SIMULATION EXPERIMENT FOR IMPROVING CONSTRUCTION PROCESSES

Shihyi Wang
Daniel W. Halpin

Division of Construction Engineering & Management
Purdue University
550 Stadium Mall Drive
West Lafayette, IN 47907-2051, U.S.A.

ABSTRACT

From the perspective of flow analysis, the more complex the project, the more wastes are prone to build up, due to the increasing number of interfaces between activities. When processes are further expanded beyond the jobsite to the scale of a supply chain, the complexity is usually beyond human perception. In order to provide a systematic approach to help management make correct and timely decisions, an analytical approach capable of efficiently and economically modeling the complicated processes and to address various managerial questions is necessary.

An analytical framework combining simulation, design of experiment, regression analysis, and mathematical programming is proposed to facilitate management to determine the optimal design for a construction process under various constraints. A simplified concrete delivery process is used as an example to illustrate how the framework is implemented. Some implementation issues and other potential applications will be discussed.

1 INTRODUCTION

Simulation is one of the most widely used operations-research and management-science techniques (Law and Kelton 2000, p. 2). It takes advantage of the efficiency and reliability of modern computers to imitate the operations of various real-world processes.

Computer simulation has been used for decades to model real-world problems ranging from designing manufacturing systems to analyzing financial or economic systems. Due to the complexity of these problems and the limitations of human intelligence models, simulation provides a virtual world where the decision makers can better understand the nature of the problem by conducting experiments in a more controllable and low-cost environment.

2 SIMULATION IN CONSTRUCTION

The use of simulation for studying construction operations was first proposed and developed by Halpin in the 1970s (Halpin 1973). He developed the CYCLONE (CYCLic Operations NEtwork) modeling system that significantly simplifies the simulation modeling process and makes it accessible to construction practitioners with limited simulation background (Sawhney et al. 2003). A computer program called MicroCYCLONE, based on the CYCLONE modeling elements was developed and became the most popular general purpose simulation program in the construction industry. The introduction of CYCLONE modeling elements and the MicroCYCLONE simulation program brought the study of construction operations to researchers' attention, and there have been many construction simulation programs developed in the past two decades based on CYCLONE concepts.

Simulation has been broadly used in analyzing construction processes as simple as a mason supply system and as complex as projects such as the Isle of Palm Connector Bridge. With the advent of the lean construction concept, a few researchers have begun to investigate the use of construction simulation in areas pertaining to work flow variability and other lean concepts (Sawhney et al. 2003). Tommelein's paper (1998) is a classic example of how simulation can be used in modeling lean construction processes.

Traditionally, research in construction simulation includes process observation, duration and resource data collection, process modeling using CYCLONE, process simulation to estimate productivity, and performing sensitivity analysis to suggest alternatives to improve productivity.

3 DEFICIENCIES IN PAST RESEARCH

In a simulation study of a construction process, the following questions are usually of interest – what is the impact of resource variation and other factors on the response of a given construction operation? What combinations of resources will maximize the productivity? Sensitivity analysis, or the so-called *one-factor-at-a-time approach*, is typically used in construction simulation to answer these questions. The general procedure is to subjectively select one of the most important factors and change the levels of

the factor while fixing the remaining factors in order to investigate system response. A similar procedure is then iteratively applied to the remaining factors based on the order of importance until the answer is obtained.

Although one-factor-at-a-time sensitivity analysis has been widely used in construction simulation studies, it has the following disadvantages when compared with the factorial design method (Wu and Hamada 2000, p. 114):

1. It requires more runs for the same precision in effect estimation.
2. It cannot estimate some interactions between factors.
3. The conclusions from this type of analysis are not general.
4. It can miss optimal settings of factors.

Another important issue overlooked in past construction simulation research is the statistical significance of process change. Suppose the output of a simulation model, e.g., productivity in construction simulation, denoted by Y, has an unknown distribution with a mean μ and a variance σ^2. Assume the result of the sensitivity analysis shows that the productivity can be increased by δ by changing the number of masons from one to two, where δ is less than σ^2. From the statistical point of view, the increase is not significant enough since it could be due to noise or random error within the model. Hence, one cannot conclude this change is effective. However, most literature in construction simulation has not addressed this issue properly.

4 COMPUTER EXPERIMENT

Design of Experiment (DOE), or experimental design, has its origin in agriculture and laboratory experiments. The purpose of an experiment is to answer the questions of interest and to understand or improve the systems. Statistical design of experiments refers to the process of planning the experiment so that appropriate data that can be analyzed by statistical methods will be collected, resulting in valid and objective conclusions (Montgomery 2001, p. 11).

Traditionally, experimental designs are applied to physical experiments, such as agriculture field experiments and controlled clinical trials. When physical experiments are practically impossible, e.g., the numbers of input variables are too large or conducting physical experiments is costly or may involve ethical issues, computer experiments become a feasible alternative. Computer experiments differ from physical experiments because they use computer codes to generate data and create virtual environments to simulate experiments of interest. While many DOE techniques were developed with physical experiments in mind, they can fairly easily be adapted for use in computer simulation experiments as well (Kelton and Barton 2003).

Among various experimental designs, 2^k full factorial designs and fractional factorial designs are most commonly used in industrial experimentation. When the number of factors is large, these two designs are used to identify important factors in the early stage of an experiment and are often called factor-screening experiments. A complete introduction to these two experimental designs can be found in classic design of experiment textbooks, such as Montgomery (2001), Wu and Hamada (2000), Oehlert (2000), and Box, Hunter, and Hunter (1978).

By performing experimental design, one can estimate *main effects* of factors and *interaction effects* between factors. Main effect measures the effect of each individual factor on the response. By assigning two opposite levels, namely high (+) and low (−) levels to each factors, one can estimate the main effect of each factor by taking the difference between the average of the responses of all observations at the high level (+) of a factor, say A, and the average of the responses of all observations at the low level (−) of A.

The major disadvantage of the one-factor-at-a-time strategy is that it fails to consider any possible interactions between factors (Montgomery 2001, p. 4). Two factors A and B are said to *interact* if the effect of factor A depends on the levels of factor B, and the interaction between A and B is denoted as $A \times B$ or AB. For example, the productivity of a worker is affected by the models of the equipment the worker uses. Productivity tends to be higher when workers use the equipment with which they are familiar or that matches their body types. If one would like to investigate the factors that will influence the production rate of a process, and both the number of workers and the models of the equipment are among the factors of interest, then the interaction between these two factors should be properly accounted for because the interaction may affect the production rate of the process. The interaction AB is called a two-way interaction because there are two factors involved, and it measures the effect of one factor depending on the levels of the other factor.

Discussion of three-way interactions, e.g., ABC, BCD, etc., or higher-order interactions can be found in the references mentioned above. The presence of interactions usually makes the interpretation of the models difficult and it is almost impossible to clearly explain the models containing three-way or higher-order interactions. Fortunately, according to the hierarchical ordering empirical principle, lower order effects are more likely to be more important than higher order effects. Therefore, in practice the effects of high-order interactions are often assumed negligible.

2^k full factorial design requires 2^k simulation runs to be performed, i.e., if $k = 4$, $2^4 = 16$ simulation runs are needed. If the number of factors in a process is large, the number of runs grows exponentially and it will become impractical and inefficient to perform all these simulation runs. For instance, if there are 10 factors of interest, i.e., $k = 10$, a 2^k full factorial design requires $2^{10} = 1024$ simulation runs, which will consume a considerable amount of time to complete. In addition, when the number of factors becomes large, most degrees of freedom will be used to estimate three-way or

higher-order interaction effects. As mentioned before, these high-order interactions are usually not significant. Therefore, it is not practical to commit a large number of degrees of freedom to estimate these insignificant effects. For these reasons, fractional factorial designs, which consist of a subset or fraction of full factorial designs, are commonly used in place of full factorial designs.

The concept of fractional factorial designs is based on the fact that if one can reasonably assume that certain high-order interactions are negligible, then the main effects and low-order interactions can be obtained by running only a fraction of the full factorial experiment. Fractional factorial designs are among the most widely used types of design in industry (Myers and Montgomery 2002, p. 155) and they are often used for factor-screening purposes.

Regression models can be used to fit the run data from the previous experimental designs. The fitted model is called a *meta model*, which is the statistical model of the simulation. The meta model acts as a proxy of the simulation model and can be used to study the effects of input factors on the response. Furthermore, mathematical programming techniques can be applied to search for the best combinations of input factors that maximize or minimize the response.

5 EXAMPLE

Assume that the productivity of Project A, measured in terms of truck-load per minute, is of interest. When a truck arrives at the jobsite, it must wait for the dump position to be available in order to proceed. Once the truck moves into the dump position, it waits for the hopper to become available for taking the concrete. After the truck finishes dumping the concrete to the hopper, it passes the counter, departs from the jobsite and returns to the batch plant. As soon as the truck departs, a dump position becomes available.

Due to the limited capacity of the concrete batch plant, it is impossible for the supplier to fulfill customers' needs all of the time. The *service rate* of Project A, the probability that the batch plant is dedicated to serving Project A, or the ratio between the number of trucks traveling to Project A and the total number of trucks during the course of simulation is defined to model this situation. For the purpose of this paper, the service rate is defined as the routing probability which indicates that a truck is allocated to Project A or Project B. In other words, service rate is the probability P that the truck moves from 12 to either 13 or 14 in Figure 1. Once the batch plant is available, a signal will be released and if a truck happens to be idle, the truck will be assigned to either Project A or Project B (collectively represents all non-A projects) based on the service rate. After the truck is loaded, it travels to the jobsite. Figure 1 is the CYCLONE network of the process.

The process focuses on the delivery and receiving of the concrete supply process. In practice, the availability of hoppers is related to the concrete placement cycles, which definitely has an impact on the truck departure time. For the sake of simplicity, this part is not taken into account in the model. In addition, the duration data used in this example is hypothetical and may not reflect the real-world situation.

6 IMPLEMENTATION

Figure 2 is the flowchart for implementing the proposed framework to study construction processes. In the following sections, the simplified concrete supply process de-
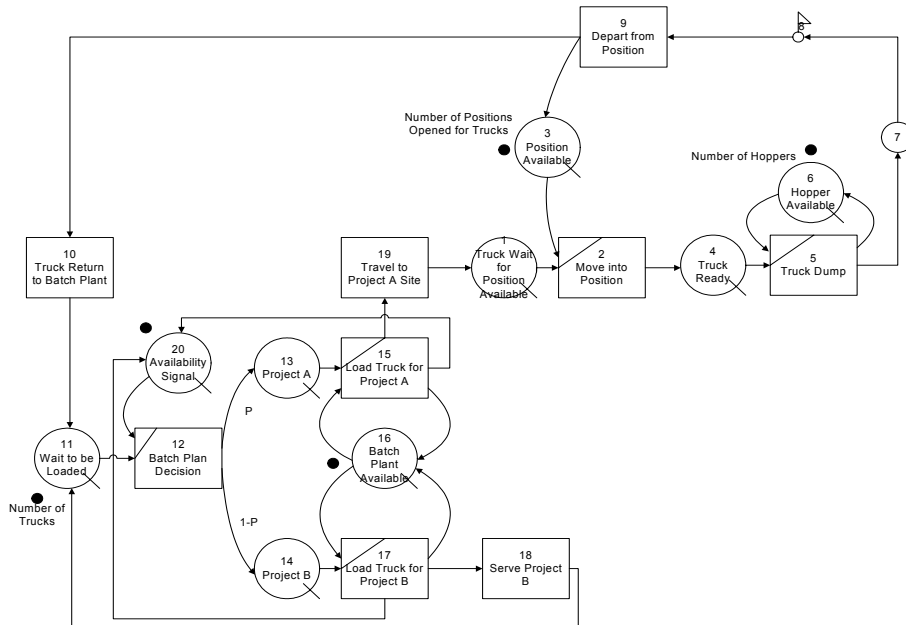

Figure 1: CYCLONE Network of the Simplified Concrete Delivery Process

scribed in Section 5 is selected to illustrate the use of construction simulation software WebCYCLONE (Halpin, Jen, and Kim, 2003) and 2^k full factorial design as a factor-screening procedure. After the important effects are identified, a regression model is fitted to the data and the fitted model will be used as an objective function in the mathematical programming model. All of the statistical analyses are performed in SAS, and the best combination of resources under constraints is obtained by using LINGO.
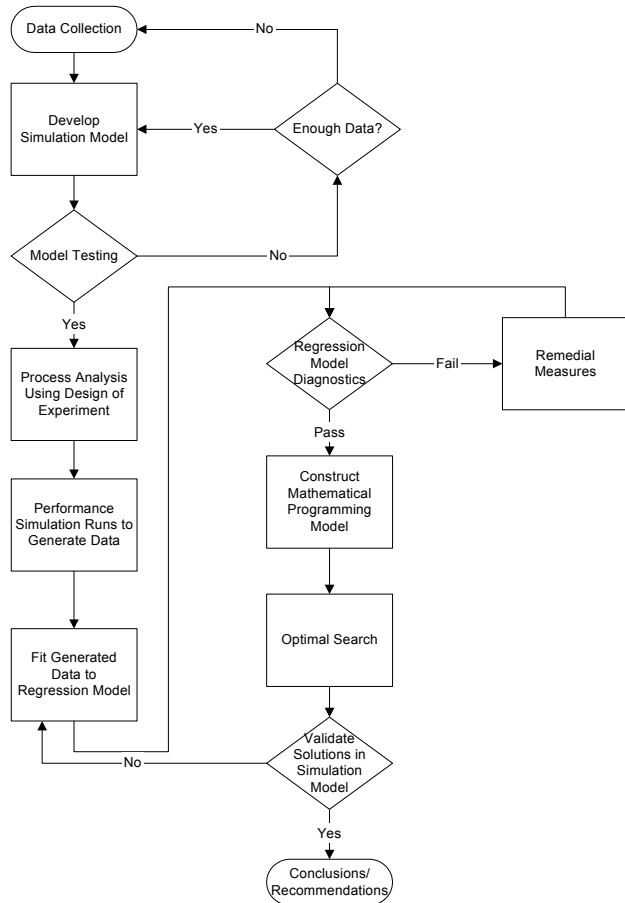

Figure 2: Flowchart for Implementing the Proposed Framework to Study Construction Processes

6.1 Data Collection and Model Development

Figure 1 is the CYCLONE model based on the observation at the batch plant and the jobsite. In addition, the duration for each activity and resource information are summarized in Tables 1 and 2 respectively. It is worth noting that the duration of Node 12 Batch Plant Decision is assumed to be 0, meaning that the assignment of trucks takes no time at all. In reality, it is possible that the management needs evaluate the priorities and consider which customer to serve first.

Table 1: Summary of Duration Information

Node	Activity	Duration
2	Move into Position	Normal (3.0, 1.0)
5	Truck Dump	Normal (10.0, 1.0)
9	Depart from Position	Normal (2.0, 0.8)
10	Truck Return to Batch Plant	Normal (12.0, 2.0)
12	Batch Plant Decision	Deterministic 0
15	Load Truck for Project A	Normal (7.0, 1.0)
17	Load Truck for Project B	Normal (7.0, 1.0)
18	Service Project B	Normal (35.0, 10.0)
19	Travel to Project A Site	Normal (15.0, 3.5)

Table 2: Summary of Initial Resources

Node	Resource	Quantity
3	Position	1
6	Hopper	1
11	Truck	5
16	Batch Plant	1
20	Availability Signal	1

6.2 Initial Simulation Run and Model Testing

After the simulation model is developed, it is important to test how well the simulation output data resembles the field data. An easy way to examine the appropriateness of the model is to compare the major summary statistics of the simulation output, e.g., mean and variance of productivity, with those of the field data to see how close they are. There are also many classical statistical tests, such as two-sample chi-square test and two-sample t-test, which can be used to validate the simulation models. However, Law and Kelton (2000, p. 280) point out that many real-world systems and simulations are nonstationary and autocorrelated so that these tests are not directly applicable. Instead, they suggest using the inspection approach, confidence interval, and time-series approaches. Detailed discussion of these three approaches can be found in Law and Kelton (2000, pp. 283 – 290).

Since there are no real-world data for comparison and validation, the simulation is assumed to be appropriate in this case study.

6.3 Design of Experiment

The model has four resource inputs and these are all factors of interest. "Signal availability" should be overlooked in this case because it only acts as a mechanism to trigger the probability assignment at the decision point. On the other hand, the "service rate" should be taken into account because the commitment of the batch plant is expected to have a critical influence on the progress of the jobsite operation. Hence, there are four factors to be investigated. Since the number of factors is not large, a 2^4 full factorial design is appropriate; otherwise, a fractional factorial design should be used instead. Table 3 summarizes the factors for the 2^4 design. Table 4 is the design matrix and re-

sponses for the 2^4 design with single replicate. The responses are generated from simulation with the same random number seed. It is worth noting that the simulation study is *deterministic* in the sense that the same set of input parameters will generate the same output. If there is more than one replicate in the design, different random number seeds must be used for each replicate. Otherwise the responses for each replicate will be exactly the same.

Table 3: Factors for the 2^4 Design

Factor	Factor Notation	Low (-) Level	High (+) Level
Number of Trucks	*A*	5.0	25.0
Service Rate	*B*	0.3	0.8
Number of Positions	*C*	1.0	8.0
Number of Hoppers	*D*	1.0	10.0

Table 4: Design Matrix and Responses for the 2^4 Design

Run #	*A* Truck	*B* Service Rate	*C* Position	*D* Hopper	*Y* (load/min) Productivity
1	-	-	-	-	0.0316
2	+	-	-	-	0.0431
3	-	+	-	-	0.0652
4	+	+	-	-	0.0667
5	-	-	+	-	0.0319
6	+	-	+	-	0.0431
7	-	+	+	-	0.0777
8	+	+	+	-	0.0994
9	-	-	-	+	0.0316
10	+	-	-	+	0.0431
11	-	+	-	+	0.0652
12	+	+	-	+	0.0667
13	-	-	+	+	0.0321
14	+	-	+	+	0.0431
15	-	+	+	+	0.0802
16	+	+	+	+	0.1153

The effect of each factor and the interactions between factors can be estimated by hand or by using statistical software packages. Table 5 is the summary of the estimated effects ranked in descending order. Often three or higher interactions are assumed to be negligible in order to estimate the mean square error. However, this assumption might pool significant interactions. A normal probability plot is created to detect those significant effects. Usually, significant effects will not lie along the same line in the normal probability plot. Based on the estimates in Table 5 and Figure 3, effects *B*, *C*, *BC*, and *A* are most significant and the further investigation should focus on them.

The presence of the interaction of *BC* makes the model interpretation difficult. By examining the interaction plot (Figure 4), the effects of *C* appear to depend on the levels of *B*. When the service rate is fixed at 0.3, the use of eight positions and one position have almost the same effect on the productivity; when the service rate raises to 0.8, the

process with eight positions has significantly higher production than that with one position.

Table 5: Estimated Effects (Descending Order)

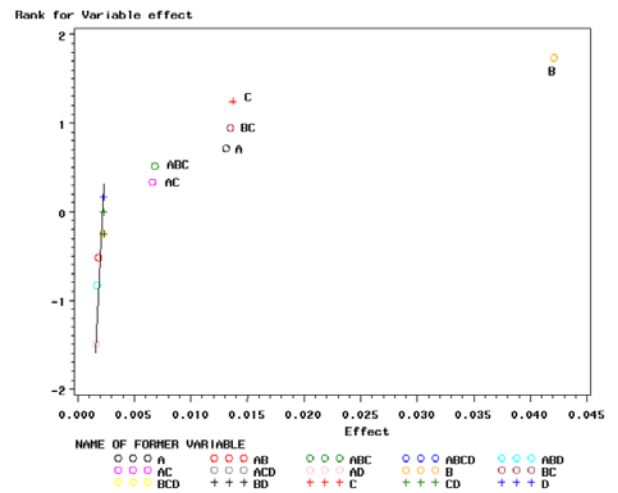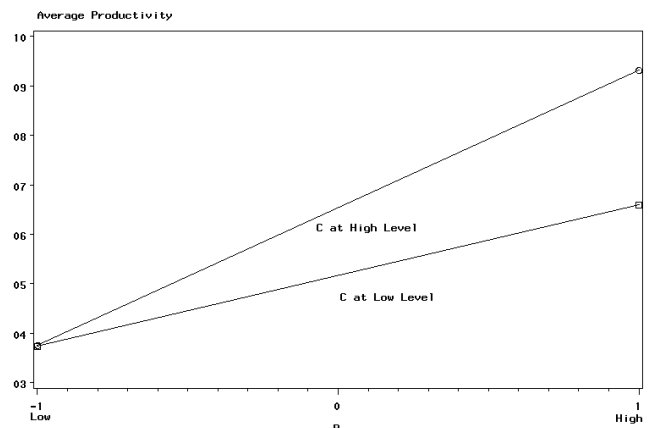| Factor	Effect	Factor	Effect
B	0.0421	BD	0.0023
C	0.0137	BCD	0.0023
BC	0.0135	AB	0.0018
A	0.0131	ABD	0.0017
ABC	0.0068	ABCD	0.0017
AC	0.0066	AD	0.0017
D	0.0023	ACD	0.0017
CD	0.0023		

Figure 3: Normal Probability Plot for the Effects

Figure 4: Interaction Plot for Factors *B* and *C*

6.4 Model Fitting and Diagnostics

Based on the assumption that three or higher interactions are negligible, along with the observations from the estimated

effects and normal probability plot, 16 observations generated from the 2^4 design were fitted to regression model

$$y = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 C + \beta_{23} BC \qquad (1)$$

and the fitted model is

$$\hat{y} = 0.0126 + 0.0007A + 0.0495B \\ -0.0023C + 0.0077BC. \qquad (2)$$

The model diagnostics showed that the assumption of constant error variance was violated. The Box-Cox variance stabilizing transformation was used to search for the best transformation to model the response. The result suggested a log transformation of the response y ($\lambda = 0$) is appropriate and the final fitted model is

$$\hat{y} = \exp[-3.7961 + 0.0117A + 1.0679B \\ -0.0270C + 0.0931BC]. \qquad (3)$$

The negative estimated parameter associated with the number of positions available may contradict the intuition that a larger C will have a positive influence on the average productivity. Due to the presence of the BC interaction term, the fitted model should be interpreted with extra care. If the service rate B is fixed at 0.3, the fitted model can be simplified as

$$\hat{y} = \exp[-3.4757 + 0.0117A + 0.0009C] \qquad (4)$$

, where C has a positive but relatively small coefficient, which suggests that by increasing one unit of the position in the jobsite when the service rate is 0.3 will only increase the average productivity by 0.09%. This extremely small amount of increase in the average productivity explains why the two lines in Figure 4 appear to intersect when B is at 0.3 (denoted by -1 in the interaction plot). When the service is 0.8, the fitted model is reduced to

$$\hat{y} = \exp[-2.9418 + 0.0117A + 0.0475C]. \qquad (5)$$

One unit increase in C will raise the average productivity by 5%. The different rates in productivity improvement show B and C are not independent of each other.

6.5 Developing Mathematical Programming Model

Suppose the objective here is to maximize the productivity of the process. Without any constraint, the best resource combination would be $A = 25$, $B = 0.8$, $C = 8$, $D = 10$, which is exactly the same as the 16th run. In practice, there are always constraints on suppliers' and contractors' resource commitments. Assume the batch plant can only have a maximum of 15 trucks operating at the same time; the management is only willing to guarantee up to 60% of

the service rate. Furthermore, the jobsite can only have five dump positions at most. Mathematically, the problem can be modeled as the following:

Objective: Maximize
$$\hat{y} = \exp[-3.7961 + 0.0117A + 1.0679B \\ -0.0270C + 0.0931BC]$$
Subject to: $\qquad\qquad\qquad (6)$
$$5 \le A \le 15$$
$$0.3 \le B \le 0.6$$
$$1 \le C \le 5$$

, where A and C are integers and B is a real number. This is a nonlinear mixed integer programming problem. The best solution obtained by using LINGO is 0.0587 truck-load per minute when $A = 15$, $B = 0.6$ and $C = 5$.

Suppose the costs of using a truck and a hopper are $100 per hour and $50 per hour respectively. There is a $30 one-time fee for preparing and cleaning a dump position, and the concrete supplier charges $10,000 for every 10% increase in service rate. Furthermore, the contractor requires 40 truck-loads of concrete to be delivered within 12 hours. The objective is to minimize the total costs. Then the problem can be modeled as shown in Table 6.

Table 6: Optimization Model in LINGO

```
MODEL:
MIN=Cost_A+Cost_B+Cost_C+Cost_D; !Objective:
Minimize the total cost;
A  <=15;
A  >=5;
B  <=0.6;
B  >=0.3;
C  <=5;
C  >=1;
Productivity=60*@exp(-3.7961+0.0117*A+1.0679*B-
0.0270*C+0.0931*B*C); !unit: truck-load/hour;
Productivity >= 3.3333; ! 40/12=3.3333 truck-
load/hour;
Time = 40/Productivity; !unit: hour;
Cost_A = 100*A*TIME;
Cost_B = 10000*(B-0.3);
Cost_C = 30*C;
Cost_D = 50*Time;
@GIN(A);@GIN(C); !Restrict A and C to integers;
```

Again this is a nonlinear mixed integer programming problem. The minimum cost of $16,840 is attained by using 11 trucks, a 60% service rate, and five dump positions. The predicted average productivity based on these parameters is 3.3601 truck-loads/hour, or 0.056 truck-loads/minute.

6.6 Validating the Solution in Simulation Model

To validate if the solution obtained from 6.5 is close to the estimated productivity from the simulation model, it is necessary to conduct one more simulation run with the best combination of resources. The average productivity ob-

tained by using $A = 11$, $B = 0.6$, $C = 5$, and $D = 1$ is 0.0836 truck-loads/minute or 5.02 truck-loads/hour, which is significantly higher than 3.36 truck-loads/hour from the fitted model. This suggests the regression model does not resemble the simulation model very well.

One of the possible reasons is due to the assumption of three or higher interactions are negligible. In fact, both Table 5 and Figure 3 show the effect of the three-way interaction ABC is ranked fifth and could be important too. By including this term in the model, the fitted equation becomes

$$\hat{y} = 0.0194 + 0.0002A + 0.0495B - 0.0023C \\ + 0.005BC + 0.0002ABC. \tag{7}$$

The output of LINGO by using the new fitted equation shows the minimum cost of \$9,619 is substantially lower and can be obtained by using five trucks, 58.69% service rate, five dump positions, and one hopper. The estimated productivity of 3.33 truck-loads/hour is much closer to the 3.47 truck-loads/hour from the simulation model, which shows the prediction of the fitted equation improves by including the three-way interaction ABC in the model.

One might argue that the effect of interaction AC is close to ABC and it should be included in the model too. By the law of parsimony, also known as Occam's Razor, a simpler or more parsimonious model that describes the data adequately is preferable to a more complicated one which leaves little of the variability unexplained (Dobson 2001, p. 34). Since the estimate from the fitted equation is good enough, there is no need to add one or more terms to increase the complexity of the model without improving the accuracy too much.

7 REMARKS

The fitted regression is a simplified equation used to model the relation between the predictor variables and the response variable. Sometimes, the estimates may not resemble the output from the simulation model. In addition, due to the characteristics of the objective function and the constraints, there is no guarantee that the global optimal solution is available in a nonlinear programming problem, which is the case here. However, the local optimal solutions can still be viewed as the *best* solution under the constraints of the computing power and the optimal searching methods. Relatively, the appropriateness of the regression model is more influential, as shown in the case study. By including the three-way interaction in the model, the minimum cost significantly decreased from \$16,840 to \$9,619. This implies the outcome of the mathematical programming could be very sensitive to the appropriateness of the regression model. Therefore, the modeler should be very cautious when making assumptions as well as performing diagnostics and remedial measures for the regression models.

More often than not, the fitted regression models are not in the form of first-order linear regression. Interactions and sometimes high-order linear terms are present in the fitted equations, which makes model interpretation difficult. In addition, as the complexity of the fitted equations increases, the global optimal solutions for the mathematical programming model may not exist. Therefore, it is always desirable to follow the law of parsimony to strike the balance between the simplicity and the predictability of the fitted models.

8 CONCLUSIONS

A methodology combining simulation, design of experiment, regression analysis, and mathematical programming is proposed to provide a systematic and analytical approach to analyze construction processes. A hypothetical case study was used to demonstrate the procedure of implementing the proposed methodology and to highlight certain issues worth noting when using this methodology.

Generally, the proposed method has the following advantages.

1. The methodology can identify the focal points of the problem. The use of planned simulation experiments is able to answer the questions of interest effectively. By estimating the effects of all of the factors and their interactions, the criticality of each effect can be ranked, which provides a starting point for management to redesign the process. For illustration purposes, the process in the case study is relatively simple and not many factors were involved. Therefore, the effectiveness of the proposed method is not obvious and it is possible to identify the most important factors by intuition or by observing run data. However, when the processes of interest have more factors, say 10, then it is practically impossible and unscientific to manually identify the factors that have a significant impact on productivity. In addition, there is no way to assess the interaction effects by just looking at the simulation outputs.

2. The methodology can integrate cost, time, and other constraints by using mathematical programming. Existing construction simulation software packages emphasize analyzing construction processes to improve productivity. Constraints on resources, time, and costs are only briefly addressed. Regression analysis and mathematical programming are used to exploit the benefits of construction simulation and at the same time take into account the constraints that cannot be explicitly modeled in simulation.

3. The obtained optimal solution sets a goal for continuous improvement. By comparing the best combination of resources maximizing the productivity with the current status and considering the existing constraints, a guideline for process improvement can emerge. Figure 5 is the contour

plot based on the final fitted equation when C is set at 5. All of the combinations of factors A and B that lead to the same productivity are linked by a contour line. In the case study the highest productivity that can be achieved is 0.08 truck-loads/minute, which is represented by the line on the top-left corner of the plot. Suppose the current status of the process is at point A. To achieve the highest productivity, there are several options, as shown in the figure. Management can consider the resource availability, as well as time and cost constraints to decide which option is more appropriate. This technique is called the Response Surface Method (RSM), which has been widely used in optimizing manufacturing processes. Finally, in light of the difference between simulations and field operations, the optimal solution may not be achievable. However, it can be viewed as a goal that management can continuously redesign the process in order to achieve.
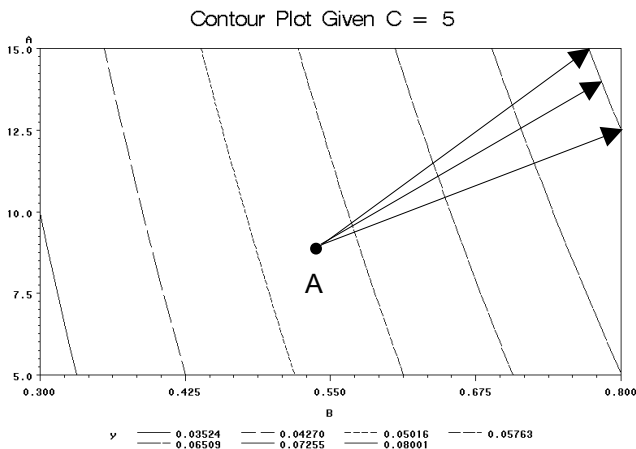

Figure 5: Contour Plot for Factors A and B, Given C is Fixed at 5 Dump Positions

REFERENCES

Box, G., Hunter, W. G. and Hunter, J. S. (1978). "Statistics for Experimenters." *John Wiley & Sons, Inc*.

Dobson, A. J. (2001). "An Introduction to Generalized Linear Models." 2nd Ed. *Chapman & Hall/CRC Press*.

Halpin, D. W. (1973). "An Investigation of the Use of Simulation Networks for Modeling Construction Operations", *Ph.D. thesis*, The University of Illinois at Urbana-Champaign, Illinois.

Halpin, D. W., Jen, H., and Kim, J. (2003). "A Construction Process Simulation Web Service." *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds., 1503 – 1509. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Kelton, W. D., and Barton, R. R. (2003). "Experimental Design for Simulation." *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds., 59 – 65. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Law, A. M., and Kelton, W. D. (2000). "Simulation modeling and analysis." *McGraw Hill*, New York, NY.

Montgomery, D. G. (2001). "Design and Analysis of Experiments." 5th Ed. *John Wiley & Sons*, Inc.

Myers, R. H., and Montgomery, D. C. (2002). "Response Surface Methodology: Process and Product Optimization Using Design Experiments." *John Wiley & Sons, Inc*.

Oehlert, G. (2000). "A First Course in Design and Analysis of Experiments." *W. H. Freeman and Company*, New York.

Sawhney, A., Bashford, H., Walsh, K., and Mulky, A. R. (2003). "Agent-based Modeling and Simulation in Construction." *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds., 1541 – 1547. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Tommelein, I. D. (1998). "Pull-driven Scheduling for Pipe-Spool Installation: Simulation of Lean Construction Technique." ASCE, *Journal of Construction Engineering and Management*, 124 (4) 279-288.

Wu, J. and Hamada, M. (2000). "Experiments: Planning, Analysis, and Parameter Design Optimization." *John Wiley & Sons, Inc*.

AUTHOR BIOGRAPHIES

SHIHYI WANG is a Ph.D. candidate in the Division of Construction Engineering and Management and a Master student in Department of Statistics at Purdue University in West Lafayette, Indiana. He received a B.S.C.E. degree from National Chiao-Tung University at Hsinchu, Taiwan in 1999 and M.S.C.E. from Purdue University in 2000. He is currently a research assistant for the Construction Safety Alliance project. His email address is <`wang16@purdue.edu`>.

DANIEL W. HALPIN is Professor and Head of the Division of Construction Engineering and Management at Purdue University in West Lafayette, Indiana. He received a B.S. degree from the U.S. Military Academy at West Point in 1961. He received M.S. and Ph.D. degrees in Civil Engineering from the University of Illinois in Champaign in 1969 and 1973 respectively. He was on the faculty at Georgia Tech (1973-1985), and held the A. J. Clark Chair Professor Position at University of Maryland (1985-1987). He developed the CYCLONE methodology for construction simulation. He has received many awards, including the Huber and Peurifoy Awards, for his contributions to computer applications in civil and construction engineering. His email address is <`halpin@ecn.purdue.edu`>