Sensitivity of Linear Construction Project Performance to Correlation in Task Durations

Ricardo Eiris Pereira and Ian Flood
Rinker School, College of Design, Construction and Planning,
University of Florida
Gainesville, United States
email: reiris@ufl.edu, flood@ufl.edu

***Abstract*– The performance of construction work and thus the optimality of a plan are impacted directly by uncertainty in the performance of productive resources. For repetitive projects, smooth work flow is necessary to minimize or eliminate interruptions and idle time with the objective of reducing costs and meeting schedule deadlines. Earlier research suggests that correlation in the duration of repeated activities (where durations are stochastic) may be important in determining the most favorable plan. This study assesses the significance of correlation in this respect, using a Linear Scheduling framework for modeling repetitive construction work. A range of levels of correlation is considered using linear correlation between immediate successor repetitions of an activity. The results provide insight into the effects of different degrees of correlation on the expected project duration, cost, crew and equipment idle times and interruptions. The potential impact of correlation on the optimality of a project plan was found to be sufficiently high to justify further investigation of the nature and impact of correlation in construction.**

Keywords - Linear Correlation; Project Planning; Repetitive Construction Work; Risk Management; Stochastic Linear Scheduling.

I. INTRODUCTION

Activity repetition is most prevalent at a low level in a work breakdown of construction tasks, such as the cycling of equipment in an earthmoving operation or the laying of bricks, but it is also common at intermediate and high levels, such as the laying of utility lines or the construction of many similar floors in a high-rise building. Modeling repetitive activities requires careful attention to ensure accuracy since a small error in the estimate of a single repetition translates to a large error over many repetitions. Effects, such as learning and forgetting [1] in repetitive activities can be dramatic and if not properly addressed can lead to significant errors in the estimation of project performance. Uncertainty in activity performance must also be taken into account since it can significantly impact the accuracy of project performance estimates. Interactions between construction processes are usually sufficiently complicated that stochastic effects can only be modeled accurately using statistical sampling techniques, the most popular of which is the Monte Carlo method.

Rachmat et al. [2] investigated stochastic simulation on repetitive projects to incorporate activity performance uncertainty in look-ahead scheduling. In this analysis, it was concluded that including uncertainty on linear schedules improves the forecasting capability of project performance and thus helps a scheduler anticipate problem areas and formulate new plans that improve project performance.

Processes that are stochastic can also demonstrate correlation between the duration of repeated activities. Positive correlation means that if one activity (or repetition of an activity) takes longer than expected then the correlated activities (or repetitions of that activity) are also more likely to take longer. Work on correlation between construction activities (repeated or otherwise) is limited, but it is easy to demonstrate that positive correlation affects the statistical performance of a project by increasing kurtosis, meaning that more of the variance in the performance of a project results from occasional larger deviations as opposed to more frequent smaller deviations. An outstanding question, however, is whether the effects of correlation significantly impact the optimality of a plan. This paper reports on ongoing research into this question, presenting a series of experiments designed to assess the potential impact of correlation on project plan optimality. If correlation is found to impact plan optimality, then this will justify further work into the nature and impact of correlation.

The paper continues in Section II with a general definition of the concept of correlation and a formal definition of correlation as it is measured in this paper. Section III introduces the modeling framework adopted for assessing the impact of correlation on the performance of a plan, and the objective functions used for measuring optimality. The results from this study and their analysis are presented in Section IV, and the conclusions along with recommendations for continuing the study are presented in Section V.

II. ACTIVITY DURARTION CORRELATION

In this study, correlation is concerned with the relationship between the duration of activities that are subject to uncertainty. That is, once the uncertainty about the duration for one of the correlated activities has been resolved (such as when the activity has been executed on site) then we can make a statistically more accurate estimate of the duration of the correlated activities. Correlation occurs when the durations of different activities are determined by common factors, such as excavation activities that operate in similar ground conditions, utilize the same crew, and/or are overseen by the same superintendent. This study is concerned with determining the significance of correlation in terms of its impact on the

 ISBN: 978-1-61208-567-8

optimality of a project schedule. For this reason, a range of levels of correlation from 0.0 (no correlation) to 1.0 (perfect correlation) will be considered. Activities that are repeated are strong candidates for demonstrating correlation since the durations of the repetitions will likely be determined by many common factors. Moreover, the impact such correlation may have on the optimality of a plan could compound over many repetitions and thus quickly become significant. For this reason, this study will consider correlation that occurs between repetitions of an activity but not between different activities.

Correlation between the durations of repeated activities will be implemented by calculating a normalized weighting of the previous repetition duration and a stochastically generated duration. That is:

$$D_n = \begin{cases} D'_n & ; n = 1 \\ k \cdot D_{n-1} + (1-k) \cdot D'_n & ; n \geq 2 \end{cases} \quad (1)$$

where: D_n = the duration for the nth repetition of the activity;

D'_n = a stochastically generated component for the duration of the nth repetition of the activity;

D_{n-1} = the duration of the (n-1)th repetition of the activity; and

k = the correlation between the durations of subsequent repetitions (ranging from 0.0 for no correlation to 1.0 for perfect correlation).

While there is no published work to support any particular model for representing correlation of durations for repetitive construction activities, the authors chose this approach since it is simple to implement. Further work in this aspect of correlation modeling is required to determine the most appropriate model. While the studies by Trofin [3] and Flood et al. [4] were concerned with how stochastic effects in repetitive activities impact project performance, they gave no consideration to correlation. Implicitly their model had correlation set to k = 1.0 in that all repetitions of a given activity had the same duration. Rachmat et al. [2] also gave no consideration to correlation, although in their case correlation was set implicitly to k = 0.0 in that a new duration was generated stochastically for each repetition of a given activity.

III. Approach

A. Modeling Methods

The Monte Carlo method has been implemented within a number of construction planning tools for modeling uncertainty, including the linear scheduling method (LSM) (see, for example, Wyrozębski, and Wyrozębska, [5]). This study will use the SciPy package from IPython [6] for Monte Carlo sampling since it provides a convenient framework for model development and analysis. For this study, the expected durations and variances of different activities will be allowed to differ, while the expected durations and variances for repetitions of a given activity will be fixed.

The LSM methodology represents repetitive activities as a production line in time and space on a two-dimensional graph, such as illustrated in part (b) of Fig. 1. Time, in this illustration, is represented on the horizontal axis and space (location of crew or repetition number) on the vertical axis. The slope of a production line represents its production rate. The slope of the line may be straight if the productivity rate is constant or variable if the productivity changes from unit to unit.

B. Objective Function and Objective Variables

The aim of the study is to determine the impact of correlation between the durations of repeated activities on the optimality of the project plan. Two objective variables were adopted: Crew Idle Time; and Missed Opportunities. Crew Idle Time refers to the total period of time that the crews spend idle or inoperative between their start and finish times. It is caused by interference between dependent crews, and requires one or other of the crews to cease operations intermittently or lower their production rate to avoid any conflict. Fig. 1.a shows two examples where Crew Idle Time would have to be introduced (to activities B and D) to avoid the conflict. In either case, the clash may have been caused by the preceding activity progressing more slowly than expected and/or the succeeding activity progressing more quickly than expected. The Crew Idle Time would be the sum of these two periods of time. Missed Opportunities refer

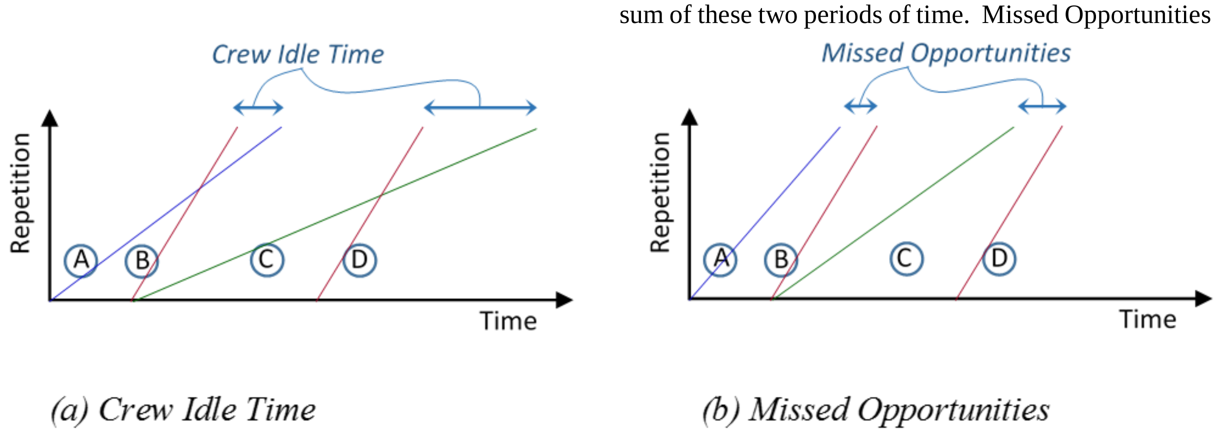

Figure 1. Examples of Inefficiencies in Performance Optimality.

to periods of time that could have been saved in the execution of a project by starting crews earlier. Fig. 1.b includes two examples of Missed Opportunities for completing work sooner, between activities A and B and between activities C and D. The Missed Opportunities would be the sum of these two periods of time.

Missed Opportunities and Crew Idle Time are caused by stochastic and correlation effects that result in crew performance rates that are different from the deterministically derived optimum base plan. This base plan assumes that each activity progresses at its expected rate and the start times for each activity are set to ensure that crews neither spend time idle nor miss any opportunities for starting and finishing work sooner. Crew Idle Time represents an additional direct cost to the project in that the crews are employed for longer periods of time to complete the specified amount of work. Missed Opportunities represent an indirect cost to the project in that they lead to a longer than necessary project duration and therefore result in unnecessary project overhead costs.

C. *Synthetically Generated Test Projects*

Investigation of the research question was completed for a sample of synthetically generated projects, similar to the approach reported by Trofin [3] for assessing the impact of uncertainty on LSM plan optimality. The number of activities in each synthetically generated project was set to 10, a large enough number to permit complicated interactions between crews. Each activity was represented by its own Beta distribution which was used to generate the stochastic component of the duration of each repetition of that activity,

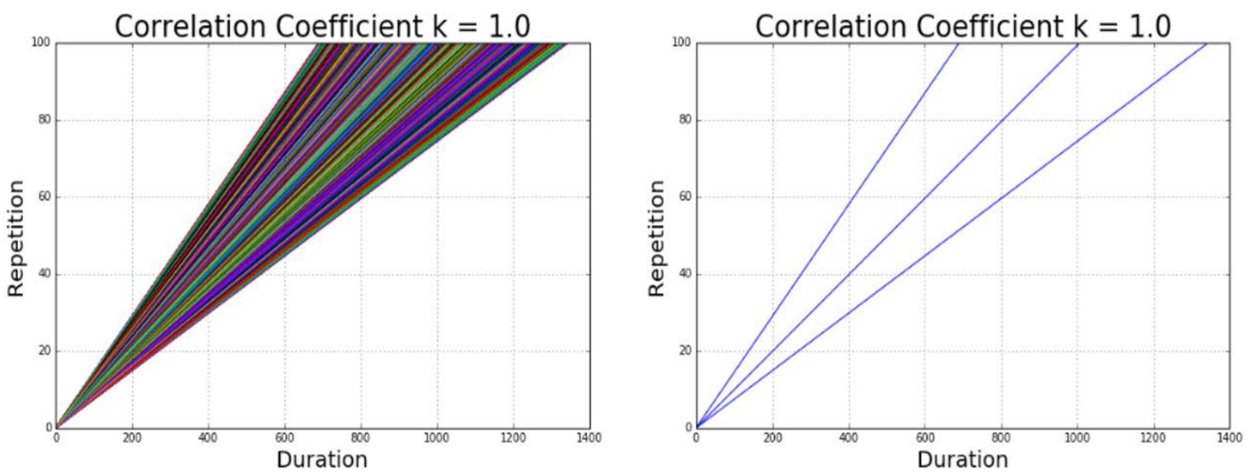

Figure 2. Activity Boundary Generation

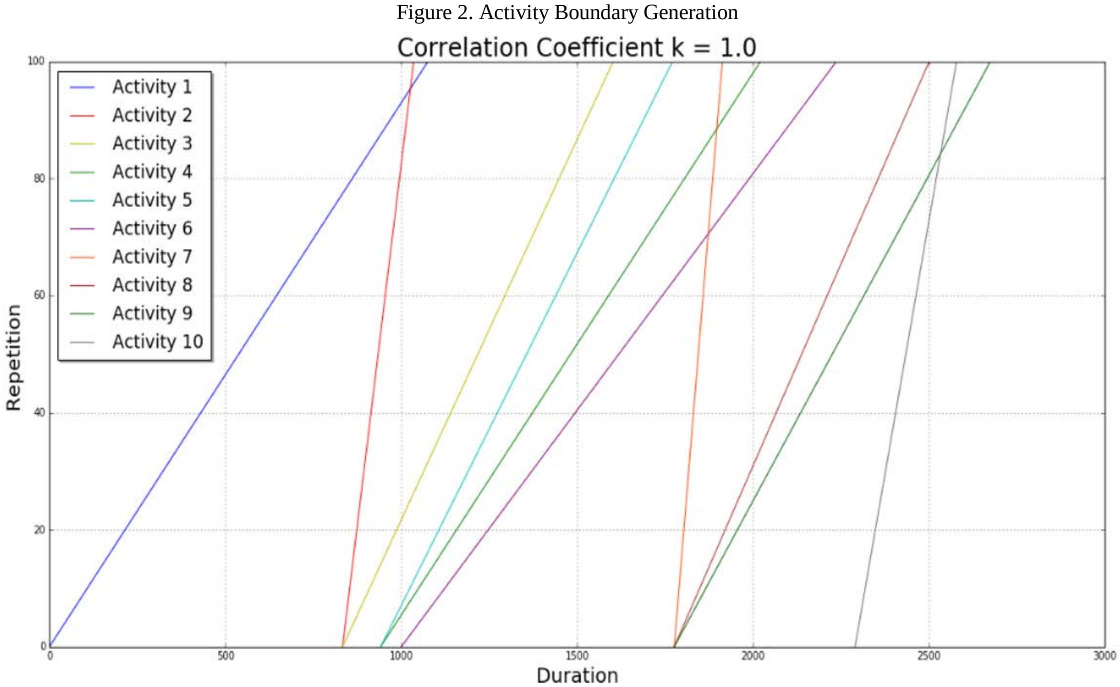

Figure 3. LSM Construction.

the parameter D'_n in (1). The Beta distribution parameters for each activity were selected using Monte Carlo sampling based on the characteristic range of values for these parameters published in AbouRizk and Halpin [7] for earthmoving operations, as detailed in Eiris [8]. For construction simulation, the Beta distribution has been found to provide a good representation of the stochastic variance apparent in construction activities [9].

The synthetically generated 10 activity projects were tested for sensitivity to changes in the level of correlation relative to the deterministically derived optimum base plan. This was undertaken for a range of correlation levels from k = 0.0 to k = 1.0. For each level of k, 1,000 alternative outcomes for the project were generated using Monte Carlo sampling, and each activity was run for 100 repetitions. Fig. 2 shows an example of 1,000 samples for one activity (for k=1.0) and its corresponding boundaries and mean. Fig. 3 shows one of the 1,000 stochastically sampled outcomes for this project, for all 10 activities, where k=1.0. The deviation from the optimal plan in Fig. 3 is apparent by both the clashes between activities (resulting in Crew Idle Time) and the gaps between succeeding activities (amounting to Missed Opportunities), which are numerous. The following section provides a detailed statistical analysis of how such inefficiencies in project performance result from the level of correlation between activity repetitions.

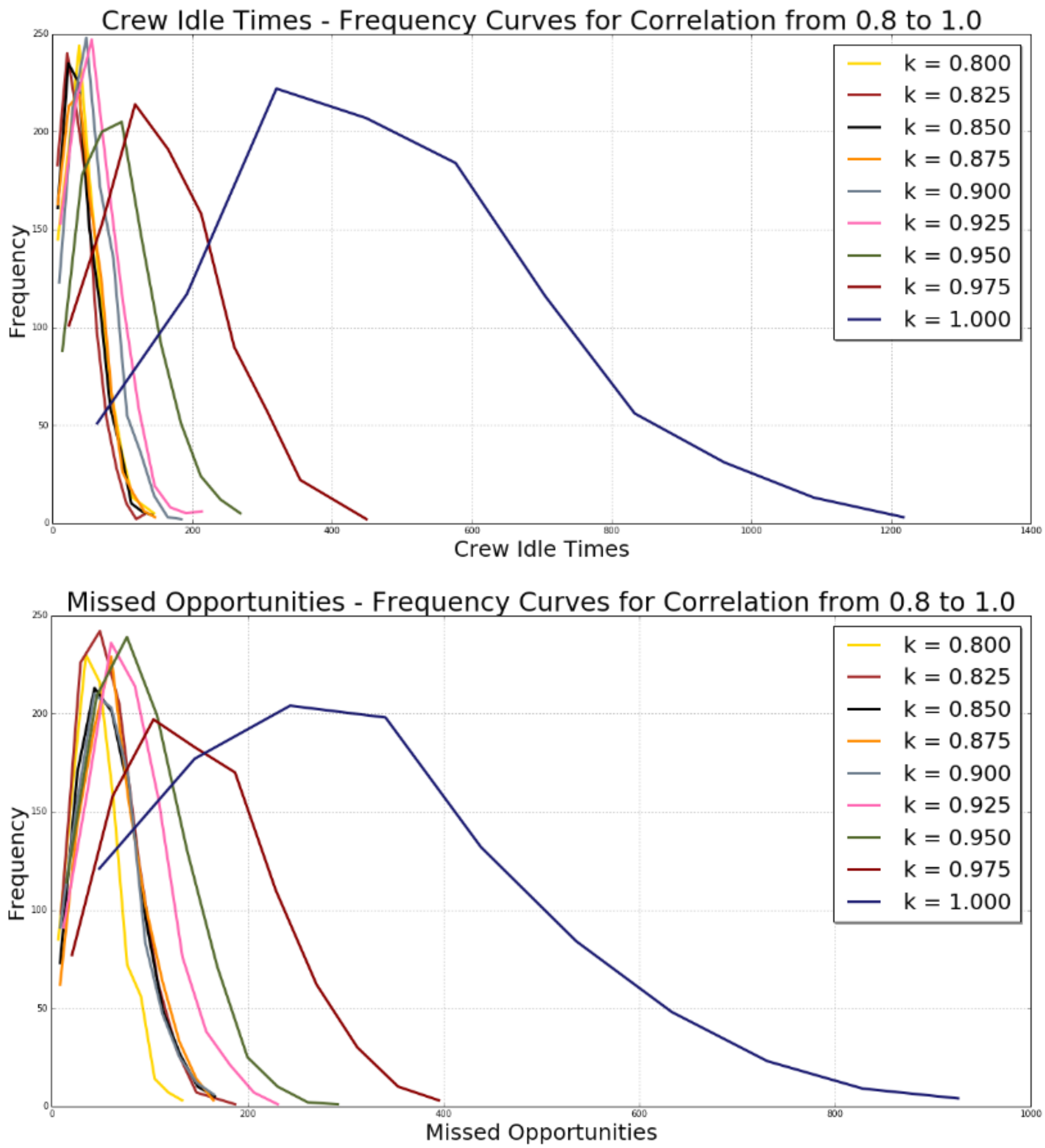

Figure 4. Distribution of Impact of Correlation, k, on *Crew Idle Time* and *Missed Opportunities*.

IV. RESULTS AND ANALYSIS

The results of the experiments described above indicated that lower levels of correlation, between 0.00 and 0.80, did not show a significant impact on either Crew Idle Time or Missed Opportunities. However, for higher levels of correlation both variables were found to increase geometrically. Therefore, an additional 1,000 LSM scenarios were generated for each level of correlation ranging from k = 0.80 to 1.00 in increments of 0.025, to provide a higher resolution in the results for the region where performance was found to change most dramatically.

The results of these experiments are presented in Fig. 4, with the first graph plotting the frequency of occurrence of different Crew Idle Time for different levels of correlation, k, and the second graph plotting the same for Missed Opportunities. Fig. 5 summarizes these results showing, how the mean Crew Idle Time and Missed Opportunities change with respect to the level of correlation. The performance of the project was most severely impacted when correlation was perfect (k=1.0), in which case the Crew Ide Time was found to be on average approximately 7% of the total time the crews were active and the Missed Opportunities were approximately 12% of the optimum base plan's project duration. Both of these values are considered to be significant, warranting further investigation of the nature and impact of correlation on project planning.

The *Crew Idle Time* and the *Missed Opportunities* data were assessed in a single sample t-test to determine the statistical margin of standard error of the sample mean. Each level of correlation was tested independently using a confidence level of 95%. Table 1 and 2 show the results of this test for different correlation levels. For *Crew Idle Time* the average margin of error from the sample mean for all the correlation levels was 3.7%. Similarly, for *Missed Opportunities* the average margin of error from the sample mean for all the correlation levels was 3.61%. Given that these margins of error are all below 5% at the 95% confidence level, the sample size of 1,000 LSM test scenarios was considered sufficiently large.

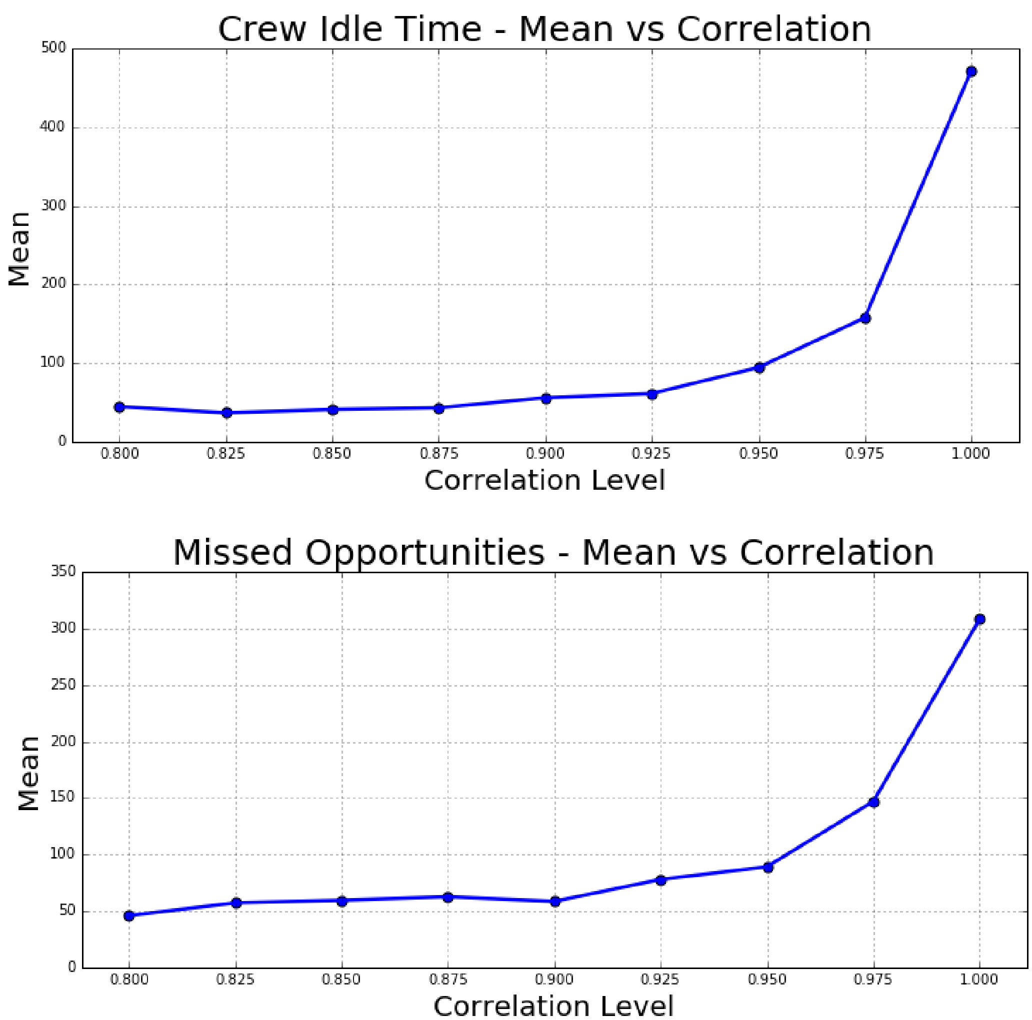

Figure 5. Summary of Impact of Correlation, k, on *Crew Idle Time* and *Missed Opportunities*.

Table 1. EXPERIMENTAL MARGIN OF ERROR AT 95% CONFIDENCE LEVEL, FOR CORRELATION LEVELS 0.0 TO 0.5.

Correlation Level (k)	0.0	0.1	0.2	0.3	0.4	0.5
% Error – Crew Idle Time	3.74	3.69	3.81	3.82	3.73	3.80
% Error – Missed Opportunities	3.76	3.62	3.73	3.65	3.58	3.74

Table 2. EXPERIMENTAL MARGIN OF ERROR AT 95% CONFIDENCE LEVEL, FOR CORRELATION LEVELS 0.60 TO 1.0.

Correlation Level (k)	0.6	0.7	0.8	0.9	1.0
% Error – Crew Idle Time	3.75	3.76	3.92	3.64	3.06
% Error – Missed Opportunities	3.67	3.47	3.32	3.45	3.73

V. CONCLUSIONS AND RECOMMENDATIONS

The impact of correlation between activities on the performance of construction projects is not well understood. Moreover, existing models of correlation are limited in sophistication and largely untested in terms of their accuracy. Before investing resources in the development of more appropriate models of correlation for construction it was decided to first test whether correlation may affect project performance significantly. Specifically, this study had the goal of determining whether the optimality of a project plan is prone to disruption by unaccounted correlation. The study was limited to correlation between repetitions of an activity in a LSM framework, using a simple model of linear compounding correlation. Project performance was assessed in terms of two objective variables: Crew Idle Time and project Missed Opportunities. The results showed that both performance indicators are significantly impacted if the level of correlation is high (between k=0.8 and k=1.0), in the worst case having an expected crew idle time of 7% of crew active time and an expected extension to the project duration of 12%.

These results provide justification for undertaking further work developing our understanding of how correlation can best be modeled in construction and its impact on project performance.

REFERENCES

[1] M. Gates and A. Scarpa. Learning and experience curves. Journal of the Construction Division, ASCE, 98(CO1), March, proceedings papers 8778, pp. 79–101, 1972.

[2] F. Rachmat , L. Song, and S. Lee. “Applying a Stochastic Linear Scheduling Method to Pipeline Construction”, 3rd International Conference on Construction Engineering and Management, ICCEM & ICCPM, May 2009, pp 154-162, 2009.

[3] I. Trofin. “Impact of Uncertainty on Construction Project Performance Using Linear Scheduling”, Master Thesis, Rinker School of Construction, University of Florida, Gainesville, FL, 2004.

[4] I. Flood, R. R. A. Issa, and I. Troffin. “Optimization of Stochastic Linear Schedules”, in proceedings of 11th International Workshop, EG-ICE, Weimar, May 2004, pp 182-189, 2004.

[5] P. Wyrozębski and A. Wyrozębska. “Challenges of Project Planning in the Probabilistic Approach Using PERT, GERT and Monte Carlo”, Journal of Management and Marketing, JMM, Vol. 1, Issue 1, November 2013, pp 1-8, 2013.

[6] F. Pérez and E. G. Granger. “IPython: A System for Interactive Scientific Computing”, Computing in Science and Engineering, Vol. 9, No. 3, pp. 21-29, doi:10.1109/MCSE.2007.53, 2007.

[7] S. AbouRizk and D. Halpin. “Statistical Properties of Construction Duration Data”, Journal of Construction Engineering and Management, 10.1061/(ASCE)0733-9364(1992)118:3(525), 525-544, 1992.

[8] R. Eiris Pereira. “Impact of Linear Correlation on Construction Project Performance Using Stochastic Linear Scheduling”, Master Thesis, Rinker School of Construction, University of Florida, Gainesville, FL, 2016.

[9] S. AbouRizk, D. Halpin, J. and Wilson. “Fitting Beta Distributions Based on Sample Data”, Journal of Construction Engineering and Management, 10.1061/(ASCE)0733-9364(1994)120:2(288), 288-305, 1994.

