Advanced Engineering Informatics 29 (2015) 380–395

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Joint probability for evaluating the schedule and cost of stochastic simulation models

Mohammed Mawlana [a], Amin Hammad [b,*]

[a] Department of Building, Civil & Environmental Engineering, Concordia University, 1515 Ste-Catherine Street West, Montréal, Quebec, H3G 2W1, Canada
[b] Concordia Institute for Information Systems Engineering, Concordia University, 1515 Ste-Catherine Street West, Montréal, Quebec H3G 2W1, Canada

ARTICLE INFO

Article history:
Received 19 August 2014
Received in revised form 13 December 2014
Accepted 21 January 2015
Available online 7 February 2015

Keywords:
Stochastic simulation
Joint probability
Conditional cumulative probability
Joint contingency estimation
Schedule generation
Seed number

ABSTRACT

The review of construction engineering and management literature shows that the occurrence of multi-performance indices in stochastic simulation models have been treated the same way as the occurrence of a single performance index. By doing so, the correlation between these indices and the impact they have on each other are ignored. Their occurrences have been treated as disjoint, which leads to errors in evaluating the probabilities of the performance indices of these models. The objectives of this paper are to present a new method that can: (1) quantify the impact of uncertainty on the project schedule and cost simultaneously; (2) calculate the conditional probability of the project cost given a specific project duration, and vice versa; (3) find the best project duration and cost that meet a specific joint probability; (4) estimate the project schedule and cost joint contingency using joint probability; and (5) generate a schedule representing a specific joint probability. The paper presents the implementation details and several case studies to demonstrate the feasibility of the proposed method. The proposed method shall provide a more accurate analysis to the output of stochastic simulation.

1. Introduction

Simulation is used to model and analyze real-world systems by mimicking their behavior in a virtual environment [26]. Simulation models capture and represent the knowledge available about the modeled system. The models' settings, input data, and simulation output represent additional information regarding the behavior of the modeled system. Understanding, analyzing, exchanging and reusing this knowledge, will lead to reliable decision-making. However, the complexity and the uncertainty of construction operations, and the limited knowledge of the behavior of the operations under different combinations of resources make it impossible to describe such operations using a closed form formula. Therefore, discrete event simulation can be used to model the complexity of the construction operations while the uncertainty can be modeled by incorporating stochastic durations and/or costs into a discrete event simulation model or into a Monte Carlo schedule network. One of the main uses of simulation in the field of construction engineering and management is estimating the performance of stochastic simulation models and schedule networks; therefore, the rest of this paper is focused on stochastic simulation models and schedule networks with multiple correlated performance indices, such as duration and cost.

Traditionally, a model is evaluated in terms of the mean value of each performance index separately. The focus is usually on quantifying the impact of uncertainty on the project schedule [29,49,4,40]. This approach provides insight on the probability of completing the project within a specific duration. In multi-performance indices models ([50,19,33,27,34]), the performance of a model is represented by the duration and cost values as a pair. However, these values are found by explicitly averaging the durations and costs of simulation replications without regard to their simultaneous occurrence. In addition, when comparing the performance of the model under different probabilities of occurrence, each performance index is separately assessed. The current practice does not provide any information on the probability of the project cost given a specific project duration, and vice versa. Without this information, the decision maker cannot quantify the impact of selecting a project duration meeting a probability of occurrence on the project cost, and vice versa. That is, the correlation between the duration and cost is not examined and the impact each performance index has on the other index is ignored. Due to the fact that there is a correlation between the project duration and cost, the analysis of the model performance must consider the simultaneous occurrence of the project duration and cost through the use of joint probability [14,48]. The current practice is valid only if a perfect

* Corresponding author.
E-mail addresses: m_mawlan@encs.concordia.ca (M. Mawlana), hammad@ciise.concordia.ca (A. Hammad).

http://dx.doi.org/10.1016/j.aei.2015.01.005

correlation between the project duration and cost exists, and the marginal and joint distributions follow a normal distribution. However, this is not always the case for construction projects.

In addition, a specific pair of performance indices selected based on the traditional approach is not necessarily achievable. The results of simulation replications may not generate a replication that has the same performance indices pair. The achievable performance indices' values for that model could be higher or lower than the values of the performance indices pair. Even if a specific performance pair is generated by the simulation replications, there is no method available for tracing this pair in order to generate the schedule of that pair. Therefore, there is a need for an analytical method to overcome these shortcomings. The objectives of this paper are to: (1) develop a method for considering the joint probability in construction simulation and quantifying its impact on the project duration and cost; (2) calculate the conditional probability of the project cost given a specific project duration, and vice versa; (3) find the best project duration and cost that meet a specific joint probability; (4) estimate the project schedule and cost joint contingency using joint probability; and (5) generate a schedule representing a specific joint probability. Applying this method is expected to have a noteworthy impact on reducing project risk and providing the decision makers with more accurate and useful information to plan and manage their projects.

2. Background

2.1. Stochastic simulation

Simulation can determine the output of a system based on the variations in the input to a system [18]. Simulation has been used in many fields such as supply chain [20], transportation [17], fluid dynamics [45], design optimization [47]; power systems [10], and biology [44]. Simulation in construction has been used for planning and resource allocation [2], comparing the outcome of alternative construction methods [41], analyzing earthmoving operations [35,32] and bridge construction operations [43]. Several methods have been proposed to collect the input data of the simulation models and generate probabilistic distributions to represent the stochasticity and correlation of these data [1,9,14,42,48]. Several researchers in the construction management field stated that time and cost are correlated [14,48]. Although there have been several advancements in construction simulation such as agent-based simulation [24], and automated knowledge discovery and data-driven simulation [3], the multiple performance indices obtained as the output of a simulation model are still treated as independent variables. Therefore, there is a need for a method that describes, analyzes and represents the knowledge of this correlation using joint probability.

Simulation can be either deterministic or stochastic. A deterministic simulation returns the same performance value for a specific configuration of decision variables whenever that configuration is evaluated as shown in Fig. 1(a). Whereas a stochastic simulation returns a different performance value every time a specific configuration of decision variables is evaluated as shown in Fig. 1(b).

The typical method of evaluating the performance of stochastic simulation models is by performing a number of simulation replications (N) [14,7]. The purpose of these replications is to capture the uncertainty and to obtain a good confidence level in the evaluation of a solution's performance.

A stochastic simulation can be described as [7]:

$$\bar{J}(\theta) = E[L(\theta, \omega)] \quad (1)$$

where $\bar{J}(\theta)$ is the mean value of the model performance, θ is a vector of all the decision variables, ω is a simulation replication with a certain uncertainty, $L(\theta, \omega)$ is the performance value with uncertainty using simulation, E is an expectation of the performance value. $L(\theta, \omega)$ can represent the cost or the duration of an operation. Since one simulation replication is not representative of a stochastic simulation model, several replications must be made to obtain a sound estimate of $E[L(\theta, \omega)]$. The sample mean is the standard approach to estimate $E[L(\theta, \omega)]$ as shown in the following equation [7]:

$$\bar{J}(\theta) = \frac{1}{N}\sum_{j=1}^{N} L(\theta, \omega_j) \quad (2)$$

where ω_j is the j-th simulation replication.

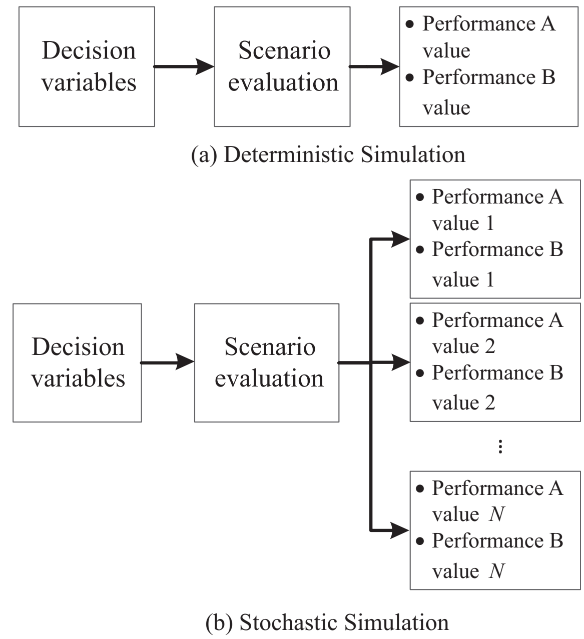

Fig. 1. Deterministic vs. stochastic simulation.

2.2. Joint probability of bivariate discrete random variables

Joint probability has been used in several applications such as integrated cost-schedule risk analysis [21,22,8], flood frequency analysis [23], system reliability analysis of flexible pavements [11], estimating extreme sea levels [28], and failure analysis of a series structural system [51].

When a stochastic experiment is performed, the outcomes of this experiment are: (1) the solution space which represents all the possible outcomes of the experiment and (2) the probability assigned to each outcome of that experiment. Each outcome can have one or more performance values. The stochastic experiment outcome is called univariate random variable, bivariate random variables, or multivariate random variables, which reflects a single performance value, two performance values, or more than two performance values, respectively. For example, the duration and cost of a project alternative are considered bivariate random variables in a time–cost tradeoff problem.

There are two main types of random variables: discrete random variables and continuous random variables. A discrete random variable is the variable that can take a finite or a countable number of values, while a continuous random variable can take any value in an interval [5]. The method of calculating the probability of an outcome of an experiment depends on the number of random variables. For the purpose of this paper, finding the joint probability of bivariate discrete random variables is the main interest. However, the concept can be extended to include multivariate random variables.