# USING NOISY RFID FOR ACCURATELY MONITORING THE ASSEMBLY LINE OF PANEL FABRICATION

**Y. Wen, D. Ogunyemi, A. S. Bouferguene, M. S. Altaf, A. Bouferguene, M. Al-Hussein**

Civil & Environmental Engineering, University of Alberta, Edmonton, Canada T6G 2G7
e-mails: ywen3@ualberta.ca, ogunyemi@ualberta.ca, boufergu@ualberta.ca, msaltaf@ualberta.ca, ahmedb@ualberta.ca, malhussein@ ualberta.ca

**Abstract**. *Over the past few years, the interest towards off-site construction as part of project delivery for residential and commercial buildings has increased dramatically. In this regard, buildings are decomposed into panels that are manufactured using assembly lines smilar to what was developped to the automotive industry. However, because these serial production systems do not have buffers that could store intermediate products when a bottleneck occurs downstream, it often happens that these products are required to wait at the current workstation before production can resume. In this contribution, we develop algorithms allowing the waiting times to be extracted from timestamps that are collected from an RFID system reading the tags on panels as they enter each workstation.*

**Keywords**: Panel manufacturing, Assembly line, Data clustering, Model mining.

## 1 INTRODUCTION

Wood based construction industry is increasingly embracing the concept of off-site manufacturing as part of its project delivery paradigm. In this context, projects are first decomposed into a set of components that are produced in a controlled indoor environment before being shipped to the construction site for assembly. This approach which is at the core of modular construction has many technical and social benefits including an improved productivity, a better quality of the end product and more importantly a working environment where accidental and occupational injuries are reduced [1]. In practice, wood-based off-site construction for residential and commercial projects is generally implemented using one of two routes: (i) modularization or (ii) panelization. In the first case, buildings are broken down into a set of functional elements, usually cuboid shaped known as modules, whereas in the second the very same buildings would be represented in a more primitive manner using panels as the fundamental objects. However, whether modularized or panelized, the elements of the end product are nowadays manufactured indoor using a production system relying on assembly lines. The immediate advantage for construction practitioners is the ability to borrow and build upon the wealth of research that was developed for the automotive industry in order to reap rapidly the benefits of off-site construction. In this respect, it is worth mentioning that assembly lines is an inherent part of lean thinking, as developed and implemented by Toyota [2], since not only do they limit workers' motion, thus waste, but are also able to unveil any issue that can disrupt the continuity of the production flow.

The research contained in this paper focuses on panelized construction. In practice, an assembly line for panel fabrication consists of a sequence of workstations that add specific components to the panel as it moves from the first station towards the last at which point it exits as a finished product and ready for shipping to the construction site. Whether the fabrication of panels is performed manually or automated using specialized machinery, the efficiency of this process requires streamlining the activities and leveling the resources in order to reduce bottlenecks. The latter is particularly important for managers since line balancing (or equivalently steady flow) is a requisite for optimal productivity. In this respect, in

order to gain better insight into the cycle time of the assembly operation as well as improve resource leveling, time studies are often required in order to accurately quantify the duration of each activity carried out at the workstations. For this purpose, data collection, i.e. activity durations, is typically collected in one of two ways: (i) by an observer operating in a live setup, in which case the experiment needs to be minutely planned in order to avoid interferences with workers or (ii) using digital devices, e.g. sensors reading RFID tags, monitoring cameras, etc [3]. While the second approach is probably the most convenient since data collection proceeds continuously without interfering with the workers, it may, however, require using special algorithms and functions in order to identify in the stream of digital data the time windows corresponding to productive, i.e. steady flow, and unproductive activities, e.g. line stoppage, bottlenecks, etc. In this respect, this contribution explores algorithms geared towards extracting durations of productive activities in an assembly line for panelized construction from data automatically collected using RFID tags and readers. These algorithms which are variations of the Random Sample Consensus (RanSaC) are analyzed from a statistical perspective in order to get better insight into their robustness [4]. If proven sufficiently accurate, these algorithms combined with RFID technology would allow managers to access almost-real time information that could be used to build simulation models, to better plan resource leveling or explore the impact of changes, e.g. floor layout, rotation of personnel, new machinery, new material, etc. on the workstation productivity [5].

## 2 INTRODUCTION

The increasingly popular panelized construction paradigm which is gradually implemented in the Canadian wood-based construction industry has drastically improved the efficiency of residential and commercial buildings projects' delivery. At this time, the construction industry can be clustered into two types of builders according to their financial abilities to acquire specialized machinery. In this respect some builders have adopted a full automation setup whereas others that could be labelled as semi-automated rely on hand held machines for the fabrication of panels. At this juncture, it is important to note that even though this research was developed using a fully automated factory, its methodological aspects can be applied to a semi-automated production insofar as an assembly line is used for production process. Before addressing the setup that is used to collect the duration of activities at each workstation, let us point out that for the purpose of maximizing the utilization of the framing workstation which uses a 12.19m (40 feet) long table, individual walls, e.g. exterior and interior, are first merged into panels of similar length, see Figure 1.

Figure 1. Walls (possibly from different projects) assembled to form a 40-feet multi-panel.

Since multi-panels are made up of multiple walls assembled to be as close as possible from the full length of the framing table, i.e. 40 feet, these walls will have different structural components according to whether they contain openings for windows and doors and/or if they are load bearing, etc. Furthermore, since projects are based on clients' orders, multi-panels will be made up from different walls and it is

unlikely to encounter frequently the exact same multi-panel. As a result, the processing time of a given multi-panel (especially at the framing stage) can vary substantially from a multi-panel to the next. Furthermore, because the configuration of multi-panels is demand driven, it is practically difficult to have a sufficiently large sample of exactly similar multi-panels from which statistics about processing times at each workstation can be obtained. In this respect, a dataset containing in excess of 7000 records corresponding to 4365 different multi-panels only 40 of these panels were encountered at least 10 times, see figure 2.

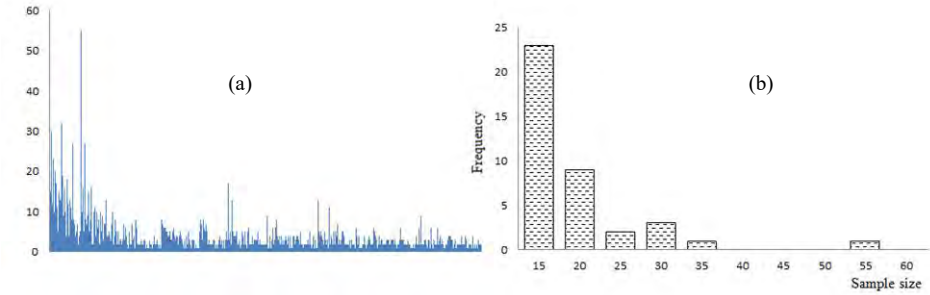


Figure 2. (a) Overall frequency of identical multi-panels; (b) Distribution of sample sizes for multi-panels encountered at least 10 times over a period of 2 years (between July 16th, 2015 and May 17th, 2017)

According to Figure 2 which depicts selected distributions of the data over a period of 2 consecutive years, it is clear that encountering the exact same panel is not very likely. In fact, a quantitative analysis of the corresponding dataset revealed that multi-panels that were encountered at least 10 times during this period represented less than 1% of the 4365 different panels that were fabricated at the partner's plant. Furthermore, even in the case of similar multi-panels, the processing time can vary substantially since any stoppage at a downstream station of the bufferless assembly line introduces an additional time, referred to as the waiting time, that is captured as part of the duration spent by the multi-panel at the current workstation. In fact, since the assembly line does not have buffers to solve the issue of downstream bottlenecks, RFID time stamps are read when a multi-panel enters a station, see Figure 3.

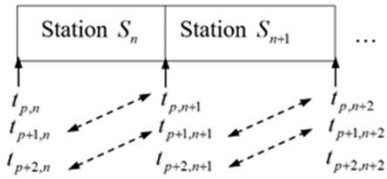


Figure 3. RFID readers record the time stamps as panels $p, (p+1),...$ enter stations $n, (n+1),...$

Based on Figure 3, the actual duration spent by panel p at workstation n is calculated defined as,

$$\Delta t_{p,n} = t_{p,n+1} - t_{p,n} = T^{P}_{p,n} + T^{W}_{p,n} \quad (1)$$

Where,

- $\Delta t_{p,n}$ is the apparent processing time (since this may be inflated by waiting time)
- $t_{p,n}$ and $t_{p,n+1}$ are the timestamps recorded for panel *p* as it enters workstations *n* and *(n+1)*;

- $T^{P}_{p,n}$ is a random variable representing the actual productive time, i.e. the time spent by workers adding value to panel *p*, and $T^{W}_{p,n}$ is a random variable describing the waiting time that panel *p* is halted at station *n*.

From equation (1), it is clear that approximating the productive time by the difference in time stamps is only valid when the assembly line is in a steady state, i.e. $T^{W}_{p,n} = 0$. Unfortunately, a depiction of the longitudinal data $\Delta t_{p,n}$ for four different multipanels exhibit large variability which is that waiting is not likely to be a Poisson process, see Fig. 4.

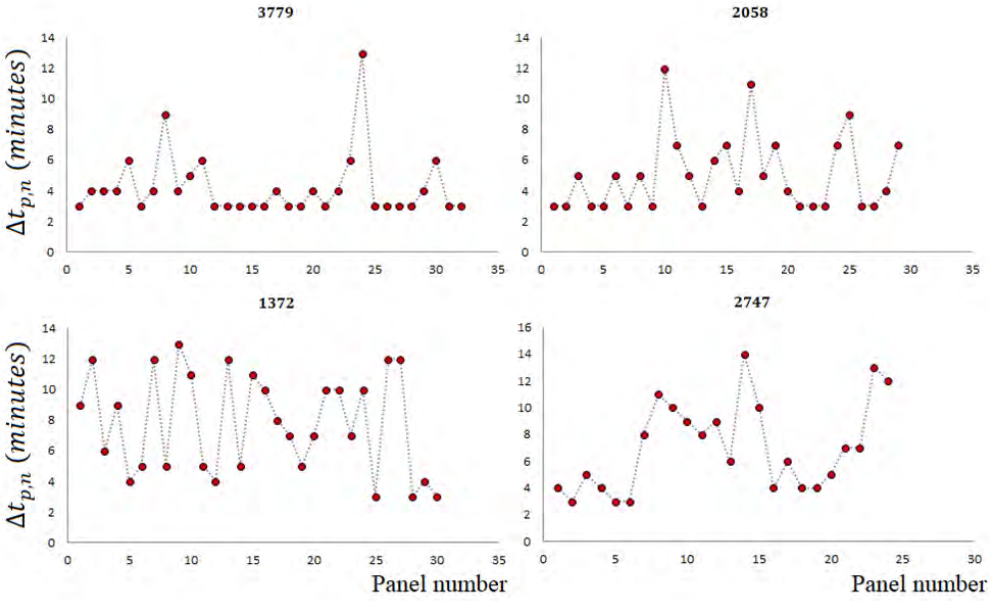


Figure 4: Framing times for 4 types of multi-panels that were encountered close to 30 times over a period of 2 years

In fact, according to Figure 4, it is clear that when production is in a steady state, processing time for a specific multi-panel is expected to be close to constant. This behavior can be seen in Figure (4) where panels 3779 and 2058 are most often processed approximately within 3 to 4 minutes whereas the framing durations for panels 1372 and 2747 show large variability resulting from the frequent occurence of waiting events. As a result, for the purpose of planning and understanding the efficiency of the production line, it is imperative to develop a computational procedure that could unmix the apparent processing time into its fundamental components, namely the actual (productive) and the waiting time.

## 3 UNMIXING THE APPARENT PROCESSING TIME

Thus far, the concept of similarity was used to identify multipanels whose processing times that could constitute datasets for statistical analysis, see Figure 4. For this purpose, similarity is evaluated using factors which differ from a station to another. For instance, the overall processing time at the framing

station can be expressed as the sum of durations associated with 10 sub-activities according to equation (2),

$$\hat{T}_{\text{Framing}} = \sum_{i=1}^{10} n_i \times \hat{T}_i \Rightarrow \left[\hat{T}_{\text{Framing}}\right] = \left[n_1\ n_2 \ldots n_{10}\right] \bullet \left[\hat{T}_1\ \hat{T}_2 \ldots \hat{T}_{10}\right]^T \quad (2)$$

Where, $n_i$ and $\hat{T}_i$ represent the number of items in sub-activity $i$ and its expected unit time respectively. As for the sub-activities in equation (2) identified by the running index $i$, they describe the type of task to be carried out on a given multi-panel, see Fig.1, and are coded as follows,

- $i = 1$ for cut zones, i.e. cuts needed to separate individual walls making up the multi-panel
- $i = 2$ for dilled holes on the multi-panel
- $i = 3$ for simple studs of the multipanels(see Fig. 5)
- $i = 4$ for L-shaped connection (see Fig. 5)
- $i = 5$ for multiple studs (see Fig. 5)
- $i = 6$ for double studs (see Fig. 5)
- $i = 7$ for window opening(s)
- $i = 8$ for doors in the multi-panel
- $i = 9$ for garage door(s)
- $i = 10$ for large door(s)

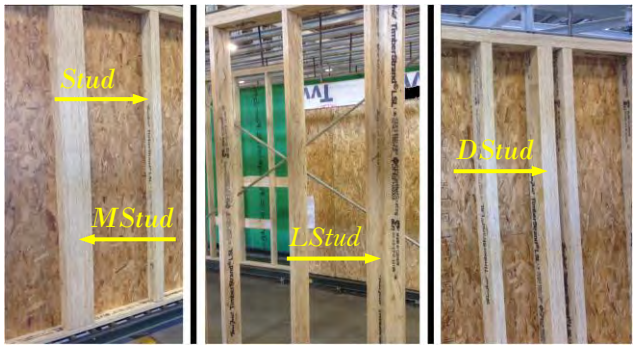


Figure 5. Types of studs encountered in a multi wall panel, referred above as a multi-panel

Clearly, if the expected (sub-activity) unit times occuring in the time vector of equation (2) were known, the average framing time for any given panel can be calculated straightforwardly. However, these unit times are not directly accessible since the overall frame times are often corrupted by waiting times as indicated in equation (1). As a result, in this work, the framing times are filtered in such a way as to estimate the productive time by means of Kalman filter. These times are then compared with values collected by an observer in the context of a time study as per the factors in equation (2). In essence, the Kalman filter tries to extract the productive processing times $T_{k,n}^{\text{Productive}}$ by correlating the (noisy) predictions of the signal with the (noisy) measurements in such a way as to minimize the error (or equivalently, the root mean square). From a computer implementation perspective, Kalman's approach is built upon two sets of equations (i.e., prediction and correction) satisfying Equation 3,

$$\begin{cases} T_{k,n}^{-} = & T_{k-1,n}^{+} \\ P_{k,n}^{-} = & P_{k-1,n} + Q \end{cases} \qquad \begin{cases} K_{k,n} = & P_{k,n}^{-}\left(P_{k,n}^{-} + R\right)^{-1} \\ T_{k,n}^{+} = & T_{k,n}^{-} + K_{k}\left(z_{k,n} - T_{k,n}^{-}\right) \\ P_{k,n} = & \left(1 - K_{k,n}\right) P_{k,n}^{-} \end{cases} \tag{3}$$

where $T_{k,n}^{-}$ is an estimate of $T_{k,n}^{+}$ which represents the filtered (or corrected) evaluation of $T_{k,n}^{\mathrm{Pr}oductive}$. Similarly, $P_{k,n}^{-}$ is the estimation of the corrected value $P_{k,n}$, which is used to define the Kalman gain $K_{k,n}$ chosen in such a way as to minimize the error between the measurements $z_{k,n}$ and the predictions. As for $Q$ and $R$, they represent the variances of the Gaussian noises contaminating the predictions and the measurements. These values are often selected empirically and adjusted to fine tune the filter. For instance, in the case where $R = (0.1)^2$ and $Q = 0.0$, multi-panel data is filtered as shown in Figure 6,

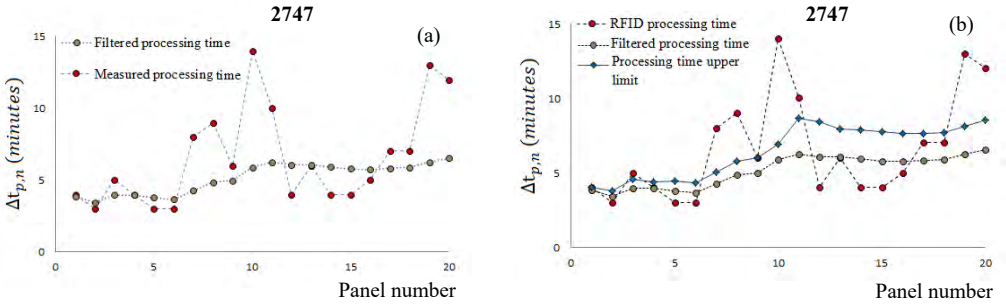


Figure 6. (a): Kalman filtered (noisy) RFID-based processing time. (b): Upper limit of processing time estimated based on the *ansatz* given in equation (4)

According to Figure 6, after 20 iterations, the corrected (true) processing time obtained from the Kalman filtered data is approximately 6.5 minutes. Note that, even though a few of the data points in Figure 6 may appear as (extreme) outliers, the presence of which is clearly indicative (lengthy) waiting times, in the long run they however did not notably deteriorate the convergence of the Kalman filter. At this juncture it is important to note that relying on a single value (6.5 minutes in the present case) to describe the productive processing time would implicitly ignore the natural variability that is inherent to manual work. As a result, it is necessary to analyze the differences $\left(T_{p,n}^{RFID} - T_{n}^{Filtered}\right)$ in order to avoid over estimating the waiting time. In the context of this work, an empirical upper limit of the productive processing is defined to satisfy Equation 4,

$$T_{p,n}^{\mathrm{Pr}oductive} \le T_{n}^{Filtered} + \min\left(Q_{\pi}, \bar{e} + t_{\alpha/2,N-1} \times \mathrm{SE}\right) \tag{4}$$

where $Q_{\pi}$ is the $\pi^{th}$ percentile of the dataset $\left\{T_{p,n}^{RFID} - T_{n}^{Filtered}\right\}_{p=1,2,\mathbb{L}}$; $\bar{e}$ is its average and SE the corresponding standard error. As for $t_{\alpha/2,N-1}$, it represents the t-distribution score associated with a two-sided $100\left(1-\alpha\right)\%$ confidence interval. Although in the current data collection setup it is practically impossible to obtain the values of $\pi$ and $\alpha$ from first principles, the values $\pi = 0.80$ and $\alpha = 0.10$ are found to lead to acceptable limits for the productive processing times. In the case of the data depicted in Figure (6a), the *ansatz* proposed in equation (4) allows the overall framing time multi-panels similar to that identified as 2747 to be bound by the envelop shown in Figure (6b). At this juncture,

it is worth mentioning that the upper bound of the processing time as defined by equation (4) is updated with every additional datum that is collected. This is particularly interesting to explore the impact of changes on the efficiency of the production line.

## 4 CASE STUDY

The methodology presented above was used for a production line of an industrial partner, ACQBUILT, who has implemented the panelized off-site construction paradigm as a means to control quality, cost and most importantly the production schedule. Although Kalman filtered data at this time is based on an empirical formulation of the upper bound of the framing time, it was found useful in estimating the cycle time for a given production plan. In addition to this, with these estimates, utilization of work stations can be analyzed. The RFID system at ACQBUILT's wall production line consists of five Motorola FX9500 RFID readers, 12 Motorola AN440 high performance dual antennas, and one Zebra ZD500R RFID label printer. There are five antennas installed at the entrance to the framing table, buffer table, sheathing table, multi-function bridge, and tilt table. These five antennas are connected to one reader (R1) located near the sheathing table. The RFID reader stores the timestamp data into the database in real time. The RFID database has four main data tables: (1) the PrintLog data table is populated at the time of RFID tag printing and stores the tag ID and associated panel name; (2) the TagDetail data table contains the initial and last timestamp data for each tag ID for each antenna; (3) the ReaderAntenna is a static data table that holds the relationship between reader and antenna and the relationship between location antenna and main antenna; and (4) the LocationTag data table contains the permanent RFID location tag information for identifying the line number. Structured query language (SQL) is used to extract the necessary panel information from the database.

The production data generated from the RFID system is used to calculate panel processing time, waiting time and workstation utilization. The RFID data is filtered using the Kalman filtering process in order to separate the noise component from the actual data. In order to apply Kalman filtering, the processing time data is clustered into different groups based on panel attributes. However, there are many different factors that can contribute to the processing time of a panel. To keep the filtering process simple, wall panels are divided into two main groups—exterior walls and interior walls. Each group is then divided into sub-groups based on their lengths. The filtered processing time is calculated based on Equation (3). Then an empirical upper limit for productive processing time is calculated based on Equation (4). Table 1 shows the Kalman filter and upper limit, of different multi-panel sub-groups for the framing station. This information is used to estimate productive processing time. If actual processing time is greater than the upper limit, the difference is regarded as waiting time/delay; otherwise. Table 2 summarizes the utilization calculation without the application of Kalman filter and utilization with Kalman filter, at the framing station for different days. The result shows that the utilization of a workstation can be over calculated if real processing time is used without applying any filtering procedure.

Table 1: Filtered value and upper limit of different sub-groups of multi-panels for framing station

| Subgroup | Kalman Filter (minutes) | $\min(Q_\pi, t_{\alpha/2,n-1} \times SE)$ | Upper Limit (minutes) |
|---|---|---|---|
| EXT >35 | 12.50 | 0.50 | 13.01 |
| EXT 30-35 | 10.59 | 0.66 | 11.25 |
| EXT 20-30 | 9.57 | 0.71 | 10.29 |
| EXT <20 | 7.84 | 0.68 | 8.52 |
| INT >35 | 10.93 | 0.27 | 11.20 |
| INT 30-35 | 10.55 | 0.64 | 11.19 |
| INT 20-30 | 8.08 | 0.61 | 8.69 |
| INT <20 | 5.85 | 0.38 | 6.23 |

Table 2: Framing station productivity detail

| Obs. | Utilization without filtering | Utilization with Kalman filtering |
|---|---|---|
| 1 | 78% | 68% |
| 2 | 77% | 71% |
| 3 | 79% | 71% |
| 4 | 81% | 71% |
| 5 | 78% | 74% |
| 6 | 68% | 53% |

Table 1 shows that the productive processing time at the framing station increases as the multipanel length increases, which validates the Kalman filtering process. However, one of the limitations of Kalman filtering is that, if the waiting time occurs frequently, it can overestimate the filtered value. The upper limit value for each type of panel can be adjusted based on simulation data in order to obtain better production data from the RFID system. As the proposed setup of the RFID system cannot distinguish the waiting time and productive processing time, the Kalman filter can provide an approximation of the processing time that can be used to calculate the workstation utilization.

## 5 CONCLUSION

At present, panelized construction is still not fully automated and many activities require manual intervention. As a result, even though the industry is slowly moving towards production systems involving assembly lines, human intervention at several stages of panel fabrication makes it extremely difficult to have an assembly line in a steady flow mode. Because waiting occurs frequently, capturing the time that is used to add value to a panel moving along the line is challenging. In this work, we have explored the potential of using a filtering technique as a means to unmix the series of total times (seen as a corrupted signal) collected at the framing station into two components: waiting and productive times. It turns out that when waiting times are not excessively long, filtering can lead to reasonnable productive times. However, frequent long waiting periods introduce bias into the filetered output which translates into overestimations of the true processing times.

## REFERENCES

[1] National Association of Home Builders. *Fast facts for panelized homes*. Retrieved October 24, 2013, from http://www.nahb.org/generic.aspx?genericContentID=10310.2009.

[2] Liker, J. K. ―The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer.” McGraw-Hill Education: New York, NY, USA. 2004.

[3] Domdouzis, K., Kumar, B., and Anumba, C. ―Radio-frequency identification (RFID) applications: A brief introduction.” Advanced Engineering Informatics, 21(4), 350–355. 2007.

[4] Raguram, Rahul, Frahm, Jan-Michael & Pollefeys, Marc. ―A Comparative Analysis of RANSAC Techniques Leading to Adaptive Real-Time Random Sample Consensus.” Springer Berlin Heidelberg. pages 500-513, ISBN 978-3-540-88688-4. 2008.

[5] Altaf, M., Liu, H., Zhang, Y., Al-Hussein, M., and Bouferguene, A. ―Discrete-event simulation modelling of prefabricated wall production line.” *Proceedings, 27th European Modeling & Simulation Symposium, Bergeggi, Italy*, Sep. 21-23, pp. 234-239. 2015.

[6] Senghore, O., Hastak, M., Abdelhamid, T. S., Abuhammad, A., & Syal, M. G. ―Production Process for Manufactured Housing.” *Journal of Construction Engineering and Management*, *130*(5), 708–718. 2004.