Time Series Analysis for Construction Productivity Experiments

By Tariq S. Abdelhamid,[1] Student Member, ASCE, and John G. Everett,[2] Member, ASCE

Abstract: Time series analysis is a powerful statistical tool that can be of great value in evaluating the results of experiments to improve construction productivity. Time series analysis explicitly recognizes the importance of the order in which experimental data are observed, and the statistical dependence of observed data. This paper presents a brief overview of time series analysis and demonstrates its application using previously published data for a series of experiments involving crane lift cycle durations. In one experiment, it was shown that despite no apparent change in productivity, a new technology changed the nature of the work operation. In another experiment, a change in productivity was observed, but little change to the overall operation was seen. In a third experiment, a large change in the complexity of the operation can be seen as well as a clear learning effect. The underlying changes are due to differences in disturbances to the overall operation. These nonproductivity changes may be of great interest to the researcher, but they would not be identified using conventional productivity comparison techniques.

INTRODUCTION

Construction engineers, managers, and researchers spend considerable effort developing new tools, equipment, and techniques in the hope of improving the productivity of construction craft workers and crews. To validate the new technology, some comparison must be made between the old methods and the new methods. Typically, a series of experiments are conducted in which several cycles of work are performed using the old method and several cycles are performed using the new method. Statistical measures such as the mean and standard deviation of the time or cost to perform a work cycle are used to determine if there is a statistically significant change in productivity.

While these statistical tools provide valuable information, they do not capture other important information—in particular, the order in which the experimental data are observed and learning effects. Time series analysis is a statistical technique that can bring out this hidden information.

The objective of the present paper is to investigate the use of time series analysis on data collected from construction productivity improvement studies and how its use may provide insights into the effects of new technologies.

Utilizing this powerful technique adds more robustness and confidence to the exercise of fitting experimental data to mathematical models (e.g., linear regression or learning curves), and to inferences about the system that produced the data. The format in which data are collected for productivity improvement studies lends itself completely to the assumptions dictated by time series analysis, especially the explicit recognition of the importance of the order in which the data are observed, and the statistical dependence of observed data. The order and dependence of the data are the exact two properties that are found in data collected for productivity improvement field experiment.

OVERVIEW OF TIME SERIES ANALYSIS

The following sections are very brief discussions of some of the important aspects of time series analysis as a powerful data analysis technique. Attempting to explain the topic in depth would require far more space than is available here. Therefore, the reader should not expect to become conversant in time series analysis by reading the brief treatment to follow; rather, the reader should expect an overall view of the technique. If interested, the reader should refer to the references on the respective topics for better familiarization.

What Is a Time Series?

When a sequence of data is observed with respect to time or other variables such as space, the collected data are called a time series. An observed time series is considered in statistical theory as one possible realization of a stochastic process (Anderson 1971; Pandit and Wu 1993). To analyze the observed time series, a special statistical methodology known as time series analysis can be used. The importance given to the order in which the data are observed, and the statistical dependence of observed data are the features that make time series analysis different and unique compared to other statistical methodologies. In fact, the theory of a stationary stochastic process or time series is essentially the theory of its correlation. According to Pandit and Wu (1993), ''The mathematical model for the dynamic system, either in continuous or discrete time, reduces the dependent or correlated time series output to the independent or uncorrelated input. The whole methodology can thus be summarized as finding such a model that accomplishes this reduction to independent data and then using standard statistical techniques for independent observation for estimation, prediction, and control.''

The basic idea behind time series analysis is to find a regression model that can represent an observation in time t, denoted as X_t, as the sum of two independent, uncorrelated, or ''orthogonal'' parts: one dependent on the preceding observed data and the other an independent sequence of unmeasured inputs (Pandit and Wu 1993). The family of regression models that may best fit the time series data are termed discrete or continuous autoregressive moving average models, denoted as ARMA(n, m) or AM(n, m) models, respectively.

A general discrete autoregressive moving average model, ARMA(n, m), would have a linear stochastic difference equation of the form

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \cdots - \phi_n x_{t-n} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_m a_{t-m} \quad (1)$$

where $a_t \approx NID(0, \sigma_a^2)$; i.e., a_t is assumed to have a normal distribution, and a_t is independent of a_{t-n}, a_{t-n-1}, ... This also implies that a_t is independent of X_{t-n-1}, X_{t-n-2}, ... The

[1]Grad. Student Res. Asst., Dept. of Civ. and Envir. Engrg., Univ. of Michigan, 1340 G. G. Brown, Ann Arbor, MI 48109-2125. E-mail: tariqa@umich.edu

[2]Asst. Prof. of Civ. and Envir. Engrg., Univ. of Michigan, 2353 G. G. Brown, Ann Arbor, MI. E-mail: everett@umich.edu

Note. Discussion open until September 1, 1999. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on November 21, 1997. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 125, No. 2, March/April, 1999. ©ASCE, ISSN 0733-9634/99/0002-0087–0095/$8.00 + $.50 per page. Paper No. 17117.

J. Constr. Eng. Manage., 1999, 125(2): 87-95

left side of (1) is the autoregressive part, and the right side is the moving average part. The general continuous autoregressive moving average model, the AM(n, m) model, would have a linear stochastic differential equation. General continuous AM(n, m) models will not be discussed in the present paper. Excellent discussions may be found on continuous AM(n, m) models in Pandit (1973) or Anderson (1971).

Fitting and Finding Adequate ARMA(n, m) Models

To fit a particular ARMA(n, m) model from an observed series, the computational effort goes into evaluating the parameters ($\phi_1, \phi_2, \ldots, \phi_n$ and $\theta_1, \theta_2, \ldots, \theta_m$). To do so, linear or nonlinear least-squares parameter estimation is performed depending on the order of the ARMA model being fitted. After fitting many ARMA(n, m) models (following some modeling strategy), statistical tests such as the F-test are carried out to compare the fitted models and to ensure that the final chosen model will accurately represent the data. Other assumptions made concerning the independence of a_ts, and the dependence of an observation X at time t on previous observations must hold. The reader is referred to Pandit and Wu (1993) and Box and Jenkins (1970) for a thorough treatment of the various computational methods and tests involved in the fitting process.

Important Characteristics of ARMA Models

After a specific ARMA(n, m) model is fitted and chosen as the adequate model to fit the data, many important characteristics of this ARMA(n, m) model can be derived to make more inferences about the system that produced the data being analyzed. In fact, it is these characteristics that could extract all the otherwise overlooked information, such as the physical characteristics of the system, and the ''memory'' in the data to disturbances imparted to the system while it is functioning. The gained information from these characteristics makes them rather intriguing and appealing. However, before any important characteristics of ARMA models may be derived, values called characteristic roots must first be calculated. For a general ARMA(n, m) model, the following may be used to calculate the autoregressive and moving average characteristic roots, respectively:

$$\begin{Bmatrix} \phi_l = (-1)^{l+1} \displaystyle\sum_{\substack{i1,i2,\ldots il=1 \\ i1\angle i2\angle\ldots\angle il}}^{n} \lambda_{i1}\,\lambda_{i2}\ldots\lambda_{il} \\ l = 1, 2, \ldots, n \end{Bmatrix} \text{ and }$$

$$\begin{Bmatrix} \theta_l = (-1)^{l+1} \displaystyle\sum_{\substack{i1,i2,\ldots il=1 \\ i1\angle i2\angle\ldots\angle il}}^{m} \nu_{i1}\,\nu_{i2}\ldots\nu_{il} \\ l = 1, 2, \ldots, m \end{Bmatrix} \tag{2}$$

Perhaps the most important characteristics of ARMA models are the ability to derive stability conditions criteria and invertibility conditions criteria for a system, the Green's function, and the autocovariance function from which the autocorrelation for the data may be readily calculated. Other characteristics such as the inverse function, partial autocorrelation, and spectrum will not be discussed in the present paper.

Stability Conditions

Stability conditions of a system refer to whether or not the system being observed will return to its ''equilibrium'' position after a disturbance a_t affecting it is complete. Depending on whether certain stability conditions are met or not, a system may be characterized as ''stable,'' ''asymptotically stable,'' or ''unstable.''

Green's Function

Green's function ''G_j'' summarizes the dependence or memory of an observed value X_t on previous a_ts. When a system is affected by the disturbances, a_ts, it is important and crucial to understand how a system response or output at time t is affected by or depends on previous values of the disturbances, a_ts, especially before and after a manipulation of some kind will be applied under some experimental setting.

In the present paper, use of the term ''Green's function'' is consistent with the terminology found in Pandit and Wu (1993). However, it should be noted that Green's function is also referred to in the literature on time series analysis as ''weighting function,'' in Wiener (1949), Pugachev (1957), and Blum (1965), or as ''ψ weights'' in Box and Jenkins (1970).

Autocovariance/Autocorrelation Function

The autocovariance function, denoted as γ_k, and autocorrelation function, denoted as ρ_k ($\rho_k = \gamma_k/\gamma_0$), are characterizations of the statistical dependence between the sequences or series of observations X_t, X_{t-1}, X_{t-2}, ... On one hand, the autocovariance function plays an important role in determining the true variance of the observed series X_t by calculating γ_0. It also provides the individual contribution of characteristic roots to the variance of the observed series X_t; i.e., how is each characteristic root affecting or contributing to the system output? On the other hand, plots of the autocorrelation function will show whether the outputs of the system observed are correlated and dependent. This may be regarded as a check on the validity of the use of the time series approach or its implied assumptions.

CRANIUM Study

Everett and Slocum (1993) introduced a new device to improve crane productivity and safety. The device, called the ''CRANIUM,'' is a video system that allows a crane operator to have direct real-time visual feedback of what is happening at the lifting point, even in situations where a direct line of sight does not exist. The CRANIUM eliminates the need for a craft worker known as the tagman, whose function is to duplicate hand signals given by a craft worker at the lifting point, thereby relaying the hand signal to the crane operator. The tagman introduces a delay in the transmission of information, causing productivity to suffer. The possibility of introducing errors in the transmission creates a safety hazard and forces crane operators to work more slowly than they would work with a direct line of sight. In the present paper, we do not address the issue of safety, but rather focus on the productivity impact results of the study.

Everett and Slocum (1993) reported on three field experiments of varying complexity that were performed with and without the CRANIUM system to validate its impact on productivity. The three experiments were labeled A, B, and C, and each required the crane operator to perform several repetitive cycles of a load lifting and unloading operation. The experiments were designed so that the crane operator could see neither the target nor the signalman. A time break of 2 h between the set of control lifts and the set of CRANIUM lifts reduced the chance for learning effects to be transferred between sets of lifts. The experiments were conducted at the Operating Engineers' Training Center in Canton, Mass. A Lorain 45 metric ton (50 ton) truck crane with 18 m (60 ft) of boom was used to perform the experiments. Two apprentice operating engineers served as signalman and tagman.

Experiment A ''simulated construction work, such as demolition or clamshell excavation, where accuracy is not critical

J. Constr. Eng. Manage., 1999, 125(2): 87-95

and high-impact landings are permissible or even desirable. The crane operator was required to position an empty concrete bucket in a target area (d = 3 m) on the ground on the far side (relative to the crane operator) of a construction trailer. The return portion of the cycle required swinging the crane and bucket 90° to a target that was in view of the operator, eliminating the need for a second signalman-tagman team during the experiment'' (Everett and Slocum 1993).

In experiment B, the crane operator was required to lower an empty concrete bucket over an embankment and touch down on a target block (0.5 m × 0.5 m) at the base of the embankment. ''The return portion of the cycle required swinging the crane and bucket 90° to target on the top of the embankment in view of the operator. In experiment B, the tagman could not anticipate hand signals from the signalman, but could only relay the hand signals after he saw them. This situation is typical of many repetitive lifts on a construction site. Everyone involved knows where the load is going, but each person in the communication chain must wait for information from someone else'' (Everett and Slocum 1993).

A lift requiring more accuracy and difficulty was required in experiment C. The lift ''required positioning a large (1 m^3) concrete block onto the ground on the far side (relative to the crane operator) of a construction trailer. The goal was to gently position one corner of the concrete block within 5 cm of a target cone'' (Everett and Slocum 1993). The target cone was moved between lifts so that the operator would not know the exact spot of the target; rather, he would have only a rough idea of where the target would be. The difficulty of the lift in experiment C was due to the need to boom up and down to adjust for the change in radial distance between the target cone and the crane. ''Booming up raises the boom, moving the load up and toward the crane operator. Booming down lowers the boom, moving the load down and away from the operator. Frequently the load had to be hoisted before booming down, to prevent dragging the load on the ground. This wastes the time previously spent carefully lowering the load. The added complexity of the lifts not only increases the mean cycle duration, but leads to learning effects. After several cycles, the signalman improves his ability to correlate booming up and down with raising and lowering the load'' (Everett and Slocum 1993).

TABLE 1. Experiment A, B, and C Collected Data and Results (Everett and Slocum 1993)

Lift number (1)	Experiment A, Lift Duration(s)		Experiment B, Lift Duration(s)		Experiment C, Lift Duration(s)	
	Control (2)	CRANIUM (3)	Control (4)	CRANIUM (5)	Control (6)	CRANIUM (7)
1	48	48	68	65	130	105
2	46	50	64	52	125	81
3	43	48	75	61	122	110
4	49	51	75	56	120	96
5	46	51	63	68	123	78
6	50	46	75	56	134	92
7	47	45	78	53	95	87
8	48	49	71	68	114	110
9	48	43	57	53	94	89
10	50	50	68	51	128	90
11	48	40	73	62	105	97
12	52	47	65	61	97	103
13	48	44	NM	56	120	76
14	46	47	—	63	88	90
15	NM	48	—	54	78	79
16	—	47	—	54	98	99
17	—	48	—	57	94	85
18	—	44	—	69	NM	78
19	—	48	—	57	—	89
20	—	46	—	52	—	75
21	—	47	—	55	—	62
22	—	48	—	59	—	74
23	—	43	—	66	—	60
24	—	44	—	56	—	77
25	—	48	—	NM	—	83
26	—	52	—	—	—	79
27	—	46	—	—	—	75
28	—	43	—	—	—	69
29	—	46	—	—	—	63
30	—	49	—	—	—	NM
31	—	49	—	—	—	—
32	—	45	—	—	—	—
33	—	52	—	—	—	—
34	—	51	—	—	—	—
35	—	47	—	—	—	—
36	—	48	—	—	—	—
37	—	46	—	—	—	—
38	—	47	—	—	—	—
39	—	48	—	—	—	—
40	—	47	—	—	—	—
Average	47.79	47.15	69.33	58.50	109.71	84.52
Standard deviation	2.19	2.65	6.26	5.59	17.02	13.51

Note: NM indicates no other measurements were performed.

Table 1 summarizes the observed data for experiments A, B, and C for both the control (normal practice using a tagman) and the CRANIUM case. From the means and standard deviations of the experimental data, Everett and Slocum concluded that for experiment A, no statistically significant change in productivity was observed. For experiment B, a 16% improvement in productivity (reduction in lift cycle duration) was observed using the CRANIUM compared to the control. For experiment C, a 21% improvement in productivity was observed.

The data collected from these experiments will be used in the present paper to demonstrate the use of time series analysis and to show how additional information can be derived beyond the statistical testing of the differences between means. For a more detailed account of the experiments, the reader is referred to Everett and Slocum (1993).

METHODS

Time Series Analysis of CRANIUM Study

The obvious question that comes to mind is: What insights and conclusions could time series analysis provide that are different from what the data reported earlier have already shown? First, we must determine if the observed data are dependent or independent by calculating the autocorrelation function, and preferably the theoretical autocorrelation. This can be done using time series analysis techniques. If the observed data are truly dependent, then what are the true values of the standard deviations that should be reported, and how may this affect the conclusions about the CRANIUM?

Second, if we consider the almost equal means and standard deviations for the control and CRANIUM cases of experiment A, does an inference based on this similarity only justify the conclusion that there is no effect on the productivity or on the workers? Similarly, if we consider the different means and standard deviations for the control and CRANIUM cases of experiments B and C, does an inference based on these differences justify the conclusion that an improvement in productivity has been accomplished? In the following sections, we address these questions using time series analysis.

Crane Operations As Stochastic Processes

To use time series analysis on the process of lifting using conventional cranes, it will be assumed that there are discrete stochastic disturbances (NOISE) that affect the operation of the crane, including wind, visibility, operator psychomotor abilities, experience of operator and crew, and so on. Thus, the lifting duration will be assumed to result from a stationary stochastic process, and the measured (observed) lifting duration data are a realization of that process. The main underlying assumptions in this modeling are that the process is modeled

J. Constr. Eng. Manage., 1999, 125(2): 87-95

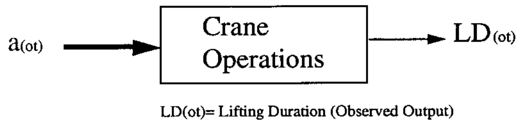

FIG. 1. Time Series Presentation of Lifting Duration While Performing Crane Operations under Disturbances

as a single series of data, and that while performing lifting operations, the crane operator and crew are subjected to the disturbances (NOISE) that cause fluctuation in the observed lifting durations.

Fig. 1 shows the formulation of these assumptions. In Fig. 1, the variable a_{0t} represents the disturbances (NOISE) to the output (observed lifting durations). More specifically, the formulation is intended to relate the observable output (the lifting duration) to its preceding values, and the nonobservable (or nonmeasured) disturbances to the lifting durations. In other words, the lifting duration at time t can be decomposed into a part that depends on its preceding values, and on the disturbances as one combined disturbance a_{0t}.

Data Representation for Time Series Analysis

As shown in Table 1, the sample size of the data collected for the control experiments A, B, and C consisted of 14, 17, and 12 data points, respectively. Similarly, the sample size of the data collected for the CRANIUM experiments A, B, and C consisted of 40, 29, and 24 data points, respectively. To analyze observed data in time series, a sufficiently large sample size (data points $\geq$ 50) is needed. Clearly, all the sample sizes are considered ''small'' for time series analysis, since each is less than 50 data points. This is typical of most construction productivity improvement studies.

One commonly used solution to enable the analysis of a small sample size using time series is to enlarge the sample size by repeating the series (Box and Jenkins 1970). The repeating process may take the form $(x_1, x_2, \ldots x_n)$, $(x_2, x_3, \ldots x_n)$, $(x_3, x_4, \ldots x_n)$, and so on for a series of n data points. This repeating process was chosen so that the effect of the induced trend caused by repeating the series would cancel out (Box and Jenkins 1970; Anderson 1971).

To compare the CRANIUM to the control for each experiment, the smaller sample size was chosen to be repeated in the manner discussed previously. For example, for experiment A, the control experiment had 14 points; thus, the first 14 data points from the CRANIUM data were used and the rest were disregarded. Consequently, the sample sizes for the control and CRANIUM experiments A, B, and C became 105, 78, and 114 observations, respectively. The reason behind choosing the smaller sample size is to avoid bias in the analysis of the CRANIUM data, which had the larger sample sizes. The bias would result from the fact that the more data you sample, the more system dynamics that are being captured in the data, which would be extracted by the time series analysis. In other words, it is possible that if a different number of observations were collected for the control cases than for the CRANIUM cases, the results would have shown a different trend in both the mean and the standard deviation.

Given this, the standard deviations for the CRANIUM data must be recalculated based on the new data set sizes, 14, 17, and 12, respectively. The recalculated standard deviations are 3.22, 6.16, and 10.78 for experiments A, B, and C, respectively. Figs. 2–4 show the plots of the repeated raw data to enable time series analysis for experiments A, B, and C, respectively.

Finding ARMA(n, m) Adequate Model

The procedure for finding the adequate autoregressive moving average model (ARMA) to represent the data was carried out by fitting the AR(1), AR(2), AR(3), and ARMA(2, 1) models. The ARMA(1, 1) model was only fitted when the ARMA(2, 1) parameter $\phi_2 = 0$ or its confidence interval contained zero. Although this procedure is slightly different from the strategy $(n, n - 1)$ or $(2n, 2n - 1)$ recommended by Pandit and Wu (1993), it was considered practical and applicable to the objective of the present paper.

In fitting the ARMA(2, 1) model, the initial values were obtained by using the inverse function approach outlined in Pandit and Wu (1993). This approach resulted also in fitting the AR(3) model ''automatically.'' The initial values were used as final values for the ARMA(2, 1) model, rather than finding them using the nonlinear least-squares estimate approach. This assumption was reasonable, again relative to the objective and scope of the present paper.

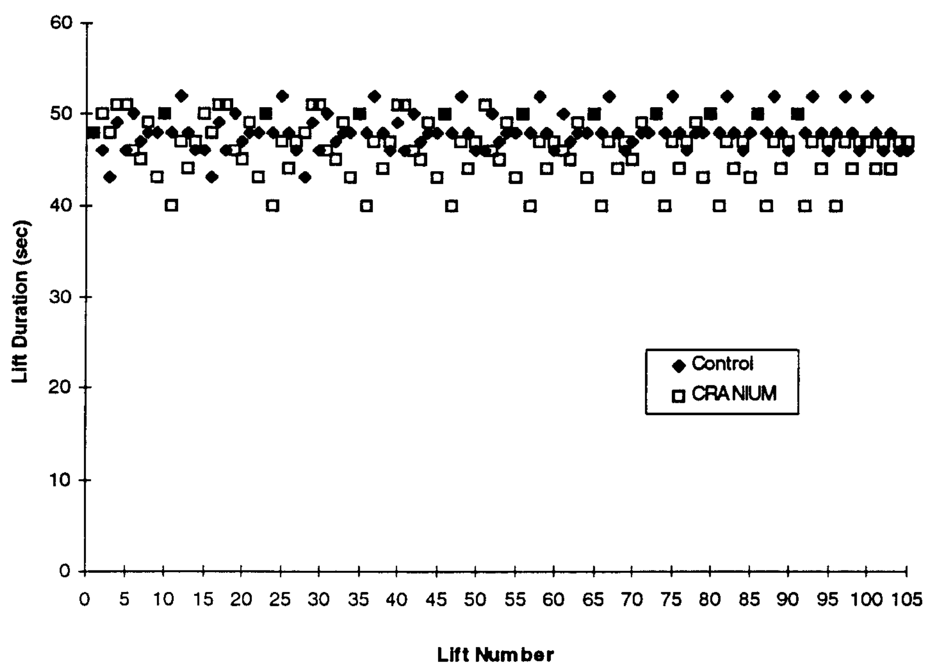

FIG. 2. Data Representation for Experiment A (Both Cases)

J. Constr. Eng. Manage., 1999, 125(2): 87-95

In addition, from the shape of the graphed observations in Figs. 2–4, it was expected that all the time series obtained were stable/stationary. Therefore, in calculating a_ts, after a model is fitted and its parameters estimated, a_t was set equal to zero for $t = 1-n$ (where n is the order of the chosen model), and the computations of a_ts were started from $t = n + 1$. This method involves some loss of information, but was considered preferable and recommendable by Pandit and Wu (1993).

After arriving at the adequate model, it was used to find out some of the characteristics of the ARMA models mentioned before. Although the characteristics are usually reported for physical systems such as machines, they may be used to understand how the crane crew were affected by the various disturbances impacting the crane lifting operations. The characteristics reported include the stability conditions, Green's function, and the autocovariance and theoretical autocorrelation functions. Green's function will play an important role in the analysis to follow, and it is important to note why.

As mentioned before, Green's function G_j summarizes the dependence or memory of an observed value X_t on previous values of a_{0t}. In performing the experiments with and without the CRANIUM, Everett and Slocum (1993) affected the disturbances, a_{0t}, by changing the crew makeup and methods employed, the overall procedure, and maybe even the skills required. Therefore, it is important to understand how an observed lifting duration at time t was affected by or depends on previous values of the disturbances, a_{0t}, before and after they were manipulated, despite the fact that we do not necessarily know which variable has been manipulated.

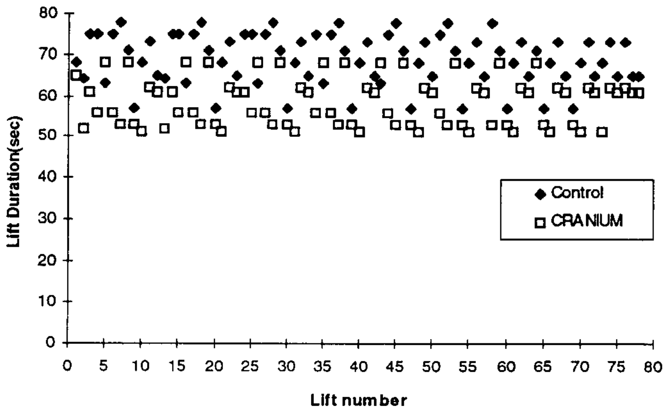

FIG. 3. Data Representation for Experiment B (Both Cases)

RESULTS AND ANALYSIS

Despite the soundness of the results obtained by Everett and Slocum (1993), it will be shown that by using the time series approach, more information can be inferred from the data through the ARMA models and, particularly, Green's function and the autocovariance and autocorrelation functions.

Detailed Time Series Analysis for Experiment A (Control Case)

Due to space limitations, the detailed time series analysis results for all experiments will not be shown; however, the

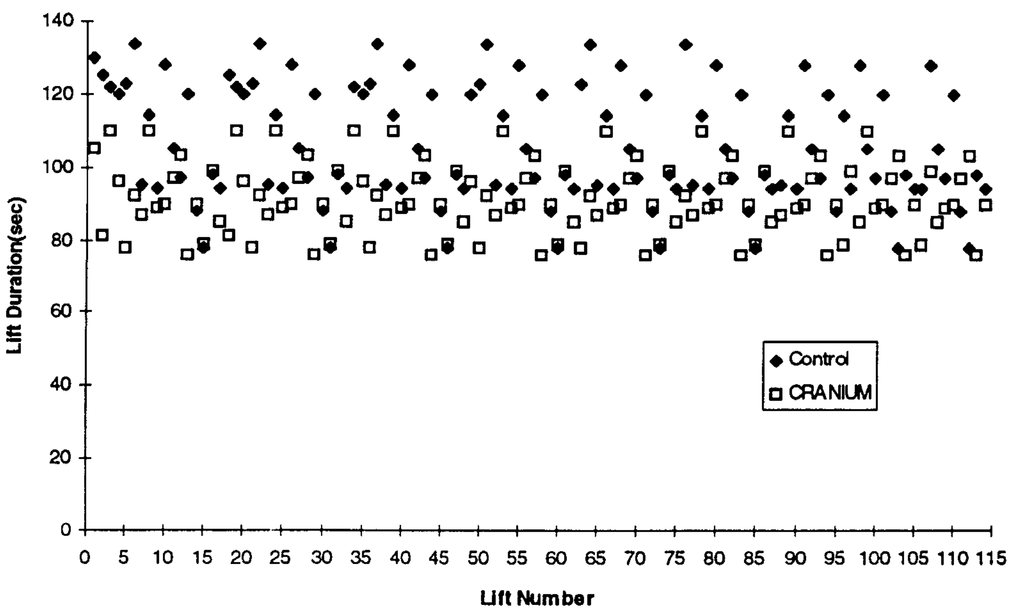

FIG. 4. Data Representation for Experiment C (Both Cases)

TABLE 2. Results of Model Fitting and Checks of Adequacy for Experiment A (Control): ARMA Models

Parameters (1)	AR(1) (2)	AR(2) (3)	AR(3) (4)	ARMA(2, 1) (5)	ARMA(1, 1) (6)
Φ1	−0.0787 ± 0.191	−0.0757 ± 0.191	−0.0975 ± 0.191	−0.78 ± 0.137	−0.431 ± 0.115
Φ2	NA	0.0536 ± 0.191	0.0397 ± 0.191	0.00716 ± 0.161	NA
Φ3	NA	NA	−0.119 ± 0.189	NA	NA
θ_1	NA	NA	NA	−0.7 ± 0.23	−0.333 ± 0.11
σ^2	4.034	3.976	3.64	4.055	4.02
RSS	419.5	413.471	378.522	421.669	437.56
Comparing RSS	—	—	Adequate	—	—
F-test	—	—	—	—	Adequate
Confidence intervals	Φ1 ≈ 0	Φ1, Φ2 ≈ 0	Φ1, Φ2, Φ3 ≈ 0	Φ2 ≈ 0	—
Chosen model	—	—	—	—	ARMA(1, 1)

Note: NA stands for "not applicable."

J. Constr. Eng. Manage., 1999, 125(2): 87-95

TABLE 3. Final Time Series Analysis Results of All Experiments Not Including Green's Functions and Autocovariance Functions

Item (1)	Experiment A: Control (2)	Experiment A: CRANIUM (3)	Experiment B: Control (4)	Experiment B: CRANIUM (5)	Experiment C: Control (6)	Experiment C: CRANIUM (7)
Adequate ARMA model, (RSS, σ_a^2)	ARMA(1, 1) (437.56, 4.2)	AR(2) (763.945, 7.346)	ARMA(2, 1) (2,232.026, 28.98)	AR(2) (1,690.02, 21.95)	ARMA(2, 1) (27,516.17, 243.51)	ARMA(1, 1) (9,725.35, 86.065)
Parameters Φ_i, θ_i	(−0.431, −0.333)	(−0.24, 0.373)	(−0.202, −0.474, −0.181)	(−0.458, −0.514)	(0.263, 0.155, 0.249)	(−0.143, 0.18)
Roots	$\lambda_1 = -0.431$	$\lambda_1, \lambda_2 = 0.5, -0.74$	λ_1, λ_2 (Complex conjugates) = $-0.101 \pm 0.6810i$	λ_1, λ_2 = Complex conjugates, = $-0.229 \pm 0.679i$	$\lambda_1, \lambda_2 = 0.55, -0.28$	$\lambda_1 = -0.143$
Stability	Stable	Stable	Stable	Stable	Stable	Stable
Autocorrelation function $\rho_k = \gamma_k/\gamma_0$	Plot displayed dependence of data	Plot displayed dependence of data	Plot displayed dependence of data	Plot displayed dependence of data	Plot displayed dependence of data	Plot displayed dependence of data

TABLE 4. Green's Functions and Autocovariance Functions for All Experiments

Experiment (1)	Setting (2)	Green's function (3)	Autocovariance function (4)
A	Control	$G_j = (-0.102)(\lambda_1)^{j-1}$	$\gamma_k = (0.43)\lambda_1^{(k)}$
A	CRANIUM	$G_j = [1/(0.5 + 0.74)][0.5^{j+1} - (-0.74)^{j+1}]$	$\gamma_k = 2.88\lambda_1^k + 7.07\lambda_2^k$
B	Control	$G_j = (0.69)^j 1.02 \cos(1.72j + 0.12)$	$\gamma_k = (18.7258 - 1.6216j)\lambda_1^k + (18.7258 + 1.6216j)\lambda_2^k$
B	CRANIUM	$G_j = (0.72)^j 1.24 \cos(1.89j + 0.325)$	$\gamma_k = (16.5004 + 1.7854j)\lambda_1^k + (16.5004 - 1.7854j)\lambda_2^k$
C	Control	$G_j = [1/(0.55 + 0.28)][0.55^{j+1} - (-0.28)^{j+1}]$	$\gamma_k = 94.68\lambda_1^k + 156.01\lambda_2^k$
C	CRANIUM	$G_j = (-0.323)(\lambda_1)^{j-1}$	$\gamma_k = (-29.11)\lambda_1^{(k)}$

detailed modeling steps will be shown for the experiment A control case (see Table 2). The final time series analysis results for all experiments are summarized in Tables 3 and 4.

Parameter Estimation

Various models were fitted (parameter estimates performed) for the data representing the experiment A control case, after which a choice of model had to be made. The choice is based on either comparing the residual sum of squares (RSS), the F-test, or considering the confidence intervals for the estimated parameters (Pandit and Wu 1993). For parameter estimates that contain zero within their confidence intervals, this implies that the parameter can be considered to equal zero. A sample of the F-test procedure, and confidence intervals calculations is shown next.

Checks of Adequacy Using F-Test, and Confidence Intervals

AR(1) and AR(2)

$$F_{\text{computed}} = \frac{A_1 - A_0}{s} \div \frac{A_0}{(N - r)} \tag{3}$$

where $A_1 = 419.5$ [RSS for AR(1), see Table 2]; $A_0 = 413.471$ [RSS for AR(2), see Table 2]; s = number of parameters that can be saved from low order model to high order model = (2 − 1) = 1; r = total number of unrestricted parameters (for model with higher order) plus mean = 2 + 0 + 1 = 3 (including mean); N = number of total observations = 105; $\therefore F_{\text{computed}} = 1.487$; $F_{\text{table}[s,(N-r)]} = F_{\text{table}[1,(102)]} = 3.92$; and $\because F_{\text{table}} > F_{\text{computed}}$; then the test shows no statistical significance, and the AR(1) model may be used.

For $\hat{\Phi}_1$, the 95% confidence interval for the AR(1) model is

$$\hat{\Phi}_1 \pm 1.96\sqrt{\frac{(1 - \hat{\Phi}_1^2)}{N}} \tag{4}$$

which comes out to -0.0787 ± 0.191.

From the foregoing information, the F-test shows no significance and AR(1) may be considered adequate. Although AR(1) seems adequate from the F-test, the confidence interval for its ϕ_1 parameter includes zero. Hence, ϕ_1 may be considered zero. This indicates the possibility of having an AR(0) model, which means that an observation at time t is independent or uncorrelated with preceding observations. However, the fitting of higher order ARMA models shows a decrease in the RSS value, indicating that an AR(0) model is not adequate. Following similar procedures with the other fitted models, the adequate model is found to be ARMA(1, 1).

Characteristic Roots

Calculation of characteristic roots for the adequate ARMA(1, 1) model is relatively simple. The characteristic root corresponding to the autoregressive part of the ARMA(1, 1) model is only shown here, since we are not interested in deriving the inverse function.

For an ARMA(1, 1) model, $\lambda_1 = \phi_1$

$$\therefore \lambda_1 = -0.431 \tag{5}$$

Stability Based on Characteristic Roots

The conditions for asymptotic stability based on the characteristic roots are

$$|\lambda_1| \le 1 \tag{6}$$

$$\because \lambda_1 = -0.431 \tag{7}$$

According to the value of the characteristic roots shown previously, the system observed is stable. This means that the crane crew will reach a fixed value of productivity (lifting duration) at some point in time and stay on it for the rest of the observed cycles.

Green's Function

For the ARMA(1, 1) model, the Green's function takes the form $G_j = (\phi_1 - \theta_1)\phi_1^{j-1}$ (Pandit and Wu 1993). By referring to the values of ϕ_1, θ_1 from Table 4

$$\therefore G_j = (-0.431 + 0.333)(-0.431)^{j-1},$$

$$\text{or, alternatively, } G_j = (-0.102)(\lambda_1)^{j-1} \tag{8}$$

J. Constr. Eng. Manage., 1999, 125(2): 87-95

Autocovariance Function

For an ARMA(1, 1) model, the autocovariance function takes the following form:

$$\gamma_k = \frac{\sigma_a^2(\phi_1 - \theta_1)(1 - \phi_1\theta_1)\phi_1^{(k-1)}}{(1 - \phi_1^2)}, \text{ for } k \geq 1 \tag{9}$$

or, alternatively, and by substituting values of ϕ_1, θ_1, and putting $\phi_1^k = \lambda_1^k$ in $\gamma_k = \sigma_a^2(\phi_1 - \theta_1)/\phi_1^2[\phi_1 - \theta_1/1 - \phi_1^2 + \theta_1]\phi_1^k$

$$\therefore \gamma_k = (0.43)\lambda_1^{(k)}, \quad \text{for } k \geq 1 \tag{10}$$

Also

$$\gamma_0 = \frac{\sigma_a^2(1 - 2\phi_1\theta_1 + \theta_1^2)}{(1 - \phi_1^2)} = 4.24 \tag{11}$$

which is the actual variance of the observed data in experiment A—control case. The value reported by Everett and Slocum was 4.84, resulting in an overestimate of 14.15% [(4.84 − 4.24)/4.24].

Results for All Experiments

In the following sections, the results of the time series analysis for all experiments will be compared to the results shown by Everett and Slocum. This comparison will be made by discussing the implications of satisfying the stability conditions, obtaining the autocovariance functions, the autocorrelation functions, and the Green's functions.

Stability and Autocorrelation Results

As shown in Table 3, the stability conditions were consistently satisfied for all the ARMA models used. This means that after a disturbance a_t affects the crane operations, productivity (lifting durations), the productivity will return to a constant value.

Fig. 5 shows the autocorrelation function ($\rho_k = \gamma_k/\gamma_0$) for the experiment A CRANIUM case. Similar plots, corresponding to the ARMA model chosen for each of the other five experiments, were developed and found to support the assumption that the observations X_t, X_{t-1}, X_{t-2}, . . . , demonstrated statistical dependence. Due to space limitations and the consistent conclusions the plots indicate, the other plots of the autocorrelation functions are not shown here. As mentioned earlier, the demonstrated statistical dependence of the data confirms the validity of using the time series approach to analyze the observed crane operations' productivity.

Autocovariance Functions

Perhaps the most important feature of the autocovariance function is its use in determining the true variance of the observed series X_t, by calculating γ_0. The value of γ_0 will be

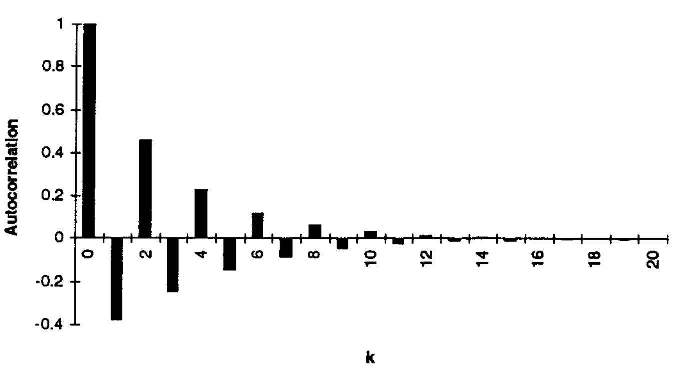

FIG. 5. Autocorrelation Function Plotted for *k* = 20 (Experiment A—CRANIUM)

TABLE 5. Comparison of Variance Reported versus Time Series Results

Experiment (1)	Variance reported in Everett and Slocum (1993) (2)	Variance from time series analysis (3)	Variance reported relative to time series analysis results (4)
A—Control	4.84	$\gamma_0 = 4.24$	Overestimated
A—CRANIUM	(7.02[a]), (10.38[b])	$\gamma_0 = 9.55$	Overestimated
B—Control	39.19	$\gamma_0 = 37.45$	Overestimated
B—CRANIUM	(31.92[a]), (37.97[b])	$\gamma_0 = 33$	Overestimated
C—Control	289.68	$\gamma_0 = 250.787$	Overestimated
C—CRANIUM	(182.25[a]), (116.28[b])	$\gamma_0 = 95.23$	Overestimated

[a]Value as originally reported in Everett and Slocum (1993).
[b]Value based on modified data size.

compared to the variance derived by squaring the standard deviation for the entire original data set from Everett and Slocum. Table 5 clearly shows that the variance in the observed data for all experiments was overestimated, compared to the values found by applying the time series analysis technique. This indicates that the use of time series techniques would give a more accurate reflection of whether improvements were truly achieved due to the use of the CRANIUM for a particular experiment.

For example, it was mentioned earlier that Everett and Slocum (1993) concluded that for experiment A, no difference in productivity (lift cycle duration) was demonstrated between the control case and the CRANIUM case. However, in light of the comparison made in Table 5, it seems that the control case had considerably less variability than the CRANIUM case. It is worth noting that experiment A was the only experiment that shows such result. This result will be mentioned again when Green's function is discussed. The results in Table 5 confirm what Everett and Slocum concluded about the improvement in productivity when using the CRANIUM.

Green's Function

Green's Function Plots for Experiment A. Figs. 6 and 7 show plots of Green's function corresponding to the experiment A—control and CRANIUM cases, respectively. From Fig. 6, it can be seen that a disturbance a_t to the lifting duration

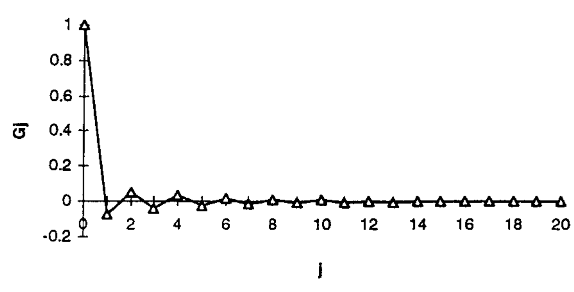

FIG. 6. Green's Function of ARMA(1, 1) (Experiment A—Control)

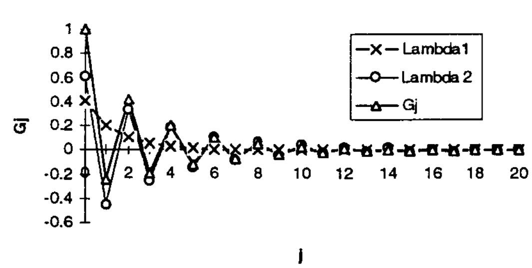

FIG. 7. Green's Function of AR(2) Model as Sum of Two Exponentials (Experiment A—CRANIUM)

J. Constr. Eng. Manage., 1999, 125(2): 87-95

in the control case is ''forgotten,'' or its effect subsides both in magnitude and duration after approximately the eighth lift. Fig. 7 shows that a disturbance a_t to the lifting duration in the CRANIUM case is forgotten, or its effect subsides both in magnitude and duration after approximately the eleventh lift. In addition, in the CRANIUM case, the structure of Green's function is a sum of two exponentially decaying parts, which is quite different from that in the control case. Also, the magnitude of the effect of a disturbance a_t is higher in the CRANIUM case than in the control case.

These differences may arise from several possibilities. For example, experiment A may have been quite well learned or required less attention in the control case than in the CRANIUM case, which was a change from what the crane operator was used to. The smooth curve showing the effect of λ_1 on Green's function indicates that the crane's crew response is affected by an exponentially decaying function. The characteristic root λ_1 could be capturing the effect of the experience of the crew in performing experiment A, or the time the crew have spent together as one crew, or many other reasons. Also, λ_1 seems to reflect a learning curve effect, in comparison to the relatively uneven effect of λ_2 shown in Fig. 7 (the CRANIUM case), which might reflect disturbances coming from wind, motor control problems, and likewise.

The only way to identify what the characteristic roots λ_1 and λ_2 represent in a more precise fashion is to be able to measure more variables as time series themselves. Under these situations, the crane operation could be considered as a multi-input, single-output process with associated disturbances, a_ts. Moreover, by simply observing and measuring more disturbing variables to the crane operation process, the disturbances, a_ts, would be accounted for more than in the case shown in Fig. 1. So, while no difference in productivity is ultimately observed for experiment A, that does not mean that there are no underlying differences in the two methods observed. We cannot tell what those differences are because we lumped all the disturbances together, but they may not all necessarily be good. For example, it is possible that work quality or safety is affected, but it just does not show up in productivity measurements.

Green's Function Plots for Experiment B. Figs. 8 and 9 show plots of Green's function for the experiment B control and CRANIUM cases, respectively. From these figures, it can be seen that a disturbance a_t to the lifting duration in both the control case and the CRANIUM case is forgotten, or its effect subsides both in magnitude and duration after approximately the tenth lift. In addition, in both cases, the structure of the Green's function is the same, resulting basically from a damped cosine wave. Also, the magnitude of the effect of a disturbance a_t is almost the same in both cases.

These observations indicate that the introduction of the CRANIUM had very little effect on the dynamics governing the disturbances to the crane operations. Based on the results shown, it seems that the introduction of the CRANIUM to settings similar to experiment B would result in improved productivity and decreased cost from eliminating the tagman without significant underlying changes to the overall process.

Green's Function Plots for Experiment C. Figs. 10 and 11 show plots of Green's function corresponding to the experiment C control and CRANIUM cases, respectively. From Fig. 10, it can be seen that a disturbance a_t to the lifting duration in the control case is forgotten, or its effect subsides both in magnitude and duration after approximately the seventh lift. From Fig. 11, it can be seen that a disturbance a_t to the lifting duration in the CRANIUM case is forgotten, or its effect subsides both in magnitude and duration after approximately the second lift. Moreover, in the control case, the structure of Green's function is a sum of two exponentially decaying parts, which is quite different from that in the CRANIUM case. Also, the magnitude of the effect of a disturbance a_t is higher in the control case than in the CRANIUM case.

These differences may arise from several possibilities. For example, since experiment C involved the most precise lifts among the three experiments performed, it seems that in the control case, the smooth curve showing the effect of λ_1 on Green's function is reflecting a learning curve effect, in comparison to the relatively uneven effect of λ_2, which might reflect disturbances coming from wind, motor control problems, and likewise.

Surprisingly, this was the exact pattern found when the CRANIUM was introduced in the experimental setting A (see Fig. 7). This implies that the difficulty of the experimental setting in experiment C has caused similar dynamics to those when a change was introduced to the crane operations in experiment A by the CRANIUM. In other words, it seems that

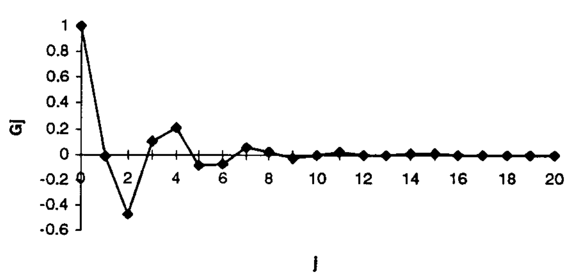

FIG. 8. Green's Function of ARMA(2, 1) Model (Experiment B—Control)

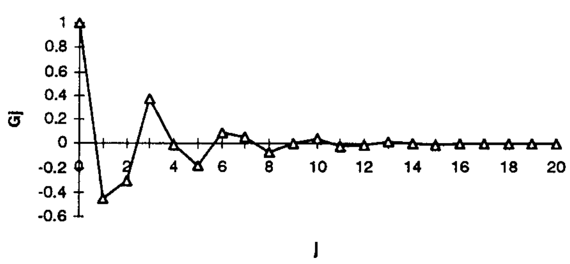

FIG. 9. Green's Function of ARMA(2, 1) Model (Experiment B—CRANIUM)

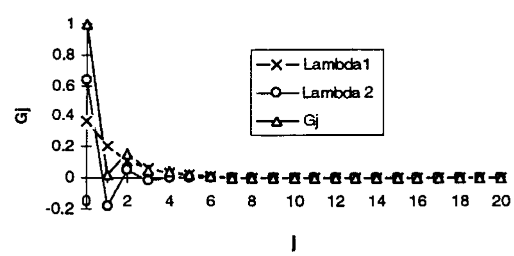

FIG. 10. Green's Function of ARMA(2, 1) Model As Sum of Two Exponentials (Experiment C—Control)

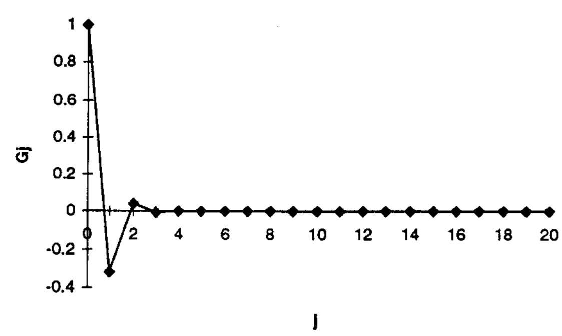

FIG. 11. Green's Function of ARMA(1, 1) Model (Experiment C—CRANIUM)

J. Constr. Eng. Manage., 1999, 125(2): 87-95

the effect of a difficult procedure would be the same regardless of the experimental setting.

By reexamining Table 3, it is evident that the value of λ_1 is almost the same for the experimental settings considered; thus, the same dynamic mode has reappeared in response to a difficult task. Despite that, the contribution of the dynamic mode λ_1 to the variance of the observed output in experiment A (CRANIUM case) was as shown in Table 4—only 2.88 versus 94.68 for experiment C (control case). Therefore, the difference is quite large between the difficulty of experiment C with respect to that of experiment A. This information is an excellent example of how time series analysis could provide better quantification for the difficulty of one experiment versus another, rather than settling for a qualitative comparison.

Based on the results shown, the introduction of the CRANIUM in settings similar to those of experiment C seems to greatly minimize the effect of disturbances on the crane operations, and, at the same time, improve productivity and reduce cost.

CONCLUSIONS

The present paper has demonstrated the value of using time series analysis for evaluating construction productivity experiments. Important information can be derived from experimental data, well beyond the comparison of means of sets of sample data. The results of the crane productivity experiments are excellent examples of the types of information that can be derived, despite the fact that a relatively small number of observations were made. In experiment A, we saw that, despite no apparent difference in productivity, there were underlying differences in the overall operation. In experiment B, we saw the reverse situation—a large difference in productivity, but little change in the overall operation. In experiment C, we saw both a change in productivity and considerable differences in the underlying operation—in particular, a distinct learning effect.

Of course, productivity differences are important when comparing a new technology to the standard method(s), but there may also be other changes to the overall operation that are not apparent when looking strictly at productivity data. It may not be possible, at first, to tell exactly what those underlying differences are, but a signal is given that additional investigation may be warranted. This may involve taking additional measurements of the disturbances to the system to see if these are the cause of the changes, or if there are other unrecognized inputs or disturbances that should be considered.

In addition, the techniques described in the present paper may be used as an evaluation tool for training programs. In the case of cranes, for example, an experienced crane operator would be less affected by disturbances than would the novice or apprentice operator. By plotting the Green's function for the experienced crane operator versus that for the novice, the difference in skill would be captured between the two. Moreover, the plots would reveal how well the novice has been trained compared to the experienced operator.

APPENDIX. REFERENCES

Anderson, T. W. (1971). *The statistical analysis of time series*. Wiley, New York.

Blum, O. M. (1965). *Theory of random functions and its applications to automatic control problems* (English translation). Pergamon, Tarrytown, N.Y.

Box, G. E. P., and Jenkins, G. M. (1970). *Time series analysis forecasting and control.* Holden-Day, San Francisco.

Everett, J. G., and Slocum, A. H. (1993). ''CRANIUM: Device for improving safety and productivity.'' *J. Constr. Engrg. and Mgmt.*, ASCE, 119(1), 23–39.

Pandit, S. M. (1973). ''Data dependent systems: Modeling analysis and optimal control via time series,'' PhD thesis, Univ. of Wisconsin—Madison.

Pandit, S. M., and Wu, S. M. (1993). *Time series and systems analysis with applications.* Krieger, Malabar, Fla.

Pugachev, V. S. (1957). *Theory of random functions and its applications to automatic control problems.* Pergamon, Tarrytown, N.Y. (in Russian).

Wiener, N. (1949). *Extrapolation, interpolation and smoothing of stationary time series.* Wiley, New York.

J. Constr. Eng. Manage., 1999, 125(2): 87-95