Methodology for Conducting Discrete-Event Simulation Studies in Construction Engineering and Management

Julio C. Martinez, M.ASCE[1]

Abstract: This paper suggests the methodology to follow when conducting discrete-event simulation (DES) studies in construction engineering and management research. Emphasis is made on the steps that, due to the uniqueness of the construction environment, are particularly important yet are not discussed extensively in the general DES literature. Guidelines are provided to determine what aspects of a DES study demand a rigorous application of the theory depending on the purpose of the study. The paper concludes with the importance of properly understanding the probabilistic concepts upon which DES relies and on coupling this understanding with engineering judgment as a key for successful use of DES in construction research.

DOI: 10.1061/(ASCE)CO.1943-7862.0000087

CE Database subject headings: Analytical techniques; Quantitative analysis; Construction management; Methodology; Simulation.

Author keywords: Analytical techniques; Discrete-event simulation; Design of operations; Quantitative analysis.

Introduction

Discrete-event simulation (DES) has been recognized as a very useful technique for the quantitative analysis of operations and processes that take place during the life cycle of a constructed facility. According to Abudayyeh et al. (2004), during 1997–2002 (the last period included in the paper) simulation was the most frequent topic among articles published in the *Journal of Construction Engineering and Management* (JCEM). While simulation is a very broad term, an inspection of all articles (348 as of this writing—see http://scholar.google.com/scholar?q=simulation&as_publication="journal of construction engineering and management") containing the word "simulation" that have been published in the JCEM reveals that the overwhelming majority of them are about the use of DES for modeling and improvement of construction operations.

The proper methodology to follow when performing a DES study is a subject that cannot be covered in a single article of limited length. Law and Kelton (2000) provided an extensive description of the process (in 768 pages) that is both formally rigorous and practical. Even that thoroughly comprehensive source refers readers to nearly 1,000 other sources for more in-depth information. Shannon (1998) provides a bird's eye view of the process for those that want a quick description. The subject of simulation methodology is so extensive, that the annual Winter Simulation Conference has dedicated tracks for this purpose (the last few conferences have included two tracks for modeling methodology and two more tracks for analysis methodology). Articles from the proceedings of the 1997 and later Winter Simulation Conferences are available online at http://wintersim.org/pastprog.htm.

In general, the process for conducting a DES study of a construction operation is not different than the one that should be followed in any other field. However, certain aspects of the construction environment are very different, and these differences sometimes dictate the use of (or need for) different tools, techniques, and precautions. Some things are only very briefly discussed in the general theory of DES, yet are very important and merit explicit attention in a construction setting.

This paper outlines the methodology for conducting DES studies of construction operations. The parts of this methodology that are no different between construction operations and general use are covered very briefly since extensive sources for this are readily available. Steps that are particularly challenging or important to construction and have not been fully discussed in the literature are discussed in more depth. The purpose for which a simulation model is built has a strong impact on which aspects of the methodology (or parts of the model in question) require the fullest rigor and which can be treated more flexibly—a discussion of this follows the presentation of the methodology.

Application of DES to Construction Operations

There are many ways in which to break up the steps needed for a proper DES study. The steps listed next should be appropriate for DES studies in construction engineering and management:

- Determine the extent to which a DES model can lead to better understanding of the system in question or to obtaining quantitative measures of performance for the problem of interest. The purpose of this first step is to answer the question: "Is DES the appropriate tool for the problem at hand?"
- Establish the scope of the model and the specific questions that the model should answer.
- Define the model for the operation. This includes establishing the level of detail of the model, selecting the elements that will be used to represent the real system (e.g., resources, activities), and capturing the logic appropriately.

[1]Associate Professor, School of Civil Engineering, Purdue Univ., 550 Stadium Mall Dr., Room 1245, West Lafayette, IN 47907. E-mail: julio@purdue.edu

Note. This manuscript was submitted on September 1, 2008; approved on May 4, 2009; published online on May 8, 2009. Discussion period open until June 1, 2010; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 136, No. 1, January 1, 2010. ©ASCE, ISSN 0733-9364/2010/1-3–16/$25.00.

- Collect and synthesize data about the operation to suit the model. In addition to actual data collection, this includes determining whether basic probabilistic assumptions hold [typically independence and identical distribution (IID)], fitting distributions, and testing for goodness of fit.
- Verification of the model and data to ensure that it matches the modeler's understanding of the system.
- Validation of the model to ensure that it matches the real or imaginary system. If a model has been verified, validation seeks to determine whether the modeler truly understood the real or proposed system. This step requires the participation of individuals that are intimately familiar with the real system.
- Analysis of simulation output for a single run.
- Design and execution of simulation experiments.
- Analyze the output of experiments to determine the performance of various system configurations or to select the best from among several alternatives.
- Document and present results.
- Use results for decision making.

Is Discrete-Event Simulation the Appropriate Tool for the Problem at Hand?

Quite often a simulation model is used for purposes that are better served by other methods of analysis that require less effort to implement properly (e.g., a spreadsheet). This often happens when DES has been newly introduced to an organization and it is incorrectly thought to be the solution to all problems. Using DES only for situations that require it is critical.

In May 2001 the writer visited *Dragados*, a Spanish construction firm that has successfully embraced DES as a tool, to learn about their experience implementing DES and to teach a short course on DES modeling using EZStrobe (Martínez 2001). Dragados trains all its engineers to recognize situations where DES can be of value. When these engineers recognize these situations, they communicate it to a team of DES specialists who look at the problem in question and determine whether indeed DES is the tool to use. Only in those cases do they perform a DES study, and they report returns on investments in their studies of 2,000%. The key to this success is that DES is applied only when appropriate. Balbontin-Bravo (1998) and Halpin and Martinez (1999) documented some of the operations where Dragados has effectively used DES.

What are the characteristics of problems where DES can be particularly valuable? To answer this question it is essential to thoroughly understand the problem at hand, what aspects of it can and cannot be represented using DES, the effort required for modeling and collecting the appropriate data, and the effectiveness and applicability of other methods of analysis. Typically, the problems that are well suited to DES:

1. Involve significant uncertainties in the time required to accomplish tasks and/or in the amount and quality of materials consumed and produced.
2. Are logistically complex with a number of context sensitive dynamic rules and decisions.
3. Have interdependent components subject to complex activity start-up conditions where many resources with distinct properties must collaborate according to highly dynamic rules (Hooper 1986).

Problems that do not involve significant uncertainties and/or logical complexity can usually be analyzed more effectively using other techniques. Often, however, even though DES is determined to not be the ideal tool, the process of attempting to build a DES model forces the engineer to think about a problem in ways that lead to its solution (without relying on actually running simulations).

Model Definition

A model is a representation of a current or proposed operation or system. If a model is valid, it is possible to experiment with the model to better understand the operation it represents. A model is considered valid only for the purpose for which it is built, and not in absolute terms. A model built to determine the number of trucks needed for an earthwork operation may be considered valid for that purpose, but may not be necessarily valid, for example, to determine the optimal fueling strategy for a fleet of trucks or to obtain a cost estimate for bidding purposes.

In order to build a model appropriately, significant effort has to be dedicated to understand the real or imaginary operation that it will represent and the specific purpose of the model. The engineer can then determine the level of detail with which to model different parts of the operation.

The model is then created in a process that is both art and science. Shannon (1998) used an analogy to oil portraiture to illustrate the artistic aspect of model building. Knowledge of perspective, color theory, media, inks, and the study of portraits from the masters does not mean that an individual can produce an oil portrait of quality. Extensive practice and talent are also needed.

A variety of tools and techniques exist for model development. Martínez and Ioannou (1999) discussed some tools that can be used to model construction operations and describes in detail some that have been created specifically for this purpose. The tool that is used is not that important, so long as the person developing the model understands it and is capable of performing the art of modeling with it. A key issue to consider while developing a model is that the aspects of reality that are relevant to the purpose of the model must be faithfully represented without simplifications and assumptions that would otherwise make the model trivial or misleading (Shannon 1999).

Data Collection and Synthesis

Collecting data for construction operations can be very challenging. Typically, construction activities are dependent on the conditions that exist at the time they are performed. The distribution of the time it takes for a truck to haul material along a path on a rainy winter day may be different than on a sunny day in the spring; it may be different even on the same day for different operators; or at different times of the day with the same operator.

Most distribution fitting techniques are based on the assumption that the observed data are IID, i.e., independent (data points are not related) and identically distributed (they all have the same distribution). However, in many instances the data as observed is not IID. In order to synthesize the data (reduce it to a probability distribution) and later generate it appropriately during simulation, the uncertainty needs to be expressed as a function of other variables that are IID. Sometimes this is not possible and in those cases it is necessary to know the conditions that existed when each data point was collected, so that when grouped by common conditions, the data are IID or can be expressed as a function of an IID variable.

In an ideal world data are collected automatically along with

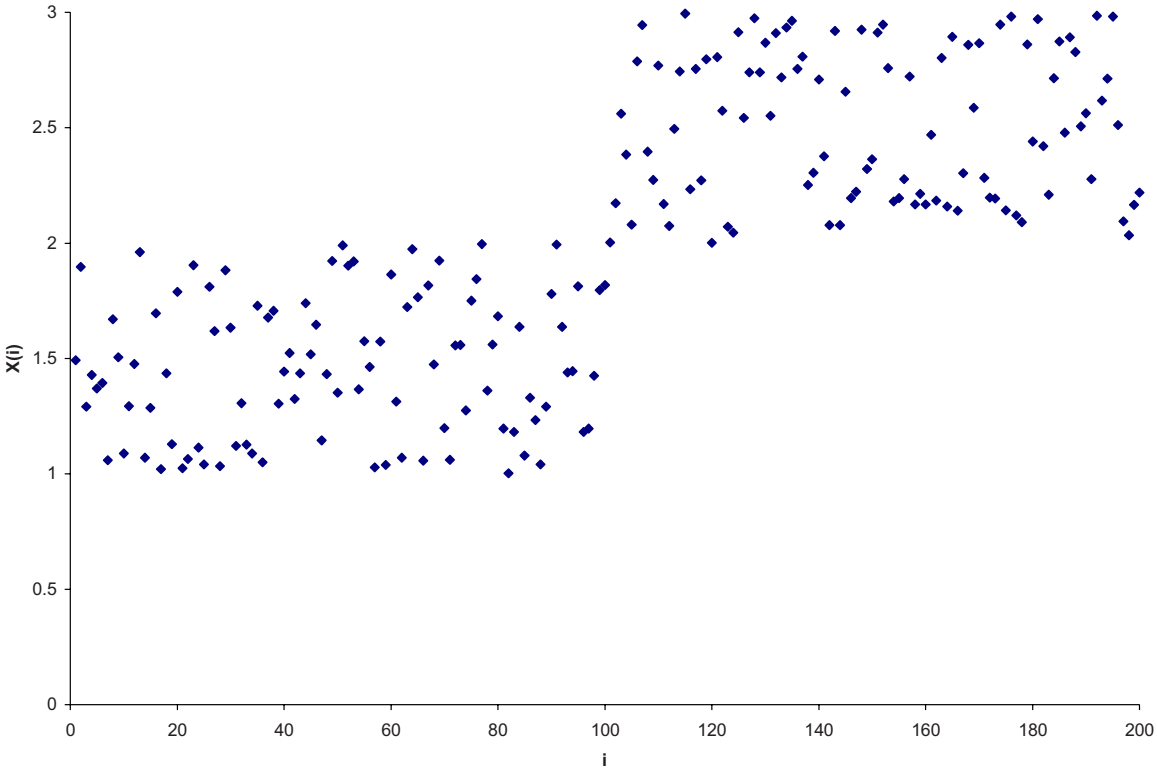

Fig. 1. Sequentially ordered plot of truck loading times

the conditions. An interesting discussion on this topic can be found in Kannan and Vorster (2000), who used the instrumentation available in off-highway trucks to collect data autonomously along with the conditions associated with each data point.

Another complication associated with collecting data for construction operations involves lack of consistency in human judgment during collection. In a discrete system, activities are assumed to have a specific start and end time. However, in the continuous real world, the point at which one activity starts and the other ends may not be very clear. Hildreth et al. (2005) describe a computerized procedure for collecting activity times involving outdoor heavy equipment from GPS traces. Using their techniques, it is possible to bring consistency to the definition of activity start and end times and to collect data autonomously.

Engineering judgment and a thorough understanding of probability and statistics are essential for proper DES modeling. The availability of data as ideally desired is rare, but the engineer needs to know that this is a difference between the model that uses the data and the real world. The engineer needs to couple that knowledge with probability and statistics to understand how the model differs from reality, and to judge its impact on the validity of the model. In construction, most of the problems with the analysis and subsequent use of the data during simulation stem from using statistical techniques that assume IID, when indeed the data are not. Some specific ways in which data are not IID, and suggestions on how to analyze and restructure models so that they can properly rely on the IID assumption, follow.

Homogeneity

Sometimes it is not obvious to the engineer whether there is a difference in conditions between data collected at different times. Other times the engineer may recognize that there are some differences in the conditions but may not know whether the IID nature of the data are affected by this difference in conditions.

To illustrate, assume that the times for a truck loading activity during day 1 (sunny, 60°F) follow a Uniform[1,2] min (i.e., uniformly distributed between 1 and 2 min). Times for the same activity during day 2 (rainy, 38°) follow a Uniform[2,3] min. One hundred samples are collected in the field on Day 1 and another hundred on Day 2. If that data are combined to form a set with two hundred observations and then fed blindly to distribution fitting software, it is likely that the software will strongly indicate that the data comes from a Uniform[1,3] min. A histogram of the data will reinforce this since it would look identical to the histogram of a Uniform[1,3] min.

However, in this example the data are not IID, the distribution of the first one hundred numbers and of the second hundred numbers are not identical. Using Uniform[1,3] in a simulation of the process will very likely produce invalid results, unless the sensitivity of the model's output to the time of the activity is extremely small.

A graph of the observed data points with the observation number (i) on the horizontal axis and the observed value $X(i)$ on the vertical axis (e.g., ordered plot), as shown in Fig. 1, makes this evident. Inlow (2006) in his excellent article about teaching engineering students to understand the concept of independence, suggested that the data points be sorted randomly and used to create a randomly ordered plot, as shown in Fig. 2. A difference in the visual character of the plots would provide indication that the data are not IID. According to Inlow, comparing a sequentially ordered plot to a randomly ordered plot "*makes it easier to determine if 1. apparent trends are systematic or simply random oscillations, and 2. if adjacent observations are related or autocorrelated since the random order plots will appear to be rougher or smoother than the original data if negative or positive autocorrelation is present*."

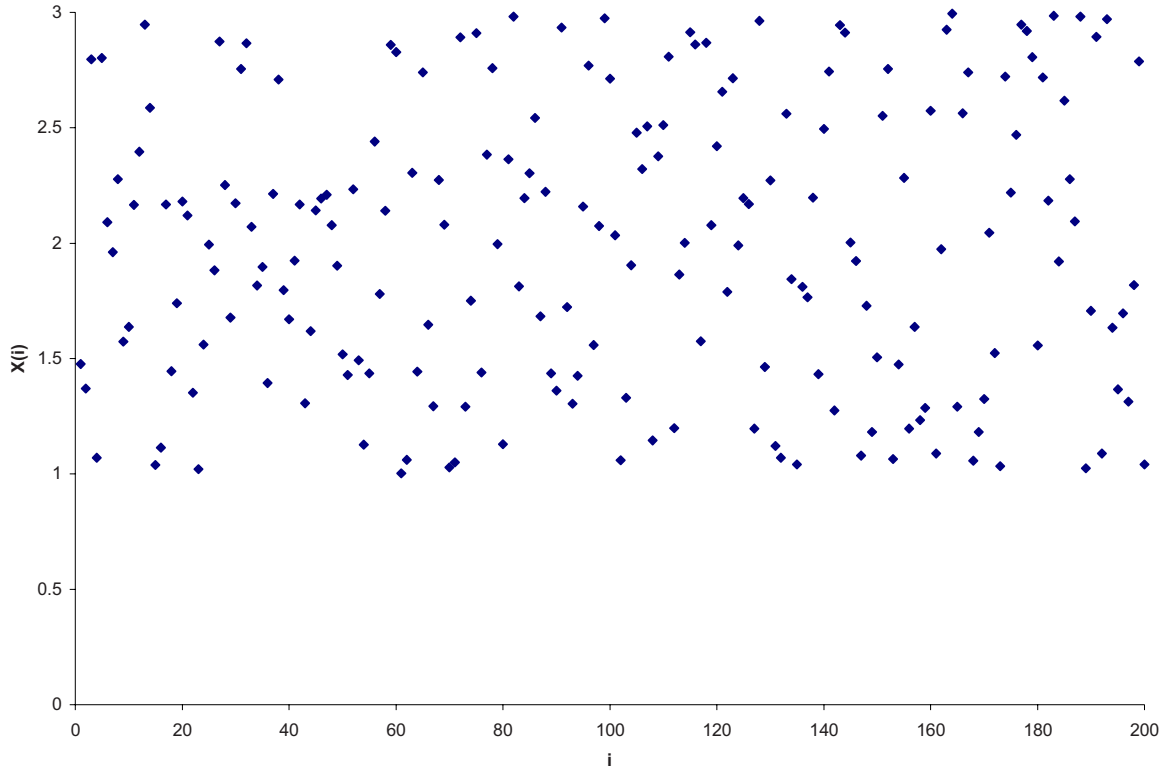

Fig. 2. DES randomly ordered plot of truck loading times

Furthermore, the graph of $X(i+1)$ versus $X(i)$ shown in Fig. 3 (e.g., a scatter diagram), clearly indicates that when an observation is between 1 and 2 the chance that the next observation is between 1 and 2 is near 100%. Similarly if the observation is between 2 and 3 the next observation is also between 2 and 3 with certainty. Quantitatively, the coefficient of correlation between $X(i)$ and $X(i+1)$ can be used to assess the degree of independence between successive observations. The lower the absolute value of

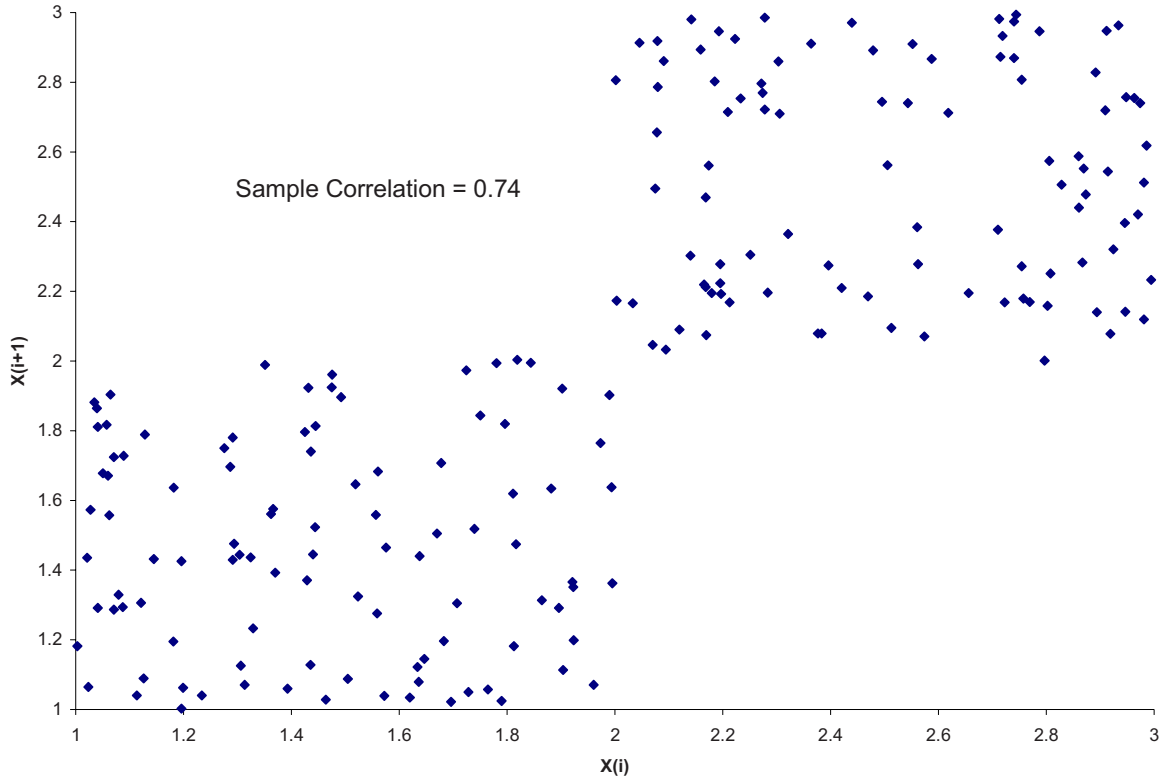

Fig. 3. Scatter diagram for truck loading times

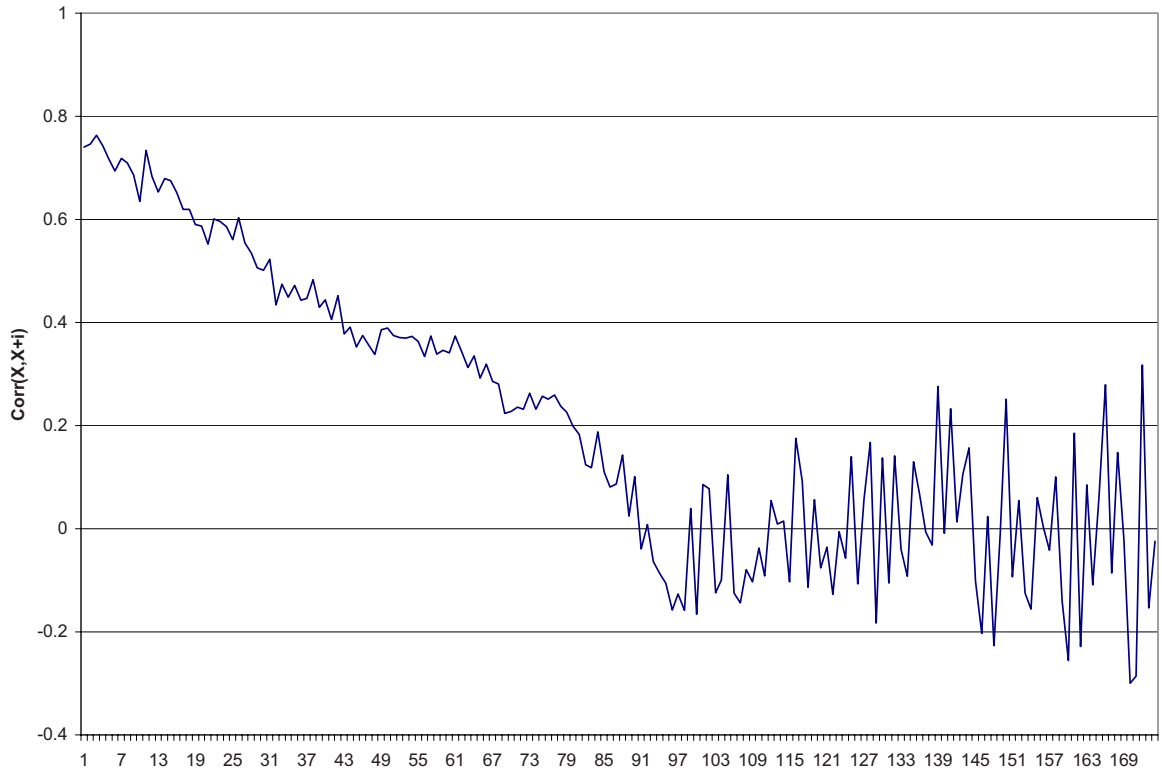

Fig. 4. Correlation plot for truck loading times

the coefficient of correlation, the greater the independence. If the data are totally independent, the true coefficient of correlation is zero. Fig. 3 shows the sample coefficient of correlation as an estimate of the true coefficient to be 0.74. This makes it pretty obvious that successive data points are not independent.

Even if successive data points are nearly independent, it is possible for data points N observations apart to exhibit some correlation, which would be an indication that the data are not independent. The correlation plot of Fig. 4 starts at 0.74 for data points one observation apart, and shows the sample coefficient of correlation for data points n observations apart. The plot shows that the sample correlation coefficient starts to oscillate between positive and negative with relatively small absolute values after n reaches 100, clearly indicating that from then on there is no correlation (i.e., the second set of 100 numbers is fully IID and exhibits no autocorrelation).

This example was artificially created to make the point. There is no real construction activity that is truly uniformly distributed, let alone showing two perfectly adjacent and equally sized uniform distributions on successive days. With real data that is not homogeneous, and that may have been collected in several batches, it may be very difficult to make a judgment by looking at an ordered plot of the observations or a scatter diagram. Sometimes the engineer, due to an understanding of what is being observed, will group the data ahead of time. Maio et al. (2000), for example, group truck travel times into several categories according to haul distance and weight before fitting distributions to each. However, often, the engineer will not know if the data are homogeneous. In these cases, the Kruskal-Wallis hypothesis test for homogeneity (Law and Kelton 2000) can be used to determine if two or more separately collected sets of data can be treated as homogenous and combined to determine a single distribution.

If separately collected data sets are not homogenous, then this is an indication that the conditions under which at least one of the data sets was collected differs from the others. If this happens, different conditional distributions should be determined for each set. Furthermore, the different conditions need to be understood and modeled. The artificial example used earlier, for example, can be modeled by some process that at the start of each day samples from a distribution to indicate good weather or bad weather (and for simplicity we are assuming here that weather is discrete and can only be either good or bad). The time for the loading activity can then be specified as conditional on the weather (e.g., Uniform[1,2] given that weather is good, Uniform[2,3] given that weather is bad).

Autocorrelated Data and Nonstationary Input Processes

Sometimes data are correlated to itself. When autocorrelation is positive a high value is likely to follow another high value, and a low value is likely to follow another low value. When autocorrelation is negative, there is tendency for data samples to alternate between high and low. In nonstationary input processes, the distribution of the data varies with time or sequence—the data exhibit a trend. Nonstationary processes tend to have some degree of autocorrelation. Examples of nonstationary inputs include the time required to perform an activity that involves learning, and hauling material at distances that are continuously increasing or decreasing.

The existence and degree of autocorrelation can be assessed graphically by comparing ordered and randomly ordered plots of the data, and by correlation plots and scatter diagrams. Formally, it is possible to assess IID using runs tests (Gibbons 1985). If the data are not IID, the underlying reason needs to be investigated, understood, and modeled (if possible). The ideal way to model

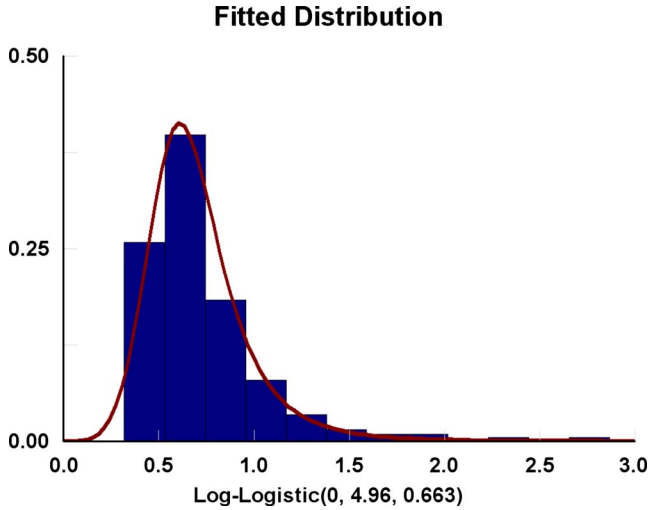

Fig. 5. Histogram of hours required for valve installation

such processes is to define the source of the data as a distribution where one or more of its parameters are a function of time, sequence, or other system state.

Consider an activity for the installation of valves, where the time required for each of the first 200 installations was collected. The histogram of the data along with the distribution fitted to it (assuming that the data are IID) by Stat::Fit (Benneyan 1998) are shown in Fig. 5. Since the fit is very good (the ChiSquare, Kolmogorov-Smirnov, and Anderson-Darling tests do not reject the hypothesis at the 0.15 level of significance), it may seem a good idea to use the fitted distribution in a model.

However, further inspection by using a sequentially ordered plot (Fig. 6), a scatter diagram (Fig. 7), and a correlation plot (Fig. 8) show that the data are not IID. In effect, the ordered plot suggests a nonstationary distribution with central tendency that follows an exponential learning curve. Fitting an exponential equation to the data using least-squares yields $\text{Hours}=2.86\times\text{Unit}^{-0.33}$, and is plotted in Fig. 6. We can then look at the deviation of the data from the learning curve to see if the deviations are IID, and if so, what distribution best explains the deviations. Fig. 9 shows the ordered plot of the deviations, Fig. 10 shows the scatter diagram, and Fig. 11 shows the correlation plot. All three indicate that the assumption of IID is likely. Fig. 12 shows the histogram and distribution fit of the deviations. Based on this analysis, the installation of valves of this type could be represented in a model using the formula $2.86\times\text{Unit}^{-0.33}+\text{Beta}[-0.233,0.194,1.49,1.26]$, where Unit is a dynamic variable that represents the sequential number of the unit being installed (the data for this example was artificially generated using the formula $3\times\text{Unit}^{-0.34}+\text{Uniform}[-0.2,0.2]$).

In many cases several activities depend on common underlying conditions. As the conditions change, so do the distributions of the activity times. When this happens, the data from the various activities are correlated. In these cases, the underlying conditions need to be identified and modeled, and the distributions used for the different processes expressed as functions of these conditions. It is also possible to determine n-variate distributions to explain the correlated data, but this is not easily done without many further assumptions regarding the underlying distributions and 1:1 correspondence among data sets.

Subjective Distributions

Experts often have a very good understanding of the distribution of some variable based on years of experience observing the vari-

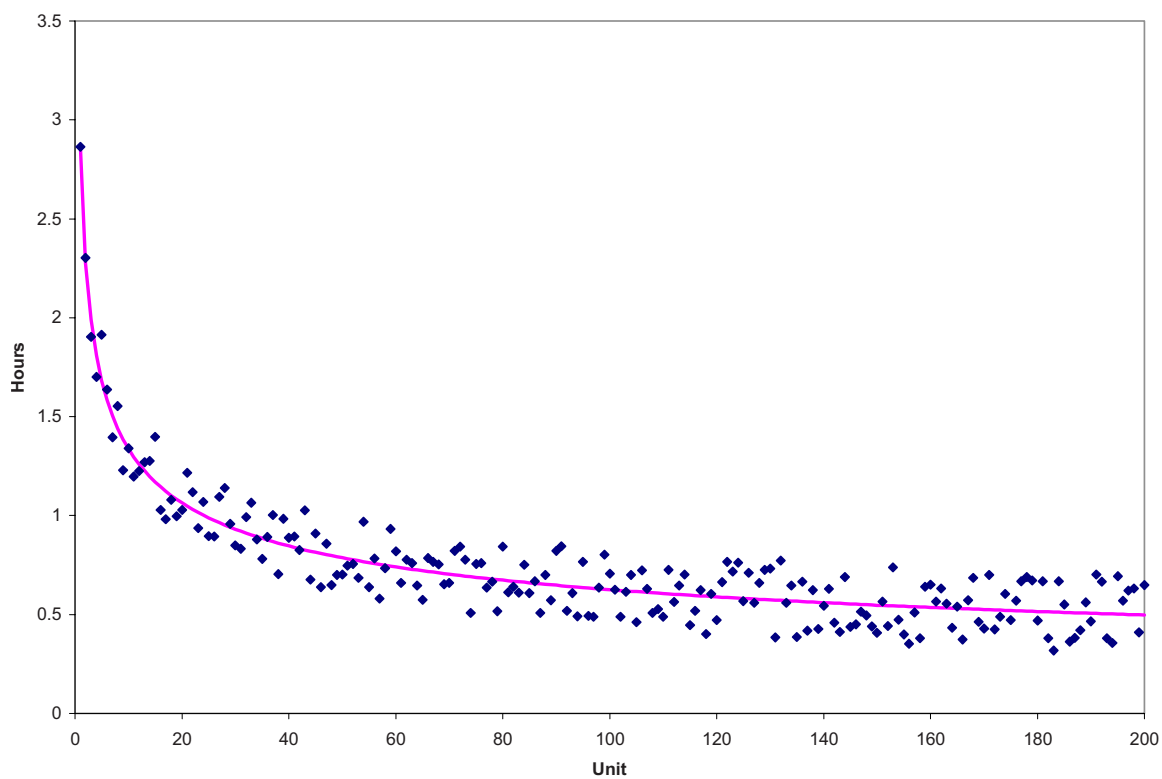

Fig. 6. Ordered plot for valve installation hours

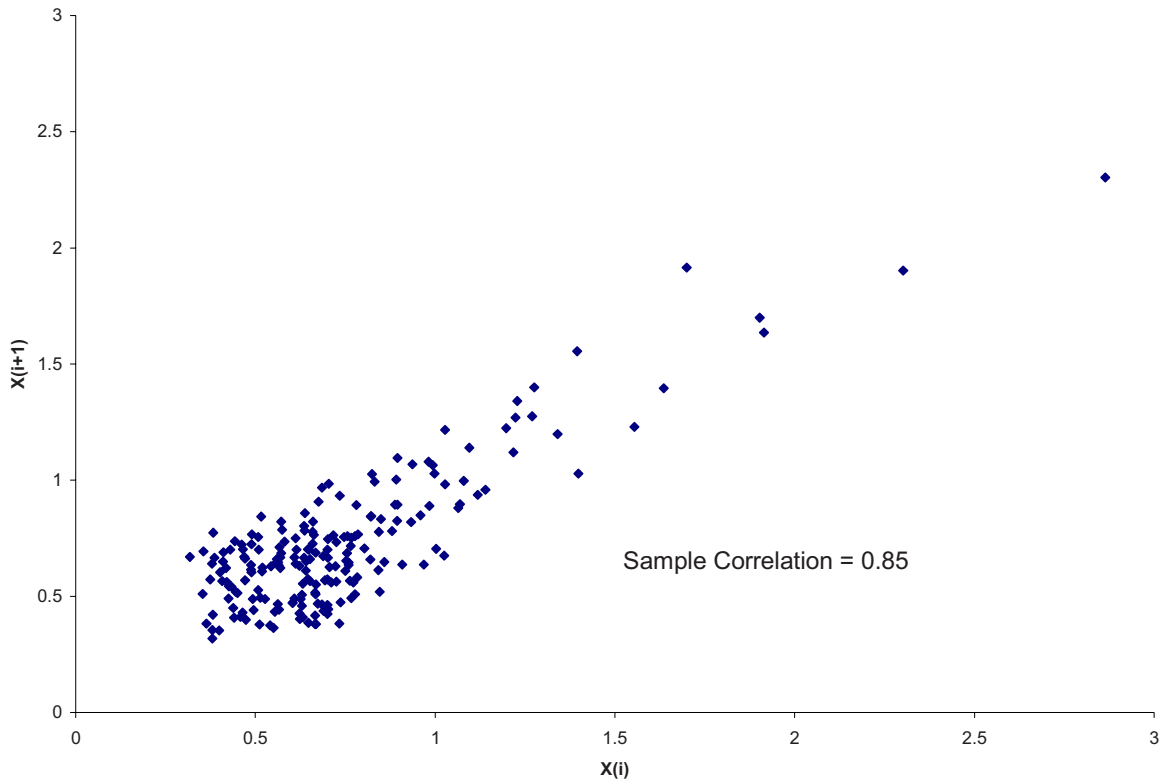

Fig. 7. Scatter diagram for valve installation hours

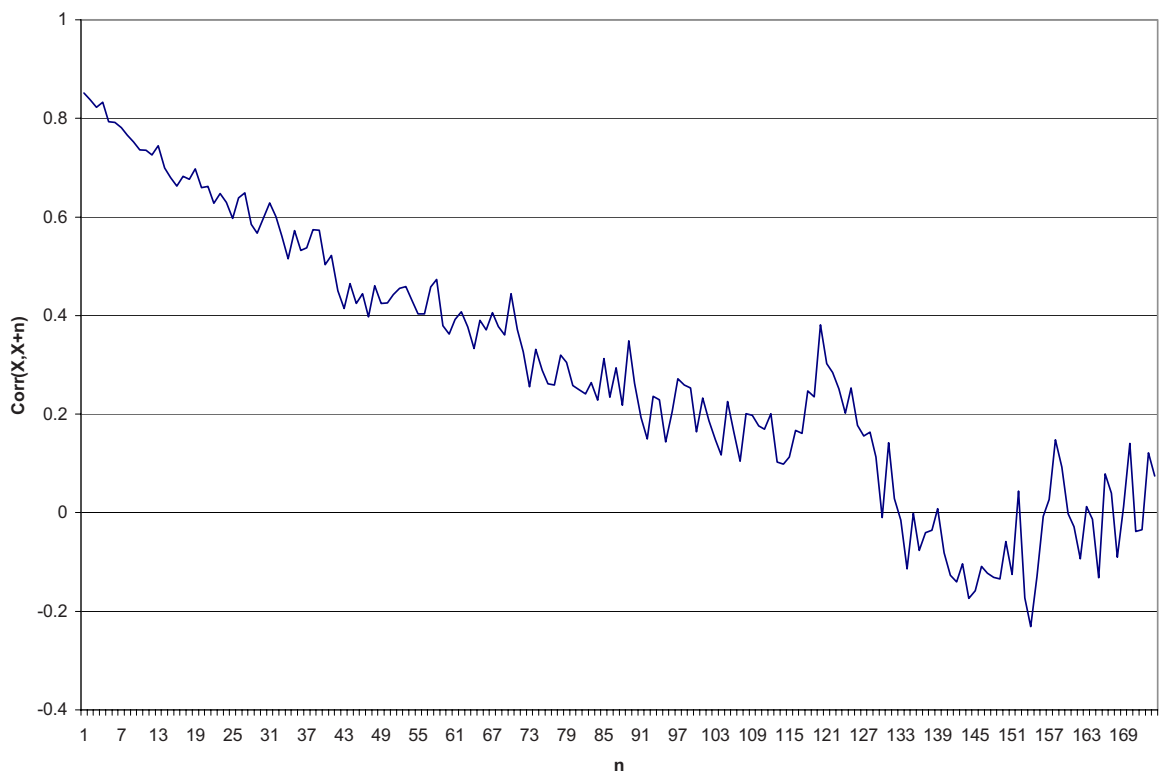

Fig. 8. Correlation plot for valve installation hours

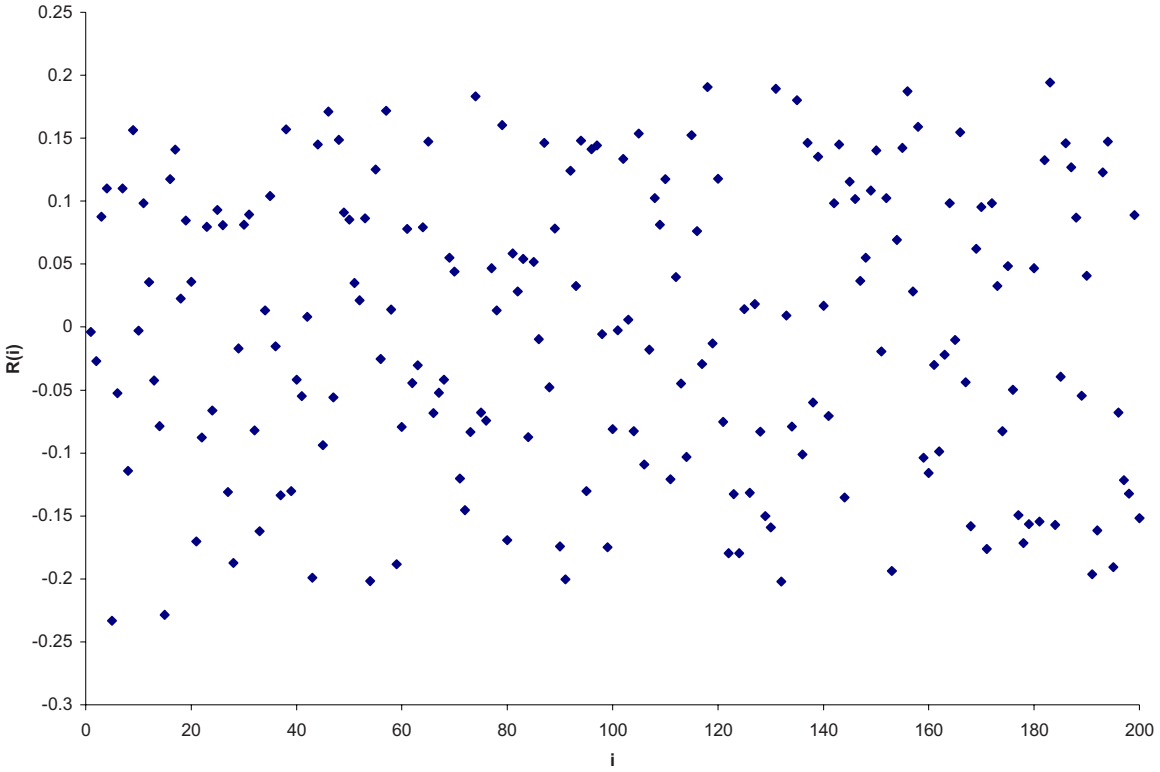

Fig. 9. Ordered plot of deviations

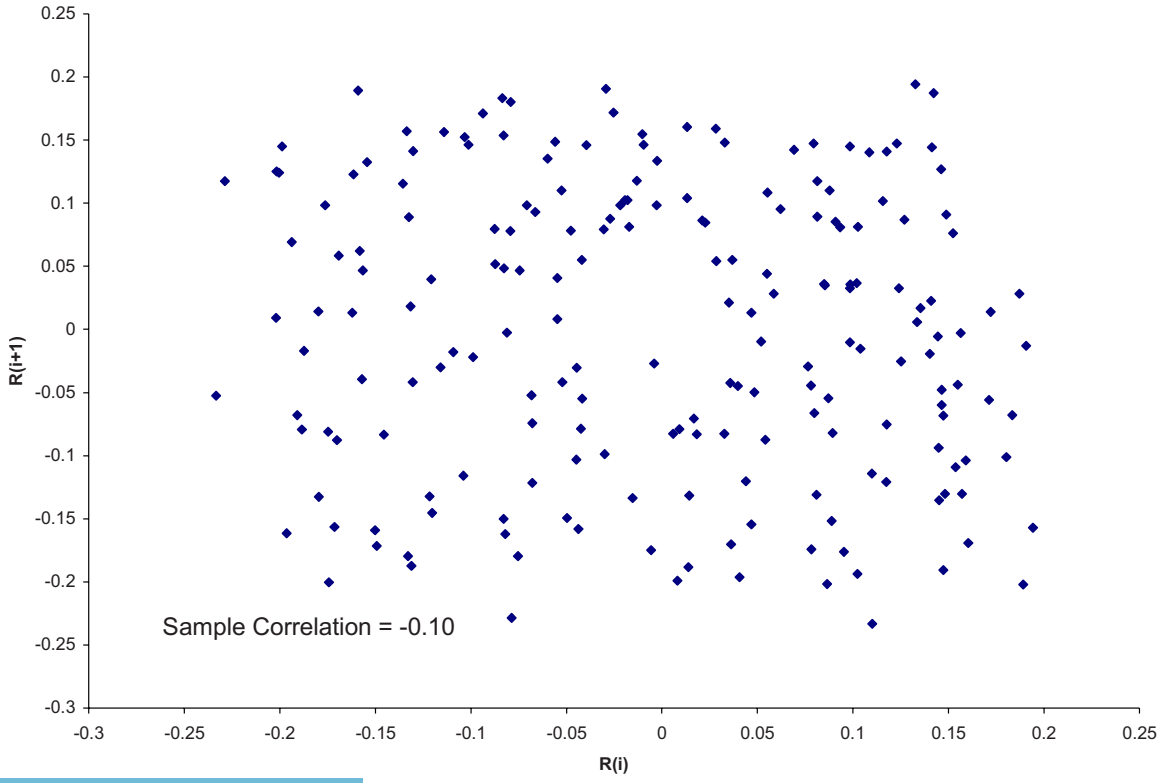

Fig. 10. Scatter diagram of deviations

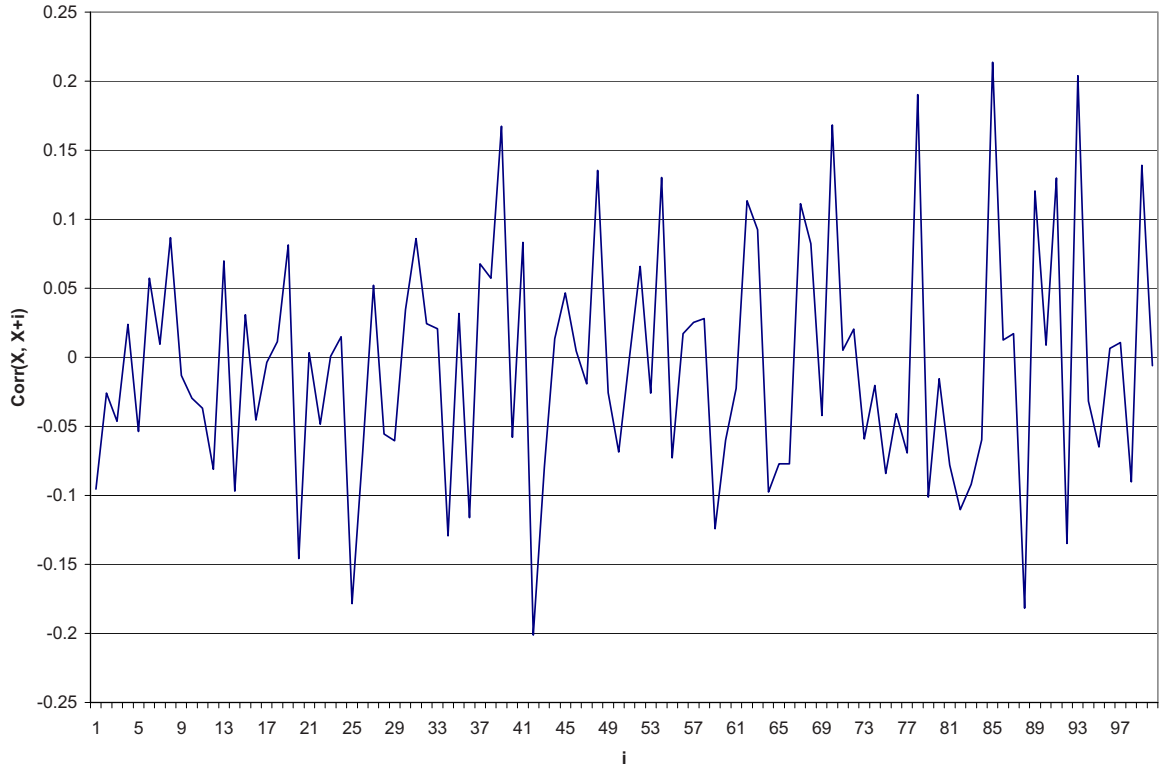

Fig. 11. Correlation plot of deviations

able in a variety of conditions. The knowledge of these experts may be significantly more extensive than what could be determined from a limited number of field observations. The problem is that this probabilistic knowledge is implicit—an expert cannot typically express this knowledge as a probability distribution. Probability encoding is a subject that has been researched extensively and that, when done properly, can be used to quantify and transfer probabilistic knowledge about uncertain quantities from experts.

At the very rudimentary level, that knowledge can be expressed by a distribution with parameters determined by eliciting from experts the most favorable, unfavorable, and commonly occurring (mode) values. According to Perry and Greig (1975), the most favorable time supplied by an expert can be assumed to be the 5th percentile and the most unfavorable time assumed to be the 95th percentile because most subjects have never actually experienced the extremes, or known so if they have. Thus, when asked for the most favorable time, the reply most experts provide is probably closer to the 5th percentile, and when asked for the most unfavorable time their reply is probably closer to the 95th percentile. Perry and Greig then make assumptions about the mean (that it is the weighted average of the mode and the 5th and 95th percentiles, with 5% less weight given to the mode) and standard deviation of the distribution (that it is the spread between the 5th and 95th percentiles divided by 3.25). These assumptions have been proven to be relatively distribution free and quite accurate. With these assumptions it is possible to determine the parameters, location, and scale, of, for example, a beta distribution. Perry and Grieg's procedure is very attractive because it requires the elicitation of only three values (as in Pert), is relatively accurate (much more than original Pert), and is implemented as a native probabilistic sampling function in some construction-oriented simulation tools (e.g., UM-Cyclone and Stroboscope).

AbouRizk et al. (1991) suggested a graphical computer based procedure to determine the beta distribution that best explains subjective expert knowledge by eliciting the minimum possible, maximum possible, and two additional values from the following pairs: mean and variance, mode and variance, mean and mode, or two arbitrary percentiles. While this procedure is 100% accurate if the four elicited quantities are 100% accurate, it is difficult for an expert to subjectively provide these quantities without biases and unconscious modes of judgment. The visual interactive display of the distribution can be of help, however, for experts who understand distribution probability concepts.

Spetzler and Von Holstein (1975) provided an excellent introduction to probability encoding that includes techniques to correct for biases and unconscious modes of judgment. More recently, Abbas et al. (2008) reviewed and compared some of the techniques that can be used to encode probabilities. These techniques

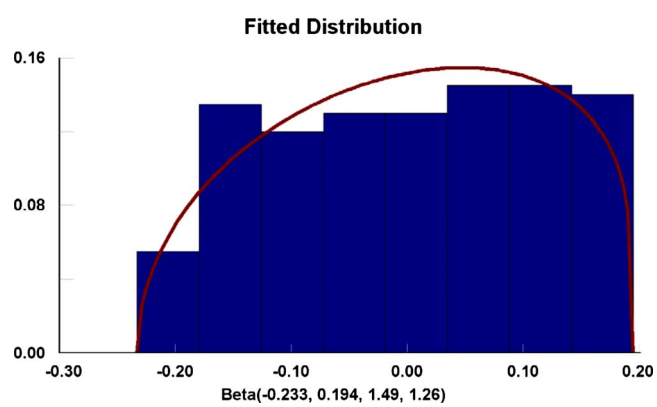

Fig. 12. Histogram and fit of deviation

should be used for important sources of uncertainty in a simulation model. Because construction operations are often condition sensitive, it makes sense to encode conditional probabilities from experts and to use these in conjunction with a simulation of the conditions.

Verification and Validation

Verification is the process by which the individual(s) that develop a model ascertain that they have modeled what they intended to model (i.e., model debugging). Some guidelines on how to do this are provided by Law and Kelton (2000) (pp. 269–273). It is very easy to make modeling mistakes in most useful models of complex systems, so verification needs to leave no doubt to the model developer regarding the correctness of the model.

In validation, the verified model and the actual or imaginary system are compared. This validation is with respect to logic, input, assumptions, and output. Since the model developers are frequently not the experts and/or decision makers (i.e., stakeholders and subject matter experts) in the actual operation, validation requires that the model be clearly communicated to the stakeholders, who can then assess many aspects of model validation. Animation is particularly valuable in this aspect of validation since it allows the stakeholders and experts to "see" a virtual version of the operation. Kamat and Martinez (2003) discussed this and presented a language and software that can be used to animate previously simulated operations in 3D.

Validation at the input level is concerned with the probability distributions used in the model. This includes ascertaining that basic probabilistic assumptions hold (e.g., the data are IID) and goodness of fit testing (e.g., Chi-Square, Kolmogorov-Smirnov, Anderson-Darling).

Validation at the output level (i.e., results validation) is where construction operations present interesting challenges when compared to other disciplines. This an area that needs considerable attention in future studies.

At the single run level, it needs to be established that the output of the model resembles the output that would be expected from the actual or imaginary system (Law and Kelton 2000, p. 279). In construction, the systems being modeled often do not exist (but occasionally do). This makes comparison with the real operation, prior to its execution, impossible. In these cases, the simulation results need to be checked for reasonableness by experts, or perhaps using Turing tests (Carson 1986). If the system exists (e.g., when the modeling has the objective of improving it), the model can potentially be compared statistically to the actual system output, but this is very difficult unless the operation will be repeated a number of times.

A single simulation experiment yields one set of outputs (i.e., cost, equipment utilization, production rates, inventory levels) that correspond to the exact sequence of events that took place in a simulation. Because simulations are probabilistic, a second independent simulation of the same model can, and should, yield different results. For this reason, a simulation model should be run several times in order to obtain (ideally) the distribution of each output. Design of experiments refers to the process of determining how many replications (and how to set up the replications) of a model simulation to perform in order to achieve a certain objective. Design of experiments also involves the creation of proper simulation conditions, ensuring that the simulation components assumed to be IID are indeed simulated as IID, and ensuring independence between different simulation runs. For example, the objective may be to obtain a 90% confidence interval on the mean cost of construction that is smaller than a certain dollar amount. To do this, a few runs of the model could be used to estimate the standard deviation of cost, and then this value could be used to calculate the total number of replications that are required. Alternatively, if the simulation software allows, it should be possible to replicate automatically until the objective is reached. Ioannou and Martínez (1996) discussed in detail the design of experiments for evaluation of alternative tunnel construction strategies. The information in that source is useful and readily adaptable to any type of simulation modeling whose objective is to compare among alternatives. For general cases of broader applicability, the reader should refer to Law and Kelton (2000).

Table 1. Number of Trucks Loads per Day in Earthmoving

Date	Loads
May 17, 1997	223
May 19, 1997	264
May 20, 1997	240
May 22, 1997	276
May 23, 1997	217
May 27, 1997	206
May 28, 1997	272
May 29, 1997	234
May 30, 1997	280
May 31, 1997	248

In contrast to most manufacturing operations, due to the uncertainties and complexities that characterize construction, outputs typically exhibit significant variance. Table 1, for example, shows the number of loads that a contractor was able to achieve each of 10 days when moving dirt for the construction of a road. The data are from Cor (1998).

Despite the nearly constant haul distance, weather, and the use of the same fleet, production varied significantly due to complexities and uncertainties related to right to use a temporary railroad crossing and to equipment breakdowns. The fact that these complexities existed is what made the use of DES particularly valuable in that case compared to the standard tools used by the contractor. A valid simulation of this operation should exhibit a similar variability between independent runs. Operations where the output has little variability are typically relatively easy to analyze using other methods (i.e., DES is overkill for them).

For most decision making purposes, it is in fact more important to understand the variability than the mean measure associated with an output. In validation, an expert should assess whether the distribution of the output is reasonable. This aspect of validation is where many DES studies in construction fail—frequently only a single run is reported and analyzed and other times the variance of the output is significantly smaller than could be expected.

Variability of output makes it also very difficult to validate a simulation model by comparing the result in the field to what the simulation had indicated. Typically the field observation is only one data point that could have come from any number of output distributions. If the nature of construction were such that the operation could be carried out independently several times (no learning effect, etc.), then it would be possible to perform statistical tests (i.e., a goodness of fit test) to determine the likelihood that the observations come from the simulation output distribution. (An ongoing earth moving operation with constant condi-

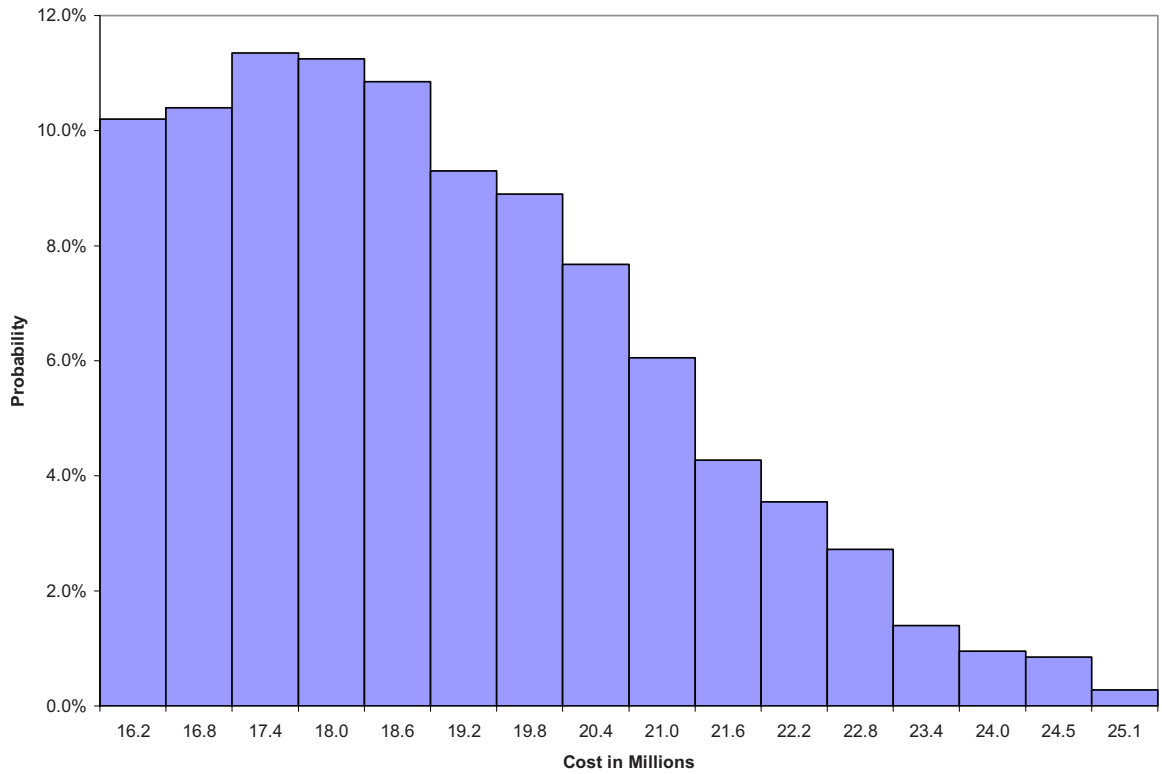

Fig. 13. Distribution of tunnel construction cost using NATM

tions could be an example to the contrary, if production is looked at on a daily basis. However, conditions rarely remain constant for long.) Fig. 13 shows the output distribution for the cost of a tunnel construction operation using the New Austrian Tunneling Method (NATM). This is from Martínez and Ioannou (1996) and was built using 4,000 independent replications. Here, the simulation output is a distribution with very high variance that ranges from a low $16 million to a high $25.4 million. The output is very useful for understanding the large risks associated with this operation. However, if the tunnel is built exactly as modeled (the model is 100% valid), the actual cost of construction will be anywhere within the range. If the contractor finds mostly good ground, the cost may be $17 million or lower. If the contractor finds mostly poor ground, the cost may rise to $23 million or higher. Whatever the actual cost, it cannot be used to prove or disprove that the model was valid. If the model was invalid, and it were possible to build (independently) several dozen identical tunnels, then it would be statistically possible to prove at some level of significance that an invalid model is indeed invalid (i.e., there would be enough evidence to reject the hypothesis that the observed costs come from the simulation model's output distribution). *It would be totally incorrect to state something similar to* "*proof of the correctness of model is that the tunnel was built at a cost of $19.4 million, which is close to the mean output of the simulation, $19.1 million.*"

Validation of construction operation studies is based predominantly on having experts convinced, *to the point that they will rely on the model to make decisions*, that the model is an accurate representation of the real or future operation for the purpose for which the model was built. In other words, a model is considered to be valid for a particular purpose if it is credible to experts for that purpose.

Relationship between the Purpose of a Study, the Rigor It Requires, and Its Validation

Performing a DES study that follows the appropriate rigor in every one of its components may represent a very substantial effort. Carrying out a DES study does not require that every single part be performed at full rigor. However, what is necessary, and cannot be circumvented, is a full understanding of the assumptions that are made and their impact in the validity of the model for the purpose for which it was built. A similar concern was expressed by Schexnayder (1997).

When using a DES study in construction practice to design an operation, sensitivity analysis can be used to determine which inputs or portions of the operation impact the performance measures of interest in a significant way. The engineer can then concentrate on modeling those parts at a reasonable level of detail and to follow all the probability theory that is pertinent rigorously. For parts that are inconsequential to the purpose of a model, flexibility may be used to model that input (e.g., not too much detail, treat non-IID data as IID). When the purpose of a model is to compare among various alternatives, emphasis should be on the parts of the model where the alternatives differ.

When DES is used to demonstrate some principle or the effectiveness of some technique, the emphasis should be on capturing the principle faithfully. Input distributions need to make sense and be reasonable to individuals knowledgeable in the area of application, but they do not need to be based on collected data. If data are available from other sources, then it can be used. Tommelein (1998), for example, used 10% as the probability of a pipe spool needing rework in models to illustrate pull scheduling, because the range 1–10% had been indicated by Howell and Ballard

(1996) in a prior study and she wanted to demonstrate her model under the largest realistic uncertainty. This was consistent with the purpose of the model used by Tommelein, which was to illustrate the impact of coordination planning and the benefits of pull over push when uncertainty is high.

Many DES models exist solely to demonstrate a research product that in some ways claims to advance the state-of-the-art in simulation modeling. This is another case where the specific inputs and details that are used are of no significant consequence (unless the research is related to data synthesis)—they must simply be reasonable in order to give credibility to the work. Claims are typically that:

1. Some aspect of the modeling process can be done more accurately without relying on simplifications that were otherwise required;
2. It can be done in a meaningful way following a thought process that differs from those that already exist; and
3. It can be done with less training or effort without reducing its other characteristics.

According to De la Garza (2007), in order for research findings to withstand the scrutiny of peers, the work needs to be described at a level of detail that will enable others to replicate it and confirm the claimed findings.

It is critical that complete data and details (input and output) of the examples used to illustrate a new tool or method be available to others (or at least to reviewers), so that they can perform alternate modeling and compare results. This has the added benefit of clarity with respect to the claims being made by the research. When the full data set, details of the model, and output are not available; even trivial endeavors can appear to be grandiose. Researchers should have developed enough confidence in the validity and true value of their contributions to not consider it a risk to provide this information. The examples used to demonstrate a procedure need not be real or confidential, or large but lacking complexity. In fact, a very significant part of the validation methodology should include the development of small cases that can be described and grasped in a reasonable amount of space, but that fully exercise the claims set forth.

Discrete-Event Simulation as Part of Higher-Level Modeling Systems

The focus of this paper has been on the methodology to follow in the types of DES research that have been traditionally published in the JCEM. Specifically, simulations of construction operations where the sequence of events, input distributions, statistical design of simulation experiments, and statistical output analysis are to a large extent controlled and defined by the modeler. In this type of modeling, whether for a case study, to prove or disprove some principle, or to demonstrate a new modeling concept; it is possible to create models that contain logic errors, that are based on incorrect statistical assumptions, or that simply do not capture the real variability of outcome that characterizes construction.

Due to the requirements for rigor, or necessity to relax rigor only when it follows a full understanding of the consequences of doing so, DES research and application of this type requires significant education and discipline. In recognition that many field engineers do not have this specialized knowledge and ability, some research effort has attempted to answer the question of how to make this technology accessible to such field personnel. Special purpose simulation (SPS) has been proposed a possible answer to this problem.

Seppanen (1990) provided early guidelines for the development of SPSs in manufacturing using general purpose simulation languages. In the construction engineering domain several SPS tools have also been developed. Systems such as Simphony (Hajjar and AbouRizk 1999) have been designed specifically for the purpose of building such tools, and general purpose programmable and extendable construction simulation systems such as Stroboscope (Martinez 1996) can also be used to create SPSs.

In the case of construction operations, SPS systems allow practitioners to define DES models by working with modeling elements from the domain of the SPS tool. In an earthmoving SPS tool, for example, it may be possible to define a model by drawing haul routes and specifying their properties (e.g., grade, rolling resistance), by selecting trucks and excavators from databases; and specifying the location of excavators, routes taken by trucks, and quantities of cut and fill. The SPS tool will use this information to create a DES model that internally contains the equivalent of activities such as load, haul, dump, and return; with parameters such as time distributions totally or partially determined from the properties of the objects used in the SPS tool. In these cases, the relationship between the SPS modeling elements and the underlying lower-level DES model is very direct. For example, deterministic travel times may be calculated from truck and haul route properties using established engineering calculations, and these may be made stochastic with input specified directly by the user (e.g., users may provide a distribution by which the calculated travel time will be multiplied to obtain the time for each haul). In any case, the ability to configure the model very precisely to the exact situation at hand is dependent on how well the processes and tools built into the SPS that are used to create the underlying model can match the current situation. Conceivably, the SPS tool can allow some degree of configuration to enable modelers to achieve this to some extent. A researcher creating an SPS answers questions related to how to define a model using domain specific elements, and how to use this information to create the corresponding lower level DES. There is a responsibility on the part of the researcher to explain very clearly how this is done, so that users can understand what is being modeled and so they can judge whether this will be satisfactory for the case at hand (i.e., users should be able judge the extent to which all the issues that should be addressed in a DES study are considered in the SPS).

It is possible to extend the SPS concept to domains that are not related to construction operations modeling. One interesting example is the virtual design team (VDT), which is a computational organization model that incorporates elements of DES, artificial intelligence, and computational organization theory (Levitt et al. 1999). When creating VDT models, users work with project-level activities [e.g., activities similar to those that would be used in a critical path method (CPM) schedule] and specify for them, for example, the required work volume, work skill, activity-level uncertainty, activity complexity, and interdependence strength to other activities. These characteristics are specified in a descriptive rather than normative manner. The knowledge embedded in the VDT then creates a DES that in addition to the actual work, models communications, disruptions, and coordination based on these parameters. The underlying DES model includes not only the direct work activities, but also the interruptions for communication and exchange of information among actors. The process by which the project-level activities and their descriptive parameters translate into DES is based on a microcontingency theory developed and validated by the VDT team for that purpose. According to Levitt et al. (1999) (p. 5), a microcontingency theory is "a theory to describe and predict behavior at the level of individual actors and activities." Because the relationship between the mod-

eling elements of a VDT model and the underlying DES is not direct, but instead based on its microcontingency theory, many of the considerations needed to properly conduct DES studies of construction operations are simply not applicable to a researcher using such a tool since there is a shift in responsibility for the modeler from the normative to the descriptive. On the other hand, the tool itself and the knowledge embedded in it need to undergo a very extensive and substantial validation process in order for users to trust it. The validation process followed for the VDT is documented in Thomsen et al. (1999).

Summary and Conclusions

The process for conducting a DES study in construction does not differ from the one used in other disciplines. Certain aspects, however, warrant significant consideration in construction as compared to studies in other disciplines.

Heavy dependency on conditions in the performance of activities, and the variability of these conditions during the simulated period, make many uncertain quantities (e.g., activity times) to not be IID. Since the assumption of IID is necessary for fitting distributions to data, in construction these uncertain quantities must be expressed as functions of other quantities that are IID and that can in turn be appropriately described by probability distributions.

Because the use of DES studies in construction practice is to make decisions prior to field execution of the modeled operations, it is not possible to validate models by comparing model output to real output before the models are used. In addition, since the operations subject of the study are frequently carried out only once and exhibit significant variability among independent outputs, it is not possible to validate models even after execution in the field. Validation must be carried out by judgment of experts, stakeholders, and decision makers who need to be convinced that the model and the operation are one and the same for the particular purpose for which the model was made. Evidence of validity is reliance on the model for making decisions.

It is not generally possible to perform a DES study of a construction operation that follows full rigor in each and every aspect. Assumptions need to be made that are known to not be 100% valid. However, this relaxation of rigor cannot take place blindly. An in-depth understanding of probability and statistics that are the basis for DES, and sound engineering judgment, are critical elements that cannot be circumvented.

In addition to use for the study of operations, DES studies can be used to prove or demonstrate theories and practices and also to illustrate research that claims to advance the state-of-the-art in construction DES. DES studies used to demonstrate new research advances should allow readers to verify, independently, the claims of the research. This often requires that significant effort be placed on preparing the modeling examples so that they are small, showcase the research claims, and are fully detailed (including all inputs and outputs).

References

Abbas, A. E., Budescu, D. V., Yu, H. T., and Haggerty, R. (2008). "A comparison of two probability encoding methods, fixed probability vs. fixed variable values." *Decis. Anal.*, 5(4), 190–202.

AbouRizk, S. A., Halpin, D. W., and Wilson, J. R. (1991). "Visual interactive fitting of beta distributions." *J. Constr. Eng. Manage.*, 117(4), 589–605.

Abudayyeh, O., Dibert-DeYoung, A., and Jaselskis, E. (2004). "Analysis of trends in construction research: 1985—2002." *J. Constr. Eng. Manage.*, 130(3), 433–439.

Balbontin-Bravo, E. (1998). "Simulation of large precast operations." *Proc., 1998 Winter Simulation Conf.*, IEEE, Piscataway, N.J., 1311–1377.

Benneyan, J.C. (1998). *Software review: Stat::Fit, OR/MS Today*, 25(1), Informs—Institute for Operations Research and the Management Sciences, Hanover, Md.

Carson, J. S. (1986). "Convincing users of model's validity is challenging aspect of modeler's job." *Industrial Engineering*, 18, 74–85.

Cor, H. (1998). "Using simulation to quantify the impacts of changes in construction work." MS thesis, Virginia Tech, Blacksburg, Va.

De la Garza, J. M. (2007). "Sponsored research and its impact on universities, faculty, and journals." *J. Constr. Eng. Manage.*, 133(9), 708–709.

Gibbons, J. D. (1985). *Nonparametric methods for quantitative analysis*, 2nd Ed., American Sciences, Columbus, Ohio.

Hajjar, D., and AbouRizk, S. (1999). "Simphony: An environment for building special purpose construction simulation tools." *Proc., 1999 Winter Simulation Conf.*, Vol. 2, IEEE, Piscataway, N.J., 998–1006.

Halpin, D. W., and Martinez, L. (1999). "Real world applications of construction process simulation." *Proc., 1999 Winter Simulation Conf.*, Vol. 2, IEEE, Piscataway, N.J., 956–962.

Hildreth, J., Vorster, M. C., and Martinez, J. C. (2005). "Reduction of short-interval GPS data for construction operations analysis." *J. Constr. Eng. Manage.*, 131(8), 920–927.

Hooper, J. W. (1986). "Strategy related characteristics of discrete-event languages and models." *Simulation*, 46(4), 153–159.

Howell, G. A., and Ballard, H. G. (1996). "Managing uncertainty in the piping process." *Rep. No. RR 47-13*, Construction Industry Institute, Univ. of Texas, Austin, Tex.

Inlow, M. (2006). "Teaching the IID assumption in engineering statistics I." *Proc., 2006 American Society of Engineering Education Annual Conf.*, American Society of Engineering Education, Washington, D.C., 2006-1397, ⟨http://soa.asee.org/paper/conference/paper-view.cfm?id=1599⟩.

Ioannou, P. G., and Martínez, J. C. (1996). "Comparison of construction alternatives using matched simulation experiments." *J. Constr. Eng. Manage.*, 122(3), 231–241.

Kamat, V. R., and Martinez, J. C. (2003). "Validating complex construction simulation models using 3D visualization." *Syst. Anal. Model. Simul.*, 43(4), 455–467.

Kannan, G., and Vorster, M. C. (2000). "Development of an experience database for truck loading operations." *J. Constr. Eng. Manage.*, 126(3), 201–209.

Law, A. M., and Kelton, W. D. (2000). *Simulation modeling and analysis*, 3rd Ed., McGraw-Hill, New York.

Levitt, R. E., Thomsen, J., Christiansen, T. R., Kunz, J. C., Jin, Y., and Nass, C. (1999). "Simulating project work processes and organizations: Toward a micro-contingency theory of organizational design." *Manage. Sci.*, 45(11), 1479–1495.

Maio, C., Schexnayder, C., Knutson, K., and Weber, S. (2000). "Probability distribution functions for construction simulation." *J. Constr. Eng. Manage.*, 126(4), 285–292.

Martinez, J. C. (1996). "STROBOSCOPE: State and resource based simulation of construction processes." Ph.D. dissertation, Civil Engineering, Univ. of Michigan, Ann Arbor, Mich.

Martínez, J. C. (2001). "EZStrobe—General-purpose simulation system based on activity cycle diagrams." *Proc., 2001 Winter Simulation Conf.*, IEEE, Piscataway, N.J., 1556–1564.

Martínez, J. C., and Ioannou, P. G. (1999). "General purpose systems for effective construction simulation." *J. Constr. Eng. Manage.*, 125(4), 265–276.

Perry, C., and Greig, I. D. (1975). "Estimating the mean and variance of subjective distributions in pert and decision analysis." *Manage. Sci.*, 21(12), 1477–1480.

Schexnayder, C. J. (1997). "Discussion: Analysis of Earth-moving systems using discrete-event simulation." *J. Constr. Eng. Manage.*,

123(2), 199–199.
Seppanen, M. (1990). "Special purpose simulator development." *Proc., 1990 Winter Simulation Conf.*, IEEE, Piscataway, N.J., 67–71.
Shannon, R. E. (1998). "Introduction to the art and science of simulation." *Proc., 1999 Winter Simulation Conf.*, IEEE, Piscataway, N.J.
Spetzler, C. S., and Von Holstein, S. (1975). "Encoding in decision analysis." *Manage. Sci.*, 22(3), 340–358.
Thomsen, J., Levitt, R. E., Kunz, J. C., Nass, C. I., and Fridsma, D. B. (1999). "A trajectory for validating computational emulation models of organizations." *Comput. Math. Org. Theory*, 5(4), 385–401.
Tommelein, I. D. (1998). "Pull-driven scheduling for pipe-spool installation: Simulation of lean construction technique." *J. Constr. Eng. Manage.*, 124(4), 279–288.