SIMULATION OF ENGINEERING SYSTEMS : SOME IMPORTANT REQUIREMENTS

by

Dr. Malik Ranasinghe
Senior Lecturer, Department of Civil Engineering
University of Moratuwa

Abstract

Simulation is often suggested to model and validate engineering systems. This paper highlights some often ignored requirements necessary for interpretation, modelling and verification of simulation models applicable to a broad class of engineering problems. Basic sampling theory is used to determine the confidence of results obtained from Monte Carlo simulation. Examples of known limiting cases to verify the simulation process are suggested. A random number modification approach to treat correlations that exist in engineering systems is suggested to obtain reality in modelling of systems.

1.0 Introduction

Simulation is often suggested to model (Dickmann, 1983; Hayes et al., 1986; Jaafari, 1988; Jayawardane, 1991; Perry and Hayes, 1985; Thompson and Wilmer 1985) and validate (Ang et al., 1975; Kottas and Lau, 1978, 1983; Ranasinghe and Russell, 1991) engineering systems. The objectives in both cases are to obtain, a realistic representation of the physical process, prediction for the performance measures and quantifications of their uncertainty. Simulation's strength lies in its ability to focus on the real characteristics of the problem as opposed to simplifications required to make it analytically tractable. Often, simulation is used to model and predict performance of engineering systems without paying proper attention to the fundamentals on which the simulation process is based. Hence, the motivation to highlight some often ignored requirements for a valid simulation process.

Jayawardane (1991), has described in detail what computer simulation is, its basics, the reasons to simulate, the process, the available languages for computer simulation and possible areas of application in construction operations. For discussion purposes herein, we focus on Monte Carlo simulation, the most popular and the simplest approach to simulate an engineering system. We explore three issues that are often ignored, but necessary requirements for a valid simulation process, namely,

1. its confidence based on the number of iterations (cycles);
2. its verification using known results; and
3. its ability to treat correlations in the system.

Various views have been expressed on the number is iterations (cycles) that should be performed (Bennett and Ormerod, 1984; Flanagan et al., 1987; Jaafari, 1988; Inyang 1983; Perry and Hayes, 1985). The recommended number ranges from 100 to 1000 iterations. We have however, returned to basic sampling theory to determine the number of iterations required to be able to explore confidence levels.

Rarely, is a simulation process verified using known results to find out whether the simulation model is working properly. We explore two limiting cases that were used to verify the simulation process for the numerical example of a project duration network used by Ranasinghe (1992a).

Treating correlations that exist in the systems is often ignored due to the inability to include them in the simulation model. We suggest a random number modification approach to treat correlations.

This paper is structured as follows. The Monte Carlo simulation and its confidence levels based on the selected number of iterations is described in the next section. Two limiting cases to verify the Monte Carlo simulation process for a project duration network are presented in the third section. The first is the duration of a parallel network of identical activities, while the second is the duration of a single path network. A random number modification approach to include correlations in a simulation model is suggested in the fourth section. Finally, we summarise and conclude the paper in the fifth section.

2.0 Monte Carlo Simulation

Conceptually, performing a Monte Carlo simulation is simple. It requires:

- deterministic model(s) to describe the performance measure(s) of the engineering system;

- identification of the random variables in the models(s);
- a cumulative distribution for each random variable;
- a random number generator; and
- a sample value from each distribution for each iteration using a random number from the uniform distribution on the interval [0.1], (i.e. U(0,1), as the entry point in a cumulative distribution function of the variables (see Fig. 1).

In each iteration (cycle), values for the random variables obtained using the random number generator and the cumulative distributions are used to evaluate a value for the performance measure of the engineering system. The larger the number of iterations, the more reliable are the results from the simulation (Eilon and Fowkes, 1973; Flanagan et al., 1987; Inyang, 1983; Jaafari, 1988; Hertz, 1964; Hull, 1977; 1980; Kryzanowski et al., 1972; Newendorp, 1976; Riggs, 1989; Van Tetterode, 1971). The procedure described above however, implies that each random variable is independent of the others. In most engineering systems variables are correlated (Cooper and Chapman, 1987; Inyang, 1983; Perry and Hayes, 1985).

The literature is diverse on the number of iterations that should be performed for an "acceptable" simulation (Flanagan et al., 1987; Jaafari, 1988; Inyang 1983; Perry and Hayes, 1985). The recommended numbers range from 100 to 1000 iterations. However, most of these recommendations are not supported theoretically or empirically, and may not be applicable in all situations (Inyang, 1983).

Bury (1975) has shown that the error band of the cumulative distribution function for the performance measure generated from a simulation can be evaluated from the Kolmogoroff statistic. Error band is the accuracy to which the cumulative distribution function for the performance measure generated from the simulation approximates to the unknown cumulative distribution function of the performance measure of the engineering system. For example, at 95% confidence level, the error band brackets the unknown cumulative distribution function for the performance measure in 95% of all simulation samples. Fig. 2 illustrates the width (E) of the error band at 95% and 90% confidence levels versus number of iterations. That is, given a cumulative distribution function generated for the performance measure Y, and a specified value y, the band on $P[Y \leq y]$ is,

$$P[Y \leq y] \pm E. \tag{1}$$

For example, at the 95% confidence level, 100 and 1000 iterations have error bands of $\pm$ 13.6% and $\pm$ 4.3% respectively. For an error band of $\pm$ 2% at least 4600 and 3800 iterations are required at 95% and 90% confidence levels respectively.

The error of the estimates for expected value (mean) and variance for performance measures from a simulation at a desired confidence limit (probability) can be approximated from the assumption that the underlying distributions for the statistics are normal (Inyang, 1983). The error is the percentage within which estimates for expected value (mean) and variance are to be of the true values. Fig. 3 illustrates the error of the mean and variance at 95% probability. For example, 1000 iterations will give an error of $\pm$ 6% and $\pm$ 8.65% for the mean and variance of the performance measures respectively.

3.0 Verification of Monte Carlo Simulation

Since Monte Carlo simulations are used to model and validate engineering systems, it is essential that the accuracy of simulation process as implemented be verified. Two limiting cases for which the analytical solutions are exact are utilised as examples to demonstrate the verification of the simulation process that was developed to validate the numerical example of a project duration network used by Ranasinghe (1992a).

3.1 First Limiting Case : Parallel Network

A parallel network consisting of thirty five identical activities in five parallel paths as shown in Fig. 4 is used as the first limiting case. Every path in a parallel network is independent of all the other paths. When using the critical path (or PERT) approach every path in the network is a critical path. Then, the cumulative distribution function for project duration should be the cumulative distribution function obtained from any path duration.

However, the bounds on the probability $p(t)$, to complete a project in time t are given by (Ang et al., 1975),

$$\prod_{i=1}^{n} P(T_i \leq t) \leq p(t) \leq \min_i P(T_i \leq t). \tag{2}$$

The lower bound of project completion time probability $p(t)$ is the upper bound for project duration for a desired probability of success (Ranasinghe, 1992a). The cumulative distribution function for project duration from a parallel network can be obtained from the lower bound of $p(t)$, because each path is independent of the others, leading to a consideration of all paths. Therefore, the cumulative distribution function for project duration for the parallel network from a valid Monte Carlo simulation process should give the same cumulative distribution function as from the lower bound of project completion time probability $p(t)$. In other words, it should be the cumulative distribution function for the upper bound from the PNET algorithm (Ang et al., 1975).

The expected value for duration for all activities is assumed as 3.64 *days* and the standard deviation as 0.67 *days*, respectively. The cumulative distribution functions for project duration from the critical path, upper bound of the PNET algorithm (Ang et al., 1975), and a Monte Carlo simulation of 20,000 iterations (cycles) are depicted in Fig. 5. Note the agreement between the cumulative distribution function generated from the simulation and that for the upper bound of the PNET algorithm and the significant underestimation of the cumulative distribution function from the critical path (or PERT) approach. The statistics for project duration from the analytical solutions and simulation are given in Table 1. Again, note the agreement between statistics from the simulation with those when all the paths are considered.

Table 1
Statistics for Project Duration for First Limiting Case

Project Duration (months)	Expected Value	Standard Deivation
Critical Path	25.51	1.76
All Paths	27.63	1.24
Monte Carlo Simulation	27.60	1.30

3.2 Second Limiting Case : Single Path

The lower bound for project duration (critical path/PERT approach) is exact only for a single path network (Ranasinghe, 1992a). Therefore, for the second limiting case we have assumed a network consisting of a single path in the parallel network with activity duration statistics as before. The cumulative distribution functions for project duration from the analytical solution and a Monte Carlo simulation of 20,000 iterations are depicted in Fig. 6. The statistics for project duration from the analytical approach and simulation are given in Table 2.

Table 2
Statistics for Project Duration for Second Limiting Case

Project Duration (months)	Expected Value	Standard Deviation
Analytical Solution	25.48	1.76
Monte Carlo Simulation	25.49	2.03

For the single path case, the analytical solution for expected value and standard deviation for project duration are exact. We also attribute the difference in the standard deviations to the sensitivity higher order simulation moments have to extreme points of the simulation (Kottas and Lau, 1983).

These limiting cases provide a verification of the Monte Carlo simulation process used for validating the numerical example of a project duration network used by Ranasinghe (1992a).

4.0 Treatment of Correlations in Simulation

The importance of treating correlations between variables in Monte Carlo simulations has been long recognised (Eilon and Fowkes, 1973; Inyang, 1983; Hertz, 1964; Hull, 1977, 1980; Kryzanowski et al., 1972; Newendorp, 1976; Thompson and Wilmer, 1986; Van Tettcrode, 1971). None of the suggested methods however, has been rigorously validated. After an extensive review of the available techniques, Inyang (1983) proposed the following approach to model correlations in Monte Carlo simulations for engineering systems.

1.Random numbers are generated for each of the variables that make up the simulation model. A column of random numbers is thus generated.

2.The correlation factors between variables have to be input as a matrix. The random numbers are modified depending on the correlation with each other. Any type of correlation factor (total, partial or no correlation) can be handled.

3.The value for a variable is obtained depending on the value of its modified random number as a result of the correlation between the variables.

We agree with Inyang (1983), that the above procedure is the most suitable approach to model correlations in simulations. However, two shortcomings have to be highlighted. First, an algorithm for the modification of random numbers was not developed by Inyang (1983). Second, the correlation matrices elicited for the simulation have to be positive definite (Ranasinghe, 1992b).

The random numbers can be modified by extending the algorithm developed by Van Tetterode (1971) to the multivariate situation. Even though the random number correction is pairwise, the use of a positive definite correlation matrix to modify the random numbers assigned to the variables recognise the multivariate situation. The random number correction suggested by Van Tetterode, (1971), is as follows.

$$RN_{ij} = RN_j + \alpha_{ij}(RN_i - RN_j) \qquad (3)$$

where RN_i and RN_j are the ith and jth random numbers in the column generated from $U(0,1)$, (step 1, Inyang, 1983), RN_{ij} is the ijth random number corrected for the correlation between variables i and j in the matrix of corrected random numbers, and α_{ij} is the correction factor given by,

$$\alpha_{ij} \doteq \frac{\rho_{ij}^2 \pm \rho_{ij}\sqrt{1 - \rho_{ij}^2}}{2\,\rho_{ij}^2 - 1} \qquad (4)$$

where ρ_{ij} is the correlation coefficient between variables i and j. The correction factor α_{ij} is lies in the interval, $0 \lesssim \alpha_{ij} \lesssim 1$ (see Fig. 7) for all correlation values. The modification from equation (3) and (4) ensures that the corrected random numbers are within the interval [0,1].

A simple numerical example is presented to demonstrate the random number modifications process. Assume a three variable model having the following positive definite correlation matrix, **R**, where

$$\mathbf{R} = \begin{bmatrix} 1.0 & -0.48 & 0.42 \\ -0.48 & 1.0 & -0.69 \\ 0.42 & -0.69 & 1.0 \end{bmatrix}$$

and the column of random numbers generated from $U(0,1)$ as

$$[0.32\ 0.75\ 0.14]^T$$

Then, the matrix of α_{ij} values from equation (4) is,

$$\alpha = \begin{bmatrix} 1.0 & 0.35365 & 0.31638 \\ 0.35365 & 1.0 & 0.48804 \\ 0.31638 & 0.48804 & 1.0 \end{bmatrix}.$$

where α is the matrix of correction factors corresponding to the correlation coefficients between variables. The matrix of random numbers corrected for the correlation values from equation (3) is,

$$\mathrm{RN} = \begin{bmatrix} 0.32 & 0.5979 & 0.1969 \\ 0.4721 & 0.75 & 0.4377 \\ 0.2631 & 0.4523 & 0.14 \end{bmatrix}$$

For each iteration, matrix RN can be computed from the generated column of random numbers. The strength of this approach is that the corrected random numbers will always remain within the interval [0,1] (i.e. will be from the uniform distribution $U(0,1)$. Then, a row selected from the matrix RN at random can be used as the random numbers for that iteration of the simulation knowing that the corrected numbers are truly random.

5.0 Summary

In this paper we have highlighted three often ignored requirements that should be considered when Monte Carlo simulations are used to model and predict performance of engineering systems.

The confidence of predictions for performance measures of engineering systems are always related to the number of iterations (cycles) of the simulation. The larger the number of iterations more accurate are the predictions for performance measures. The confidence on the generated cumulative distribution functions, expected values and variances can be described respectively in terms of error bands and percentages within which the predictions are to the true values. The necessity to verify the simulation process using known analytical solutions was highlighted as the second requirement. Examples of limiting cases used to verify the simulation process for a project duration network were demonstrated. The importance of treating correlations to obtain realistic representations was highlighted. A random number modification approach to include correlations in the Monte Carlo simulation process was developed.

Simulation is a powerful tool to model and validate engineering systems. It is therefore essential that proper attention is paid to the fundamentals on which the process is based, to obtain maximum benefits from simulating engineering systems.

Acknowledgements

The continuing support and facilities provided by the University of Moratuwa to carry out this research is gratefully acknowledged.

References

Ang, A.H-S., Abdelnour, J., & Chaker, A.A. (1975), "Analysis of Activity Networks under Uncertainty", Journal of Engineering Mechanics Division,. ASCE, Vol. 101 No. EM4, pp.373-387.

Bennett, J., & Ormerod, R.N. (1984), "Simulation Applied to Construction Projects", Construction Management and Economics, Vol. 2, pp. 225-263.

Bury, K.V. (1975), Statistical Models in Applied Science, John Wiley & Sons, New York.

Cooper, D.F. and Chapman, C.B. (1987), Risk Analysis for Large Projects, Models, Methods and Cases, John Wiley & Sons, New York.

Diekmann, J.E. (1983), "Probabilistic Estimating : Mathematics and Applications", Journal of Construction Engineering & Management, Vol. 109, No. 3, pp. 297-307.

Eilon, S. and Fowkes, T.R. (1973), "Sampling Procedures for Risk Simulation", Operational Research Quarterly, Vol. 24, No. 2, pp. 241-252.

Flanagan, R., Kendall, A., Norman, G., & Robinson, G.D. (1987), "Life Cycle Costing and Risk Management", Construction Management and Economics, Vol. 5, pp.S53-S71.

Hayes, R.W., Perry, J.G., Thompson, P.A., & Wilmer, G. (1986), "Risk Management in Engineering Construction, Implications for Project Managers", Project Management Group, UMIST, Thomas Telford Ltd., London, pp.5-36.

Hertz, D.B. (1964), "Risk Analysis in Capital Investment", Harvard Business Review, Vol. 42, No. 1, pp.95-106.

Hull, J.C. (1977), "Dealing with Dependence in Risk Simulation", Operational Research Quarterly, Vol. 28, No. lii, pp.201-213.

Hull, J.C. (1980), The Evaluation of Risk in Business Investment, Pergamon Press, U.K.

Inyang, E.D. (1983), Some Aspects of Risk Analysis for Decision Making in Engineering Project Management, PhD Thesis, University of Manchester, UMIST, UK.

Jaafari, A. (1988), "Project Viability and Economic Risk Analysis", Journal of Management in Engineering, ASCE, Vol. 4, No. 1, pp.29-45.

Jayawardane, A.K.W. (1991), "Computer Simulation Applied to Construction Operations", Transactions, The Institution of Engineers, Sri Lanka, Vol. 1, pp.11-30.

Kottas, J.F., & Lau, H-S. (1978), "Stochastic Breakeven Analysis", Journal of the Operational Research Society, Vol. 29, No. 3, pp. 251-257.

Kottas, J.F. & Lau, H-S. (1983), "A Four-Moments Alternative to Simulation for a Class of Stochastic Management Models", Management Science, Vol. 28, No. 7, pp.749-758.

Kryzanowski, L., Lusztig, P., and Schwab, B. (1972), "Monte Carlo Simulation and Capital Expenditure Decisions - A Case Study", Engineering Economist, Vol. 18, No. 1, pp.31-48.

Newendorp, P.D. (1976), "A Method for Treating Dependencies between Variables in Simulation Risk-Analysis Models", Journal of Petroleum Technology, Oct. 1976, pp.1145-1150.

Perry, J.G. and Hayes, R.W. (1985) "Risk and its management in Construction Projects", Proceedings, Institution of Civil Engineers, UK, Vol. 78, No. 1, pp.499-521.

Ranasinghe, M. (1992a), "Quanytification and Management of Uncertainty in Activity Duration Networks", Construction Management and Economics, E & FN Spon, U.K. (to appear).

Ranasinghe, M. (1992b), "Correlations for Engineering Systems", submitted to Engineer, Journal of The Institution of Engineers, Sri Lanka.

Ranasinghe, M. & Russell, A.D. (1991), "Economic Risk Analysis of Large Engineering Projects", Canadian Journal of Civil Engineering, Vol. 18, No. 3, pp.416-427.

Riggs, L.S. (1989), "Numerical Approach for Generating Beta Random Variates", Journal of Computing in Civil Engineering, Vol. 3, No. 2, pp.183-191.

Thompson, P.A. and Wilmer, G. (1985), "CASPAR - A Program for Engineering Project Appraisal and Management", 2nd International Conference in Civil & Structural Engineering Computing", Vol. 1, pp.75-81.

Van Tetterode, L.M. (1971), "Risk Analysis.... or Russian Roulette?", Transactions, American Association of Cost Engineers, pp. 185-191.

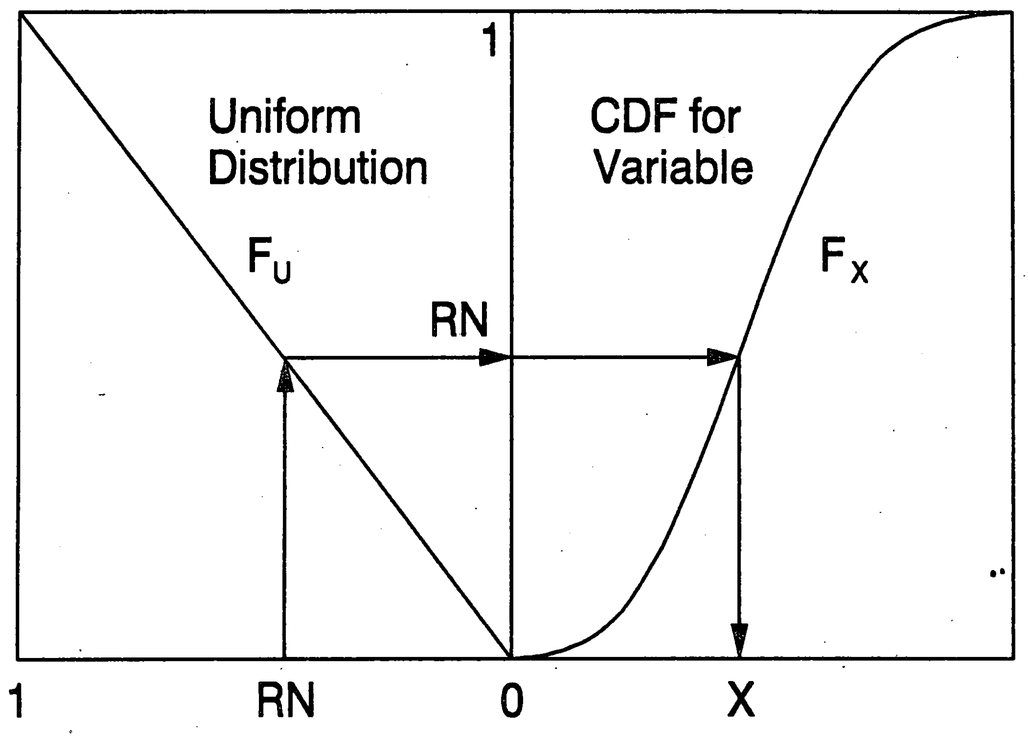

Fig. 1. Random Variate Generation

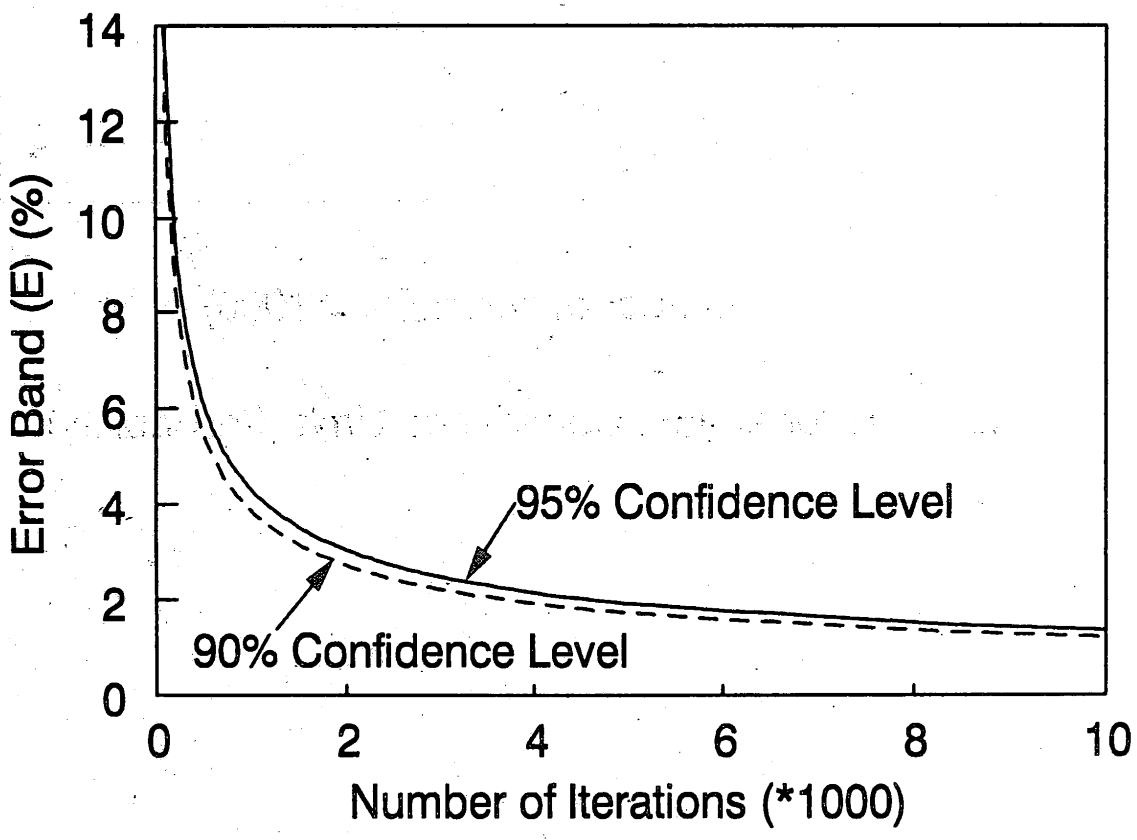

Fig. 2. Error Bands for the CDFs from Simulation

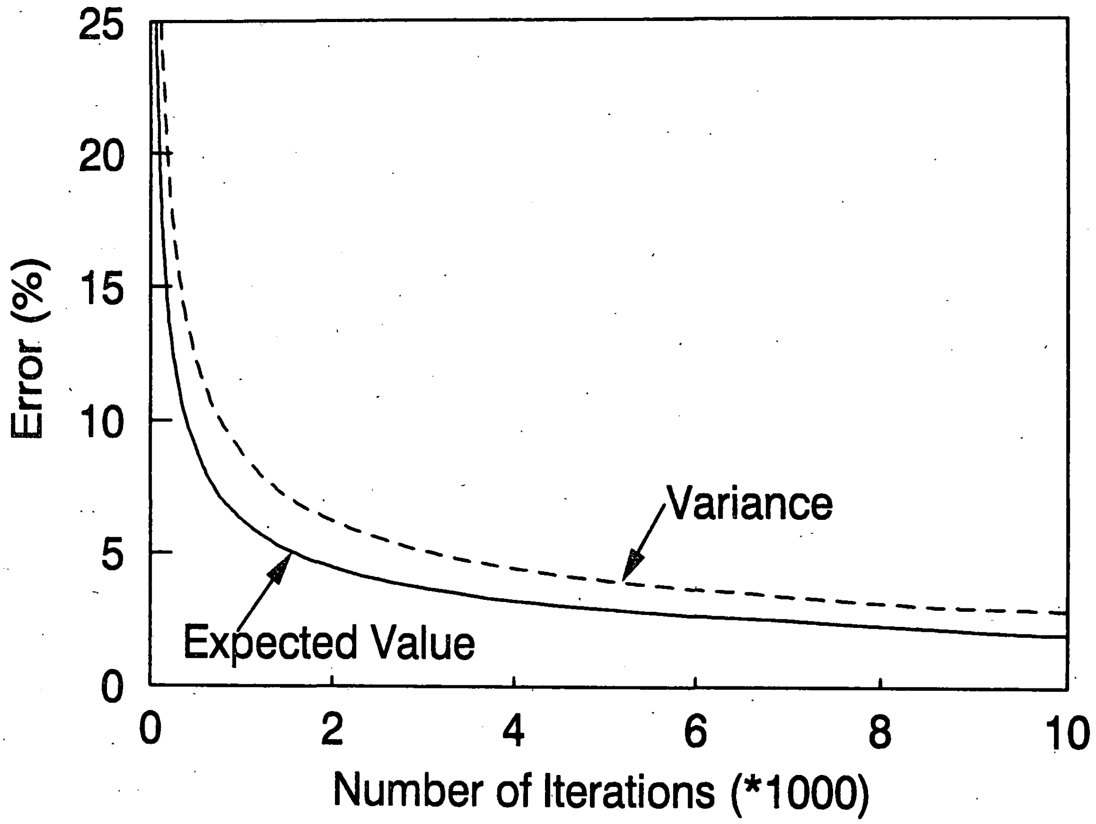

Fig. 3. Error at 95% Confidence Limit (Probability)

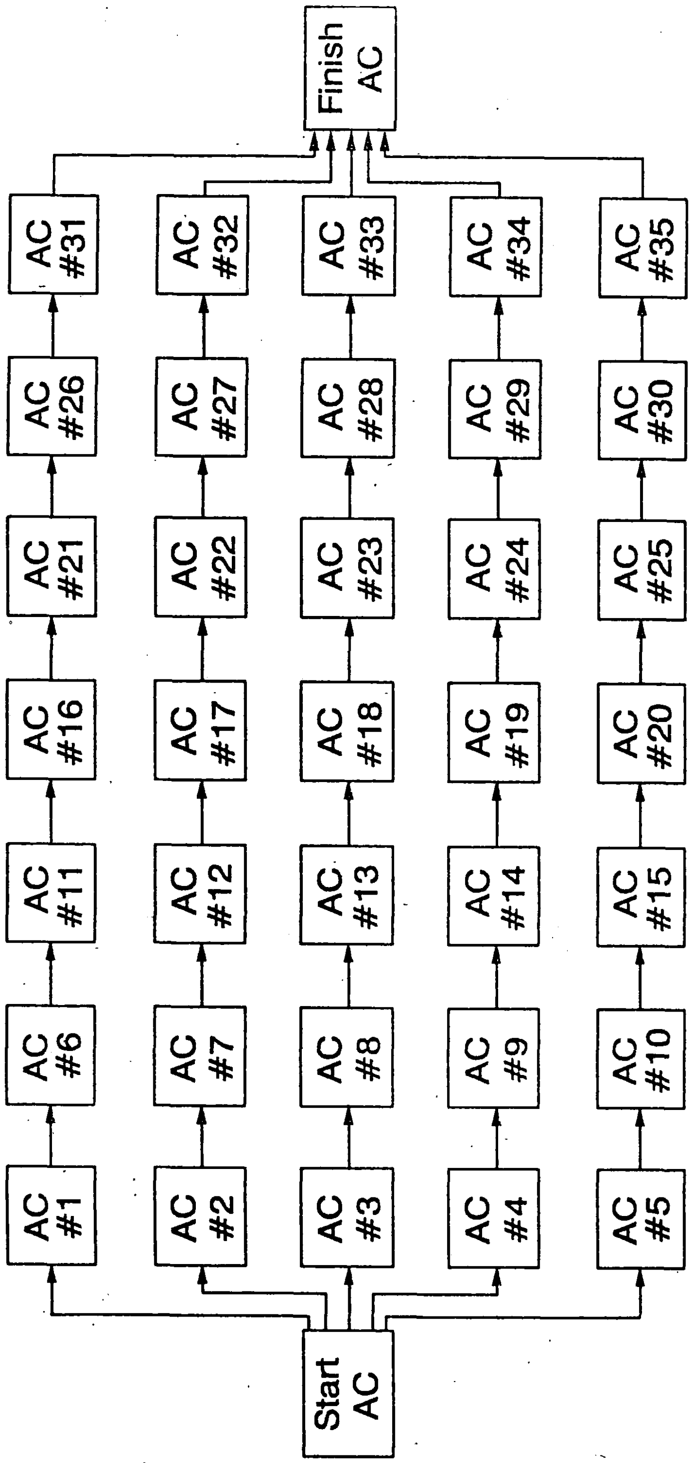

Fig. 4. The Parallel Network

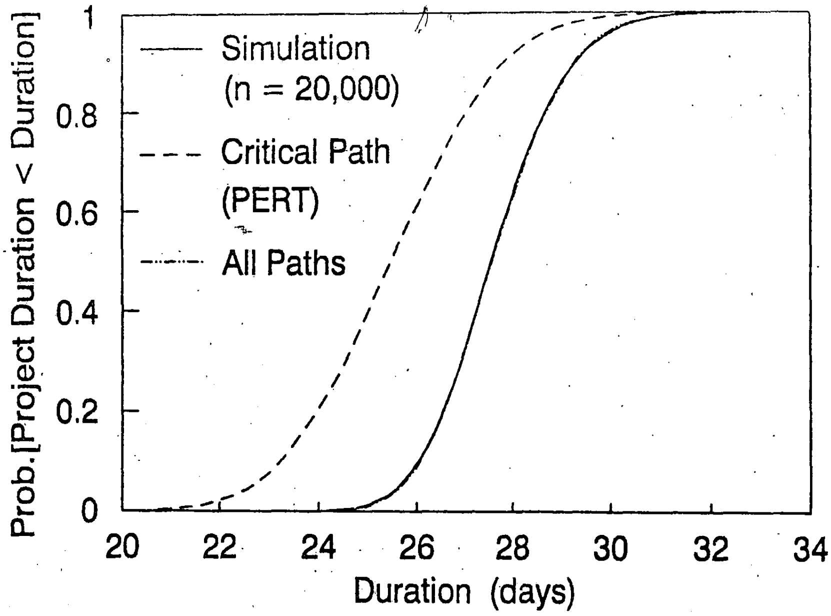

Fig. 5. CDFs for Project Duration for the Parallel Network

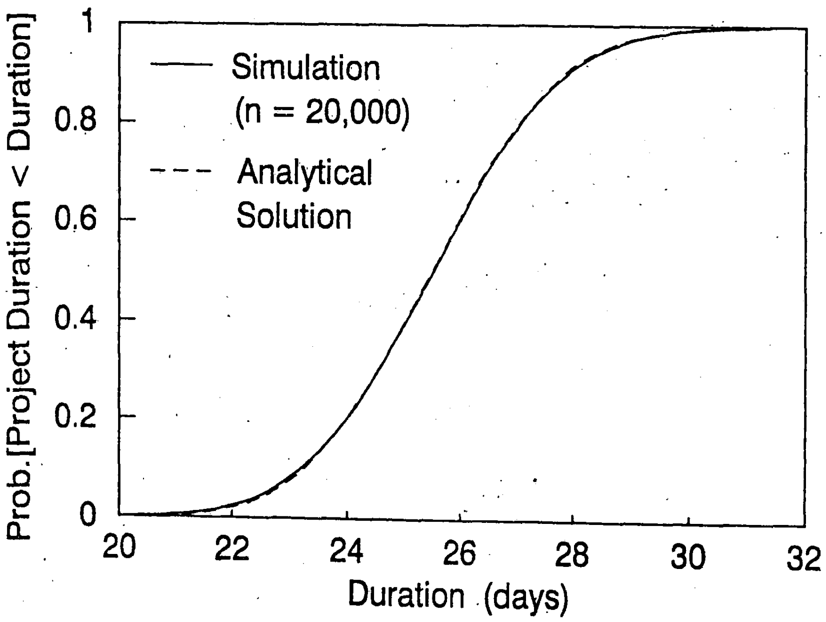

Fig. 6. CDFs for Project Duration from a Single Path

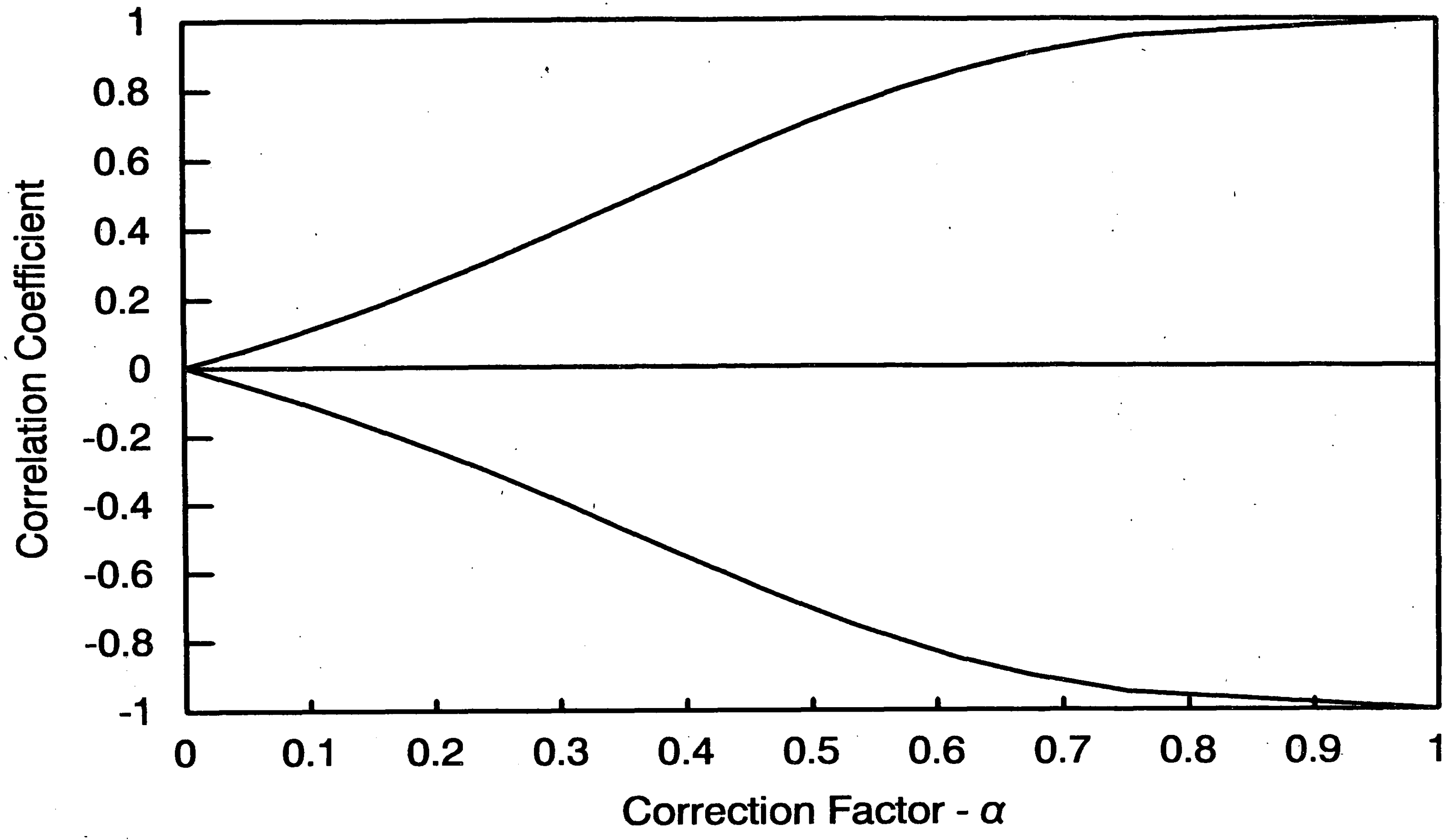

Fig. 7 Correction Factor α for Different Correlation Coefficient Values