CHAPTER 2

LITERATURE REVIEW AND SCOPE OF THE INVESTIGATION

2.1 GENERAL

Simulation tools can provide more realistic scenarios of complex problems which cannot be experimented on a mathematical model or physical model. When applied to construction projects, they can predict unexpected variations in conditions which influence cost and duration associated with the activities. Simulation models backed with accurate probabilistic models for activity costs and durations can provide more realistic representation of progress of the project. Additionally, delay and cost overrun study can provide a more activity specific technique to combat project duration and cost overrun - the two problems to be tackled for effective management of construction projects.

Efficient project management is a combination of various components and factors like understanding the objectives, knowledge of the activities, experience in economic usage of resources, suitability of tools used, proficiency in the handling of tools, awareness of planned time-cost relationship and the commitment of the firm to achieve the objectives. The common planning techniques used for scheduling engineering project are
• Bar charts
• Closely resembling network analysis like
  ▪ Precedence Diagram Method
  ▪ Critical Path Method or Critical Path Analysis
  ▪ Project Evaluation and Review Technique
  ▪ Line of Balance Method

In all the above techniques, duration of the different activities associated with the project is one of the major uncertainties. Estimation of the activity time is frequently established by norms of past performance of the firm. The activity duration is presumed either as some deterministic value or assumed to follow some probability distribution (generally Beta) with reasonable parametric values. The activity durations require continuous revision through the progress of the project depending on the site conditions and project delays. The conditions prevalent in developing countries like India are highly uncertain for realistic prediction of activity duration because activity durations are seriously influenced by acute variations in socio-economic, political, geographical and climatic scenarios existing in the country. These aspects influence highway construction projects more so because highways run across the length and breadth of the country traversing varied geographic and geologic conditions. Since the above-mentioned aspects influence cost and duration of highway construction projects considerably, it becomes essential to develop new techniques for accurate prediction of cost and duration of the different activities as the existing approaches disregard these factors. Such an approach is possible only by an analysis of existing related literature. The available literature can be categorised into
• Literature related to characterisation of probability distribution functions for activity durations.
• Literature related to characterisation of probability distribution function for activity cost.
• Literature related to development of simulation models for construction process.
• Literature related to development of decision support tools.

Recent developments in this field are summarised in the following sections

2.2 DURATION ANALYSIS

Studies were carried out on the analysis of activity durations for various construction processes like dozer cycle time and truck payload experiment etc. The analysis revealed that activity durations involved in these processes required more flexible probability distributions for statistical representation than eventually developed.

Abouri/k and Halpin (1990) reviewed different methods of characterising the probability distribution functions for activity durations in repetitive construction projects using probabilistic simulation. Kolmogrov-Smirnov Test and Chi square test were used for analysing the goodness of fit and Quartile - Quartile (Q-Q) plot was generated for visual inspection of the accuracy of the function. The method calculated parameters of distribution, associated with data, using moment matching method, maximum likelihood method and percentile matching technique. Normality checks for simulation
output were carried out using Shapiro-Wilk test. The analysis of simulation output included determination of whether the simulation model was deterministic or stochastic and whether it reflected static, transient or steady state simulation. The model was validated by comparing the results with historical data. This methodology is highly reliable for simulation of construction operations where the uncertainty in the parameters can be easily identified. However, this approach cannot be adopted where the uncertainties in the project parameters are unidentical and constraints exist on the availability of data. Further, this study was conducted with an assumption that duration of all the activities followed Beta distribution.

Simaan M. Abourizk, Halpin and Wilson (1991) formulated visual interactive fitting of beta distribution (VIBES) for arriving at a statistical model for activity durations using minimum time, maximum time, mean and variance or selected percentiles. This study specified end target distribution functions together with mean and variance or selected percentiles. The shape parameter for Beta distribution function was iterated until it matched with that of the sample data. User was given option to revise the curve till a satisfactory fit was obtained. It was also found that the distribution was continuous between the limits that were distinct, finite and non-negative and had unique mode. User was given flexibility to experiment with parameters of the distribution function so that user could modify the parameters of Beta distribution obtained from the software. This methodology provides a more realistic approach to fit Beta distribution function using four characteristic parameters. However, in any construction project Beta distribution function for activity durations is not sacrosanct under all situations and hence a different methodology may be adopted to obtain a more flexible probability distribution function.
Javier Fente, Cliff Schexnayder and Kraig Knutson (2000) in defining a probability distribution function for construction simulation presented a different technique to determine Beta distribution function for construction activities by adopting 75th percentile values. This method related minimum activity duration to the mode and 75th percentile with the maximum duration using factors or ratios. These parameters for the probability distribution were calculated from the available data using the ordinary least square error minimisation method. This method assumed that data followed Beta distribution and shape parameters found out by this method lay between the corresponding parameters obtained by Virtual Interactive Beta Estimation Software (VIBES) and moment matching method but the standard parameters like the mean, variance, kurtosis, and skewness were found to be the same as those obtained as maximum likelihood estimate. This method provides a unique Beta distribution function, which is having the parametric values positive and greater than one, for each activity.

In all the above studies the basic assumption was that the activity duration follows Beta distribution and that the shape parameters of the Beta distribution would be functions of the site uncertainties. This assumption might become debatable if the uncertainties tend to be highly randomised. In fact literature is available wherein other types of distribution are assumed for some construction activities.

Simaan M Abourizk and Halpin (1992) studied the statistical distributions for construction activity durations using dozer cycle time. This study provided a practical approach for fitting the probability distribution curve. The correct distribution was arrived at by constructing a frequency distribution
for data, relating the shape to a known distribution and comparing sample statistics to those of the distribution function. Skewness and Kurtosis coefficients were used to calculate the shape factors of the underlying distribution. The study concluded that more flexible distributions other than Beta were required to ensure proper modelling of diversified characteristics of construction duration data. The study concluded that distributions like Lognormal, Normal and Pearson distribution could be employed.

Christino Maio, Cliff Schexnayder, Kraig Knutson, Sandra Webber (2000) formulated probability distribution function for construction simulation which showed quantity of data required was important for probability distribution fitting. The goodness of fit was carried out using Kolmogrov Smirnov test (KS), Anderson Darling test (AD) and chi square test using Bestfit software. The study was carried out on truck payload monitoring system for 5400 sets of data collected using automatic data recorders on the construction equipment. It was found that, in most of the cases, data followed Extreme value distribution function. This paper clearly showed that the sample size was an important parameter for defining the probability distribution and pointed out that the KS test was the best suited if the sample size was large.

In the methodology put forward by Wie Chih Wang and Laura A Deistez (2000) they took into account correlation between the activity durations. This study identified various parameters like labour availability, material availability, weather conditions, etc that influence the activity durations. The influence of these parameters was quantified and characterised as probability distribution functions. The study had also taken account the uncertainties in project completion as well as the target completion dates of the
activities. This algorithm provided flexibility to account the uncertainties in the project by formulating child distribution functions which eventually contributed to the parent distribution function representing the activity duration. A factor based approach was used for the correlation between the child probability distributions which helped in evaluating the effect of parameters in defining the critical path of the project.

2.2.1 Inference

Most of the researchers in the area of construction scheduling assumed that activity duration follows Beta distribution. But in situations where parameters influencing the activity duration are subject to large degree of uncertainty due to unforeseen complexities it is likely that probability distribution need not follow Beta distribution. Some of the researches which were carried out for verifying the suitability of Beta distribution functions for construction processes emphasised the need for a more flexible probability distribution function to represent activity durations. Hence, it is felt that a thorough analysis is required for computation of probability distribution function of activity durations of projects, more so for highway projects where the uncertainties are more acute.

2.3 DELAY ANALYSIS

Studies carried out on the project duration and cost overrun emphasised the need for delay analysis to incorporate it into the project schedule model so that variations in the project duration and cost can easily be evaluated. Jonathan Jingsheng Shi, S.O. Cheung and David Arditi (2001)
formulated a method for computing the contribution of activity variations to project delay. The method consists of a set of equations, which can be easily coded into a computer program so that the project delay information of individual activities is easily obtainable. This method is applicable for intermediate construction process for evaluating in-progress project delay. Unlike the traditional approach this method is not based on the criticality of activities. The as-planned schedule provides a basis of comparison for evaluating the variation in the project duration and does not need to be updated. This computation method is based on simple finish-to-start relationship. Even though it is not applicable to other relationships such as start-to-start relationship and finish-to-finish relationships in the current form, it can be expanded to all these relationships.

David T Hulett (1996) carried out studies on schedule risk analysis using Monte Carlo simulation of project network. This study was carried out by simulating the project network and assuming suitable probability distribution function for the activity durations. The probability distribution function was arrived at based on the three time estimates viz, pessimistic, most likely and optimistic time and assuming it to be a triangular distribution. Monte Carlo simulation was employed to simulate the whole network. The total project duration thus obtained was stochastically analysed to obtain the cumulative frequency function for it. This analysis emphasised the need for thorough analysis of probability distribution function for the activity duration overrun and variation of activities from critical path that need management attention to prevent delay of the project. This study was conducted for industrial projects.
2.3.1 Inference

From the studies carried out, it has been observed that an integration of overrun analysis with simulation model for project schedule will provide a more accurate tool for predicting actual cost and duration of the project. The studies also indicate the necessity for the analysis of variation in critical path due to different combinations of activity durations in various simulation runs.

2.4 COST ANALYSIS

Activity cost has two components viz., direct cost and the indirect cost. Direct cost is dependent on site conditions like the type of soil, level of water table, machineries used, construction materials used, etc. The indirect cost is dependent mainly on the delay associated with the activity. Delay is mainly attributed to the various unpredictable site conditions, climatic conditions, non-availability of raw materials and machineries, socio economic problems, etc.

Project cost overruns are studied independently without considering the effect of delay on the project. However, in the field, it has been observed that delays are one of the major causes of cost overrun. This phenomenon is observed in many of the Indian construction projects that are having relatively large life span due to the labour-intensive technology adopted in construction due to economic reasons even though the nation is supporting mechanisation of industries. From the literature pertaining to analysis of project construction costs it was observed that the cost and time interaction of project activities was
not carried out in any of the studies. Recent research was limited to the study of various parameters that are likely to influence the project cost.

Barraza and Edward Back and Fernando Matta (2001) put forward an efficient project monitoring technique using probabilistic monitoring of project performance using stochastic S (SS) Curves. The method adopts simulation technique for obtaining SS curves. Final duration and cost are derived stochastically from the SS curve which is simulated using activity durations and associated cost parameters. Project monitoring was performed by comparing most likely budget and activity duration, obtained from respective probability distributions for actual progress with project’s actual data on duration and cumulative cost using both time based S curve and progress based S curve. This study provided a graphical representation of project cost and duration thereby integrating it to the project progress. Project performance was evaluated by comparing budgeted cost and planned duration with actual cumulative cost and time elapsed. Since SS curve provides probability distribution of the budgeted cost and time elapsed at any point of time, performance evaluation is carried out by comparing expected budgeted cost and planned duration with the actual cumulative cost and the time elapsed.

Tarek Hegazy and Nageb Wassef (2001) developed cost optimisation in project with repetitive nonserial activities. This model adopted full integration of line of balance and critical path methodologies with crew synchronisation and work continuity among nonserial activities to minimise direct cost, indirect cost, interruption cost, incentives and liquidated damages. This study incorporated an integrated approach for time cost trade off using genetic algorithm technique to identify combination of construction method,
crew size and work interruption for each activity to meet project dead line with minimum cost criteria. A software tool was developed using spreadsheet, which provided user with transparency to use in any project with repetitive activities. Experimentations carried out varying crew size produced a schedule, which was shorter than project deadline with minimum cost. This methodology adopted a non-traditional approach for optimising cost and could provide a more realistic estimate of cost and duration associated with each method of construction. However, this technique cannot be adopted to analyse projects, which have activities with more than three successor and predecessor activities as the computations become more complex to be handled by spreadsheets.

Wei Chih Wang (2002) formulated utility theory for determining project ceiling price based on utility theory and facilitated by cost simulation approach. Utility theory was applied to reflect owner’s perspective regarding price determination criteria while simulation approach was used to generate a more objective approach to support execution of utility theory. In this study, simulation model is used to generate project cost and associated probability. The cumulative probability distribution function was formulated using output of the simulation model. This was used to determine the utility values for the various project costs. The highest utility value was provided for the project cost with highest probability and the lowest utility for the cost with probability zero. The threshold criteria were determined as the cost that the contractor may have fifty percent chance of cost overrun or underrun. Depending on the value of maximum price, minimum price and threshold price, ceiling price function was developed based on owner’s perception. Based on these and the project criteria utility value, the corresponding probability of each cost was computed. The ceiling price for a project was found using the cumulative probability
distribution. In this approach, the option for user’s preferences regarding selection of probability distribution for simulated cost and the normalisation approach for the cost created a more realistic method of determining the project ceiling price.

2.4.1 Inference

Various studies carried out for determination of project and activity cost emphasised the need for a more accurate modelling technique for predicting project cost. This modelling is very complex as it depends mainly on the external factors like the delay, technology adopted, economic and social status of the employees, etc. However researches indicated that cost associated with the activities is related with the duration and it is necessary to have a combined analysis of cost and duration of activity as well as of the project.

2.5 SIMULATION MODELS

Simulation is one of the sophisticated tools available for modelling complex situations. In construction projects, activity cost and duration are highly unpredictable and uncertain which lead to over estimation or under estimation of the project cost and duration. This results in continuous rescheduling of the construction project depending on the site conditions. Even though PERT takes into account uncertainties in durations to some extent, it is found to be inadequate to account for variation in cost and duration due to variation in the critical path resulting out of combining activity durations with different probabilities. Also this analysis may be cumbersome.
Simulation approach for project scheduling involves determination of probability distribution function for activity duration and cost and then, simulating the network for various scenarios. In all the research models provided, activity cost and duration are not jointly analysed. Observation from various construction sites indicated that cost variation and delay are having close relationship.

Probabilistic optimal cost scheduling put forward by Leroy J Isidore and Back (2001) used discrete system simulation to simultaneously perform range estimation and probabilistic scheduling. The result obtained was analysed statistically to arrive at a compressed schedule and obtain a minimised cost estimate. Activity cost and duration were assumed to follow uniform probability distribution function, based on assumption that both cost and duration are likely to take any value between the extremes with equal probability. ABCsim Software simulated projects for both activity cost and duration and it was found that even though there was not much significant difference between project durations obtained by least cost method and probabilistic optimal cost scheduling technique, there existed a wide variation in minimum cost of the project obtained by the above methods. Moreover, this methodology provided a range of cost associated with each project completion duration, thus allowing the decision maker a more user friendly tool for predicting the project cost and duration. This model did not specify the type of probability distribution function followed by the cost and duration of the activities and the interaction between them. Moreover, the variation in critical path for cost and duration estimates was not considered
Automated construction simulation optimisation put forward by Abourizk and Jingsheng Shi (1994) provided optimum allocation of resources, from initial trial allocation using various simulation runs. This method adopted a technique for minimising the cost of resource per unit time and maximising the production rate. In this system, the user provides boundaries for resource inputs and simulation conditions. The system experiments various scenarios and analyses performance and evaluates, whether the scheme is acceptable or not under the set of objectives. This approach used activities and delays as basic elements. The activities consume resource and time while queues consume time for resources to be used up. This methodology provided optimum allocation of resources from initial trial allocation by various iterations within the set boundaries, hence minimising cost of resource per unit time and maximising production rate. The improvement in production rate in last two runs was used for obtaining the optimum allocation.

Integration of range estimation and probabilistic scheduling was put forward by Leroy J Isidore and Back (2000) In their study, multiple simulation analysis for probabilistic cost and schedule integration were employed. Linear analysis was used to quantify the values of project cost estimate and probabilistic schedule. This technique integrated discrete system simulation, regression analysis and numerical analysis to relate stochastic cost estimate and schedule data. In this model, the direct cost and duration of each activity was assumed to follow some probability distribution while indirect cost was considered as a constant. The simulation model was formulated to replicate the project cost and duration. Linear regression analysis was adopted to form a relationship between project cost and duration. This model identified a range of values for the project duration, with some probabilistic percentile values. This
technique did not take into account joint probability function between activity cost and duration and the variation in parameters of probability distribution of cost and time due to variation in the contributing parameters associated with the project. This methodology did not take into account the variation in the critical path with the variation in combination of activity durations. In this study, activity durations and associated cost were taken from independent probability distribution functions. Moreover, this study did not specify the type of probability distribution function that are followed by activity duration and cost and their interrelationship in construction projects.

Activity based construction modelling and simulation method adopt one single element modelling for general construction process instead of multiple elements as in current simulation systems. The three stages by which the model proceeds are activity selection by the system, simulation advance and entity release. The method put forward by Jonathan Jingsheng Shi (1999) used a unique simulation method with a major advantage of simulating dynamic and random behaviour of construction process. This methodology differentiates between idle and active states by using different modelling elements and represents static model of the construction process. The simulation model provided the platform to experiment with the ABC model so that dynamic behaviour of construction process could be observed by tracing movement and interaction of the entities in model. This model provided system production of each activity, which is defined as product of iteration count of the activity and the production of several activities in each operation. Apart from the system, production the model provided the utilisation of resource and the time distribution of resource that could provide the amount of resource that was utilised fully and duration for which resource had been kept idle before it was
used. The activity nodes for simulation were based on the construction process and do not standardise probability distribution function for the activity durations.

David J. Harmelink and Maria Andrea Bernal (1999) provided a technique to use various types of probability distribution functions for each activity in simulating linear schedules. The study presented an innovative technique for resource scheduling of the various equipment required in highway projects, in which the activities are linear in nature. In this study the productivity rate of truck operation was analysed for the activities to be carried out by various machineries. Various types of distribution functions for the activity durations in the projects were taken into account. This methodology did not take into account various cost and cost overruns due to delays. This technique assumes different probability distribution functions for various activity durations so as to provide a more realistic representation of activity durations for the construction process.

In Simplified CPM/PERT simulation model put forward by Ming Lu, Abourizk (2000) the project was modelled as a network of merge node and burst nodes and an entity was designed to leave merge node only after arrival of all the entities. The nodes were made to delay the entity in accordance with the activity duration. The departing entity from merge node would start only when all the incoming entities had reached the node. In forward pass, an entity was created at the origin node at control time and as it passed forward and reached the merge node with n branches, all the n entities were batched into single entity and released to the next node. This study assumed Beta distribution for activity duration and provided the project manager with some flexibility by
providing option for objective approach to fit the sample. This technique provided a set of mathematical equations to calculate the criticality index of the activity, which had to be computed in the backward pass. However this study dispensed the importance of critical path and computed the criticality index to various activities only.

Chua and Li (2002) formulated resource integrated simulation modelling for construction, which was used for project scheduling at two abstraction levels, viz. process level and resource level. At resource level, it conceived the project as a collection of resources involved and their interaction. Logic of operations was represented as internal resource flow between complex resources. Process level provided logic necessary and process duration was based on the attribute value. Interaction between resource and process was achieved by four kinds of mechanisms

- simple resource flow;
- internal complex resource flow;
- common process and;
- interactive signals

The classification of resources varied with operations depending on number of attributes of each resource used.

2.5.1 Inference

From most of the researches it is observed that there is a need for formulation of a simulation model which can analyse jointly the duration and cost of the activities. The cost-duration overrun simulation models if formulated
can provide a more realistic estimate of project cost and estimation. Simulation model study has to be carried out for a more realistic analysis of project completion time because the project completion time which is likely to occur with maximum frequency may not occur for a unique critical path. Total project duration and critical path are likely to vary in each set of simulation run. Moreover, it is likely that with more number of simulation runs more than one path might become critical depending on the project duration and combination of activity durations.

2.6 DECISION SUPPORT TOOL

Research on development of decision support tool integrating simulation studies was found to be very limited in the field of construction scheduling as compared to the process engineering. The developments in this field included formulation of simulation languages, simulation of construction processes and integration of artificial intelligence in construction project scheduling. Some simulation languages are used as decision support tools for various construction processes. The use of simulation languages in construction management is evident from CYCLONE, STROBOSCOPE, etc. Considerable research is carried out in modification and development of add-ons to the simulation languages.

Photios G Ioannou and Julio C Martinez (1998) put forward a probabilistic scheduling technique with the aid of an add-on program to STROBOSCOPE using state based probabilistic decision networks. This approach provides correlation between activity duration, precedence relationship and controlling path. The software developed assumes standard
Beta distribution for activity durations and incorporates Central limit theorem to compute total project duration, its probability and standard deviation from thousand simulation runs. The methodology simulates the PERT network and computes average duration for each activity. This duration is taken for the calculation of total project duration and correspondingly its probability and standard deviation are computed. This study indicated that there was no single critical path for the entire project network. All paths were likely to become critical depending on the combination of activity durations in each simulation run. It also computed criticality index associated with each of the paths and also indicated that all paths are likely to become critical depending on various scenarios.

Ming Lu (2002) in his research on enhancing Project Evaluation and Review Technique simulation through artificial neural network based input modelling has proposed an artificial neural network (ANN) model to estimate errors in calculating statistical parameters for Beta distribution. In this model, distribution defined is by minimum, maximum values and values along the upper and lower quartiles and the range of distribution parameters are obtained by observing their effect on the final distribution curve. The various criteria that are taken into account for training ANN in characterising the distribution function are

- distribution should be continuous, non-negative and finite between the lower and upper bound
- distribution should have unique mode and most likely value and should be able to accommodate various degrees of skewness.
2.6.1 Inference

From the limited literature it was observed that most of the decision support tools were formulated for analysing the construction process or for verifying the probability distribution functions. Many of the decision support tools were formulated as simulation languages. Though these simulation languages were efficient in modelling the construction process, studies were not conducted on combined analysis of cost and duration of the construction project.

2.7 OBSERVATIONS FROM THE LITERATURE

The review of literature presented in the previous sections indicated the following aspects

- Though Beta distribution is the most advocated probability distribution to represent activity duration in construction process, there is evidence from certain researches that alternative distributions may be more apt in certain cases.
- There is not enough information available regarding the probabilistic nature of activity costs or their correlation to activity durations.
- There does not exist adequate literature to define clearly the relation between delays and their influence on costs and duration overruns when these are assumed to be stochastic in nature.
Project scheduling through simulation of network would provide more accurate information regarding project duration and cost. However these models tend to get complicated when the stochastic information of activity durations and cost are introduced into the models.

There is scant literature on decision support tools which can help project managers to estimate project duration, cost and overrun and the risk involved with a particular decision.

2.8 OBJECTIVES OF THE PRESENT STUDY

In recent years India has taken up infrastructure development especially highway development on a large scale involving thousands of crores of Rupees (billions of dollars). India is a vast country with vastly diverse geographic, climatic, topographic, political and social features. Since highways run through such diverse states, the influence of these features on the project parameters can be highly constrictive and can create highly uncertain environment for the construction process. Hence more reliable methods of estimation of the project parameters like duration and cost and some decision support system shall be developed if considerable savings in time and money are to be ensured. With these aspects in mind and based on the indicators of earlier section, the following objectives are set forth for this current investigation.

Collect large amount of data from various highway construction projects across the country and analyse these data and
• describe accurately the probabilistic nature of the durations and costs of activities associated with highway construction
• conduct a survey to determine the causes of delay in highway construction projects and the activities affected by these delays
• perform a delay analysis to obtain the correlation between the duration and cost overruns and the causes of delay
• introduce the above probabilistic models for activities into project schedule network along with activity precedence and succedence relationship
• develop simulation models for network analysis involving duration, cost, causes of delay and overruns as project parameters
• develop a decision support tool, which uses the results of simulation runs as knowledge base, to help estimate project duration and cost and the risk associated and overruns if any due to delay, so that a project manager through a scenario analysis can appreciate the impact of any decision he/she makes on the project duration and cost.

While majority of the results of this study may be applicable only for highway projects and some of them only under conditions as exist in India, the process decisions are expected to be universally applicable.