A Data Processing and Control System to Support Remote Infrastructure Monitoring

Principal Investigator: Professor Pablo L. Durango-Cohen

A final report submitted to the Infrastructure Technology Institute

A Data Processing and Control System to Support Remote Infrastructure Monitoring

Yikai Chen

David J. Corr

Pablo L. Durango-Cohen

Department of Civil & Environmental Engineering and Northwestern University Transportation Center, 2145 Sheridan Road, A332, Evanston, IL, 60208-3109, Phone: 847-491-4008, Fax: 847-491-4011, Email: pdc@northwestern.edu .

1. Background and Objective

The Hurley Bridge (Wisconsin Structure B-26-7) carries westbound traffic on US Route 2 over the Montreal River from Ironwood, Michigan to Hurley, Wisconsin. The bridge is a three-span continuous five-girder steel structure with a composite deck. In cooperation with WisDOT, the Research Engineering Group (REG) of ITI has installed a continuous remote structure health monitoring system on the bridge consisting of strain gauges, thermocouples, accelerometers and displacement transducers at selected locations, in conjunction with a weigh-in-motion system installed by a third-party contractor.

The bridge is subject to heavy loads from daily truck traffic, many of which are believed to be overweight logging trucks travelling from Michigan into Wisconsin. WisDOT is concerned about that the observed traffic will cause premature degradation of the structure due to fatigue and overstress conditions. Therefore, the main objectives of the system include:

- Process measurements related to structural health, traffic loads, and environmental conditions in an integrated fashion, thereby yielding comprehensive, facility-level condition assessment and forecasting capabilities;
- Provide real-time, reliable alerts when potential damage or risk of structural change in the facilities is detected; and
- Determine the nature of the detected changes, i.e., infer underlying structural properties, and identify possible assignable causes.

In line with the USDOT's research goals, this work involves development and implementation of cutting-edge, transformative research tools to support information management, and decisions related to the management/renewal of surface transportation infrastructure to ensure that it operates in a "state of good repair". In addition, this work is also complementary to the expertise of Northwestern University's Infrastructure Technology Institute (ITI) in developing advanced remote monitoring systems.

2. Scope of Work

This research complements earlier work funded by ITI. We outlined its major parts below:

2.1 Data collection and pre-processing

The SHM system includes 2 displacement transducers and 13 strain gauges at critical locations on the bridge (shown in Figure 1) to measure real-time structural conditions. It also involves a cabinet that collects the environmental conditions including ambient temperature and humidity, along with a weigh-in-motion system that records the traffic loading information. Data collection initiated on April 1, 2010 and is projected through August 2013 at a frequency of 100Hz.

In this report, we include results from the analysis of the longitudinal and transverse displacements at the east end of the bridge (denoted Displ-L & Displ-T) as an example to validate the effectiveness of the data processing and control system. The analysis period spans from April 1st, 2010 through June 30th, 2011. Hourly averages of measurements are transmitted by a computer on site, which we use to further group into daily averages over the analysis period.

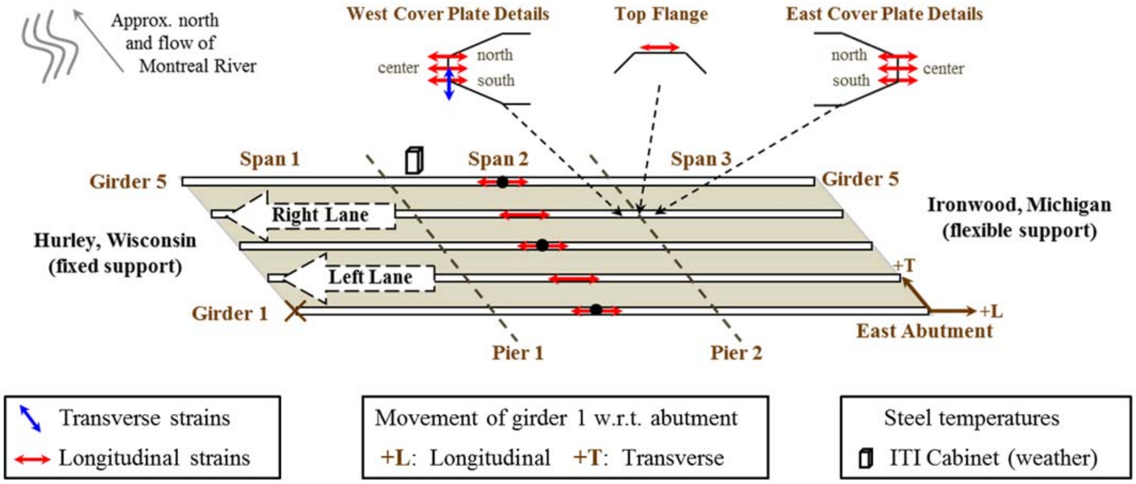

Figure 1: Structural Health Monitoring System on the Hurley Bridge

2.2 Statistical performance modeling

In this part we implement regression analysis to formulate and estimate performance conditions of the facility components on the bridge. It decomposes the dynamic bridge performance into long-term trends, seasonal patterns and a sequence of residuals. The regression process generates reliable estimates of the true structural conditions as well as how they respond to environmental factors. Such estimates further lend us to the capability of predicting future conditions of the bridge components.

2.3 Statistical process monitoring

The work in this part is built on the notion of Statistical Process Control (Shewhart 1931, Deming 1975) originally used in manufacturing processes. In particular, we adapt techniques described in Alwan and Roberts (1990) and expand it to involve the following technical areas:

1) **Control for serial dependence and heteroscedasticity**. Residuals from the performance models are influenced by carry-over effects and heteroscedasticity (i.e. changing variance) that are unavoidable in field measurements. We formulate time series models to capture and control these effects in the residual sequences. The results will be used to refine prediction models as well as to generate a sequence of innovations for detecting unusual changes below.
2) **Detect transitory anomalies**. We construct Shewhart control charts to detect transitory anomalies in the measurement sequence with pre-determined confidence levels.
3) **Detect persisting shifts**. We use supplementary control charts with memory properties to detect subtle and persisting shifts in the measurement sequence.
4) **Monitor unusual changes in variability**.
5) **Coordinate multiple data sequences**. We employ multivariate control charts to monitor multiple measurements simultaneously, taking into account their inter-correlations, so as to reduce the overall data complexity and ensure more reliable results.
6) **Interpret detected changes**. We identify plausible causes for the detected outliers and shifts/drifts in the measurements on the Hurley Bridge.

3. Content of Research and Result

The thrust of this research is to develop a rigorous and generally applicable statistical framework to process data from the long-term remote monitoring of the Hurley Bridge as well as similar transportation infrastructures. According to Deming (1975), changes in the output of any process consist of three components:

- Common-cause variation, which can be explained and predicted by usual, and typically recurrent, operational characteristics;
- Special-cause variation attributed to unusual or extraordinary events with assignable causes;
- Random variation, which is the cumulative effect of many small, inherent and essentially unavoidable causes in the process (Montgomery, 2009). Such variation can be statistically described as a series of random variables that are independent and identically distributed.

In the case of the Hurley Bridge, operational characteristics include those associated with usual traffic and weather conditions. Extraordinary events could refer to either exogenous and even observable events, such as extreme traffic or weather, or endogenous (and sometimes latent) changes, such as the advent of certain types of damage.

In our framework, we use statistical performance models to capture common-cause variation and control charts (univariate and multivariate methods respectively) to detect special-cause variations. Accordingly, the content of this research consists of three parts:

3.1 Estimation of statistical performance models

This part develops performance models to describe and predict common-cause variation in the deterioration of infrastructure facilities in the measurement sequences. Given the longitudinal nature of the data, the use of state-space specifications of time series models to capture dynamic effects of exogenous factors, e.g., weather, traffic, maintenance, seems promising. We estimate and validate such models, and compare them based on goodness-of-art and predictive capabilities.

We consider a two-step approach to model and analyze common-cause variation in the response data from the Hurley Bridge. The first step consists of using regression analysis to explain the effect of the exogenous variables, including linear trend terms as supplementary predictors, on the structural response measurements. The second step consists of estimating sophisticated time series models to simultaneously control for serial dependence, autocorrelation and conditional heteroscedasticity in the residual series.

The regression models can be used to decompose the measurement sequences into level, linear trend, seasonal components and residuals. Figure 2 provides a graphical illustration for Displ-T. The residuals, presented at the bottom of the figure, are used subsequently in the analysis of special-cause variation. We observe that during summer, the bridge moves North (positive value in measurement); whereas in the winter the bridge moves South. After controlling for the error, the net position is shifting to the North at a rate of 0.03 mils per day.

Serial-dependence, i.e., carryover effects from earlier measurements, influenced by earlier changes in the explanatory variables, can be a significant source of unexplained, common-cause variation. Modeling serial-dependence, therefore, leads to refinements in forecasting the progression of the measurements. Moreover, the analysis of special-cause variation is predicated

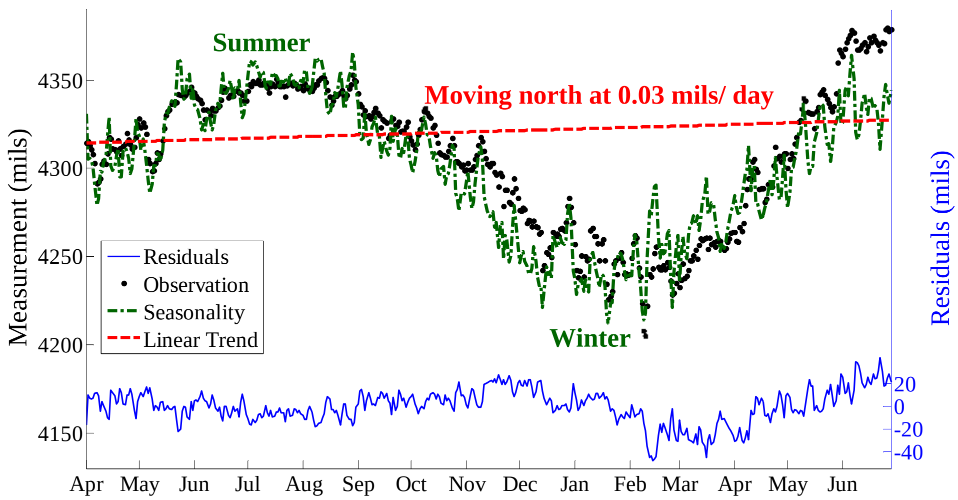

Figure 2: Structural Decomposition of the Transverse Displacement on the Hurley Bridge

on the assumption of a state of statistical control, where variation is random and, in turn, motivates the need to control for sources of systematic variation such as serial-dependence. We formulate and estimate an ARIMA-GARCH model to control for these effects, and to construct series of standardized innovations that satisfy the assumption of a state of statistical control.

3.2 Univariate analysis of special-cause variation

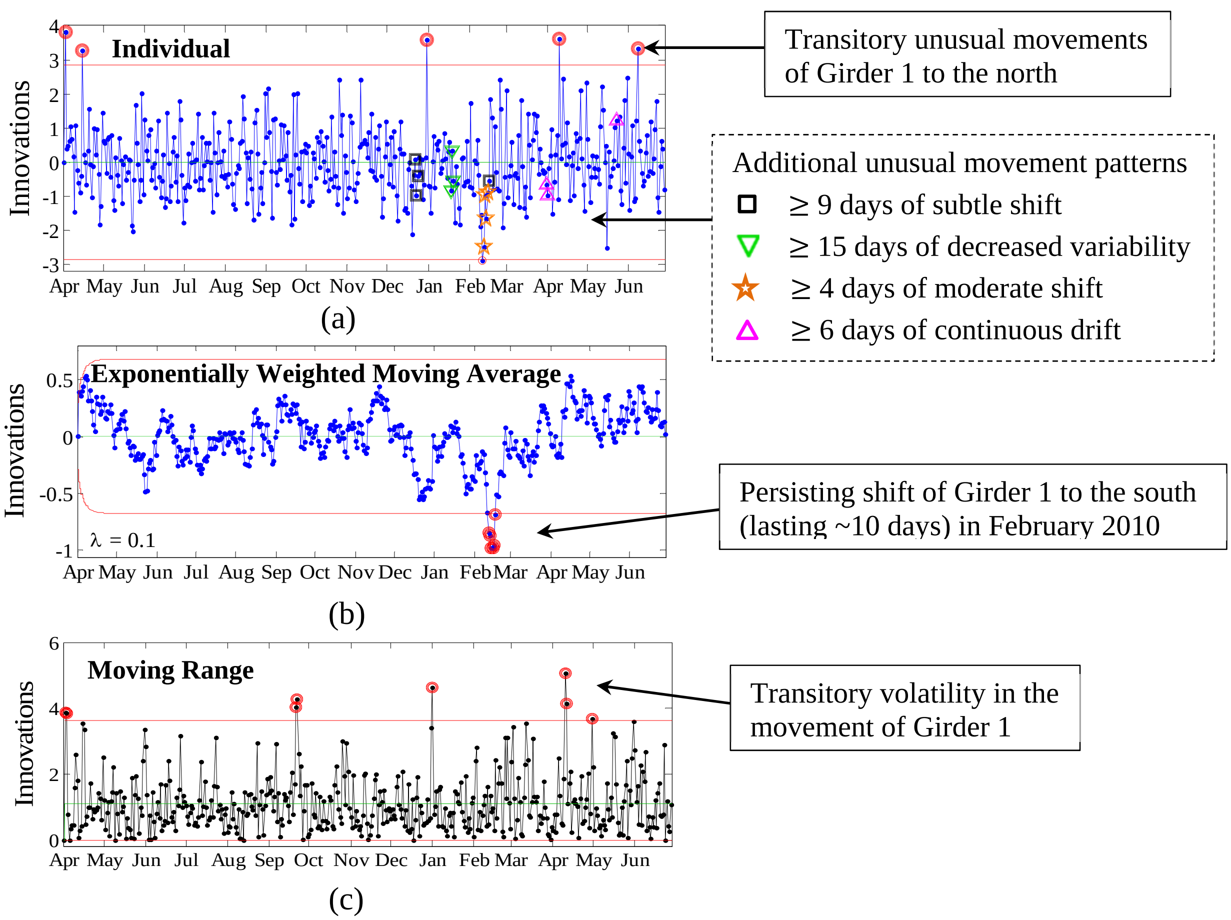

Figure 3: Outliers in Control Charts for the Transverse Displacement on the Hurley Bridge

Shewhart Control Charts serve as the main tool to detect special-cause variation in stationary random processes. These charts provide a representation of a time series, e.g., the standardized innovations, relative to the capabilities of the generative process: mean and variability. Shewhart Control Charts are effective in detecting large transitory changes in either mean or variance.

Figure 3(a) shows the *Individual Control Chart* for the transverse displacement. The chart includes 3 horizontal reference levels: a centerline set at the level of the mean of the given standardized innovation series; and upper and lower control limits set at ± 3 standard deviations from the centerline. Points outside of the control limits signal departures from a state of statistical control with a prescribed confidence. Additional triggering rules are also used to identify other patterns of unusual changes (Chen et al, 2012a).

Moreover, identification of persisting, possibly with small magnitude, shifts can be of great interest in field monitoring. Supplementary control charts with memory properties, such as the *Exponentially Weighted Moving Average (EWMA) Control Chart*, are deployed for this purpose. At each sampling period, the EWMA statistic is a weighted combination of past observation values and the current observation value:

$$EWMA_t = \lambda \cdot x_t + (1 - \lambda) \cdot EWMA_{t-1} \quad (1)$$

where x_t is the innovation value of the measurement at time t, and λ is the tuning parameter. This cumulative effect increases the chart's sensitivity to changes with small magnitude that are not detectable using the Individual Chart, such as the shift event shown in Figure 3(b).

Figure 3(c) shows the *Moving Range Control Chart* that monitors variability in the process, where the monitored statistic is the absolute difference between adjacent observations:

$$MR_t = |x_t - x_{t-1}| \quad (2)$$

The logic in setting the upper and lower control limits is analogous to the Individual Chart.

3.3 Multivariate analysis of special-cause variation

The advances in sensing technology have made simultaneous monitoring of multiple measurements possible. The key motivation of multivariate analysis is to reduce data complexity and enhance the overall reliability of results. In fact, the SHM system on the Hurley Bridge collects 15 response measurements. In addition to the computational and managerial burden, the use of separate univariate control charts makes it difficult to control the overall false alarm rate. From an engineering perspective, challenges in calibrating sensors, in addition to their high sensitivity and relatively low reliability, also motivate the need to create a multivariate framework that can exploit redundancies between the measurements. Recent deployments of dense networks of wireless sensors with dozens of instruments will only increase its importance. To address these needs, this part consists of the following technical steps:

3.3.1 Construction of multivariate control charts

In the first step, we use analogs to the Shewhart control charts to analyze special-cause variation in multidimensional sets of measurement sequences. Various statistics have been devised to monitor mean, variability and correlation in multivariate processes. The Hotelling T^2 Control Chart, which monitors the divergence of a multidimensional set of measurements from its mean, is a representative example:

$$T^2 = (X_t - \mu)' \cdot \Sigma^{-1} \cdot (X_t - \mu) \tag{3}$$

where X_t represents the innovations vector of multidimensional measurements collected at time t; μ and Σ denote the mean vector and covariance matrix. Since the T^2 statistic is a scalar variable and follows an F distribution, we construct a Shewhart-like control chart and set an upper control limit to trigger outliers.

Chen et al (2012b) use T^2 Control Charts to monitor 15 measurements on the Hurley Bridge. In this research, we collected one year of data starting from April 1st, 2010 and averaged daily observations into weekly subgroups for each measurement sequence. As an example, Figure 4 illustrates the T^2 control chart for the biaxial displacements of the bridge on the east end. The chart indicates an obvious outlier in the middle of February 2011. Meanwhile, we also observe a spike in mid-November 2010 which rises closely to the triggering limit.

To understand the results graphically, we present a two-dimensional data plane that incorporates a projection of both univariate and bivariate T^2 control charts (Figure 5). The horizontal and vertical axes correspond to Displ-L and Displ-T respectively. The rectangular region corresponds to the control limit range of the two separate univariate charts, and the ellipsoid corresponds to the range of the T^2 control limit.

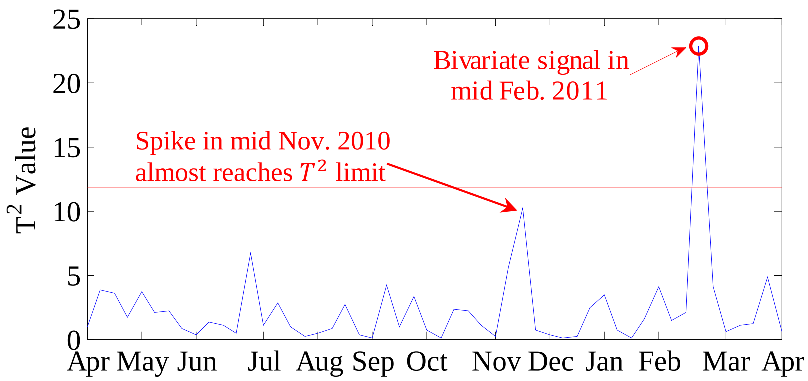

Figure 4: T^2 Control Chart of Biaxial Displacements on the Hurley Bridge

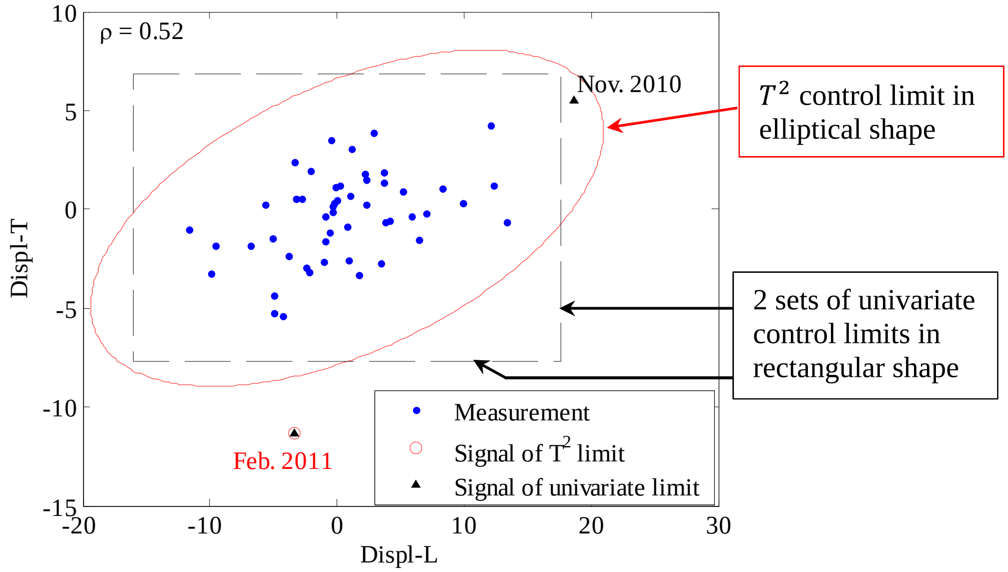

Figure 5: Univariate and T^2 Control Region for Biaxial Displacements on the Hurley Bridge

Figure 5 illustrates that the outlier in Feb. 2010 falls outside both the rectangle and the ellipsoid to the downer side, indicating that an unusual movement towards the south triggered the T^2 signal. On the other hand, the observation in Nov. 2010 falls barely outside the rectangle to the right while it stays within the ellipsoid. This explains why this sample does not trigger a signal in the T^2 chart. Importantly, the existence of such difference between the rectangular and elliptical control regions reflects the different utility of the two types of control charts, which is briefly summarized in the table below.

Table 1: A Comparison between T^2 and Univariate Control Chart

Control Chart	**Capability in detecting outliers**		**Data manage and illustration**	**Overall false alarm rate**
	Within single variable	**Cross-variables**		
$\boldsymbol{T^2}$	Subtle tolerance	Strict	One central chart	Controllable & rigorous
Univariate	Strict	None	Separate charts	Hard to compute

3.3.2 Numerical interpretation of T^2 outliers

A practical difficulty encountered with the T^2 control chart is the rigorous interpretation of outlier signals, i.e., to identify a subset of measurements that contribute significantly to the signals with certain confidence level. This is particularly critical in SHM of transportation infrastructure where the detected signals trigger costly inspections or service restrictions.

We adopt the Mason-Young-Tracy (MYT) method to justify reliable results on the Hurley Bridge. The MYT method is a numerical decomposition of the T^2 value into mutually independent components that respectively reflect the contribution of individual variables. Taking the biaxial displacements above as an example, the T^2 value at each sampling period can be decomposed as:

$$T^2 = T_L^2 + T_{T|L}^2 = T_T^2 + T_{L|T}^2 \tag{4}$$

where T_L^2 and T_T^2 are the component values of T^2 that are solely contributed by the longitudinal displacement (Displ-L) and transverse displacement (Displ-T) respectively. $T_{T|L}^2$ is the conditional contribution of Displ-T to T^2 when the value of Displ-L is fixed; vice versa for $T_{L|T}^2$. Statistically speaking, T_L^2 and $T_{T|L}^2$ are independent to each other and both follow an F distribution. Therefore, we can set critical value at desired confidence level to identify which variable(s) contribute significantly to the outlier.

Table 2 shows that the transverse displacement is the major cause for the T^2 outlier in mid-February of 2011, and this corroborates with the conclusion from graphical illustrations above. Compared with graphical approaches, the numerical scheme provides more reliable evidences as well as lends its capability to address high-dimension measurements when graphical illustration is not visually convenient.

Table 2: MYT Decomposition of the Outlier Sample in Biaxial Displacements on Hurley

Outlier	$\boldsymbol{T^2}$	**Marginal**		**Conditional**	
		T_L^2	T_T^2	$T_{T\|L}^2$	$T_{L\|T}^2$
Mid Feb. 2011	**22.77**	0.5	**19.29**	**22.27**	3.48
Critical Value [1]	11.84	8.97	-	-	-

[1] *At 99.7% confidence level*

4. Conclusion

We developed a control system to process the long-term remote monitoring data on the Hurley Bridge. The system demonstrates its capability to:

- Formulate statistical models to estimate and predict long-term performance conditions;
- Construct control charts to detect, characterize, and quantify the effect of unusual changes and trigger reliable alerts when potential risks occur;
- Identify plausible causes of the detected changes and provide valuable information for maintenance and repair.

In terms of performance conditions on the bridge, our analysis shows that:

1) Seasonal effects, i.e., temperature and humidity, and a linear trend, included as a supplementary predictor, account for a large percentage of the overall variation in the response measurements (40-90%);
2) Small, but significant, linear trends indicate permanent displacement of the bridge, i.e., after controlling for seasonal adjustments, the net position is drifting. This is likely an indication of deterioration of the bridge supports. Linear trends in the strain measurements could be influenced by changes in the materials, i.e., fatigue, sensors or traffic loads;
3) There is evidence that serial dependence, i.e., carryover effects, is a significant source of common-cause variation.

Further implementation of control charts detected 43 special-cause events over the experiment period of April 1st 2010 through June 30th, 2011. Details of these events can be found in Chen et al (2012a) and were reported to engineers at ITI. Events occurring on May 10th-20th, June 12th-20th, October 14th-20th of 2010, and January 24th-31st, February 8th-16th of 2011, are worthy of additional investigation due to their significance, duration and magnitude. In particular, we found that the unexplained changes in strain at multiple locations of the bridge over May 10th-20th of 2010 indicate high correlation with the maximum, seasonally-adjusted daily traffic load.

While unusual in occurrence, none of the detected changes represent an immediate threat to the safety or serviceability of the Hurley Bridge. Instead, they demonstrate the ability of the data processing system to efficiently and reliably detect out-of-the ordinary events as well as subtle long-term changes. Future work will enable similar data processing systems to provide alerts of structural performance changes in near-real time.

Throughout this project we disseminated the results of our work at academic conferences and research seminars. In addition, all the aforementioned models are formulated with computer programs written in Matlab. Finally, we would like to acknowledge the contributions of:

– David Kosnik, Infrastructure Technology Institute

– Weizeng Zhang, PhD candidate in the Department of Civil & Environmental Engineering

5. Reference

Alwan, L. and Roberts, H. (1988), "Time-series modeling for statistical process control". *Journal of Business & Economic Statistics* 6 (1), 87-95.

Chen, Y., Corr, D., and Durango-Cohen, P. (2012a), "Analysis of common and special causes of variation in the deterioration of transportation infrastructure: A field application of statistical process control for structural health monitoring". Submitted to *Transportation Research Part B: Methodological*.

Chen, Y., Corr, D., and Durango-Cohen, P. (2012b), "Multivariate statistical process control procedure for structural health monitoring: An analytical framework". Preprint.

Deming, W. (1975), "On probability as a basis for action". *The American Statistician*, 29(4), 146-152.

Montgomery, D. (2009), *Introduction to Statistical Quality Control*, *6th Edi*. Wiley, Hoboken, NJ.

Shewhart, W. (1931), *Economic Control of Quality of Manufactured Products*. Van Nostrand Company, New York, NY.