuzzy membership sets. The intersection of fuzzy sets *A and B* with membership functions: $f_a(x)$ *and* $f_b(x)$ is a fuzzy set *C* written: $C = A \cap B$. In abbreviated form the intersection operation is: $f_c = f_a \wedge f_b$ where $\wedge$ indicates a min function. Representation of a union of two fuzzy sets *A and B* with membership functions $f_a(x)$ *and* $f_b(x)$ is a fuzzy set C written: $C = A \cup B$. In abbreviated form the union operation is written $f_c = f_a \vee f_b$, where $\vee$ indicates a max function.

Rule evaluation results in a fuzzy output. The methodology used in this research applies the fuzzy output (in the form of a "crisp" solution after defuzzification) to the top level of the cost estimate. During development of a cost estimate each of the individual elements within the CES could be assessed technical and/or schedule uncertainty. In that case rather than applying fuzzy output to the top level it would be applied to any (or many) of the sub-level cost elements within the CES. The top level of a life cycle cost estimate would be CES 0.0 Total as depicted in Appendix A and the top level of an estimate for RDT&E or Procurement funded elements would be 1.0 or 2.0 respectively. This "top level" CES is the focus for application of the fuzzy approach to quantifying technical and schedule uncertainty in this research. During the actual development of a cost estimate any of the individual elements within the CES could have schedule uncertainty, technical uncertainty, or both, applied. Note that the CES displayed in Appendix A goes to the third level of indenture called a summary level. It is common in a cost estimate for most of these individual cost elements to have sub-elements that are indentured as much as six levels below summary, to the ninth level. This degree of indenture in effect "tailors" the CES for a particular cost estimate. The fuzzy output of rule evaluation in this methodology applies to the top level of an estimate.

SCHEDULE \ TECHNICAL	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
VERY LOW	VL	VL	VL	VL	VL
LOW	VL	L	L	L	L
MEDIUM	VL	L	M	M	M
HIGH	VL	L	M	H	H
VERY HIGH	VL	L	M	H	VH

Table 3.5-1 Fuzzy Rule Strengths for Min Function

		TECHNICAL				
		VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
SCHEDULE	VERY LOW	VL	L	M	H	VH
	LOW	L	L	M	H	VH
	MEDIUM	M	M	M	H	VH
	HIGH	H	H	H	H	VH
	VERY HIGH	VH	VH	VH	VH	VH

Table 3.5-2 Fuzzy Rule Strengths for Max Function

3.6 Defuzzification

The selection of a methodology for defuzzification is the final step in the demonstration of the basic fuzzy approach for this research. Two of the most common techniques are: (1) the composite maximum and the (2) composite moment or centroid (Cox, 1992:61). The output of the uncertainty rule base was a fuzzy set with a linguistic label. The process of defuzzifying derives a "crisp" output which may be applied as a multiplier to the "uncertainty free" portion of a cost estimate to quantify uncertainty. The "crisp" multiplier may be applied to the top level of a cost estimate as in the case of this research or to individual elements within the CES. The centroid technique considers the contribution of all elements of the fuzzy output and is a preferred technique in the literature base (Cox, 1992:61; Voit, 1993:45; Self, 1990:42; Schwartz and Klir, 1992:33). In this research defuzzification will be accomplished via a modified approach to the centroid technique which uses the abscissa value represented by the apex of the triangular function. Figure 3.6-1 illustrates the Very Low and Low technical uncertainty fuzzy set triangular functions developed in Section 3.3 and depicts the variables used in the methodology for deriving a "crisp" solution. Discrete values of the variables represented in the figure were provided in the same Section. The area of the triangle in the negative region of the abscissa does not impact calculation but is shown for completeness.

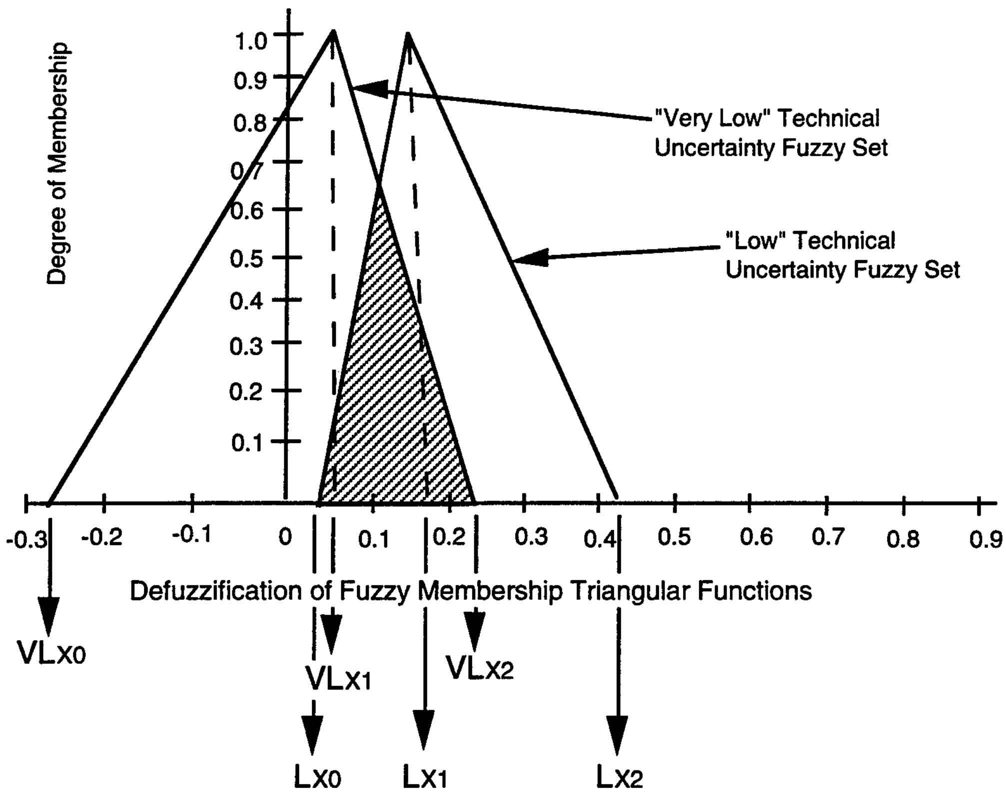


Figure 3.6-1 Triangular Function Variables Associated With Defuzzification

Derivation of a "crisp" solution methodology utilizes the triangular function discrete abscissa value in combination with the strengths of the contribution weights generated by the simulation (degree of membership). The methodology will be described using the "Low" and "Medium" triangular functions. Figure 3.6-2 depicts the triangular discrete variables.

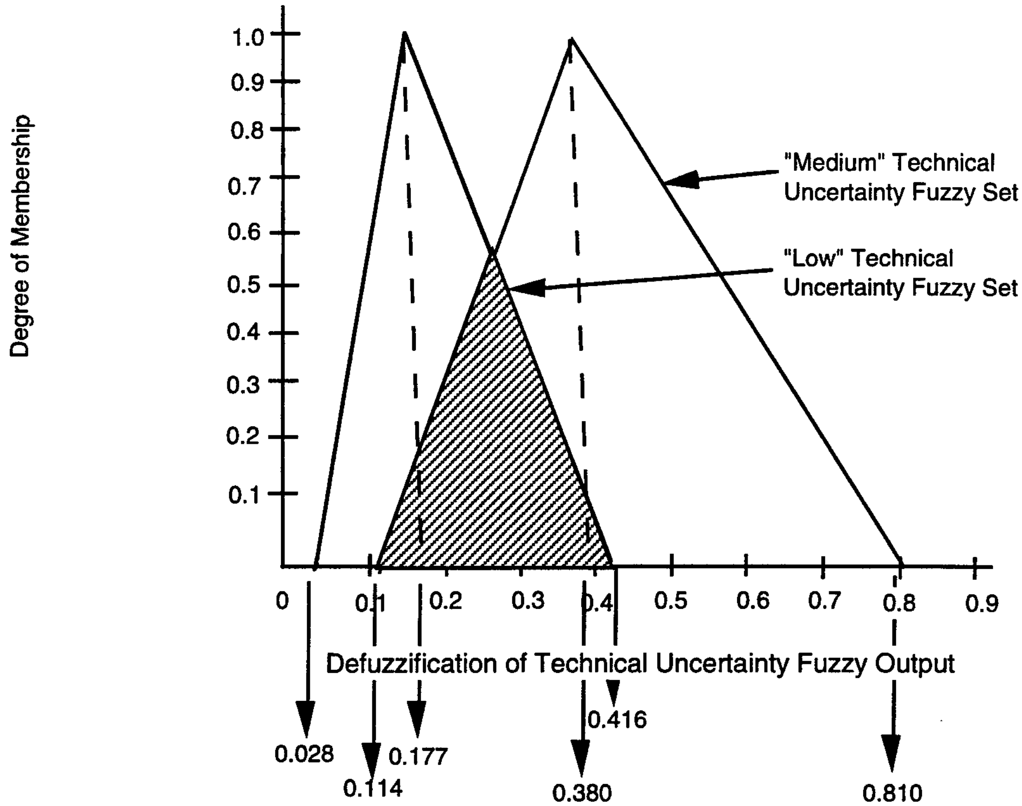


Figure 3.6-2 Triangular Function Discrete Variables of Overlapping Technical Uncertainty Fuzzy Sets

These variables in combination with simulation contribution weights provide inputs for the methodology to derive a weighted average "crisp" solution which will be used as a multiplier of the "uncertainty free" portion of the top level cost estimate for quantification of technical uncertainty.

In general the methodology is specified:

$$D_{wt_i} = \sum_1^n (C_{mt_i} * X_{t_i})$$

Where;

D_{wt_i} = Weighted average of the "crisp" technical uncertainty solution

C_{mt_i} = Weighting contribution from simulation

X_{t_i} = Abscissa value of the apex of triangular function

The "crisp" solution may be calculated either from an input value for uncertainty derived from "earned value" data via a performance measurement system or from a qualitative assessment of uncertainty. A qualitative assessment may derived from a Delphi or modified Delphi methodology from qualitative input such as assessments from an expert panel (low, medium etc.)

In general the methodology is specified:

$$D_{qt_i} = \sum_1^n (X_{qt_i} * R_{qt_i})$$

Where;

D_{qt_i} = Weighted average, qualitative technical input

X_{qt_i} = Abscissa value of the apex of the triangular function for technical uncertainty

R_{qt_i} = Ratio of the range represented by individual technical fuzzy membership set to the total of the ranges of the sets specified by the Delphi technique

3.7 Earned Value Concepts

A fundamental requirement of acquisition management for major weapon systems is visibility into progress made on contracts. Earned value management ensures government visibility into a contractor cost and schedule performance data which properly relates cost, schedule and technical accomplishments. The objective of the integrated cost schedule management system defined in the C/SCSC is for the government to be able to rely on timely and auditable reports for determining product-oriented contract status (Cost Schedule Management Guide, 1995:1-2).

C/SCSC provides the methodology for capturing data relative to significant technical and schedule variances during specified periods of time. Generally the reporting level of the contract work breakdown structure (CWBS) is level 3, the summary level. High cost and high risk items can be tracked at a lower level. All variances that breach a stated percentage or dollar threshold are tracked closely. The tracking methodology includes cumulative budgeted cost for work scheduled (BCWS), cumulative budgeted cost of work performed (BCWP), actual cost of work performed (ACWP), and estimated cost at completion and budgeted cost at completion.

Analysis of BCWS, BCWP and ACWP provides insight into schedule and cost variance. Variance may be analyzed at the level of effort within the CWBS including work package levels to gain insight into areas of technical difficulty. Assessment of the information provided in the C/SCSC enables the determination of the number of weeks (or months) of schedule variance. Unfavorable schedule and technical variances may be quantified into a percent of contract duration/level of technical effort and applied as a percent factor for input to the fuzzy simulation algorithm.

3.8 Cost and Price Data

Historical data in the form of government cost estimates for specific contracts and subsequently negotiated prices for the respective contracts were collected for this research. Data represents twenty six separate contracts with one of the contracts having twenty one major contract modifications for a total of forty six separate contract actions with an independent government cost estimate for each.

Top level government cost estimate data were obtained in two parts: (1) the portion of the cost estimate which did not include uncertainty ("uncertainty free") and (2) that portion of the cost estimate which did quantify uncertainty. These two elements combined, "uncertainty free" plus quantification of uncertainty, represent the total cost estimate, independent government estimate (IGE), for each contract. Each procurement action, request for proposal (RFP), requires an IGE as part of the documentation package approval process. Generally the cost elements of these estimates are an abbreviated form of the CES that appears in Appendix A.

This research develops an approach for quantification of uncertainty which may be used as a multiplier to the "uncertainty free" portion of the IGE to derive a "revised" top level cost estimate for each contract. The methodology makes use of the negotiated contract prices to "re-quantify" uncertainty. It is not critical to know if the original quantification of uncertainty was accomplished by detailed analysis of WBS/CES or accomplished at a top level. Since the data were obtained at the top level it is assumed that uncertainty quantification was accomplished at a top level during the IGE process using percentage factors such as will be the subject of the analysis to follow.

The data represent IGEs and contracts which cover a period of several years and have a very wide dispersion in absolute dollar value ranging from $200-300K (thousand) to over $300M (million). Indexing will be used to adjust absolute values to a common base so inflation effects over time are not a factor of the analysis. Dispersion will be dealt with by reducing the deltas in absolute value to percentage deltas by normalizing and then calculating appropriate statistics: mean, variance and standard deviation etc.

Certain assumptions are made relative to available data in order to complete the methodology and accommodate the analysis that follows. A background understanding of the data is important to the assumptions. IGEs are composed of two data points each: one data point is the "uncertainty free" portion of the estimate the other data point is the quantified uncertainty. Negotiated contract prices are represented by one data point, uncertainty is not separately quantified. The contract negotiation process, supported by technical and administrative personnel on government and contractor teams operating at "arms length," provides the expert arena in which a "price" is derived not only for the technical work to be accomplished but inherent in the process within the quantification of "price" is recognition of and quantification of uncertainty. A single contract value results from negotiation but that value includes uncertainty about which both parties recognized and negotiated. Inherent in this methodology is the assumption that the unknown uncertainty portion of the negotiated contract price is represented by the difference between the "uncertainty free" portion of the IGE and negotiated contract price which all expert parties to the negotiation realized included uncertainty. This difference is called the "negotiated uncertainty." Implicit in this assumption is the further assumption that although the "uncertainty free" portion of the IGE was known only to the government team it is the counterpart of the risk free portion of the negotiated contract price. Using this assumption a "negotiated uncertainty percentage" will be calculated as a ratio and normalized;

(Negotiated Price-IGE "uncertainty free")/IGE "uncertainty free"

Appropriate statistics will be computed: mean, variance and standard deviation etc. for comparison with statistics developed for the IGE "Uncertainty percentage." IGE Uncertainty percentages for certain contract prices will be processed through the fuzzy membership sets and defuzzified for a "crisp" solution for each. The "crisp" solutions, "observed" values, are compared to the "truth" values represented by the "negotiated uncertainty percentages." Appropriate statistics will be calculated for the deltas between observed minus expected normalized values. "Goodness of Fit" and significance testing are accomplished to provide a basis for hypothesis testing.

4. ANALYSIS

Analysis of data presented in the methodology encompasses the survey instrument, characterization of the fuzzy membership functions and defuzzifying outputs to derive a "crisp" solution. From the perspective of the hypotheses, quantification and testing is demonstrated.

4.1 Context Specific Survey Instrument

Design of the survey instrument considered several objectives. Primary concern was to develop and administer an instrument which would be applicable to a context specific target group. Within the target group the objectives were two fold: identify respondents currently performing work for a specific weapon system and include multiple functional areas and educational backgrounds. For the instrument itself the objectives were to explicitly differentiate the two subject areas, technical and schedule, of the survey and to accomplish the differentiation in a straight forward manner by constraining overall definitions to a minimum of points of differentiation.

The completed survey instrument was received from thirty personnel of various educational experiences, varying number of years experience with the specific weapon system and various project office functional expertise. The common ingredient was that all personnel were full time employees in professional or managerial categories currently supporting the specific weapon system for which this research was generated. From a perspective of context specificity the goal of the survey instrument was realized.

From an educational background perspective personnel with the following diplomas or degree types completed the questionnaire: high school, Bachelor of Science, Bachelor of Arts, Master of Science, Master of Arts, and Doctor of Engineering. Degree types included: Physics, Engineering (systems, electrical, mechanical, industrial), Math, Computer Science, Logistics, Business Administration, Accounting, Aviation Management, Management Information Systems, History and Psychology. Their number of years experience with this specific weapon system ranged from four to twenty three with a mean of slightly over ten years. Functional areas currently being supported by these personnel included: testing, program management, hardware, software, product assurance, system engineering, system integration, safety, foreign military sales and simulation and analysis. All respondents were current full time professionals supporting the specific weapon system project office of which this research was targeted. Every major functional area defined on the project office organization chart was represented in the survey responses with the exception of Configuration Management. Educational backgrounds of respondents covered a broad spectrum of technical and non-technical fields.

The survey instrument was directed at two areas, technical and schedule uncertainty. Explicit definitions constrained the focus of technical uncertainty to that deriving from impacts to requirements: for schedule uncertainty to that deriving from impacts to contract duration. From these minimum points of differentiation, requirements and contract duration, the definitions of the schedule uncertainty linguistic variables sought to depict those impacts which would as clearly as possible distinguish in the minds of respondents those events that would have "relatively" more schedule impacts while having "relatively" less technical impacts. Twelve descriptors were provided that assisted the respondents in further distinguishing potential events which would constitute relatively more schedule and relatively less technical impacts. The definitions supplemented by the descriptors succeeded in providing the necessary differentiation in a straight forward manner.

4.2 Characterization of Fuzzy Membership Functions

Compilation of the survey input data was the basis for characterization of the fuzzy membership sets. Raw data from the questionnaire was input unsorted into the DEGMEMB.BAS QuickBASIC program. Data presented in Tables 3.2-1 and 3.2-2 are listed by respondent e.g. respondent number one provided input listed for the linguistic variables in row number one (raw unsorted data).

The program TRIFIT.BAS generated a trial and error "pattern search" for the set of technical values of the four parameters: x_0, x_1, x_2, x_3 that yields a minimum value of the sum of the squared errors. Table 4.2-1 summarizes the results for the best-fit triangular functions for technical data.

Parameters for Best-Fit Triangular Functions
For Technical Uncertainty Data

Fuzzy Se	x_0	x_1	x_2	y_1	Q	S	ρ
Very Lo	-0.2534210	0.0671320	0.2190150	0.7250604	.08792462	4.54501096	.9806547
Low	0.0275000	0.1769575	0.4164018	0.7971067	.16935897	5.88915526	.9712422
Medium	0.1139300	0.3792579	0.8100909	0.7872364	.16385444	7.03462223	.9767074
High	0.2081770	0.7309583	0.9533138	0.5708506	.18982908	3.75951109	.9495070
V. High #1	0.4821642	1 (fixed)	(N/A)	1 (fixed)	.30309475	11.80222192	.9743188
V. High #2	0.4979121	1 (fixed)	(N/A)	1.0556102	.26573971	11.80222192	.9774839
V. High #3	0.4995735	0.9704265	$\rightarrow \infty$	1.0000000	.26261135	11.80222192	.9777490

Table 4.2-1

The negative value in the Very Low fuzzy set represents the extension of the left side of the triangle to intersection with the x axis. In practical application negative values would not be used. The minimum value on the x axis would be 0.0 however all values for the parameters are depicted for completeness. At the opposite end of the fuzzy sets the very high curve increases to a maximum membership level as x approaches the value one. At 1.0 and beyond the Very High fuzzy set curve has a continuous slope of zero.

There were at least three potential options for fitting the very high technical uncertainty curve. Option 1 made "very high" a triangular distribution that proceeded along a linear function from $y = 0$ at some x value to $y = 1$ at $x = 1$. Under this option the value of x_2 is irrelevant since only the left side of the triangle is practical. Option 2 considered a triangular distribution which proceeded linearly from $y = 0$ at some x value to whatever value of y at $x = 1$ provides the best fit. Under this option as under Option 1 x_2 becomes irrelevant. Option 3 effectively considered the function a trapezoid with x_2 moved very far to the right ($x_2 = 1.0 * 10^{10}$) allowing the right side of the function effectively to become a line with zero slope extending from the point (x_1, y_1). Note the results of the best-fit value of y_1 approaches one as x_2 approaches ∞. Plots for the apparent best-fit triangular functions are depicted in Figures 4.2-1-2-3-4-5-6 and-7.

Degree of Membership for "Very Low" Technical Uncertainty Fuzzy Set

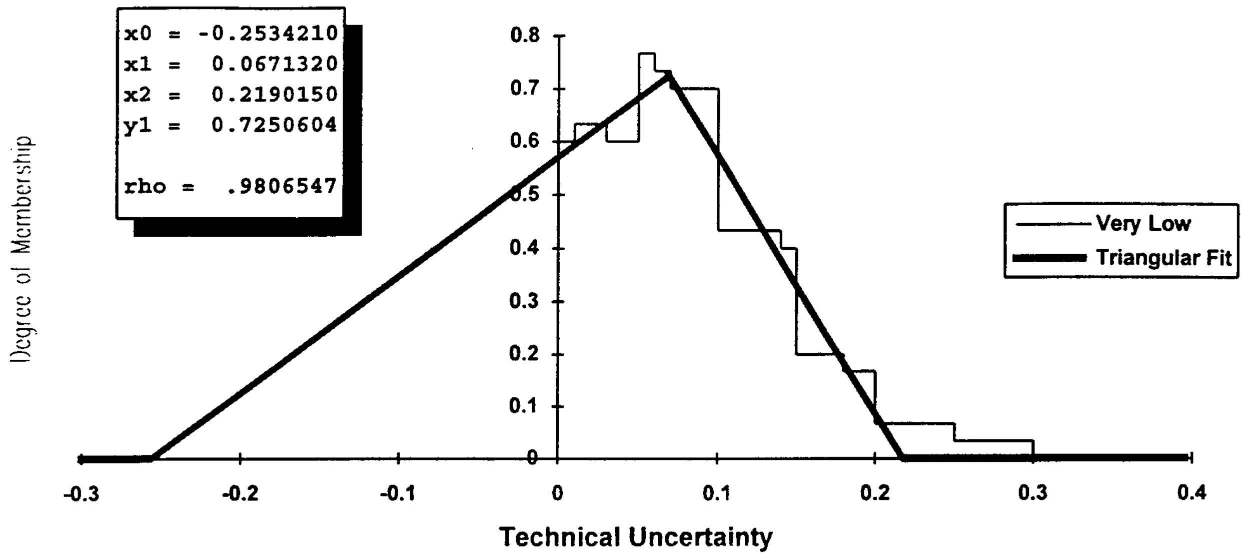


Figure 4.2-1

Degree of Membership for "Low" Technical Uncertainty Fuzzy Set

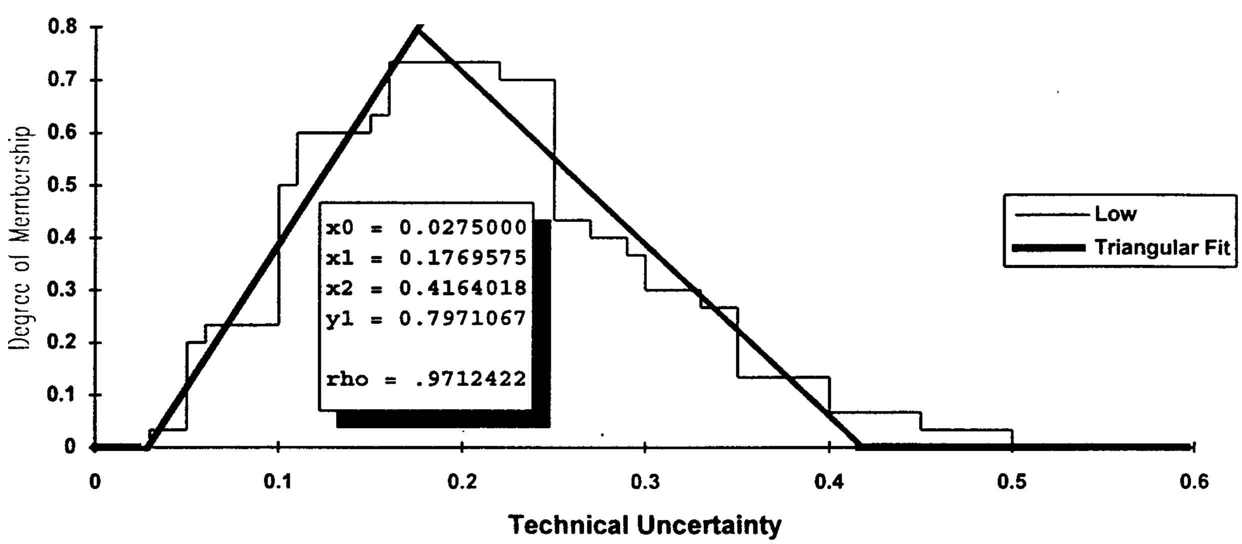


Figure 4.2-2

Degree of Membership for "Medium" Technical Uncertainty Fuzzy Set

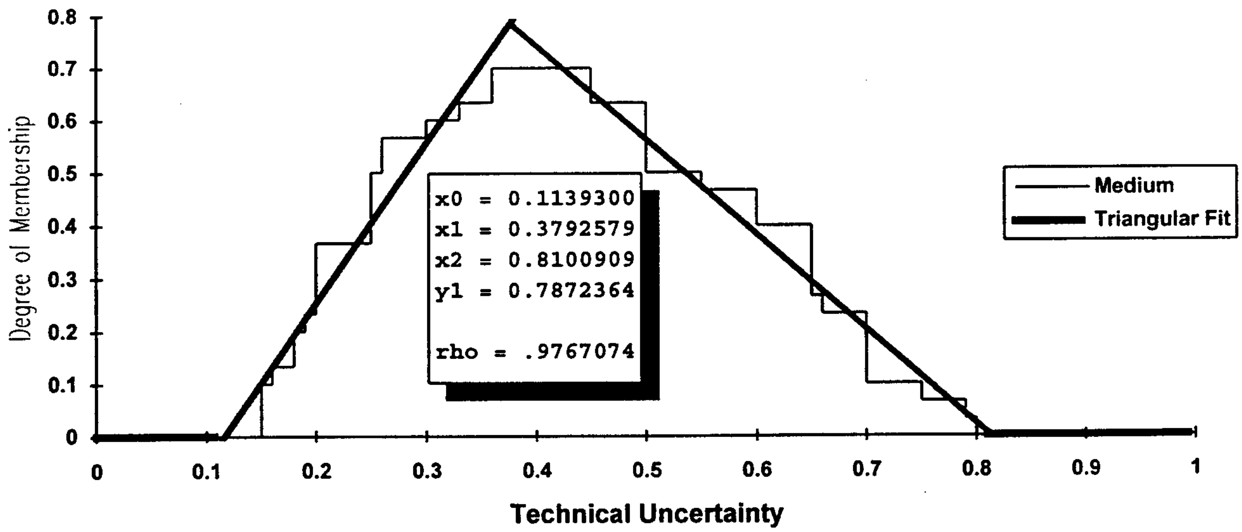


Figure 4.2-3

Degree of Membership for "High" Technical Uncertainty Fuzzy Set

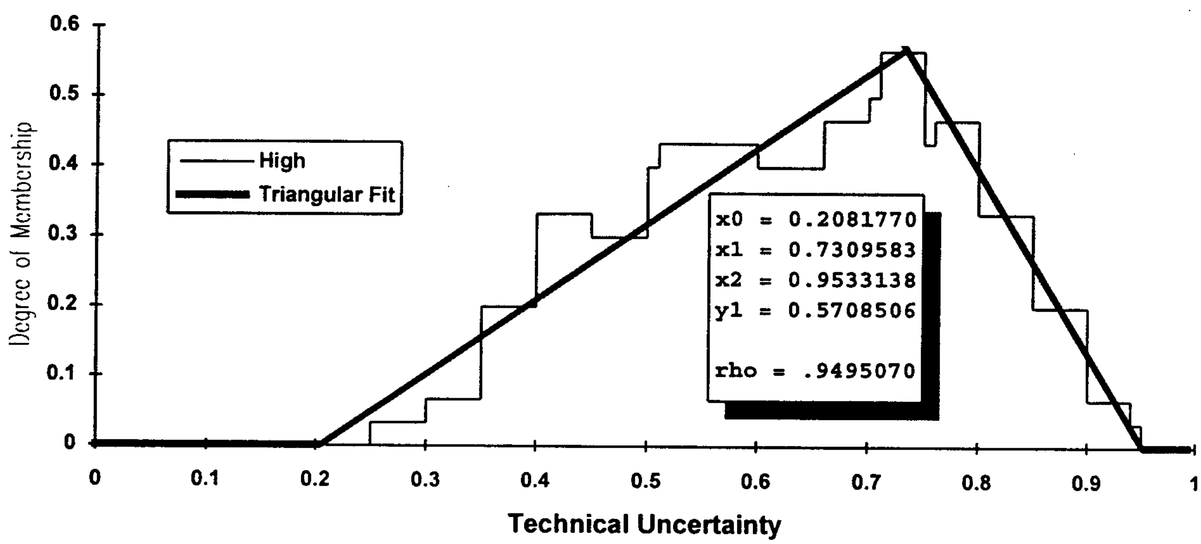


Figure 4.2-4

Degree of Membership for "Very High" Technical Uncertainty Fuzzy Set, with "Triangular" Fit #1

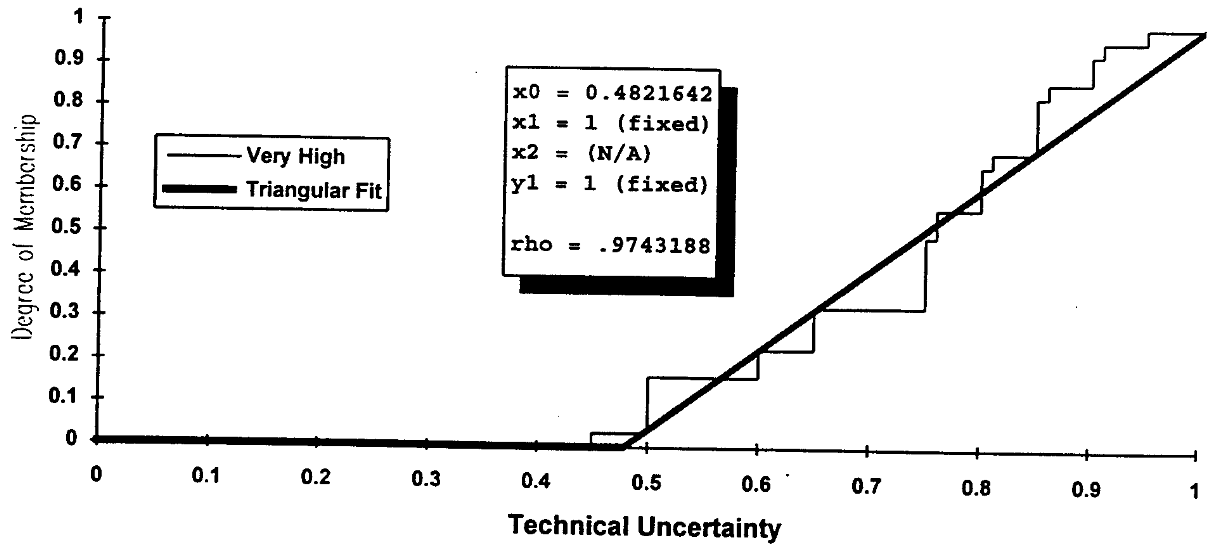


Figure 4.2-5

Degree of Membership for "Very High" Technical Uncertainty Fuzzy Set, with "Triangular" Fit #2

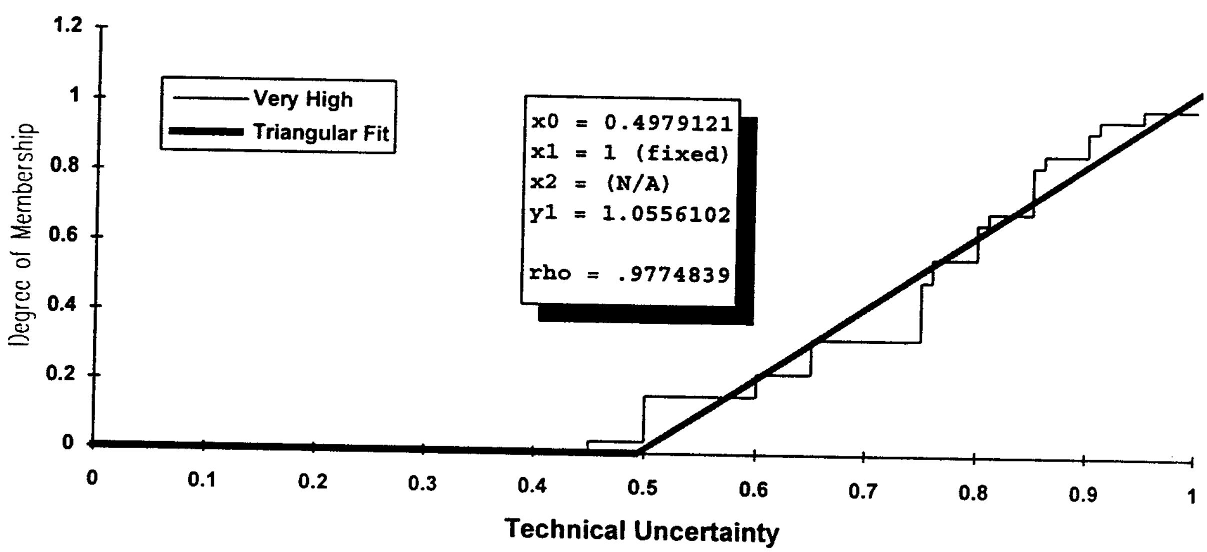


Figure 4.2-6

Degree of Membership for "Very High" Technical Uncertainty Fuzzy Set, with "Triangular" Fit #3

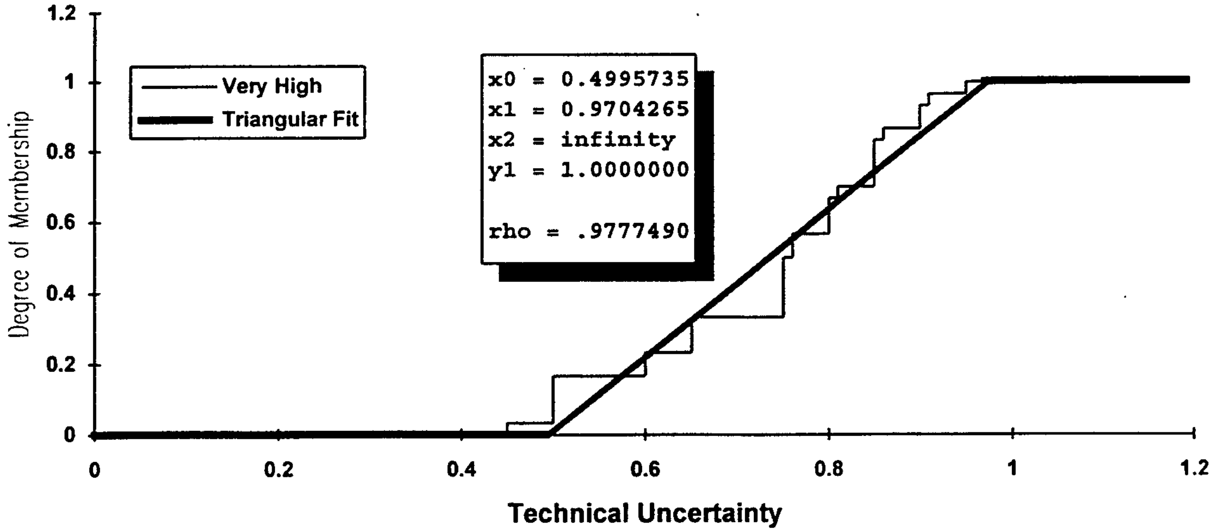


Figure 4.2-7

The program TRIFIT.BAS also generated a trial and error "pattern search" for the set of schedule uncertainty values of the four parameters; x_0, x_1, x_2, x_3 that yields a minimum value of the sum of the squared errors. Table 4.2-2 summarizes the results for the best-fit triangular functions for schedule data.

Fuzzy Se	x_0	x_1	x_2	y_1	Q	S	ρ
Very Lo	-0.1039474	0.0717230	0.1974927	0.7946992	.14221613	4.77798901	.9702351
Low	0.0286019	0.1412642	0.4728496	0.7203559	.11321082	5.19048903	.9781888
Medium	0.0157469	0.4809522	0.7580241	0.6619697	.46626052	5.22222253	.9107161
High	0.1473750	0.7742851	0.9088662	0.5304369	.19006971	3.15923346	.9398368
V. High #1	0.4405798	1 (fixed)	(N/A)	1 (fixed)	.46944417	11.77382146	.9601281
V. High #2	0.4529128	1 (fixed)	(N/A)	1.0382551	.45012591	11.77382146	.9617689
V. High #3	0.4563741	0.9769519	$\to \infty$	1.0000000	.44907583	11.77382146	.9618581

Table 4.2-2

As in the case of technical uncertainty fuzzy membership sets the negative value in the Very Low set represents the extension of the left side of the triangle to intersection with the x axis and in practical application negative values would not be used. The minimum value on the x axis would be 0.0 however, all values for the parameters are depicted for completeness. At the opposite end of the fuzzy sets the very high curve increases to a maximum membership level as x approaches the value one. At 1.0 and beyond the Very High fuzzy set curve has a continuous slope of zero.

There were three potential options considered for fitting the very high schedule uncertainty curve. Option 1 made Very High a triangular distribution that proceeded along a linear function from y = 0 at some x value to y = 1 at x = 1. Under this option the value of x_2 is irrelevant since only the left side of the triangle is practical. Option 2 considered a triangular distribution which proceeded linearly from y = 0 at some x value to whatever value of y at x = 1 provides the best fit. Under this option as under Option 1 x_2 becomes irrelevant. Option 3 effectively considered the function a trapezoid with x_2 moved very far to the right ($x_2 = 1.0 * 10^{10}$) allowing the right side of the function effectively become a line with zero slope extending from the point (x_1, y_1). Note the results of the best-fit value of y_1 approaches one as x_2 approaches ∞. Plots for the apparent best-fit triangular functions are depicted in Figures 4.2-8-9-10-11-12-13 and-14.

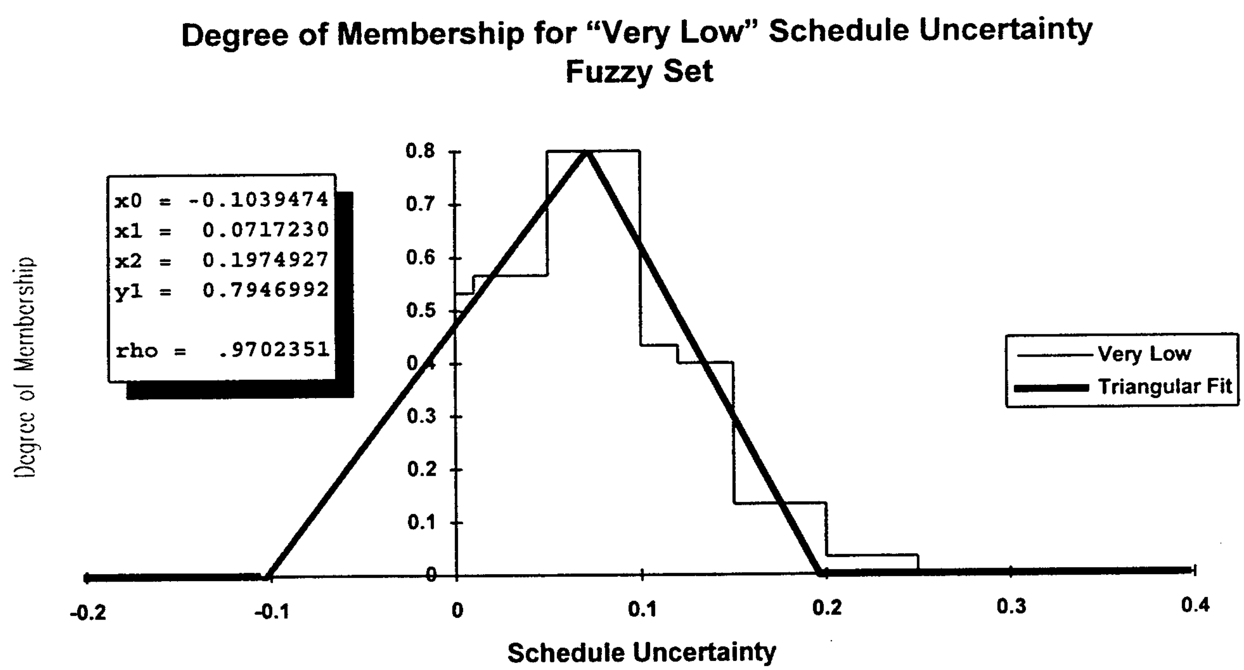


Figure 4.2-8

Degree of Membership for "Low" Schedule Uncertainty Fuzzy Set

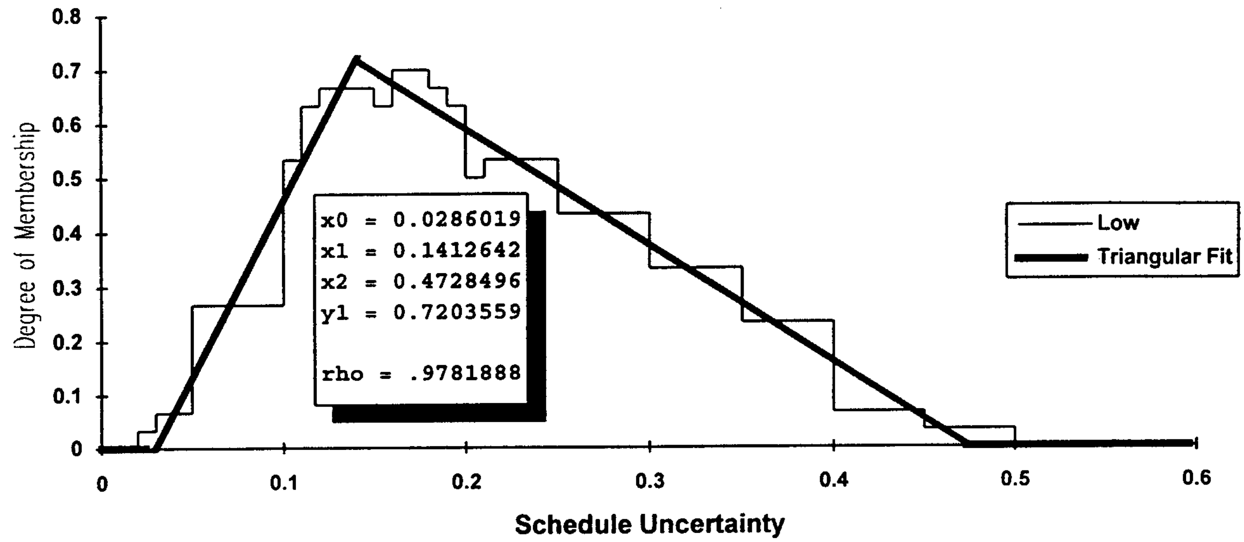


Figure 4.2-9

Degree of Membership for "Medium" Schedule Uncertainty Fuzzy Set

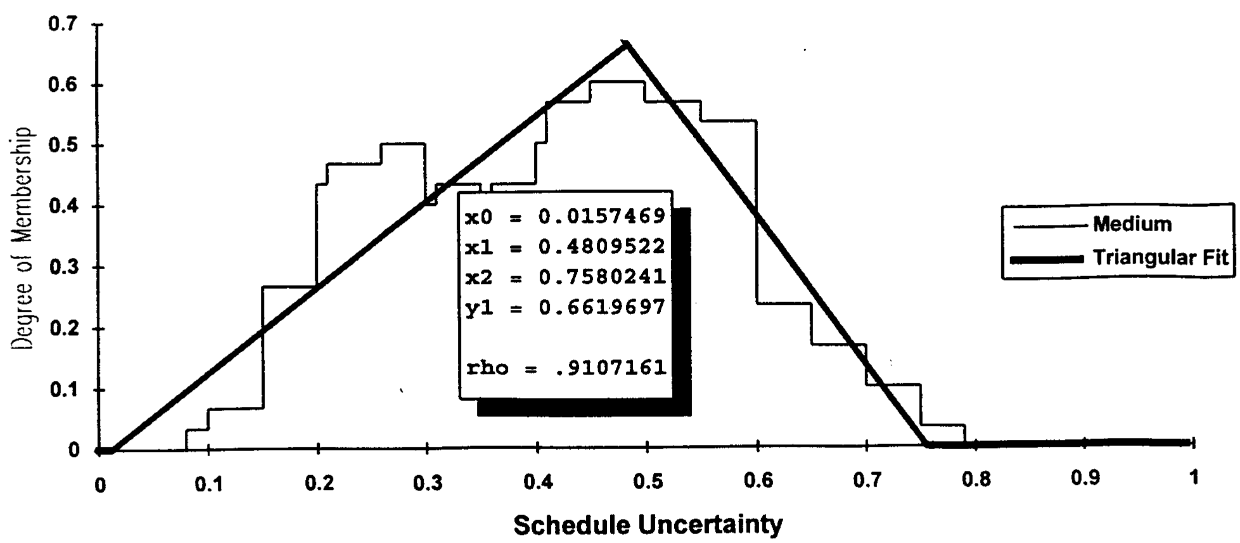


Figure 4.2-10

Degree of Membership for "High" Schedule Uncertainty Fuzzy Set

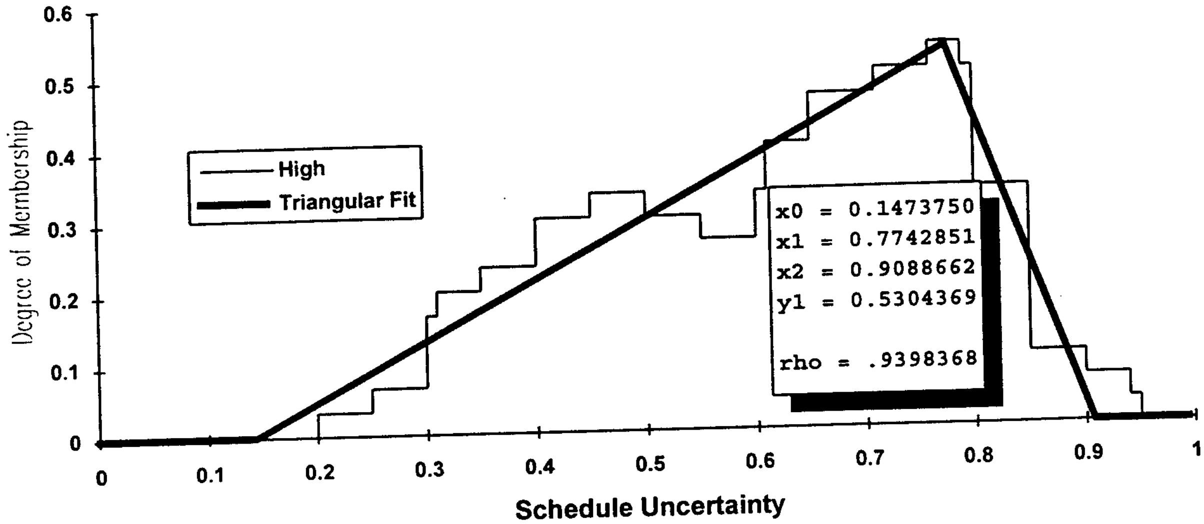


Figure 4.2-11

Degree of Membership for "Very High" Schedule Uncertainty Fuzzy Set, with "Triangular" Fit #1

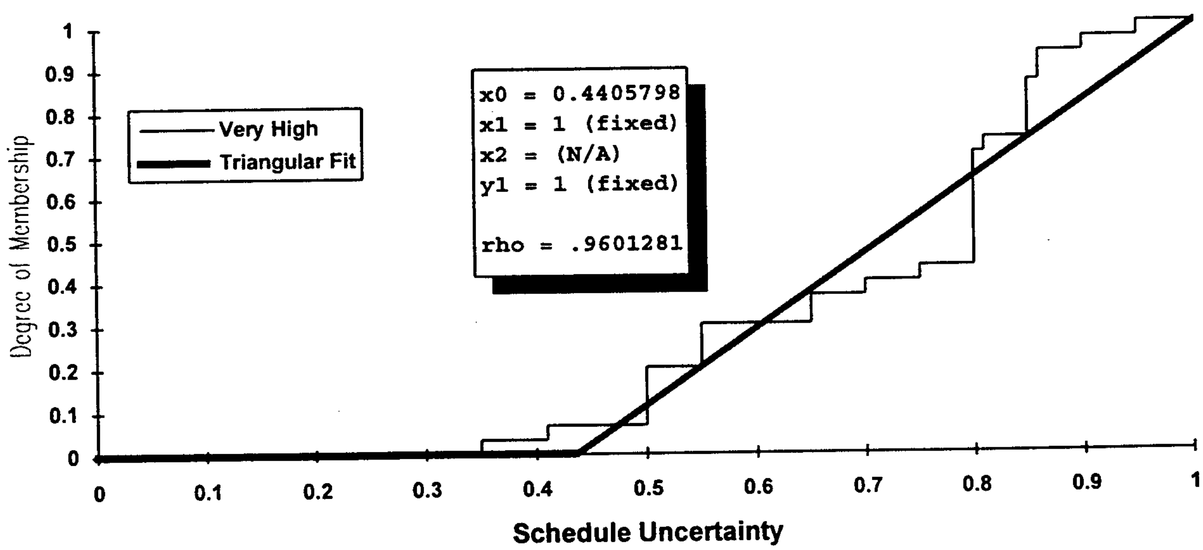


Figure 4.2-12

Degree of Membership for "Very High" Schedule Uncertainty Fuzzy Set, with "Triangular" Fit #2

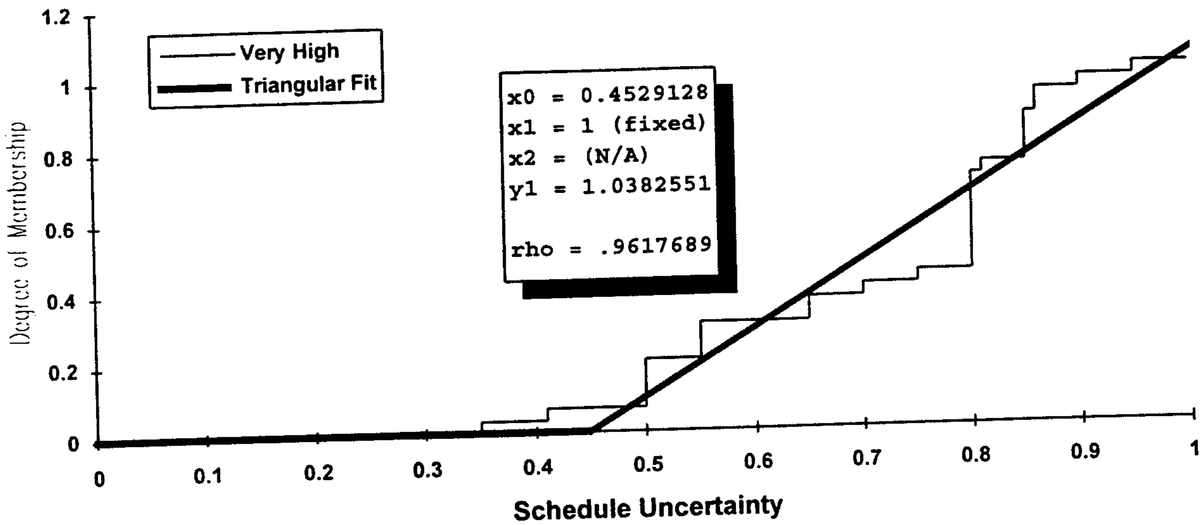


Figure 4.2-13

Degree of Membership for "Very High" Schedule Uncertainty Fuzzy Set, with "Triangular" Fit #3

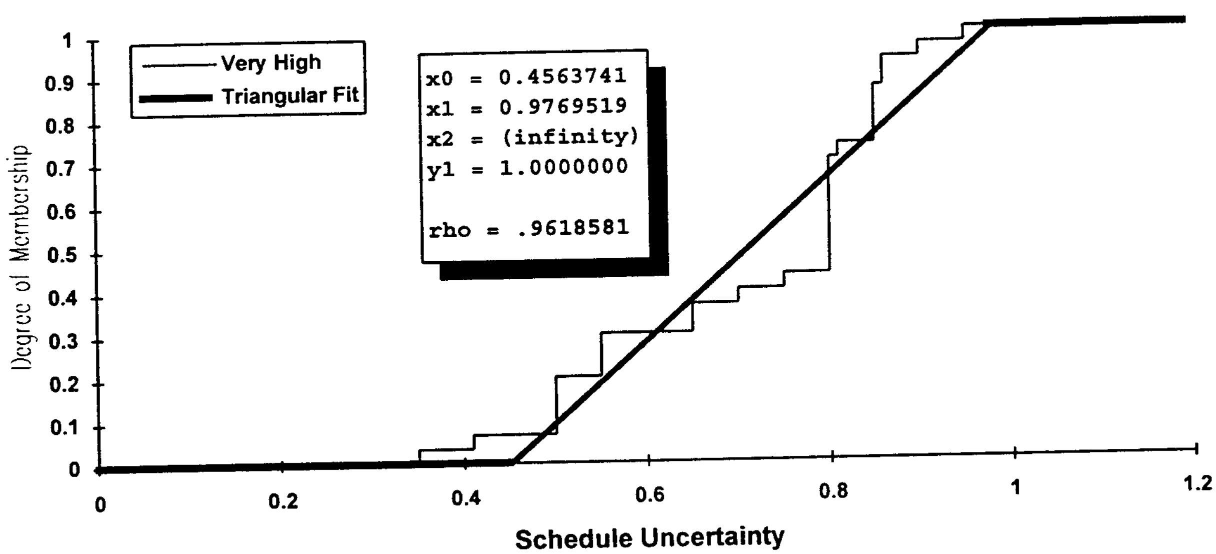


Figure 4.2-14

Using the program TRAPFIT.BAS a trial and error "pattern search" was generated for the set of technical and schedule values of the five parameters; x_0, x_1, x_2, x_3 and y_1 to determine if using trapezoidal functions would yield significantly better "fit" results. The pattern search program yields a minimum value of the sum of the squared errors. Table 4.2-3 summarizes the results for the best-fit trapezoidal functions for technical and schedule data.

Set	x_0	x_1	x_2	x_3	y_1	Q	S	ρ
V. Low	-.2534210	.0650660	.0681109	.2190150	.7203873	.08792462	4.54501096	.9806547
Low	.0296053	.1613158	.2005639	.4112604	.7333333	.16272345	5.88915526	.9723690
Medium	.1307735	.3098214	.4581616	.8017042	.6688889	.11423800	7.03462223	.9837606
High	.2577461	.5476808	.7791675	.9506098	.4594203	.15139997	3.75951109	.9597288

Set	x_0	x_1	x_2	x_3	y_1	Q	S	ρ
V. Low	-.1039474	.0723579	.0744190	.1907993	.7975715	.14093130	4.77798901	.9705041
Low	.0291121	.1319159	.1709146	.4700826	.6666667	.10796559	5.19048903	.9791993
Medium	.1021571	.2404691	.5521597	.7505610	.5096774	.29625951	5.22222253	.9432695
High	.1473750	.7716176	.7748577	.9088662	.5281799	.19006971	3.15923346	.9398368

Table 4.2-3

Best-Fit Trapezoidal Functions; Technical (top) Schedule (bottom)

As in the cases of the triangular functions negative values in the very low fuzzy sets for technical and schedule uncertainty represent the extension of the left side of the trapezoid to intersection with the x axis. In practical application negative values would not be used. The techniques described for handling the Very High data sets of the triangular functions included functions that were virtually trapezoidal (extension of right side at zero slope) therefore no further insight could be gained by fitting additional trapezoids to the Very High data sets for technical and schedule uncertainty. The trapezoidal cases analyzed for both technical and schedule uncertainty include only the Very Low, Low, Medium and High fuzzy membership sets.

The following plots, four technical and four schedule uncertainty, depict the best-fit trapezoidal function plotted on top of the apparent best-fit triangular function. Actual data for each for the four parameters specified for triangular and five parameters specified for trapezoidal functions are compared for each fuzzy set (eight total).

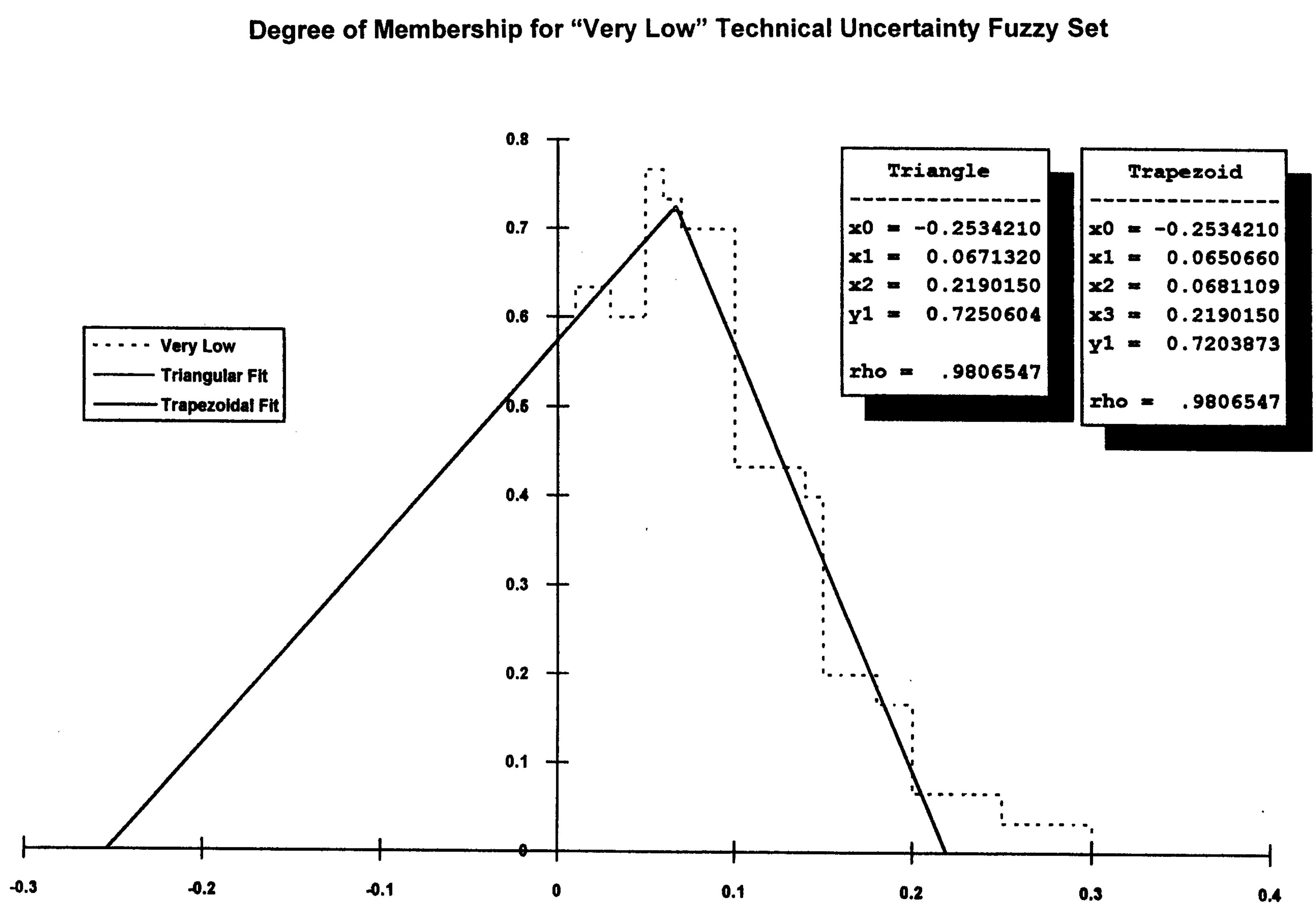

Degree of Membership for “Very Low” Technical Uncertainty Fuzzy Set
Very Low
Triangular Fit
Trapezoidal Fit
Triangle
x0 = -0.2534210
x1 = 0.0671320
x2 = 0.2190150
y1 = 0.7250604
rho = .9806547
Trapezoid
x0 = -0.2534210
x1 = 0.0650660
x2 = 0.0681109
x3 = 0.2190150
y1 = 0.7203873
rho = .9806547
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4

Figure 4.2-15

Degree of Membership for "Low" Technical Uncertainty Fuzzy Set

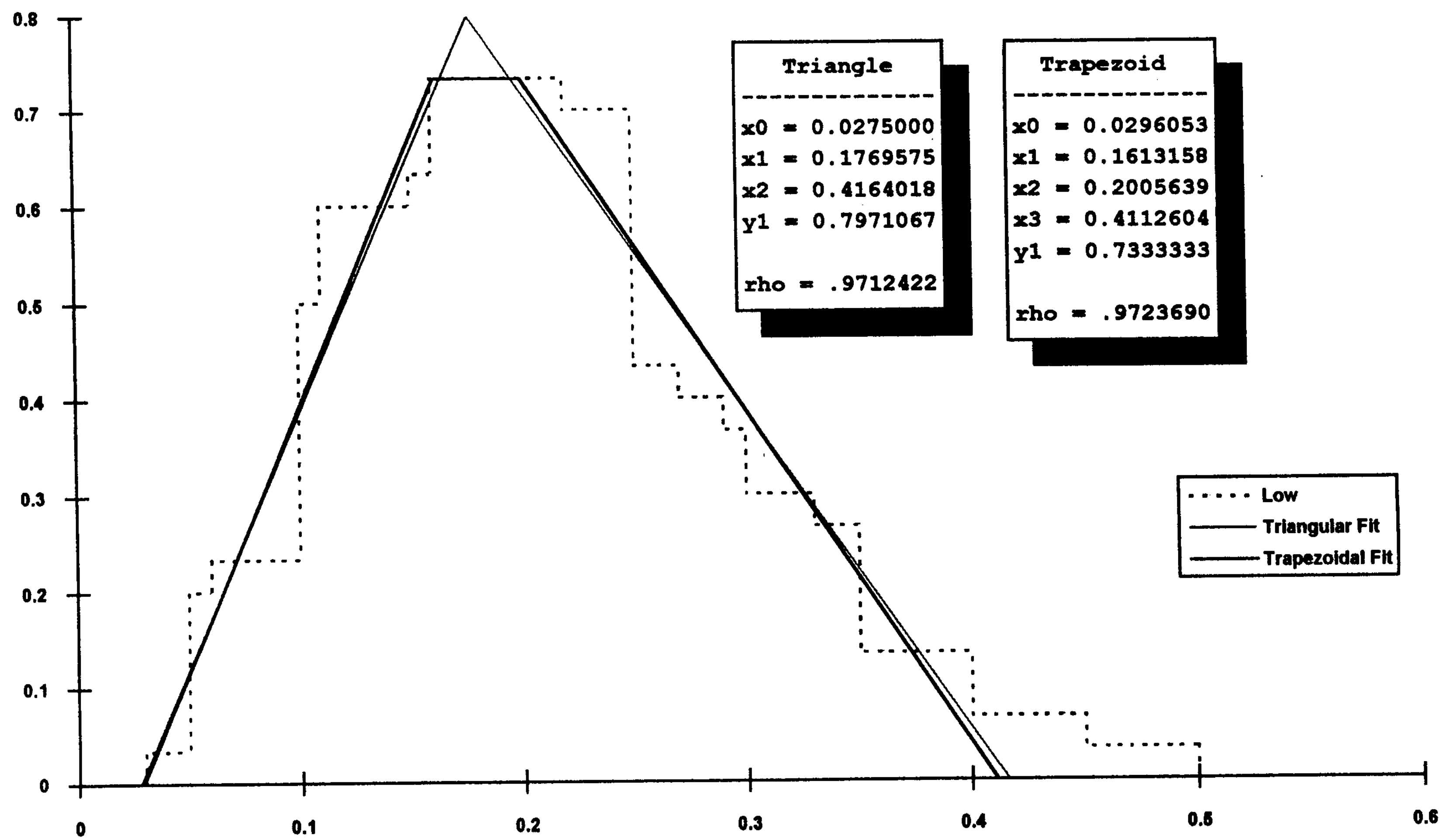
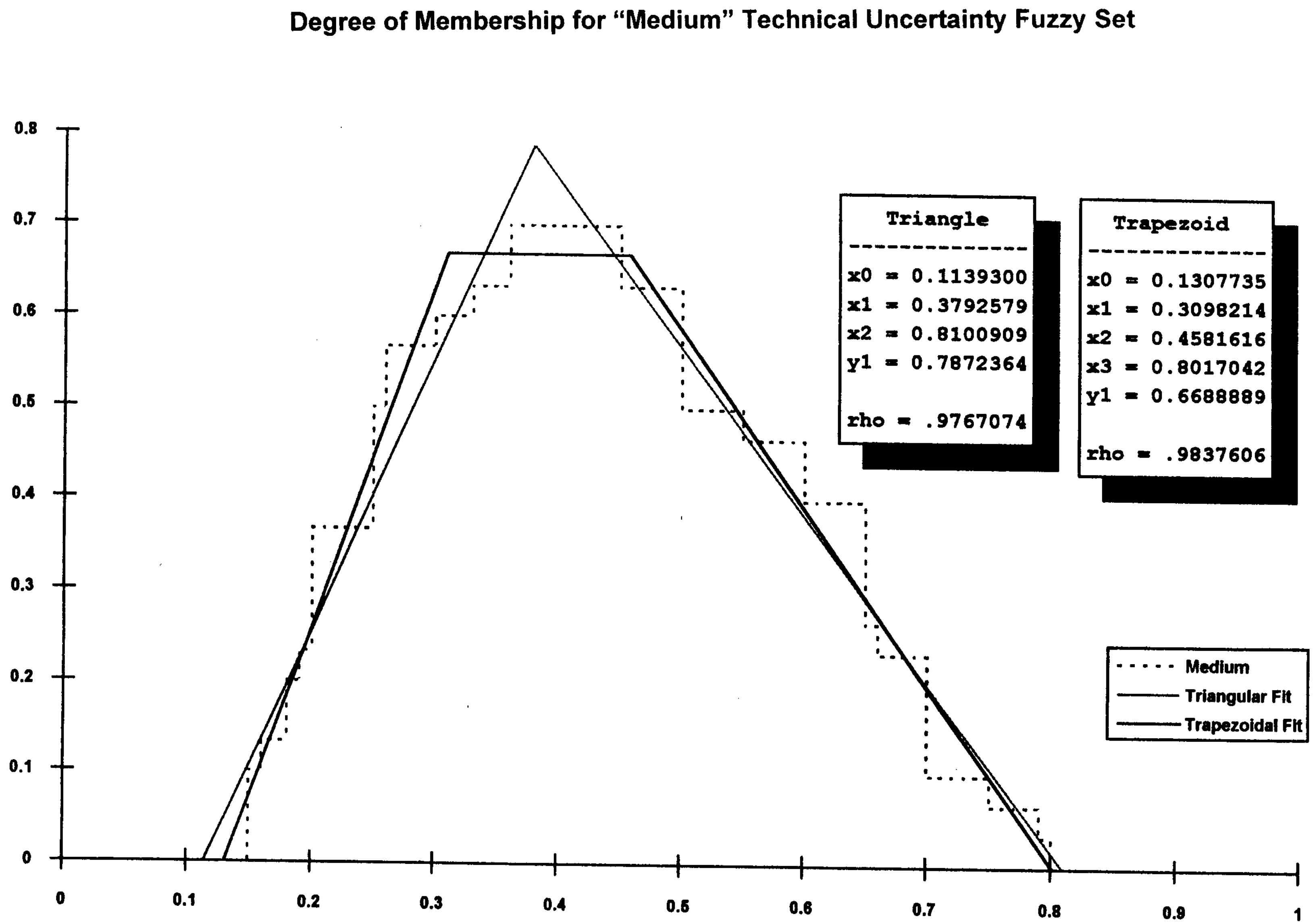


Figure 4.2-16



Degree of Membership for "Medium" Technical Uncertainty Fuzzy Set
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Triangle
x0 = 0.1139300
x1 = 0.3792579
x2 = 0.8100909
y1 = 0.7872364
rho = .9767074
Trapezoid
x0 = 0.1307735
x1 = 0.3098214
x2 = 0.4581616
x3 = 0.8017042
y1 = 0.6688889
rho = .9837606
Medium
Triangular Fit
Trapezoidal Fit

Figure 4.2-17

Degree of Membership for "High" Technical Uncertainty Fuzzy Set

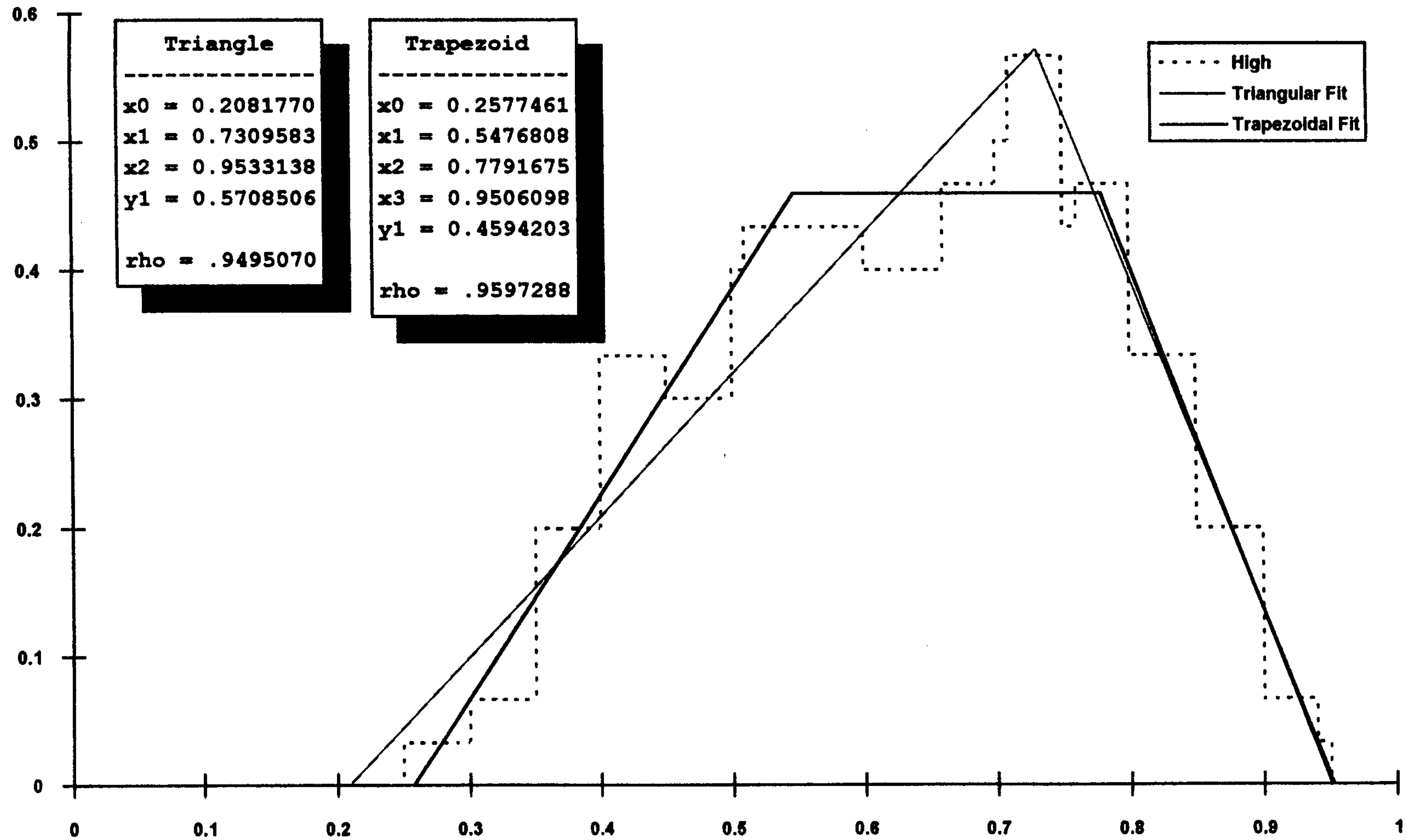


Figure 4.2-18

Degree of Membership for “Very Low” Schedule Uncertainty Fuzzy Set

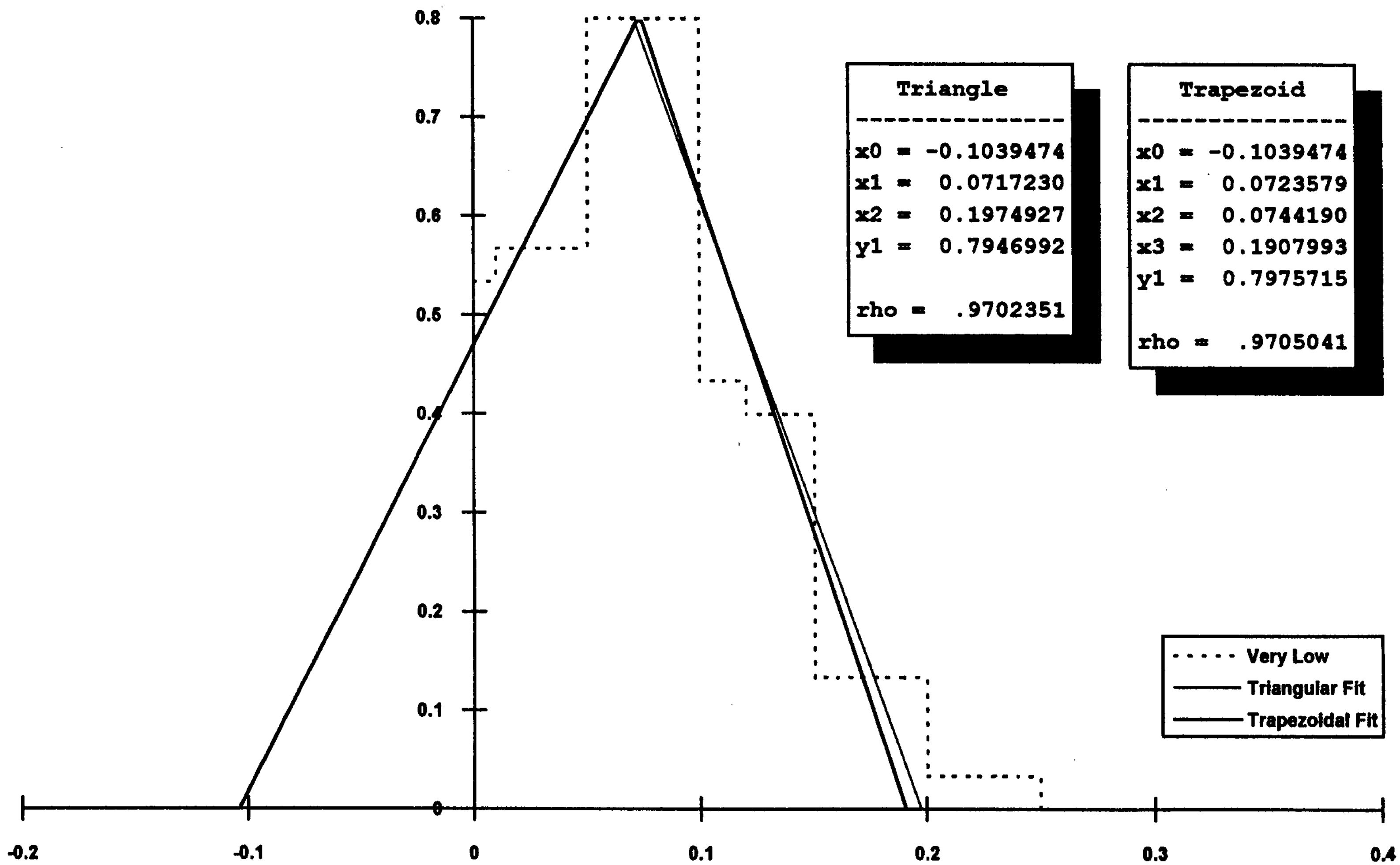


Figure 4.2-19

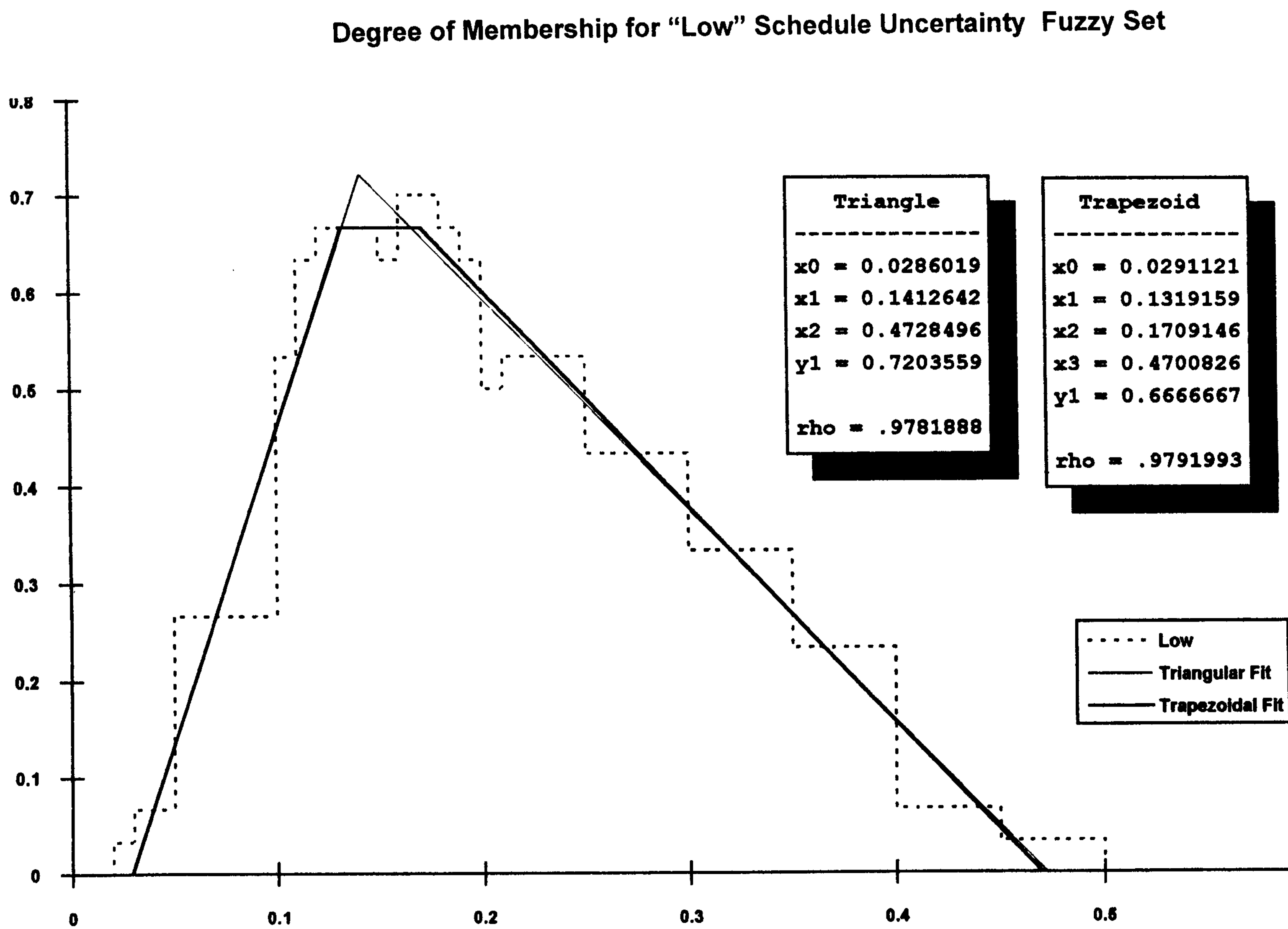

Degree of Membership for "Low" Schedule Uncertainty Fuzzy Set
Triangle
x0 = 0.0286019
x1 = 0.1412642
x2 = 0.4728496
y1 = 0.7203559
rho = .9781888
Trapezoid
x0 = 0.0291121
x1 = 0.1319159
x2 = 0.1709146
x3 = 0.4700826
y1 = 0.6666667
rho = .9791993
Low
Triangular Fit
Trapezoidal Fit
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

Figure 4.2-20

Degree of Membership for "Medium" Schedule Uncertainty Fuzzy Set

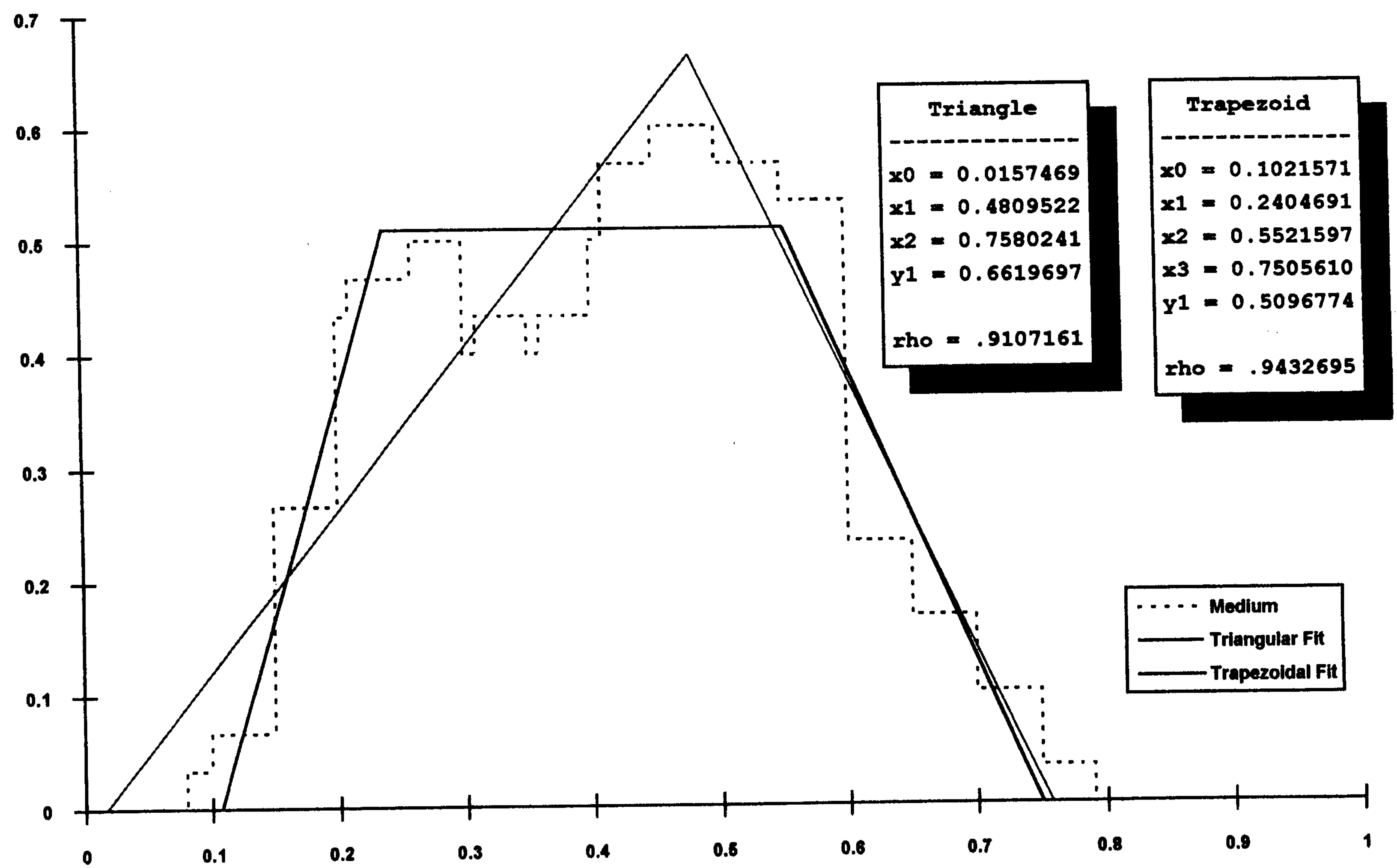


Figure 4.2-21

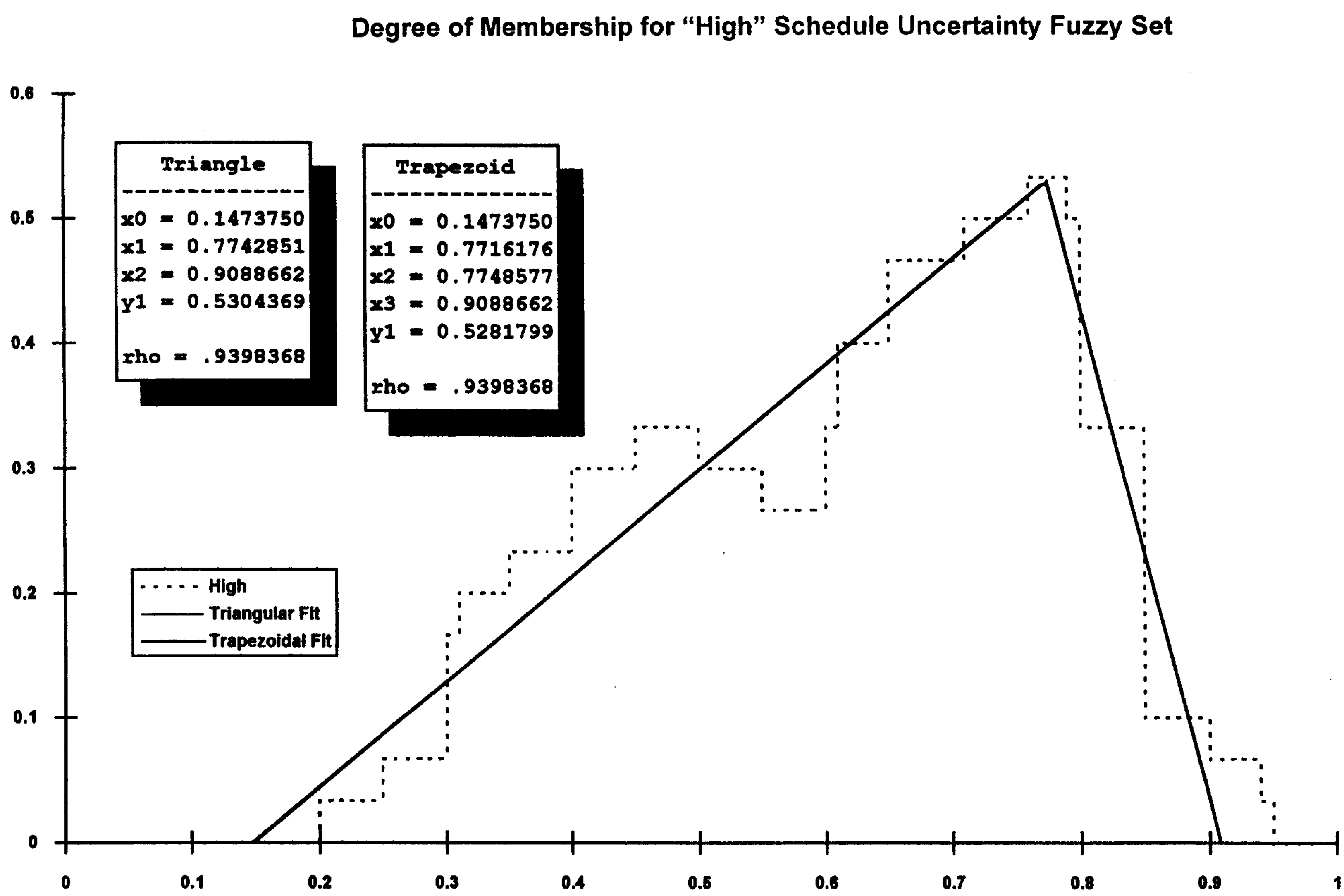

Degree of Membership for "High" Schedule Uncertainty Fuzzy Set
Triangle
x0 = 0.1473750
x1 = 0.7742851
x2 = 0.9088662
y1 = 0.5304369
rho = .9398368
Trapezoid
x0 = 0.1473750
x1 = 0.7716176
x2 = 0.7748577
x3 = 0.9088662
y1 = 0.5281799
rho = .9398368
High
Triangular Fit
Trapezoidal Fit
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 4.2-22

4.2.1 Findings

Analysis of the calculated parameter values for the triangular functions for technical uncertainty data reveals that all were acceptable. Four parameters were estimated for each of the triangular functions. A low value of .9495 was observed for the high technical uncertainty triangular function with all other values being .9712 or above. Schedule uncertainty data were also deemed acceptable although the overall goodness of fit was not as high as the technical data. A value of .9107 for Medium and .9398 for High represented the low end of the fit while values for Very Low, Low, and Very High were .9601 or above.

Analysis of the trapezoidal functions for technical and schedule uncertainty data also revealed acceptability. Five parameters were estimated for each of the trapezoidal functions making these functions somewhat more resource intensive. Goodness of fit for the trapezoidal functions for technical uncertainty revealed the High fuzzy set was the low end of the fit with a value of .9597. Values for Very Low, Low and Medium were .9723 or above. Goodness of fit for the schedule uncertainty revealed that the High fuzzy set was the low end of the fit with a value of .9398 with the Medium set being .9432. The remaining two sets were .9706 or above.

4.2.2 Conclusions

The original assumption made when deciding to test fit for both triangular and trapezoidal functions was that no appreciable difference in "goodness of fit" was expected. Analysis of the resulting data confirms this assumption. Although both functions generally tested high for "goodness of fit" for those sets tested (no test for trapezoidal Very High technical or schedule uncertainty) the triangular function was judged to be overall the best function for purposes of this research. Two basic reasons led to this conclusion: (1) due to the finding that the triangular fit was as good as and generally slightly higher than the trapezoidal fits and (2) the triangular functions required one less parameter to estimate making these functions some what less resource intensive.

4.3 Defuzzification

The defuzzification process can initiate with a value, such as a percent from the C/SCSC system (or from the Delphi technique) or via linguistic fuzzy output of an uncertainty rule base e.g., the uncertainty is medium. Using an input percentage value this research demonstrates a simulation that derives a category (Very Low, Low, etc.) and calculates a share value which represents the degree of membership within each fuzzy set of which there is partial inclusion. When defuzzification initiates from fuzzy linguistic output of a rule base an algorithm for a "crisp" solution is defined separate from the simulation approach.

4.3.1 "Crisp" Solutions

Deriving a "crisp" solution from the simulation uses an algorithmic approach and several inputs. Input data consists of the number of categories to be manipulated, upper and lower boundaries of each category and use of a midpoint for each boundary (apex of the triangle rather than midpoint). An analysis of the algorithm provides insight into the following rules:

a. If input value > Apex value
then; degree of membership = (high value - input)/(high value - apex)
b. Else; if input = Apex value
then degree of membership = 1.0
c. Else; If input value < Apex value
then degree of membership = (input value - low value)/(mid value -low)

Use of the Apex in lieu of the midpoint recognizes the triangular functions are characterized by some degree of skewness for each function. Figure 4.3.1-1 depicts overlap and skewness for the five technical uncertainty fuzzy sets.

Technical Uncertainty Fuzzy Membership Sets Overlaping Regions

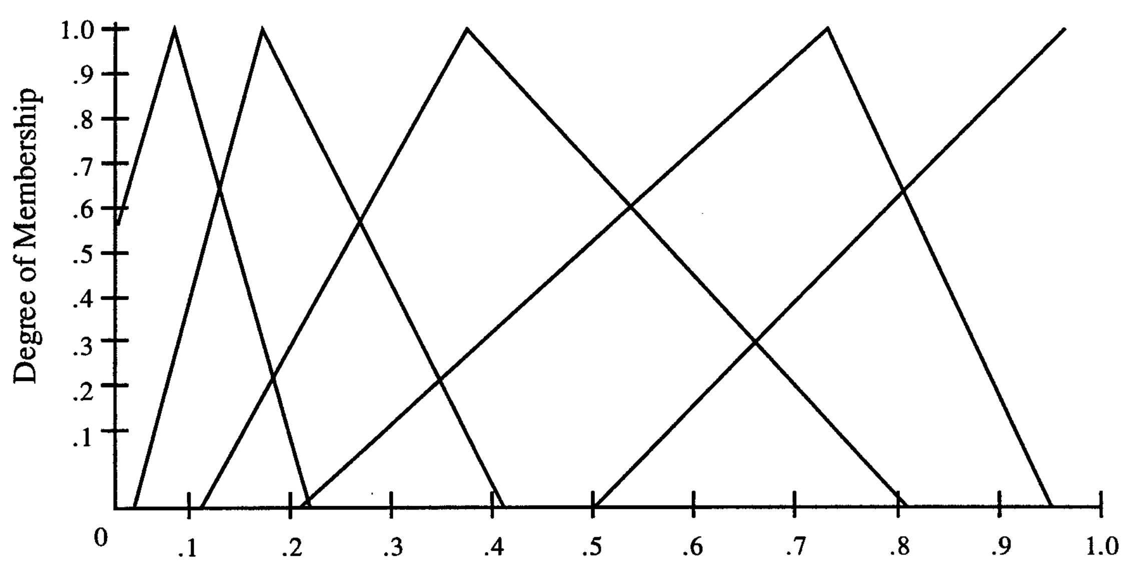


Figure 4.3.1-1

The following analysis uses the algorithm to calculate three input values 0.10, 0.20, and 0.45 for the technical uncertainty fuzzy set. Values depicted in Table 4.3.1-1 indicate the parameters calculated for the triangular functions. In the first step of the algorithm these parameters are used to determine the degree of membership during initial processing of data in the simulation. The Low value, x_0, for the Very Low Fuzzy Set is truncated at a value of 0.00 rather than the actual parameter negative value of -0.2534 developed using the best-fit triangular functions because this research analyzes only those values which are have practical application. A second step uses the weighting contribution of the degree of membership plus the triangular apex to provide the "crisp" solution.

Technical Uncertainty Fuzzy Set
Triangular Functions

Boundary Value:	Low (X0)	Apex (X1)	High (X2)
Linguistic Variable:			
VeryLow	0.00	0.067	0.219
Low	0.028	0.177	0.416
Medium	0.114	0.380	0.810
High	0.208	0.731	0.950
Very High	0.500	0.978	∞

Table 4.3.1-1

In the first step of the algorithm this analysis processes three input values and derives a degree of membership for each fuzzy set which contains the value. The simulation determines the number of categories to be processed, in this case five, one for each of the fuzzy sets. It then determines which fuzzy sets the input value has membership within. The analysis that follows depicts a value of 0.10 that has membership in two fuzzy sets, 0.20 that has membership in three fuzzy sets and 0.45 which has membership in two fuzzy sets.

Processing value of 0.10: = Very Low Fuzzy Set
= Input > Apex
= [(0.219-0.10)/(0.219-0.067)]
= 0.783 Degree of Membership
= Low Fuzzy Set

= Input < Apex
= [(0.10-0.028)/(0.177-0.028)]
= 0.483 Degree of Membership

Processing value of 0.20 = Very Low Fuzzy Set
= Input > Apex
= [(0.219-0.20)/0.219-0.067)]
= 0.125 Degree of Membership
= Low Fuzzy Set
= Input > Apex
= [(0.416-0.20)/(0.416-0.177)]
= 0.904 Degree of Membership
= Medium Fuzzy Set
= Input < Apex
= [(0.20-0.114)/(0.380-0.114)]
= 0.323 Degree of Membership

Processing value of 0.45 = Medium Fuzzy Set
= Input > Apex
= [(0.810-0.45)/(0.810-0.380)]
= 0.837 Degree of Membership
= High Fuzzy Set
= Input < Apex
= [(0.45-0.208)/(0.731-0.208)]
= 0.463 Degree of Membership

The next step in the analysis is to derive a weighted average of the fuzzy sets for technical uncertainty using degree of membership calculated in the first step and the x_1 parameter value calculated in the characterization of the triangular functions. As previously defined the following equation is used in the calculation;

$$D_{wt} = \sum_{1}^{n} (C_{mt_i} * X_{t_i})$$

where;

D_{wt} = Weighted average of "crisp" technical uncertainty solution

C_{mt_i} = Weighting contribution from simulation

X_{t_i} = Abscissa value of the apex of triangular function

$$= [(0.783 * 0.067) + (0.483 * 0.177)]$$
$$= 0.138$$

This value is the "crisp" solution derived from the fuzzy sets, Very Low and Low, both sets in which the input value of 0.10 had a degree of membership. Continuing the analysis for the input of 0.20:

$$= [(0.125 * 0.067) + (0.904 * 0.177) + (0.323 * 0.380)]$$
$$= 0.2907$$

This "crisp" solution represents the fuzzy sets Very Low, Low and Medium. Continuing the analysis for an input of 0.45:

$$= [(0.837 * 0.380) + (0.463 * 0.731)]$$
$$= 0.6565$$

This "crisp" solution represents the fuzzy sets Medium and High and completes the analysis for inputs in the form of a percentage that may be obtained via such sources as the C/SCSC.

"Crisp" solutions may be calculated from a qualitative assessment of uncertainty which is common using the Delphi technique. The following equation was defined for derivation of a "crisp" solution from fuzzy input such as "the uncertainty is judged to be medium."

$$D_{qt} = \sum_{1}^{n} (C_{qt_i} * R_{qt_i})$$

Where;

D_{qt} = Weighted average, qualitative technical input

C_{qt_i} = Abscissa value of the apex of the triangular function for technical uncertainty

R_{qt_i} = Ratio of the range represented by individual technical fuzzy membership sets to the sum of the ranges of the sets specified by the Delphi technique

With fuzzy input, linguistic descriptions, as the basis for initiation of deriving a "crisp" solution the analysis begins with an additional degree of vagueness as opposed to initiating the analysis with a "hard number." Characterized parameters, x_0, x_1, and x_2, that have been previously defined provide the data for quantification of the fuzzy sets for Low and Medium technical uncertainty as follows;

$$= [(0.177 * 0.3579) + (0.380 * (1-(0.3579)]$$
$$= 0.3073$$

The value of 0.3579, defined as the ratio of the range represented by individual technical fuzzy membership sets to the total of the ranges of the sets specified by the Delphi technique, is derived in this analysis as follows:

Range of abscissa values for Low fuzzy set;
= 0.416-0.028
= 0.388

Range of abscissa values for Medium fuzzy set;
= 0.810-0.114
= 0.696

Therefore the sum of ranges specified in this example;
= 0.388 + 0.696
= 1.084

and the ratios of the individual sets to the sum of the sets;
= 0.388/1.084

= 0.3579 for the Low fuzzy set and;

= 0.696/1.084

= 0.6421 for Medium fuzzy set or as detailed above;

= 1- 0.3579

= 0.6421

4.3.2 Combining Rule Outputs

The look up table matrix provides rules for combining fuzzy outputs under circumstances of simultaneous occurrences of technical and schedule uncertainty. Conflicting fuzzy outputs may generate from the simulation or from a group of "experts" whose opinions represents differences in assessment of uncertainty relative to either one element of the CES or to the top level of a cost estimate. Matrices developed in the methodology specified two rules, for the "if-and-then" and for the "if-or-then" cases. In each case the rule base executes to a previously defined "truth value" represented by either a min or max function.

Each matrix output was created by 25 rules each of which have some possibility of occurrence. Certain levels of technical uncertainty may or may not have an impact on schedule uncertainty depending on available "work-a-rounds." Medium technical uncertainty may be driven by pushing the state of the art in a technical area while creative solutions emerge that negate schedule impact. In this case technical impacts are possible without a corresponding schedule impact. Technical uncertainty could however have a direct and significant impact on schedule. For those situations in which a possibility of simultaneous technical and schedule uncertainty exists the rule base provides an approach for mitigating conflicting rule outputs e.g. technical uncertainty is judged to be high but solutions are available which should reduce the impact to schedule uncertainty thereby allowing schedule uncertainty to be considered low.

Although possibility exists for all matrix combinations, experience suggests that certain regions within each matrix have a near zero potential of occurrence. Factoring this realization into the analysis allows rule base modification without reducing robustness. Experience suggests it would not be practical to "sell" a program with known high or very high technical uncertainty as one which has very low or low schedule uncertainty. Conversely, very low or low technical uncertainty would probably not drive schedule uncertainty to medium, high or very high levels. This rationale suggests appropriate modifications to the min and max function rule bases as follows.

Technical

SCHEDULE	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
VERY LOW	VL	VL	VL		
LOW	VL	L	L		
MEDIUM	VL	L	M	M	M
HIGH	VL	L	M	H	H
VERY HIGH	VL	VL	M	H	VH

Table 4.3.2-1

Modified Fuzzy Rule Strengths for Min Function

TECHNICAL

SCHEDULE	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
VERY LOW	VL	L	M		
LOW	L	L	M		
MEDIUM			M	H	VH
HIGH			H	H	VH
VERY HIGH			VH	VH	VH

Table 4.3.2-2

Modified Fuzzy Rule Strengths for Max Function

4.4 Cost and Price Data

A multi-step approach was undertaken for analysis of the 46 pair of original cost and price sample data plus the re-quantification of 14 pair of original IGE uncertainty data. That process included quantifying uncertainty percentages and generating statistics for both IGE and negotiated contract prices. Fourteen points of the IGE data were processed through the fuzzy algorithm to derive a "crisp" solution for comparison with the historical cost data. The 14 IGEs were generated for recent procurement actions for which contracts were dated February 1992 and forward. Recent contracts were chosen since it was necessary to verify how the uncertainty quantification was derived for the IGEs in order to properly apply a "re-quantified uncertainty" via the fuzzy algorithm approach. It was determined that all uncertainty estimates were quantified using multiplicative factors applied at high levels within the CES by cost analysts using their judgment after discussion with technical experts. This methodology of application is consistent with techniques commonly applied throughout the missile weapon system community.

Descriptive statistics were calculated for the 46 pair of data and the 14 pair of data including histogram distributions and analysis of variance (ANOVA). Data from the 14 pair was separately analyzed using a paired t-Test.

4.4.1 Forty Six Pair Descriptive Statistics

Significant dispersion of absolute values for both IGE and negotiated contract prices negated the utility of statistical distance measures including the mean absolute deviation. Magnitude of the absolute values justified use of percent ratios for each set of data rather than actual values. Adjusting input data to ratios constituted a normalizing procedure that aided in data comparison. All statistical measures for comparison were made with normalized data by creating ratio factors (percentages) rather than the absolute values of the IGEs and negotiated contract price data. Histograms were generated for both IGE uncertainty ratios and "negotiated price uncertainty ratios" developed for the negotiated contract prices. Tables 4.4.1-1 and 4.4.1-2 depicts IGE and "negotiated price" uncertainty ratio descriptive statistics respectively. Adjusted data from which the ratios were derived are provided for completeness.

ADJUSTED DATA	ADJUSTED DATA	ADJUSTED DATA		
IGE	Uncertainty	Uncert %		
110226576.9	8450704.225	0.083	Column 1	
104102878.1	6123698.714	0.063		
229493297	16473528.74	0.077	Mean	0.072562128
83805476.86	5193578.848	0.066	Standard Error	0.003137997
1121340.888	108593.0123	0.107	Median	0.069683051
31242899.34	2385821.404	0.083	Mode	0.0625
272816916.8	12380848.03	0.048	Standard Deviatio	0.021282934
795273.8014	51124.74438	0.069	Variance	0.000452963
124971597.4	8634401.272	0.074	Kurtosis	-0.33026736
198312698.6	13366933.27	0.072	Skewness	0.489197786
241043059.1	9860852.416	0.043	Range	0.083163386
42730360.47	4273036.047	0.111	Minimum	0.037421267
158948818.5	10490622.02	0.071	Maximum	0.120584653
1031048.003	99607.92625	0.107	Sum	3.337857886
133044554.5	7219471.947	0.057	Count	46
122731023.1	5981848.185	0.051		
13923267.33	866336.6337	0.066		
120668316.8	9591584.158	0.086		
49270990.45	2916038.21	0.063		
86475615.89	5279034.691	0.065		
1124694.377	81662.59169	0.078		
58320764.2	4424333.836	0.082		
43237807.94	2312719.96	0.057		
51282051.28	3016591.252	0.063		
23127199.6	1131221.719	0.051		
256410.2564	10055.30417	0.041		
3520782.396	132029.3399	0.039		
1564792.176	65525.67237	0.044		
2815485.168	101558.5721	0.037		
28154851.68	1860231.272	0.071		
3318250.377	226244.3439	0.073		
103572786	6081337.894	0.062		
69926650.37	6748166.259	0.107		
24132730.02	1960784.314	0.088		
75305623.47	4645476.773	0.066		
49410870.39	2945648.043	0.063		
1407742.584	99547.51131	0.076		
1955990.22	102689.4866	0.055		
17578867.35	902698.5937	0.054		
2660585.329	237552.2615	0.098		
596577.0171	50855.74572	0.093		
44243338.36	3318250.377	0.081		
18529076.4	1083238.312	0.062		
135941320.3	14278728.61	0.117		
47510452.3	4133409.35	0.095		
8997555.012	968215.1589	0.121		

Table 4.4.1-1 Adjusted Data and Descriptive Statistics
for IGE Uncertainty Ratios

ADJUSTED DATA	Adj Data			
Uncertainty	Neg %	IGE		
Price-IGE UF	Uncert.	Uncert. %		
18371096.14	0.200	0.083	*Column 1*	
9797917.942	0.104	0.063		
36355373.78	0.188	0.077	Mean	0.212
-5901794.145	-0.066	0.066	Standard Error	0.020
177053.8244	0.188	0.107	Median	0.199
3976369.007	0.146	0.083	Mode	#N/A
53686863.15	0.245	0.048	Standard Deviatior	0.137
227221.0861	0.400	0.069	Variance	0.019
17041581.46	0.158	0.074	Kurtosis	0.750
20817355.1	0.117	0.072	Skewness	0.270
-8765202.147	-0.035	0.043	Range	0.697
8765202.147	0.258	0.111	Minimum	-0.097
41326692.8	0.351	0.071	Maximum	0.600
240542.5453	0.304	0.107	Sum	9.735
21658415.84	0.194	0.057	Count	46.000
26815181.52	0.280	0.051		
2578382.838	0.227	0.066		
16501650.17	0.158	0.086		
8547008.547	0.210	0.063		
5530417.295	0.068	0.065		
146699.2665	0.150	0.078		
12066365.01	0.261	0.082		
7038712.921	0.194	0.057		
10055304.17	0.244	0.063		
6535947.712	0.394	0.051		
20110.60835	0.085	0.041		
586797.066	0.200	0.039		
391198.044	0.333	0.044		
-301659.1252	-0.097	0.037		
3016591.252	0.120	0.071		
703871.2921	0.269	0.073		
17103762.83	0.198	0.062		
5378973.105	0.083	0.107		
9049773.756	0.600	0.088		
7823960.88	0.116	0.066		
4751045.23	0.106	0.063		
100553.0417	0.077	0.076		
488997.555	0.333	0.055		
4751045.23	0.370	0.054		
190041.8092	0.077	0.098		
88019.5599	0.173	0.093		
14077425.84	0.467	0.081		
4275940.707	0.300	0.062		
25427872.86	0.230	0.117		
11402508.55	0.316	0.095		
2738386.308	0.438	0.121		

Table 4.4.1-2 Adjusted Data and Descriptive Statistics for Negotiated Price Uncertainty Ratios

IGE cost estimate data was obtained in two parts: "uncertainty free" and quantified uncertainty. The quantified uncertainty part was normalized by calculating the ratio of its value relative to the "uncertainty free" value. Statistics were calculated on the data in the Uncert % column of Table 4.4.1-1 and the Neg % Uncert column of Table 4.4.1-2.

Inspection of the sample means indicate a distinct difference which is expected of this data. IGE data ranges are much more narrow than the negotiate price ranges as indicated by summary measures of dispersion. The IGE Uncertainty ratio range is 0.084 while the Negotiated Uncertainty percent range is 0.697. Tables 4.4.1-3 and 4.4.1-4 provide histograms for individual and cumulative frequencies. The number of bins was calculated as the square root of the number of input values evenly distributed between the data's minimum and maximum values. The IGE uncertainty distribution is not particularly well behaved indicating the form of the underlying sample distribution contains a slight degree of positive skewness. Note the negotiated price uncertainty distribution indicates mesokurtic tendency. Skewness is calculated as follows;

$$\frac{n}{(n-1)(n-2)}\sum\left(\frac{x_i-\bar{x}}{s}\right)^3$$

The platykurtic tendency of the IGE frequency distribution compared to a normal distribution is noted in the negative kurtosis value. Positive kurtosis is noted in the negotiated price distribution as indicated by its relative peakedness. Kurtosis is calculated as follows;

$$\frac{n}{(n-1)(n-2)(n-3)}\sum\left(\frac{x_i-\bar{x}}{s}\right)^4-\frac{3(n-1)^2}{(n-2)(n-3)}$$

Due to the characterization of platykurtic tendency and positive skewness in the sample distribution of IGE uncertainty data an analysis was undertaken to determine the proper choice of sample size needed for various tests of significance that occur later in the analysis. Operating characteristic curves for values of n for a two sided F test for a level of significance of $\alpha = 0.05$ use the following guideline to determine sample size;

$$\gamma=\frac{\sigma_1}{\sigma_2}$$

The utility and results of this statistic will be provided later in this analysis.

Bin	Frequency	Cumulative %
0.037421267	7	15.22%
0.051281831	12	41.30%
0.065142396	12	67.39%
0.07900296	6	80.43%
0.092863524	3	86.96%
0.106724089	6	100.00%

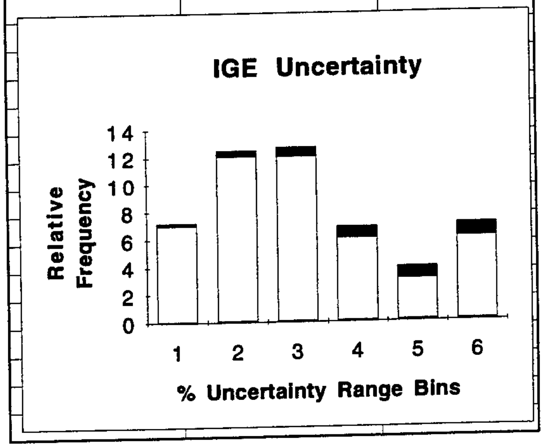


Table 4.4.1-3 IGE Uncertainty Frequency Distribution

Bin	Frequency	Cumulative %
-0.097	3.000	6.52%
0.019	10.000	28.26%
0.135	17.000	65.22%
0.252	10.000	86.96%
0.368	5.000	97.83%
0.484	1.000	100.00%

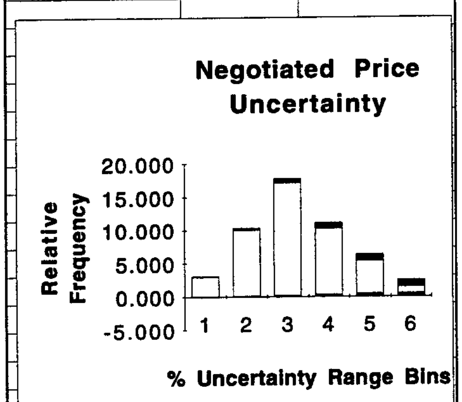


Table 4.4.1-4 Negotiated Price Uncertainty Frequency Distribution

ANOVA comparison of the means of the IGE and negotiated price distributions using the sample variances verifies the underlying distributions are not identical. The ratio of explained to unexplained variation is 46.45. This value is significantly higher than required for the $F_{0.05}$ level for one degree of freedom in the numerator and ninety degrees of freedom in the denominator. F distribution values for this level are specified as 3.96 and 3.94 respectively for 80 and 100 degrees of freedom for the denominator. Table 4.4.2-5 provides ANOVA for the 46 pair of data.

ANOVA: Single-Factor						
Summary						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Column 1	46	9.734558158	0.21162083	0.018697	Neg Price	
Column 2	46	3.337857886	0.072562128	0.000453	IGE	
ANOVA						
Source of Variation						
	SS	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.444758417	1	0.444758417	46.45077	1.03724E-09	3.946865945
Within Groups	0.861735035	90	0.009574834			
Total	1.306493452	91				

Table 4.4.1-5 ANOVA for IGE and Negotiated Price

4.4.2 Fourteen Pair Descriptive Statistics

A subset of the total 46 data points of the IGE, representing more recent contract prices and IGEs, was re-quantified for additional statistical analysis. The initial statistical issue with this analysis was choice of sample size. To determine a sample size for recalculation, operating characteristic curves for different values of n for a two sided F test for a level of significance $\alpha = 0.05$ was utilized. Using the following guideline a value of 2.19 was calculated which indicated that a sample size of 14 would represent less than a 25 percent probability of accepting H_0 when H_0 is false:

$$\gamma = \frac{\sigma_1}{\sigma_2}$$

The 14 "crisp" values of technical and schedule uncertainty processed via the fuzzy algorithm were used to develop descriptive statistics ("re-quantified") that were compared to statistics for the same 14 values prior to the fuzzy process ("historical"). Data range increased as indicated in the change of mean value. Inspection of the sample means indicate a distinct difference in data sets. Kurtosis, that existed in the 46 pair historical data, was magnified significantly in the 14 pair re-quantified uncertainty and can be characterized as platykurtic with the tendency to concavity. Skewness shifted remarkably from positive to negative. A large contributor to the reversal in skewness was the change in value of n from 46 to 14. The effect of this change was to increase the front end of the equation by a factor of 15. Table 4.4.2-1 provides a comparison of historical and re-quantified IGE uncertainty descriptive statistics.

IGE "Requantified" Uncertainty		IGE "Historical" Uncertainty	
14 Data Points		14 Data Points	
Column 1		*Column 1*	
Mean	0.242642046	Mean	0.079961
Standard Error	0.011840732	Standard Error	0.00540641
Median	0.252493818	Median	0.0785882
Mode	0.205933905	Mode	#N/A
Standard Deviatio	0.044303961	Standard Deviatio	0.02022893
Variance	0.001962841	Variance	0.00040921
Kurtosis	-1.06817394	Kurtosis	-1.06596647
Skewness	-0.240315121	Skewness	0.36278528
Range	0.141477667	Range	0.06323229
Minimum	0.17035912	Minimum	0.05413105
Maximum	0.311836787	Maximum	0.11736334
Sum	3.396988645	Sum	1.11945401
Count	14	Count	14

Table 4.4.2-1 Comparative Descriptive Statistics
for the 14 Point IGE "Re-quantified" Uncertainty and IGE "Historical" Uncertainty

Histograms, frequency distributions, were developed for the 14 data points of re-quantified and historical uncertainty. Both distributions are characterized as "non-normal." A notable change is the reduction in number of range bins from 6 to 3 driven by reduction in the number of data input values from 46 to 14. The number of range bins was calculated as the square root of the number of input values.

Bin	Frequency	Cumulative %
0.054131054	6	42.86%
0.075208484	5	78.57%
0.096285914	3	100.00%

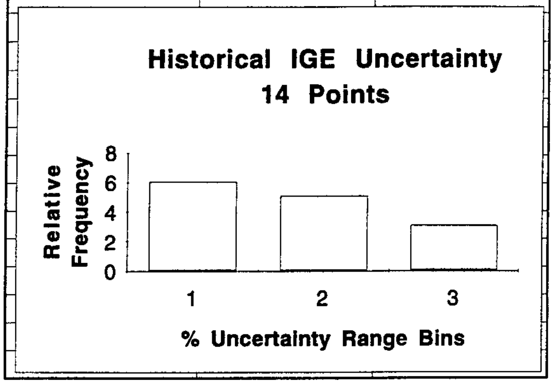


Table 4.4.2-2 Historical IGE Uncertainty
Frequency Distribution for 14 Points

Bin	Frequency	Cumulative %
0.17035912	5	35.71%
0.217518343	3	57.14%
0.264677565	6	100.00%

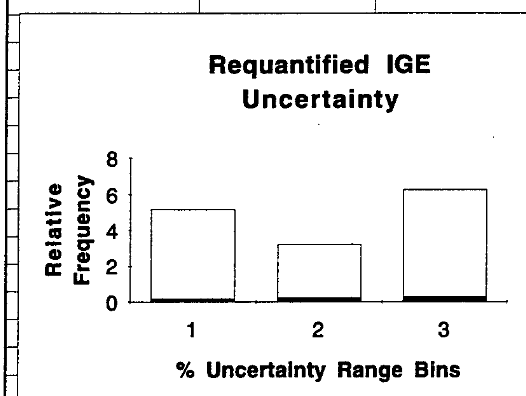


Table 4.4.2-3 Re-quantified IGE Uncertainty
Frequency Distribution for 14 Points

ANOVA for the 14 points involved recalculation, via a fuzzy algorithm process, of the uncertainty quantification portion of historical IGE data. Results of the ANOVA for the 14 data points of IGE recalculated uncertainty appear in Table 4.4.2-4 below.

Anova: Single-Factor						
Summary						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Column 1	14	1.119454012	0.079961001	0.000409	Historical IGE Uncertainty	
Column 2	14	3.396988645	0.242642046	0.001963	Requantified IGE Uncertainty	
ANOVA						
Source of Variation						
	SS	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.185255857	1	0.185255857	156.1989	1.69326E-12	4.225199746
Within Groups	0.030836656	26	0.001186025			
Total	0.216092513	27				

Table 4.4.2-4 ANOVA for IGE Historical Uncertainty
and Re-quantified IGE Uncertainty 14 Points

To provide ANOVA for the 14 data points the Historical IGE uncertainty values were processed through the fuzzy logic algorithm. Tables 4.4.2-5 and 4.4.2-6 provide data and rules for calculating the "crisp" output of the fuzzy process.

	Technical Uncertainty Triangular Parameter Data				
	Low (x0)	Apex (X1)	High (x2)		
Very Low	0.000	0.067	0.219		Rule:
Low	0.028	0.177	0.416		Input > Apex
Medium	0.114	0.380	0.810		Input < Apex
High	0.208	0.731	0.953		Input = Apex = 1
Very High	0.500	0.978	Infinity		
Uncertainty:	Fuzzy Sets:		c_{mt_i}	x_{t_i}	"Crisp":
0.062	VL, L	Input < Apex	0.925373134	0.067	0.102389262
		Input < Apex	0.228187919	0.177	
0.107	VL, L	Input > Apex	0.736842105	0.067	0.143214059
		Input < Apex	0.530201342	0.177	
0.088	VL, L	Input > Apex	0.861842105	0.067	0.129018589
		Input < Apex	0.402684564	0.177	
0.066	VL, L	Input < Apex	0.985074627	0.067	0.11114094
		Input < Apex	0.255033557	0.177	
0.063	VL, L	Input < Apex	0.940298507	0.067	0.104577181
		Input < Apex	0.234899329	0.177	
0.076	VL, L	Input > Apex	0.940789474	0.067	0.120053029
		Input < Apex	0.322147651	0.177	
0.055	VL, L	Input < Apex	0.820895522	0.067	0.087073826
		Input < Apex	0.181208054	0.177	
0.054	VL, L	Input < Apex	0.805970149	0.067	0.084885906
		Input < Apex	0.174496644	0.177	
0.098	VL, L	Input > Apex	0.796052632	0.067	0.136489889
		Input < Apex	0.469798658	0.177	
0.093	VL, L	Input > Apex	0.828947368	0.067	0.132754239
		Input < Apex	0.436241611	0.177	
0.081	VL, L	Input > Apex	0.907894737	0.067	0.123788679
		Input < Apex	0.355704698	0.177	
0.062	VL, L	Input < Apex	0.925373134	0.067	0.102389262
		Input < Apex	0.228187919	0.177	
0.117	VL, L, M	Input > Apex	0.671052632	0.067	0.154971073
		Input < Apex	0.597315436	0.177	
		Input < Apex	0.011278195	0.380	
0.095	VL, L	Input > Apex	0.815789474	0.067	0.134248499
		Input < Apex	0.44966443	0.177	

Table 4.4.2-5 Technical Uncertainty Fuzzy Algorithm
for 14 Points

	Schedule Uncertainty Triangular Parameter Data				
	Low (x0)	Apex (X1)	High (x2)		
Very Low	0.000	0.072	0.197		Rule:
Low	0.029	0.141	0.473		Input > Apex
Medium	0.156	0.481	0.758		Input < Apex
High	0.147	0.774	0.909		Input = Apex = 1
Very High	0.441	0.977	Infinity		
Uncertainty:	Fuzzy Sets:		c_{mi_i}	x_{i_i}	"Crisp":
0.062	VL, L	Input < Apex	0.86111111	0.072	0.103544643
		Input < Apex	0.29464286	0.141	
0.107	VL, L	Input > Apex	0.72000000	0.072	0.150036429
		Input < Apex	0.69642857	0.141	
0.088	VL, L	Input > Apex	0.87200000	0.072	0.137060786
		Input < Apex	0.52678571	0.141	
0.066	VL, L	Input < Apex	0.91666667	0.072	0.112580357
		Input < Apex	0.33035714	0.141	
0.063	VL, L	Input < Apex	0.87500000	0.072	0.105803571
		Input < Apex	0.30357143	0.141	
0.076	VL, L	Input > Apex	0.96800000	0.072	0.128865643
		Input < Apex	0.41964286	0.141	
0.055	VL, L	Input < Apex	0.76388889	0.072	0.087732143
		Input < Apex	0.23214286	0.141	
0.054	VL, L	Input < Apex	0.75000000	0.072	0.085473214
		Input < Apex	0.22321429	0.141	
0.098	VL, L	Input > Apex	0.79200000	0.072	0.143890071
		Input < Apex	0.61607143	0.141	
0.093	VL, L	Input > Apex	0.83200000	0.072	0.140475429
		Input < Apex	0.57142857	0.141	
0.081	VL, L	Input > Apex	0.92800000	0.072	0.132280286
		Input < Apex	0.46428571	0.141	
0.062	VL, L	Input < Apex	0.86111111	0.072	0.103544643
		Input < Apex	0.29464286	0.141	
0.117	VL, L	Input > Apex	0.64000000	0.072	0.156865714
		Input < Apex	0.78571429	0.141	
0.095	VL, L	Input > Apex	0.816	0.072	0.141841286
		Input < Apex	0.58928571	0.141	

Table 4.4.2-6 Schedule Uncertainty Fuzzy Algorithm for 14 Points

Collection of the two pair of sample data output processed through the fuzzy algorithm, historical and re-quantified IGE Uncertainty, allowed the application of a special case of two-sample t-Tests. A t-Test is appropriate because it serves as an appropriate guideline for making probability statements regarding the hypotheses. A paired t test was performed on the difference between the historical and re-quantified IGE uncertainty to determine whether the mean of the differences was zero. The initial assumption for this test includes that the differences are normally and independently distributed random variables with a mean of μ_D and a variance σ_D^2. Paired t-Testing on the differences is equivalent to testing the hypothesis $H_0 : u_1 = u_2$. The appropriate test statistic:

$$t_0 = \frac{\overline{D}}{S_D / \sqrt{n}}$$

Guidelines from the Student t distribution table for degrees of freedom, N-1 = 13, indicates a value of 1.771. These guidelines are used for making probability statements regarding the hypotheses. Data and descriptive statistics for the paired t-Test are depicted in Table 4.4.2-7.

Paired t Test Input Data				
Historical IGE	Requantified IGE	Difference		
0.062	0.206	0.143	*Column 1*	
0.107	0.293	0.216		
0.088	0.266	0.200	Mean	0.167836102
0.066	0.224	0.116	Standard Error	0.012703125
0.063	0.210	0.128	Median	0.172485581
0.076	0.249	0.201	Mode	#N/A
0.055	0.175	0.106	Standard Deviat	0.047530743
0.054	0.170	0.096	Variance	0.002259172
0.098	0.280	0.208	Kurtosis	-1.763663327
0.093	0.273	0.231	Skewness	-0.156903252
0.081	0.256	0.145	Range	0.134435269
0.062	0.206	0.135	Minimum	0.09614037
0.117	0.312	0.205	Maximum	0.230575639
0.095	0.276	0.219	Sum	2.349705432
			Count	14

Table 4.4.2-7 Paired t-Test of Historical and Re-quantified IGE Uncertainty

4.5 Hypothesis Testing

The objective of this research was to demonstrate a methodology of improving the quantification of technical and schedule uncertainty in DOD weapon system acquisition specifically related to a U.S. Army missile system. Use of linguistic variables and membership sets, elements of fuzzy logic, was integral to the methodology because these elements represent a new application of an important existing theory, the Theory of Fuzzy Logic. A multi level hypothesis was envisioned for this demonstration which consisted of a test of the means of various sets of sample data. One of the samples was a 46 point data set, the other was a 14 point data set that was a subset of the larger (46 data point) data set. Hypotheses were defined for: (1) the 46 point set, (2) the 14 point set, (3) a paired t test on the differences of two 14 point data sets and (4) on the re-quantification of the top level historical IGE compared to the historical negotiated prices.

4.5.1 The Forty Six Pair Set

There were 46 pair of data consisting of IGEs and Negotiated Contract Prices. Each data set represented independent samples of quantified cost composed of two elements; an uncertainty free portion and a quantified uncertainty portion. This hypothesis focused on the quantified uncertainty portion of each data set.

H_0: $\mu_1 = u_2$

The means of the two 46 point uncertainty data sets are equal.

H_1: $\mu_1 \neq u_2$

The means of the two 46 point uncertainty data sets are not equal.

The mean of the 46 point IGE uncertainty data set was calculated to be 0.072 while the mean of the Negotiated Price Uncertainty data set was calculated to be 0.212. Single factor ANOVA indicated the underlying distributions were not identical. The F distribution critical value corresponding to an $\alpha = 0.05$ and 91 degrees of freedom was calculated to be 3.94 and the actual F = 46.45. The conclusion was to reject H_0.

4.5.2 The Fourteen Pair Set

A subset of the 46 pair set of data consisting of IGEs and Negotiated Contract Prices was chosen for comparative purposes. Each data set represented independent samples of quantified cost composed of two elements: an uncertainty free portion and a quantified uncertainty portion. This hypothesis also focused on the quantified uncertainty portion of each data set.

H_o: $\mu_1 = u_2$

The means of the two 14 point uncertainty data sets are equal.

H_1: $\mu_1 \neq u_2$

The means of the two 14 point uncertainty data sets are not equal.

The mean of the 14 point IGE uncertainty data set was calculated to be 0.079 while the mean of the Negotiated Price Uncertainty data set was calculated to be 0.242. Single factor ANOVA indicated the underlying distributions were not identical. The F distribution critical value corresponding to an $\alpha = 0.05$ and 27 degrees of freedom was calculated to be 4.225 and the actual F = 156.198. The conclusion was to reject H_o.

4.5.3 <u>The Paired t-Test</u>

The t test performed on the 14 point data set focused on the differences between historical IGE uncertainty ratios and Re-quantified IGE ratios (using fuzzy algorithm). The objective of the paired t-Test was to determine if the two sets of IGE data were likely to produce the same mean, which is equivalent to testing:

H_o:$\mu_d = 0$

The mean of the difference between the two 14 point IGE uncertainty data sets is zero.

H_1:$\mu_d \neq 0$

The mean of the difference between the two 14 point IGE uncertainty data sets is not zero.

The mean of the difference was calculated to be 0.1678. The t table critical value for $t_{0.05,14} = 1.761$ and the calculated $t_0 = 13.212$. The conclusion was the two sets of data do not produce a zero mean therefore H_o was rejected.

4.5.4 <u>Top Level IGE and Negotiated Price</u>

The final hypothesis test was on the top level IGE cost and Negotiated Contract Price to determine results of applying the "crisp" output of the fuzzy algorithm. The "crisp" output was applied to the historical IGE "uncertainty free" cost to provide a comparison with the negotiated contract price. If the difference between the IGE cost and negotiated contract price was reduced it would represent an improved cost estimate. This implies there should be a spread of the data about the calculated mean. This spread is noted in Table 5.1-3 which compares the difference between negotiated price percent uncertainty and re-quantified IGE percent uncertainty and in Table 5.1-4 which compares absolute values of the negotiated contract prices with re-quantified IGE top level estimates. A paired t-Test was used to test the hypothesis.

$H_o: \mu_d = 0$

The mean of the difference between the 14 point IGE uncertainty data sets is zero.

$H_1: \mu_d \neq 0$

The mean of the difference between the two 14 point IGE uncertainty data sets is not zero.

The mean of the difference was calculated to be 0.0035. The t table critical value for $t_{0.05,14} = 1.761$ and the calculated $t_{0.05,14} = 0.077055$ The conclusion was the two sets of data could produce a zero mean therefore we fail to reject H_o. The mean of the difference was calculated to be very close to but not exactly zero and the critical value of $t_{0.05,14} = 0.077055 < t_{table}$ therefore H_1 cannot be accepted. This indicates the quantified uncertainty percent of the IGE moved in the direction of being much closer to the quantified negotiated price uncertainty thereby representing an improved cost estimate.

5. SUMMARY AND CONCLUSIONS

5.1 Discussion of Results

Application of certain elements of fuzzy logic, linguistic variables and fuzzy membership sets, have demonstrated in this research to produce an improved cost estimate. Further comparison is provided to support this conclusion. Table 5.1-1 compares the fourteen IGE historical top level estimates with the comparable fourteen historical contract negotiated prices. A percent delta is calculated to maintain the same "ratio" basis with which this research has proceeded. The mean is calculated to be 0.15159 with a standard deviation of 0.0683. This sub set of 14 data points represent IGEs and negotiated contract prices for procurements since 1992. Note that in each case the negotiated price is greater than the IGE, this was not the case for all 46 pair of data.

ADJUSTED DATA		% Delta		
IGE Historical	Negotiated Price	Contract-IGE		
103572786	120676548.841	0.142	*Column 1*	
69926650.37	75305623.472	0.071		
24132730.02	33182503.771	0.273	Mean	0.151599946
75305623.47	83129584.352	0.094	Standard Error	0.018272602
49410870.39	54161915.621	0.088	Median	0.149654021
1407742.584	1508295.626	0.067	Mode	#N/A
1955990.22	2444987.775	0.200	Standard Deviat	0.068369817
17578867.35	22329912.581	0.213	Variance	0.004674432
2660585.329	2850627.138	0.067	Kurtosis	-1.165318815
596577.0171	684596.577	0.129	Skewness	0.213100905
44243338.36	58320764.203	0.241	Range	0.206060606
18529076.4	22805017.104	0.188	Minimum	0.066666667
135941320.3	161369193.154	0.158	Maximum	0.272727273
47510452.3	58912960.851	0.194	Sum	2.122399247
			Count	14

Table 5.1-1 Comparison of IGE Historical Top Level Estimate with Historical Negotiated Contract Price

Two additional Paired t-Tests were performed: (1) Negotiated Price percent uncertainty Vs Historical IGE percent uncertainty and (2) Negotiated Price percent uncertainty Vs Re-quantified IGE percent uncertainty. Table 5.1-2 provides the results of the first Paired t-Test that indicates a mean value of 0.1662. Table 5.1-3 provides results of the second Paired t-Test that indicates a mean value of 0.0035. These statistics indicate the mean values of the re-quantified IGE percent uncertainty are much closer to the mean values of the negotiated price percent uncertainty. Large differences are still noted in three cases where the negotiated contract percent uncertainty is 20 percent or greater than the re-quantified IGE percent uncertainty.

Paired t Test Input Data				
Negotiated Price % Uncertainty	Historical IGE % Uncertainty	Difference		
0.198	0.062	0.135	*Column 1*	
0.083	0.107	-0.023		
0.600	0.088	0.512	Mean	0.166226996
0.116	0.066	0.050	Standard Error	0.043777573
0.106	0.063	0.043	Median	0.124074591
0.077	0.076	0.001	Mode	#N/A
0.333	0.055	0.278	Standard Deviat	0.163800678
0.370	0.054	0.316	Variance	0.026830662
0.077	0.098	-0.021	Kurtosis	-0.273208772
0.173	0.093	0.080	Skewness	0.711639437
0.467	0.081	0.386	Range	0.535042438
0.300	0.062	0.238	Minimum	-0.023477812
0.230	0.117	0.113	Maximum	0.511564626
0.316	0.095	0.220	Sum	2.327177944
			Count	14

Table 5.1-2 Paired t-Test of Negotiated Price Percent Uncertainty Vs Historical IGE Percent Uncertainty

Paired t Test Input Data				
Negotiated Price % Uncertainty	Requantified IGE % Uncertainty	Difference		
0.198	0.206	-0.008	*Column 1*	
0.083	0.293	-0.210		
0.600	0.266	0.334	Mean	0.003545951
0.116	0.224	-0.108	Standard Error	0.046022692
0.106	0.210	-0.104	Median	-0.044939999
0.077	0.249	-0.172	Mode	#N/A
0.333	0.175	0.159	Standard Deviat	0.172201146
0.370	0.170	0.200	Variance	0.029653235
0.077	0.280	-0.203	Kurtosis	-0.889782274
0.173	0.273	-0.100	Skewness	0.503767061
0.467	0.256	0.211	Range	0.543837779
0.300	0.206	0.094	Minimum	-0.209917154
0.230	0.312	-0.082	Maximum	0.333920625
0.316	0.276	0.040	Sum	0.049643311
			Count	14

Table 5.1-3 Paired t-Test of Negotiated Price Percent Uncertainty Vs Re-quantified IGE Percent Uncertainty

To summarize comparison of the re-quantified IGE, which included the "crisp" output of the fuzzy algorithm, and the historical negotiated contract prices Table 5.1-4 was developed. The mean changed remarkably to a negative value of -0.0548 as compared to the mean value provided in Table 5.1-1 of 0.15159. Review of the table indicates a much different pattern of calculated ratio values about the mean. This pattern indicates the calculated IGE values range above and below the historical negotiated contract values in a random manner. This data provides a sharp contrast with data in Table 5.1-1 which indicate in each case the IGE values were below negotiated contract values.

ADJUSTED DATA						
IGE Uncertainty Free	Requantified IGE Uncertainty	IGE Requantified Total	Negotiated Contract Price	% Delta Contract-IGE		
103572786.013	1.206	124901934.2	120676548.8	-0.035	*Column 1*	
69926650.367	1.293	90432674.66	75305623.47	-0.201		
24132730.015	1.266	30553951.72	33182503.77	0.079	Mean	-0.0548398
75305623.472	1.224	92153095.21	83129584.35	-0.109	Standard Erro	0.02674592
49410870.391	1.210	59805966.49	54161915.62	-0.104	Median	-0.0696106
1407742.584	1.249	1758155.999	1508295.626	-0.166	Mode	#N/A
1955990.220	1.175	2297908.985	2444987.775	0.060	Standard Dev	0.10007408
17578867.351	1.170	20573587.73	22329912.58	0.079	Variance	0.01001482
2660585.329	1.280	3406560.137	2850627.138	-0.195	Kurtosis	-1.4285322
596577.017	1.273	759579.5571	684596.577	-0.110	Skewness	0.00176162
44243338.361	1.256	55572684.21	58320764.2	0.047	Range	0.28009045
18529076.397	1.206	22344841.45	22805017.1	0.020	Minimum	-0.2008755
135941320.293	1.312	178332824.8	161369193.2	-0.105	Maximum	0.079215
47510452.300	1.276	60627602.84	58912960.85	-0.029	Sum	-0.7677568
					Count	14

Table 5.1-4 Comparison of Re-quantified IGE Top Level Estimate with Historical Negotiated Contract Price

5.2 Limitations

This research is constrained in scope and has certain limitations which should be enumerated. DOD weapon systems acquisition processes is the context specific environment to which this research is directed. It is recognized that the U.S. Military-Industrial complex is a unique environment. Within this complex the research was further directed toward the acquisition management processes of a specific type of system; a weapon system, and a specific type of weapon system; a missile weapon system. The survey instrument was context specific directed toward definition of two aspects of uncertainty: technical and schedule. Context specificity extended to individual interpretation of definitions contained within the survey instrument. Limitations of the definitions associated with the linguistic variables focused on one specific factor for technical uncertainty, requirements, and one specific factor for schedule uncertainty, contract duration.

The method of the study could not address situations in which cost estimates were consistently above negotiated contract prices. A major shortcoming of the estimating process is that IGEs are consistently below subsequently negotiated contract prices. Given this fact the direction undertaken in the research was to develop definitions that addressed increases in technical difficulty (requirements) and increases in schedule (contract time/duration). Survey size was adequate, 30 participants, but not exceptionally large.

Finally, use of fuzzy logic concepts were limited to two elements: linguistic variables and membership sets. There was no attempt to expound on the calculus of fuzzy logic since the mathematical basis of the theory is well founded and has been the subject of other research efforts.

5.3. Future Research

There are several lucrative areas for research relating concepts of fuzzy logic to quantification of cost for DOD weapon systems. Three areas deserve mention in concluding this research: (1) reliability, (2) learning curves and man loading (level of effort) and (3) linking fuzzy logic to cognitive maps and neural networks.

Specification of weapon system technical requirements is dependent on individual human judgment which is the initial source of imprecision in the requirements development process. The relationship of technical requirements to uncertainty in this research was depicted in the context of reliability engineering related to hardware system effectiveness and design specification. In each of these reliability areas the inclusion of vagueness and imprecision suggests the opportunity for development of a methodology for quantification of uncertainty. Improving the credibility, efficiency and effectiveness, of failure data bases and failure analysis is a source for application of fuzzy logic techniques. Bazovsky and Benz (1993:372-373) suggest an application of fuzzy logic to reliability data bases in which combinations of synonyms are matched to failure symptoms noted in hardware failure reports. Specific application relates to developing membership sets of linguistic variables such as "works well" for explicitly defined criteria related to quantitative measurements stipulated in the specification and acceptance test procedures. The relationship of failure experience to cost estimating exists in the area of warranty cost estimating. Factors related to warranty uncertainty include: (1) the characteristic addressed under warranty, (2) price, (3) operational factors, (4) self sufficiency, (5) equipment design, (6) transition, and (7) administrative complexity. Each factor can be related to fuzzy logic through the elements of linguistic variables and membership sets.

Learning curves are a major factor in cost estimates that are not likely to loose significance. However, learning curve application has undergone much re-analysis during the past fifteen years due to changes in manufacturing technologies and materials. Data relating to current technologies: agile manufacturing, stereolithography, robotics and changes in quality such as Motorola's six sigma technique have not made an impact in application of learning curves for DOD weapon systems. The evolution of manufacturing technologies is impacting cost estimating in ways which directly affect changes to learning curve applications which could benefit from use of linguistic variables and use of membership sets. Another major area of manufacturing, stereolithography, has drastically reduced the time involved in a prototyping. This change alone impacts level of effort related to research and development in such a significant manner that some estimates indicate cost can be reduced by as much 90 percent for certain applications (Comerford, 1993:29).

Fuzzy logic concepts could be applied in the DOD environment where vagueness and imprecision in cost modeling is the rule rather than the exception. Quantification of difficult concepts related to reliability, learning curves, and manufacturing technology can be handled with linguistic variables and fuzzy membership sets in a manner which could move cost estimating in the direction of the cost estimating/modeling state-of-the art. Further, there is a natural linking of fuzzy cognitive maps, fuzzy neural network techniques, to fuzzy logic for example in defining weights from fuzzy sets for processing in neural networks.

The complexity of the cost estimating process can be modeled as a neural network utilizing fuzzy cognitive maps. Such techniques as fuzzy logic and neural networks are ripe for modeling current DOD procurement environment. Remarkable changes have occurred in such areas as international teaming agreements, associative contractor arrangements rather than the traditional prime-sub relationship. Understanding the current environment combined with utilization of robust techniques for "handling" such complexities are a major challenge facing the DOD cost analysis community. Based upon these technology examples, further investigation into the application of fuzzy logic concepts to cost estimating for DOD weapon system acquisition is needed.

REFERENCES

Ahmed, A.N. 1991. "Characterization of Beta, Binomial, and Poisson Distribution." IEEE Transactions On Reliability, 40.3 (Augus), pp. 290-95.

ARINC Research Corporation. 1964. Reliability Engineering (New Jersey: Prentice-Hall), p. 14.

Bai, D.S., Kim, M.S. and Lee, S.H. 1989. "Optimum Simple Step-Stress Accelerated Life Tests with Censoring." IEEE Transactions On Reliability, 38.5 (December), pp. 528-532.

Bazovsky, Igor, Benz, G.E. 1993. "An Application of Fuzzy Logic to Reliability." Proceedings of the 1993 Annual Reliabilty and Maintainability Symposium, 25 (January), pp. 372-373.

Bowles, John B. 1992. "A Survey Of Reliability-Prediction Procedures For Microelectronic Devices." IEEE Transactions On Reliability, 41.1 (March), pp. 2-12.

Butts, George. 1989. Retired Chrysler Corporation Vice President, Verbal Remarks at TQM Seminar, Huntsville, Alabama, 5 September.

Clements, J. 1995. Fuzzify. Custom Computer Software, Disk.

Comerford, Richard. 1993. "A Quick Look at Rapid Prototyping." IEEE Spectrum, September, pp. 28-29.

Coppola, Anthony. 1984. "Reliability Engineering of Electronic Equipment A Historical Perspective." IEEE Transactions On Reliability, 33.1 (April), pp. 29-35.

Cox, Earl. 1992. "Fuzzy Fundamentals." IEEE Spectrum, October, pp. 58-61.

Cox, Earl. 1993. "Adaptive Fuzzy Systems." IEEE Spectrum, February, pp. 27-31.

Defense System Management College. 1983. Risk Management Concepts and Guidance. Update has no date, pp. 5-11, 5-38, F-2.

Department of the Air Force. 1995. Draft Guide to Cost/Schedule Management. Headquarters Space and Missile Systems Center, January, p. 1-3.

Department of Defense. 1991. Defense Acquisition Documentation and Reports. Under Secretary of Defense for Acquisition, February, pp. 8-8 - 8-9.

Department of the Army. 1992. Cost Analysis Manual. U.S. Army Cost and Economic Analysis Center, August, pp. 35, 48-50.

Department of Defense. 1985. Assistant Secretary of Defense Acquisition and Logistics. Transition From Development To Production, Solving the Risk Equation, pp. 1-3.

Department of the Navy. 1986. Best Practices. Reliability, Maintainability, and Quality Assurance Directorate, pp. 4-3.

Dorr, Oscar J. 1990. "Logistics Peace Dividend With Interest." Logistics Spectrum. 24.1 (Spring), pp. 5-8.

Garvey, Paul R. 1990. "A General Analytic Approach to System Cost Uncertainty Analysis." Edited by Willis R. Greer and Daniel A Nussbaum. Cost Analysis and Estimating (New York: Springer-Verlag), 161-181.

Hisdale, Ellen. 1994. "Interpretative Versus Prescriptive Fuzzy Set Theory." IEEE Transactions on Fuzzy Systems, 2.1 (February), pp.22-25.

Holland, John. 1989. "Using Classifier Systems to Study Adaptive Nonlinear Networks. "Lectures In The Sciences Of Complexity Edited by Daniel L. Stein. Sante Fe Institute Studies In the Sciences Of Complexity, (California: Addison-Wesley Publishing Company), pp. 463-499.

Kamke, E. 1950. Theory of Sets (New York, Dover Publications, Inc.), p. 1.

Kuratowski, Kazimierz. 1972. Introduction to Set Theory and Topology (Oxford: Pergamon Press), pp. 27-33.

Lindh, Lars. 1993. "A Fuzzy Logic Engine for Turbo Pascal." AI Expert, October, p. 36.

McCauley, J. J. 1994. "A Fuzzy-Logic Toruqe Servo." The C Users Journal, March, p.49.

McNichols, Gerald M. 1984. "The State of the Art of Cost Uncertainty Analysis." The Journal Of Cost Analysis, 1.1 (Spring), pp. 149-173.

Nikhil, R. Pal, and Bezdek, James C. 1994. "Measuring Fuzzy Uncertainty." IEEE Transactions On Fuzzy Systems, 2.2 (May), pp. 107-118.

Schoenfield, J. 1995. QuickBASIC Program for Generating Degrees of Membership. Custom Computer Software, Disk.

Schwartz, D.G., and Klir, G.J. 1992. "Fuzzy Logic Flowers in Japan." IEEE Spectrum, July, pp.32-35.

Self, K. 1990. "Designing with Fuzzy Logic." IEEE Spectrum. November, pp. 42-44, 105.

Sperry, Tyler. 1993. "Gaining a Fuzzy Perspective." Embedded Systems Programming, 6.1 January, pp. 31-34.

Swartz, Daniel G., and Klir, George J. 1992. "Fuzzy Logic Flowers in Japan." IEEE Spectrum, July, pp. 32-35.

Valluru, B. Rao and Hayagriva, V. Rao. 1993. C++ Neural Networks and Fuzzy Logic. (New York: Management Information Source, Inc. A subsidiary of Henry Holt and Company, Inc.), p. 26.

van Dalen, D. 1972. Sets and Integration An Outline of the Development. (Groningen, The Netherlands, Wolters-Noordhoff Publishing), pp. 3-4.

Verma, Dinesh. 1994. "Application Of Fuzzy Logic In The Assurance Sciences." 1994 Proceedings Annual Reliability and Maintainability Symposium. (Anaheim, California), pp. 436-439.

Vinod, H.D., and Basu, Parantap. 1990. Cost Analysis and Estimating Tools and Techniques., A Series Of Technical Papers edited by Willis R. Greer, Jr., and Daniel A. Nussbaum. (New York: Springer-Verlag), pp. 3-29.

Voit, Greg. 1993. "Fuzzy Logic: Concepts to Constructs." AI Expert November, pp. 26-33.

Watson, Stephen R., Weiss, Jonathan J., and Donnell, Michael L. 1979. "Fuzzy Decision Analysis." IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-9 No. 1, (January), pp. 1-9.

Wehrle, Robert A. 1990. "Focus on the Process Closing of the Military Design." Logistics Spectrum, 24.1 (Spring), pp.21-22.

Wiskerchen, M.J. and Pittman, R. Bruce. 1989. "Dynamic Systems Engineering Process: The Application of Concurrent Engineering." Engineering Management Journal, 1.2, pp. 27-34.

Zadeh, L.A. 1965. "Fuzzy Sets." Information and Control, 8 (June), pp. 338-353.

Zadeh, L.A. 1973. "Outline of a New approach to the Analysis of Complex Systems and Decision Processes." IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3 No. 1 (January), pp. 28-44.

Zadeh, L.A. 1984. "Making Computers Think like People." IEEE Spectrum, August pp. 26-32.

Zimmerman, H. J. 1987. Fuzzy Sets, Decision Making and Expert Systems (Boston, Dordrecht, Lancaster, Kluwer Academic Publishers), pp. 10-11.

APPENDIX A

APPENDIX A

COST ELEMENT STRUCTURE

0.0 Total

1.0 Research, Development, Test, and Evaluation (RDT&E) Funded Elements

1.01 Development Engineering

1.02 Producibility Engineering and Planning (PEP)

1.03 Development Tooling

1.04 Prototype Manufacturing

1.05 System Engineering/Program Management

1.051 Project Management Administration (PM Civ/Mil)

1.052 Other

1.06 System Test and Evaluation

1.07 Training

1.08 Data

1..09 Support Equipment

1.091 Peculiar

1.092 Common

1.10 Development Facilities

1.11 Other RDT&E

2.0 Procurement Funded Elements

2.01 Nonrecurring Production

2.011 Initial Production Facilities (IPFs)

2.012 PRoduction Base Support (PBS)

2.013 Other Nonrecurring Production

2.02 Recurring Production

2.021 Manufacturing

2.022 Recurring Engineering

2.023 Sustaining Tooling

2.024 Quality Control

2.025 Other Recurring Production

2.03 Engineering Changes

2.04 System Engineering/Project Management

2.041 Project Management Administration (PM Civ/Mil)

2.042 Other

2.05 SystemTest and Evaluation

2.06 Training

2.07 Data

2.08 Support Equipment

2.081 Peculiar

2.082 Common

2.09 Operational/Site Activation

2.10 Fielding

2.101 Initial Depot-Level Reparables (Spares)

2.102 Initial Consumables (Repair Parts)

2.103 Initial Support Equipment

2.104 Transportation (Equipment to Unit)

2.105 New Equipment Training (NET)

2.106 Contractor Logistics Support

2.11 Training Ammunition/Missiles

2.12 War Reserve Ammunition/Missiles

2.14 Other Procurement

3.0 Military Construction (MC) Funded Elements

3.01 Development Construction

3.02 Production Construction

3.03 Operational/Site Activation Construction

3.04 Other MC

4.0 Military Personnel (MP) Direct Funded Elements

4.01 Crew

4.02 Maintenance (MTOE)

4.03 System-Specific Support

4.04 System Engineering/Project Management

4.041 Project Management Administration (PM Mil)

4.042 Other

4.05 Replacement Personnel

4.051 Training

4.052 Permanent Change of Station (PCS)

4.06 Other MP

5.0 Operations and Maintenance (O&M) Funded Elements

5.01 Field Maintenance Civilian Labor

5.02 System-Specific Base Operations

5.03 Replenishment Depot-Level Reparables (Spares)

5.04 Replemenishment Consumables (Repair Parts)

5.05 Petroleum, Oil and Lubricants (POL)

5.06 End-Item Supply and Maintenance

5.061 Overhaul (P7M)

5.062 Integrated Materiel Management

5.063 Supply Depot Support

5.064 Industrial Readiness

5.065 Demilitarization

5.07 Transportation

5.08 Software

5.09 System Test and Evaluation, Operational

5.10 System Engineering/Project Management

5.101 Project Management Administration (PM Civ)

5.102 Other

5.11 Training

5.12 Other O&M

APPENDIX B

APPENDIX B

TECHNICAL AND SCHEDULE UNCERTAINTY
SURVEY

1. Purpose of the Survey: To determine if there is potential for improving the quantification of uncertainty (sometimes used interchangeably with risk) related to cost estimates. Your input is critical to the final results of a cost research effort which is in process. Thank you for your participation. It is not necessary to identify yourself.

2. Participant Data:

Status	Functional Organization	Education (Degree/Major)
A. Government:	____________________	1. ____________________
B. Contractor:	____________________	2. ____________________
		3. ____________________

3. Overview:

The survey is in two parts, technical and schedule, please complete both parts.

Please begin with a review of the definitions of the five categories of uncertainty, then apply the definitions, based on your experience and opinion, by establishing a percent range for each category.

What does this percent range represent?

A. Technical: The percent range of uncertainty is your opinion of the **increase in technical difficulty** imposed by the definitions as opposed to a program where virtually no technical uncertainty (or risk) exists.

B. Schedule: The percent range of uncertainty is your opinion of the **increase in contract time/duration** imposed by the definitions as opposed to a program where virtually no schedule related uncertainty (or risk) exists.

ANNEX B (Continued)

SURVEY QUESTIONNAIRE

TECHNICAL UNCERTAINTY (Risk) Definitions:

Very Low; Components or subsystems affecting **REQUIREMENTS** have been developed and are based on existing technology. Analysis, simulation and testing have demonstrated a high probability of successfully meeting requirements.

Low; Components or subsystems affecting **REQUIREMENTS** have been developed and are based on existing technology. Limited analyses indicate a high probability of successfully meeting requirements.

Medium; Some components or subsystems affecting **REQUIREMENTS** have been developed and it is judged based on analysis, simulation or testing that the requirements can be met. Components to be developed are based on existing technology.

High; Many components or subsystems affecting **REQUIREMENTS** represent new technology or have not been developed, or have not been integrated into subsystems, or have not been field tested.

Very High; Majority of components or subsystems affecting **REQUIREMENTS** represent new technology or have not been developed, or have not been integrated into subsystems or have not been field tested.

NOTE: Overlapping of percent range from one category to another is allowed.

Example:	Very Low	5% to 15%
	Low	12% to 20%

CATEGORY OF UNCERTAINTY	% RANGE OF UNCERTAINTY (From___ To ___)	
Very Low	_______	_______
Low	_______	_______
Medium	_______	_______
High	_______	_______
Very High	_______	_______

SURVEY QUESTIONNAIRE

SCHEDULE UNCERTAINTY Definitions:

It is acknowledged that schedule uncertainty is not independent of technical uncertainty. However, an attempt is made in this portion of the survey to focus attention on the occurrence of activities or events which are "relatively" less technical and "relatively" more schedule direct. "Schedule" is to be associated with the contract duration (number of months) of a program. "Early" development of plans should be interpreted as desirable. The following list, not comprehensive, is intended to provide an idea of activities or events whose occurrence would have relatively less technical and relatively more schedule direct impacts.

a. Schedule realism (adequate time to accomplish program requirements)
b. Staffing for contract execution (quantity of personnel and skill levels)
c. Test programs; scheduling of assets, range availability etc.
d. Completeness of formal documentation (TEMPs, ILSPs, Specifications)
e. Management Plans; System Engineering, Risk, Configuration Management
f. Interrelated schedules developed from top down to work-package level.
g. Work Breakdown Structure (WBS) expanded to successively lower levels and associated with both a particular organization and a budgeted task
h. Data Item Description (DID) requirements sufficient for contract performance measurement and cost estimating/analysis
i. Major vendors identified and qualified
j. Material schedules and lead times developed
k. No unfunded program requirements
l. Program Advocacy stable (support for program at highest levels)

Very Low; Realistic schedule, adequate contractor staffing. Formal program plans and documentation (including contract related) developed and adequate to alleviate government concerns throughout **Contract Duration** (items c-h) . Contractor accounting system adequate for performance measurement.. Material planning adequate, all requirements funded, stable program advocacy. No known problems.

Low; Realistic schedule, adequate contractor staffing. Formal program plans and documentation (including contract related) developed and adequate to alleviate government concerns throughout **Contract Duration** (items c-f) . WBS expanded to organizational level, budgets not yet allocated. DIDs sufficient for measuring performance but lack detail level for complete cost estimating/analysis. Contractor accounting system adequate for performance measurement.. Most major vendors identified and qualified. Material planning underway but not complete, all requirements funded, stable program advocacy.

Medium; Schedule realism questionable, adequate contractor staffing for contract initiation, staff build up judged questionable. Formal program plans and documentation (including contract related) partially developed for **Contract Duration** (items c-f) . WBS expanded to organizational level, budgets not yet allocated. DIDs sufficient for measuring performance but lack detail level for complete cost estimating/analysis. Contractor accounting system adequate for performance measurement.. Several (not most) major vendors identified but not qualified. Material planning underway but not complete, all requirements funded, stable program advocacy.

High; Schedule realism highly questionable, adequate contractor staffing for contract initiation, staff build up judged questionable. Formal program plans and documentation (including contract related) partially developed for **Contract Duration** (items c-f) . WBS not expanded to organizational level, budgets not yet allocated. DIDs questionable for measuring performance, no detail level for cost estimating/analysis. Major vendors not completely identified. Material planning underway but not complete. Contract requirements funded, discussions with contractor to expand scope of work prior to contract definitization. Stable program advocacy.

Very High; Schedule realism highly questionable, adequate contractor staffing for contract initiation, staff build up judged questionable. Formal program plans and documentation (including contract related) partially developed for **Contract Duration** (items c-f) . WBS not expanded to organizational level, budgets not yet allocated. DIDs questionable for measuring performance, no detail level for cost estimating/analysis. Major vendors not completely identified. Material planning underway but not complete. Initial contract requirements funded (first 3 years of 5 year contract). Due to constrained budget environment, stable program advocacy cannot be assured.

NOTE: Overlapping of percent range from one category to another is allowed.

Example:	Very Low	5% to 15%
	Low	12% to 20%

CATEGORY OF UNCERTAINTY	% RANGE OF UNCERTAINTY (From___ To ___)	
Very Low	______	______
Low	______	______
Medium	______	______
High	______	______
Very High	______	______

APPENDIX C

APPENDIX C

Contract Price Data

Contracts	Performance	Negotiated Amount
DAAH01-87-C-	Dec86-Oct94	
Multi-Year Missiles		
Ground Support Eq.		
Engineering Svcs.		
FMS GSE		
FMS GSE		
Mod1	Jun86-May88	$128,597,673
Mod2	Sept86-Nov88	$113,900,796
Mod3	Apr87-Apr90	$265,848,671
Mod4	Jul87-Jul89	$77,903,683
Mod5	Sept87-Aug88	$1,298,395
Mod6	Dec87-Dec91	$35,219,268
Mod7	Jun88-Jun91	$326,503,780
Mod8	Aug88-Aug89	$1,022,495
Mod9	Nov88-Nov89	$142,013,179
Mod10	Feb89-Feb91	$219,130,054
Mod11	May89-May91	$232,277,857
Mod12	Dec89-Dec91	$51,495,563
Mod13	Mar90-Feb92	$200,275,511
Mod14	Jul90-Jul91	$1,271,591
Mod15	Nov90-Nov93	$154,702,970
Mod16	Jan91-Jan93	$149,546,205
Mod17	Jun91-Jun93	$16,501,650
Mod18	Aug91-Aug93	$137,169,967
Mod19	Apr92-Apr94	$57,817,999
Mod20	Sep92-Sep94	$92,006,033
Mod21	Jun93-Dec94	$1,271,394
DAAH01-91-C-0625	Sept91-Feb94	$70,387,129
DAAH01-91-C-0602	Sept91-Sept94	$50,276,521
DAAH01-89-C-0458	Sept91-Feb95	$61,337,355
DAAH01-92-C-0134	Feb92-Aug94	$29,663,147
DAAH01-92-C-0338	Sept92-Sept94	$276,521
DAAH01-92-C-0301	Jul92-Sept97	$4,107,579
DAAH01-92-C-0112	Jan92-Dec95	$1,955,990
DAAH01-92-C-0079	Jan92-Feb94	$2,513,826
DAAH01-93-C-0033	Nov92-Jan94	$31,171,443
DAAH01-93-C-0096	Nov92-Feb94	$4,022,122
DAAH01-92-C-0363	Jul92-Sept97	$120,676,549
DAAH01-92-C-0251	May92-Jun96	$75,305,623

DAAH01-92-C-0157	Feb92-Aug94	$33,182,504
DAAH01-93-C-0149	Feb93-Nov95	$83,129,584
DAAH01-94-C-0107	Dec93-Aug97	$54,161,916
DAAH01-93-C-0126	Feb93-Feb94	$1,508,296
DAAH01-93-C-0124	Feb93-Oct94	$2,444,988
DAAH01-93-C-0294	Jul93-Aug96	$22,329,913
DAAH01-94-C-0014	Oct93-May95	$2,850,627
DAAH01-93-C-0346	Sep93-Aug95	$684,597
DAAH01-89-C-0458	Feb92-Aug94	$58,320,764
DAAH01-91-C-0602	Apr93-Jun96	$22,805,017
DAAH01-92-C-0006	Mar94-Feb95	$161,369,193
DAAH01-94-C-0105	Jan94-Dec97	$58,912,961
DAAH01-89-C-0008	Jan94-Dec94	$11,735,941

APPENDIX D

APPENDIX D

Quickbasic Program for Generating

Degrees of Membership

Program, DEGMEMB.BAS:

```
' QuickBASIC program to compute degree of membership in each of five
' fuzzy sets . . . .

DEFINT I-N

CONST NUMPARTS = 30   ' number of participants
CONST NUMSETS = 5     ' number of fuzzy sets (very lo, lo, med, hi, very hi)

DIM R(NUMPARTS, NUMSETS, 2)

' R(I, J, 1) is the low end of the range of values assigned by
'              participant I to fuzzy set J
'
' R(I, J, 2) is the high end of the range


OPEN "TECH.DAT" FOR INPUT AS 1   ' contains only the Technical Risk data
OPEN "TECH.OUT" FOR OUTPUT AS 2  ' will contain degree-of-membership data

PRINT #2, "Echo-check of data:"
PRINT #2,


INPUT #1, dummy

FOR I = 1 TO NUMPARTS
  INPUT #1, dummy1, dummy2
  FOR J = 1 TO NUMSETS
    INPUT #1, R(I, J, 1), R(I, J, 2)
    PRINT #2, USING "#.## "; R(I, J, 1); R(I, J, 2);
  NEXT J
  PRINT #2,
NEXT I

PRINT #2,
PRINT #2,
PRINT #2,
PRINT #2, "Technical      Degree of Membership times 30"
PRINT #2, "  Risk      VL     L     M     H     VH"
PRINT #2, "---------   ------ ------ ------ ------ ------"

FOR IX = 1 TO 100
  X = (IX - .5) / 100
  PRINT "Checking degree of membership for X = "; X
  XLO = (IX - 1) / 100
```

```
  XHI = IX / 100
  PRINT #2, USING "#.##"; XLO;
  PRINT #2, "-";
  PRINT #2, USING "#.##"; XHI;
  FOR J = 1 TO NUMSETS
    MemberCount = 0
    FOR I = 1 TO NUMPARTS
      IF X > R(I, J, 1) AND X < R(I, J, 2) THEN
        MemberCount = MemberCount + 1
      END IF
    NEXT I
    PRINT #2, USING "########"; MemberCount;
  NEXT J
  PRINT #2,
NEXT IX
```

Echo-check of data:

```
0.10 0.20 0.15 0.40 0.35 0.60 0.55 0.80 0.75 1.00
0.00 0.03 0.03 0.25 0.25 0.35 0.35 0.60 0.60 1.00
0.05 0.20 0.20 0.35 0.35 0.65 0.65 0.85 0.85 1.00
0.00 0.15 0.10 0.20 0.15 0.80 0.80 0.95 0.90 1.00
0.00 0.10 0.05 0.45 0.40 0.70 0.70 0.90 0.85 1.00
0.05 0.10 0.10 0.25 0.25 0.50 0.50 0.75 0.75 1.00
0.01 0.07 0.05 0.15 0.15 0.30 0.25 0.45 0.45 1.00
0.00 0.10 0.10 0.20 0.18 0.50 0.50 0.90 0.75 1.00
0.00 0.10 0.10 0.33 0.33 0.66 0.66 0.90 0.90 1.00
0.00 0.10 0.10 0.30 0.30 0.50 0.50 0.75 0.75 1.00
0.00 0.10 0.11 0.25 0.26 0.50 0.51 0.75 0.76 1.00
0.00 0.20 0.20 0.30 0.30 0.45 0.40 0.65 0.65 1.00
0.00 0.15 0.15 0.35 0.35 0.55 0.55 0.85 0.85 1.00
0.00 0.20 0.20 0.40 0.40 0.60 0.60 0.80 0.80 1.00
0.20 0.30 0.40 0.50 0.50 0.65 0.65 0.80 0.80 1.00
0.00 0.10 0.11 0.35 0.36 0.79 0.80 0.94 0.95 1.00
0.05 0.10 0.10 0.25 0.25 0.40 0.35 0.65 0.60 1.00
0.05 0.15 0.16 0.25 0.26 0.75 0.76 0.85 0.86 1.00
0.00 0.05 0.05 0.15 0.15 0.70 0.70 0.85 0.85 1.00
0.05 0.10 0.10 0.25 0.20 0.35 0.35 0.55 0.50 1.00
0.05 0.15 0.16 0.35 0.36 0.65 0.66 0.80 0.81 1.00
0.05 0.15 0.15 0.25 0.25 0.40 0.40 0.55 0.50 1.00
0.10 0.25 0.20 0.40 0.40 0.65 0.65 0.80 0.75 1.00
0.00 0.10 0.10 0.20 0.20 0.35 0.35 0.65 0.65 1.00
0.00 0.14 0.11 0.25 0.19 0.45 0.40 0.70 0.65 1.00
0.00 0.15 0.16 0.29 0.30 0.70 0.71 0.90 0.91 1.00
0.00 0.10 0.05 0.20 0.20 0.40 0.40 0.60 0.50 1.00
0.00 0.06 0.05 0.27 0.20 0.50 0.50 0.80 0.80 1.00
```

APPENDIX D (Continued)

Quickbasic Program for Generating

Degrees of Membership

0.10 0.18 0.15 0.22 0.18 0.30 0.30 0.50 0.50 1.00
0.00 0.05 0.06 0.15 0.16 0.70 0.71 0.75 0.76 1.00

Results of Technical Uncertainty Data

Technical Risk	Degree of Membership times 30 VL	L	M	H	VH
---------	------	------	------	------	------
0.00-0.01	18	0	0	0	0
0.01-0.02	19	0	0	0	0
0.02-0.03	19	0	0	0	0
0.03-0.04	18	1	0	0	0
0.04-0.05	18	1	0	0	0
0.05-0.06	23	6	0	0	0
0.06-0.07	22	7	0	0	0
0.07-0.08	21	7	0	0	0
0.08-0.09	21	7	0	0	0
0.09-0.10	21	7	0	0	0
0.10-0.11	13	15	0	0	0
0.11-0.12	13	18	0	0	0
0.12-0.13	13	18	0	0	0
0.13-0.14	13	18	0	0	0
0.14-0.15	12	18	0	0	0
0.15-0.16	6	19	3	0	0
0.16-0.17	6	22	4	0	0
0.17-0.18	6	22	4	0	0
0.18-0.19	5	22	6	0	0
0.19-0.20	5	22	7	0	0
0.20-0.21	2	22	11	0	0
0.21-0.22	2	22	11	0	0
0.22-0.23	2	21	11	0	0
0.23-0.24	2	21	11	0	0
0.24-0.25	2	21	11	0	0
0.25-0.26	1	13	15	1	0
0.26-0.27	1	13	17	1	0
0.27-0.28	1	12	17	1	0
0.28-0.29	1	12	17	1	0
0.29-0.30	1	11	17	1	0
0.30-0.31	0	9	18	2	0
0.31-0.32	0	9	18	2	0
0.32-0.33	0	9	18	2	0
0.33-0.34	0	8	19	2	0
0.34-0.35	0	8	19	2	0
0.35-0.36	0	4	19	6	0
0.36-0.37	0	4	21	6	0
0.37-0.38	0	4	21	6	0

APPENDIX D (Continued)

Quickbasic Program for Generating Degrees of Membership

Results for Technical Uncertainty Data

0.38-0.39	0	4	21	6	0
0.39-0.40	0	4	21	6	0
0.40-0.41	0	2	21	10	0
0.41-0.42	0	2	21	10	0
0.42-0.43	0	2	21	10	0
0.43-0.44	0	2	21	10	0
0.44-0.45	0	2	21	10	0
0.45-0.46	0	1	19	9	1
0.46-0.47	0	1	19	9	1
0.47-0.48	0	1	19	9	1
0.48-0.49	0	1	19	9	1
0.49-0.50	0	1	19	9	1
0.50-0.51	0	0	15	12	5
0.51-0.52	0	0	15	13	5
0.52-0.53	0	0	15	13	5
0.53-0.54	0	0	15	13	5
0.54-0.55	0	0	15	13	5
0.55-0.56	0	0	14	13	5
0.56-0.57	0	0	14	13	5
0.57-0.58	0	0	14	13	5
0.58-0.59	0	0	14	13	5
0.59-0.60	0	0	14	13	5
0.60-0.61	0	0	12	12	7
0.61-0.62	0	0	12	12	7
0.62-0.63	0	0	12	12	7
0.63-0.64	0	0	12	12	7
0.64-0.65	0	0	12	12	7
0.65-0.66	0	0	8	12	10
0.66-0.67	0	0	7	14	10
0.67-0.68	0	0	7	14	10
0.68-0.69	0	0	7	14	10
0.69-0.70	0	0	7	14	10
0.70-0.71	0	0	3	15	10
0.71-0.72	0	0	3	17	10
0.72-0.73	0	0	3	17	10
0.73-0.74	0	0	3	17	10
0.74-0.75	0	0	3	17	10
0.75-0.76	0	0	2	13	15
0.76-0.77	0	0	2	14	17
0.77-0.78	0	0	2	14	17
0.78-0.79	0	0	2	14	17
0.79-0.80	0	0	1	14	17
0.80-0.81	0	0	0	10	20
0.81-0.82	0	0	0	10	21
0.82-0.83	0	0	0	10	21

APPENDIX D (Continued)

Quickbasic Program for Generating

Degrees of Membership

Results for Technical Uncertainty Data

0.83-0.84	0	0	0	10	21
0.84-0.85	0	0	0	10	21
0.85-0.86	0	0	0	6	25
0.86-0.87	0	0	0	6	26
0.87-0.88	0	0	0	6	26
0.88-0.89	0	0	0	6	26
0.89-0.90	0	0	0	6	26
0.90-0.91	0	0	0	2	28
0.91-0.92	0	0	0	2	29
0.92-0.93	0	0	0	2	29
0.93-0.94	0	0	0	2	29
0.94-0.95	0	0	0	1	29
0.95-0.96	0	0	0	0	30
0.96-0.97	0	0	0	0	30
0.97-0.98	0	0	0	0	30
0.98-0.99	0	0	0	0	30
0.99-1.00	0	0	0	0	30

Results for Schedule Uncertainty

Echo-check of data:

0.05 0.10 0.10 0.25 0.20 0.55 0.50 0.85 0.80 1.00
0.03 0.15 0.25 0.45 0.45 0.70 0.65 0.85 0.85 1.00
0.05 0.20 0.20 0.40 0.40 0.60 0.60 0.80 0.80 1.00
0.00 0.10 0.05 0.20 0.15 0.50 0.70 0.95 0.85 1.00
0.00 0.05 0.05 0.20 0.45 0.65 0.65 0.85 0.80 1.00
0.05 0.12 0.10 0.25 0.20 0.60 0.50 0.75 0.75 1.00
0.03 0.10 0.05 0.18 0.15 0.30 0.30 0.50 0.50 1.00
0.00 0.03 0.02 0.08 0.08 0.30 0.25 0.50 0.50 1.00
0.00 0.10 0.10 0.40 0.40 0.60 0.60 0.80 0.80 1.00
0.00 0.15 0.15 0.35 0.35 0.60 0.60 0.80 0.80 1.00
0.00 0.20 0.21 0.40 0.41 0.60 0.61 0.80 0.81 1.00
0.00 0.05 0.05 0.15 0.15 0.25 0.30 0.50 0.55 1.00
0.00 0.15 0.10 0.35 0.35 0.60 0.60 0.80 0.80 1.00
0.00 0.25 0.25 0.50 0.50 0.75 0.75 0.90 0.90 1.00
0.10 0.20 0.15 0.40 0.40 0.60 0.60 0.85 0.85 1.00
0.00 0.10 0.11 0.35 0.36 0.79 0.80 0.94 0.95 1.00
0.05 0.15 0.10 0.30 0.25 0.45 0.40 0.60 0.55 1.00
0.05 0.15 0.16 0.25 0.26 0.75 0.76 0.85 0.86 1.00
0.00 0.05 0.05 0.15 0.15 0.60 0.60 0.80 0.80 1.00
0.05 0.10 0.10 0.25 0.25 0.35 0.35 0.55 0.50 1.00
0.05 0.15 0.16 0.40 0.41 0.70 0.71 0.85 0.86 1.00

APPENDIX D (Continued)

Quickbasic Program for Generating

Degrees of Membership

Echo Check of Data

```
0.05 0.10 0.08 0.15 0.15 0.25 0.20 0.35 0.35 1.00
0.05 0.15 0.10 0.30 0.30 0.65 0.65 0.85 0.85 1.00
0.00 0.10 0.10 0.20 0.20 0.35 0.35 0.65 0.65 1.00
0.05 0.15 0.12 0.25 0.20 0.50 0.45 0.70 0.65 1.00
0.00 0.10 0.11 0.30 0.31 0.60 0.61 0.79 0.80 1.00
0.00 0.10 0.05 0.20 0.20 0.40 0.40 0.60 0.50 1.00
0.00 0.03 0.03 0.10 0.10 0.35 0.40 0.60 0.70 1.00
0.01 0.10 0.10 0.19 0.15 0.30 0.30 0.60 0.55 1.00
0.00 0.10 0.11 0.20 0.21 0.30 0.31 0.40 0.41 1.00
```

Results for Schedule Uncertainty Data

Schedule	Degree of Membership times 30				
Risk	VL	L	M	H	VH
---------	------	------	------	------	------
0.00-0.01	16	0	0	0	0
0.01-0.02	17	0	0	0	0
0.02-0.03	17	1	0	0	0
0.03-0.04	17	2	0	0	0
0.04-0.05	17	2	0	0	0
0.05-0.06	24	8	0	0	0
0.06-0.07	24	8	0	0	0
0.07-0.08	24	8	0	0	0
0.08-0.09	24	8	1	0	0
0.09-0.10	24	8	1	0	0
0.10-0.11	13	16	2	0	0
0.11-0.12	13	19	2	0	0
0.12-0.13	12	20	2	0	0
0.13-0.14	12	20	2	0	0
0.14-0.15	12	20	2	0	0
0.15-0.16	4	19	8	0	0
0.16-0.17	4	21	8	0	0
0.17-0.18	4	21	8	0	0
0.18-0.19	4	20	8	0	0
0.19-0.20	4	19	8	0	0
0.20-0.21	1	15	13	1	0
0.21-0.22	1	16	14	1	0
0.22-0.23	1	16	14	1	0
0.23-0.24	1	16	14	1	0
0.24-0.25	1	16	14	1	0
0.25-0.26	0	13	14	2	0
0.26-0.27	0	13	15	2	0
0.27-0.28	0	13	15	2	0
0.28-0.29	0	13	15	2	0

Results for Schedule Uncertainty Data

0.29-0.30	0	13	15	2	0
0.30-0.31	0	10	12	5	0
0.31-0.32	0	10	13	6	0
0.32-0.33	0	10	13	6	0
0.33-0.34	0	10	13	6	0
0.34-0.35	0	10	13	6	0
0.35-0.36	0	7	12	7	1
0.36-0.37	0	7	13	7	1
0.37-0.38	0	7	13	7	1
0.38-0.39	0	7	13	7	1
0.39-0.40	0	7	13	7	1
0.40-0.41	0	2	15	9	1
0.41-0.42	0	2	17	9	2
0.42-0.43	0	2	17	9	2
0.43-0.44	0	2	17	9	2
0.44-0.45	0	2	17	9	2
0.45-0.46	0	1	18	10	2
0.46-0.47	0	1	18	10	2
0.47-0.48	0	1	18	10	2
0.48-0.49	0	1	18	10	2
0.49-0.50	0	1	18	10	2
0.50-0.51	0	0	17	9	6
0.51-0.52	0	0	17	9	6
0.52-0.53	0	0	17	9	6
0.53-0.54	0	0	17	9	6
0.54-0.55	0	0	17	9	6
0.55-0.56	0	0	16	8	9
0.56-0.57	0	0	16	8	9
0.57-0.58	0	0	16	8	9
0.58-0.59	0	0	16	8	9
0.59-0.60	0	0	16	8	9
0.60-0.61	0	0	7	10	9
0.61-0.62	0	0	7	12	9
0.62-0.63	0	0	7	12	9
0.63-0.64	0	0	7	12	9
0.64-0.65	0	0	7	12	9
0.65-0.66	0	0	5	14	11
0.66-0.67	0	0	5	14	11
0.67-0.68	0	0	5	14	11
0.68-0.69	0	0	5	14	11
0.69-0.70	0	0	5	14	11
0.70-0.71	0	0	3	14	12
0.71-0.72	0	0	3	15	12
0.72-0.73	0	0	3	15	12
0.73-0.74	0	0	3	15	12
0.74-0.75	0	0	3	15	12
0.75-0.76	0	0	1	15	13
0.76-0.77	0	0	1	16	13

APPENDIX D (Continued)

Results for Schedule Data

0.77-0.78	0	0	1	16	13
0.78-0.79	0	0	1	16	13
0.79-0.80	0	0	0	15	13
0.80-0.81	0	0	0	10	21
0.81-0.82	0	0	0	10	22
0.82-0.83	0	0	0	10	22
0.83-0.84	0	0	0	10	22
0.84-0.85	0	0	0	10	22
0.85-0.86	0	0	0	3	26
0.86-0.87	0	0	0	3	28
0.87-0.88	0	0	0	3	28
0.88-0.89	0	0	0	3	28
0.89-0.90	0	0	0	3	28
0.90-0.91	0	0	0	2	29
0.91-0.92	0	0	0	2	29
0.92-0.93	0	0	0	2	29
0.93-0.94	0	0	0	2	29
0.94-0.95	0	0	0	1	29
0.95-0.96	0	0	0	0	30
0.96-0.97	0	0	0	0	30
0.97-0.98	0	0	0	0	30
0.98-0.99	0	0	0	0	30
0.99-1.00	0	0	0	0	30

APPENDIX E

APPENDIX E

Pattern Search Algorithm for Triangular
Least Squares Error Fit

```
DECLARE SUB ShowVals (Ecurr#)
DECLARE FUNCTION SumErr2# (J%)
' QuickBASIC program to perform four-part curve fit (to sides of a triangle
'   and lines left and right of it) characterizing each interval along the
'   Risk axis by its midpoint
'
' Uses a pattern search algorithm to find the best fit (least squared error)


DEFINT I-N
DEFDBL A-H, O-Z

CONST NUMINTS = 100    ' number of intervals
CONST NUMSETS = 5     ' number of fuzzy sets (very lo, lo, med, hi, very hi)

DIM Rmid(NUMINTS)     ' risk values at midpoints of intervals
DIM YMembership(NUMINTS, NUMSETS)    ' degree of membership of each midpt in
each
                        '  fuzzy set
DIM YSum(NUMSETS)     ' sum of YMembership values for each fuzzy set
DIM Y2Sum(NUMSETS)    ' sum of squares of YMembership vals for each fuzzy set
DIM TotVar(NUMSETS)   ' sum of squares of (Y-mean(Y)) terms (as a measure
                  '   of variability of the data

DIM P(4)    ' four parameters for triangle:  X0, X1, X2, Y1
        ' (left side from (X0,0) to (X1,Y1); right from (X1,Y1) to (X2,0)
DIM DP(4)     ' increments for pattern search
DIM Pold(4)  ' old values for pattern search

OPEN "TECH.MID" FOR INPUT AS 1   ' contains only the Technical Risk data

CLS

' Initialize summing variables

FOR J = 1 TO NUMSETS
   YSum(I) = 0
   Y2Sum(I) = 0
NEXT J

PRINT "Echo check of input data:"

FOR I = 1 TO NUMINTS

  INPUT #1, Rmid(I)
  PRINT USING "##.###    "; Rmid(I);

  FOR J = 1 TO NUMSETS
```

```
        INPUT #1, Ytemp
        YMembership(I, J) = Ytemp
        PRINT USING "##.###"; YMembership(I, J);
        YSum(J) = YSum(J) + Ytemp
        Y2Sum(J) = Y2Sum(J) + Ytemp ^ 2

    NEXT J

    PRINT

NEXT I


FOR J = 1 TO NUMSETS

    PRINT "For triangle number"; J

    INPUT "Enter 1st guess for X0: ", P(1)
    INPUT "Enter 1st guess for X1: ", P(2)
    INPUT "Enter 1st guess for X2: ", P(3)
    INPUT "Enter 1st guess for Y1: ", P(4)

    Ecurr = SumErr2(J)
    Eold = Ecurr

    FOR Kparam = 1 TO 4
      Pold(Kparam) = P(Kparam) ' "previous value" (for extrapolating to next)
      DP(Kparam) = .1#        ' initial step size
    NEXT Kparam

    Retrench = -1    ' true

    DO

      FOR Kparam = 1 TO 4

        Psave = P(Kparam)                  ' save if need to reset
        P(Kparam) = P(Kparam) + DP(Kparam)  ' positive perturbation
        Etest = SumErr2(J)

        IF Etest < Ecurr THEN   ' new location is better
          Ecurr = Etest
          CALL ShowVals(Ecurr)
        ELSE
          P(Kparam) = Psave - DP(Kparam) ' negative perturbation
          Etest = SumErr2(J)
          IF Etest < Ecurr THEN   ' this new location is better
            Ecurr = Etest
```

```
    CALL ShowVals(Ecurr)
   ELSE              ' neither one is better
    P(Kparam) = Psave     ' reset to unperturbed val
   END IF
  END IF

NEXT Kparam


IF Ecurr < Eold THEN   ' establish new base

  PRINT "New Base:"

  Retrench = 0  ' false

  FOR Kparam = 1 TO 4
    Ptemp = 2 * P(Kparam) - Pold(Kparam)' extrapolate
    Pold(Kparam) = P(Kparam)
    P(Kparam) = Ptemp
  NEXT Kparam

  Eold = Ecurr

  Ecurr = SumErr2(J)
  CALL ShowVals(Ecurr)

ELSE

  PRINT "Pattern broken"

  FOR Kparam = 1 TO 4
    P(Kparam) = Pold(Kparam)   ' old was better; reset
  NEXT Kparam

  Ecurr = Eold    ' reset current best value

  CALL ShowVals(Ecurr)

  IF Retrench THEN

    PRINT "Retrenching"

    IF DP(1) < .00000001# THEN    ' small enough; get out
      EXIT DO
    END IF

    FOR Kparam = 1 TO 4
      DP(Kparam) = DP(Kparam) / 10#   ' smaller steps
```

```
        NEXT Kparam

      END IF

      Retrench = -1   ' true

    END IF

  LOOP

  TotVar(J) = Y2Sum(J) - (YSum(J) ^ 2) / NUMINTS
  PRINT : PRINT "For triangle number"; J; ":"
  PRINT "S = "; TotVar(J)
  PRINT "Rho = "; (TotVar(J) - Ecurr) / TotVar(J)
  PRINT

NEXT J

SUB ShowVals (Ecurr)

SHARED P()

FOR Kparam = 1 TO 4
  PRINT USING "##.########"; P(Kparam);
NEXT Kparam

PRINT USING "#####.##########"; Ecurr

END SUB

FUNCTION SumErr2 (J)

SHARED Rmid(), YMembership(), P()

IF P(1) > P(2) OR P(2) > P(3) THEN   'invalid parameter set

  TempSum = 1D+30    ' big number

ELSE

  TempSum = 0

  FOR I = 1 TO NUMINTS

    X = Rmid(I)

    IF X < P(1) THEN  ' beyond left end of triangle
      Ypred = 0
```

```
    ELSEIF X < P(2) THEN   ' left side of triangle
      Ypred = (P(4) / (P(2) - P(1))) * X - P(1) * P(4) / (P(2) - P(1))
    ELSEIF X < P(3) THEN   ' right side of triangle
      Ypred = -(P(4) / (P(3) - P(2))) * X + P(3) * P(4) / (P(3) - P(2))
    ELSE    ' beyond right end of triangle
      Ypred = 0
    END IF

    TempErr = YMembership(I, J) - Ypred
    TempSum = TempSum + TempErr * TempErr

  NEXT I

END IF

SumErr2 = TempSum

END FUNCTION
```

APPENDIX F

APPENDIX F

Pattern Search Algorithm For

Trapezoidal Least Square Error Fit

The modified program for fitting the Technical Uncertainty data to a trapezoid (TRAPFIT1.BAS) is as follows:

```
DECLARE SUB ShowVals (Ecurr#)
DECLARE FUNCTION SumErr2# (J%)
' QuickBASIC program to perform five-part curve fit (to sides and top of
'   a trapezoid and lines left and right of it) characterizing each interval
'   along the Risk axis by its midpoint
'
' Uses a pattern search algorithm to find the best fit (least squared error)


DEFINT I-N
DEFDBL A-H, O-Z

CONST NUMINTS = 100    ' number of intervals
CONST NUMSETS = 5      ' number of fuzzy sets (very lo, lo, med, hi, very hi)

DIM Rmid(NUMINTS)     ' risk values at midpoints of intervals
DIM YMembership(NUMINTS, NUMSETS)   ' degree of membership of each midpt in
                                    ' each fuzzy set
DIM YSum(NUMSETS)     ' sum of YMembership values for each fuzzy set
DIM Y2Sum(NUMSETS)    ' sum of squares of YMembership vals for each fuzzy set
DIM TotVar(NUMSETS)   ' sum of squares of (Y-mean(Y)) terms (as a measure
                '   of variability of the data

DIM P(5)    ' five parameters for trapezoid:  X0, X1, X2, X3, Y1
         ' (left side from (X0,0) to (X1,Y1); top from (X1,Y1) to (X2, Y1);
         '  right side from (X2,Y1) to (X3,0)
DIM DP(5)     ' increments for pattern search
DIM Pold(5)  ' old values for pattern search

OPEN "TECH.MID" FOR INPUT AS 1   ' contains only the Technical Risk data

CLS

' Initialize summing variables

FOR J = 1 TO NUMSETS
  YSum(I) = 0
  Y2Sum(I) = 0
NEXT J

PRINT "Echo check of input data:"

FOR I = 1 TO NUMINTS
```

```
    INPUT #1, Rmid(I)
    PRINT USING "##.###    "; Rmid(I);

    FOR J = 1 TO NUMSETS

      INPUT #1, Ytemp
      YMembership(I, J) = Ytemp
      PRINT USING "##.###"; YMembership(I, J);
      YSum(J) = YSum(J) + Ytemp
      Y2Sum(J) = Y2Sum(J) + Ytemp ^ 2

    NEXT J

    PRINT

  NEXT I


  FOR J = 1 TO NUMSETS

    PRINT "For trapezoid number"; J

    INPUT "Enter 1st guess for X0: ", P(1)
    INPUT "Enter 1st guess for X1: ", P(2)
    INPUT "Enter 1st guess for X2: ", P(3)
    INPUT "Enter 1st guess for X3: ", P(4)
    INPUT "Enter 1st guess for Y1: ", P(5)

    Ecurr = SumErr2(J)
    Eold = Ecurr

    FOR Kparam = 1 TO 5
      Pold(Kparam) = P(Kparam) ' "previous value" (for extrapolating to next)
      DP(Kparam) = .1#        ' initial step size
    NEXT Kparam

    Retrench = -1    ' true

    DO

      FOR Kparam = 1 TO 5

        Psave = P(Kparam)                ' save if need to reset
        P(Kparam) = P(Kparam) + DP(Kparam)  ' positive perturbation
        Etest = SumErr2(J)

        IF Etest < Ecurr THEN   ' new location is better
          Ecurr = Etest
```

```
    CALL ShowVals(Ecurr)
  ELSE
    P(Kparam) = Psave - DP(Kparam) ' negative perturbation
    Etest = SumErr2(J)
    IF Etest < Ecurr THEN   ' this new location is better
      Ecurr = Etest
      CALL ShowVals(Ecurr)
    ELSE                    ' neither one is better
      P(Kparam) = Psave     ' reset to unperturbed val
    END IF
  END IF

NEXT Kparam


IF Ecurr < Eold THEN   ' establish new base

  PRINT "New Base:"

  Retrench = 0  ' false

  FOR Kparam = 1 TO 5
    Ptemp = 2 * P(Kparam) - Pold(Kparam)' extrapolate
    Pold(Kparam) = P(Kparam)
    P(Kparam) = Ptemp
  NEXT Kparam

  Eold = Ecurr

  Ecurr = SumErr2(J)
  CALL ShowVals(Ecurr)

ELSE

  PRINT "Pattern broken"

  FOR Kparam = 1 TO 5
    P(Kparam) = Pold(Kparam)   ' old was better; reset
  NEXT Kparam

  Ecurr = Eold   ' reset current best value

  CALL ShowVals(Ecurr)

  IF Retrench THEN

    PRINT "Retrenching"
```

```
with Text_IO; use Text_IO;
with Category; use Category;
with Ada.Numerics.Float_Random; use Ada.Numerics.Float_Random;
procedure fuzzify is
  Random_Number_Generator: Generator;

  subtype How_Many_Categories_Type is integer range 2..integer'last;
  subtype Range_Boundary_Type is float;

  package How_Many_Categories_IO is new Integer_IO(How_Many_Categories_Type);
  package Range_IO is new Float_IO(Range_Boundary_Type);
  use How_Many_Categories_IO, Range_IO;

  Number_Of_Categories, Number_Of_Data_Values: How_Many_Categories_Type;

  Test_Data: File_Type;
begin
  Reset(Random_Number_Generator);
  Open(Test_Data, In_File, "test.dat");
  while not End_Of_File(Test_Data) loop
    Put("=I=I=I=I=I"); -- Indentation
    Put_Line("=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=");
    Put("          "); -- Indentation
    Put("Number of categories: ");
    Get(Test_Data,Number_Of_Categories);
    Put("          "); -- Indentation
    Put(Number_Of_Categories, WIDTH => 5);
    New_Line;

    PROCESS_DATA:
    declare
      type Category_Task_Pointer is access Category_Task_Type;
      type All_Categories is array(1..Number_Of_Categories)
        of Category_Task_Pointer;
      Category: All_Categories;
      Line_Length: Natural;
      Total: float := 0.0;
      Relative_Probability: array(1..Number_Of_Categories) of float;
      Random_Number_In_Total: float;
      Running_Total: float;
      Range_Boundary_1, Range_Boundary_2: Range_Boundary_Type;
      Data_Value: float;
      Share: float := 0.0;
      Current_Category : integer;
    begin
```

```
for I in Category'range loop
  Category(I) := new Category_Task_Type;
end loop;
Get_Line(Test_Data,Category_Name, Line_Length);
Put("        "); -- Indentation
Put_Line("  Lower   Upper     Category");
for I in 1..Number_Of_Categories loop
  Get_Line(Test_Data,Category_Name, Line_Length);
  Category(I).Set_Name(Category_Name, Line_Length);
  Get(Test_Data, Range_Boundary_1);
  Skip_Line(Test_Data);
  Get(Test_Data, Range_Boundary_2);
  Skip_Line(Test_Data);
  Category(I).Set_Value(Range_Boundary_1,Range_Boundary_2);
  Put("        "); -- Indentation
  Put(Range_Boundary_1,FORE => 6, AFT => 2, EXP => 0);
  Put(Range_Boundary_2,FORE => 6, AFT => 2, EXP => 0);
  Put("  ");
  Put(Category_Name(1..Line_Length));
  New_Line;
end loop;
New_Line;
Get(Test_Data, Number_Of_Data_Values);
Skip_Line(Test_Data);
Put("        "); -- Indentation
Put("Number of data values to process: ");
Put(Number_Of_Data_Values, WIDTH => 5);
New_Line(2);

for I in 1..Number_Of_Data_Values loop
  Get(Test_Data, Data_Value);
  Put("Processing data value of: ");
  Put(Data_Value, FORE => 6, AFT => 2, EXP => 0);
  New_Line;
  Total := 0.0;
  for J in Category'range loop
    Category(J).Get_Share(Data_Value,Share);
    Relative_Probability(J) := 100.0 * Share;
    Total := Total + Relative_Probability(J);
  end loop;

  if Total = 0.0 then
    Put_Line("Error: Data out of range...");
  end if;
```

```
      Random_Number_In_Total := Random(Random_Number_Generator) * Total;

      Running_Total := Relative_Probability(Category'first);
      Current_Category := Category'first;
      while (Running_Total    < Random_Number_In_Total) and
          (Current_Category < Category'last        ) loop
        Current_Category := Current_Category + 1;
        Running_Total := Running_Total +
                    Relative_Probability(Current_Category);
      end loop;
      Put("Category selected is: ");
      Category(Current_Category).Get_Name(Category_Name,Line_Length);
      Put(Category_Name(1..Line_Length));
      New_Line;
      Put_Line("Membership        Category");
      Put_Line("======================================");
      for I in Category'range loop
        Category(I).Get_Name(Category_Name, Line_Length);
        Put(Relative_Probability(I)/total, EXP => 0);
        Put("    ");
        Put(Category_Name(1..Line_Length));
        New_Line;
      end loop;
      New_Line;
    end loop; -- Number_Of_Data_Values
  end PROCESS_DATA;
 end loop;
 Close(Test_Data);
end fuzzify;
```

APPENDIX G
FUZZIFY Simulation Program

```
package category is
    subtype Category_Name_Type is string(1..30);
    Category_Name, Name: Category_Name_Type;
    subtype Share_Type is float range 0.0 .. 1.0;

    task type Category_Task_Type is
        entry Set_Name(Name: in  Category_Name_Type;  String_Length: in
Natural);
        entry Get_Name(Name: out Category_Name_Type; String_Length: out
Natural);
        entry Set_Value(Number_1, Number_2: in float);
        entry Get_Low_Value(Low_Number: out float);
        entry Get_Mid_Value(Mid_Number: out float);
        entry Get_High_Value(High_Number: out float);
        entry Get_Share(Input: in  float; Share: out Share_Type);
    end Category_Task_Type;
end category;


package body Category is
    task body Category_Task_Type is
        My_Name: Category_Name_Type;
        My_Length: Natural;
        Low_Value, Mid_Value, High_Value, Mid_2_Low, High_2_Mid: float;
    begin
        loop
            select
                accept Set_Name (Name: in Category_Name_Type;
                                String_Length: in Natural) do
                    My_Name := Name;
                    My_Length := String_Length;
                end Set_Name;
            or
                accept Get_Name(Name: out Category_Name_Type;
                                String_Length: out Natural) do
                    Name := My_Name;
                    String_Length := My_Length;
                end Get_Name;
            or
                accept Set_Value(Number_1, Number_2: in float) do
                    if Number_1 > Number_2 then
                        Low_Value := Number_2;
```

```
                    Mid_Value := Number_2 + ((Number_1 - Number_2) /
2.0);
                    High_Value := Number_1;
                else
                    Low_Value := Number_1;
                    Mid_Value := Number_1 + ((Number_2 - Number_1) /
2.0);
                    High_Value := Number_2;
                end if;
            end Set_Value;
        or
            accept Get_Low_Value(Low_Number: out float) do
                Low_Number := Low_Value;
            end Get_Low_Value;
        or
            accept Get_Mid_Value(Mid_Number: out float) do
                Mid_Number := Mid_Value;
            end Get_Mid_Value;
        or
            accept Get_High_Value(High_Number: out float) do
                High_Number := High_Value;
            end Get_High_Value;
        or
            accept Get_Share(Input: in  float; Share: out Share_Type) do
                Mid_2_Low  := (Mid_Value  - Low_Value);
                High_2_Mid := (High_Value - Mid_Value);
                if (Input <= Low_Value) or (Input >= High_Value) then
                    Share := 0.0;
                else
                    if Input  > Mid_Value then
                        Share := (High_Value - Input)/High_2_Mid;
                    else
                        if Input = Mid_Value then
                            Share := 1.0;
                        else
                            Share := (Input - Low_Value)/Mid_2_Low;
                        end if;
                    end if;
                end if;
            end Get_Share;
        or
            terminate;
        end select;
    end loop;
end Category_Task_Type;
```

```
end category;


with Text_IO; use Text_IO;
with Category; use Category;
with Ada.Numerics.Float_Random; use Ada.Numerics.Float_Random;
procedure fuzzify is
    Random_Number_Generator: Generator;
    subtype How_Many_Categories_Type is integer range 2..integer'last;
    subtype Range_Boundary_Type is float;
    package How_Many_Categories_IO is new
Integer_IO(How_Many_Categories_Type);
    package Range_IO is new Float_IO(Range_Boundary_Type);
    use How_Many_Categories_IO, Range_IO;
    Number_Of_Categories, Number_Of_Data_Values:
How_Many_Categories_Type;
    Test_Data: File_Type;
begin
    Reset(Random_Number_Generator);
    Open(Test_Data, In_File, "test.dat");
    while not End_Of_File(Test_Data) loop
        Put("=I=I=I=I=I"); -- Indentation

    Put_Line("=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I=I="
);
        Put("               "); -- Indentation
        Put("Number of categories: ");
        Get(Test_Data,Number_Of_Categories);
        Put("               "); -- Indentation
        Put(Number_Of_Categories, WIDTH => 5);
        New_Line;
        PROCESS_DATA:
        declare
            type Category_Task_Pointer is access Category_Task_Type;
            type All_Categories is array(1..Number_Of_Categories) of
Category_Task_Pointer;
            Category: All_Categories;
            Line_Length: Natural;
            Total: float := 0.0;
            Relative_Probability: array(1..Number_Of_Categories) of float;
            Random_Number_In_Total: float;
            Running_Total: float;
            Range_Boundary_1, Range_Boundary_2: Range_Boundary_Type;
            Data_Value: float;
            Share: float := 0.0;
```

```
    Current_Category : integer;
begin
    for I in Category'range loop
        Category(I) := new Category_Task_Type;
    end loop;
    Get_Line(Test_Data,Category_Name, Line_Length);
    Put("                "); -- Indentation
    Put_Line(" Lower  Upper      Category");
    for I in 1..Number_Of_Categories loop
        Get_Line(Test_Data,Category_Name, Line_Length);
        Category(I).Set_Name(Category_Name, Line_Length);
        Get(Test_Data, Range_Boundary_1);
        Skip_Line(Test_Data);
        Get(Test_Data, Range_Boundary_2);
        Skip_Line(Test_Data);
        Category(I).Set_Value(Range_Boundary_1,Range_Boundary_2);
        Put("                "); -- Indentation
        Put(Range_Boundary_1,FORE => 6, AFT => 2, EXP => 0);
        Put(Range_Boundary_2,FORE => 6, AFT => 2, EXP => 0);
        Put("  ");
        Put(Category_Name(1..Line_Length));
        New_Line;
    end loop;
    New_Line;
    Get(Test_Data, Number_Of_Data_Values);
    Skip_Line(Test_Data);
    Put("                "); -- Indentation
    Put("Number of data values to process: ");
    Put(Number_Of_Data_Values, WIDTH => 5);
    New_Line(2);
    for I in 1..Number_Of_Data_Values loop
        Get(Test_Data, Data_Value);
        Put("Processing data value of: ");
        Put(Data_Value, FORE => 6, AFT => 2, EXP => 0);
        New_Line;
        Total := 0.0;
        for J in Category'range loop
            Category(J).Get_Share(Data_Value,Share);
            Relative_Probability(J) := 100.0 * Share;
            Total := Total + Relative_Probability(J);
        end loop;
        if Total = 0.0 then
            Put_Line("Error: Data out of range...");
        end if;
```

```
            Random_Number_In_Total :=
Random(Random_Number_Generator) * Total;
            Running_Total := Relative_Probability(Category'first);
            Current_Category := Category'first;
            while              (Running_Total           <
Random_Number_In_Total           ) and
               (Current_Category                      < Category'last )
loop
               Current_Category := Current_Category + 1;
               Running_Total := Running_Total +
                              Relative_Probability(Current_Category);
            end loop;
            Put("Category selected is: ");

   Category(Current_Category).Get_Name(Category_Name,Line_Length);
            Put(Category_Name(1..Line_Length));
            New_Line;
            Put_Line("Membership                       Category");

   Put_Line("========================================");
            for I in Category'range loop
               Category(I).Get_Name(Category_Name, Line_Length);
               Put(Relative_Probability(I)/total, EXP => 0);
               Put("     ");
               Put(Category_Name(1..Line_Length));
               New_Line;
            end loop;
            New_Line;
         end loop; -- Number_Of_Data_Values
      end PROCESS_DATA;
   end loop;
   Close(Test_Data);
end fuzzify;
```