Multiple Simulation Analysis for Probabilistic Cost and Schedule Integration

Leroy J. Isidore[1] and W. Edward Back[2]

Abstract: One of the major goals of the construction industry today is the quantification and minimization of the risk associated with construction engineering performance. When specifically considering the planning of construction projects, one way to control risk is through the development of reliable project cost estimates and schedules. Two techniques available for achieving this goal are range estimating and probabilistic scheduling. This paper looks at the integration of these techniques as a means of further controlling the risk inherent in the undertaking of construction projects. Least-squares linear regression is first considered as a means of relating the data obtained from the application of these techniques. However, because of various limitations, the application of linear regression was not considered the most appropriate means of relating the results of range estimating and probabilistic scheduling. Integration of these techniques was, therefore, achieved through the development of a new procedure called the multiple simulation analysis technique. This new procedure combines the results of range estimating and probabilistic scheduling in order to quantify the relationship existing between them. Having the ability to accurately quantify this relationship enables the selection of high percentile level values for the project cost estimate and schedule simultaneously.

DOI: 10.1061/(ASCE)0733-9364(2002)128:3(211)

CE Database keywords: Scheduling; Costs; Construction industry; Risk management; Simulation.

Introduction

Risk management, as it relates to construction projects, is vital to the successful undertaking and completion of any construction process. One way to effectively manage project risk is to develop more reliable means of accounting for the time and cost variability existing in construction operations. Since such operations typically involve the coordination of resources (i.e., labor, materials, and equipment) to achieve a desired cost, schedule, quality, and safety objective, sources of variability can be primarily attributed to such things as labor productivity, regional wage rates, and the availability and cost of materials and equipment. In addition to the coordination of project resources, additional variability in construction processes is due to the fact that they are performed in an environment that is subject to the unique conditions specific to that environment, e.g., the project location, topography, geology, and climate.

Two established tools that are commonly employed in the planning process to manage the level of risk associated with undertaking construction projects are the project cost estimate (CE) and the project schedule (PS). It is, therefore, important to develop a methodology to accurately predict the construction cost and schedule of a project while, at the same time, recognizing the inherent complexity, and variability, within all construction processes. Without a probabilistic technique for estimating and scheduling construction projects, it is not possible to account for this variability or determine the percentile level (PL) associated with a chosen cost estimate and schedule. Hence, it would not be possible to reliably determine the project cost and schedule, with a sufficient degree of certainty, to confidently minimize risk.

Recent attempts to more reliably quantify the risk inherent in construction projects have focused on range estimating and stochastic scheduling (also referred to as probabilistic estimating and probabilistic scheduling, respectively). These typically involve modeling the duration and cost of the activities that make up construction projects as stochastic quantities as opposed to fixed, deterministic values. A simulation technique, e.g., Monte Carlo simulation, is then used to generate a range of possible cost and schedule values through which the project can be expected to reasonably vary.

While it is widely assumed that the estimate and schedule of construction projects are related, presently these two planning tools (i.e., a range estimate and stochastic schedule) are often applied to construction projects independently of each other. This, therefore, makes it difficult to determine how the estimate and schedule for a specific project are related. Additionally, the independent application of these planning tools does not make it possible to understand the impact that selecting a high confidence level cost-estimate value has on the schedule value selected to execute a project, and vice versa. Finding a way to analytically relate the estimate and schedule, in order to exploit the information they provide, would supply additional information to the project planning process, and significantly aid in the minimization of the risk involved in undertaking construction projects.

This paper, therefore, presents an analytical procedure that enables decision makers to combine the techniques of range estimating and stochastic scheduling, such that the interrelationships between them can be understood and effectively used to enhance

[1]Visiting Assistant Professor, Dept. of Civil Engineering, Texas A&M Univ., College Station, TX 77843-3136.

[2] Assistant Professor, Dept. of Civil Engineering, Clemson Univ., Clemson, SC 29634-0911.

Note. Discussion open until November 1, 2002. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on May 23, 2000; approved on August 25, 2000. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 128, No. 3, June 1, 2002. ©ASCE, ISSN 0733-9364/2002/3-211–219/$8.00+$.50 per page.

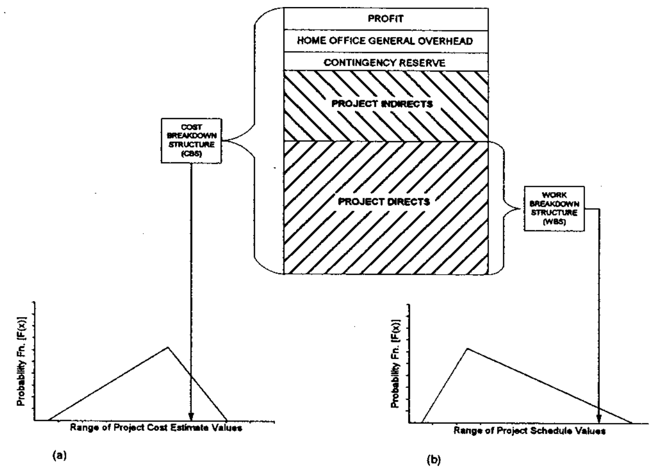

Fig. 1. Data generation approach for traditional nonintegrated range estimating and probabilistic scheduling. Triangular representation of: (a) project's estimate and its variability; (b) project's schedule and its variability.

the decision-making process. This will involve combining probabilistic scheduling and range estimating techniques to reliably quantify their complex interactions.

Definition of Cost and Schedule Integration

The phrase "cost-schedule integration," as used throughout this paper, refers to the preestablished planning structure used to provide the information needed for a specific project in order to develop a probabilistic control estimate and schedule. The control estimate, and schedule, represents the baseline against which actual performance will be measured throughout the execution phase of the project. This integrated system will also be considered for its potential to provide additional information to the planning process, and, ultimately, its ability to minimize the risk associated with the undertaking of construction projects.

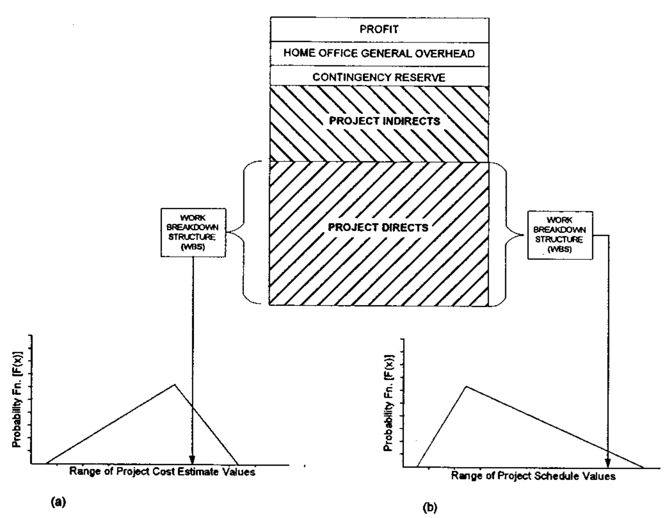

Fig. 2. Data generation approach for integrated range estimating and probabilistic scheduling. Triangular representation of: (a) project's estimate and its variability; (b) project's schedule and its variability.

Table 1. Abridged Listing of Simulation Output

Random run number	Project schedule (days)	Cost estimate (dollars $\times 10^5$)
366	751	1,583
367	742	1,546
368	767	1,639
369	736	1,651
370	714	1,513
371	747	1,684
[Average]	740	1,618
[Standard deviation]	17.63	47.98

Range Estimating and Probabilistic Scheduling

Range estimating and probabilistic scheduling are estimating and scheduling techniques used to generate the cost estimate and schedule of a project as probability distributions. They involve defining the activity cost and durations as probability distributions rather than fixed deterministic quantities. Once distributions for the activity cost and duration have been defined, a Monte Carlo simulation algorithm (using the generation of random numbers) is applied to allow random sampling of these distributions. After a random cost and duration value, per activity, has been generated, all of the activity cost values and all of the activity duration values, respectively, are added together to determine the overall cost and schedule of the project. This process, when repeated a large number of times, results in a probability distribution for the total project cost and another for the total project schedule. Hence, having the total project cost and schedule represented as probability distributions allows for easy selection of a project cost estimate and schedule having a low probability of being exceeded.

While it is evident that the estimate and schedule of construction projects are somehow related, range estimating and probabilistic scheduling are often *separately and independently* applied to develop the cost estimate and schedule of construction projects. This is illustrated in Fig. 1, which shows that range estimating is frequently carried out at the cost breakdown structure level, while probabilistic scheduling is most effectively completed at the work breakdown structure (WBS) level.

In such cases, it is not possible to account for the effect that selecting a particular project schedule, at a stated percentile level, has on the construction cost of a project, or vice versa. Hence, decisions may be made without any consideration being given to the true relationship that exists between a project's cost estimate and schedule. These decisions, while more reliable than those made using traditional deterministic methods, still fall short of the

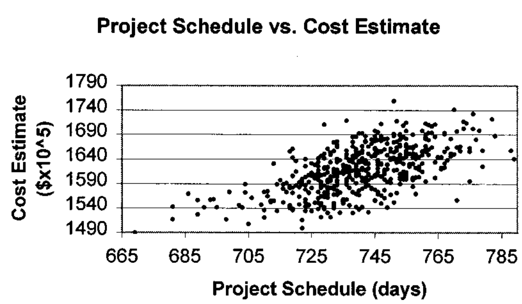

Fig. 3. Scatter diagram of project schedule and cost estimate data from ABC-Sim

J. Constr. Eng. Manage., 2002, 128(3): 211-219

goal of truly capturing the interdependency of the cost and schedule associated with construction projects.

Activity based Costing Simulation (ABC-Sim)

A relatively new approach, based on discrete event simulation, has recently been implemented to simultaneously produce the probability distributions for the project cost estimate and schedule. This tool is called ABC-Sim, and was developed as part of a research study at Texas A&M Univ.

ABC-Sim has the added advantage over traditional stand-alone range estimating and probabilistic scheduling applications of being able to simultaneously perform range estimating and probabilistic scheduling for an appropriately modeled construction project. This is illustrated graphically in Fig. 2, which shows that ABC-Sim performs both range estimating and probabilistic scheduling at the WBS level for a construction project.

Having the ability to combine these tools has two primary advantages. First, it results in probability distributions for both the project cost estimate and the project schedule values being simultaneously generated. Second, for each iteration of the simulation process, a project schedule value, and its corresponding cost-estimate value, is produced.

To make use of the information resulting from the simultaneous execution of range estimating and probabilistic scheduling for construction projects, it is necessary to find some adequate analytical means of relating their respective data sets. This will make it possible to determine the relationship that exists between a preferred project schedule and its associated cost estimate, thereby relating a selected schedule to the appropriate cost estimate of that project. Once an adequate means of relating these data sets can be obtained, it will be possible to use this information in the planning and control of construction projects in order to minimize the risk involved with the undertaking of these projects.

Traditional Analysis of Stochastic Cost Estimate and Schedule

A construction project for which suitable stochastic data were available was selected as the test case in order to demonstrate the differences between the application of the traditional probabilistic estimating and scheduling technique and the new technique presented in this paper. This project was simulated using 500 iterations. A partial list of the simulation output data, along with summary statistics, is presented in Table 1. A scatter diagram of the paired cost estimate and project schedule values is shown in Fig. 3, and represents the 500 possible combinations of the cost estimate and project schedule values generated using ABC-Sim. Once the simulation run was completed, the data generated were first summarized in the traditional way in order to enable the selection of the cost estimate and project schedule values that both had a high probability of not being exceeded.

From these data, the project schedule and cost estimate values at the 80th, 85th, 90th, and 95th percentiles were identified, and are shown in columns 2 and 5 of Table 2. These represent some of the typical high percentile values that would likely be considered in a traditional range estimating and/or probabilistic scheduling exercise. Once these values are identified, the selection of an appropriate value to represent the project cost estimate and schedule would be made based on the level of risk, as determined by the characteristics of the project, the amount of additional project-specific information available, experience, and the engineering judgment of the project planners.

Table 2. Cost Estimate and Schedule Percentiles for High Percentile Schedule and Cost Estimate Values, Respectively

Percentile level (%)	Project schedule (days)	Corresponding cost estimate (dollars$\times 10^5$)	Cost estimate percentile level (%)	Cost estimate (dollars$\times 10^5$)	Corresponding project schedule (days)	Project schedule percentile level (%)
80	755	1,613	47	1,658	788	98
85	758	1,570	18	1,667	741	52
90	762	1,649	73	1,682	745	61
95	769	1,646	71	1,694	763	91

Table 3. 95% Prediction Interval for Predicted Cost Estimate Values

Project schedule percentile ranking	Project schedule (days)	CE predicted value (dollars$\times 10^5$)	CE lower 95% predicted value (dollars$\times 10^5$)	CE upper 95% predicted value (dollars$\times 10^5$)
80th	755	1,642	1,568	1,717
85th	758	1,647	1,573	1,722
90th	762	1,654	1,580	1,729
95th	769	1,666	1,591	1,740

Table 4. 95% Prediction Interval for Predicted Project Schedule Values

Cost estimate percentile ranking	Cost estimate (dollars$\times 10^5$)	PS predicted value (days)	PS lower 95% predicted value (days)	PS upper 95% predicted value (days)
80th	1,658	750	722	777
85th	1,667	752	724	779
90th	1,682	755	728	782
95th	1,694	758	730	785

J. Constr. Eng. Manage., 2002, 128(3): 211-219

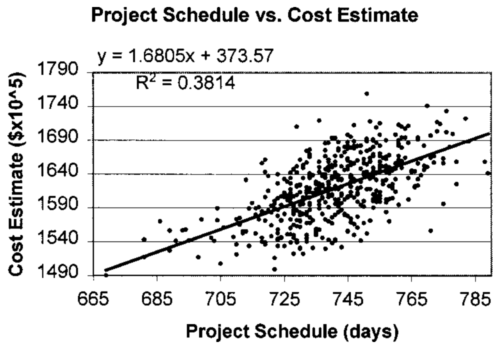

Fig. 4. Linear regression of project schedule and cost estimate

One major limitation of a nonintegrated approach can be demonstrated by choosing the 95th percentile level value for both the project schedule (i.e., 769 days) and the cost estimate (i.e., $\$1{,}694\times10^5$). While it is possible that this combination of project schedule and cost estimate values could occur for this project given a very specific set of conditions, it is not reasonable to expect them to actually occur simultaneously during project execution, particularly for projects that may be either cost driven or schedule driven. Additionally, the variability in the activity durations and the execution logic of the process flow diagram may not permit such a combination to occur. This reasoning can also be extended to any other construction project where similar nonintegrated range estimating and probabilistic scheduling exercises are undertaken.

Table 2 lists the high confidence project schedule values with their actual, simulation-generated cost estimate values. Note that these values vary considerably. If, for example, the 95% confidence level of the project schedule is chosen (i.e., 769 days), the actual observed corresponding cost estimate value of $\$1{,}646\times10^5$ only provides a confidence level of 71%, with respect to all cost values.

In a similar manner, if the high confidence cost estimate values are listed with their simulation-generated project schedule values, as shown in Table 2, they do not necessarily correspond to project schedule values with high confidence levels. In other words, choosing a schedule value having a high confidence level (i.e., low likelihood of being exceeded) does not guarantee that the associated cost estimate, corresponding to the selected schedule value, will also have a comparably high level of confidence.

The objective is to find an analytical way to relate the cost estimate and project schedule data such that values having a low probability of being exceeded can be selected for both of these tools. Being able to relate values for the cost and schedule, in a logical and systematic way, will eliminate the practice of arbitrarily choosing a high confidence level cost estimate and project schedule value during the planning process. This will also allow the actual decision to be based on the possible combinations of the cost estimate and project schedule data, the ranges of the true probability distributions for these planning tools as determined by Monte Carlo simulation, the activity data ranges, and the execution sequence of those activities. Least-squares linear regression was first considered as a possible means of relating these values.

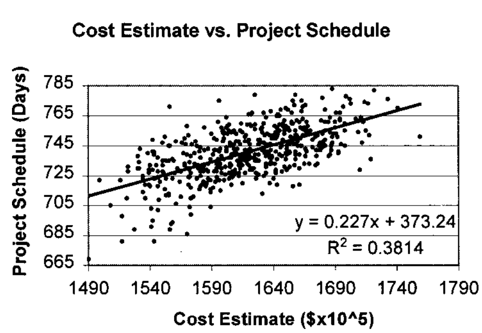

Fig. 5. Linear regression of cost estimate and project schedule

Table 5. Abridged List of 95th Percentile Level Project Schedule Data Obtained from 25 Project Simulations

Simulation number	Project schedule (days)	Cost estimate (dollars$\times10^5$)	Range estimates percentile	Conditional cost estimate percentile
1	769	1,683	0.92	0.69
2	766	1,672	0.88	0.59
12	775	1,606	0.42	0.08
13	767	1,670	0.87	0.57
24	765	1,633	0.64	0.23
25	764	1,629	0.60	0.20
[Average]	768.84	1,662.88	0.77	0.50
[Standard deviation]	3.73	40.16	0.21	0.32

Least-Squares Linear Regression Application

Fig. 4 is a scatter diagram of the probabilistic schedule values for the example construction project plotted against their corresponding range estimate values. Since the data indicated a linear trend, they were fitted with a linear regression equation using the PS values as the independent variable and the CE values as the dependent variable.

Based on least-squares minimization, the regression relationship for this data set is given by

$$CE_i = 1.6805PS_i + 373.57 \quad (1)$$

The coefficient of determination, R^2, is equal to 0.3814. This means that based on a linear regression relationship, the project schedule explains approximately 38% of the variability (or scatter) observed in the cost estimate data. Also, based on the scatter observed in the data and this low R^2 value, it can be concluded that the scatter in the data points, and the presence of numerous outliers, has a considerable influence on the least-squares minimization estimates of the model parameters.

In addition to relating the project schedule to the cost estimate, there are some situations where it is equally important to consider relating the cost estimate to the project schedule. Considering such a situation, a linear regression relationship was fitted with the cost estimate data as the explanatory variable and the project schedule data as the dependent variable. This resulted in a new linear regression relationship, which was different from that obtained previously [Eq. (2)]. The linear regression relationship is shown graphically on the scatter diagram of the cost estimate versus the project schedule data shown in Fig. 5

$$PS_i = 0.227CE_i + 373.24 \quad (2)$$

Summary of Linear Regression Application to Simulated Data

The regression equations [Eqs. (1) and (2)] were used to predict values for the project schedule and cost estimate at the 80th, 85th,

J. Constr. Eng. Manage., 2002, 128(3): 211-219

Table 6. Comparison of Cost Estimate Means for Example Construction Project

Percentile ranking	Conditional cost estimate: Cost Estimate (dollars$\times 10^5$)	Conditional cost estimate: Standard Deviation (σ)	Range estimate: Cost Estimate (dollars$\times 10^5$)	Range estimate: Standard Deviation (σ)	Range estimate: Statistically Significant Difference
80th	1,701	1.104	1,656	2.746	Yes
85th	1,707	1.197	1,669	2.325	Yes
90th	1,713	1.297	1,683	2.057	Yes
95th	1,719	1.401	1,699	2.353	Yes

90th, and 95th percentile rankings. These results are shown in Tables 3 and 4, respectively. Since linear regression estimates are not very useful without some idea of their accuracy, prediction intervals were developed for the cost estimate and project schedule values predicted from Eqs. (1) and (2) at the 80th, 85th, 90th, and 95th percentile values of the project schedule. These values are also shown in Tables 3 and 4.

These intervals in Table 3 can be interpreted as follows: ''With 95% confidence, the average cost estimate values that were predicted, based on the respective project schedule, lie somewhere between the lower 95% prediction value and the upper 95% prediction value.'' A similar interpretation for the project schedule data in Table 4 can be stated as follows: ''With 95% confidence, the average project schedule values that were predicted, based on the respective cost estimate, lie somewhere between the lower 95% prediction value and the upper 95% prediction value.'' Since these prediction intervals included most of the data observed from the simulation run, it was concluded that the excessive scatter observed in the simulation output data makes it difficult to obtain a satisfactory regression equation.

Based on the results observed from the application of linear regression, it was concluded that linear regression is not a suitable means of relating the stochastic data for the project schedule to those of the cost estimate. This is because the best-fit linear regression line is unable to adequately account for the variability in the output data that is produced by stochastic simulation modeling. Thus, fitting a regression line to such data makes it difficult to use high percentile project schedule values to adequately predict high percentile values for the cost estimate. For the same reason (i.e., the wide scatter in the data), it is difficult to use high percentile cost estimate values to adequately predict high percentile project schedule values. This conclusion is confirmed by analyzing the width of the prediction interval obtained for the estimated values. Additionally, based on the scatter of the data generally observed when Monte Carlo simulation is applied to construction project models, it can safely be concluded that the application of linear regression to a similar data set, for a different project, would have produced similar results.

Multiple-Simulation Analysis Technique (MSAT)

The MSAT is a new technique that was created as a means of relating the project cost estimate and schedule data obtained from simultaneously performing range estimating and probabilistic scheduling for a project. It was specifically developed to address the problem of relating the probabilistic cost estimate and project schedule data such that high percentile values selected for both of these tools were related in some meaningful way.

The technique combines discrete event simulation, regression, and numerical analysis in order to develop a model that explains the relationship between the stochastic cost estimate and schedule data. This allows much more detailed and integrated project planning than was possible in the past, when range estimating and probabilistic scheduling were independently applied to construction projects. The steps outlining this technique will now be presented.

Using the example project, the first step in this procedure requires fixing the percentile level of either the cost estimate or the project schedule. Fixing either the project schedule or the cost estimate at a predetermined percentile allows the analysis to proceed in a systematic way, such that the other can be determined. In this particular case, the percentile ranking of the project schedule was fixed at the 95% probability level (i.e., there was only a 5% chance that the observed value would be exceeded).

The next step involves simulating the project using ABC-Sim and sorting the resulting data based on the project schedule values. With respect to the example project, the 95th percentile value was then located and recorded along with its corresponding cost estimate. The data were then used to determine the percentile ranking of this cost estimate value, which was also recorded. The project was resimulated 25 times using 100 runs per simulation, and for each simulation, the 95th percentile project schedule value, its corresponding cost estimate value, and the percentile ranking of the cost estimate value were recorded. Once these data were recorded, the mean cost estimate and its standard deviation were then used to determine the conditional percentile ranking of the cost estimate values. A list of these data is presented in Table 5, along with the relevant summary statistics. Note that the results from each of the respective simulation runs are independent from each other.

Table 7. Abridged List of Modified 95th Percentile Cost Estimate Data

Simulation number	Cost estimate (dollars$\times 10^5$)	Project schedule (days)	Probabilistic schedule percentile	Conditional project schedule percentile
1	1,692	760	0.81	0.55
2	1,679	778	0.71	0.93
3	1,685	750	0.27	0.27
4	1,695	732	0.77	0.02
15	1,685	749	0.96	0.24
16	1,689	778	0.87	0.93
20	1,676	748	0.91	0.22
21	1,673	737	0.52	0.06
[Average]	1,689.92	758.32	0.72	0.50
[Standard deviation]	8.35	13.42	0.18	0.29

J. Constr. Eng. Manage., 2002, 128(3): 211-219

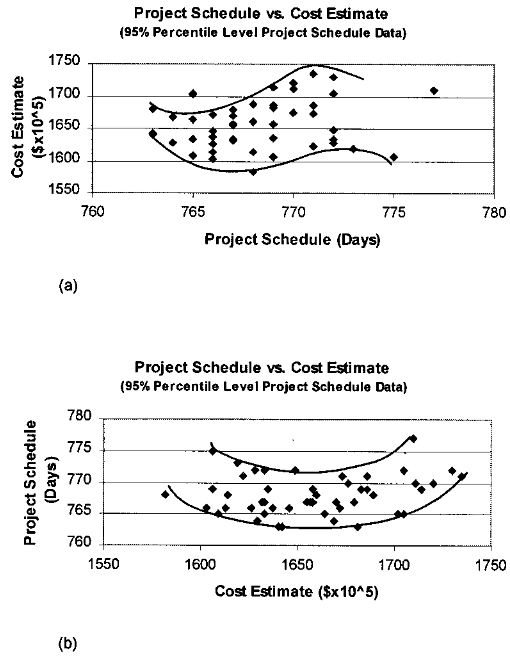

Fig. 6. Scatter diagrams based on cost estimate data generated at 95th percentile level of project schedule: (a) cost estimate values plotted against 95th percentile project schedule values; (b) 95th percentile project schedule values plotted against cost estimate values

From the data shown in Table 5, it was observed that the 95th percentile project schedule value was not a fixed value, but fluctuated about a mean value of 769 days. However, the standard deviation was only four days; thus, it may be concluded that this observed fluctuation was simply due to the randomness inherent in the simulation process.

Since the percentile level for the project schedule values had been preestablished at 95% and the observed values do not vary significantly about the mean value of 769 days, this value was used as the best estimate of the actual project schedule at the 95th percentile. Once the project schedule value was fixed in this way, the important issue became finding a way to select the cost estimate value that would also have a high percentile level and was related to this project schedule value in a meaningful way.

For the 95th percentile project schedule data, Table 5 shows that the cost estimate values fluctuated and ranged from $\$1{,}606\times 10^5$ with a percentile ranking of 42% to $\$1{,}730\times 10^5$ with a percentile ranking of 99%. Since these observations were based on a representative sample, they were not expected to contain the absolute minimum and maximum values.

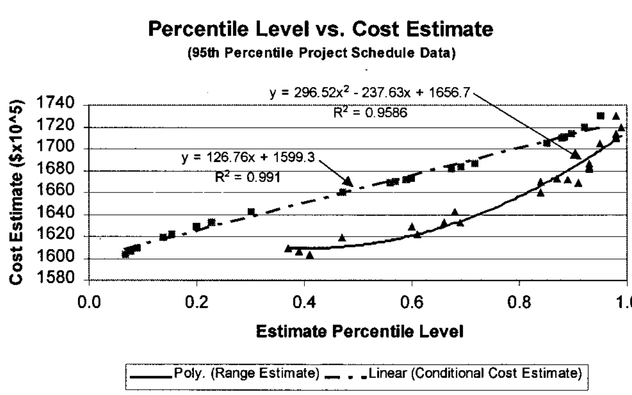

Fig. 7. Percentile levels versus cost estimate

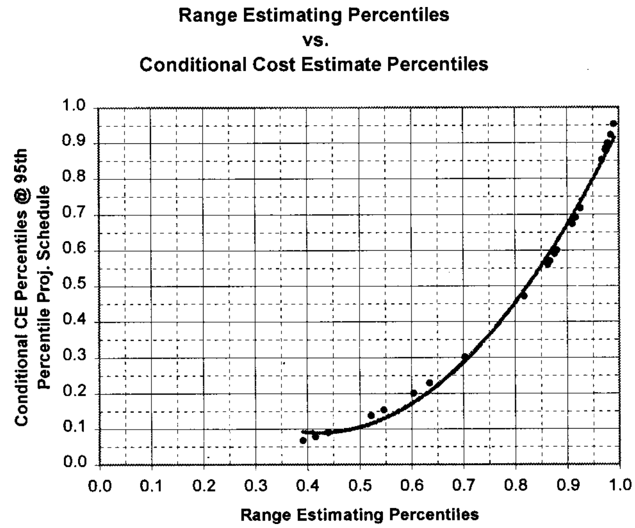

Fig. 8. Comparison of percentiles between range estimating percentiles and conditional cost estimating percentiles

95th Percentile Project Schedule Data

One obvious way of relating the 95th percentile project schedule data with the resulting cost estimate data is in a manner similar to the technique used previously—i.e., relating the project schedule and cost estimate values directly. Also, as was previously discussed, while finding a way to relate the project schedule to the cost estimate is a separate problem from that of relating the cost estimate to the project schedule, both problems were considered to be of equal importance, and will be considered simultaneously in this analysis.

Relating the 95th percentile project schedule data presented in Table 5 resulted in the scatter diagrams shown in Fig. 6. In Fig. 6(a), the 95th percentile project schedule values are plotted against their corresponding cost estimate values; Fig. 6(b) shows the same data with the axis reversed. From this figure, it is possible to see that there was a distinct nonlinear pattern present in the data. This observed pattern in the data indicates that the project cost estimate and schedule values are related when the data are examined at a specific percentile level—in this particular case, the 95th percentile. Finding some means of extracting this relationship would be a key step in relating the project cost estimate and schedule data.

Predicting Cost Estimate using Range Estimate Percentiles

As was previously stated, the cost estimate values were preselected at the 95th percentile project schedule, and based on this percentile level, the data in Table 5 were generated. Based on this fact, these cost estimate values all had one common characteristic; i.e., they were all obtained using the 95th percentile project schedule values. Having this in common meant that it was acceptable to investigate these values as a separate subset, to see if they provided any additional information that would allow for the se-

J. Constr. Eng. Manage., 2002, 128(3): 211-219

Table 8. Comparison of Project Schedule Means for ODP Project

	Conditional project schedule		Probabilistic schedule		
Percentile ranking	Project Schedule (days)	Standard Deviation (σ)	Project Schedule (days)	Standard Deviation (σ)	Statistically Significant Difference
80th	772	0.683	757	0.525	Yes
85th	774	0.744	760	0.562	Yes
90th	776	0.809	763	0.690	Yes
95th	778	0.876	767	0.916	Yes

lection of a high percentile cost estimate value at the 95th percentile project schedule data. Thus, the CE values were plotted against their respective range estimate PLs and fitted using regression analysis. This resulted in the plot shown in Fig. 7 and the polynomial regression relationship shown in Eq. (3). The R^2 value of 0.96 indicates that using the percentile level to predict the cost estimate at the 95th percentile project schedule data resulted in approximately 96% of the variability in the cost estimate being explained

$$CE_i = 296.52PL_i^2 - 237.63PL_i + 1{,}656.7; \quad (0 \leqslant PL \leqslant 1) \quad (3)$$

Predicting Project Cost Estimate using Conditional Percentiles

The next step in the MSAT procedure is to fit a regression equation to the cost estimate data, generated at the 95th percentile project schedule data, and the conditional percentiles of this data set. This regression relationship, which represents a good fit to the data, is given by Eq. (4) and is illustrated in Fig. 7

$$CE_i = 126.76PL_i + 1{,}599.3; \quad (0 \leqslant PL \leqslant 1) \quad (4)$$

Thus, by using the range estimating and conditional percentile level associated with the cost estimate values generated at the 95th percentile project schedule data, it was possible, through Eqs. (3) and (4), to relate the cost estimate to the project schedule at a predetermined percentile level. Having this relationship, therefore, makes it possible to determine the cost estimate based on a desired project schedule percentile level. Additionally, this procedure provides the opportunity to quickly consider other cost estimate options throughout the entire range of cost estimate percentiles, given that the project schedule confidence level was fixed at the 95th percentile.

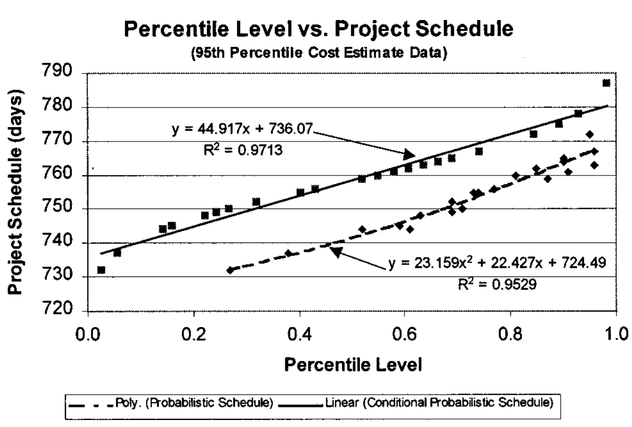

Fig. 9. Percentile level versus project schedule data

Comparison of Mean Cost Estimate Values

A statistical test for the comparison of means was performed at the 80th, 85th, 90th, and 95th percentile levels for the cost estimate values generated in order to determine whether or not there were statistically significant differences between the values obtained from Eqs. (3) and (4). The values being considered for comparison are presented in Table 6. These data represent the mean cost estimate values obtained from the application of range estimating and conditional estimating at the 80th, 85th, 90th, and 95th percentile levels.

Based on the statistical comparison of these values, the final column of Table 6 indicates that there was a statistically significant difference between the observed values from the application of range estimating [Eq. (3)] and the conditional cost estimates generated using the linear regression relationship of Eq. (4). As a result of this significant difference, it was thus possible to apply the MSAT procedure, which is based on the conditional probability [i.e., Eq. (4)], in order to determine the desired percentile level cost estimate value, given that the 95th percentile level project schedule value is the one of interest to the planner.

Comparative Analysis for Cost Estimate Determination

Once it had been determined that there was a statistically significant difference between the cost estimate values determined using range estimating and the conditional cost estimates, the next step was to determine to what extent these approaches for determining the project cost estimate differed. To make this comparison, a plot

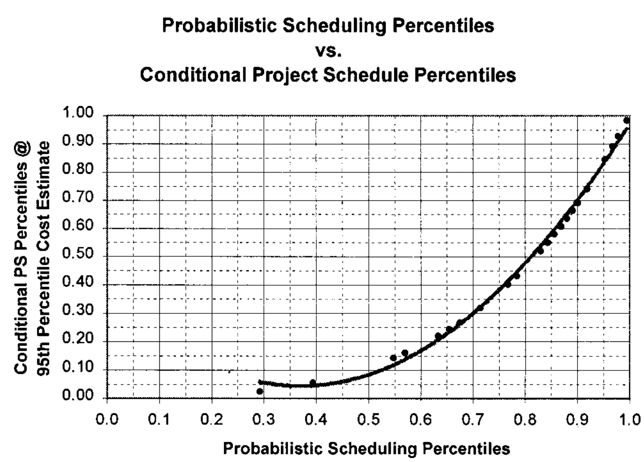

Fig. 10. Comparison of percentiles between project scheduling percentiles and conditional probabilistic scheduling percentiles

J. Constr. Eng. Manage., 2002, 128(3): 211-219

of the range estimate percentiles against the conditional cost estimate percentiles was made. This plot, shown in Fig. 8, was then used to determine the change in cost estimate percentiles when using traditional range estimating versus using the conditional cost estimate percentiles associated with the 95th percentile project schedule data.

If, for example, it is desired to have a cost estimate with an overall probability of overrun of only 10%, then using the traditional range estimating approach would have resulted in the selection of the 90th percentile cost estimate value. For the example project, this represents a cost estimate of $\$1{,}683\times10^5$. However, considering Fig. 8, the 90th percentile cost estimate obtained using range estimating corresponds to the 65th percentile level, based on the conditional percentiles developed for the 95th percentile project schedule data. Hence, not taking the project schedule into consideration when determining the project cost estimate resulted in a 25% underestimation of the true probability level associated with the cost estimate. This meant that using the traditional range estimating approach would have increased the potential for lost revenues by 25%. Such a loss would have a significant adverse impact on any organization undertaking the execution of construction projects.

To be truly confident that the 90th percentile cost estimate is obtained, after taking the impact of the project schedule into consideration, Fig. 8 indicates that a higher overall percentile level range estimate would have to be considered. For the example project, the 90th percentile cost estimate value, taking the 95th percentile project schedule data into consideration, was $\$1{,}713\times10^5$. This cost estimate value thus provides an overall higher level of confidence than would have been obtained had the project schedule been ignored.

Predicting Project Schedule using MSAT

The entire analysis just presented was predicated on fixing the project schedule at the 95th percentile level. If instead of fixing the project schedule value, the cost estimate value was fixed at the 95th percentile, an analysis similar to that undertaken in the previous section could be performed. Simulating the example project and fixing the cost estimate at the 95th percentile resulted in project schedule values and their corresponding probabilistic scheduling percentile level were generated (see Table 7). Fitting regression models to the data set resulted in Eqs. (5) and (6), and as can be seen from Fig. 9, these models provided a very good fit to these data

$$PS_i=59.611PL_i^2-19.21PL_i+734.99;\quad (0\leqslant PL\leqslant 1) \tag{5}$$

$$PS_i=63.402PL_i^2+709.04;\quad (0\leqslant PL\leqslant 1) \tag{6}$$

Comparison of Mean Project Schedule Values

Eqs. (5) and (6) were then used to generate project schedule values using the traditional probabilistic scheduling approach and the conditional approach developed by fixing the cost estimate at the 95th percentile. A comparative analysis of the resulting project schedule values (see Table 8) shows that the schedule results obtained from both approaches were statistically significantly different. This means that some significant differences will occur when an integrated versus a traditional, nonintegrated approach is used to predict the schedule of a project.

Comparative Analysis for Project Schedule Determination

Having determined that there is indeed a statistically significant difference between the project schedule values determined from probabilistic scheduling and the MSAT approach, the next step was to determine to what extent the project schedule values provided by these approaches differed from each other. Considering the goal of selecting a project schedule value that has a 10% probability of overrun (i.e., a 90% probability level), the traditional probabilistic scheduling approach indicated that the 90th percentile project schedule value was 763 days. Examination of Fig. 10 indicates that this schedule value corresponds to the 70th percentile conditional project schedule value associated with fixing the cost estimate at the 95th percentile. This means that not taking the cost estimate into account resulted in a 20% underestimation of the true probability level associated with the project schedule and would have increased the potential for schedule overrun by 20%. Having this much overrun on the project schedule would again financially impact any construction organization in an adverse manner.

To obtain the overall 90th percentile project schedule value, taking the impact of the project cost estimate into consideration, it is necessary to select a project schedule of 776 days. This project schedule value, thus, provides an overall higher confidence level than would have been obtained had the cost estimate been ignored.

Conclusion

The multiple-simulation analysis technique is a new technique that was developed as part of this research study on integrated range estimating and probabilistic scheduling. It was developed specifically to address the problem of relating the probabilistic cost estimate and project schedule data such that high percentile values for both of these tools could be selected that were related in some meaningful way.

MSAT combines discrete event simulation, regression analysis, and numerical analysis in order to develop a model that explains the relationship between the stochastic cost estimate and the schedule data. This allows much more detailed and integrated project planning than was possible in the past, when range estimating and probabilistic scheduling were independently applied to construction projects.

As was demonstrated using the example project data, the ability to perform an integrated approach to project planning allowed for appropriate consideration to be given to the selection of the cost estimate by specifying the project schedule desired, and vice versa. This resulted in the selection of a cost estimate and project schedule value having an overall higher probability of not being exceeded. Ignoring this integrated approach and using the traditional range estimating approach would have resulted in a substantial underestimation of the cost estimate at the 90th percentile level, given that the 95th percentile project schedule had been chosen. In a similar manner, selecting the project schedule value based on a traditional probabilistic scheduling exercise would have significantly underestimated the project duration, given that the 95th percentile level cost estimate had been selected for this project.

Using the MSAT procedure allows cost estimate and project schedule values, both having high percentile levels, and which are related to each other in some meaningful way, to be selected.

J. Constr. Eng. Manage., 2002, 128(3): 211-219

MSAT was applied to several example projects, and was found to provide consistent results in all cases. It is, therefore, recommended as a reliable means of truly integrating the results of range estimating and probabilistic scheduling.

Bibliography

Carr, R. I. (1993). ''Cost, schedule, and time variances and integration.'' *J. Constr. Eng. Manage.*, 119(2), 245–265.

Construction Industry Institute. (1987). ''Work packaging for project control.'' *Source Document No. 28*, Univ. of Texas at Austin, Austin, Tex.

Hendrickson, C. T., and Au, T. (1989). *Project management for construction: Fundamental concepts for owners, engineers, architects, and builders*, Prentice-Hall, Englewood Cliffs, N.J.

Ibbs, W. C., Ashley, D. B., Niel, J. M., and Feiler, F. W. (1989). ''An implementation strategy for improving project control systems.'' *Project controls: Needs and solutions, Proc., Speciality Conf.*, ASCE, New York, 101–112.

Navon, R. (1994). ''Cost-schedule integration for cash-flow forecasting.'' *J. Comput. Civ. Eng.*, 1536–1539.

Rasdorf, W. J., and Abudayyeh, O. Y. (1991). ''Cost- and schedule-control integration: Issues and needs.'' *J. Constr. Eng. Manage.*, 117(3), 486–502.

J. Constr. Eng. Manage., 2002, 128(3): 211-219