IDENTIFICATION OF UNUSUAL EVENTS IN MULTI-CHANNEL BRIDGE MONITORING DATA

Piotr Omenzetter (**corresponding author**)

School of Civil and Environmental Engineering, Nanyang Technological University

50 Nanyang Avenue, Singapore 639798, Singapore

E-mail: cpiotr@ntu.edu.sg

Tel: +65-6790-4853

Fax: +65-6791-0046

James Mark William Brownjohn

School of Civil and Environmental Engineering, Nanyang Technological University

50 Nanyang Avenue, Singapore 639798, Singapore

Pilate Moyo

School of Civil and Environmental Engineering, Nanyang Technological University

50 Nanyang Avenue, Singapore 639798, Singapore

Abstract

Continuously operating instrumented structural health monitoring (SHM) systems are becoming a practical alternative to replace visual inspection for assessment of condition and soundness of civil infrastructure such as bridges. However, converting large amounts of data from an SHM system into usable information is a great challenge to which special signal processing techniques must be applied. This study is devoted to identification of abrupt, anomalous and potentially onerous events in the time histories of static, hourly sampled strains recorded by a multi-sensor SHM system installed in a major bridge structure and operating continuously for a long time. Such events may result, among other causes, from sudden settlement of foundation, ground movement, excessive traffic load or failure of post-tensioning cables. A method of outlier detection in multivariate data has been applied to the problem of finding and localizing sudden events in the strain data. For sharp discrimination of abrupt strain changes from slowly varying ones wavelet transform has been used. The proposed method has been successfully tested using known events recorded during construction of the bridge, and later effectively used for detection of anomalous post-construction events.

Running title: Identification of Unusual Events in Monitoring Data

Key words: bridges, multivariate statistics, novelty detection, outlier analysis, wavelet transform

1. INTRODUCTION

Bridges, which are the key components of any transportation system, have expected life cycles as long as 50 or 100 years; it is however not uncommon to see even older structures which are still in service. During its long service life a bridge would be subject to aging processes, harsh environmental conditions and excessive loads, leading to deterioration. Consequences of impaired physical condition can be as severe as the collapse of the structure, causing great financial loss or even casualties, but even major protective maintenance and upgrading works or replacement of the bridge are expensive and troublesome. Thus, it is important to have updated information on structural condition and performance of bridges in order to early detect any worrying signs of decline and undertake protective countermeasures.

Transportation infrastructure authorities have long recognized the need to keep their bridges healthy and to this end have implemented various inspection and management programs. The current health monitoring practice is primarily based on visual inspection. Due to high manpower demand such inspections cannot be performed frequently. Other drawbacks of visual inspection based condition assessment include inaccessibility of critical parts of the structure and lack of information on actual loading. These shortcomings lead to subjective and inaccurate evaluations of bridges' safety and reliability [1]. As a result some bridges may be retrofitted or replaced, while in fact they are sound; on the other hand, existing damages in other bridges may not be identified until they become expensive to repair or dangerous for structural integrity.

An alternative to the periodic visual inspection can be continuously operating instrumented structural health monitoring (SHM) systems. While their application to in-service civil engineering structures, with exception for unique constructions such as super-long bridges, is still rare, they are becoming more and more feasible due to availability of affordable hardware and advances in signal processing techniques [2-4]. A long-term,

continuous SHM system consists of a host computer, an array of sensors (and possibly actuators), and communication hardware and software for data collection and interpretation. The system provides continuous measurements of structural response (strains, stresses, accelerations etc.) due to ambient loadings (temperature, wind, traffic, ground motion etc.) or purposely exerted test loadings. Measurement data can be used for verification if the structure is loaded and behaves as designed, and for warning on excessive loads, unusual behavior or damage. According to Sikorsky [5], SHM should be broadly defined as "the use of in-situ, nondestructive sensing and analysis of structural characteristics, including the structural response, for the purpose of identifying if damage has occurred (Level 1 of diagnostics), determining the location of damage (Level 2), estimating the severity of damage (Level 3), and evaluating the consequences of damage on structural serviceability, reliability and durability (Level 4)". While there have been many works on SHM monitoring in relation to, for example, mechanical or aerospace structures addressing the aforementioned various stages of damage-related studies, monitoring of civil infrastructures is still at a preliminary phase. To the authors' knowledge few researchers attempted to face the challenge of converting data from monitoring into a deep and profound knowledge of structural behavior and condition. The reason for this can be the complexity and uniqueness of civil structures, their size, and exposure to various ambient inputs that are difficult to quantify. In view of the absence of similar studies, the present work sets up a modest goal of developing an analytical method for detection of unusual, abrupt events in the time series of data obtained during long-term, continuous operation of an SHM system. With the above definition by Sikorsky, the proposed method can be classified within Level 1 of structural health diagnostics as it identifies events that could possibly manifest an onset of structural degradation or damage. In particular, hourly sampled, static strain time histories recorded on a major bridge structure are analyzed using wavelet transform in order to discriminate abrupt strain changes from slow ones. Then,

statistical analysis of wavelet coefficient time series is conducted to detect outliers, i.e. data that significantly protrude from the remainder. Preliminary work incorporating this approach was reported by Moyo and Brownjohn [6]. However, their analysis is basically applicable to data from a single sensor. Also, identification of outliers was based on a rather ad-hoc assumed threshold without studying statistical properties of the wavelet-transformed measurement data. In view of these limitations of the previous study, the present paper extends applicability of the method proposed in [6] to handle correlated, multi-channel measurement data. Coefficients of wavelet-transformed strains are investigated using vector autoregressive moving average models and outliers are identified via multivariate statistics approach. The monitoring program, theoretical basics, and analysis of measurement data are presented in the subsequent sections in a self-contained manner.

2. MONITORING OF THE BRIDGE STRUCTURE

The subject of the reported monitoring was the Singapore-Malaysia Second Link, also referred to as the Tuas Link (Figure 1). The bridge serves as a vehicular crossing between the Island of Singapore and Malaysian Peninsula, and is located in the Western side of the island. The bridge was completed in 1997 and opened to traffic in the same year. This section of the paper is intended to give only the basic outline of the structure and the monitoring program for the sake of self-containment, and readers interested in more details can find them in [6] and [7]. The bridge is about 1.9 km long and comprises 27 spans; the Singapore side is about 170 m long and the main span of this section is 92 m long. The bridge was cast in-situ using the balanced cantilever method, to enable the navigation channel to be kept free throughout the construction, and post-tensioned. The cross-section of the post-tensioned, continuous box girder varies in depth from 2.6 m to 6.5 m along the bridge length.

A suite of instruments was installed in the bridge in order to monitor its short-term and long-term performance under environmental and traffic loads. The instrumentation consists of four data loggers, twelve vibrating wire strain gauges, twelve pressure cells, forty four thermocouples and one tri-axial accelerometer, distributed in three segments (Segments 23, 27 and 31) of the main span (Figure 2). Three data loggers are responsible for static measurements, that is strains, stresses and temperature, and one data logger is responsible for dynamic measurements. All the data loggers are connected to a host computer, resident in the bridge. The host computer can be accessed remotely via a modem.

The monitoring program was divided into two components, namely static and dynamic monitoring, the results of the former one being the focus of this paper. Static monitoring involved the acquisition of stress, strain and temperature data from embedded sensors. The magnitude and variation of these parameters depend upon structural geometry and boundary conditions, material properties, and the environment and live loads. The environmental effects include ambient temperature, solar radiation and wind. All static data were recorded hourly.

3. THEORY

Moyo and Brownjohn [6] concentrated on analysis of static strains and concluded that the major challenge was to convert the large amount of data from the SHM system into usable information. They paid special attention to identification of abrupt, anomalous events in the time histories of strains. Such events may result, among other causes, from sudden settlement of foundation, ground movement, excessive traffic load or failure of post-tensioning cables. As these events can potentially be onerous for structural integrity, it is of interest to discriminate them in the whole body of SHM-generated data. Achieving this goal by visual inspection of time series can be at best only qualitative, and in many cases is difficult and inaccurate, if not impossible. To overcome these difficulties, the analytical tool used here for

identification of abrupt events is the wavelet transform (WT). Another important issue, not yet adequately addressed in the previous work, is the simultaneous inspection of data from an array of synchronized sensors. Such an analysis can be used to assess whether the identified changes in performance affect the whole structure, or whether their impact is localized. Here, statistical outlier analysis of multivariate data is conducted in order to account for correlation of signals recorded at various locations. In short, the strategy adopted consists of the following steps:

- Wavelet transform of strain time series

- Formulation of a vector dynamic regression model for wavelet coefficients

- Identification of outliers among dynamic regression model driven wavelet coefficients.

3.1 Discrete wavelet transform (DWT)

The wavelet transform is one of rapidly developing signal processing tools with new theoretical and applied works appearing every day. One of the most useful aspects of WT is the time-frequency localization, which enables zooming into data at both time and frequency, and close inspection of the signal's minute features. This section of the paper gives a brief description of DWT's fundamentals that are relevant for this study.

The basic idea of wavelet analysis is to represent general functions in terms of simple building blocks at different scales and positions. In the majority of practical signal processing applications, data decomposition is carried out using the discrete wavelet transform, which has an orthonormal basis and uses fast numerical algorithms. A given discrete signal in time $y(n)$ ($n \in \mathbb{Z}$, where $\mathbb{Z}$ denotes the set of integers) can be decomposed as follows [8]:

6

$$y(n) = \sum_{k \in \mathbb{Z}} c_{J,k} \phi_{J,k}(n) + \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(n) \tag{1}$$

where $\phi_{J,k}(n)$ and $\psi_{j,k}(n)$ are discrete scaling and wavelet functions, respectively. They are derived through translation and dilation from the basic scaling function $\phi(n)$ and the mother wavelet $\psi(n)$:

$$\phi_{j,k}(n) = 2^{-j/2} \phi\left(2^{-j} n - k\right), \quad j,k \in \mathbb{Z} \tag{2}$$

$$\psi_{j,k}(n) = 2^{-j/2} \psi\left(2^{-j} n - k\right), \quad j,k \in \mathbb{Z} \tag{3}$$

Coefficients $c_{J,k}$ and $d_{j,k}$ are termed scaling and detail coefficients, respectively, and are defined through the following standard inner products:

$$c_{J,k} = \left\langle \phi_{J,k}(n), y(n) \right\rangle \tag{4}$$

$$d_{j,k} = \left\langle \psi_{j,k}(n), y(n) \right\rangle \tag{5}$$

Physical meaning of scaling and wavelet coefficients is such that the scaling coefficient $c_{J,k}$ shows average values of $y(n)$ at scale $s_J = 2^J$ in the vicinity of time $n = 2^j k$, whereas the wavelet coefficient $d_{j,k}$ is associated with changes of $y(n)$ at scale $s_j = 2^j$ in the vicinity of time $n = 2^j k$.

Under normal conditions, static strains in a bridge girder vary slowly and smoothly with time, the main cause of their change being the diurnal temperature variation. Thus, abrupt changes would cause sudden, localized discontinuity in the variation of strains. Wavelet coefficients in the neighborhood of the abrupt change in the signal or its derivative significantly differ from others [9], hence wavelet analysis comes as the natural tool for detection of these discontinuities. In what follows we will restrict the study to real-valued WT.

3.2 Statistical model for wavelet coefficients

Wavelet coefficients of signals from m measurement channels at scale j and time k can be considered as an m-dimensional random vector:

$$\mathbf{d}_k^{(j)} = \begin{bmatrix} d_{j,k}^{(1)} & d_{j,k}^{(2)} & \dots & d_{j,k}^{(m)} \end{bmatrix}^T \tag{6}$$

where "T" denotes transposition. To simplify notation, hereafter the superscript j will generally be omitted and used only when necessary to avoid confusion. Without loss of generality it can be assumed that random process $\{\mathbf{d}_k\}$ has a zero mean:

$$E\mathbf{d}_k = \mathbf{0} \tag{7}$$

where E denotes the expected value operator and the symbol "$\mathbf{0}$" is used to indicate a zero matrix or vector of dimensions to be understood by the context. Consequently, hereafter any references to mean values of $\{\mathbf{d}_k\}$ are omitted. The covariance matrix function for process $\{\mathbf{d}_k\}$ is defined as follows [10]:

$$E\mathbf{d}_{k-l}\mathbf{d}_k^T = \mathbf{\Sigma}_l^{(d)} \tag{8}$$

where superscript l is the time lag. Strain signals simultaneously measured at several spatial locations will often be correlated and so will be their corresponding wavelet coefficients, and hence, the dispersion matrix $\mathbf{\Sigma}_0^{(d)}$ [11] will be non-diagonal. This suggests that multivariate statistics should be used in the analysis of wavelet coefficient time series rather than univariate statistics. In addition, the covariance function will also be non-zero for some lags l different than zero, i.e. $\mathbf{\Sigma}_l^{(d)} \neq \mathbf{0}$, $l \in \mathbb{Z}$. This is so, because the wavelet coefficients $\mathbf{d}_{k-l}$ and $\mathbf{d}_k$, as defined in Equation (5), are linear combinations of several values of the analyzed signal in the possibly overlapping neighborhoods of times $k-l$ and l, respectively. Such

correlation among the contemporary signal and its lagged version is characteristic for the autoregressive moving average (ARMA) process, thus suggesting that process $\{\mathbf{d}_k\}$ should be modeled using a vector ARMA representation of order (p, q) (referred to as ARMA(p, q)):

$$\boldsymbol{\Phi}_p(B)\mathbf{d}_k = \boldsymbol{\Theta}_q(B)\mathbf{a}_k \tag{9}$$

$\boldsymbol{\Phi}(B)$ and $\boldsymbol{\Theta}(B)$ are matrix polynomials in the back shift operator B:

$$\boldsymbol{\Phi}_p(B) = \mathbf{I} - \sum_{i=1}^{p} \boldsymbol{\Phi}_i B^i \tag{10}$$

$$\boldsymbol{\Theta}_q(B) = \mathbf{I} - \sum_{i=1}^{q} \boldsymbol{\Theta}_i B^i \tag{11}$$

where $\boldsymbol{\Phi}_i$ ($i = 1, \ldots, p$) and $\boldsymbol{\Theta}_i$ ($i = 1, \ldots, q$) are $m \times m$ matrices, and $\mathbf{I}$ is the identity matrix. The action of the back shift operator on a vector time series $\{\mathbf{x}_k\}$ is defined as follows:

$$B\mathbf{x}_k = \mathbf{x}_{k-1} \tag{12}$$

Process $\{\mathbf{a}_k\}$ is a stationary zero-mean multivariate Gaussian white noise process with dispersion matrix $\boldsymbol{\Sigma}^{(a)}$, i.e. $\{\mathbf{a}_k\}$ has an m-variate zero-mean normal distribution, in short $\{\mathbf{a}_k\} \sim N_m\left(\mathbf{0}, \boldsymbol{\Sigma}^{(a)}\right)$.

Assuming that the process defined in Equation (9) is invertible [10], we can rewrite it in the following autoregressive (AR) from:

$$\boldsymbol{\Pi}(B)\mathbf{d}_k = \mathbf{a}_k \tag{13}$$

where

$$\boldsymbol{\Pi}(B) = \left[\boldsymbol{\Theta}_q(B)\right]^{-1} \boldsymbol{\Phi}_p(B) = \mathbf{I} - \sum_{i=1}^{\infty} \boldsymbol{\Pi}_i B^i \tag{14}$$

and $\mathbf{\Pi}_i$ ($i = 1,\dots\infty$) are $m \times m$ matrices. In practical applications the infinite sum in Equation (14) need be replaced by a finite one, resulting in the following AR model of order r (in short AR(r)):

$$\mathbf{\Pi}(B) \approx \mathbf{I} - \sum_{i=1}^{r} \mathbf{\Pi}_i B^i \tag{15}$$

It follows from Equation (9) and Gaussian distribution of $\{\mathbf{a}_k\}$ that process $\{\mathbf{d}_k\}$ is also normally distributed. Thus, detection of sudden changes in the analyzed strain signals, manifested by unduly large values of corresponding wavelet coefficients, can be achieved by examining the actual statistical distribution of wavelet coefficients for the presence of outliers.

3.3 Outlier detection in ARMA process

Detection of outliers in multivariate random data is more difficult than among univariate observations because outliers have more room to hide in the bulk of multidimensional data. The difficulty is even higher when observations are not independent from one another but are governed by an ARMA process, and the analysis needs to account for correlation between samples. The detection of time series outliers was first studied by Fox [12] for the univariate ARMA process. The theory presented here is an extension of Fox's concepts to handle multivariate situation.

Assume that $\{\mathbf{d}_k\}$ is an outlier-free process described by ARMA model of Equation (13), whereas $\{\mathbf{f}_k\}$ is an outlier-including process. Fox [12] considers two types of outliers that lead to spurious observations: the additive outliers (AOs) and the innovative outliers (IOs). They are respectively defined as additions to $\{\mathbf{d}_k\}$:

$$\mathbf{f}_k = \mathbf{d}_k + \boldsymbol{\omega} I_k^{(K)} \tag{16}$$

10

$$\mathbf{f}_k = \mathbf{d}_k + \left[\mathbf{\Pi}(B)\right]^{-1}\mathbf{\omega}I_k^{(K)} \tag{17}$$

where $\mathbf{\omega}$ is the amplitude of an outlier and $I_k^{(K)}$ is the indicator variable signaling the presence or absence of an outlier at time K and is defined as follows:

$$I_k^{(K)} = \begin{cases} 1, & k = K \\ 0, & k \neq K \end{cases} \tag{18}$$

It can be seen respectively from Equations (16) and (17) that an AO affects only the K-th observation, whereas the effect of an IO propagates to the subsequent observations due to the memory of the system defined by operator $\mathbf{\Pi}(B)$. Assuming that a time series contains several AOs and IOs, it can be described by the general model:

$$\mathbf{f}_k = \mathbf{d}_k + \sum_{i=1}^{M}\mathbf{v}_i(B)\mathbf{\omega}_i I_k^{(K_i)} \tag{19}$$

where $\mathbf{v}_i(B) = \mathbf{I}$ for an AO and $\mathbf{v}_i(B) = \left[\mathbf{\Pi}(B)\right]^{-1}$ for an IO, respectively.

Wei [10] describes a method for estimation of one outlier's magnitude when all parameters in Equation (19), i.e. the ARMA model as well as outlier's type and time of occurrence, are known. Extending his results to a multivariate case, we define

$$\mathbf{e}_k = \mathbf{\Pi}(B)\mathbf{f}_k \tag{20}$$

and from Equations (16) and (17) have respectively for an AO and an IO:

$$\mathbf{e}_k = \mathbf{\Pi}(B)\mathbf{\omega}I_k^{(K)} + \mathbf{a}_k \tag{21}$$

$$\mathbf{e}_k = \mathbf{\omega}I_k^{(K)} + \mathbf{a}_k \tag{22}$$

The least square estimations of the magnitude of an AO and IO occurring at time K, respectively denoted by $\hat{\mathbf{\omega}}_K^{(A)}$ and $\hat{\mathbf{\omega}}_K^{(I)}$, are then:

$$\hat{\mathbf{\omega}}_K^{(A)} = \mathbf{T}^{-1}\mathbf{\Pi}^*(F)\mathbf{e}_K \tag{23}$$

$$\hat{\boldsymbol{\omega}}_K^{(I)} = \mathbf{e}_K \tag{24}$$

where

$$\boldsymbol{\Pi}^*(F) = \mathbf{I} - \sum_{i=1}^{r} \boldsymbol{\Pi}_i^T F^i \tag{25}$$

and F is the forward shift operator acting on a time series $\{\mathbf{x}_k\}$ such that

$$F\mathbf{x}_k = \mathbf{x}_{k+1} \tag{26}$$

Matrix $\mathbf{T}$ in Equation (23) is defined as

$$\mathbf{T} = \mathbf{I} + \sum_{i=1}^{r} \boldsymbol{\Pi}_i^T \boldsymbol{\Pi}_i \tag{27}$$

The dispersion matrices of outlier magnitude estimators in Equations (23) and (24) can be shown to be respectively equal to

$$\boldsymbol{\Sigma}^{(A)} = E\hat{\boldsymbol{\omega}}_K^{(A)}\hat{\boldsymbol{\omega}}_K^{(A)^T} = \mathbf{T}^{-1}\left(\boldsymbol{\Sigma}^{(a)} + \sum_{i=1}^{r} \boldsymbol{\Pi}_i \boldsymbol{\Sigma}^{(a)} \boldsymbol{\Pi}_i^T\right)\mathbf{T}^{-1} \tag{28}$$

$$\boldsymbol{\Sigma}^{(I)} = E\hat{\boldsymbol{\omega}}_K^{(I)}\hat{\boldsymbol{\omega}}_K^{(I)^T} = \boldsymbol{\Sigma}^{(a)} \tag{29}$$

The outlier test is based on the Mahalanobis distances [11] defined as:

$$\Delta_K^{(A)} = \left(\hat{\boldsymbol{\omega}}_K^{(A)^T}\boldsymbol{\Sigma}^{(A)^{-1}}\hat{\boldsymbol{\omega}}_K^{(A)}\right)^{1/2} \tag{30}$$

$$\Delta_K^{(I)} = \left(\boldsymbol{\omega}_K^{(I)^T}\boldsymbol{\Sigma}^{(I)^{-1}}\boldsymbol{\omega}_K^{(I)}\right)^{1/2} \tag{31}$$

The Mahalanobis distance differs from the Euclidian distance in that it accounts for the relative dispersions and correlations among vector elements. It is a probabilistic distance in the sense that equal distances imply equal likelihoods.

The null hypothesis of the outlier test, H_0, assumes that $\mathbf{f}_K$ is neither an AO nor an IO. The two hypotheses competing with H_0 are: H_A assuming that $\mathbf{f}_K$ is an AO, and H_I assuming that $\mathbf{f}_K$ is an IO, respectively. The likelihood test statistics for H_A versus H_0 and

12

H_I versus H_0 are the squared Mahalanobis distances, $\Delta_K^{(A)^2}$ and $\Delta_K^{(I)^2}$, respectively. Under the null hypothesis the magnitudes of outliers are normally distributed [10], $\left\{ \boldsymbol{\omega}_K^{(A)} \right\} \sim N_m \left(\mathbf{0}, \boldsymbol{\Sigma}^{(A)} \right)$ and $\left\{ \boldsymbol{\omega}_K^{(A)} \right\} \sim N_m \left(\mathbf{0}, \boldsymbol{\Sigma}^{(I)} \right)$, and thus both squared Mahalanobis distances follow a chi-square distribution with m degrees of freedom [11], $\left\{ \Delta_K^{(A)^2} \right\} \sim \chi_m^2$ and $\left\{ \Delta_K^{(I)^2} \right\} \sim \chi_m^2$.

While testing whether a particular observation is an outlier, the value of its squared Mahalanobis distance is compared against an assumed threshold C. The knowledge of statistical distributions of the squared Mahalanobis distances of outlier magnitudes enables relating the threshold's value to the upper 100α-th percentile of the m-degree-of-freedom chi-square distribution, denoted by $\chi_m^2 (\alpha)$:

$$C = \chi_m^2 (\alpha) \tag{32}$$

where

$$\alpha = 1 - F_{\chi_m^2} \left[\chi_m^2 (\alpha) \right] \tag{33}$$

and $F_{\chi_m^2}$ is the cumulative distribution function for the m-degree-of-freedom chi-square distribution. Thus, significance levels can be assigned to the decision whether a tested observation is an inlier or an outlier: an observation is declared to be an inlier if $\Delta_K^{(A)^2} \leq C$ and $\Delta_K^{(I)^2} \leq C$, or the null hypothesis H_0 is rejected with the significance level of $\alpha\%$ in favor of H_A or H_I if $\Delta_K^{(A)^2} > C$ or $\Delta_K^{(I)^2} > C$, respectively.

It should be noted that the above procedure for outlier detection assumes that all the parameters of the outlier-free AR model of Equation (13), including the white noise

dispersion matrix $\mathbf{\Sigma}^{(a)}$, are known. In practice, however, they usually need to be estimated from available data, and presence of outliers can lead to a serious bias in the estimates. Quite often, as is also true in the case of the analyzed bridge monitoring data, none of the portions of measured time series can be declared outlier-free beforehand and used for an unbiased estimation of the AR model. To deal with this problem, a multivariate extension of the iterative outlier detection procedure of Chang and Tiao [13] reported in [10] is used. Details of the procedure are given in Appendix.

Having identified outliers in multivariate data it is of interest to examine which of the vector components contribute most to the Mahalanobis distance. A convenient measure was proposed by Worden et al. [14], who used individual "diagonal" components, $\mathbf{\Sigma}_{ii}^{-1}\mathbf{\omega}_K^{(i)2}$ ($i = 1, 2, \ldots, m$), that sum together with "off diagonal" components, $\mathbf{\Sigma}_{ij}^{-1}\mathbf{\omega}_K^{(i)}\mathbf{\omega}_K^{(j)}$ ($i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, m$; $i \neq j$), to form the squared Mahalanobis distance. Here, $\mathbf{\Sigma}_{ij}^{-1}$ denotes element (i, j) of an inverse of dispersion matrix $\mathbf{\Sigma}$ and $\mathbf{\omega}_K^{(i)}$ i-th component of vector $\mathbf{\omega}_K$. In the present study, normalized square roots of diagonal components defined as follows

$$D_K^{(i)} = \frac{\sqrt{\mathbf{\Sigma}_{ii}^{-1}\mathbf{\omega}_K^{(i)2}}}{\max\limits_{i=1,2,\ldots,m}\left\{\sqrt{\mathbf{\Sigma}_{ii}^{-1}\mathbf{\omega}_K^{(i)2}}\right\}} \tag{34}$$

are used to assess the relative contribution of each channel to the examined outlier.

4. APPLICATION

4.1 Method verification using monitoring data recorded during bridge construction

In order to verify the proposed analytical method we first use it to analyze strain time histories recorded during construction works. As mentioned before, the bridge was built using the balanced cantilever method and in-situ concreting [15]. In this bridge erection technique

the construction process departs from a pier, and consists of repetitive steps during which two segments of the bridge girder, symmetrically located with respect to the pier, are first concreted, then stressed using tendons, and then concreting form traveler is shifted. At the end of construction of the span the closure segment is cast and continuity or integration tendons are stressed. Then the girder changes from separate cantilevers into a continuous beam. Each of these events, i.e. concreting, tensioning, shifting of form traveler last for a few hours and are separated by a few days' time intervals. They are performed according to a construction schedule and are well documented. At the same time, these events can be expected to produce significant, sudden changes in strains. Thus, the proposed analytical method to be effective should be able to identify the concreting, tensioning and form traveler shifting events. This section of the paper is devoted to examining the performance of identification method as well as explaining and illustrating its details. Another useful result of analysis of known events can be the ability to compare them to identified future unknown events. Noticing similarities between known and unknown events may be helpful in understanding the causes and consequences of the latter ones.

A portion of strain time series recorded by four strain gauges located in Segment 31 was chosen for the purpose of verification of the identification method. The strain gauges are denoted as SG31-1, SG31-2, SG31-3 and SG31-4 (see Figure 2); SG31-1 and SG31-3 are placed close to the top of the girder, whereas SG31-2 and SG31-4 close to the bottom. The strain time series starts on April 29, 1997 at 00:00 hours, corresponding to time $n = 1$, and comprises 1500 hourly sampled measurements. These time series for the fours strain gauges are shown in Figure 3. The sign convention is such that positive strains denote contraction; in addition, the numerical values of strains, given here in the microstrain units ($1\mu\varepsilon = 10^{-6}$), correspond to relative rather then absolute strains and are measured from unknown reference

levels, which are generally different for different channels. Table 1 lists major stages of the construction schedule together with the times when they were carried out.

Firstly, the time series of strains were wavelet-transformed using the Daubechies orthonormal wavelet systems of order N and support length $2N-1$, denoted by $D(N)$ [8]. In particular, results presented here were obtained using the $D(4)$ system. This order was chosen as a trade-off between good time and poor frequency localization of lower order systems and poor time and good frequency localization of higher order ones. Figure 4 shows time series of wavelet coefficients at the lowest scale, $\mathbf{d}_k^{(1)}$, for the four strains. For easier comparison of strain and wavelet coefficient time series, the latter ones were plotted with time interval of 2, to compensate for the down-sampling inherent in DWT. Several significant excursions from the mean value of zero can be seen in the wavelet coefficient time series in all channels, and the subsequent analysis aims at classifying them as inliers or outliers.

To choose an adequate order of AR process governing the wavelet coefficients, several AR(r) models with different r were checked and it turned out that a low order model - AR(2) ensures sufficient accuracy. The rationale behind this choice was that all values of AR model matrix coefficients $\mathbf{\Pi}_i$ for $i \geq 3$ did not have significant values, and process $\{\mathbf{a}_k\}$ became white noise for AR(2). The AR model parameter estimation and simultaneous outlier analysis ware conducted using all available data. The initial estimation of parameters of AR(2) and their standard deviations computed under the assumption that there are no outliers were obtained through the least squares method [10]. The results are shown below where numbers in parenthesis are the standard deviations; (only these entries that exceed 1.5 of their respective standard deviations are retained, while the other are set to zero):

$$\hat{\mathbf{\Pi}}_1 = \begin{bmatrix} -0.42 \ (0.07) & 0.15 \ (0.07) & -0.11 \ (0.07) & 0 \\ 0 & -0.37 \ (0.05) & 0 & -0.11 \ (0.05) \\ -0.22 \ (0.07) & 0.14 \ (0.07) & -0.24 \ (0.07) & 0 \\ -0.12 \ (0.05) & 0 & 0.10 \ (0.05) & -0.33 \ (0.05) \end{bmatrix} \tag{35}$$

$$\hat{\mathbf{\Pi}}_2 = \begin{bmatrix} 0 & 0 & -0.14 \ (0.07) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.08 \ (0.05) \end{bmatrix} \tag{36}$$

and the initial estimation of the white noise dispersion matrix is

$$\hat{\mathbf{\Sigma}}^{(a)} = \begin{bmatrix} 0.58 & 0.07 & 0.34 & -0.12 \\ & 0.24 & -0.11 & -0.05 \\ & & 0.53 & 0.09 \\ sym. & & & 0.27 \end{bmatrix} \tag{37}$$

By examining the diagonal and off-diagonal coefficients of AR(2) model parameter matrices, it can be seen that wavelet coefficients in all four channels are related to their own past values as well as cross-related to past values in other channels. The relationship is strong between the current observations and the observations immediately preceding them, but much smaller for earlier observations. By examining $\hat{\mathbf{\Sigma}}^{(a)}$, which has significant values of off-diagonal entries, it can be concluded that the signals' wavelet coefficients are also contemporaneously related.

An important step in the outlier detection procedure is the choice of the threshold value C. This threshold has to be judicially chosen because too low a value would lead to an excessive number of outliers and false alarms, whereas too high a value could leave some outliers undetected. After several trials the threshold to be used in this study was chosen as $C=24.50$. The value of α for the upper 100α-th percentile of the 4-degree-of-freedom chi-square distribution, $\chi_4^2(\alpha)$, [Equation (33)], is this case very low: $\alpha = 6.3 \times 10^{-5}$. To better

17

appreciate the order of this threshold value it can be explained as follows. The χ_4^2 variable, which describes statistical distribution of the Mahalanobis distances, is a sum of four independent variables with a zero-mean unit-variance normal distribution, $N(0,1)$. Taking the value of probability of exceeding ± 4 standard deviations for an $N(0,1)$ variable and finding the value of a χ_4^2 variable corresponding to this probability yields the assumed value of C. It appears that the chosen threshold is relatively high.

The outlier detection procedure classified about 14% of observations as outliers. It should be noticed that, given a relatively high value of the threshold, the original distribution of noise $\{\mathbf{a}_k\}$ before outlier identification was quite heavily tailed as compared to the normal distribution. This could be explained by the fact that during construction the bridge experienced many unusual, abrupt events due to construction works and corresponding loads, such as excitation from construction machines and equipment. Durations of those events often extended for several hours influencing a number of wavelet coefficients in their vicinity. In addition, some outlying observations could be just noise due to, for example, electromagnetic interference. It should also be noticed that the wavelet transformation is a trade off between time and frequency localization. For a given sharp jump in signal at time instant n not only a single corresponding wavelet coefficient at that time is affected but also a few coefficients nearby. That means an identified event is quite often manifested not by a single outlier but rather by a few outliers.

The final estimation of parameters of AR(2) after outliers have been removed is

$$\hat{\mathbf{\Pi}}_1 = \begin{bmatrix} -0.37\ (0.03) & 0 & 0 & 0 \\ -0.03\ (0.02) & -0.31\ (0.02) & 0.03\ (0.02) & -0.11\ (0.02) \\ -0.20\ (0.03) & 0.08\ (0.03) & -0.14\ (0.03) & 0 \\ -0.06\ (0.02) & -0.03\ (0.02) & 0.02\ (0.02) & -0.40\ (0.02) \end{bmatrix} \tag{38}$$

$$\hat{\mathbf{\Pi}}_2 = \begin{bmatrix} 0 & \text{-0.07 } (0.03) & \text{-0.05 } (0.03) & 0.06 \ (0.03) \\ \text{-0.05 } (0.02) & \text{-0.06 } (0.02) & 0 & \text{-0.05 } (0.02) \\ 0 & \text{-0.06 } (0.03) & \text{-0.09 } (0.03) & 0 \\ 0 & \text{-0.04 } (0.02) & \text{-0.03 } (0.02) & \text{-0.08 } (0.02) \end{bmatrix} \tag{39}$$

and the dispersion matrix of the white noise process is

$$\hat{\Sigma}^{(a)} = \begin{bmatrix} 0.12 & -0.01 & 0.10 & 0.00 \\ & 0.04 & -0.01 & 0.02 \\ & & 0.12 & 0.01 \\ sym. & & & 0.03 \end{bmatrix} \tag{40}$$

It can be noticed that the white noise dispersion is now markedly smaller. As for the changes in AR(2) model parameters, comparing their estimated values with standard deviations, it can be concluded that they do not vary much from the initial estimates.

After describing the details of the outlier identification method we will now concentrate on examining the method's efficiency in detecting abrupt events experienced by the bridge. Figure 5 shows the Mahalanobis distances for all identified additive and innovative outliers and compares them to the square root of threshold, $C^{1/2}$. As can be seen in Figure 5 and Table 1 the outlier detection procedure was able to detect all but one (F25) known events during construction. Zooming into the strain time series revealed, however, that event F25 did actually not cause visually noticeable strain fluctuations. In addition to the known events, many other, unknown ones were discovered. Figure 6 shows examples of strain time series in vicinity of tensioning (T24), concreting (C24), shifting of form traveler (F26) as well as an unknown event (UN1). Events T24, C24 and F26 were chosen because they share common features with the remaining tensioning, concreting and form shifting events. The tensioning events are the most evident in Figure 5, followed by concreting, and finally form shifting. The latter two types give in fact similar Mahalanobis distances to many detected unknown events. Indeed, as seen in Figure 6, the tensioning event is characterized by

an abrupt jump of strain, while concreting and form shifting last longer and do not cause such strong changes in strain. On the other hand, some of the unknown events turned out to give origin to strain changes of sharpness comparable to the tensioning events. Figure 7 shows in the left hand side column the wavelet coefficients corresponding to the events of Figure 6, and in the right hand side column it shows bar plots explaining the relative contribution of each signal to the Mahalanobis distance as defined in Equation (34). Regarding event T24, it is most noticeable in bottom signal SG31-4 followed by bottom signal SG31-2; top signal SG31-1 is also quite strong whereas top signal SG31-3 much weaker. This pattern is similar for the other tensioning events as well, which were characterized by large values of wavelet coefficients (relative to their variances) of the bottom signals, whereas the top signals, although possibly strong, did not show any consistent trend. Concerning the concreting events all of them were also easily visible in the bottom signals, while the top signals did not contribute much to their detection. As to the form shifting events, usually all four channels contributed significantly to their detection. Eventually, the unknown event UN1 affected mainly the bottom strains, similarly to the concreting events. In fact, comparing the strain time series of events C24 and UN1 (see Figure 6) reveals their close resemblance: both bottom strains experience large, step-like changes, while the top strains are not much affected. Regarding the rest of identified unknown events, histograms of relative contribution of each signal to the Mahalanobis distance are shown in Figure 8. The histograms show the percentage of relative contributions laying in the intervals marked on the vertical axis, and high values in the interval 0.9-1.0 demonstrate that a particular signal was an important factor in detection of a large number of outliers. Thus, it can be noticed that signals SG31-1, 3 and 4 contributed almost equally to identification of outliers, while the bottom signal SG31-2 did not play such an important role.

4.2 Bridge monitoring after construction

In the previous section, the detection of outliers during construction period was conducted inclusively, i.e. the estimation of the AR model and dispersion matrix of white noise were conducted simultaneously with outlier detection, using the proposed iterative procedure described in the Appendix. To reduce computational burden associated with the iterative procedure, it is desirable to specify "training data" and use it for estimation of an AR model and noise dispersion matrix. Once these are computed, they can be regarded as known, and outlier detection in newly recorded data, referred to as the monitoring data, can be carried out exclusively. This approach will be used for analysis of post-construction strain data.

For the purpose of monitoring during post-construction stage, a portion of strain time series was chosen which was recorded by eight strain gauges located in Segments 23 and 31, denoted as SG23-1, SG23-2, SG23-3, SG23-4, SG31-1, SG31-2, SG31-3, and SG31-4 (see Figure 2). The record starts on May 19, 1999 at 17:00 hours, corresponding to time $n=1$, and comprises 2492 hourly sampled measurements. The training data ware chosen as the first 600 wavelet coefficients and AR(2) model was fitted to this data. The inclusive outlier detection procedure with a threshold C=32.93 found that about 6% of training data were outliers. The threshold value used here corresponds to the same significance level $\alpha = 6.3 \times 10^{-5}$ as in Section 4.1. The dispersion matrix after outlier removal is

$$
\hat{\boldsymbol{\Sigma}}^{(a)} =
\begin{bmatrix}
0.11 & 0.11 & 0.09 & 0.11 & 0.07 & 0.04 & 0.07 & 0.04 \\
 & 0.20 & 0.09 & 0.15 & 0.08 & 0.04 & 0.08 & 0.04 \\
 & & 0.16 & 0.11 & 0.06 & 0.04 & 0.07 & 0.05 \\
 & & & 0.26 & 0.08 & 0.05 & 0.09 & 0.05 \\
 & & & & 0.07 & 0.02 & 0.05 & 0.02 \\
 & & & & & 0.09 & 0.02 & 0.05 \\
 & & & & & & 0.08 & 0.02 \\
\textit{sym.} & & & & & & & 0.06
\end{bmatrix}
\tag{41}
$$

21

It is interesting to notice that even wavelet coefficients of strains from different segments are quite strongly contemporaneously correlated.

An exclusive outlier detection procedure using the estimated AR model and white noise dispersion matrix found about 9% outliers among the remainders of wavelet time series or the monitoring data, and the total number of outliers detected during post-construction stage amounted to 7.6% of the whole data. It can be seen that this number is smaller by a factor of almost two as compared to the number of outliers found in the data recorded during construction. This can be explained by the fact that during normal operation the bridge did not sustain so many strong, abrupt events such as tendon tensioning or segment concreting. Figure 9 shows the Mahalanobis distances for the identified outliers and compares them to the square root of threshold, $C^{1/2}$. It can be seen in this figure too, that post-construction events are not that strong as those during construction (cf. Figure 5).

Two examples of identified events, respectively referred to as Events E1 and E2, are shown in Figure 10, while their Mahalanobis distances are marked in Figure 9. Figure 11 shows bar plots explaining the relative contribution of each channel to the Mahalanobis distances for the two considered events. It can be noticed that Event E1 was registered by all the eight sensors, though the sensors in Segment 31 contributed more. Indeed, sharp changes in strain can be seen in Figure 10 for strain time series of all the channels, and it can be concluded from the sign of this changes that the girder was subjected to a strong compressive force. On the other hand, Event E2 was more localized in space and it was registered only by bottom strain gauges in the middle span, i.e. at Segment 23, and almost not felt at Segment 31. Figure 12 shows histograms of relative contribution of each signal to the Mahalanobis distances of all identified events. It is interesting that the bottom strain gauges SG21-2, SG23-4, SG31-2 and SG31-4 contributed much stronger to the detection of abnormal events, which is clear from high values of respective histograms in the 0.9-1.0 interval. It can be also

22

noticed that the histograms for all eight sensors have relatively large values in the interval 0.0-0.2. This shows that for a significant number of identified events some channels contributed strongly while the remaining ones did not. Thus, many of the events were localized in space, affecting only one of the instrumented girder segments or just selected sensors from a segment.

5. CONCLUSIONS AND FUTURE RESEARCH

A method of outlier detection in multivariate data has been applied to the problem of finding and localizing abnormal, sudden events in the strain data recorded during continuous, long-term operation of a multi-sensor SHM system installed on a full-scale bridge structure. For separation of abrupt strain changes from slowly varying ones wavelet transform has been used. Wavelet decomposition significantly enhances ability to detect sudden events, which are typically well hidden among other strain fluctuations when one wants to find them by visual inspection of strain time series. The multivariate statistics approach for detection of abrupt events enables joint, systematic analysis of mutually correlated signals from an array of sensors, and yields information about spatial localization of events. The proposed method has been successfully tested through application to the strains measured during construction of the bridge, when abrupt events are known beforehand, and later effectively used for detection of abnormal post-construction events.

While the proposed method enables detection of events which could signalize an onset of structural change or damage, it does not decisively tell that any form of damage has actually been sustained. Ongoing investigations will try to address this issue through application of the multivariate intervention analysis of time series. Intervention analysis enables examination of what happened to the system after an unusual event and checking if this event had a permanent impact or only a transient one. Furthermore, the proposed method

is able to spatially localize points in the structure where abrupt changes in strains occur, but it cannot point out what caused those strain fluctuations. It is planned to formulate an analytical bridge model and use numerical simulations of different damage scenarios, such as post-tensioning tendon failure or foundation settlement, to obtain some insight into the possibility of linking patterns in measured strain to the physical events and processes experienced by the structure.

Eventually, as the currently used SHM system merely collects the measurement data, it is of interest to augment its software with a module that would be able to perform signal processing described in this paper on site, immediately after strain readings are available. Such an extended SHM system could operate autonomously issuing warnings to an operator, and would be a further step towards intelligent civil engineering structures.

Equation Section (Next)

6. APPENDIX. PROCEDURE FOR ITERATIVE OUTLIER DETECTION

The procedure shown here enables iterative estimation of parameters of an AR model of Equation (13) as well as detection of outliers when the number of AOs and IOs is not known beforehand. The steps of the procedure are as follows.

Step 1. Model time series $\{\mathbf{f}_k\}$ as if there were no outliers; the estimated model is denoted by $\hat{\mathbf{\Pi}}(B)$. Compute the residuals from the estimated model:

$$\hat{\mathbf{e}}_k = \hat{\mathbf{\Pi}}(B)\mathbf{f}_k \tag{A.1}$$

Compute initial estimate of dispersion matrix of $\{\mathbf{a}_k\}$:

$$\hat{\mathbf{\Sigma}}^{(a)} = \frac{1}{N}\sum_{k=1}^{N}\hat{\mathbf{e}}_k \tag{A.2}$$

where N is the size of the sample.

24

Step 2. Calculate $\hat{\Delta}_k^{(A)^2}$ and $\hat{\Delta}_k^{(I)^2}$ from Equations (30) and (31) using the estimated model and

define

$$\hat{\Delta}_K^2 = \max_{k=1,2\ldots,N} \max_{\{A,I\}} \left\{ \hat{\Delta}_k^{(A)^2}, \hat{\Delta}_k^{(I)^2} \right\} \tag{A.3}$$

If $\hat{\Delta}_K^2 = \hat{\Delta}_K^{(A)^2} > C$ then there is an AO at time K with an estimated magnitude of $\hat{\boldsymbol{\omega}}_K^{(A)}$.

Modify the data as follows:

$$\tilde{\mathbf{f}}_k = \mathbf{f}_k - \hat{\boldsymbol{\omega}}_K^{(A)} I_k^{(K)} \tag{A.4}$$

and compute modified residuals:

$$\tilde{\mathbf{e}}_k = \hat{\mathbf{e}}_k - \hat{\mathbf{\Pi}}(B) \hat{\boldsymbol{\omega}}_K^{(A)} I_k^{(K)} \tag{A.5}$$

If $\hat{\Delta}_K^2 = \hat{\Delta}_K^{(I)^2} > C$ then there is an IO at time K with an estimated magnitude of $\hat{\boldsymbol{\omega}}_K^{(I)}$. Modify

the data as follows:

$$\tilde{\mathbf{f}}_k = \mathbf{f}_k - \left[\hat{\mathbf{\Pi}}(B)\right]^{-1} \hat{\boldsymbol{\omega}}_K^{(I)} I_k^{(K)} \tag{A.6}$$

and compute modified residuals:

$$\tilde{\mathbf{e}}_k = \hat{\mathbf{e}}_k - \hat{\boldsymbol{\omega}}_K^{(I)} I_k^{(K)} \tag{A.7}$$

Compute a new estimate for the dispersion matrix, $\tilde{\mathbf{\Sigma}}^{(a)}$, from the new residuals.

Step 3. Recompute $\hat{\Delta}_k^{(A)^2}$ and $\hat{\Delta}_k^{(I)^2}$ based on the modified residuals $\tilde{\mathbf{e}}_k$ and dispersion matrix

$\tilde{\mathbf{\Sigma}}^{(a)}$, and repeat Step 2 until all outliers are identified. The initial estimate of $\hat{\mathbf{\Pi}}(B)$ remains

unchanged.

Step 4. Suppose that in Step 3 M outliers have been tentatively identified at times K_1, K_2,

…, K_M. Estimate the outlier magnitudes and the time series parameters simultaneously using

the model

$$\mathbf{f}_k = \sum_{i=1}^{M} \mathbf{v}_i(B)\boldsymbol{\omega}_i I_k^{(K_i)} + \left[\boldsymbol{\Pi}(B)\right]^{-1} \mathbf{a}_k \tag{A.8}$$

where $\mathbf{v}_i(B) = \mathbf{I}$ for an AO and $\mathbf{v}_i(B) = \left[\boldsymbol{\Pi}(B)\right]^{-1}$ for an IO, respectively. Compute new residuals

$$\hat{\mathbf{e}}_k = \hat{\boldsymbol{\Pi}}(B)\left[\mathbf{f}_k - \sum_{i=1}^{M} \hat{\mathbf{v}}_i(B)\hat{\boldsymbol{\omega}}_i I_k^{(K_i)}\right] \tag{A.9}$$

and revised estimate of $\hat{\boldsymbol{\Sigma}}^{(a)}$.

Recalculate $\hat{\Delta}_K^{(A)^2}$ and $\hat{\Delta}_K^{(I)^2}$ for the M outlier candidates and compare with threshold C. Discard of these candidates whose magnitudes do not exceed the threshold value. Repeat Step 4 until all outlier candidates' magnitudes are above the threshold value.

Repeat Steps 2 through 4 until all outliers are identified and their magnitudes as well as the time series parameters are simultaneously estimated. Doing so finally leads to the following estimated model:

$$\mathbf{f}_k = \sum_{i=1}^{M} \hat{\mathbf{v}}_i(B)\hat{\boldsymbol{\omega}}_i I_k^{(K_i)} + \left[\hat{\boldsymbol{\Pi}}(B)\right]^{-1} \mathbf{a}_k \tag{A.10}$$

REFERENCES

1. B. M. Phares, D. D. Rolander, B. A. Graybeal and G. A. Washer 2001 <u>Public Roads 64(5)</u>, 22-29. Reliability of Visual Bridge Inspection.

2. F. K. Chang (Ed.) 1997 <u>Structural Health Monitoring. Current Status and Perspectives</u>. Lancaster: Technomic.

3. F. K. Chang (Ed.) 1999 <u>Structural Health Monitoring 2000</u>. Lancaster: Technomic.

4. D. L. Balageas (Ed.) 2002 <u>Structural Health Monitoring 2002</u>. Lancaster: Destech Publications.

5. C. Sikorsky 1999 <u>Structural Health Monitoring 2000</u>, F. K. Chang (Ed.), 68-81. Development of a Health Monitoring System for Civil Structures Using a Level IV Non-destructive Damage Evaluation Method.

6. P. Moyo and J. M. W. Brownjohn 2002 <u>Mechanical Systems and Signal Processing 16(2/3)</u>, 429-445. Detection of Anomalous Structural Behavior Using Wavelet Analysis.

7. J. M. W. Brownjohn and P. Moyo 2000 <u>Second International Conference on Experimental Mechanics, Proceedings of SPIE 4317</u>, 528-533. Monitoring of Malaysia-Singapore Second Link During Construction.

8. G. Kaiser 1994 <u>A Friendly Guide to Wavelets</u>. Boston: Birkhäuser.

9. F. Abramovich and A. Samarov 2000 <u>Tel Aviv University, Israel, Technical Report RP-SOR-00-02</u>. On One-sided Estimation of a Sharp Cusp Using Wavelets.

10. W. W. S. Wei 1993 <u>Time Series Analysis</u>. Redwood City: Addison-Wesley.

11. G. A. F. Seber 1984 <u>Multivariate Observations</u>. New York: John Wiley & Sons.

12. A. J. Fox 1972 <u>Journal of Royal Statistical Society Ser. B, 43</u>, 350-363. Outliers in Time Series.

13. I. Chang and G. C. Tiao 1983 <u>University of Chicago, Statistics Research Center, Technical Report 8</u>. Estimation of Time Series Parameters in the Presence of Outliers.

14. K. Worden, G. Manson and N. R. J. Fieller 2000 <u>Journal of Sound and Vibration 229(3),</u> 647-667. Damage Detection Using Outlier Analysis.

15. J. Mathivat 1983 <u>The Cantilever Construction of Prestressed Concrete Bridges</u>. New York: John Wiley & Sons.

Table 1. Construction schedule and results of event detection.

Activity	Abbreviation	Time	Detection result
Concreting of Segment 27	C27	91-95	+
Tensioning of Segment 27	T27	154-156	+
Shifting of form off Segment 27	F27	180-181	+
Concreting of Segment 26	C26	285-289	+
Tensioning of Segment 26	T26	326-330	+
Shifting of form off Segment 26	F26	350-355	+
Concreting of Segment 25	C25	499-502	+
Tensioning of Segment 25	T25	539-541	+
Shifting of form off Segment 25	F25	572-576	-
Concreting of Segment 24	C24	717-723	+
Tensioning of Segment 24	T24	779-781	+
Shifting of form off Segment 24	F24	950-953	+
Concreting of Segment 23	C23	1234-1240	+
Tensioning of continuity tendons	TC	1359-1363	+
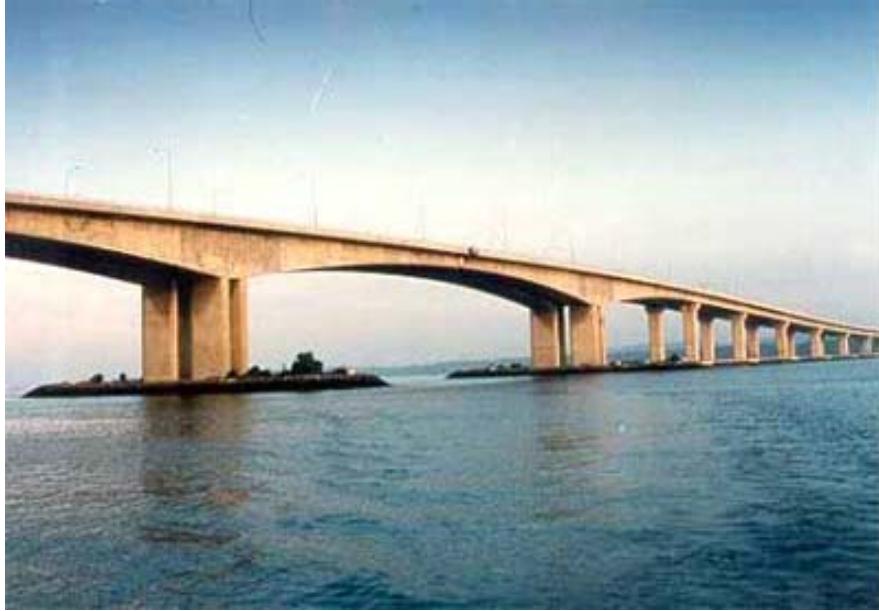
Figure 1. The Singapore-Malaysia Second Link.

Omenzetter et al.
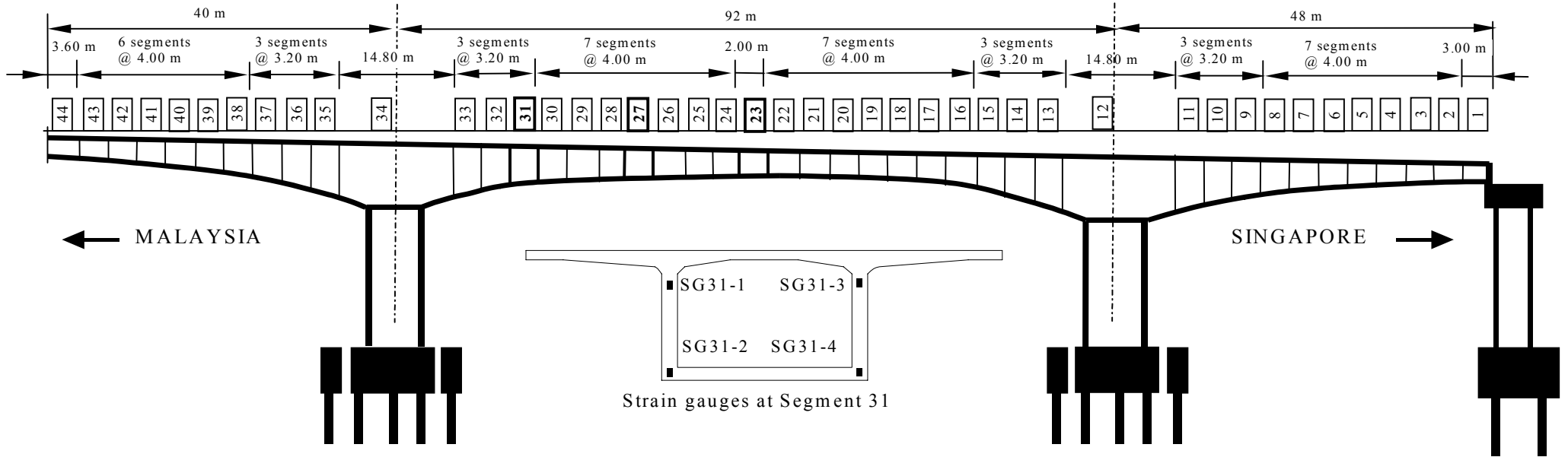
Figure 2. Bridge layout and strain gauges locations.
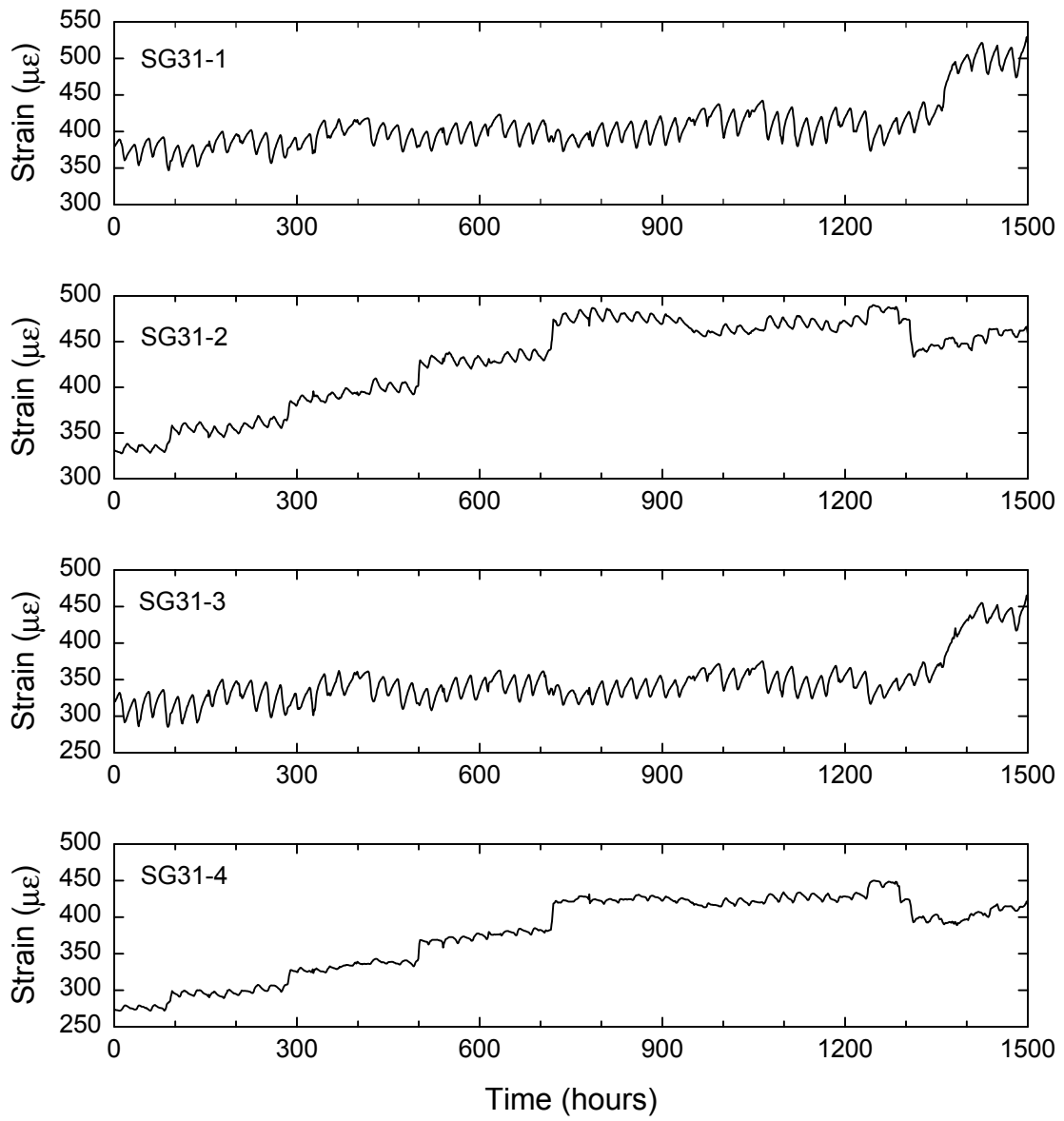
Omenzetter et al.

Figure 3. Strains recorded during construction at Segment 31.
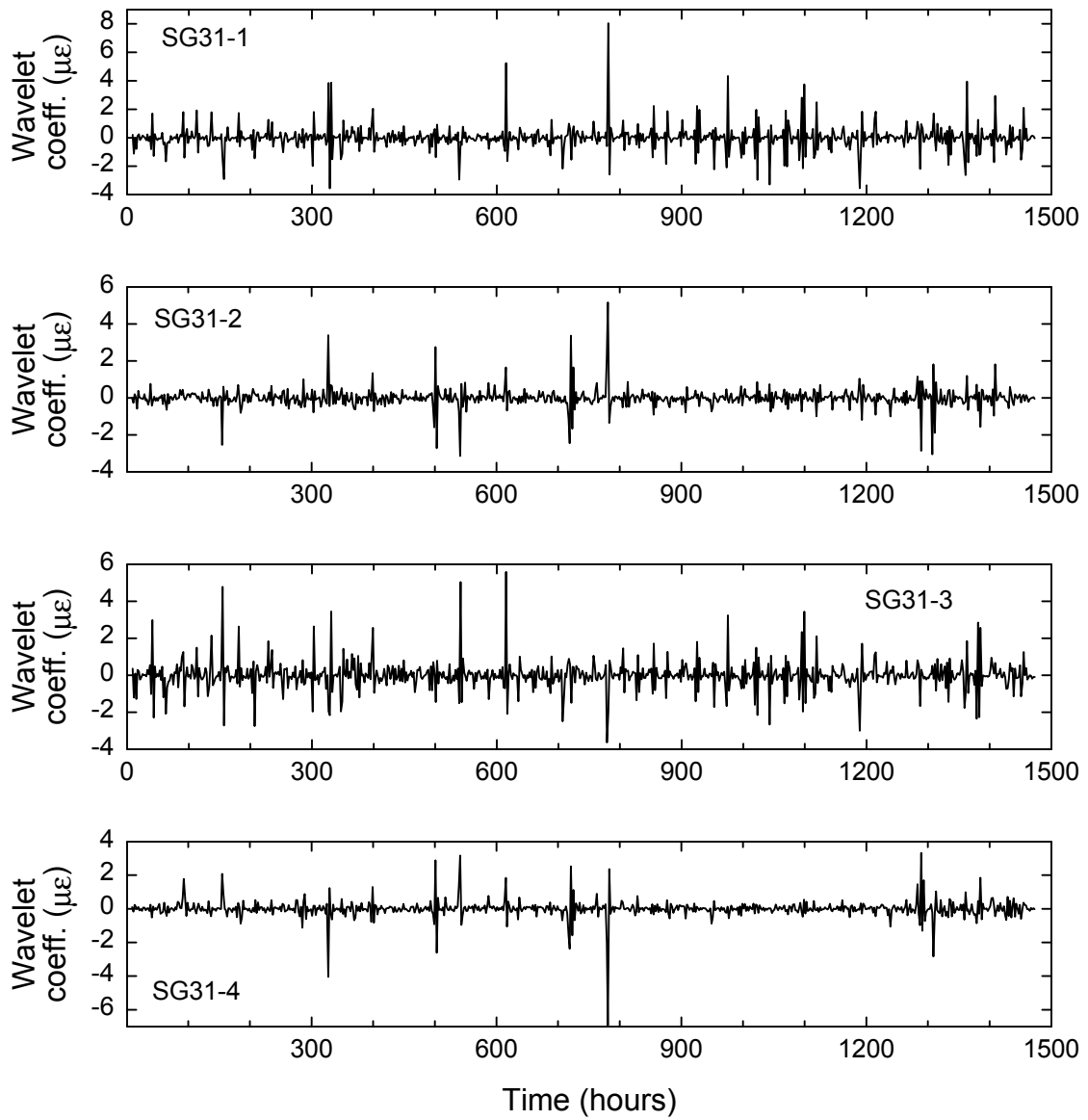
Omenzetter et al.

Figure 4. Wavelet coefficients at lowest scale of strains during construction at Segment 31.
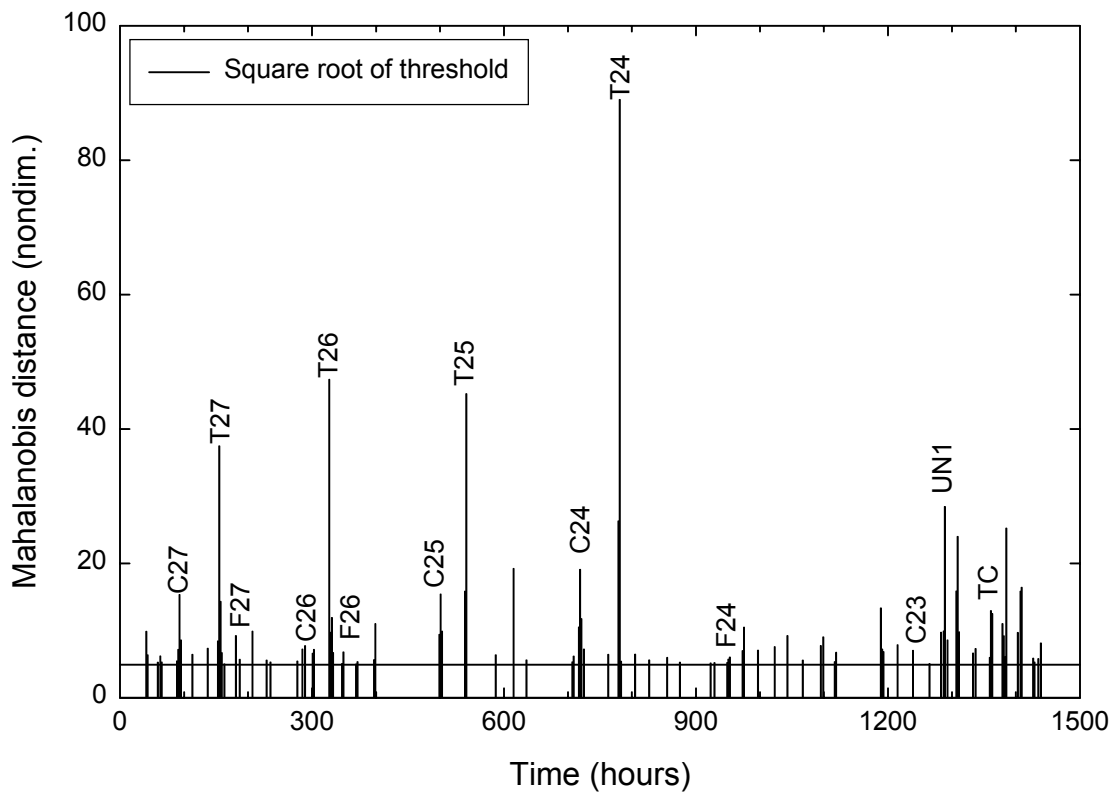
Omenzetter et al.

Figure 5. Mahalanobis distances for events identified during construction.
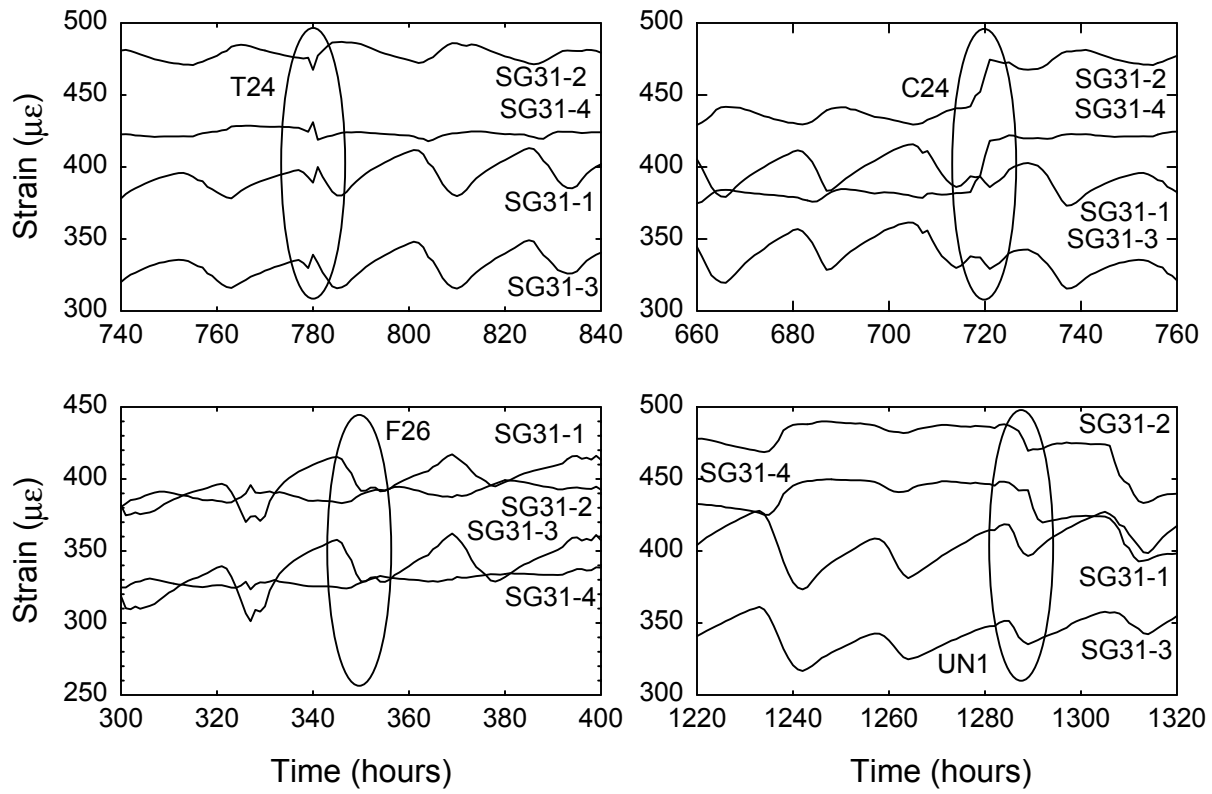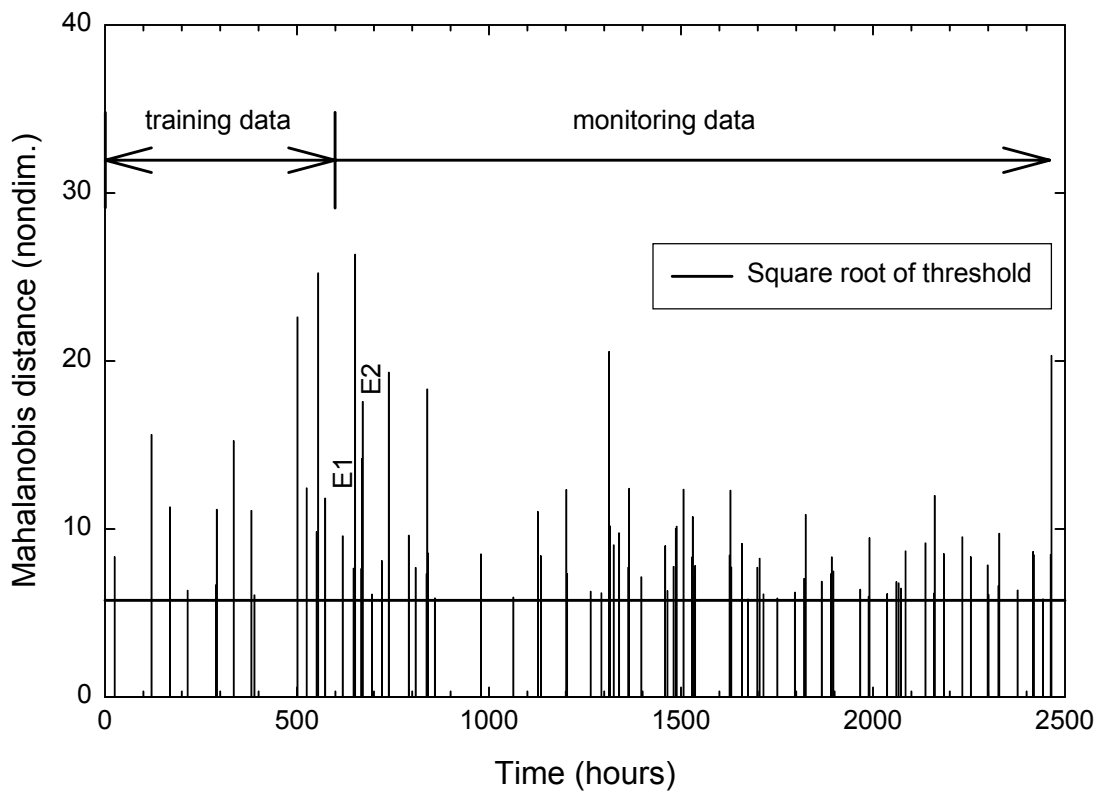
Omenzetter et al.

Figure 6. Strain time series for selected events identified during construction.

Omenzetter et al.
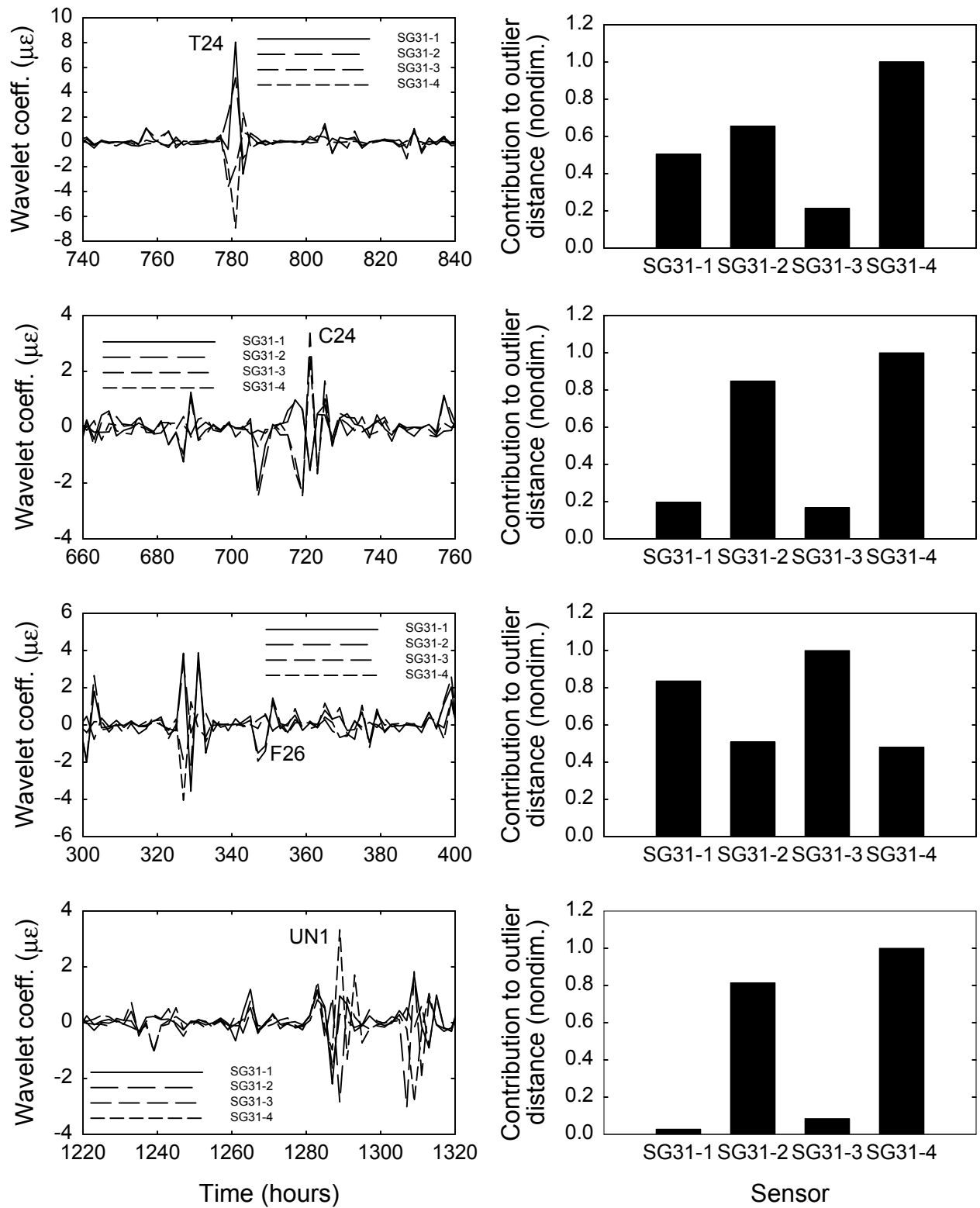
Figure 7. Wavelet coefficient time series and sensor contribution to outlier distance
for selected events identified during construction.
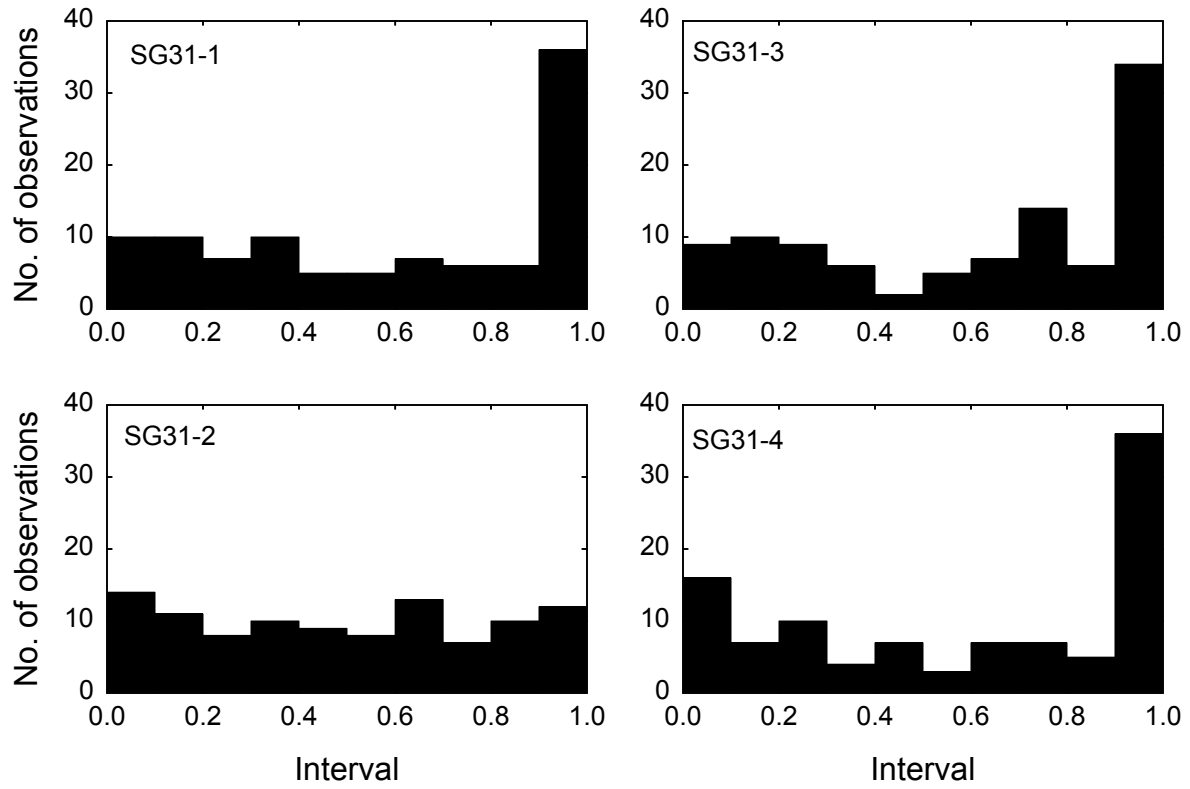
Omenzetter et al.

Figure 8. Histograms of sensor contribution to outlier distance for all events identified during construction.

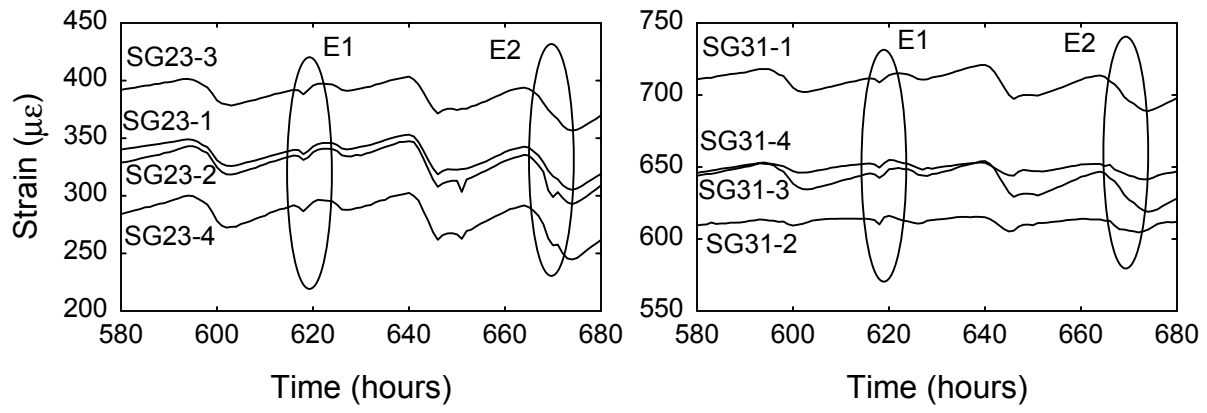Figure 9. Mahalanobis distances for events identified after construction.

Omenzetter et al.

Figure 10. Strain time series for selected events identified after construction.
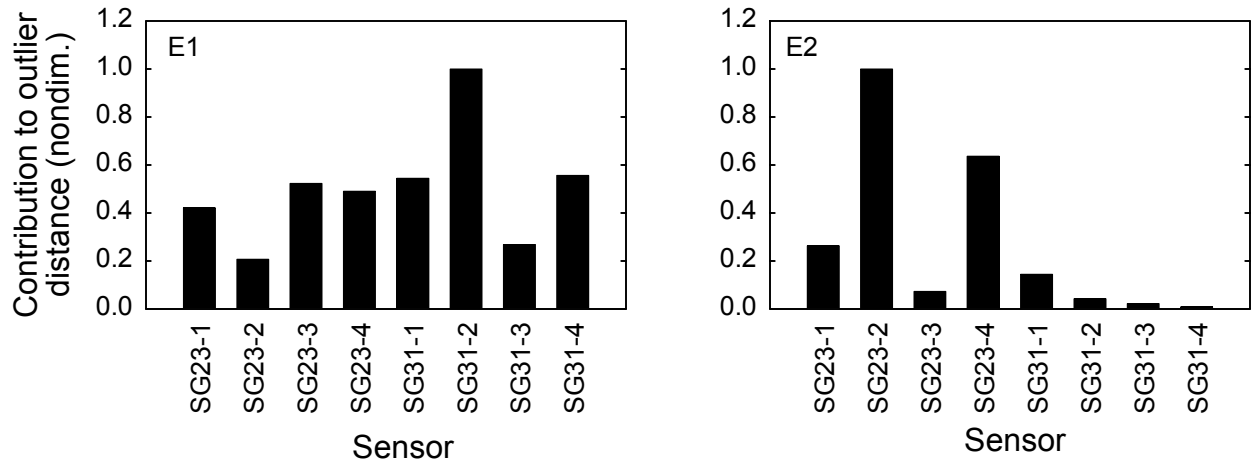
Omenzetter et al.

Figure 11.Sensor contribution to outlier distance for selected events identified after construction.
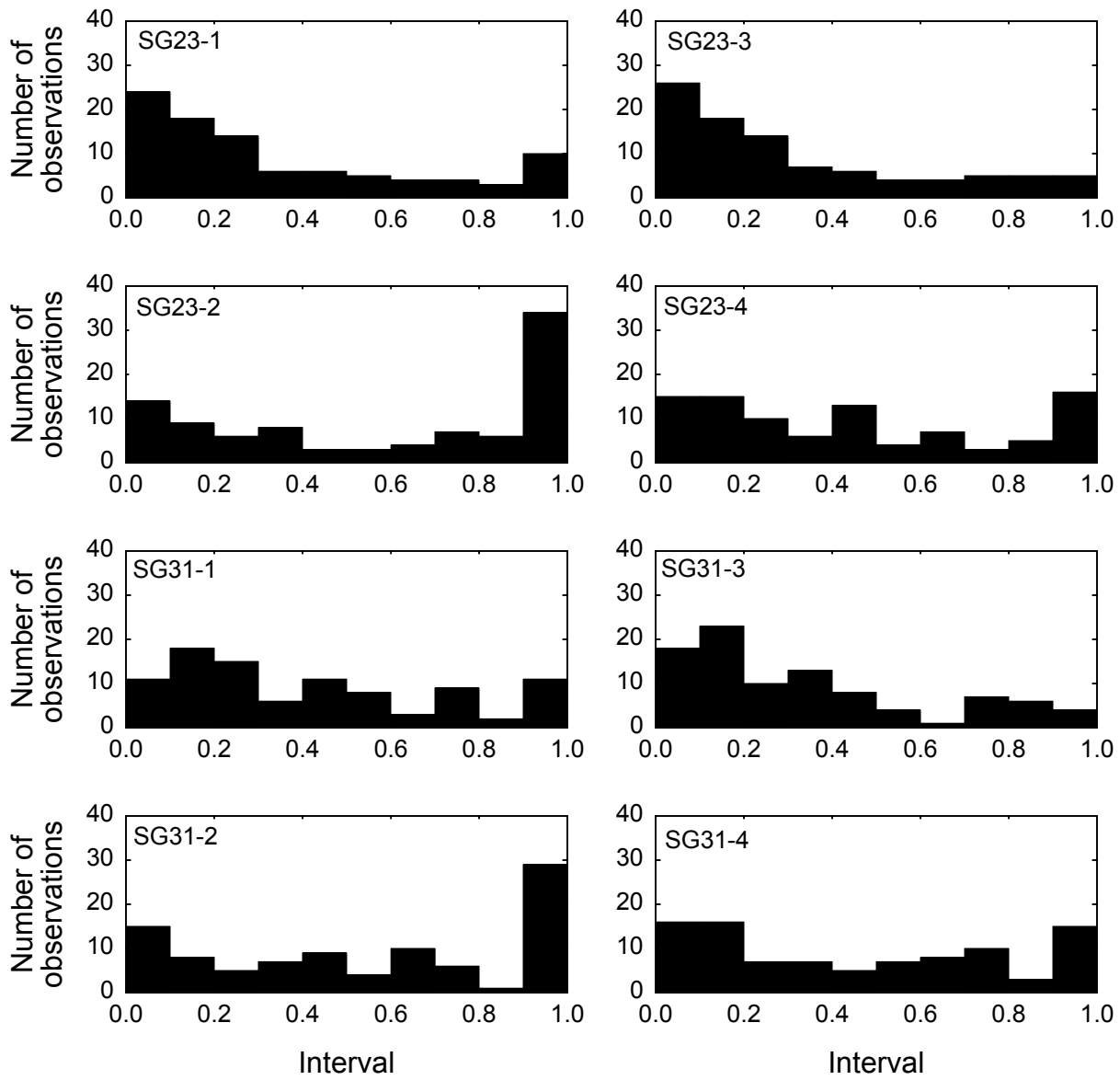
Figure 12. Histograms of sensor contribution to outlier distance for all events identified after construction.