Model for Evaluating Networks under Correlated Uncertainty—NETCOR

By Wei-Chih Wang[1] and Laura A. Demsetz[2]

Abstract: Construction activities are often influenced by factors such as weather, labor, and site conditions. When several activities are influenced by the same factor, their durations may be correlated. If many activities along a path are correlated, the variability of path duration will increase, possibly increasing the uncertainty of completing the project by a target date. This paper presents the simulation-based model NETCOR (NETworks under CORrelated uncertainty) to evaluate schedule networks when activity durations are correlated. Based on qualitative estimates of the sensitivity of each activity to each factor, uncertainty in an activity's duration distribution (grandparent) is distributed to several factor subdistributions (parents). Each subdistribution is broken down further into a family of distributions (children), with each child corresponding to a factor condition. Correlation is captured by sampling from the same-condition child distributions for a given iteration of the simulation. NETCOR integrates the effect due to each factor at the path level. Awareness of the factors to which a path is sensitive can provide management with a better sense of what to control on each path, particularly on large projects.

INTRODUCTION

Factors such as weather, labor skills, site conditions, and management quality can influence the duration of construction activities. The factors will often influence multiple activities on a particular project and may cause activity durations to be correlated (Carr 1979; Woolery and Crandall 1983; Ahuja and Nandakumar 1985; Levitt and Kunz 1985; Padilla and Carr 1991). For example, if bad weather occurs during concrete forming, it also is likely to influence other activities taking place at the same time (but perhaps in different locations) on the project. Bad weather will increase the duration of each weather-sensitive activity. Similarly, the duration of each weather-sensitive activity will decrease when the weather is good. If this correlation effect occurs for many activities along a network path, the variability of the path's duration may be significantly increased. If the path is critical or near critical, the variability in its duration will lead to a variability in the project's duration. Increased variability in the project's duration increases the uncertainty of completing the project by a target date. Therefore, the correlation effect has the potential to create an unexpected schedule overrun.

A better understanding of the effects of correlation would make the schedule a more useful management tool by providing a better estimate of uncertainty in the project's duration and helping to focus attention on the factors that have the greatest impact on the project's duration. When a path consists of several activities that are highly sensitive to the same factor, the path tends to be highly sensitive to this factor. In current practice, project managers may informally keep track of the factors that influence particular paths through the project. Better knowledge of the factors to which a path is sensitive and of the paths that are most sensitive to a particular factor could give management a better sense of what to control on each path, especially on large projects.

In the critical-path method (CPM) and probabilistic models such as program evaluation and review technique (PERT) and Monte Carlo simulation, the duration of each activity is entered or evaluated independently of the durations of other activities in the network. As they are currently used, these approaches will not capture the correlation that may exist between the durations of different activities in a schedule network. This paper presents the simulation-based model NETCOR (NETworks under CORrelated uncertainty), which incorporates the effect of correlation in network schedules and provides factor-sensitivity information to support schedule risk management. In the next section, previous research on correlation in network schedules is reviewed. The third section presents the NETCOR model. The effect of correlation for a small example network is evaluated in the fourth section. Results are summarized in the final section. The application of NETCOR to a recent construction project is described in a companion paper (Wang and Demsetz 2000).

[1]Asst. Prof., Dept. of Civ. Engrg., Nat. Chia-Tung Univ., 1001 Ta-Hsueh Rd., Hsin-Chu 300, Taiwan.

[2]Assoc. Prof., Engrg. Dept., Coll. of San Mateo, 1700 W. Hillside Blvd., San Mateo, CA 94402-3784.

Note. Discussion open until May 1, 2001. Separate discussions should be submitted for the individual papers in this symposium. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on November 4, 1997. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 126, No. 6, November/December, 2000. ©ASCE, ISSN 0733-9634/00/0006-0458–0466/$8.00 + $.50 per page. Paper No. 15243.

PREVIOUS WORK

Several models have been proposed to treat correlation in either a project's duration or cost. These models can be broadly divided into two groups: indirect and direct elicitation. Indirect-elicitation models evaluate the effect of correlation without explicit specification of correlation coefficients. Direct-elicitation models require correlation coefficients as input.

Indirect-Elicitation Models

MUD/DYNASTRAT

Carr (1979) developed the simulation-based MUD (Model for Uncertainty Determination) to evaluate a project network under uncertainty. The MUD simulation recognizes that the durations of activities are correlated when the activities share the same factors, such as site condition, crew efficiency, and equipment performance, which are independent of the calendar date, and the effect of weather, which is dependent on the calendar date. The MUD simulation was further refined into a component of the DYNASTRAT (DYNAmic-STRATegy) model for dynamically allocating resources (Padilla and Carr 1991). In DYNASTRAT, daily progress for an activity is the product of the work crew's standard productivity, a weather correction factor (based on historical data), and duration modifying factors, which are the combined effect of factors that are independent of calendar dates. Correlation is introduced by using the same sample drawn from a shared factor in each scheduling day. The evaluation of uncertainty is factor based, and uncertainty is treated as having both favorable and adverse effects.

J. Constr. Eng. Manage. 2000.126:458-466.

PRODUF

PRODUF (Project Duration Forecast) was developed to generate more objective duration distributions of activities before performing conventional Monte Carlo-simulation procedures (Ahuja and Nandakumar 1985). In generating a probabilistic activity duration, single-day progress for an activity is taken to be the product of terms that adjust for workday loss due to several factors. In each scheduling day, correlation is captured by setting workday loss caused by a factor as the same for all activities that share the factor. PRODUF, a factor-based model, captures positive correlation resulting from sharing of factors. For most factors, only the adverse effect of uncertainty is considered. A great deal of historical data would be required to construct appropriate distributions to describe the impact of several factors for each activity.

PLATFORM

PLATFORM, a rule-based method developed by Levitt and Kunz (1985), updates the durations of uncompleted activities based on the durations of completed activities. Each activity has associated risk factors that affect its duration. A risk factor is labeled a ''knight'' if it is shared by two or more ''short'' activities (completed activities whose actual durations were less than expected). Similarly, risk factors associated with two or more ''long'' activities (completed activities whose durations were greater than expected) are considered to be ''villains.'' Correlation is captured by reducing the durations of uncompleted activities associated with a knight and increasing the durations of uncompleted activities associated with a villain.

CEV

In the CEV (Conditional Expected Value) model, proposed by Ranasinghe and Russell (1992), the correlation coefficient between two variables, x and y, is derived from the conditional expected value, which is found by asking experts the question ''What is the anticipated value for y when $x = F$?'' The outcome of the elicitation process largely dominates the results (Ranasinghe and Russell 1992]; therefore, the model requires excellent quality of inputs. Uncertainty can have both favorable and adverse outcomes in the model and is not treated as being factor based.

Direct-Elicitation Models

Exact Simulation

To conduct an exact simulation analysis incorporating the effect of correlation, a proper assessment of the joint probability density function (PDF) for the correlated variables is need (Touran and Wiser 1992). The only joint PDF for which a well-organized theory of statistical inference currently exists is the multivariate normal distribution (Law and Kelton 1991). If variables are assumed to follow a normal distribution, then one needs only to have the multivariate normal distribution to generate correlated variables, given that the correlation coefficients between variables are known (Touran and Wiser 1992). Due to the difficulty of quantitatively assessing the correlation coefficient, qualitative estimates may be adopted (Touran 1993).

Quantile Simulation

A facility in commercially available Monte Carlo-simulation software may be used to capture the effect of correlation when the correlation coefficient is known (Chau 1995). The sampling procedure increases the probability of sampling the same quantiles from two PDFs when the correlation coefficient is positive. Similarly, when the correlation coefficient is negative, there will be a higher probability of sampling the nth percentile and $100 - n$th percentile from the two PDFs.

MSRN

In the MSRN (Modified Second Random Number) simulation, the value of the correlation coefficient is again used to influence the selection of random numbers. For example, Van Tetterode 91971) modifies the second random number by a proportion of the difference between the first and second random numbers.

Factored Simulation

A stochastic network model dealing with correlated durations was developed by Woolery and Crandall (1983). In this model, the duration of an activity consists of a time distribution for the activity duration under optimal conditions and a series of time distributions for various problems (factors) that may lengthen the activity duration. For a given activity, these problems are assumed to be independent. As an example, the delays for weather problems and labor and material shortages are assumed to be independent of each other. However, the effect of the same problem on multiple activities is assumed to be correlated. For example, the weather delay of one activity is assumed to be correlated with the weather delay of another activity. This correlation can be perfect or partial. The use of a base duration modified by a series of factor-related distributions is a logical way to evaluate the effect of uncertainty. However, because the base duration of an activity is assumed to be optimal, uncertainty can only have an adverse effect.

Summary of Past Work

For a model to be of use as a management tool, the following characteristics are desirable. A factor-based approach should be used to most naturally capture the correlation caused by the influence of factors on multiple activities. The required input should be such that it can be reasonably provided by project management. In particular, the method of introducing correlation should not require the user to express the extent to which activities are correlated because this information is not readily available. Finally, because the effect of uncertainty can be either adverse of favorable, the model should be capable of representing both increases and decreases in duration.

Among previously developed correlation-capturing models with respect to the desired characteristics, all four duration-focused models (i.e., MUD/DYNASTRAT, PRODUF, PLATFORM, and factored simulation) assume a positive correlation and adopt a factor-based approach. The indirect-elicitation simulation models (i.e., MUD/DYNASTRAT and PRODUF) require extensive inputs or historical data. Only PLATFORM meets all desired characteristics. However, PLATFORM relies on the performance of completed activities and treats all factors as having the same effects. Therefore, a new model that can more comprehensively meet the desired characteristics to capture correlation is desirable. Previous research has not reported on the significance of considering correlation. The work presented here and in a companion paper (Wang and Demsetz 2000) begins to address this issue.

NETCOR MODEL

General Description of NETCOR Model

This section provides an overview of a simulation-based model for evaluating NETCOR. In NETCOR, a factor-based

J. Constr. Eng. Manage. 2000.126:458-466.

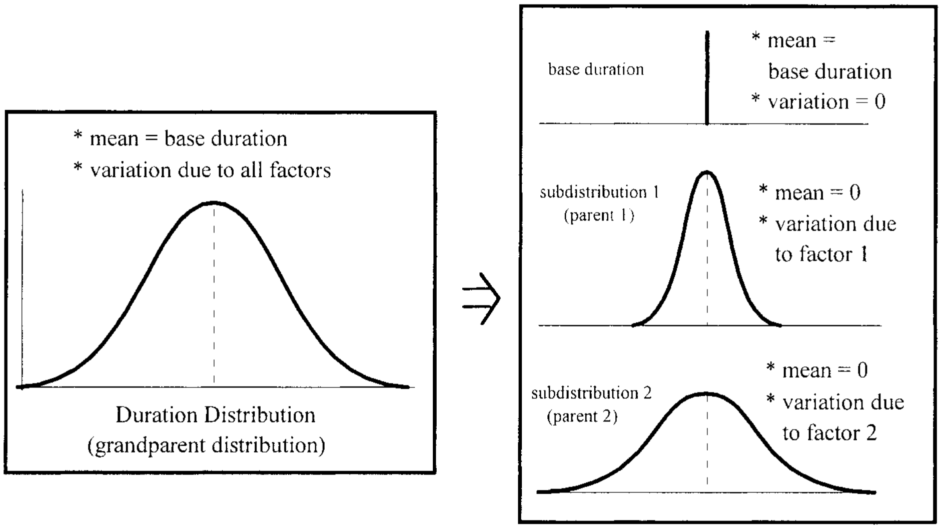

FIG. 1. Breakdown of Uncertainty

procedure to indirectly elicit positive correlation is used. This procedure is facilitated by an activity duration model that disaggregates the effect of uncertainty from a factor. The activity duration model considers both favorable and adverse effects of uncertainty. Although the effects of uncertainty for each factor are qualitatively estimated by the user and evaluated at the activity level, NETCOR integrates these individual effects of a factor at the path level.

Breakdown of Uncertainty

In NETCOR, the duration of an activity is considered to be a random variable. The duration distribution is represented by a ''grandparent'' distribution that is a combination of a base duration and variations due to different factors. The variations due to a particular factor are represented by a duration subdistribution, or ''parent'' distribution. The base duration is assumed to be deterministic, whereas the parent distribution for each factor is assumed to be a zero-mean random variable. This approach, shown schematically in Fig. 1, is appealing, both intuitively and mathematically. The base duration is treated as the user's best guess of an activity's duration under the expected factor conditions and is the expected value of the overall duration distribution (grandparent) for the activity. Deviations from the expected value due to various factors are introduced through the parent distributions.

Qualitative Estimates of Uncertainty Sensitivity

The derivation of parent distributions is based on subjective information. Project planners are asked to qualitatively estimate the level of influence that each factor has on the duration of each activity. For example, if the duration of an activity can vary greatly depending on the weather, the activity would be considered to have a high sensitivity to weather. It is believed that this approach of qualitative estimates is practical because the impact of uncertainties is easily expressed in linguistic terms (Chang 1987). There is no inherent restriction on the number of levels of influence used for each factor. The examples included in this paper use four levels of influence: high, medium, low, and no influence.

Factor-Based Correlation

The NETCOR model assumes that the duration of activities are correlated only through the impact of shared factors. Different factors are assumed to cause independent effects. For example, assume Activity 1 is sensitive to weather and labor

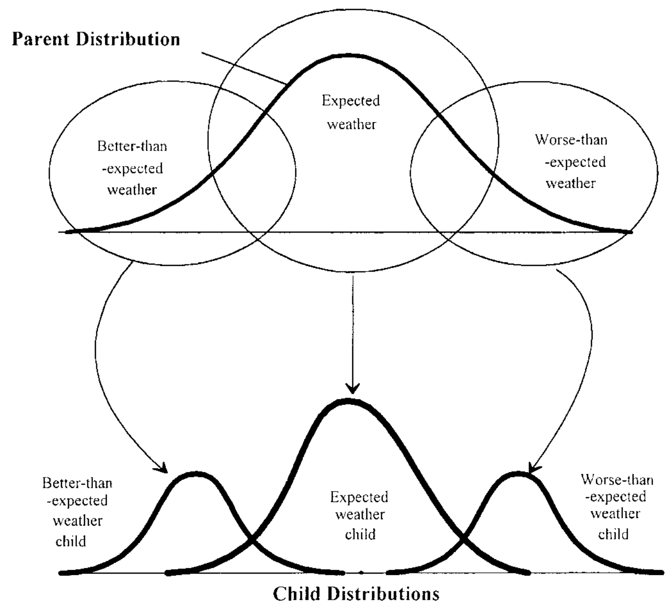

FIG. 2. Decomposition of Parent into Child Distributions

and Activity 2 is sensitive to weather and equipment. Only the weather-related parent distributions are correlated; the variations caused by labor and equipment are assumed to be independent.

The NETCOR model captures correlation by drawing duration samples from related portions of the parent distributions for activities that are sensitive to a given factor. For example, the upper part of Fig. 2 shows weather conditions classified as ''better-than-expected,'' ''expected,'' and ''worse-than-expected.'' Based on these three different weather conditions, the parent distribution due to weather is disaggregated into three child distributions (shown in the lower half of Fig. 2): better-than-expected, expected, and worse-than-expected. When a simulation is run under better-than-expected weather, sample durations will be independently drawn from the better-than-expected weather child of any weather-sensitive activities, and likewise for expected and worse-than-expected weather conditions. Therefore, better-than-expected weather tends to produce a shorter duration for each weather-sensitive activity, whereas worse-than-expected weather tends to produce a longer duration.

Child distributions may overlap. That is, the duration of an activity may be the same under both better-than-expected and

J. Constr. Eng. Manage. 2000.126:458-466.

expected weather conditions or the duration under expected weather may be shorter than the duration under better-than-expected weather. An extreme case is when child distributions are perfectly overlapped; then, the duration samples will always be drawn from the same child distribution under any conditions and there will be no correlation. On the other hand, when an activity is highly sensitive to weather, the child distributions will be distinct. The use of a child distribution is appealing because it is anticipated that NETCOR would be used to capture only the most important factors. Therefore, for a given set of factor conditions, there should be uncertainty in the duration. Furthermore, even when two activities are affected by the same factors, they may not be perfectly correlated.

In summary, the core of the NETCOR model is the two-step breakdown of uncertainty. The first breakdown separates uncertainty by a factor for each grandparent distribution; that is, grandparent distribution = base duration + parent distributions. The second breakdown separates uncertainty by a condition for each parent distribution; that is, parent distribution = family of child distributions. Correlation is introduced by sampling from the child distribution representing a given factor condition (such as worse-than-expected weather).

Path Analysis

When several activities along a network path are sensitive to particular factors, it is likely that the performance of the path will be dominated by these factors. With knowledge of factor-sensitive paths, management will have a better sense of what to control. Consider, for instance, the network shown in Fig. 3. Suppose the foundation, steel, and concrete paths are dominated by equipment performance, labor skill, and weather behavior, respectively. If the (equipment-sensitive) foundation path is critical, management effort should focus on ensuring the availability of equipment. If the (labor-sensitive) steel path is critical, management should focus on the quality and availability of labor. Controlling those factors that affect performance may offer far greater potential for improving performance than modifying or changing the work method (Thomas et al. 1990).

Development of NETCOR

Activity Duration Modeling

A model of the activity duration in which the effect of uncertainty is broken down by factors may be derived from the productivity perspective. Productivity is expressed as the amount of time required to finish a unit of work (Thomas et al. 1990).

$$PM = \text{time/quantity} \tag{1}$$

where PM = productivity measure.

In a deterministic environment, the estimated productivity measure for an activity i, $PM_{i(\text{estimated})}$, can be mathematically represented in the following form (Thomas et al. 1999):

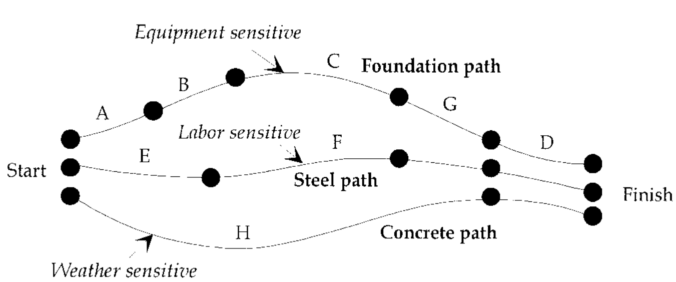

FIG. 3. Factor-Sensitive Paths

$$PM_{i(\text{estimated})} = PM_{i(0)} + \sum_{j=1}^{J} X_{i(j)} \tag{2}$$

where $PM_{i(0)}$ = estimated productivity measure for activity i under estimated outcomes of factors; and $X_{i(j)}$ = estimated constant representing the variation in the productivity measure of activity i with respect to factor j. The value of $X_{i(j)}$ can be negative, zero, or positive.

In a probabilistic environment, however, (2) should be rewritten

$$PM'_{i(\text{estimated})} = PM'_{i(0)} + \sum_{j=1}^{J} X'_{i(i)} = PM'_{i(0)} + X'_{i(1)} + X'_{i(2)} + \cdots + X'_{i(j)} + \cdots + X'_{i(J)} \tag{3}$$

where $PM'_{i(\text{estimated})}$, $X'_{i(1)} + X'_{i(2)}, \ldots, X'_{i(J)}$ = random variables. Each realization of $X'_{i(j)}$ represents the increase or decrease in the productivity measure of activity i due to factor j. Based on (3), D_i, the duration of activity i, may be expressed

$$\begin{aligned} D_i &= \text{quantity} \times PM'_{i(\text{estimated})} \\ &= \text{quantity} \times (PM'_{i(0)} + X'_{i(1)} + X'_{i(2)} + \cdots + X'_{i(J)}) \\ &= d_{i(0)} + d_{i(1)} + d_{i(2)} + \cdots + d_{i(J)} = d_{i(0)} + \sum_{j=1}^{J} d_{i(j)} \end{aligned} \tag{4}$$

in which $d_{i(0)}$ = estimated (or base) duration; and the random variables $d_{i(j)}$ = duration parent distributions of activity i due to factor j, $j = 1, \ldots, J$.

Eq. 4 shows the variations in the duration of an activity as a base duration and a series of parent duration distributions for various factors that may lengthen or shorten the activity duration. In NETCOR, the following assumptions are applied to this activity duration model:

- The value of $d_{i(0)}$ is assumed to be deterministic. In other words, the value of $d_{i(0)}$ is equal to the duration that is estimated under the expected conditions of all factors.
- The expected values of $d_{i(j)}$ are assumed to be zero; i.e., $m_{i(1)} = m_{i(2)} = \cdots = m_{i(J)} = 0$. Each sample of $d_{i(j)}$ thus represents a change from the expected duration.
- Values of $d_{i(1)}, d_{i(2)}, \ldots, d_{i(J)}$ are assumed to be independent of each other. That is, for a given activity, the impact of weather, labor skills, and other factors are assumed to be independent of each other.

Then, regardless of the type of marginal distribution of $d_{i(j)}$, the mean and variance of the duration of activity i are (Benjamin and Cornell 1970)

$$M_i = m_{i(0)} + m_{i(1)} + m_{i(2)} + \cdots = m_{i(0)} \tag{5}$$

$$\begin{aligned} \sigma_i^2 &= SD^2_{i(0)} + SD^2_{i(1)} + SD^2_{i(2)} + \cdots + SD^2_{i(J)} \\ &= SD^2_{i(1)} + SD^2_{i(2)} + \cdots = SD^2_{i(J)} \end{aligned} \tag{6}$$

in which M_i and σ_i = mean and standard deviation, respectively, for D_i (the grandparent duration distribution for activity i); and $m_{i(j)}$ and $SD_{i(j)}$ = mean and standard deviation, respectively, for $d_{i(j)}$ (the parent duration distribution for activity i due to factor j), with $SD_{i(0)} = 0$.

NETCOR finds M_i and σ_i for activity i and then determines $SD_{i(j)}$. In the example presented in this paper, the three-point estimates of PERT are used to calculate M_i and σ_i. However, there is no inherent restriction on the use of other methods (e.g., the direct assignment of a particular distribution to each activity) as long as the values of M_i and σ_i can be found.

J. Constr. Eng. Manage. 2000.126:458-466.

Scale System to Break Down Uncertainty by Factor

Based on qualitative estimates of the uncertainty sensitivity of activity i to factor j, a scale system is used to distribute the uncertainty associated with the grandparent distribution to the parent distributions. That is

$$\sigma_i^2 = \sum_{j=1}^{J} SD_{i(j)}^2 = SD_{i(1)}^2 + SD_{i(2)}^2 + \cdots + SD_{i(J)}^2 \quad (7a)$$

$$\sigma_i^2 = (w_1[Q_{i(1)}] + w_2[Q_{i(2)}] + \cdots + w[Q_{i(J)}]) \times K_i \quad (7b)$$

$$\sigma_i^2 = \left(\sum_{j=1}^{J} w_j[Q_{i(j)}]\right) \times K_i \quad (7c)$$

$$SD_{i(j)}^2 = w_j[Q_{i(j)}] \times K_i \quad (7d)$$

where $Q_{i(j)}$ = qualitative estimate (such as high, medium, low, or no) of the sensitivity of activity i to factor j; and $w_j[Q_{i(j)}]$ = scale for each level of influence. For example, the values of the estimates of high, medium, low, and no sensitivity for factor j can be represented by w_j[high], w_j[medium], w_j[low], and w_j[no], respectively. The constant K_i is an adjustment that ensures that σ_i^2 is preserved. Because $w_j[Q_{i(j)}]$ is fixed for a given factor j, K_i will be different for each activity. The value of w_j[no] is always zero. When $Q_{i(j)}$ represents a higher level of influence, the value of $w_j[Q_{i(j)}]$ is higher. Consequently, a larger portion of the variance will be distributed to a parent distribution that has a higher sensitivity. The value of $w_j[Q_{i(j)}]$ is determined by the user according to the relative importance of the factors. For example, if the user thinks that Factor 1 causes more uncertainty than other factors, then values of w_j[high], w_j[medium], and w_j[low] for Factor 1 should be higher than the corresponding values for other factors.

Breakdown of Uncertainty by Condition

In constructing a family of child distributions to represent changes in duration due to factor conditions, one goal is to preserve the mean and variance of the parent distribution. In other words, the mean and variance of the combination of the child distributions for a family should be the same as the mean and variance of the parent distribution. Mathematically, this relationship can be represented

$$m_{i(j)} = \sum_{c=1}^{C} p_{j(c)} \times o_{i[j(c)]} = 0 \quad (8)$$

$$SD_{i(j)}^2 = \sum_{c=1}^{C} p_{j(c)} \times (sd_{i[j(c)]}^2 + o_{i[j(c)]}^2) \quad (9)$$

in which C = number of child distributions; $p_{j(c)}$ = probability of occurrence for child distribution c of factor j; and $o_{i[j(c)]}$ and $sd_{i[j(c)]}$ = mean and standard deviation, respectively, for child distribution c of factor j for activity i. Eqs. 8 and 9 are valid for any type of statistical distribution. Steiner's theorem can be directly applied to justify (9) (Kreyszig 1983). Note that the mean and variance of the combination of a base duration and parent distributions have been preserved for the grandparent distribution [(5) and (6)].

Properties of Child Distributions

The properties associated with the child distributions should be selected such that the mean and standard deviation of the parent distributions are maintained.

Number of Child Distributions. How many child distributions should be used to capture the impact of factor conditions that are different than expected? It seems reasonable to use an odd number of child distributions; the central child distribution will then be associated with the expected condition. A different number of child distributions may be appropriate for different factors. For the networks examined to date, it has been found that, given other constraints, the number of child distributions has relatively little impact on correlation (Wang 1996).

Probability of Occurrence. The probability that each condition actually occurs may be different. Fig. 4 shows two of the many possible combinations. On the left, the expected condition has been assigned a higher probability than either of the extreme conditions. On the right, each condition has been assigned the same value.

Mean. The mean of the child distribution for a given condition is the expected deviation from the mean of the parent distribution when the activity is performed under the given condition. Means of child distributions are expressed through a variable x, the mean placement. Fig. 5 shows the means of three child distributions as represented by ($-x$, 0, and x) or ($-0.5x$, $-x$, and $2x$). The mean of each child distribution should be confined to a range that maintains the variance of the parent distribution. Consider a family of three child distributions, as shown in Fig. 6. As the mean placement x approaches the limit, the standard deviations of child distributions must become smaller if the variance of the parent is to be preserved. When x is equal to the limit, the child distributions will have zero standard deviations.

Standard Deviation. The standard deviation of each child distribution is determined as a multiple of sd. For example, the standard deviation of a child distribution can be equal to sd or $2sd$. Specifying the mean placement x and requiring the variance of the combination of child distributions to equal the variance of the parent yields the value of sd.

Constructing Child Distributions

To construct a family of child distributions is to determine their means and standard deviations. Consider a parent distribution that is sensitive to factor j and has a variance of 4 days. Assume that the user chooses the categories of better-than-expected, expected, and worse-than-expected conditions to describe the conditions of the factor. Then a family of three child distributions should be constructed. Assume that the probabilities of occurrence for the child distributions are equal; that is, $p_1 = p_2 = p_3 = 1/3$. Thus, based on (8) and (9), the mean and variance, respectively, of the combined child distributions are

$$(1/3)o_1 + (1/3)o_2 + (1/3)o_3 = 0 \quad (10)$$

$$(1/3)(sd_1^2 + o_1^2) + (1/3)(sd_2^2 + o_2^2) + (1/3)(sd_3^2 + o_3^2) = 4 \quad (11)$$

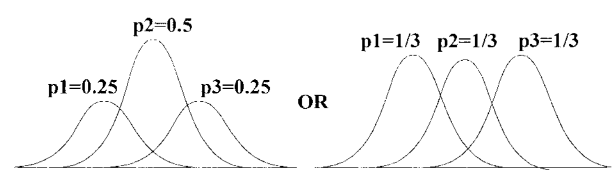

FIG. 4. Child Distributions for Different Probabilities of Occurrence

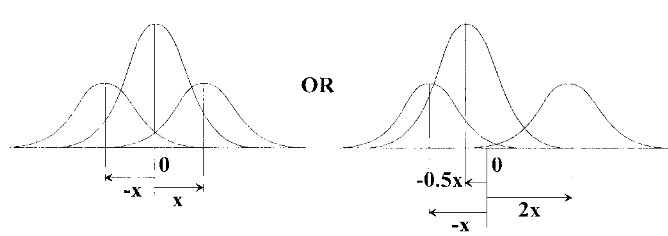

FIG. 5. Child Distributions for Different Mean Placements

J. Constr. Eng. Manage. 2000.126:458-466.

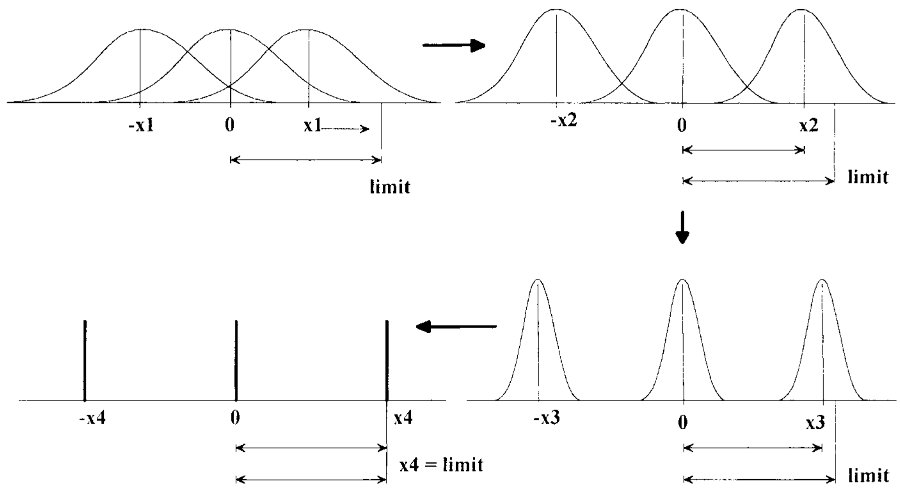

FIG. 6. Effect of Mean Placement

Assume $-o_1 = o_3 = x$ and $o_2 = 0$ so that (10) is satisfied, and let the child distributions have equal standard deviations, then (11) can be rewritten

$$sd^2 + (2/3)x^2 = 4 \quad (12)$$

The limit of the value of x is found by requiring that the variance of the child distribution be nonnegative. That is

$$sd^2 = 4 - (2/3)x^2 \geq 0 \quad (13)$$

Thus, the limit in this case is $x \leq \sqrt{6} = 2.45$ (limit = 2.45). In other words, the values of 2.45 and −2.45 are the two extreme means for Child Distributions 1 and 3, respectively. The next step is to select the value of x between 0 and 2.45. Instead of specifying the exact value of x, the NETCOR model suggests that the value of x be selected based on the level of influence of the factor under consideration on the activity under consideration. In this example, assume x is set to one-half of the limit. Then x is equal to 1.27. The properties of this family of three child distributions are thus Child 1 ($p_1 = 1/3$, $o_1 = -1.27$, and $sd_1 = 1.71$), Child 2 ($p_2 = 1/3$, $o_2 = 0$, and $sd_2 = 1.71$), and Child 3 ($p_3 = 1/3$, $o_3 = 1.27$, and $sd_3 = 1.71$). A comprehensive analysis of the input distribution (such as the analysis of the characteristics of the implicitly integrated parent and grandparent distributions) is included in Wang (1996).

Further Comments on Correlation

The occurrence of correlation between parent distributions is due to the sharing of the same factor(s). However, it may not be true that parent distributions that are sensitive to the same factor must be correlated. This is because the outcomes of any shared factor may not be the same over a long period of time. In other words, as recognized in earlier works [e.g., Carr (1979) and Padilla and Carr (1991)], correlation can be time dependent. Two outdoor activities that are each highly sensitive to weather may not be correlated if they are scheduled over different days or in different seasons. Other factors may require refinement as well. For example, site conditions may vary across a project. Thus two activities that are each highly sensitive to site conditions may have have correlated durations. Nevertheless, in a broad view, it is reasonable to attribute correlation to the sharing of factors.

Uncertainty Sensitivity along Path

In this paper, the uncertainty sensitivity for factor j along a path is measured using the coefficient of variation (CV). Mathematically, the value of CV along a path for factor j, denoted as CV_j, is given

$$CV_j = \sqrt{\text{variance}_j}/\text{mean} \quad (14a)$$

and the value of CV along a path for all factors, denoted as CV, is given

$$CV = \sqrt{\text{variance}_{\text{all}}}/\text{mean} \quad (14b)$$

In (14) mean = expected duration of a path; variance_j = variance of the path when only factor j is evaluated; and $\text{variance}_{\text{all}}$ = variance of the path when all factors are evaluated. A path with a high value of CV for a factor is considered to be highly sensitive to the factor. High sensitivity to a factor indicates that a path duration is more likely to be affected by this factor (if the factor condition is other than expected).

Expected Delay Penalty

Several indicators can be used to measure the performance of a project in terms of duration. Because one of the main reasons to finish a project on time is to prevent enforcement of a delay penalty, NETCOR uses a measure of the expected delay penalty to gauge the risk of a path or project. The expected delay penalty EDP is defined

$$EDP = \sum_{T_d=1}^{\infty} R_{T_d} \times T_d \times DP \quad (15)$$

in which T_d = number of days beyond the target duration; R_{T_d} = probability that the project duration is T_d days beyond the target; and DP = daily delay penalty.

Implementation of NETCOR

The implementation strategy for NETCOR's duration and correlation modeling procedure can be found in Wang (1996). A newly developed simulation language, STROBOSCOPE (STate and ResOurce Based Simulation of COnstruction ProcEsses) (Martinez 1996), is adopted to execute the simulation-relevant algorithms described in the NETCOR model. In addition to STROBOSCOPE's powerful capabilities to dynamically access the state of the simulation and properties of the resources involved in construction operations, it has an add-on that allows the definition of CPM networks with stochastic durations and calculation of various statistics about the project and activities. STROBOSCOPE can be run under the environment of the 32-bit version of Windows (e.g., Windows 95) or 16-bit version of Windows (e.g., Windows 3.1 plus the Win32s operating system extension).

J. Constr. Eng. Manage. 2000.126:458-466.

ANALYSIS OF CORRELATION EFFECT

This section demonstrates the use of NETCOR. A sample project is used to begin to investigate the correlation effect on a project's duration, an analysis of factor sensitivities along paths is conducted, and the performance of NETCOR, PERT, and Monte Carlo-simulation models are compared. The main objective of these exercises is to examine whether NETCOR works as expected.

Sample Network

Fig. 7 presents an example nine-activity project network. The network consists of three paths. The first path (1 → 2 → 3) is the conventional CPM with no float available. The second (4 → 5 → 6) and third paths (7 → 8 → 9) have 1 and 4 days of float, respectively. In NETCOR, two types of input are required: information that can be used to estimate the mean and standard deviation of each activity and a qualitative estimate of the sensitivity of each activity to each factor.

Duration Mean and Standard Deviation

For this example, the mean and standard deviation for each activity were derived from the entered values for the optimistic duration l, mode t, and pessimistic duration u of each activity, as shown on the left side of Table 1. These values were used to calculate the mean and standard deviation σ of the grandparent distribution for each activity by assuming a beta distribution with shape parameters α and β. This is the same procedure used to find mean and standard deviation in a traditional PERT analysis.

Qualitative Estimates

The factors considered in this example project are labor skills, weather, and equipment. The right side of Table 1 shows the assumed sensitivity of each activity to each factor. For example, Activity 1 has a low sensitivity to labor and medium sensitivities to weather and equipment.

Evaluation

Based on the inputs presented in Table 1, NETCOR then uses a family of child distributions to represent the grandparent distribution for each activity. This is done in two steps: (1) Calculate the standard deviation of the parent distribution based on a scale system; and (2) calculate the mean and standard deviation for each child distribution. During this two-step process, the mean and standard deviation of the grandparent distribution are always preserved. The scale system used is w_1[high] = w_2[high] = w_3[high] = 8, w_1[medium] = w_2[medium] = w_3[medium] = 5, and w_1[low] = w_2[low] = w_3[low] = 1.

Three child distributions are used for each family. The probabilities of occurrence and standard deviations of child distributions for each family are assumed to be the same; child distribution means are symmetrically placed. The value of x is set to 0.7, 0.5, or 0.3 limit when the family has a high, medium, or low sensitivity, respectively. Then, based on the calculation procedures described previously, the mean and standard deviation of child distributions with respect to labor, weather, and equipment can be obtained.

Evaluation of NETCOR's results is complicated by the fact that the true impact of correlation is unknown. In the discussion presented below, NETCOR's results for the sample network are compared with four analyses that do not take correlation into account: standard PERT analysis, Monte Carlo simulation carried out using normally distributed activity durations with the same mean and variance used in NETCOR's grandparent distribution (Normal Grand), Monte Carlo simulation carried out using beta-distributed activity durations with the same mean and variance used in NETCOR's grandparent distribution (Beta Grand), and Monte Carlo simulation carried out directly on NETCOR's child distributions (Normal Child). The PERT and Normal Grand analyses are included because these are the most commonly used means of incorporating uncertainty. It has frequently been suggested that construction activity duration distributions can be best modeled using a beta distribution. However, because NETCOR models activity duration as a sum of durations due to several factors, the resulting grandparent duration is, by the central-limit theorem, normally distributed [see Wang (1996) for a further discussion]. The Beta Grand analysis is included to see whether this limitation is important for the example network. The Normal Child analysis is included to further isolate the effect of correlation as captured by NETCOR. Child distributions identical to those

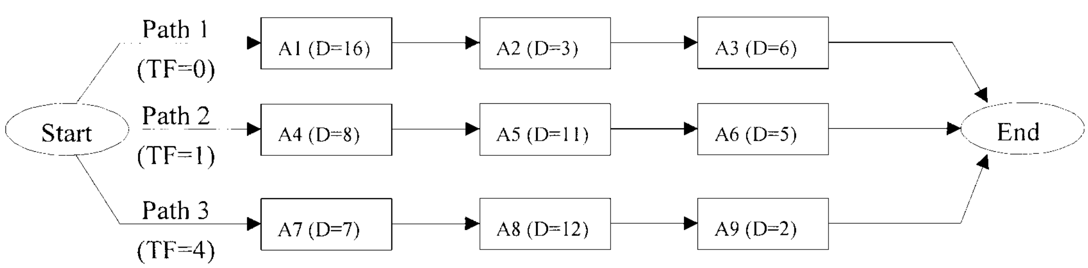

FIG. 7. Example Project Network

TABLE 1. Three-Point Estimates and Uncertainty Sensitivity for Example Network

Activity (1)	l (2)	t (3)	u (4)	α (5)	β (6)	Mean (7)	σ (8)	Labor (9)	Weather (10)	Equipment (11)
A1	13	15	25	2	6	16	1.73	L	M	M
A2	1	2	7	1.3	2.7	3	1.26	H	L	M
A3	3	4	10	1.1	1.4	6	1.85	M	M	L
A4	3	5	20	1.3	3.1	8	3.35	M	H	H
A5	7	8	18	1.1	1.9	11	2.65	M	M	H
A6	3	4	8	1.2	1.8	5	1.22	H	M	L
A7	4	5	18	1.3	4.7	7	2.17	M	H	H
A8	4	6	24	1.1	1.6	12	5.12	H	M	L
A9	1	1.5	8	1.7	10	2	0.68	L	H	M

Note: l = optimistic duration, t = mode, u = pessimistic duration, σ = standard deviation, α and β = shape parameters of beta distribution, H = high, M = medium, and L = low.

J. Constr. Eng. Manage. 2000.126:458-466.

TABLE 2. Comparison of Without- and With-Correlation Analyses

Project duration (1)	PERT (2)	Normal Child (3)	Beta Grand (4)	Normal Grand (5)	With-correlation (6)
Mean	25	27.22	26.79	27.14	27.02
Standard deviation	2.83	3.08	3.15	2.93	4.20
EDP	1.12*DP*	2.88*DP*	2.66*DP*	2.89*DP*	3.19*DP*
90% confidence intervals	—	—	—	22.6–32.2 (Δ = 9.6)	20.6–34.2 (Δ = 13.6)
Probability of overrun	0.43	0.76	0.69	0.76	0.66

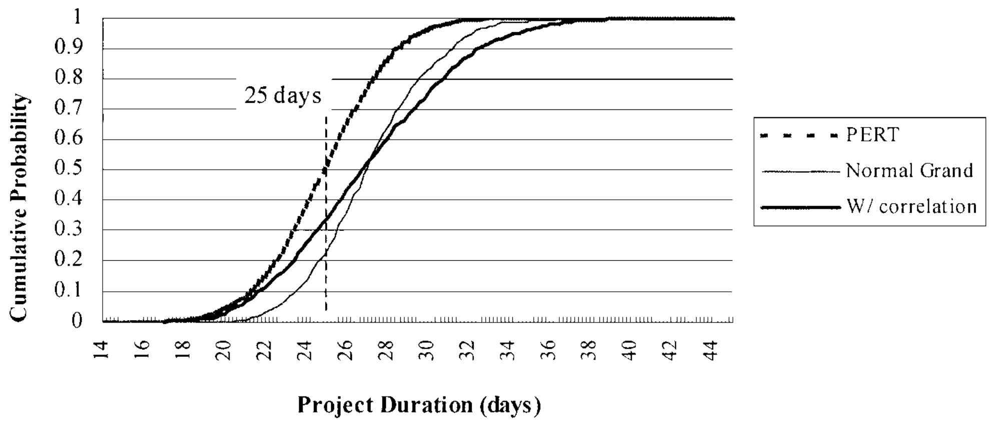

FIG. 8. Project Duration for With- and Without-Correlation Analyses

used in NETCOR form the basis of this simulation; the only difference between the Normal Child analysis and NETCOR analyses is the introduction of correlation.

Results—Project Duration

In the discussion below, project duration for the various analyses (PERT, Normal Grand, Beta Grand, Normal Child, and with correlation) is compared using several metrics: the mean and standard deviation of the overall duration distribution, expected delay penalty *EDP*, range of 90% confidence intervals for project duration, and probability of schedule overrun (the probability that the project duration exceeds the CPM duration).

PERT versus Without-Correlation Analyses

The left side of Table 2 summarizes the results of project duration for the four without-correlation analyses. As expected, the Normal Grand and Normal Child distributions generate similar results in every category. The Beta Grand distribution generates slightly different results than the other two simulation analyses. Compared with the PERT analysis, each of the without-simulation analyses results in increases in the expected project duration of about 2 days (8%), standard deviation of about 0.22 days (8%), *EDP* of about 1.69*DP* (150%), and probability of overrun of about 0.31 (62%). PERT underestimates variation because it ignores the uncertainty associated with other noncritical paths (Moder et al. 1983).

Without-Correlation versus With-Correlation Analyses

Fig. 8 graphically summarizes the cumulative probability distribution of the project's duration for PERT, without-correlation Normal Grand analysis, and with-correlation analysis. The effect of correlation can be clearly observed from the figure; the with-correlation distribution has extended the tails of the distribution.

The right side of Table 2 shows that the with-correlation analysis results in a small decrease in the project's expected duration. With respect to standard deviation and *EDP*, the values under with-correlation analysis are 4.20 days and 3.19*DP*, respectively. These represent 43 and 10% increases as compared with the Normal Grand analysis. As expected, the 90% confidence intervals for the with-correlation analysis are larger than those of the without-correlation analysis. The probability of schedule overrun depends on the target date. In Fig. 8, the without-correlation and with-correlation distributions intersect each other when the project's duration is about 27 days. If the project's completion date is set to 27 days, then both analyses have the same probability of schedule overrun. Detailed discussion of the effect of various scale systems on project duration can be found in Wang (1996).

TABLE 3. Uncertainty Sensitivity by Path

Path (1)	Float (days) (2)	Labor (3)	Weather (4)	Equipment (5)
1	0	0.080	0.077	0.065
2	1	0.122	0.133	0.143
3	4	0.205	0.191	0.108

Results—Uncertainty Sensitivity along Path

In a network, the uncertainty sensitivity (measured by *CV*) along a path should be examined from two perspectives: between paths (Which path is most sensitive to a particular factor?) and along a path (To which factor is a given path most sensitive?). Table 3 presents the uncertainty sensitivity with respect to each factor for each path. Path 3 (4 days of float) is the path that is most sensitive to labor, weather, and all factors combined. Path 2 (1 day of float) is the path that is most sensitive to equipment. Although Path 1 is the critical path, it is the least sensitive path in every comparison. Table 3 also shows that the factor that has the greatest influence on Paths 1 and 3 is labor and equipment has the greatest influence on Path 2.

With knowledge of both the uncertainty sensitivity along a path from a factor and the amount of float associated with that path, management can focus attention on two types of paths, critical and near critical, that are highly sensitive to factors. Because there is no float time available for critical paths to react to changes and even small variations in the duration of

J. Constr. Eng. Manage. 2000.126:458-466.

critical path directly affect project duration, critical paths should receive particular attention. Near-critical paths with a large variance that is caused primarily by a single factor warrant special attention, because poorer than expected conditions for that factor can lead to a large increase in the path duration.

CONCLUSIONS

This paper has presented a simulation-based model, NETCOR, that allows correlation between activity durations to be considered in network analysis. The incorporate of correlation in network analysis should lead to improved estimates of project duration. The strengths of NETCOR arise from two sources. First, the use of qualitative estimates (e.g., high, medium, and low) to describe the effect of factor-based uncertainty should allow the user to be more comfortable in providing input than would be the case with other approaches (particularly direct-elicitation approaches that require correlation coefficients as input). Second, in NETCOR, correlation is evaluated based on a grandparent-parent-child structure that systematically breaks down the effects of uncertainty by factor and condition. This systematic structure eases the evaluation of the effects of individual factors on path and project duration, making it possible to identify the factors that have the greatest impact on a project. NETCOR relies on a variety of assumptions and simplifications to allow the use of the grandparent-parent-child structure. For example, it is assumed that there is no interaction between factor conditions in NETCOR; thus, the results generated by NETCOR are only approximate.

A small example network is presented here to demonstrate NETCOR. In a companion paper (Wang and Demsetz 2000), NETCOR is applied to a recent construction project. Future work on NETCOR may include exploring ways to capture nonnormal grandparent distributions [see Wang (1996) for a detailed discussion]; implementing time-dependent and non-time-correlated duration variables; collecting field data to justify child distributions and parent distributions, and values of mean placement x, and correlation coefficient; investigating applications of the path approach (e.g., managing subcontractors or work packages from the viewpoint of a path); applying NETCOR to additional construction projects; and applying NETCOR to other areas. For example, NETCOR could be used to investigate correlation in cost. NETCOR also could be used in resource allocation, with resources assigned first to activities that are highly sensitive to particular factors.

ACKNOWLEDGMENTS

The writers thank the reviewers of this paper for their careful evaluation and thoughtful comments. The writers are indebted to Julio Martinez for making STROBOSCOPE available and generously providing assistance in its use.

APPENDIX. REFERENCES

Ahuja, H. N., and Nandakumar, V. (1985). ''Simulation model to forecast project completion time.'' *J. Constr. Engrg. Mgmt.*, ASCE, 111(4), 325–342.
Benjamin, J., and Cornell, A. (1970). *Probability, statistics and decision for civil engineers*, McGraw-Hill, New York.
Carr, R. I. (1979). ''Simulation of construction project duration.'' *J. Constr. Div.*, ASCE, 105(2), 117–128.
Chang, T. C. (1987). ''Network resource allocation using an expert system with fuzzy logic reasoning.'' PhD dissertation, University of California, Berkeley, Calif.
Chau, K. W. (1995). ''Monte Carlo simulation of construction costs using subjective data.'' *Constr. Mgmt. and Economics*, 13, 369–383.
Kreyszig, E. (1983). *Advanced engineering mathematics*, 5th Ed., Wiley, New York.
Law, A. M., and Kelton, W. D. (1991). *Simulation modeling and analysis*, 2nd Ed., McGraw-Hill, New York.
Levitt, R. E., and Kunz, J. C. (1985). ''Using knowledge of construction and project management for automated schedule updating.'' *Proj. Mgmt. J.*, 16(5), 57–76.
Martinez, J. C. (1996). ''STROBOSCOPE: State and resource based simulation of construction processes.'' PhD dissertation, University of Michigan, Ann Arbor, Mich.
Moder, J. J., Philips, C. R., and Davis, E. W. (1983). *Project management with CPM, PERT and precedence diagramming*, 3rd Ed., Van Nostrand Reinhold, New York.
Padilla, E. M., and Carr, R. I. (1991). ''Resource strategies for dynamic project management.'' *J. Constr. Engrg. and Mgmt.*, ASCE, 117(2), 279–293.
Ranasinghe, M., and Russell, A. D. (1992). ''Treatment of correlation for risk analysis of engineering projects.'' *Civ. Engrg. Syst.*, 9, 17–39.
Thomas, H. R., Maloney, W. F., Horner, R. M. W., Smith, G. R., Handa, V. K., and Sanders, S. R. (1990). ''Modeling construction labor productivity.'' *J. Constr. Engrg. and Mgmt.*, ASCE, 116(4), 750–726.
Touran, A. (1993). ''Probabilistic cost estimating with subjective correlations.'' *J. Constr. Engrg. and Mgmt.*, ASCE, 119(1), 58–71
Touran, A., and Wiser, E. D. (1992). ''Monte Carlo technique with correlated random variables.'' *J. Constr. Engrg. and Mgmt.*, ASCE, 118(2), 258–272.
Van Tetterode, L. M. (1971). ''Risk analysis . . . or Russian roulette?'' *Trans. Am. Assn. of Cost Engrs.*, 124–129.
Wang, W.-C. (1996). ''Model for evaluating networks under correlated uncertainty—NETCOR.'' PhD dissertation, University of California, Berkeley, Calif.
Wang, W.-C., and Demsetz, L. A. (2000). ''Application example for evaluating networks considering correlation.'' *J. Constr. Engrg. and Mgmt.*, ASCE, 126(6), 467–474.
Woolery, J. C., and Crandall, K. C. (1983). ''Stochastic network model for planning scheduling.'' *J. Constr. Engrg. and Mgmt.*, ASCE, 109(3), 342–354.

J. Constr. Eng. Manage. 2000.126:458-466.