Stochastic Project Scheduling Simulation System (SPSS III)

Dong-Eun Lee*

Abstract

This paper introduces a Stochastic Project Scheduling Simulation system (SPSS III) developed by the author to predict a project completion probability in a certain time. The system integrates deterministic CPM, probabilistic PERT, and stochastic Discrete Event Simulation (DES) scheduling methods into one system. It implements automated statistical analysis methods for computing the minimum number of simulation runs, the significance of the difference between independent simulations, and the confidence interval for the mean project duration as well as sensitivity analysis method in What-if analyzer component. The SPSS III gives the several benefits to researchers in that it (1) complements PERT and Monte Carlo simulation by using stochastic activity durations via a web based JAVA simulation over the Internet, (2) provides a way to model a project network having different probability distribution functions, (3) implements statistical analyses method which enable to produce a reliable prediction of the probability of completing a project in a specified time, and (4) allows researchers to compare the outcome of CPM, PERT and DES under different variability or skewness in the activity duration data.

Keywords : Critical path method, probability, scheduling, simulation, sensitivity analysis.

1. Introduction

The Critical Path Method (CPM) has been used extensively due to its simplicity. It computes deterministic project duration by analyzing which sequence of activities has least float. Early start and finish times are computed by using a forward-pass-algorithm, and late start and finish times using a backward-pass-algorithm (Moder et al. 1983, Patrick 2004). The Program Evaluation and Review Technique (PERT) has been considered as a complement to CPM by using probabilistic activity durations (Patrick 2004), by estimating the uncertainty in a project network, and by computing the probability to complete the project within a specified time. PERT, however, is limited in that it treats the expected activity durations as deterministic durations, not as random variables.

* Ph.D., Construction Engineering and Management Program, Department of Civil and Architectural Engineering, Illinois Institute of Technology, 3201S. Dearborn St., Chicago, IL 60616-3793, USA. Email: leedong1@iit.edu

To complement the limitations of PERT scheduling methods, simulation has been applied to schedule networks by considering random activity durations for estimating project durations (Douglas 1978, Ahuja and Nandakumar 1985) with the recognition that simulation is the most promising approach to eliminate these problems (Douglas 1978, Martinez and Ioannou 1997), to predict total behavior of a network that displays probabilistic and stochastic features, and to enhance the scheduling of project activities (Crandall 1977, Douglas 1978).

In recent years, Discrete Event Simulation (DES) has been extended to the modeling and analysis of project scheduling simulation with the recognition that simulation deals more efficiently project-scheduling problems with the stochastic and random nature of activity durations (Simphony 2003, Lee 2005).

After SPSS I (Lee 2005; ASCE), which addresses the stochastic nature of construction activities with a flexible modeling approach, had been proposed as a distributed JAVA application for project scheduling simulation over the Internet, SPSS I was extended (1) to integrate the CPM, PERT, and DES methods into a single system, (2) to determine the minimum number of simulation runs, (3) to tests the statistical difference between independent simulations, and (4) to computes the confidence interval for the mean project duration resulting in SPSS II (Lee 2004; WSC).

The SPSS II dealt with two important issues in simulation application in project scheduling. It assures the validity of a simulation by computing the minimum number of simulation runs. If a simulation is terminated prematurely, the output data analysis may not objectively represent the project network simulation under study and therefore be invalid. Properly determining the number of runs needed for a given project helps determine the minimal computing time and ensure valid results. Moreover, it verifies if output data obtained from a simulation can be considered as a same population with the data obtained from other simulations. When multiple simulation experiments are conducted for the same schedule network, different results may be obtained due to different streams of random numbers. Therefore, it is essential to examine if a simulation result is statistically identical or different.

This article includes add-on component (i.e., What-if analyzer) that make use of two important descriptive statistical measures, skewness and variability of activity durations, for sensitivity analysis in scheduling simulation. Automated JAVA components are implemented in the SPSS III system for carrying out the add-on functions. The What-If Analyzer allows user to change skewness and variability of activity durations. When the descriptive measures change, how the project duration behaves can be simulated. It provides deeper insight into the project schedule's behavior.

The differences among SPSS I, II, and III are briefly presented in Table 1 along with comparisons of CPM, PERT, and other operation oriented simulation applications (e.g., Micro-Cyclone, ABC, etc).

The purpose of SPSS III is to complement the problems which exist in CPM, PERT, and DES applications. First of all, CPM uses deterministic fixed terms for activity durations. PERT computes probabilistic fixed terms and deals them as deterministic manner in computing the probability to compete a project. PERT considers the uncertainty in estimating activity durations. However, it has been known that PERT systematically underestimates the total project duration. That is why SPSS III is required to increase the accuracy of scheduling project activities by treating the activity duration as random variables. Existing simulation applications (e.g., CYCLONE: Halpin 1977, UM-CYCLONE: Ioannou; 1989, DISCO: Huang et al.; 1994, ABC: Shi; 1999, STROBOSCOPE: Martinez and Ioannou 1999, Symphony: Hajjar and AbouRizk 1999) are mostly oriented into operational research and analysis, not into project scheduling. Therefore, this research aims to apply simulation modeling and analysis techniques directly to project scheduling application.

Ensuring that the contractor has sufficient capacity to meet contractual requirements is valuable, particularly at the time of bidding phase. SPSS III may be quite effective in supporting construction planning in the bidding phase. It provides a better prediction of the completion probability when the collection of historic data and data analysis are not available as in the bidding phase. Since selecting input probability distributions of activity durations requires extensive amount of historical data, SPSS III is a good fit for accommodating expert opinions in the absence of data. Even if it is burdensome to input the three time estimates for each activity, the system can reflect the uncertainty involved in activity durations and the stochastic nature of the completion time.

Currently, all the functions for predicting the final completion duration are completely implemented. However, cost aspects have

Table 1. The differences among CPM, PERT, DES, SPSS I, SPSS II, and SPSS III

Category	CPM	PERT	Exist. Sim. App. (Micro-cyclone, ABC, etc)	SPSS I	SPSS II	SPSS III
Feature	D[1]	P[2]	S[3]	D, S	D, P, S	D, P, S
Domain	Scheduling	Scheduling	Operation Research	Scheduling	Scheduling	Scheduling
Method	CPM[4]	PERT[5]	DES[6]	CPM, DES	CPM, PERT, DES	CPM, PERT, DES
Unique functions	BPA[7], FPA[8], LPA[9]	BPA[7], FPA[8], LPA[9], PDG[10]	SMA[11]	WBS[12]	MinNo[13], TA[14]	SA[15]
Activity duration	DCV[16]	PCV[17]	SRV[18]	SRV[18]	SRV[18]	SRV[18]
Terminating rule	NA	NA	Unknown	Not exist	Exist	Exist
Validity testing	NA	NA	Unknown	Not exist	Exist	Exist
Data attributes for modeling	Non-redundant	Non-redundant	Including redundant	Non-redundant	Non-redundant	Non-redundant

D[1] : Deterministic, P[2]: Probabilistic, S[3]: Stochastic, CPM[4]: Critical Path Method, PERT[5]: Program Evaluation Review Technique, DES[6]: Discrete Event Simulation, BPA[7]: Backward Path Algorithm, FPA[8]: Forward Path Algorithm, LPA[9]: Longest Path Computation Algorithm, PDG[10]: PERT Duration Generation, SMA[11]: Simulation Modeling and Analysis, WBS[12]: Web Based Simulation modeling MinNo[13]: Computing the minimum Number of simulation runs, TA[14]: t-test analysis, SA[15]: Sensitivity Analysis using skewness and variability, DCV[16]: Deterministic Constant Values, PCV[17]: Probabilistic Constant Values, SRV[18]: Stochastic Random Variables not been considered. The only disadvantage is that the network modeling GUI is limited in accommodating the node and arrow elements in a screen. Using spreadsheet modeling component such as in Primavera Project Planner (P3) may be the alternative for the disadvantage.

The benefits to researchers and practitioners are (1) to provide a methodological approach to establish sensitivity analyses method in project scheduling simulation so that one can discover more heuristics in a project schedule under study.

2. SPSS III

The SPSS III system developed by the author is coded using JAVA programming language. It applies simulation modeling and analysis methods on project scheduling problem. After the first version of the SPSS, called SPSS I (Lee 2004; ASCE), is introduced, the system has been advanced in version as automates advanced functions for statistical analyses and integrates them into earlier version. The second version is called SPSS II (Lee 2004; WSC). The third version, SPSS III, is presented in this paper. SPSS III implements sensitivity analysis method to predict the total project duration in accordance to the changes in the input probability distributions of activity durations. The versioning represents the distinct advancement in add-ons. The functions in each earlier version have been integrated into later versions. To make this paper standalone by acquiring the independence and completeness of the paper, it includes some explanations which already available in those earlier versions.

Existing simulation applications have different modeling elements, representation, and strengths. Which means they have redundant data attributes as shown in Table 1. However, they share similar data attributes such as activity name, duration, logical relationship (AND or OR), resources, production measure, activity characteristic, and processing entity, etc. While those systems may be used for project scheduling, the interfaces of which those systems have do neither regulate efficient human interaction (between the systems and the scheduler) nor provide efficient project schedule modeling.

That is why SPSS is developed to reflect the basic and minimal characteristics of project scheduling by achieving a significant simplification of the simulation modeling. The required simplification is achieved by retaining the simplicity of the CPM representation and eliminating the redundant data attributes that are not required in regular scheduling operations.

SPSS I (Lee 2005; ASCE) analyzes a project schedule network by using deterministic CPM and stochastic DES methods over the Internet.

SPSS II (Lee 2004; WSC) fully automate and integrate CPM, PERT, and DES-based scheduling methods in a single system. In many simulation studies, little effort has been made to analyze the simulation output data appropriately. A very common mode of operation is to make a single simulation run of some arbitrary length and then accept the resulting simulation estimates as the true model characteristics. However, these estimates are just particular realizations of random variables that may have large variance (Law and Kelton 2000). To complement these problems, SPSS II implements terminating rule to determine the minimum number of simulation runs and validity testing method to verify the statistical difference between independent simulation experiments.

As shown in Figure 1, SPSS III retains the rendering feature of SPSS II that shows all the information involved in each node including early and late activity start and finish times, and total float in a node object. It also implements sensitivity analysis components which allow to simulate alternative system configurations and to compare their results automatically.

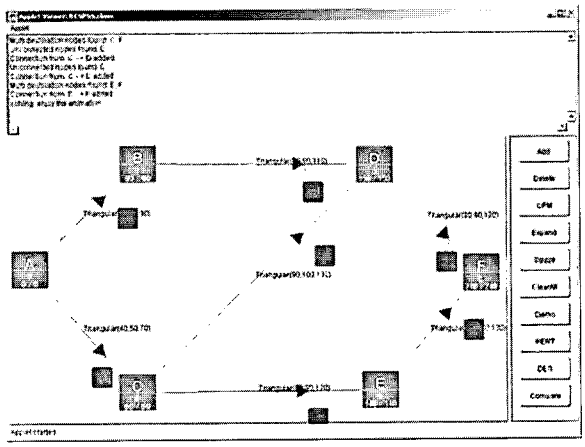

Figure 1. Project Network Modeling in SPSS III

In PERT mode, SPSS III generates expected activity durations by making use of the optimistic, most likely, and pessimistic time estimates. It then computes project duration, variance, standard deviation, and the probability to complete the project within a specified number of days (Figure 2). While PERT provides a probabilistic approach for representing activity durations, the PERT statistics shows that it is conservative than CPM. The range of the mean project duration at 95% confidence is first calculated for a simulation and presented to the user (Figure 3). The output data are then used to calculate the minimum number of simulation runs, which is also presented to the user in the same screen (Figure 3). The overall mean project duration and its range at 95% confidence are finally calculated by conducting several simulations each with the minimum number of runs. The next following sections highlight the features of SPSS III.

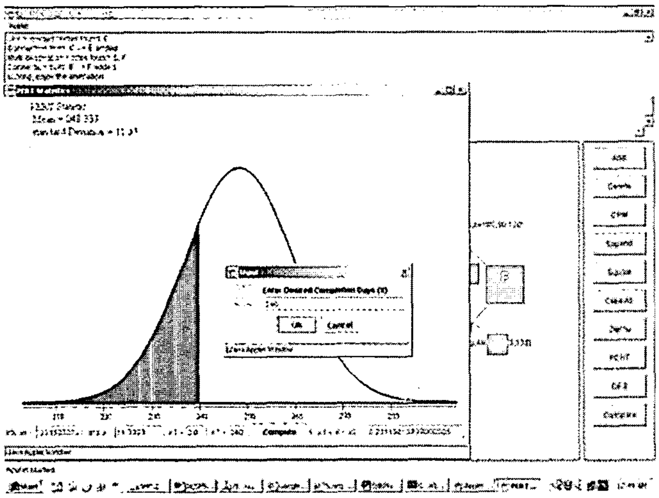

Figure 2. Probabilistic PERT in SPSS III

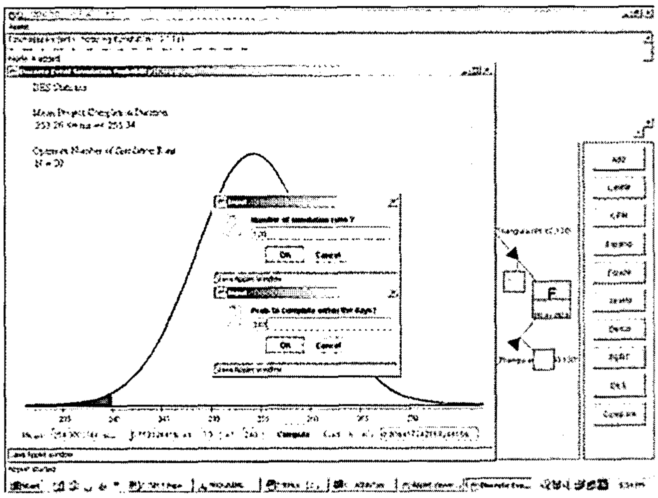

Figure 3. DES Mode in SPSS III

2.1 Computing the Confidence Interval for the Mean Project Duration in a Simulation

In DES mode, SPSS III generates N project durations that correspond to the N runs conducted by the system in a simulation. The mean project duration (mx) is the mean of the N durations. The descriptive statistics obtained from simulation runs in DES mode (e.g., mean project durations and standard deviations) are used for computing the confidence interval for the mean project duration as shown in Figure 3. The confidence interval is established by automating the Central limit theorem (Ang and Tang 1975, Khisty and Mohammadi 2001).

2.2 Determining the Minimum Number of Simulation Runs

A simulation cannot be executed indefinitely. Generally speaking, a large number of simulation runs leads to a statistically valid simulation results but it will consume a long computing time. Although new computing technology has made computing time not a major concern for small to medium projects, computing time may be still crucial for large projects, e.g., with thousands of activities. It is well accepted in the literature that 10,000 to 30,000 simulation runs are in general quite enough for simulation experiment to realize a desirable level of confidence (Van Slyke 1963). Nevertheless, picking the easy way and ran the simulations with arbitrary numbers (e.g., 10,000 runs) is not an acceptable engineering practice. Therefore, the SPSS III optimizes the project scheduling simulation by computing a minimum number of simulation runs that would produce reliable results at a given confidence (e.g., 95 %).

In simulation-based project scheduling, the minimum number of simulation runs is project-specific. Since the number of simulation runs is a critical factor for a successful simulation, a statistical analysis method was retained in SPSS III to calculate the minimum number of simulation runs (Lee 2004; WSC).

2.3 Testing the Significance of the Difference between Two Simulations

Once the minimum number of simulation runs N is determined according to the section 2.2, several simulations (each consisting of N runs) are conducted. The results of the different simulations may produce a variety of project durations depending on the variability and distribution of activity durations. The Student's t-test is used to assess if the mean project durations generated by two simulations (each consisting of N runs) are statistically different from each other (Wilcox 1997). Since the numbers of simulation runs (N) in all simulations are the same, the paired t-statistic is calculated.

To test the difference between the mean project duration values of two independent simulation experiments, t-Test is implemented in SPSS III. The t-test was conducted to see whether the mean project durations generated by a pair of simulation experiments (each consisting of 120 runs) are statistically different from each other by using the automated t-test analysis component as shown in Figure 4. Testing the difference between the variances of two independent simulation runs can also be conducted to check if they are from the same population.

2.4 Analyzing the Sensitivity of Input Parameters on Simulation Output Data Analysis

The skewness and variability are useful standard measures in engineering statistic that characterize the location and shape of activity duration. However, those characterization methods have not been applied as analysis method in project scheduling simulation. Each activity in a project network may have different combinations of the measures. They jointly effect on the project duration. However, it is always questionable how a change of the measures (e.g., changes in input data) affect on the project duration (e.g., results in output data). Although how sensitively the change of a measure affect on project duration is project specific, an automated methodology can be a great help to predict the total behavior of a network. The "What-If Analyzer" implemented in SPSS III has two analyses methods, Variability analysis (Figure 5) and Skewness analysis (Figure 6) as follows.

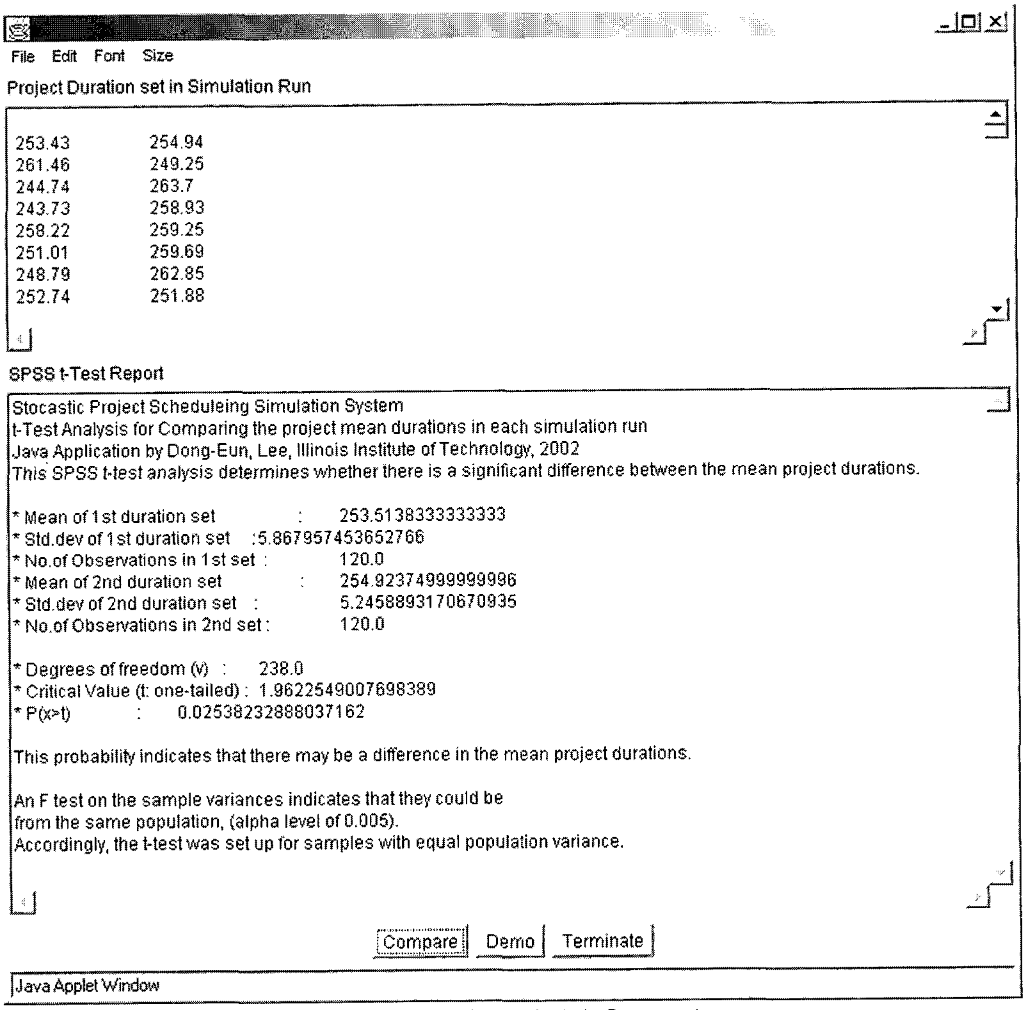

Figure 4. Automated t-test Analysis Component

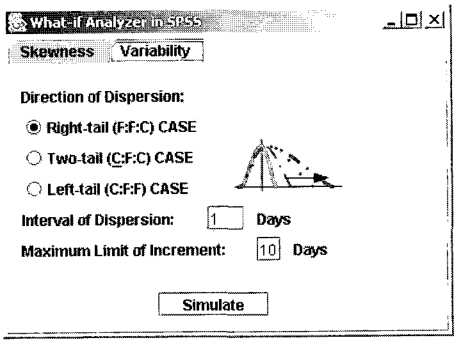

Figure 5. Variability Analysis in What-If Analyzer

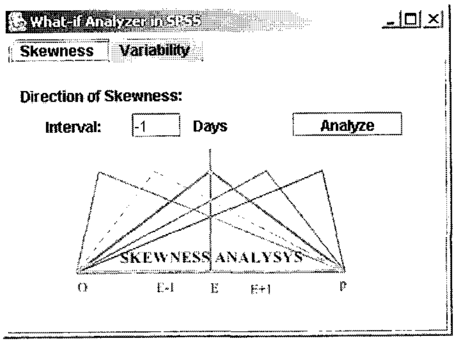

Figure 6. Skewness Analysis in What-If Analyzer

The automated component provides a mean to change the location and shape of activity durations in each simulation experiment. The automation of methodology can lead to easiness in data creation, collection, and analysis for finding heuristic in a project network.

Sensitivity analysis (Ossenbruggen 1984) measures the impact of changing three time estimates (optimistic, expected, and pessimistic time) on project durations. Skewness analysis could be performed to see how the project duration changes as the most likely time estimate is considered with other two (i.e., optimistic and pessimistic time estimates) are hold as same. Variability analysis, in turn, could be performed to see how the project duration changes as the two-time estimates (i.e., optimistic and pessimistic time estimates) are changed in an interval while holding the most likely time estimate as it.

The What-If Analyzer allows changing the skewness and variability of activity duration by assigning different time estimates in different simulation experiment. The automated component is implemented to define different location and shape of activity data set. It provides an efficient ways to estimate the sensitivity of simulation output (e.g., project durations) in response to change in input parameters (e.g., three time estimates).

3. Feasibility and Validity of the system

When applied to an example (Figure 1), the SPSS III system calculates (1) a deterministic project duration of 240 days in CPM mode, (2) a probability of 23% of completing the project in 240 days in PERT mode, and (3) a project duration of 253.26 to 255.34 days with 95% certainty in DES mode. Because all paths in the network modeled in Figure 2 compete with each other to be a critical path, the project duration tends to be longer in DES mode. As mentioned earlier, it is well known that PERT is always more conservative than CPM (MacCrimmon and Ryavec 1962, Halpin and Riggs 1992, Shi 2001). In the other hand, DES may be more or less conservative than PERT (Lee 2004) as shown in Figure 7.

The variability of PERT' mean project durations is greater than that of DES's mean project duration. In other word, PERT assumes that the output data are more widely scattered than those of DES. As shown in Figure 7, 260 days is the break-even point where DES is more conservative when the queried completion duration is less than 260. On the other hand, PERT is more conservative when the queried completion duration is greater than 260. In PERT, The probability to complete the project within 267 is 95%. In DES, the probability to complete the project within 262.7 and 267 days are 95%, and 99.20%. Therefore, DES is more optimistic than PERT, when the queried project duration is greater than 260.

4. Conclusion

The SPSS III provide a comprehensive means to analyze a project network by using deterministic CPM, probabilistic PERT, and stochastic DES methods in a comparative way. It automatically determines the minimum number of simulation runs necessary to get valid simulation result as well as tests the validity of a simulation experiment by using paired t-test analysis method which identifies

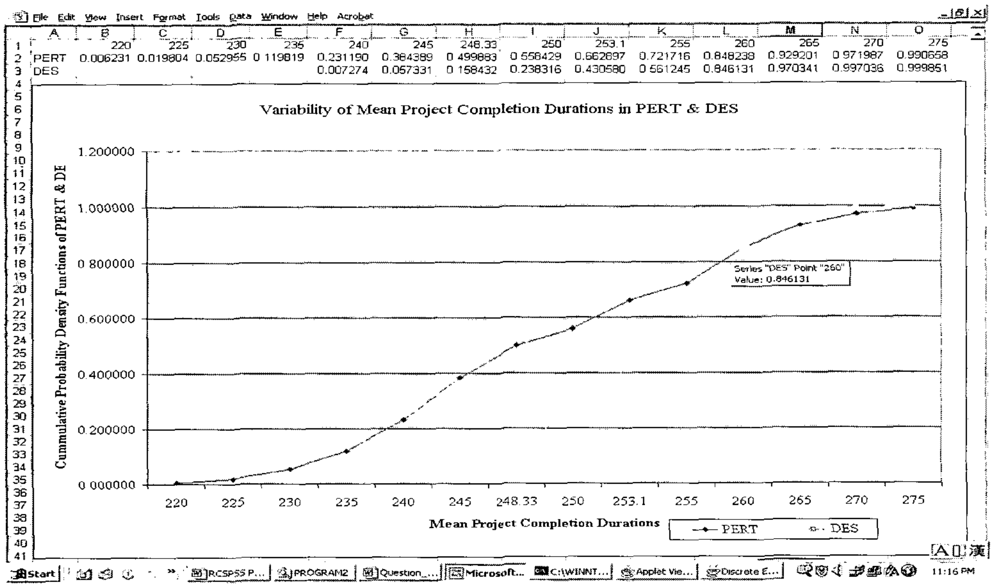

Figure 7. Variability of Mean Project Completion Durations in PERT and DES

and eliminates outlier cases that do not conform to the simulation population.

The What-if Analyzer is useful to measure the sensitivity of variability and skewness on project completion duration. High variability in activity durations tends to result in longer project durations and therefore more conservative results in DES. However, it is found that the interval between expected and pessimistic times is the most important predicates, which jointly effect on the project completion duration. Therefore, the very interval between expected and pessimistic times should be paid more attention in project schedule network modeling. The automated sensitivity analysis component helps to investigate how the skewness and variability of activity durations affect project duration.

References

Ang, Alfredo. H-S., and Tang, Wilson. H. (1975). "Probability Concepts in Engineering Planning and Design: Volume I Basic Principles." Analytical Models of Random Phenomena, John Wiley & Sons, New York. 189-191, 231-235

Barrie, D. S., and Paulson, B. C. (1992). Professional Construction Management including CM, Design-Contract and General contracting. 3rd ed, McGraw-Hill, New York, NY.

Bennett, F. L., (2001). "Discussion: Simplified CPM/PERT Simulation Model." J. Constr. Eng. Mgnt, ASCE, 127(6), 513-515

Crandall, C. K., (1977). "Analysis of Schedule Simulation." J. Construction Division, ASCE, 133(2), 387-394.

Douglas, E. D. (1978). "PERT and simulation." Proceedings of the 10th conference on Winter simulation - Volume 1, Miami Beach, FL, 89 ? 98

Evans, M., Hastings, N., and Peacock, B. (2000). Statistical Distributions. Wiley, New York, NY.

Hajjar, Dany., and AbouRizk, Simaan. M. (1999). "Simphony: An Environment for Building Special Purpose Construction Simulation Tools." Proc. of the 1999 Winter Simulation Conference. IEEE, Piscataway, N.J: 998-1006.

Halpin, Daniel. W. (1977). "CYCLONE - A Method for Modeling Job Site Processes." J.Construction Division, ASCE, 103(3), 489-499.

Halpin, D. W. and Riggs L. S. (1992). Planning and Analysis of Construction Operations. Wiley, New York, NY.

Khisty, C. Jotin., and Mohammadi, Jamshid. (2001). "Fundamental of System Engineering with Economics, Probability, and Statistics." Principles of probability, Prentice Hall. Upper Saddle River, N.J. 208-214.

Huang, R., Grigoriadis, A.M., and Halpin, Daniel. W. (1994). Simulation of Cable-stayed Bridges Using DISCO. Proc. of 1994 Winter Simulation Conference. IEEE, Piscataway, N.J. 1130-1136.

Ioannou, P.G. (1989). "UM_CYCLONE User's Guide." Dept. of Civil Engineering, The University of Michigan, Ann Arbor. Mich.

Law, M. A., and Kelton, W. D. (2000). Simulation Modeling and Analysis. 3rd ed, McGraw-Hill, Boston, MA.

Lee, Dong-Eun. (In-Press 2005). "Probability of Project Completion Using Stochastic Project Scheduling Simulation (SPSS)." J. Constr. Eng. Mgnt, ASCE. 131(3).

Lee Dong-Eun. (2004). "Statistical Analyses for Simulating Schedule Networks" Proceedings of the 2004 Winter Simulation Conference. R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, eds., IEEE Press, Washington.D.C. p.1283-1289.

MacCrimmon, K. R., and Ryavec, C. A. (1962). An Analytical Study of the PERT Assumptions. Memorandum RM-3408-PR.

Martinez, C. J. and Ioannou, G. P. (1997) "State-Based Probabilistic Scheduling Using STROBOSCOPE's CPM Add-On" Proceedings, Construction Congress V, Stuart D. Anderson,ed, ASCE, Minneapolis, MN, 438-445

Moder, J. J., Phillips, R. C. and Davis, W. E. (1983). Project management with CPM, PERT, and precedence diagramming. 3rd ed., Van Nostrand Reinhold, New York, NY.

O'Brien, J. J. (1993) CPM in Construction Management. 4th ed., McGraw-Hill, New York, NY.

Ossenbruggen, P.J., (1984). System Analysis for Civil engineers, Wiley.

Shi,. (1999). Activity-Based Construction (ABC) Modeling and Simulation Method." J. Constr.Eng. Mgnt, ASCE, 125(5), 354- 360.

Shi, J. (2001). Activity-Based Construction (ABC) Modeling and Simulation. Dept. of Civil Engineering, Illinois Institute of Technology, Chicago, IL.

Simphony. (2003). "Template for Project Scheduling Risk" NSERC/Alberta Construction Industry <http://irc.construction.ualberta.ca/html/research/PERT.html>

Van Slyke, M. R. (1963). "Monte Carlo Methods and The PERT Problem." Operation Research,11. 839?860

Wilcox, R. R. (1997). "Introduction to Robust Estimation and Hypothesis Testing." Academic Press. San Diego, CA.