Proceedings of the 1999 Winter Simulation Conference
P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, eds.

DEVELOPING THE STATISTICAL PARAMETERS FOR SIMULTANEOUS VARIATION IN FINAL PAYLOAD AND TOTAL LOAD TIME

Govindan Kannan
Michael C. Vorster
Julio C. Martinez

Charles E Via Jr. Department of Civil Engineering,
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0105, U.S.A.

ABSTRACT

It is argued that the two principal components of simulating earthmoving operations are modeling and providing input data. The knowledge and proficiency in modeling has been substantially improved through general purpose simulation systems such as CYCLONE and STROBOSCOPE. However, the inability to provide statistically reliable input data to the models is source of concern. Automated data collection through instrumented vehicles provides the first opportunity to collect cycle time data in a continuous manner and hence, provide a statistically-reliable data set. This paper presents the use of automated data to fit probabilistic distributions that are used as input to simulation models. This paper also describes a methodology to develop statistical parameters for simultaneous variation in payload and load time through the concept of a payload time (PLT) Map. Field data from on-going earthmoving projects is used to illustrate these concepts.

1 INTRODUCTION

Earthmoving operations can be divided into four elemental activities, namely, load, haul, dump and return. Each activity is associated with a probabilistic estimate of the duration. The load activity is the only one that is associated with two variables – load time and payload. The inter-relationship between payload and load time makes the load activity more complicated than the other activities. The challenge is to model the simultaneous variation in payload and load time. This paper addresses the simultaneous variation in payload and load time by developing a parametric form of payload in terms of load time.

Simulating the load activity requires the ability to model the process as well as the data reduction capability to support the modeling constructs. The data reduction process for the simultaneous variation in payload and load time involves the development of a joint probability function of payload and load time. The direct development of the joint probability function is complicated. Therefore, the joint probability function is described using the marginal probability of load time and the conditional probability of payload given load time. The modeling of the correlated payload and load time is demonstrated using STROBOSCOPE, a general-purpose simulation system.

2 BACKGROUND

Necessary background research can be grouped under three areas: a) modeling environment, b) modeling the load activity, and c) vehicle instrumentation. Previous research pertaining to each one of these areas of research is discussed under the appropriate sub-heading.

2.1 Modeling Environments

One of the earliest developments in the modeling environment for construction simulation was CYCLONE [Halpin 1977]. CYCLONE can model and simulate repetitive construction processes that are cyclical. A number of researchers have extended the capability of CYCLONE since then, the notable enhancements being the development of MicroCYCLONE — a microcomputer edition [Lluch and Halpin 1982] and UM-CYCLONE at University of Michigan [Ioannou 1990]. INSIGHT [Kalk 1980] is based on CYCLONE and includes time-lapse photography for data acquisition, graphics support and an interactive environment [Paulson *et al.* 1987]. Since then, there have been several other simulation programs developed. A recent addition to this list is the development of STROBOSCOPE [Martinez 1996], a comprehensive and sophisticated tool that incorporates end-user programmability and extensibility. Martinez and Ioannou [1995], and, Ioannou and Martinez [1996] illustrate the principles of construction simulation and capabilities of STROBOSCOPE.

Most of the modeling environments mentioned here use an activity-based approach which is conducive to modeling construction operations [Martinez and Ioannou 1999]. The typical approach to modeling an operation involves specifying an activity cycle diagram and the corresponding probabilistic estimates for the durations of the activities. Modeling the load activity, however, involves correlated variability among load time and payload. Modeling this requires programmability and thus necessitates the use of STROBOSCOPE.

It must also be pointed out that there are commercially available simulation environments such as GPSS/H and ProModel which allow for programmability. However, most of these environments are based on a process interaction (PI) strategy wherein a model is written from the point of view of entities that flow through a system. The PI strategy is very effective in modeling industrial applications but presents substantial challenges when used to model construction applications where there is a heavy interaction between resources [Martinez and Ioannou 1999].

The above account suggests that current knowledge in modeling is capable of realistically replicating construction activities. The implicit assumption made by various researchers in formulating their problem is that sufficient data is readily available. AbouRizk and Halpin [1992] have shown that input data is of paramount importance in simulating construction operations.

2.2 Modeling the Load Activity

The load activity as part of the earthmoving cycle is applicable to both scrapers and loader-truck operations. Day [1973] presents a load-growth curve wherein the payload is described as a function of the load time and is used as the basis of a non-linear optimization process. These curves have been specifically used for scraper operations although their scope can be extended to loader-truck systems [Gransberg 1996]. The load-growth curve for a loader-truck operation would have step-increments instead of a continuous curve as in the case of the scrapers. Although the load-growth curves describe the relation between the incremental payload and the incremental load time, they do not include the variation involved in either payload or load time.

Martinez [1996] and Ioannou and Martinez [1996] show how to model correlated variation in payload and load time in STROBOSCOPE by using the marginal of payload and the conditional of load time given the payload. For illustration purposes, those examples use empirical mathematical formulas that apply to a loading unit pass rather than to the entire multi-pass loading activity of a loader-truck operation.

2.3 Vehicle Instrumentation

Vehicle instrumentation presents a viable opportunity for continuous productivity data collection. The vehicle instrumentation could include machine health diagnostic systems, global positioning systems and radio frequency identification systems. Machine health diagnostic systems were used to collect performance data for the purpose of providing input to simulation models. Performance records refer to the activity durations of different component of the earthmoving cycle. Different equipment manufacturers have their own proprietary onboard instrumentation systems. Examples of proprietary systems currently in use include VIMS® & TPMS® (Caterpillar [http://www.cat.com]), CONTRONICS® (Volvo [http://www.vce.volvo.se]) and HMS® (Komatsu [http://www.komatsuamerica.com]).

3 MODELING THE LOAD ACTIVITY

The load activity can be measured using two quantitative measures, namely, the final payload and the total load time. Final payload refers to the payload of material placed on the truck as it leaves the load area and the total time reflects the time taken to achieve the final payload. Measuring the final payload and the total load time for the purpose of modeling the uncertainty in these measures requires a continuous form of data collection. Vehicle instrumentation provides the first opportunity to continuously and to autonomously collect data on final payload and the total load time. Therefore, the primary assumption is that instrumented data on the load activity are available.

3.1 Payload and Load Time as Recorded by the Vehicle Instruments

Two possible scenarios can be considered in modeling the load activity. The first scenario is that the final payload and the total load time are independent of one another and the second scenario is that the final payload and load time are correlated. The second scenario may appear more intuitive because of the knowledge of load-growth curves as described for a scraper operation. The data used by the concept of load-growth curves reflects the incremental time and the corresponding payload. The values for payload and load time recorded by the vehicle instruments are registered at the time when the truck leaves the load area. It is not possible to capture the “growth” process using the instrumented data.

3.2 Independent PDFs for Payload and Load Time

The two variables, payload and load time are treated as independent variables and an appropriate probability density function (PDF) is defined for each variable. A standard distribution that best defines the payload is determined using standard procedures [Law and Kelton 1991]. The probability of obtaining a particular range of final payload is modeled by Eq (1).

$$P(U_p > p > L_p) = \int_{L_p}^{U_p} f(p)dp \tag{1}$$

where *P(x)* is the probability of a random variable *x*, L_p and U_p are lower and upper limits of the range of payload, *p* is the payload, and *f(p)* is the PDF that defines the variation in payload for the entire range.

Figure 1 shows a graphical representation of Eq(1) wherein the integral of the PDF is calculated as the area enclosed by the cumulative frequency curve within the specified limits. Since all discrete values of payload are theoretically possible, the variable payload is treated as a continuous random variable. The probability of obtaining a single value of payload is zero; hence, the need to use a range to determine the probability.

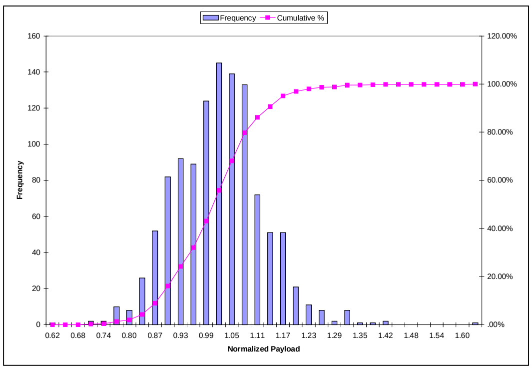

Figure 1: Frequency Distribution of Payload

The probability of obtaining a particular range of total load time can be developed in a manner similar to that used for payload and is given by Eq (2).

$$P(U_t > t > L_t) = \int_{L_t}^{U_t} f(t)dt \tag{2}$$

where L_t and U_t are lower and upper limits of the range of load time required, *t* is the load time, and *f(t)* is the PDF that defines the variation in load time for the entire range.

Figure 2 presents the graphical form of Eq (2) wherein the integral of the PDF is calculated as the area enclosed by the cumulative frequency curve within the specified limits. It is again noted that since all discrete values of load time are theoretically possible, the variable load time is also treated as a continuous random variable.

The joint probability of obtaining a combination of payload and load time is the product of the individual probabilities under the assumption of independence. The joint probability calculated in this manner will depend on the distributions that define the payload and load time. It is also possible that more than one unique distribution can be used to define specific intervals of payload and load, which makes the situation more complicated. Moreover, it is not intuitive to understand the simultaneous variation in payload and load time if defined by the two independent PDFs. A powerful graphical format called a PLT Map is used to graphically present the simultaneous variation in payload and load time.

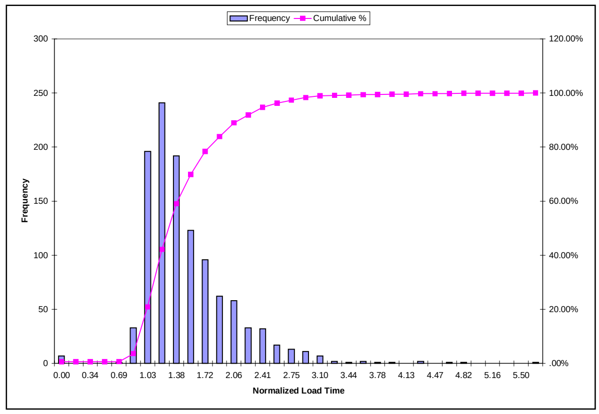

Figure 2: Frequency Distribution of Load Time

3.3 PLT Map

A PLT Map (Payload- Load Time Map) describes the joint frequency of payload and load time. An example of a PLT Map is shown in Figure 3. The abscissa is represented by final payload and the ordinate by total load time. The values are normalized to maintain a consistent scale. The contour plot uses color or pattern to indicate areas with equal probability, the legend for which is shown above the plot. For example, the payload-time pair (1-1.1, 1.1-1.3) occurs with a probability between 9% and 10%. The payload-time pairs represented by a particular color or pattern have equal probability of occurrence. The shape or footprint of the plot, the location of the modal region and the area occupied by each color or pattern provide interesting insights into the operation.

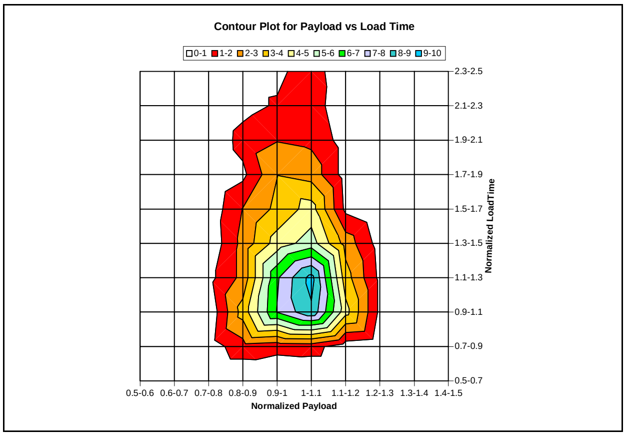

Figure 3: An Example of a PLT Map

The joint probability of obtaining a particular range of payload and load time is given by Eq (3).

$$P(U_t > t > L_t, U_p > p > L_p) = \int_{L_t}^{U_t}\int_{L_p}^{U_p} f(t,p)dpdt \qquad (3)$$

where $f(t,p)$ is the surface which is depicted by the PLT Map that represents the joint frequency distribution of load time and payload. The frequency of a pair of payload and load time is normalized and so the volume under the surface is equal to one. Therefore, the color or pattern corresponding to a particular combination of payload and load time represents the probability of occurrence of that pair.

It is possible to imagine that the PLT Map is a series of continuous payload distributions, each conditioned on a given range of load time as shown in . A similar set of distributions can be created for load time conditioned on ranges of payload.

3.4 Joint Probability Distribution in a Parametric Form

The significance of the independence assumption is that both payload and load time can be independently sampled from their respective distributions. It is possible to generate a high value for load time and a low value for payload for an instance of the load activity and vice versa. It is possible that a longer load time could generate a higher payload. The most elemental way to determine if a correlation exists is to observe the PLT Map. If the PLT Map has elliptical regions and the ellipses makes an angle with the payload axis, then there exists a relation between final payload and total load time. This relation is strong if the ellipses make a 45° angle. The discussion under this heading makes use of the assumption that payload is dependent on load time.

The procedure under this assumption is based on marginal distributions of load time and conditional distributions of payload given load time (discretized). This is because the probabilistic models needed to describe the joint distribution using this approach are more tractable than an approach in which the marginal of load time and the conditionals of payload are used even though the latter is directly supported by STROBOSCOPE.

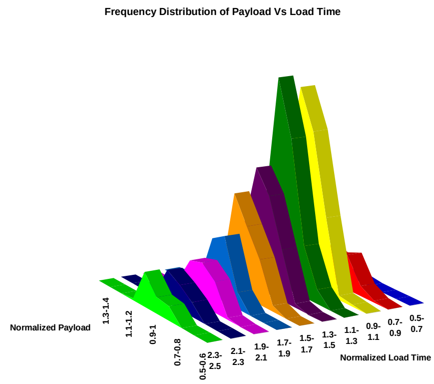

Figure 4: Payload Distribution for Each Discrete Interval of Load Time

First, the load time — treated as the independent variable — is modeled as a continuous distribution. AbouRizk and Halpin [1992] used a transformed method to determine the statistical properties of construction data. They used the values of skewness and kurtosis to determine the type of distribution associated with various construction activities. From the field data shown in Figure 5, it is clear that most instances of the load time of a truck can be modeled by a Beta distribution, confirming the findings of AbouRizk and Halpin [1992]. Let it be assumed that load time is modeled as a Beta distribution as shown in Eq (4).

$$P(t) = Beta\ [a, b, s_1, s_2] \qquad (4)$$

where $P(t)$ is the PDF describing the load time, a is the lower bound of the distribution, b is the upper bound of the distribution, s_1 and s_2 are scale and shape factors respectively of a Beta distribution.

The second step is to parametrically define payload in terms of the load time and at the same time capture the variance in payload. Figure 6 suggests the distribution corresponding to payload could be modeled as a Normal distribution. For reference, a line corresponding to Log-Normal distribution is shown in the figure. It can be noted that the Log-Normal offers the next best alternate distribution to model the data.

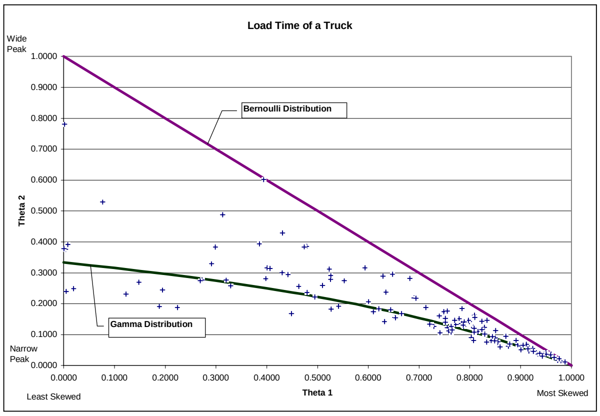

Figure 5: Distribution to Model Load Time Data

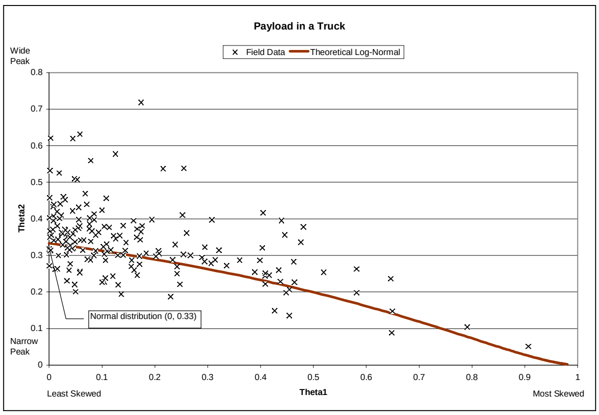

Figure 6: Distribution to Model Payload Data

The intent is to develop equations that define the mean and standard deviation of a Normal distribution that defines payload in terms of load time. The equations for the mean and standard deviation are developed by constructing a grid of payload and load time and calculating the mean and standard deviation of payload for each interval of load time. The data from the grid is used to generate a graph as shown in Figure 7 where both average and standard deviation of payload is plotted in terms of load time.

The best-fit equations for the mean and standard deviation of payload are shown in Figure 7. The next step is to define a Normal distribution with a mean and standard deviation using these best-fit equations. Eq(5) shows the form of the Normal distribution.

$$P(p) = Normal\ [-0.0363t^2 + 0.1239t + 0.8933, -0.0015t^2 - 0.013t + 0.1221] \qquad (5)$$

where $P(p)$ is the PDF that describes the payload, t is an instance of load time that is generated by the Beta distribution defined in Eq (4). A pair of load time-payload coordinates is generated by sampling from Eq (4) and Eq (5) for modeling each instance of the load activity during a simulation run.

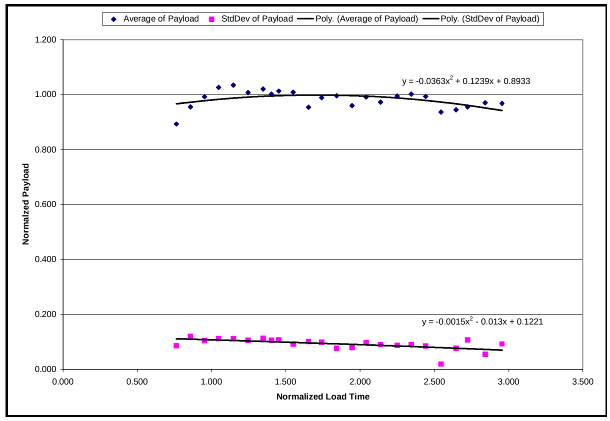

Figure 7: Parametric Form of Payload and Load Time

4 DEMONSTRATING THE JOINT PROBABILITY DISTRIBUTION

Most of the general purpose simulation systems allow the duration of the activity to be modeled using a PDF. The implementation of the joint probability distribution requires programmability. STROBOSCOPE offers the user programmability to accomplish this requirement and hence it will be used to demonstrate the methodology. The syntax and other details of STROBOSCOPE are available in Martinez [1996].

STROBOSCOPE has direct, two-statement, support for modeling a joint distribution where time (load time) is a function of resource usage (payload) because in STROBOSCOPE activity times are determined after resources have been acquired. In this case, however, it was more convenient to reduce the data to a format in which payload is a function of load time. Modeling this requires the more involved STROBOSCOPE snippet of code shown in Figure 8. The duration of the load truck activity is determined by sampling from a *ScaledBeta* distribution. At the start of the load truck activity, the conditional probability of *Payload* is determined as a function of load time. This two-statement approach in STROBOSCOPE allows one to model the simultaneous variation in payload and load time.

```
DURATION LoadTruck
        ScaledBeta[0,1.23,11.1,4.21];

ONSTART LoadTruck ASSIGN
        LoadTruck.Truck.Payload
        Normal[0.0363*LoadTruck.Duration^2 +
        0.1239*LoadTruck.Duration + 0.8933,
        -0.0015* LoadTruck.Duration ^2-0.013*
        LoadTruck.Duration +0.1221];
```

Figure 8: A Snippet of Code in STROBOSCOPE to Implement the Joint Probability Distribution of Load Time and Payload

A model was created incorporating the code shown in Figure 8; the results of the payload - load time pairs created during run time is shown in the form of a PLT Map in Figure 9. It is interesting to note the following similarities between the PLT Map created using field data as shown in Figure 3 and the PLT Map simulated using the PDFs as shown in Figure 9.

The range of payload represented by the field data and the simulated data correspond to the same range (0.7 to 1.3).

The modal region for the field data and the modal region for the simulated data fall in the same region.

An important point to note regarding the field data is that the trucks are loaded in integer passes. Schexnayder *et al.* [1999] reported that there appears a good match between volumetric load and full shovel buckets cycles. Such granularity in the data is smoothened out when describing the data using a PDF. Moreover, a unique Normal distribution is used for the entire range of payload. It is clear from that different distributions can be modeled for each range of load time. Similarly, it is noted that the distribution that describes the load time and the equations that describe the mean and standard deviation will differ between ranges of data. The data can be dissected into as many granular slices as needed by the application. However, for the purpose of illustration, the payload and load time are defined by a unique PDF models.

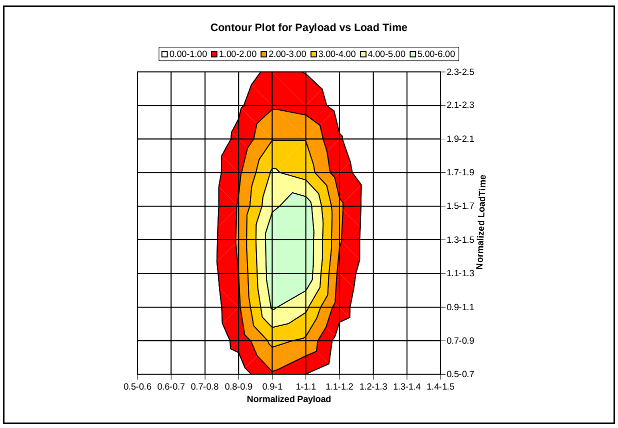

Figure 9: PLT Map Simulated Using the Parametric Form of the Joint Probability Distributions

5 CONCLUSION

This paper focussed on modeling the load activity in an earthmoving operation involving a loader and a truck. The load activity is characterized by two quantitative measures, namely, final payload and total load time and hence requires the knowledge of variation in both parameters. This paper described a methodology to develop the statistical parameters necessary to define the simultaneous variation in payload and load time.

The statistical parameters in the form of probability distributions can be used to define the load time as well as the payload of the load activity. Two scenarios are possible in terms of payload and load time. The first scenario is that the payload may be independent of the load time in which case a product of the individual probabilities of payload and load time can be used to determine the joint probability of a payload-load time combination. The second scenario is that payload may be dependent on the load time in which case a parametric form of payload in terms of load time must be used to determine the joint probability. The second scenario was discussed in detail using field data.

The probabilistic distributions define the input to simulation models. However, they are not intuitive enough to provide an immediate response regarding the operation. A PLT Map was developed to provide a graphical representation of the simultaneous variation in payload and load time. The PLT Maps can improve on-the-job communications and hence facilitate productivity improvement.

Finally, it is pointed out that the performance records collected by the vehicle instrumentation are not characterized, that is, there is no record of the operating conditions under which the data is collected. The effect of the operating conditions on the variance of the activity's performance is not captured automatically. Recording the operating condition requires manually data collection because it involves subjective data. If the operating conditions are available, the performance records collected by the vehicle instruments can be filtered based on the conditions. It then becomes possible to generate a PLT Map and hence the PDF for each set of operating conditions.

REFERENCES

AbouRizk, S. M. and D. W. Halpin, (1992), Statistical properties of construction duration data, *Journal of Construction Engineering and Management*, ASCE, 118(**3**), pp525-544.

Day, D. A., (1973), *Construction Equipment Guide*, John Wiley & Sons, New York, NY.

Gransberg, D. G., (1996), Optimizing haul unit size and number based on loading facility characteristics, *Journal of Construction Engineering and Management*, ASCE, 123(**3**), pp248-253.

Halpin, D. W., (1977), CYCLONE- Methods for modeling job site processes, *Journal of Construction Division*, 103(**3**), pp489-499.

Ioannou, P. G. and J. C. Martinez, (1996) Comparison of Construction Alternatives Using Matched Simulation Experiments, *Journal of Construction Engineering and Management*, ASCE, 112(**3**), ASCE, Reston, VA.

Ioannou, P. G., (1990), *UM-CYCLONE- Discrete Event Simulation System—User's Guide*, UMCEE Report

No. 89-12, Dept. of Civil Engineering, University of Michigan, Ann Arbor, MI.

Kalk, A., (1980), INSIGHT: Interactive simulation of construction operations using graphical Techniques, Technical Report No. 238, Dept. of Civil Engineering, Stanford University, Stanford, CA.

Law, A.M. and W.D. Kelton, (1991), *Simulation Modeling and Analysis* 2nd Edition, McGraw-Hill, Inc, New York.

Lluch, J. and D. W. Halpin, (1982), Construction operations and microcomputers, *Journal of Construction Division*, ASCE, 108(CO1), pp129-145.

Martinez, J. C. and P. G. Ioannou, (1999), General Purpose Systems for Effective Construction Simulation, *Journal of Construction Engineering and Management*, ASCE, (July/August 1999), Reston, VA.

Martinez, J. C. and P. G. Ioannou, (1995), Advantages of the Activity Scanning Approach in the Modeling of Complex Construction Processes. In *Proceedings of the 1995 Winter Simulation Conference*, San Diego, CA.

Martinez, J. C., (1996), *STROBOSCOPE: State and resource based simulation of construction processes*, Ph.D. Thesis submitted to the University of Michigan.

Ioannou, P. G., and Martinez, J. C. (1996). Simulation of Complex Construction Processes, *Proceedings of the 1996 Winter Simulation Conference*, San Diego, CA, 1321-1328.

Paulson, Jr., B. C., W. T. Chan and C. C. Khoo, (1987), Construction operations simulation by microcomputer, *Journal of Construction Engineering and Management*, ASCE, 113(**2**), pp302-314.

Schexnayder, C., Weber, S. L., and Brooks, B. T., (1999), Effect of truck payload weight on production, *Journal of Construction Engineering and Management*, ASCE, 125(**1**), pp1-7.

AUTHOR BIOGRAPHIES

GOVINDAN KANNAN received his Ph.D. degree from the Via Department of Civil Engineering at Virginia Tech in May 1999. He received a Masters Degree in Civil Engineering from the National University of Singapore in 1995 and a Bachelors Degree in Civil Engineering from the Regional Engineering College (Tiruchirapalli, India) in 1991. He worked as a site engineer for a large construction company in India for two years. His interests include construction simulation, optimization, scheduling, and estimating.

MICHAEL C. VORSTER is the David H Burrows Professor of Construction Engineering and Management in the Via Department of Civil Engineering at Virginia Tech. He received his Ph.D. from the University of Stellenbosch in 1981; an M.B.A. from the University of Capetown in 1971; and a B.Sc. from the University of Capetown in 1965. He has published articles as well as served as a reviewer for various construction-related journals. To his credit, he has more than 14 years of experience in the construction industry. His research interests include construction scheduling, dispute resolution, construction methods, and equipment economics.

JULIO C. MARTINEZ is an Assistant Professor in the Via Department of Civil Engineering at Virginia Tech. He received his Ph.D. in Civil Engineering at the University of Michigan in 1996; an MSE in Construction Engineering and Management from the University of Michigan in 1993; an M.S. in Civil Engineering from the University of Nebraska in 1987; and a Civil Engineer's degree from *Universidad Catolica Madre y Maestra* (Santiago, Dominican Republic) in 1986. He designed and implemented the STROBOSCOPE simulation language as part of his Ph.D. dissertation research. In addition to discrete event simulation, his research interests include construction process modeling and decision support systems for construction.