

Phase Space Dissimilarity Measures for Structural Health Monitoring

August 2011

Prepared by
J. A. Bubacz
H. T. Chmielewski
A. J. Depersio
A. E. Pape
L. M. Hively
R. K. Abercrombie
S. D. Boone

DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via the U.S. Department of Energy (DOE) Information Bridge.

Web site <http://www.osti.gov/bridge>

Reports produced before January 1, 1996, may be purchased by members of the public from the following source.

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone 703-605-6000 (1-800-553-6847)
TDD 703-487-4639
Fax 703-605-6900
E-mail info@ntis.gov
Web site <http://www.ntis.gov/support/ordernowabout.htm>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE) representatives, and International Nuclear Information System (INIS) representatives from the following source.

Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831
Telephone 865-576-8401
Fax 865-576-5728
E-mail reports@osti.gov
Web site <http://www.osti.gov/contact.html>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Computational Sciences and Engineering Division

**PHASE SPACE DISSIMILARITY MEASURES FOR STRUCTURAL HEALTH
MONITORING**

J. A. Bubacz
H. T. Chmielewski
A. J. Depersio
A. E. Pape
L.M. Hively
R.K. Abercrombie
S. D. Boone

Date Published: August 2011

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6283
managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
Under contract DE-AC05-00OR22725

CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	v
ABSTRACT	vi
1. INTRODUCTION	1
2. BACKGROUND	3
2.1 FUNDAMENTAL CHARACTERISTICS	3
2.2 ANALYSIS AND PROCESSING CHARACTERISTICS	4
2.3 APPLICATIONS OF SHM	5
3. ANALYSIS APPROACH	5
3.1 OVERVIEW OF PSDM APPROACH	5
3.2 ANALYSIS APPROACH FOR DEMONSTRATING ALGORITHM SCALABILITY	8
3.3 ANALYTICAL APPROACH FOR TRIAXIAL VX. UNIAXIAL DATA	8
3.4 ANALYTICAL APPROACH FOR STATISTICAL VALIDATION	9
4. DATA	10
5. RESULTS	14
5.1 RESULTS FOR ALGORITHM SCALABILITY TEST	14
5.2 RESULTS FOR SENSOR PLACEMENT	14
5.3 RESULTS FOR MULTIAXIAL OPTIMIZATION	15
5.4 RESULTS FOR STATISTICAL VALIDATION	20
6. DISCUSSION	22
7. CONCLUSIONS AND FUTURE WORK	23
8. REFERENCES	25
APPENDIX A. SENSING AND NETWORK CHARACTERISTICS	A-3
APPENDIX B. EXPERIMENTAL SHM SYSTEMS	B-3
APPENDIX C. TECHNICAL DETAILS	C-3
APPENDIX D. TESTING THE ALGORITHM	D-3
APPENDIX E. SHM RESEARCHERS CONTACTED FOR DATA	E-3

LIST OF FIGURES

Figure	Page
1. Bridge deficiencies by count	1
2. Number of bridges in 2006 vs age	2
3. Z24 bridge deck instrumentation with sensors chosen for analysis circled in yellow	11
4. Z24 bridge pier instrumentation with sensors chosen for analysis circled in red	12
5. Two elevation views and a plan view illustrating pier settlement on Z24 bridge.....	13
6. Algorithm output vs number of processors and computation time	14
7. PSDM vs pier offset level for sensor 229 (Channel 3)	16
8. PSDM vs pier offset level for sensor 233 (Channel 4)	17
9. PSDM vs pier offset level for uniaxial data, sensor 229 (Channel 3).....	18
10. PSDM vs pier offset level for uniaxial data, sensor 233 (Channel 4)	19
D.1. Probability density function and cumulative probability function shapes	D-5
D.2. Cumulative frequency of actual data (series 2 vs theoretical exponential distribution (series 1)).....	D-7

LIST OF TABLES

Table	Page
1. Longest Battery Lives for WSN as Reported in Literature	4
2. Triaxial Damage Detection Strength of Each Sensor Location at Each Damage Level	20
3. Triaxial Extent-of-Damage prediction Capabilities of PSDM and at Each Sensor Location	21
4. Uniaxial Damage Detection Strength of Each Sensor Location at Each Damage Level	22
5. Uniaxial Extent-of-Damage Prediction Capabilities of Each PSDM and at Each Sensor Location	22

ABSTRACT

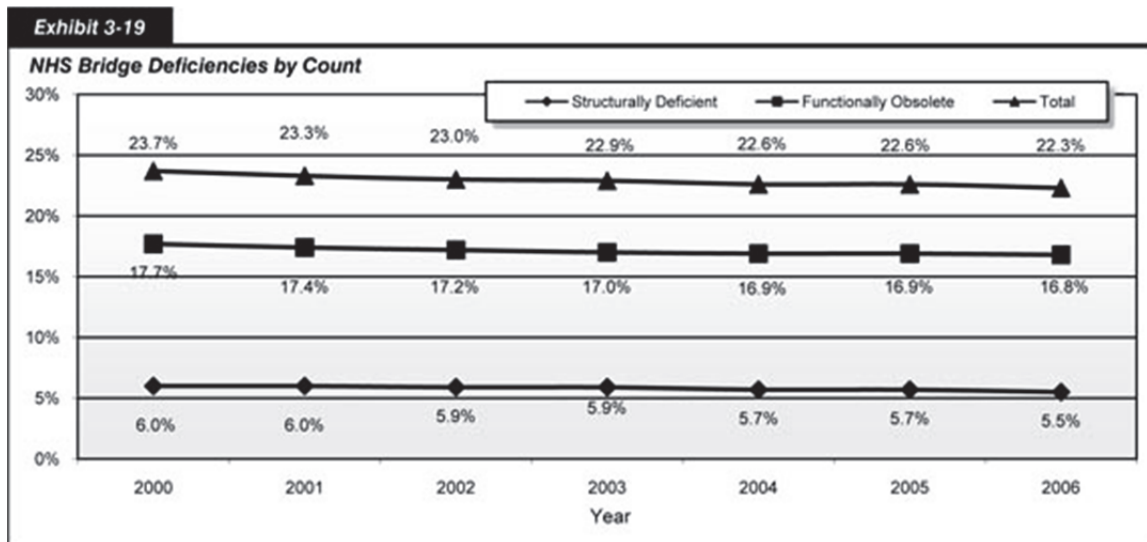
A novel method for structural health monitoring (SHM), known as the Phase Space Dissimilarity Measures (PSDM) approach, is proposed and developed. The patented PSDM approach has already been developed and demonstrated for a variety of equipment and biomedical applications. Here, SHM of bridges was investigated via analysis of time serial accelerometer measurements. This work has four aspects. The first is algorithm scalability, which was found to scale linearly from one processing core to four cores. Second, the same data were analyzed to determine how the use of the PSDM approach affects sensor placement. It was found that a relatively low-density placement sufficiently captures the dynamics of the structure. Third, the same data were analyzed by unique combinations of accelerometer axes (vertical, longitudinal, and lateral with respect to the bridge) to determine how the choice of axes affects the analysis. The vertical axis was found to provide satisfactory SHM data. Fourth, statistical methods were investigated to validate the PSDM approach for this application, yielding statistically significant results.

1. INTRODUCTION

The ability to provide accurate and timely forewarning of failure in complex systems would yield outstanding benefits toward improving the safety of the public while lowering maintenance costs and repair time. Typical failures in critical-infrastructure include random events, design flaws, and instabilities triggered by physical and/or cyber-attacks (Abercrombie et al. 2011). The aging infrastructure of the United States, specifically the worsening condition of America’s bridges, and the improvement of a wireless sensor network for structural health monitoring (SHM) are of particular interest in this report.

According to the latest American Society of Civil Engineers (ASCE) infrastructure report card completed in 2009, over 26% of the United States bridges are either structurally deficient or functionally obsolete, and this statistic continues to grow. Seventeen billion dollars are required for notable improvement while only \$10.5 billion is currently spent on the construction and maintenance of bridges (ASCE 2009). Although the term ‘structurally deficient’ does not immediately declare a structure as unsafe, it does indicate that the bridge’s condition is deteriorating and will require more attention, repair, and eventually complete rehab or replacement. Figure 1 presents the Federal Highway Administration’s graphical display of bridge deficiencies over the past six years. If the nation’s bridges were divided into subsets by age, one can conclude that the older the subset, the larger the percentage of deficient and obsolete bridges. Figure 2 displays bridges in this division by age. The point of this explanation is that the number of bridges in the United States requiring repair and the severity of damage will only increase with time resulting in the escalation of safety risks and maintenance costs.

Further reinforcement for the need of dependable SHM systems is the unreliability of the current inspection policies in place today, involving visual inspection and non-destructive evaluation methods (ultrasonic/acoustic wave propagation, thermal imaging, x-ray radiography, eddy-current methods). One FHWA report, *The reliability of visual inspections for highway bridges*, indicates that ratings assigned to the same bridge by several different inspection teams are variable (Moore et al. 2001). For example, among 49 inspectors from 25 states conducting a routine inspection, an average of between four and five different Condition Ratings (on a ten-point scale) were assigned to the same primary bridge element (Moore et al. 2001, p. 453). Further, this report indicates that even in-depth inspection (as opposed to routine inspection), can often fail to detect small weld cracks. A tradeoff clearly exists between replacement of aging infrastructure and maintaining that infrastructure with the aid of SHM.



Source: National Bridge Inventory.

Fig. 1. Bridge deficiencies by count. (<http://www.fhwa.dot.gov/policy/2008cpr/chap3.htm#6>)

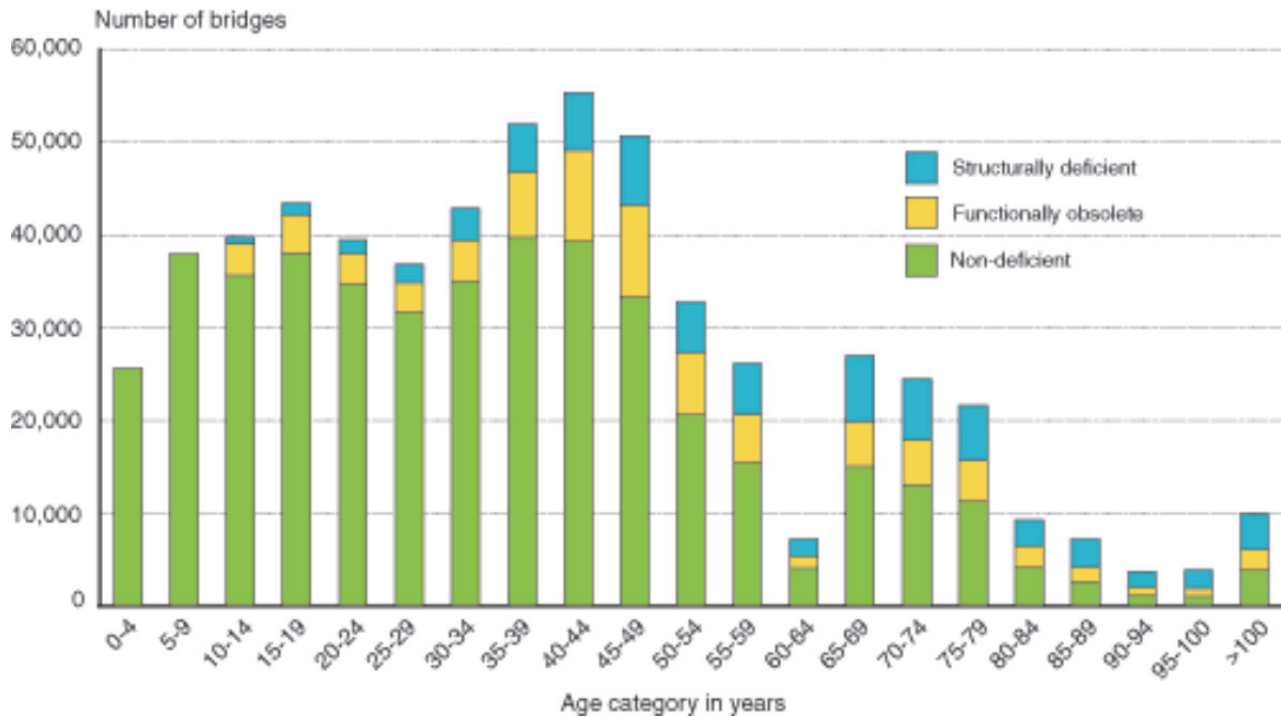


Fig. 2. Number of bridges in 2006 vs age [as reported on Memmot (2007)].
http://www.bts.gov/publications/bts_special_report/2007_09_19/html/figure_03_table.html

The final goal is to construct a wireless sensor network (WSN) that will be inexpensive to manufacture, easy to install, and long lasting which will provide better results than traditional nondestructive evaluation and visual inspection. Hopefully, these goals will be achieved through the use of the Phase Space Dissimilarity Measures (PSDM), which are discussed in detail in the Methods section. The PSDM approach was developed at Oak Ridge National Laboratory (ORNL) and is related to a number of patents: Clapp and Hively (1997), Hively et al. (1998), Hively and Ng (1998), Hively et al. (1999), Welch et al. (2002), Hively et al. (2002), Hively (2006), Hively (2007), Hively (2010), Hively (2009), Abercrombie and Hively (2010). The details of this method are discussed in Sect. 3.

One major roadblock in the development of a failure forewarning algorithm is the procurement of satisfactory data. Avoidance of ‘garbage in-garbage out’ becomes a strict policy when completing data quality checks prior to calculation. The difficulty of obtaining proper data was further increased for civil engineering applications because small-scale, inexpensive models are inadequate in representing the complex, non-linear dynamics of a real structure. Moreover, data acquisition from a full-scale structure from a healthy state to failure was a near impossibility.

This work focuses on four issues. First, a test of algorithm scalability was conducted to demonstrate the potential for improving data processing and storage efficiency within a wireless sensing network. Second, the same data were analyzed to determine how the use of the PSDM approach affects sensor placement. This was an important consideration, as the cost of a WSN is a per-sensor cost. Third, the same data were analyzed using only the vertical component of acceleration and the results were compared to those from the analysis of all three axes. This analysis is of interest because fewer axes, and thus fewer channels, reduce the amount of computation required. Fourth, real data from a real bridge damage test were analyzed and statistical methods are investigated for the purpose of validating the PSDM method for this application.

The work is presented as follows. In Sect. 2, characteristics of and previous work on SHM via WSN is presented. In Sect. 3, the PSDM approach is summarized. In Sect. 4, the source and nature of the data

used for the analysis is discussed. In Sect. 5, results of our analysis are presented. In Sect. 6, the content and quality of the results are discussed. In Sect. 7, conclusions are drawn from those results and future work is outlined.

2. BACKGROUND

The purpose of SHM is to identify and characterize damage to structures such as aircraft, buildings, and bridges (Farrar and Worden, 2007). SHM by means of wireless WSN has gained a great deal of attention in recent years and is of particular interest. The following background serves as an illustrative summary of SHM research to date. The characteristics and applications of SHM through WSN with examples of specific implementations are outlined.

2.1 FUNDAMENTAL CHARACTERISTICS

Six distinct fundamental characteristics of all SHM systems were identified. First and foremost is cost, both in operations and maintenance. WSN are intended to be an improvement over traditional wired systems in this respect. Though relatively few technical papers give specific cost figures, those that do indicate per-sensor costs are typically \$100–600.

The second characteristic is whether the system is active or passive. A passive system is one that senses and analyzes only the behavior of the structure under whatever conditions the structure is subjected, whereas an active system is one that can also impose a signal by means of some actuation device. According to Lynch and Loh (2006), piezoelectric sensors or ultrasonic transducers can be used to provide the input signal.

Third, physical scalability is the ability of a system paradigm to be efficiently and effectively applied to a wide range of structures (such as bridges) of different sizes, and also to a multitude of structures across a geographic region, while maintaining a high level of usability. This is not to be confused with computational scalability, which is discussed below. Physical scalability considerations affect the network and communication (intra-structure and remote) design. This effect is especially true of the routing protocol, because some systems rely on a peer-to-peer network topology to pass data from one node to the next until it reaches a central node.

The fourth characteristic is maintainability, typically involving physical maintenance and data management. Ideally, a system would operate autonomously for an indefinite period of time, only requiring human intervention for important decision-making. In reality, some degree of maintenance is required. One of the most significant maintenance needs is battery replacement, since most systems are battery-powered. Table 1 gives a summary of the longest battery lives reported in the literature.

The fifth characteristic is adaptability to new paradigms. Recent examples are impedance monitoring (Overly et al. 2008) and mobile systems (Zhu et al. n. d.; Taylor et al. 2009). Adaptability of a SHM system allows the use of new sensors or data processing.

The sixth characteristic is how adaptable the system is for other structural applications. As discussed below, SHM is used in many different engineering applications, and it is desirable to have a system that, once successful in one area (such as bridge monitoring) can be easily adapted for use in another area (such as aircraft monitoring).

A significant number of other characteristics pertaining to sensing and networking can be found in Appendix A.

Table 1. Longest Battery Lives for WSN as Reported in Literature

Researcher(s)	System	Data Type	Battery Life
Harms et al., (2009a, 2009b)	SmartBrick	Acceleration, temperature, tilt, and strain	5 years (projected)
Galbreath et al. (2003)	[unnamed]	Strain	5 years (projected)
Overly et al., (2008)	WID 2.0	Impedance	5 years (projected)
Bilbao et al. (2011)	[unnamed]	Acceleration	3 years (projected)

2.2 ANALYSIS AND PROCESSING CHARACTERISTICS

The vibration-based damage detection for civil engineering applications began in the 1980s. Typical approaches are resonant frequencies, mode shape vectors, and quantities derived from these parameters (Farrar and Worden 2007; Farrar et al. 2007). Damage identification is commonly divided into four categories of increasing complexity: detection; localization; assessment; and prognosis (time estimate to failure). Of these four, there has been increasing success in the determination of the first three levels. One method in particular uses a data-driven approach through machine learning theory (Worden 2006). Prognosis is notably harder and may only be conceivable as a statistical theory. This difficulty arises, because knowledge of material properties, structural geometry, and future loading are required but very difficult to obtain with a high level of certainty (Worden 2006).

When expressing the increased difficulty of damage detection/failure forewarning in civil structures, it is often useful to divide SHM algorithms into supervised and unsupervised classifications. Supervised assessment requires data for undamaged and damaged states. Unsupervised learning has no data for damage/failure states. Because large structures are one-of-a kind and highly expensive, unsupervised learning is often used (Farrar and Worden, 2007). However, many researchers have used experimental structures to develop and test models for use on other structures.

Many damage detection algorithms use Fourier analysis to extract modal properties or damage-sensitive features for change detection. Fan (2010) provides a comprehensive comparative study of modal parameter-based damage methods. The vibrational characteristics of damaged structures are highly time-dependent on failure progression. Numerous models (e.g., time-frequency and time-scale analysis) have been developed to account for damage dynamics. Fast Fourier transforms and wavelet analysis have been used to analyze structural dynamics to improve damage detection in SHM applications (Staszewski 2007).

Real structural damage (crack growth, loose connections, delamination, etc.) typically produces nonlinear effects, while many of the above methods model this damage linearly. Linear models are inadequate for capturing the nonlinear dynamics of a complex structure, making accurate representation of damage and boundary conditions impossible (Farrar 2007). Even worse, the modal properties of a large structure are more sensitive to temperature and loading variations, than to damage. This sensitivity causes an excess of false positives (Worden 2006). Discrimination between environmental and damage effects via modal frequencies has been demonstrated by covariance-driven stochastic subspace identification with ‘fuzzy clustering’ algorithm (Carden 2008).

Progress in nonlinear and chaotic dynamics has yielded three indicators of damage detection: (i) the first minimum in the mutual information function as a measure of decorrelation time; (ii) the correlation dimension as a measure of complexity; and (iii) the Kolmogorov entropy as a measure of predictability (Protopopescu and Hively 2005). These traditional nonlinear measures (TNM) are successful in capturing the nonlinear dynamical features of a complex system, but lack the necessary sensitivity in detecting dynamical change that is required to give a forewarning of failure (Protopopescu and Hively 2005). A number of researchers have investigated techniques for SHM similar to the one presented here, namely the use of phase-space based analysis. Trendafilova (2003) discusses the analysis of the probability density function of the attractor. Moniz et al. (2005) and Overbey and Todd (2008) discuss the phase-space based use of data from multiple sensors for SHM. The U.S. Naval Research Laboratory has been successful in detecting damage on model offshore structures via stiffness reduction (Nichols 2003).

2.3 APPLICATIONS OF SHM

A wide variety of applications exist for SHM with WSN in general, and for this algorithm in particular. These can roughly be divided into two groups: mechanical structures and civil structures. Mechanical structure applications include aircraft, spacecraft, vehicles, machinery, and marine vessels. The algorithm discussed in this paper has already been successfully applied to data from a helicopter gearbox failure (Hively 2009). Boller and Buderath (2007) discuss SHM for aircraft, and Lynch and Loh (2006) give a thorough review of work on that application. SHM can also be applied to marine vessels. Swartz et al. (2010) present the implementation of a structural health monitoring system in an FSF-1 Sea Fighter, a relatively small naval vessel with an aluminum catamaran hull. Perez et al. (2010) discuss general implementation of SHM in the U.S. Navy.

However, civil structures are the focus of this report. Civil applications include pipelines, dams, offshore structures, communication towers, buildings, and bridges. Stoianov et al. (2007) present a WSN for SHM of pipelines, called “PIPENET.” Saftner et al. (2008) discuss SHM in geotechnical work, including monitoring of underground pipelines. Brownjohn (2007) covers SHM of dams in the UK. Nichols (2003) discusses passive SHM for offshore structures using a phase-space approach, similar to the approach presented in this paper. Elshafey et al. (2010) also discuss SHM for offshore structures, though from a physics-based approach rather than a phase-space based approach. Lam et al. (2011) investigates damage detection for electrical transmission towers.

SHM of buildings is an application that is inherently more complicated than many other applications, due to the uniqueness and importance of the structure being monitored. SHM of historic buildings was investigated by researchers in Rome, Italy (Abruzzese et al. 2009). Wu et al. (2010) present the monitoring of a historic tower in Trento, Italy, which provided valuable information about the structure and its behavior. Rainieri et al. (2010) discuss integration of SHM systems with seismic early warning systems (SEWS). This concept is of special interest because the algorithm presented in this paper is potentially applicable to both types of monitoring, which would make the integration of those two systems much easier than with other algorithms. Rainieri et al. (2010) further discuss the possibility of using a city-wide system not only for monitoring and forewarning, but also for emergency management after a seismic event. Fabbrocino et al. (2009) investigate SHM for retaining walls by means of instrumenting piles. Duron (2003) presents extensive testing of a system to forewarn of the collapse burning buildings for firefighter safety.

Lin et al. (2005) and Bahar et al. (2010), among others, discuss integration of semi-active SHM systems with magnetorheological (MR) dampers. MR dampers are ones whose damping fluid has a viscosity that is influenced by the presence of a magnetic field. By dynamically changing the magnetic field (based on real-time SHM data), the stiffness of the damper, and consequently the dynamic behavior of the structure of which it is a part, can be controlled.

A list of bridge SHM experimental systems can be found in Appendix B.

3. ANALYSIS APPROACH

3.1 OVERVIEW OF PSDM APPROACH

ORNL has developed a method using phase-space dissimilarity measures that are able to “quantify condition change from complex, time-serial data” (Protopopescu 2005). One goal of the present work is development of this algorithm for SHM. The algorithm has been successful in failure forewarning in a tactical quiet generator (Hively 2008), a helicopter gearbox (Hively 2009), and epileptic seizure episodes (Hively 2005).

Previous work (Hively 2009) determined that vibration power is an excellent indicator of damage.

Consequently, this work focuses on instantaneous vibration power, which is proportional to

$$\mathbf{a} \cdot \int \mathbf{a} dt, \quad (1)$$

where a is the acceleration vector. This choice is certainly not unique. For example, one component of acceleration might be adequate. The advantage of vibration power is a single channel of data (rather than several channels) with all of the three-dimensional dynamics and with a proportionate reduction in computational effort to quantify the condition change. The data, e , are sampled at equal time intervals, τ , starting at an initial time, t_0 , yielding a time-serial set of N points in each cutset, $e_i = e(t_0 + i\tau)$.

Conversion of the triaxial acceleration data to vibration power uses an artifact-filter to eliminate the secular term. We remove this artifact with a novel zero-phase quadratic filter, which retains the nonlinear amplitude and phase relationships (Hively et al. 1995). This filter uses a moving window of $2w + 1$ points of e_i -data with the same number of data points, w , on either side of a central point. We fit the e_i -data to a quadratic equation, $F(t) = a_1 T^2 + a_2 T + a_3$. The notation here is $T = t - t_c$, with t_c as the time at the central point of the moving window. The fit minimizes $\Psi = \sum_i [F(t) - e_i]^2$, where the sum is over the moving window of $2w + 1$ points. The minimum in Ψ occurs for $\partial\Psi/\partial a_k = 0$, which yields three linear equations in three unknowns. The window-averaged artifact is given by the fitted value at the central point, $F(t_c = t_i) = a_3$. The sums over odd powers of T_i are zero. Symmetric sums over even powers of T_i for $-w \leq i \leq w$ become sums from 1 to w , which can be evaluated explicitly by standard formulae. The window-averaged solution for the artifact signal at the central point then is:

$$f_i = F(t = t_c) = \frac{3(3n^2 + 3n - 1)(\sum_i i^2 e_{i+c}) - 15(\sum_i i^2 e_{i+c})}{(4n^2 + 4n - 3)(2n + 1)}. \quad (2)$$

The sums in Eq. (2) are over $-w \leq i \leq w$. The computations for this equation can be reduced further by one initial sum with $c = w + 1$, and recursions thereafter (Hively et al. 1995). Application of this filter to the N -point analysis window of e_i -data yields $N - 2w$ points of artifact data, f_i . The residual (artifact-filtered) signal, $g_i = e_i - f_i$, has essentially no low-frequency artifacts. We chose $w = 250$ to remove the slowly varying secular term, yielding 65,036 points for each damage state.

The g_i -data from the previous paragraph are symbolized into S discrete values, s_i , namely $0 \leq s_i \leq S - 1$. Equiprobable symbols can be formed by ordering data in the first baseline cutset from the smallest to largest value. The first N/S of these ordered values correspond to the first symbol, 0; data values $(N/S) + 1$ through $2N/S$ correspond to the second symbol, 1, and so on. Alternatively, uniform symbols use g_x and g_n , which are the maximum and minimum in the first baseline cutset, respectively:

$$s_i = INT \left[\frac{S(g_i - g_n)}{(g_x - g_n)} \right]. \quad (3)$$

Here, the INT function converts a decimal number to the closest lower integer. Equiprobable symbols are a poor choice [TQG 2008] for the present analysis, which uses uniform symbols.

The symbolized data can be converted into such a phase space representation by standard reconstruction of the dynamics via time-delay phase-space vectors (Eckmann and Ruelle 1985):

$$y(i) = [s_i, s_{i+\lambda}, \dots, s_{i+(d-1)\lambda}]. \quad (4)$$

Here, λ is the time delay, and d is the dimension of the phase space, which has S^d possible discrete states. Theorems by Takens (1981) and by Sauer *et al.* (1991) guarantee faithful topological reconstruction from time-serial observations in a sufficiently high-dimensional phase space to capture the

unique dynamical signature. Base- S arithmetic uniquely labels the i -th discrete phase-space state in Eq. (4) by $J(i)$:

$$J(i) = \sum_m s_{i+(m-1)\lambda} S^{m-1}. \quad (5)$$

Eq. (5), the sum is over the d components of the time-delay vector of Eq. (4), $1 \leq m \leq d$. A statistical distribution function is then constructed by counting the number of visitations to each discrete phase-space state, where Q_J and R_J denote the population of the J -th state for the baseline and test cases, respectively, which are then compared by dissimilarity measures, namely:

$$\chi_n^2 = \sum_J \frac{(Q_J - R_J)^2}{(Q_J + R_J)}, \quad (6)$$

$$L_n = \sum_J |Q_J - R_J|. \quad (7)$$

Here, J is the base- S label for the phase-space states from Eq. (5). The dynamical flow (Abarbanel 1996) is $y(i) \rightarrow y(i + \mu)$, and is represented by a connected-phase-space vector of the form, $Y(i) = [y(i), y(i + \mu)]$. Base- S arithmetic also uniquely labels the i -th connected-phase-space state by $K(i)$:

$$K = J(i) + J(i + \mu)S^d. \quad (8)$$

The corresponding connected-phase-space dissimilarity measures are χ_c^2 and L_c :

$$\chi_c^2 = \sum_K \frac{(Q_K - R_K)^2}{(Q_K + R_K)}, \quad (9)$$

$$L_c = \sum_K |Q_K - R_K|. \quad (10)$$

As before, K is the base- S label for the connected-phase-space states from Eq. (8). In this context, χ^2 in Eqs. (5) and (8) is not an unbiased statistic to accept or reject a null statistical hypothesis, but rather a *relative* measure of dissimilarity (Hively et al. 1999) to quantify changes in geometry and visitation frequency. These dissimilarity measures are based on summing the absolute value of differences, and give better discrimination than traditional nonlinear measures, which are based on a difference of averages.

The dissimilarity measures from Eqs. (5), (6), (8), and (9) are normalized to account for their disparate range and variability. The normalized form is:

$$U_i(V) = \frac{(V_i - \bar{V})}{\sigma}. \quad (11)$$

Previous analysis used an absolute value in Eq. (11). The mean of each dissimilarity measure, \bar{V} , is obtained by comparison among the $B(B-1)/2$ unique combinations of the B baseline cutsets, with a corresponding sample standard deviation, σ . Each contiguous, non-overlapping test cutset is subsequently compared to each of the B baseline cutsets to obtain the corresponding average dissimilarity, V_i , of the i -th analysis window for each dissimilarity measure. U_i is the number of standard deviations that the i -th test

case (unknown dynamics) deviates from the baseline (nominal-state). Limitations on the number of data samples for each damage state (65,036) imply that only a few baseline cutsets are appropriate ($B = 5$), yielding a cutset length of $N = 13,007$.

Success can be measured as an increase in the dissimilarity measures, proportional to the damage level. The analysis flow then proceeds as follows: (a) choose specific values for each of the parameters in the set, $\{d, S, \mu, \lambda\}$; and (b) search exhaustively over these parameters to find a linear increase in dissimilarity measures versus the damage state.

3.2 ANALYSIS APPROACH FOR DEMONSTRATING ALGORITHM SCALABILITY

In previous work with this algorithm (Protopopescu and Hively 2005), the authors noted that one practical barrier to the phase-space reconstruction methodology is the analyst-intensive process of selecting appropriate parameters during the data training phase. One avenue for making progress over previous applications of the algorithm is to demonstrate the code's scalability, or the effectiveness of the algorithm in processing more data as more hardware is made available to the network. One advantage of the necessity for multiple wireless sensors to monitor one bridge is the addition of processing units. If the algorithm code scales effectively, then the availability of these processing resources will make the training phase and subsequent computations in vibrational monitoring more time-efficient—one step in showing the applicability of the non-linear algorithm to (near) real-time vibrational bridge data.

In order to demonstrate the scalability of the failure forewarning algorithm, a virtual machine containing the code was uploaded to an ORNL cluster computer, with access to a cluster of processors where multiple instantiations of the forewarning algorithm could be run simultaneously, as well as the computing and storage capacity for an increasing volume of input data, (for more details about the implementation of the virtual machine, see Appendix C). On the ORNL cluster computer, the virtual machine and the code could be run with various configurations. Specifically, the scalability of the code could be tested by running the virtual machine with different numbers of processors dedicated to the code's execution. The idea being tested was twofold. The first test would demonstrate whether the FORTRAN algorithm would result in linear scalability: as the number of processors running the code was increased, the code output produced in the same execution time would increase linearly. The second test would demonstrate that regardless of the amount of data being analyzed, the program would still maintain its scalability.

The tests would be performed as follows. First, a single processor would be dedicated to the virtual machine. A preliminary run of the code would be executed, and the number of parameter sets computed would be recorded. The virtual machine would then be reconfigured to have access to successively more processors (1,2,3,4, etc.), and the process would be repeated. The code would be executed in this way until a clear trend emerged between the cumulative processing time across all CPUs and the amount of output. The second test would demonstrate the second part of scalability (the volume of data) by following essentially the same process, except this time, by increasing the amount of data given to the virtual machine along with the RAM. If the output produced still increased linearly with the amount of total processing time, regardless of the addition of multiple channels of data, then scalability would be effectively demonstrated.

The scalability of the algorithm, as determined by the above tests, is shown in the Results section. See Appendix C for more details on the test implementation and creating the virtual machine environment.

3.3 ANALYTICAL APPROACH FOR TRIAXIAL VS UNIAXIAL DATA

In the process of choosing what axes of accelerometer data to analyze, the smallest combination of accelerometer axes that yield satisfactory data for PSDM analysis is sought. The axes are vertical, longitudinal, and lateral with respect to the bridge. Since each of these three axes can be either on or off

(that is, being analyzed or not), there are $2^3 = 8$ different combinations. However, the case in which all three are off is of no interest, so there are effectively seven combinations to be considered. For this phase of the project, we wish to compare the results of analysis of all three axes of accelerometer data to that of solely the vertical axis of each sensor.

3.4 ANALYTICAL APPROACH FOR STATISTICAL VALIDATION

ORNL's nonlinear algorithm is capable of two different modes of detecting phase-space changes that are useful in SHM: the damage detection mode and the failure forewarning mode. The damage detection mode used in the analyses performed in this paper is based on stationary vibrational data collected at distinct damage states. This mode has been used in previous analyses by the author to detect damage in a helicopter gearbox with seeded faults (Protopopescu and Hively 2005). The failure forewarning mode is used on continuous, non-stationary data. This mode has been used by the authors in previous applications, including predicting failure in a tactical quiet generator (Hively 2008) and a helicopter gearbox (Protopopescu and Hively 2005; Hively 2009), and forewarning epileptic seizure events from EEG data (Hively et al. 2005). The statistical test methodologies for each mode of the algorithm are summarized below.

The two analyses conducted in this paper are meant to show the implications of the algorithm's effectiveness in relation to: 1) sensor placement; and 2) multiaxial optimization. In both analyses, the presence and severity of damage are detected by comparing the phase-space dissimilarity measures at different damage levels to those found within the different cutsets at the undamaged state. For the sensor placement problem, a series of one-sided two-sample t-tests of unequal variance is conducted to determine the effectiveness of each sensor in detecting damage of each level with each dissimilarity measure. The two-sample t-test is a way to compare the means of two independent samples. The assumption of unequal variance is made due to the observation that the PSDM results at each of the damage levels has a much higher standard deviation than the PSDMs in the undamaged state. In each of the tests in the series, the mean and standard deviation of each sample are computed and tested against a null hypothesis of zero (i.e. there is "no difference" between the two samples). For every dissimilarity measure at each damage level for every sensor location, the five PSDM values for the undamaged and damaged sample are compared in succession. Sample sizes greater than five would produce more statistically conclusive results; however, because there was only a limited amount of data, and reducing the length of the cutsets to obtain more PSDM values per damage level would result in a sacrifice in PSDM precision, the sample sizes in this analysis must remain small. Such a test demonstrates whether the difference in means of the pre-damage and post-damage PSDMs obtained from the available cutsets is centered about zero.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}, \quad (12)$$

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2 + s_2^2}{n}}, \quad (13)$$

$$H_0: \mu_2 - \mu_1 = 0, \quad (14)$$

$$H_1: \mu_2 - \mu_1 > 0. \quad (15)$$

\bar{X}_1 and \bar{X}_2 are the means of each combination of damaged and undamaged PSDM values. The p-values are recorded in the tables in the Results section. Four different dissimilarity measures and four damage levels at each of five sensors produce a total of 80 hypothesis tests for each analysis.

A second table contains the results of another series of tests designed to determine the algorithm's extent-of-damage quantification capabilities. The tests are performed by constructing a prediction interval from the linear regression line calculated from each PSDM at each sensor. Since quantifying extent-of-damage would use four PSDM inputs (on the y-axis) to predict damage level (on the x-axis), for each PSDM plot the standard deviation of the x-axis residuals was used to construct a prediction interval for each point along the regression line. These prediction intervals were used to determine the damage level prediction sensitivity (true positives divided by total damaged cases) and specificity (true negatives divided by total undamaged cases) of each PSDM at each sensor.

$$x_{min,max} = x_i \pm t_{\alpha} \cdot s_{X_1} \cdot \sqrt{1 + \frac{1}{n}}. \quad (16)$$

In the multi-axial optimization analysis, an analogous series of tests is performed. The purpose of such an analysis would be to determine the minimum number of axes required to obtain the same damage detection rate. In this analysis, two similar tables are filled with the results of the same analysis as above, conducted with only the data from the vertical axis at each sensor. These results can be compared to the results from the triaxial data analysis, to determine whether the uniaxial analysis returns inferior, equal, or superior damage detection results (equal or greater numbers of significant test results within each cell of the results table).

For the results of the statistical tests outlined above, please refer to Sect. 5, Results of the Statistical Validation. For a review of the use of statistics techniques and statistical validation in vibration-based SHM, as well as a description of the statistical tests and procedures utilized in this paper, please see Appendix D.

4. DATA

An extensive search for data was conducted on which to test and refine the PSDM approach for bridge structures. See Appendix E for a summary of the contacts. Satisfactory data from a damage test of the Z24 bridge in Switzerland was found.

The 3-span bridge was constructed in 1961–1963 and the spans measure 14, 30, and 14 m. Condition prior to damage testing was qualified as 'relatively good' with the only visible damage being exposed rebar in a few places. Accelerometers recorded bridge dynamics for ambient and forced vibration tests (AVT and FVT). During the FVT series, the bridge was excited by two mechanical shakers: one 20 kN dead weight (max 10 kN dynamic load); and one 25 kN dead weight (max 20kN dynamic load). The shakers were driven with uncorrelated random noise, band-limited to 3–30Hz while nine setups of up to 15 sensors each captured 65,536 samples with a sampling frequency of 100 Hz. Varying damage states were imposed on the structure including: (i) settlement of a column; (ii) tilting foundation by lowering one column; (iii) spalling of reinforcement concrete ; (iv) failure of tendon anchor; (v) failure of tendon wires, involving a 35% loss pre-stress force; (vi) landslide; and (vii) failure of concrete hinges at abutment piers (Krämer and de Smet, 1999). Figure 5 shows the plan and elevation view of the bridge, with the pier settlement indicated.

Figures 3–4 show nine arrays of accelerometers that were placed along the length of the bridge. Each array had a 3-dimensional set of $5 \times 3 \times 4$ accelerometers also as shown in Figures 3–4. For the analysis in this report, five triaxial accelerometers were strategically chosen and their FVT data files were concatenated for the progressive settlement of a column (i) from 0 to 20, 40, 80, and 95 mm. Then, for each of the individual sensors, we created seven combinations of the axes to analyze by isolating each component of the triaxial data with the seven combinations being vertical (V), transverse (T), longitudinal

(L), VL, VT, LT, and VLT. As explicitly demonstrated in the ‘Methodology’ section, the algorithm takes the acceleration vectors and converts them to power, a scalar value, before calculating PSDM. This approach allows analysis with minimal changes to the original software for each test case.

The analysis involves comparison of damage levels with the PSDM from uni, bi, and triaxial accelerometer data. The objective is how well PSDM correlate to damage levels by sensor location, and by the number and choice of the accelerometer axes. We expect uniaxial accelerometer data in the vertical direction to suffice, because the amplitude of longitudinal and transverse bridge motion is typically much smaller than vertical deflection.

Data quality analysis verifies important features in the data: proper number of data points; any intervals with unchanged amplitude; saturation at high or low limits as an indicator of improper data scaling; consistent amplitude across datasets in the test sequence; adequate sampling rate; excessive periodic content; and excessive noise. An adequate sampling rate should span the de-correlation time with a sufficient number of time samples. The decorrelation-time measures are the first minimum in the mutual information function (<4 time steps) and the first zero in the autocorrelation function (<4 time steps). Excessive periodicity obscures the underlying nonlinear dynamics and has more than 50% of the total area under the two largest peaks in the Fourier spectrum vs frequency. Excessive noise obliterates the useful information with disorderly signal values. Consequently, a measure of order in the signal (Shannon entropy, E) vs the number of uniform data symbols (S) allows determination of the average number of bits of information (b) in the data as the maximum in E vs $S = 2b$; less than five bits of information corresponds to excessive noise. The garbage-in-garbage-out syndrome is avoided by rejection of data that fails one or more of these tests (Hively 2008).

The sensors chosen for analysis are 213, 223, 229, 233, and 541, all of which are triaxial. These numbers correspond to Figs. 3–4 below. Sensors 229 and 233 were selected because they were nearest to the lowered column. Sensor 223 was one of the closest sensors to bridge mid-span; 213 was chosen to study the underlying dynamics at the opposite column. Finally, sensor 541 was unique because of its location on the damaged column rather than the bridge deck.

See Fritzen and Kraemer (2009) for a discussion of two approaches to the same damage case on the Z24 Bridge, referred to as the Inverse Eigensensitivity Method and the Modal Kinetic Energy Method.



Fig. 3. Z24 bridge deck instrumentation with sensors chosen for analysis circled in yellow. (Krämer and de Smet 1999)

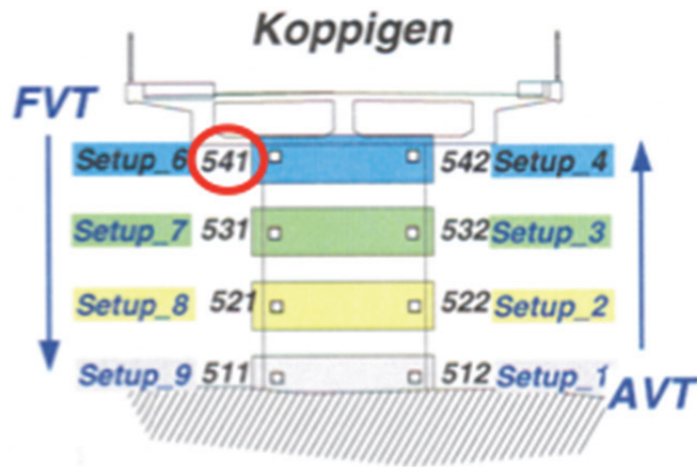


Fig. 4. Z24 bridge pier instrumentation with sensor chosen for analysis circled in red. (Krämer and de Smet 1999)

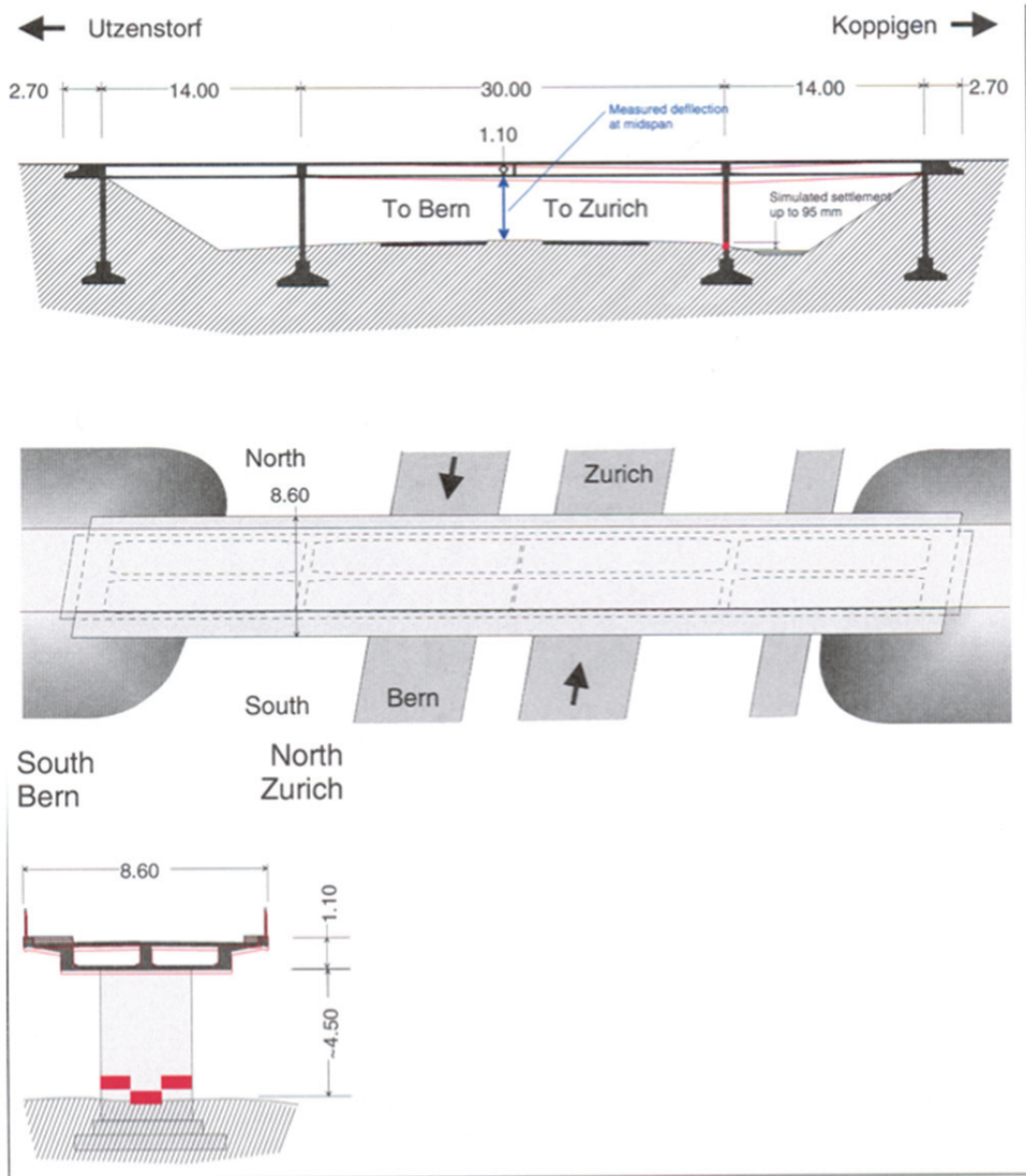


Fig. 5. Two elevation views and a plan view illustrating pier settlement on Z24 bridge. (Krämer and de Smet 1999)

5. RESULTS

5.1 RESULTS FOR ALGORITHM SCALABILITY TEST

The results of the scalability test on the ORNL cluster computer virtual machine are shown in Fig. 6 below.

Computation output is measured by the number of parameter sets computed. Computational power is measured by the number of processors, or equivalently, the total computational time. This test shows a linear increase in computational output vs the number of processors. Thus, the computational output scales with the number of processors, as expected. In the application of wireless sensors on a bridge, this type of scalability will be a necessary feature of the structural health algorithm utilized in the network, as each sensor will need to be a processing center capable of analyzing and managing its own data.

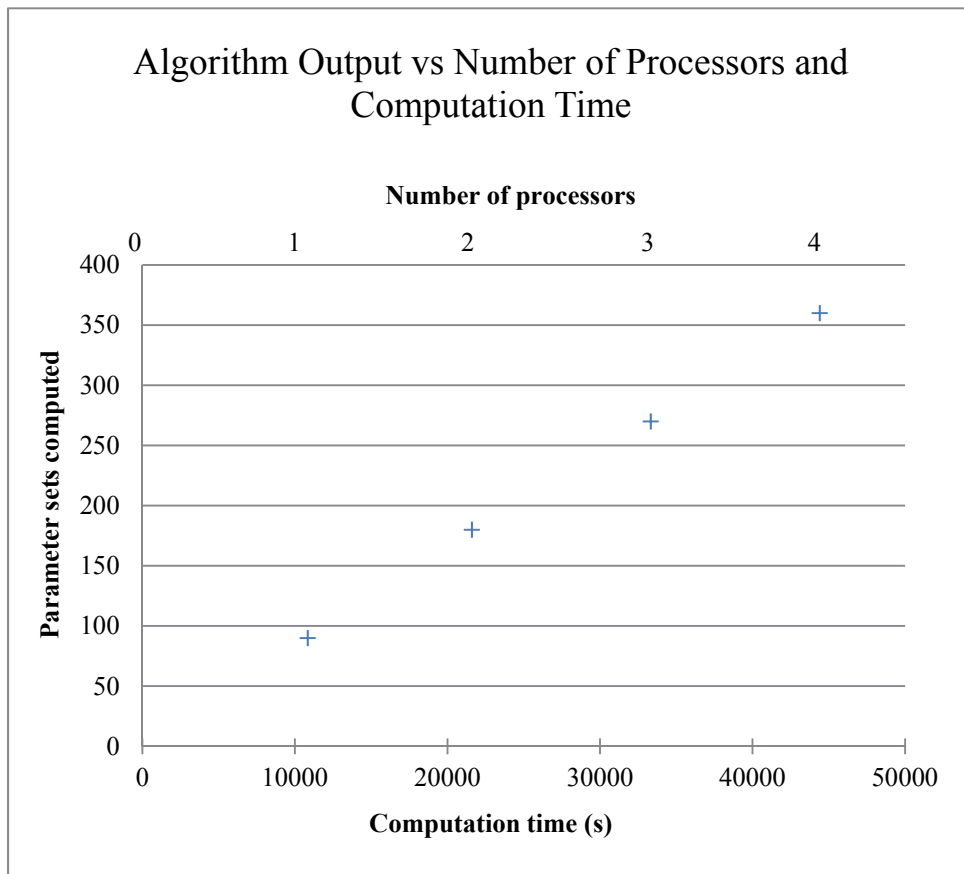


Fig. 6. Algorithm output vs number of processors and computation time.

5.2 RESULTS FOR SENSOR PLACEMENT

Figures 7–8 below present some of the results of the damage detection algorithm. Each figure represents one channel (or sensor location), and contains four plots—one for each PSDM. The five means with error bars correspond to the PSDMs calculated from each of the five data cutsets at each damage level. Here, “damage level” refers to the offset (in millimeters), corresponding to the amount of pier settlement.

The detection of damage by each sensor can be evaluated by comparing the PSDMs calculated at each offset level to those at the undamaged state (with zero offset). This is exactly what is described in the series of statistical tests whose results are shown in the tables in the following sections. The strongest damage detection (the changes of the greatest magnitudes, compared with the mean and standard deviation at the undamaged level) seems to occur in channels 3 (sensor 229), 4 (sensor 233), and to some extent 5 (sensor 541). Even if the positive linear correlation is not strong, however, the algorithm may still detect damage through changes in the PSDMs from those detected in the undamaged state.

The next level of SHM involves detecting the extent of damage. The linear regressions superimposed on the plots below serve as a way to judge the strength of the algorithm in predicting damage level from the PSDMs calculated. The second table in the Results for Statistical Validation section below illustrates the effectiveness of the algorithm in predicting damage level at each sensor location. As it can be seen in the plots and from inspecting the resultant tables, channels 3 and 4 (sensors 229 and 233—corresponding to the deck sensors closest to the damaged pier) show the most consistent linear correlation for the purposes of PSDM prediction, with the highest sensitivity and specificity levels, as shown in the tables below.

The location of damage has not been pinpointed by the algorithm, although the largest PSDM changes do occur at the local deck sensors nearest the pier, as might be expected. The fact that damage is detected by all PSDMs in each location does suggest that a more minimalistic sensor configuration could have been used, at least in this case. Further analyses to consider more types of damage, which are beyond the scope of the current paper, would be needed in order to determine an optimal sensor placement configuration.

5.3 RESULTS FOR MULTIAXIAL OPTIMIZATION

Figures 9–10 represent the PSDMs produced from just one dimension of accelerometer data—the vertical direction. Again, channels 3 and 4 produced the best linear correlations, and their plots are presented here. In the majority of cases, the magnitude of the PSDMs in the uniaxial analysis are less than the magnitude of the PSDMs produced from the triaxial data. However, the regression line is evidently a more accurate fit in the uniaxial analyses than in the triaxial analyses.

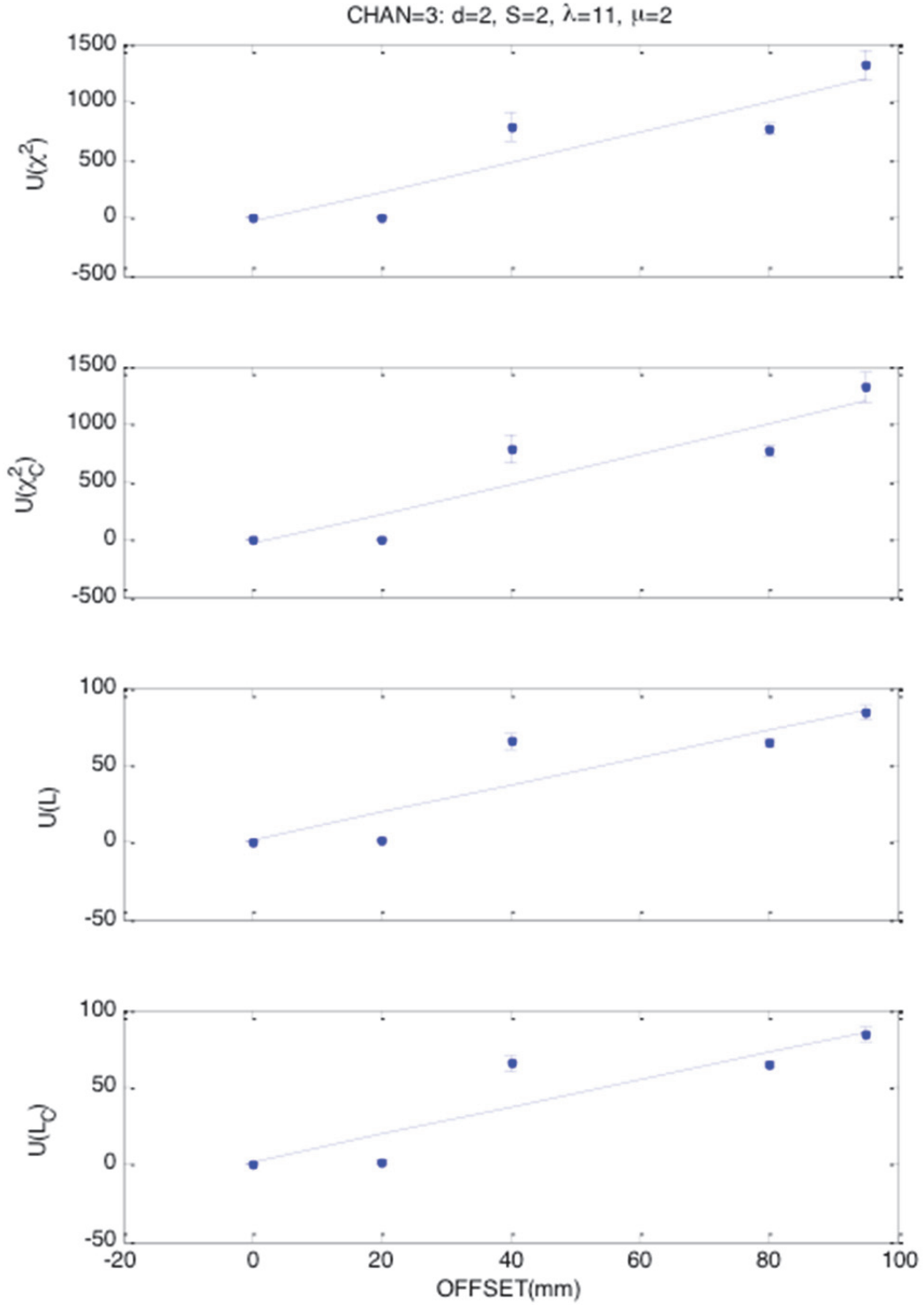


Fig. 7. PSDM vs pier offset level for sensor 229 (Channel 3).

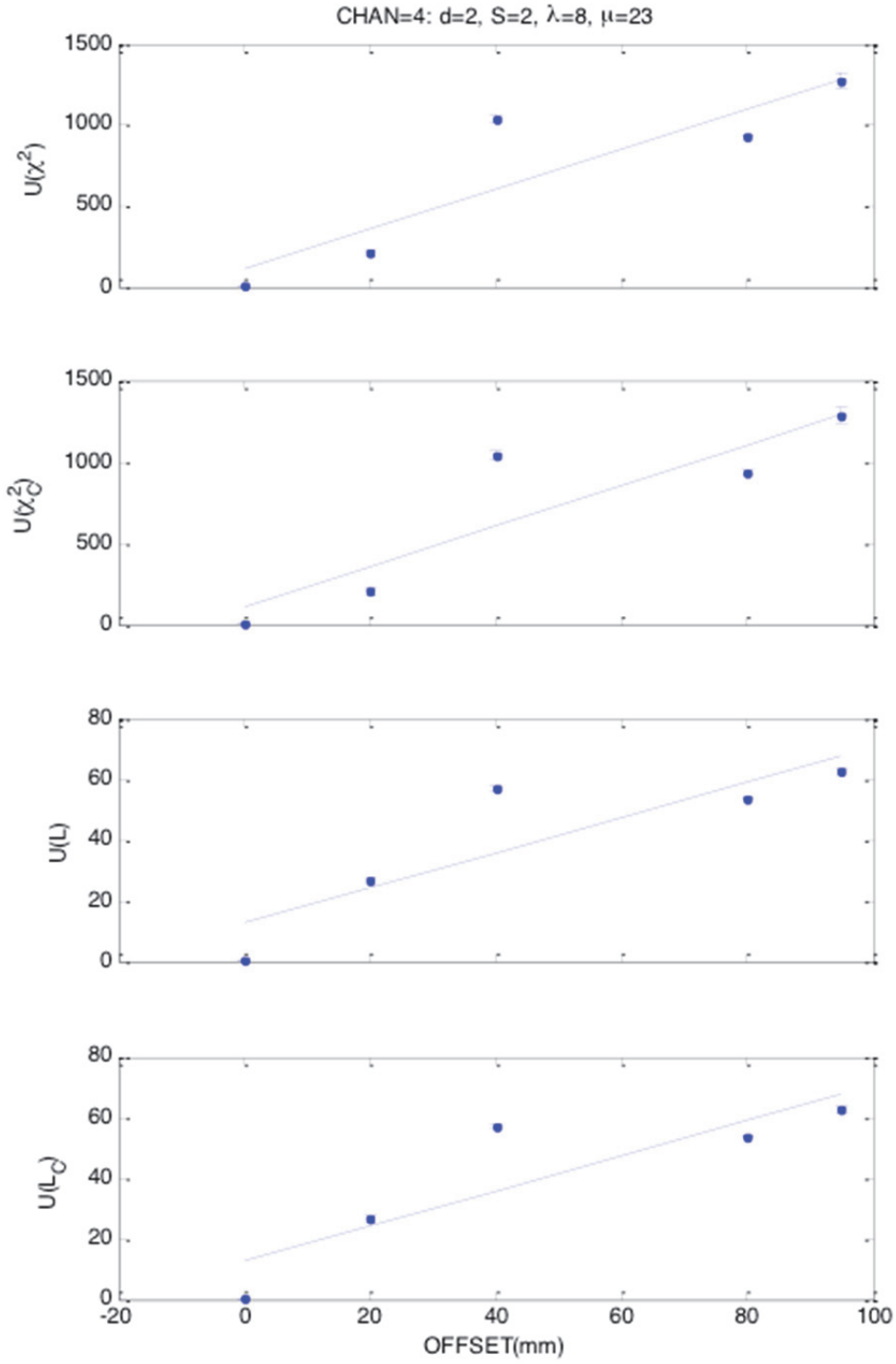


Fig. 8. PSDM vs pier offset level for sensor 233 (Channel 4).

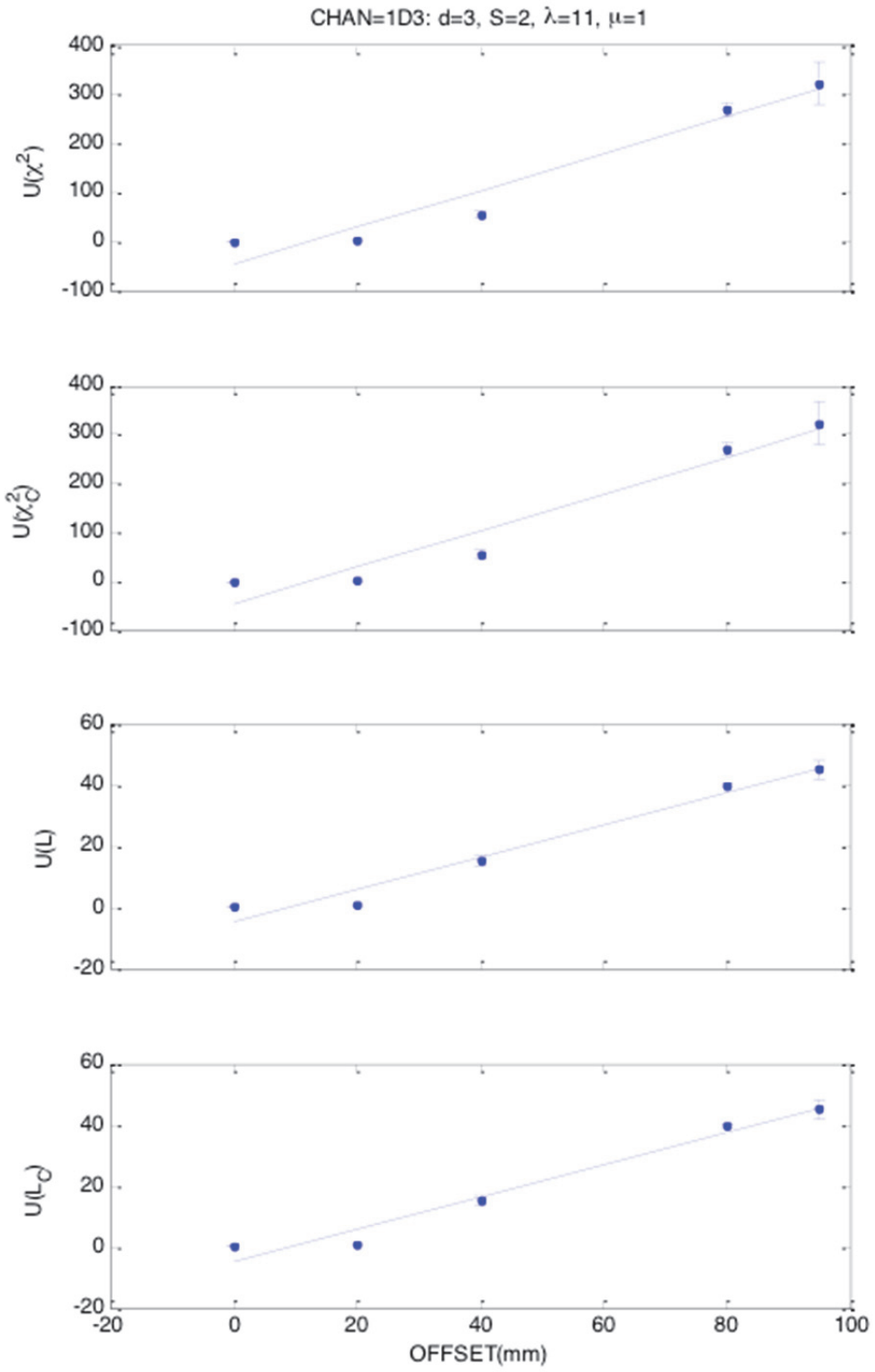


Fig. 9. PSDM vs pier offset level for uniaxial data, sensor 229 (Channel 3).

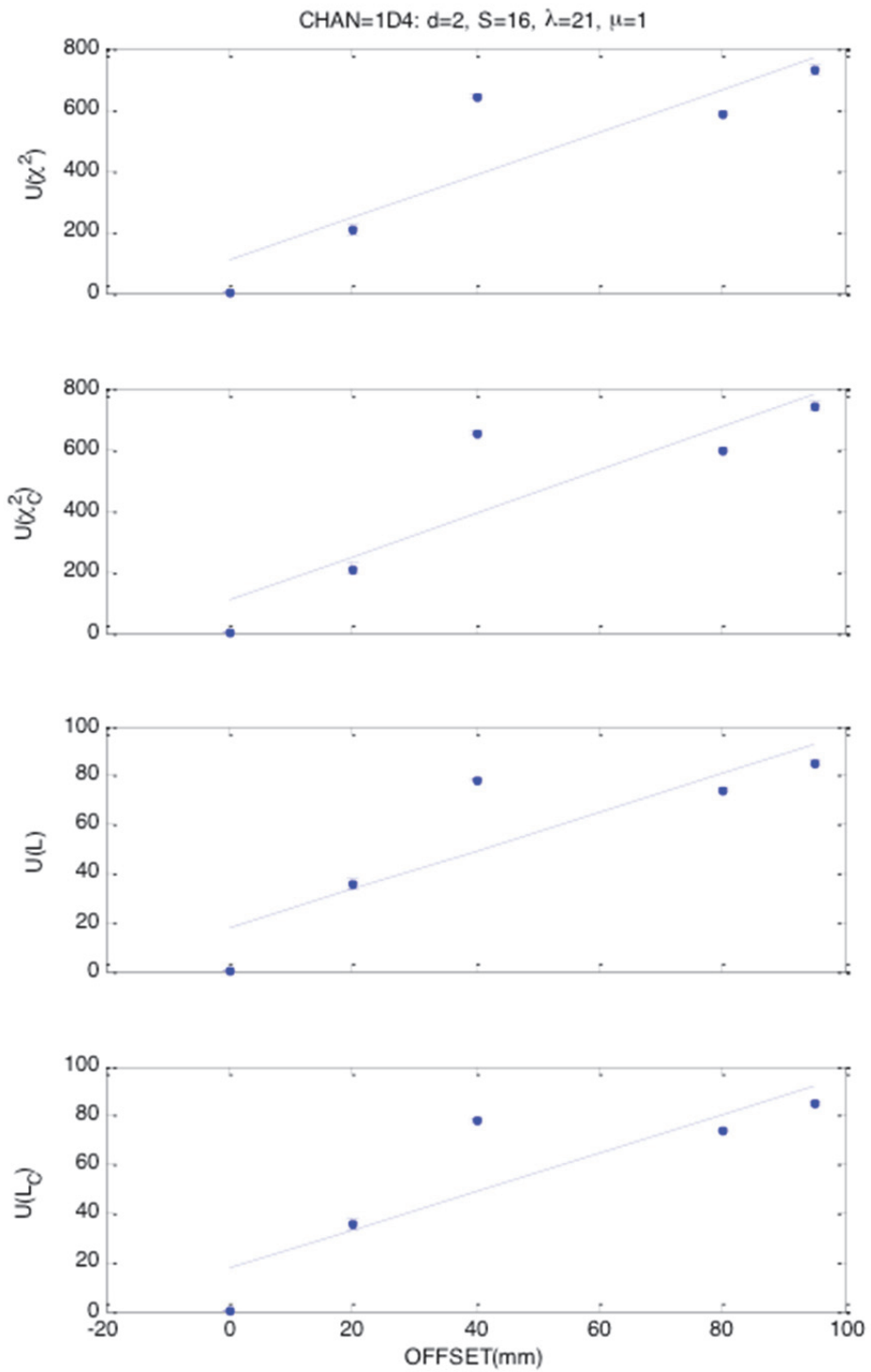


Fig. 10. PSDM vs pier offset level for uniaxial data, sensor 233 (Channel 4).

5.4 RESULTS FOR STATISTICAL VALIDATION

Table 2 summarizes the results of two series of tests. The first consists of 80 one-sided, two-sample student-t hypothesis tests. In each case, the test sample of five PSDMs was compared to the control sample of five PSDMs at the undamaged level. In Table 2 below, the numbers in each cell represent the significance level, with which each PSDM correctly detected damage. The second series of tests involves the construction of prediction intervals around the damage levels predicted for each PSDM by the linear regression equation. Table 3 contains the values for the sensitivity (true positives divided by total damaged cases) and specificity (true negatives divided by total undamaged cases) of each PSDM at each sensor location, as well as the combined values for all PSDMs at each sensor location. Here, “true positive” is taken to mean that for the cases in which damage level is not zero, the prediction interval contains the actual damage level *and* does not contain zero. “False negatives” include the damaged cases in which either the actual damage level is not within the prediction interval or zero is within the interval. “True negatives” include the undamaged cases in which zero is within the prediction interval, and “false positives” include the undamaged cases in which zero is not within the interval.

Table 2. Triaxial Damage Detection Strength of Each Sensor Location at Each Damage Level

		Damage Level 1 (20 mm)	Damage Level 2 (40 mm)	Damage Level 3 (80 mm)	Damage Level 4 (95 mm)
Sensor 213 (Channel 1)	χ^2	1.57 E -06	5.54 E -07	1.83 E -07	2.37 E -08
	χ_c^2	1.42 E -06	5.35 E -07	1.86 E -07	2.12 E -08
	L	2.18 E -07	5.22 E -08	3.68 E -09	1.12 E -10
	L_C	2.17 E -07	5.15 E -08	4.30 E -09	1.13 E -10
Sensor 223 (Channel 2)	χ^2	4.59 E -05	7.55 E -05	1.18 E -05	4.20 E -05
	χ_c^2	4.47 E -05	7.07 E -05	1.11 E -05	0.000324
	L	6.20 E -06	1.19 E -05	3.92 E -06	9.45 E -06
	L_C	6.17 E -06	1.18 E -05	3.88 E -06	9.32 E -06
Sensor 229 (Channel 3)	χ^2	0.034104	7.14 E -05	1.51 E -06	1.42 E -05
	χ_c^2	0.034112	7.15 E -05	1.51 E -06	0.000239
	L	0.011057	4.50 E -06	5.25 E -08	1.19 E -06
	L_C	0.011057	4.50 E -06	5.25 E -08	1.19 E -06
Sensor 233 (Channel 4)	χ^2	8.51 E -06	5.59 E -08	1.26 E -07	1.99 E -07
	χ_c^2	8.59 E -06	5.56 E -08	9.32 E -06	9.32 E -06
	L	5.11 E -07	1.86 E -09	5.24 E -09	5.24 E -09
	L_C	5.11 E -07	1.86 E -09	5.24 E -09	5.24 E -09
Sensor 541 (Channel 5)	χ^2	7.16 E -08	7.58 E -09	2.91 E -05	1.15 E -08
	χ_c^2	7.28 E -08	5.77 E -09	2.59 E -05	2.45 E -07
	L	2.75 E -10	3.42 E -16	3.77 E -06	1.82 E -13
	L_C	2.83 E -10	3.48 E -16	3.67 E -06	1.50 E -13

Table 3. Triaxial Extent-of-Damage Prediction Capabilities of PSDM and at Each Sensor Location

		True Positive	False Negative	True Negative	False Positive	Sensitivity (TP/(TP+FN))	Specificity (TN/(TN+FP))
Sensor 213 (Channel 1)	χ^2	0	20	5	0	0.00	1.00
	χ_c^2	0	20	5	0	0.00	1.00
	L	0	20	5	0	0.00	1.00
	L _C	0	20	5	0	0.00	1.00
	Total	0	80	20	0	0.00	1.00
Sensor 223 (Channel 2)	χ^2	8	12	5	0	0.40	1.00
	χ_c^2	8	12	5	0	0.40	1.00
	L	8	12	5	0	0.40	1.00
	L _C	8	12	5	0	0.40	1.00
	Total	32	48	20	0	0.40	1.00
Sensor 229 (Channel 3)	χ^2	15	5	5	0	0.75	1.00
	χ_c^2	15	5	5	0	0.75	1.00
	L	15	5	5	0	0.75	1.00
	L _C	15	5	5	0	0.75	1.00
	Total	60	20	20	0	0.75	1.00
Sensor 233 (Channel 4)	χ^2	15	5	5	0	0.75	1.00
	χ_c^2	15	5	5	0	0.75	1.00
	L	15	5	5	0	0.75	1.00
	L _C	15	5	5	0	0.75	1.00
	Total	60	20	20	0	0.75	1.00
Sensor 541 (Channel 5)	χ^2	0	20	5	0	0.00	1.00
	χ_c^2	0	20	5	0	0.00	1.00
	L	4	16	5	0	0.20	1.00
	L _C	4	16	5	0	0.20	1.00
	Total	8	72	20	0	0.10	1.00

The tables below contain the results of the same testing scheme, this time applied to uniaxial data from sensors 223, 229, and 233. Comparison of these results to those summarized in the tables above will determine the damage detection strength and extent-of-damage prediction capabilities of uniaxial data, as compared to triaxial data at each sensor location.

Table 4. Uniaxial Damage Detection Strength of Each Sensor Location at Each Damage Level

		Damage Level 1 (20 mm)	Damage Level 2 (40 mm)	Damage Level 3 (80 mm)	Damage Level 4 (95 mm)
Sensor 223 (Channel 2)	χ^2	7.24 E -05	4.53 E -05	7.87 E -06	5.19 E -05
	χ_c	7.19 E -05	4.44 E -05	7.22 E -06	0.000352
	L	1.23 E -05	9.36 E -06	2.96 E -06	1.23 E -05
	L_c	1.23 E -05	9.40 E -06	2.87 E -06	1.23 E -05
Sensor 229 (Channel 3)	χ^2	0.046381	4.75 E -05	6.90 E -07	3.84 E -05
	χ_c	0.046351	4.74 E -05	6.91 E -07	0.000403
	L	0.032973	1.83 E -05	3.70 E -09	2.41 E -06
	L_c	0.032973	1.83 E -05	3.70 E -09	2.41 E -06
Sensor 233 (Channel 4)	χ^2	8.19 E -06	1.09 E -09	3.48 E -09	2.05 E -08
	χ_c	7.25 E -06	1.14 E -09	3.39 E -09	1.95 E -08
	L	1.78 E -06	1.86 E -11	1.40 E -11	1.11 E -09
	L_c	1.74 E -06	1.97 E -11	1.31 E -11	1.09 E -09

Table 5. Uniaxial Extent-of-Damage Prediction Capabilities of Each PSDM and at Each Sensor Location

		True Positive	False Negative	True Negative	False Positive	Sensitivity (TP/(TP+FN))	Specificity (TN/(TN+FP))
Sensor 223 (Channel 2)	χ^2	9	11	5	0	0.45	1.00
	χ_c	9	11	5	0	0.45	1.00
	L	8	12	5	0	0.40	1.00
	L_c	8	12	5	0	0.40	1.00
Total		34	46	20	0	0.425	1.00
Sensor 229 (Channel 3)	χ^2	15	5	5	0	0.75	1.00
	χ_c	15	5	5	0	0.75	1.00
	L	15	5	5	0	0.75	1.00
	L_c	15	5	5	0	0.75	1.00
Total		60	20	20	0	0.75	1.00
Sensor 233 (Channel 4)	χ^2	15	5	5	0	0.75	1.00
	χ_c	15	5	5	0	0.75	1.00
	L	15	5	5	0	0.75	1.00
	L_c	15	5	5	0	0.75	1.00
Total		60	20	20	0	0.75	1.00

6. DISCUSSION

From examining the tables in Sect. 5, strength of the selected methods in damage detection and extent-of-damage quantification was inferred. All sensors have relatively low sensitivity while having perfect specificity levels of 1.00. This is because the relatively poor linear correlations have caused the 95% prediction intervals to be so wide that they are more likely to contain zero, suggesting that the damage might be non-existent, and therefore one of the criteria lowering sensitivity to extent-of-damage.

Moreover, the damage detection mode gave highly statistically significant results in all four PSDMs at all damage levels and sensors.

A direct reason for the strength of the algorithm in damage detection and not in extent-of-damage quantification is the way in which the algorithm was trained to calculate the PSDMs from the data provided. The phase-space parameters were selected specifically in order to maximize the sum of the slopes of the regression lines; however, this maximization is accompanied by the unfavorable sum of squares of the regression line residuals. The difference in the means of the PSDMs at each damage level and the undamaged state are being intentionally maximized (causing the strongest possible results in simple damage detection, with p-values many orders of magnitude more extreme than the selected 0.05 alpha level in most cases). The unintentional result is that prediction error in inferring extent-of-damage from the regression line increases. A possible remedy for this issue is to optimize the selection of phase-space parameters during the training process to produce a sum of slopes that is still significant at the specified alpha level (but not orders of magnitude more significant), and also a sum of squares of regression line residuals that is small enough to provide more sensitive extent-of-damage detection from the regression line of predicted values.

As is apparent in the plots, the uniaxial analysis produced PSDMs with a slightly better linear correlation. Consequently, the prediction intervals constructed from the standard deviation of the horizontal residuals (the amount of error in the damage level predictions) were smaller. As a result, the uniaxial analyses produced the same or slightly better sensitivity levels in the extent-of-damage prediction by decreasing the number of false negatives in the prediction intervals, and also made the damage level prediction more precise. Mean damage level prediction errors for Channels 2, 3, and 4 were 22–23 mm, 13–14 mm, and 14–16 mm, respectively, for the triaxial data, and smaller for the uniaxial data: 21–23 mm, 6–9 mm, and 7–8 mm.

7. CONCLUSIONS AND FUTURE WORK

A phase-space-based methodology for SHM analysis via structural vibrations was presented. The primary advantages of this approach are

1. The theoretical basis in the time-delay embedding theorem for detection of condition change;
2. Low-complexity algorithm that scales linearly with the number of processors;
3. Use of real, noisy, finite-length data from a full-size bridge with representative damage; and
4. Statistical demonstration of condition change measures with the bridge damage.

A number of issues remain to be addressed in future work. Identified and discussed below are some notable ones. First, those that directly follow from this work are listed, then more general challenges are listed.

- **Multiaxial Optimization.** Analyses from single-axis vertical accelerometer data with that from all three axes of each accelerometer was compared, but it would be worthwhile to understand how analyses from all seven different axis combinations compare to each other.
- **Investigation of other damage cases.** Only one damage case and only five sensors for the Z24 bridge was analyzed. Further analysis would be useful for statistical validation and sensor placement evaluation. It was difficult based on the limited analysis to determine how sensor location affected sensitivity to damage, since there are multiple factors that can influence sensitivity of a particular sensor, such as distance from damage location, location within span, location of excitation, etc.
- **Management of signal variability.** The present analysis used data from a bridge under forced vibration, not ambient or in-service vibration. Also, only one very specific type and location of damage existed. An important area of future study would be use of data from different loading

and damage scenarios. The challenge would be to minimize the frequency of false positive and false negative results under the uncertainty of whether changes in bridge's dynamics are due to damage. Examples of sources of such variability not caused by damage include change in temperature, traffic loading, and wind loading. Examples of damage that may have similar effects are scouring, weld cracking, and corrosion. Briaud et al. (2011) have thoroughly investigated the use of accelerometers and other instruments and techniques to address the problem of scouring.

- Parameter Set Optimization Methodology. The choice of parameter set for the PSDM analysis was based on trials of many different combinations of parameter values; the number of combinations for this analysis is on the order of one hundred thousand. This process required retrospective data to obtain a parameter set in order to conduct prospective forewarning analysis. Clearly, retrospective failure data is difficult to obtain for all possible types and sizes of bridges, so could be desirable to have a methodology to be able to make a more intelligent choice of parameter set with less or no retrospective data.

8. REFERENCES

- R. K. Abercrombie, L. M. Hively, S. J. Prowell, B. G. Schlicher, and F. T. Sheldon, "Forewarning of Failure in Complex Systems" (2011).
- R. K. Abercrombie and L. M. Hively, "Method for Forewarning of Critical Conditions in Monitoring Civil Structures," patent pending (invention disclosure #201002369) submitted to USPTO (December 8, 2010).
- D. Abruzzese, M. Angelaccio, B. Buttarazzi, R. Giuliano, L. Miccoli, and A. Vari, "Long Life Monitoring of Historical Monuments via Wireless Sensors Network," *Wireless Communication Systems*, **6**, 570–574 (2009).
- D. Adams and C. R. Farrar, "Classifying Linear and Nonlinear Structural Damage Using Frequency Domain ARX Models," *Structural Health Monitoring*, **1**(2), 185–201 (2002).
- American Society of Civil Engineers (ASCE), "Report Card for America's Infrastructure" www.asce.org/reportcard (2009).
- A. Bahar, F. Pozo, L. Acho, J. Rodellar, and A. Barbat, "Hierarchical Semi-active Control of Base-isolated Structures Using a New Inverse Model of Magnetorheological Dampers," *Computer and Structures*, **88**, 483–496 (2010).
- A. Bilbao, D. Hoover, J. Rice, and J. Chapman, "Ultra-low Power Wireless Sensing for Long-term Structural Health Monitoring," *Sensors and Smart Technologies for Civil, Mechanical, and Aerospace Systems, proceedings of SPIE*, **7981** (2011).
- J. L. Briaud, S. Hurlebaus, K. A. Chang, C. Yao, H. Sharma, O. Y. Yu, C. Darby, B. E. Hunt, and G. R. Price, "Realtime Monitoring of Bridge Scour Using Remote Monitoring Technology," (Report 0-6060-1) College Station, TX: Texas Transportation Institute (February 2011).
- J. M. W. Brownjohn, "Structural Health Monitoring of Civil Infrastructure," *Phil. Trans. R. Soc. A*, **365**, 589–622 (2007).
- P. J. Boltryk, C. J. Harris, and N. M. White, "Intelligent Sensors – A Generic Software Approach," *J. of Physics*, Conference Series 15, 155–160 (2005).
- M. Bocca, E. I. Cosar, J. Salminen, and L. M. Eriksson, "A Reconfigurable Wireless Sensor Network for Structural Health Monitoring," *Proceedings from SHMII-4: Fourth International Conference on Structural Health Monitoring of Intelligent Infrastructure*, (Zurich, Switzerland) 2009.
- C. Boller and M. Buderath "Fatigue in Aerostructures – Where Structural Health Monitoring Can Contribute to a Complex Subject," *Phil. Trans. R. Soc. A*, **365**, 561–587 (2007).
- E. P. Carden and J. M. W. Brownjohn, "Fuzzy Clustering of Stability Diagrams for Vibration-Based Structural Health Monitoring," *Computer-Aided Civil and Infrastructure Engineering*, **23**, 360–372 (2008).
- K. Chang, T. Lin, Y. Lin, and L. Wang, "FBG Sensors for Structural Health Monitoring," *Advanced Nondestructive Evaluation for Structural and Biological Health Monitoring, Proceedings of SPIE*, **4335** (2001).
- K. Chebrolu, B. Raman, N. Mishra, P. K. Valiveti, and R. Kumar, "BriMon: A Sensor Network System for Railway Bridge Monitoring," *Proceedings of Sixth International Conference on Mobile Systems, Applications, and Services*, Breckenridge, Colorado (June 17–20, 2008).
- S. Cheng and F. A. Spiring, "A Test to Identify the Uniform Distribution, with Applications to Probability Plotting and Other Distributions," *IEEE Transactions on Reliability*, **R36**, 98–105 (1987).
- N. E. Clapp and L.M. Hively, "Method and Apparatus for Extraction of Low-Frequency Artifacts from Brain Waves for Alertness Detection," *U.S. Patent #5,626,145* (May 6, 1997).
- Crossbow, "Crossbow Announces Crossbow Imote2, the Most Powerful Mote Platform for Wireless Sensor Network Research," Crossbow Technology, Inc., San Jose, CA (2007).

S. Doebling, F.M. Hemez, and W. Rhee, "Statistical Model Updating and Validation Applied to Nonlinear Transient Structural Dynamics," Proceedings of the European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, 409–418 (June 6–9, 2000).

A. Dunkels, B. Gronvall, and T. Voigt, "Contiki - a Lightweight and Flexible Operating System for Tiny Networked Sensors," *Proceedings of the 29th annual IEEE International Conference on Local Computer Networks (LCN'04)* (November 16–18, 2004).

Z. H. Duron, "Early Warning Capabilities for Firefighters: Testing of Collapse Prediction Technologies" [NIST GCR 03-846], National Institute of Standards and Technology, Gaithersburg, MD (2003).

A. A. Elshafey, M. R. Haddara, and H. Marzouk, "Damage Detection in Offshore Structures Using Neural Networks," *Marine Structures*, **23**, 131–145 (2010).

C. Krämer and C. de Smet, "Results of Tasks A1 and A2: Long-term Monitoring and Bridge Tests," SIMCES Tasks A1 and A2 Internal Report [No. 168'349/20e], Dubendorf, Switzerland:EMPA (May 25, 1999).

ERI Distance Learning Center. Appendix – Table VII: Kolmogorov-Smirnov Test. Accessed June 17, 2011.

<http://www.eridlc.com/onlinetextbook/index.cfm?fuseaction=textbook.appendix&FileName=Table7>

G. Fabbrocino, C. Laorenza, C. Rainieri, and F. S. de Magistris, "Seismic Monitoring of Structural and Geotechnical Integrated Systems," *Materials Forum*, **33**, 404–419 (2009).

W. Fan and P. Qiaon, "Vibration-based Damage Identification Methods: A Review and Comparative Study," *Structural Health Monitoring*, **10**, 83–111 (2010).

M. O. Farooq and T. Kunz, "Operating Systems for Wireless Sensor Networks: A Survey," *Sensors*, **11**, 5900–5930 (2011).

C. R. Farrar, P. J. Cornwell, S. W. Doebling, and M. B. Prime, "Structural Health Monitoring Studies of the Alamosa Canyon and I-40 Bridges," *Los Alamos National Laboratory Document LA-13635-MS* (July 2000).

C. R. Farrar, S. W. Doebling, and D. A. Nix, "Vibration-based Structural Damage Identification." *Philosophical Transactions of the Royal Society A*, **359**, 131–149 (2001).

C. R. Farrar and N. A. J. Lieven, "Damage Prognosis: The Future of Structural Health Monitoring." *Phil. Trans. R. Soc. A*, **365**, 623–632 (2007).

C. R. Farrar and K. Worden, "An Introduction to Structural Health Monitoring." *Phil. Trans. R. Soc. A*, **365**, 303-315 (2007).

C. R. Farrar, K. Worden, M. D. Todd, G. Park, J. Nichols, D. E. Adams, M. T. Bement, and K. Farinholt, "Nonlinear System Identification for Damage Detection," *Los Alamos National Laboratory Document LA-14353* (2007).

S. D. Fassois and J. S. Sakellariou, "Time-series Methods for Fault Detection and Identification in Vibrating Structures," *Phil. Trans. R. Soc. A*, **365**, 411–448 (2007).

C. P. Fritzen and P. Kraemer, "Self-diagnosis of Smart Structures Based on Dynamical Properties," *Mechanical Systems and Signal Processing*, **23**, 1830–1845 (2009).

J. H. Galbreath, C. P. Townsend, S. W. Mundell, M. J. Hamel, B. Esser, D. Huston, and S. W. Arms, "Civil Structure Strain Monitoring with Power-Efficient, High-Speed Wireless Sensor Networks," *Presented at Fourth International Workshop on Structural Health Monitoring*, Stanford University, Stanford, CA (September 15–17, 2003).

M. V. Gangone, M. J. Whelan, and K. D. Janoyan, "Wireless Monitoring of a Multispan Bridge Superstructure for Diagnostic Load Testing and System Identification," *Computer-Aided Civil and Infrastructure Engineering*, **26**, 1–20 (2011).

G. Hackmann, F. Sun, N. Castaneda, C. Lu, and S. Dyke, "A Holistic Approach to Decentralized Structural Damage Localization Using Wireless Sensor Networks," *Real-Time Systems Symposium*, 35–46 (2008).

C. C. Han, R. Kumar, R. Shea, E. Kohler, and M. Srivastava, "A Dynamic Operating System for Sensor Nodes," *Proceedings of the Third International Conference on Mobile Systems, Applications, and Services (MobSys'05)*, Seattle, WA (June 6–8, 2005).

T. Harms, F. Bastianini, and S. S. Sarvestani, "An Embedded Wireless System for Remote Monitoring of Bridges," *Sensors and Smart Technologies for Civil, Mechanical, and Aerospace Systems 2008, Proceedings of SPIE*, **6932** (2008).

T. Harms, B. Banks, S. S. Sarvestani, and F. Bastianini, "Design and Testing of a Low-Power Wireless Sensor Network for Structural Health Monitoring of Bridges," *Sensors and Smart Technologies for Civil, Mechanical, and Aerospace Systems 2009, Proceedings of SPIE*, **7292** (2009[a]).

T. Harms, F. Bastianini, and S. Sedigh, "Recent Enhancements to the SmartBrick Structural Health Monitoring Platform," *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MO (October 3–7, 2009[b]).

F. M. Hemez, S. W. Doebling, and M. C. Anderson, *Model Verification and Validation in Engineering Mechanics: Theory and Applications of Uncertainty Quantification and Predictive Accuracy* (Wiley, New York) 2005.

J. Hill, M. Horton, R. Kling, and L. Krishnamurthy, "The Platforms Enabling Wireless Sensor Networks," *Communications of the ACM*, **47**(6), 41–46 (2004).

L. M. Hively, "Methods for Improved Forewarning of Condition Changes in Monitoring of Physical Processes," patent pending (Invention disclosure #201002321) (2010).

L. M. Hively, "Failure Prediction of Complex Structures Under Arbitrary /time-serial Loading Condition," patent pending (Invention disclosure #200802045) submitted to USPTO (Aug. 2009).

L. M. Hively, N. E. Clapp, C. S. Daw, W. F. Lawkins, "Apparatus and Method for Epileptic Seizure Detection using Nonlinear Techniques," *U.S. Patent #5,743,860* (April 28, 1998).

L. M. Hively and E. G. Ng, "Integrated Method for Chaotic Time Series Analysis," *U.S. Patent #5,815,413* (September 29, 1998).

L. M. Hively, N. E. Clapp, C. S. Daw, W. F. Lawkins, "Epileptic Seizure Prediction By Non-linear Methods," *U.S. Patent #5,857,978* (January 12, 1999).

L. M. Hively, P. C. Gailey, V. A. Protopopescu, "Condition Assessment of Nonlinear Processes," *U.S. Patent #6,484,132* (November 19, 2002).

L. M. Hively, "Methods for Consistent Forewarning of Critical Events Across Multiple Data Channels," *U.S. Patent #7,139,677* (November 21, 2006).

L. M. Hively, "Methods for Improved Forewarning of Critical Events Across Multiple Data Channels," *U.S. Patent #7,209,861* (April 24, 2007).

L. M. Hively, "Methods for Improved Forewarning of Critical Events Across Multiple Data Channels," *European Patent #1665131* (June 10, 2009).

L. M. Hively, V. A. Protopopescu, and N. B. Munro, "Epilepsy Forewarning via Phase-Space Dissimilarity," *J. Clin. Physiol.*, **22**, 402–409 (December 2005).

L. M. Hively, "Prognostic Analysis of the Tactical Quiet Generator," ORNL/TM-2008-012455, Oak Ridge National Laboratory, Oak Ridge, TN (September 2008).

L. M. Hively, "Prognostication of Helicopter Failure," ORNL/TM-2009-244, Oak Ridge National Laboratory, Oak Ridge, TN (November 2009).

N. R. Jennings, K. Sycara, and M. Wooldridge, "A Roadmap of Agent Research and Development." *Autonomous Agents and Multi-Agent Systems*. **1**, 7–38 (1998).

S. Kim, S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser, and M. Turon, "Health Monitoring of Civil Infrastructures Using Wireless Sensor Networks," *Int. Conf. on Information Processing in Sensor Networks* (April 25–28, 2007).

H. F. Lam and T. Yin, "Dynamic Reduction-based Structural Damage Detection of Transmission Towers: Practical Issues and Experimental Verification," *Engineering Structures*, **33**, 1459–1478 (2011).

P. Y. Lin, P. Roschke, and C. H. Loh, "System Identification and Real Application of a Smart Magneto-Rheological Damper," *Proceedings of the 2005 IEEE International Symposium on Intelligent Control*, Limassol, Cyprus (June 27–29, 2005).

J. P. Lynch and K. J. Loh, "A Summary Review of Wireless Sensors and Sensor Networks for Structural Health Monitoring," *The Shock and Vibration Digest*, **38**(2), 91–128 (2006).

Y. Marhuenda, D. Morales, and M. C. Pardo, "Power Results of Tests for the Uniform Distribution," *Centro de Investigacion Operativa* (2005)
http://cio.umh.es/es/publicaciones/ficheros/CIO_2005_09.pdf

K. Mechitov, W. Kim, G. Agha, and T. Nagayama, "High-frequency Distributed Sensing for Structure Monitoring." *Trans. of the Society of Instrument and Control Engineers*, **E-S-1** (1), 109–114 (2006).

MEMSIC, "MICAz Wireless Measurement System," San Jose, CA: MEMSIC, Inc.

J. Memmott, "Highway Bridges in the United States—an Overview," Bureau of Transportation Statistics Special Report [SR-003], USDOT Research and Innovative Technology Administration (September 2007). Figure Retrieved January 31, 2010, from
http://www.bts.gov/publications/special_reports_and_issue_briefs/special_report/2007_09_19/html/figure_03.html

F. Miller (1975) Statistics for Testing Uniformity on the Unit Interval. Institute of Statistics Mimeograph Series No. 983. http://www.stat.ncsu.edu/library/mimeo.archive/ISMS_1975_983.pdf

L. Moniz, J. M. Nichols, C. J. Nichols, M. Seaver, S. T. Trickey, M. D. Todd, L. M. Pecora, and L. N. Virgin, "A Multivariate, Attractor-based Approach to Structural Health Monitoring," *J. of Sound and Vibration*, **283**, 295–310 (2005).

Moteiv, "Tmote Sky: Datasheet," San Francisco, CA: Moteiv Corporation (February 6, 2006).

M. Moore, B. Phares, B. Graybeal, D. Rolander, and G. Washer, "Reliability of Visual Inspection for Highway Bridges," *Publication No. FHWA-RD-01-020*, McLean, VA: Federal Highway Administration (June 2001).

National Institute of Standards and Technology. 1.3.5.16. Kolmogorov-Smirnov Goodness-of-Fit Test. *Engineering Statistics Handbook*. Accessed June 17, 2011.
<http://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm>

National Institute of Standards and Technology. 1.3.5.14. Anderson-Darling Test. *Engineering Statistics Handbook*. Accessed June 17, 2011.
<http://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm>

J. M. Nichols, "Structural Health Monitoring of offshore Structures Using Ambient Excitation," *Applied Ocean Research*, **25**, 101–114 (2003).

J. M. Nichols, M. D. Todd, M. Seaver, and L. N. Virgin, "Use of Chaotic Excitation and Attractor Property Analysis in Structural Health Monitoring," *Physical Review E*, **67**, 016209 (2003).

L. A. Overbey and M. D. Todd, "Damage Assessment using Generalized State-Space Correlation Features," *Structural Health Monitoring*, **7**, 347–363 (2008).

T. G. S. Overly, G. Park, K. M. Farinholt, and C. R. Farrar, "Development of an Extremely Compact Impedance-Based Wireless Sensing Device," *Smart Materials and Structures*, **17**, 065011-065019 (2008).

J. Paek, O. Gnawali, K. Y. Jang, D. Nishimura, R. Govindan, J. Caffrey, M. Wahbeh, and S. Masri, "A Programmable Wireless Sensing System for Structural Monitoring," *Fourth World Conference on Structural Control and Monitoring*, San Diego, CA (July 11–13, 2006).

J. Paek, B. Greenstein, O. Gnawali, K. Y. Jang, A. Joki, M. Vieira, J. Hicks, D. Estrin, R. Govindan, and E. Kohler, "The Tenet Architecture for Tiered Sensor Networks," *ACM Transactions on Sensor Networks*, **6**(4), Article 34 (July 2010).

S. N. Pakzad, G. L. Fenves, S. Kim, and D. E. Culler, "Design and Implementation of Scalable Wireless Sensor Network for Structural Monitoring," *Journal of Infrastructure Systems*, **14**(1), 89–101 (2008).

- G. Park, C. R. Farrar, F. L. Di Scalea, and S. Coccia, "Performance Assessment and Validation of Piezoelectric Active-Sensors in Structural Health Monitoring," *Smart Materials and Structures*, **15**, 1673–1683 (2006).
- G. Park, C. R. Farrar, M. D. Todd, W. Hodgkiss, and T. Rosing, "Energy Harvesting for Structural Health Monitoring Sensor Networks," *Los Alamos National Laboratory Document LA-14314-MS*, (2007).
- I. Perez, M. DiUlio, S. Maley, and N. Phan, "Structural Health Management in the NAVY," *Structural Health Monitoring*, **9**(3), 199–207, (2010).
- V. Protopopescu and L.M. Hively, "Phase-space Dissimilarity Measures of Nonlinear Dynamics: Industrial and Biomedical Applications," *Recent Res. Devel. Physics*, **6**, 649–688 (2005).
- M. Rahman, L. Pearson, and H. Heien, "A Modified Anderson-Darling Test for Uniformity," *Bulletin of the Malaysian Mathematical Sciences Society (2)*, **29**(1), 11–16 (2006).
<http://math.usm.my/bulletin/pdf/v29n1/v29n1p2.pdf>
- C. Rainieri, G. Fabbrocino, and E. Cosenza, "Integrated Seismic Early Warning and Structural Health Monitoring of Critical Civil Infrastructures in Seismically Prone Areas," *Structural Health Monitoring*, **10**(3), 291–308 (2010).
- J. Roscoe and J. Byars, "An Investigation of the Restraints with Respect to Sample Size Commonly Imposed on the Use of the Chi-Square Statistic," *J. of the American Statistical Assoc.*, **66**, 755–759 (1971).
- M. E. Ruiz-Sandoval, "*Smart*" Sensors for Civil Infrastructure Systems (Doctoral Dissertation), University of Notre Dame, Notre Dame, IN (May 2004).
- D. A. Saftner, R. D. Hryciw, R. A. Green, J. P. Lynch, and R. L. Michalowski, "The Use of Wireless Sensors in Geotechnical Field Applications," *Proceedings of the 15th Annual Great Lakes Geotechnical/Geoenvironmental Conference*, Indianapolis, IN (2008).
- H. Sohn and C. R. Farrar, "Damage Diagnosis Using Time Series Analysis of Vibration Signals," *Smart Materials and Structures*, **10**, 446–451 (2001).
- S. Sridhar, K. Ravisankar, P. Sreeshylam, S. Parivallal, K. Kesavan, and S. G. N. Murthy, "Development of a Real-time Remote Structural Monitoring Scheme for Civil Infrastructural Systems," *Structural Health Monitoring*, **8**(6), 0509–0521 (2009).
- W. J. Staszewski and A. N. Robertson, "Time-Frequency and Time-Scale Analysis for Structural Health Monitoring," *Phil. Trans. R. Soc. A*, **365**, 449–477 (2007).
- I. Stoianov, L. Nachman, and S. Madden, "PIPENET: A Wireless Sensor Network for Pipeline Monitoring," *International Conference on Information Processing in Sensor Networks* (April 25–27, 2007).
- R. A. Swartz, A. T. Zimmerman, J. P. Lynch, J. Rosario, T. Brady, L. Salvino, and K. H. Law, "Hybrid Wireless Hull Monitoring System for Naval Combat Vessels," (2010).
- S. G. Taylor, K. M. Farinholt, E. B. Flynn, E. Figueiredo, D. L. Mascarenas, E. A. Moro, G. Park, M. D. Todd, and C. R. Farrar, "A Mobile-agent-based Wireless Sensing Network for Structural Monitoring Applications," *Measurement Science and Technology*, **20**, 045201–045214 (2009).
- TinyOS. <http://www.tinyos.net/>.
- M. D. Todd, J. M. Nichols, S. T. Trickey, M. Seaver, C. J. Nichols, and L. N. Virgin, "Bragg Grating-based Fibre Optic Sensors in Structural Health Monitoring," *Phil. Trans. R. Soc. A*, **365**, 317–343 (2007).
- S. Torkamani, E. A. Butcher, M. D. Todd, and G. Park, "Damage Assessment Using Hyperchaotic Excitation and State-space Geometry Changes," *Proceedings of the ASME 2010 Conference on Smart Materials, Adaptive Structures and Intelligent Systems*, Philadelphia, PA, USA, **2**, 599–608 (September 28 – October 1, 2010).
- I. Trendafilova, "A State Space Based Approach to Health Monitoring of Vibrating Structures," *Materials Science Forum*, **440–441**, 203–210 (2003).

- I. Trendafilova, "Vibration-based Damage Detection in Structures Using Time Series Analysis," *J. of Mech. Eng. Sci.*, Proceedings of the Institution of Mechanical Engineers Part C, **220**(3), 261–272 (2006).
- D. E. Welch, L. M. Hively, and R. F. Holdaway, "Nonlinear Structural Crack Growth Monitoring," *U.S. Patent #6,460,012* (October 1, 2002).
- M. J. Whelan and K. D. Janoyan, "Design of a Robust, High-Rate Wireless Sensor Network for Static and Dynamic Structural Monitoring," *J. of Intelligent Material Systems and Structures*, **20**, 849–863 (2009).
- M. J. Whelan and K. D. Janoyan, "In-Service Diagnostics of a Highway Bridge from a Progressive Damage Case Study," *J. of Bridge Engineering*, **15**(5), 597–607 (2010).
- M. J. Whelan, M. V. Gangone, K. D. Janoyan, and R. Jha, "Operational Modal Analysis of a Multi-span Skew Bridge Using Real-time Wireless Sensor Networks," *J. of Vibration and Control*, **0**(0), 1–12 (2010).
- K. Worden and G. Manson, "The Application of Machine Learning to Structural Health Monitoring," *Phil. Trans. R. Soc. A*, **365**, 515–537 (2007).
- H. Wu, D. Zonta, M. Pozzi, P. Zanon, and M. Corra, "Historic Buildings: Long Term Stability Evaluation Using Wireless Sensor Networks," *Advanced Materials Research*, **133–134**, 235–240 (2010).
- D. Zhu, X. Yi, Y. Wang, K. M. Lee, and J. Guo, "A Mobile Sensing System for Structural Health Monitoring: Design and Validation," *Smart Materials and Structures*, **19**, 055011–055021 (2010).
- ZigBee Alliance. <http://www.zigbee.org/> (2011).

APPENDIX A. SENSING AND NETWORK CHARACTERISTICS

APPENDIX A. SENSING AND NETWORK CHARACTERISTICS

A.1 SENSING CHARACTERISTICS

Beyond the fundamental characteristics, a SHM system can be characterized by how it senses, communicates, and processes data. In terms of hardware, many systems, such as those outlined in Lynch and Loh, (2006), utilize common sensing and wireless communication platforms, such as the MICAz, Moteiv Tmote Sky, or Imote2 wireless modules. All three of these types of modules include processing capabilities with some form of TinyOS operating system, are equipped with flash memory, communicate wirelessly via the IEEE 802.15.4 standard (discussed below), and require external sensors to actually obtain the acceleration data. These units differ in terms of specifications like range and memory capacity (Memsic; Moteiv 2006; Crossbow 2007). Lynch and Loh (2006) give a tabulated summary of the most popular academic and commercial wireless modules from 1998–2005. These tables include processor models, bus size, clock speed, and program and data memory for each of these prototypes. Bilbao et al. (2011), and Aygun and Gungor (2011) also discuss other modules.

A very important characteristic is the type of data being collected, such as acceleration, strain, tilt, impedance (Overly et al. 2008), temperature (Harms et al. 2009b), and others (Perez, et al. 2010; Lynch and Loh 2006), including fiber optic sensors (Todd et al. 2007, Chang et al. 2001). One example of particularly intensive sensing is that of the new I35 bridge in Minneapolis, Minnesota, which was built to replace the I35 bridge that collapsed in 2007 (French et al. 2011). The reinforced concrete bridge is instrumented with 219 strain gauges, 12 displacement sensors, 243 temperature sensors, 26 accelerometers, 12 potentiometers, and 4 corrosion sensors (French et al. 2011).

Another significant characteristic of every WSN is its power source. Many run on common batteries, but alternative energy (energy harvesting) systems have been proposed (Park et al. 2007). Energy harvesting is of particular interest for WSN on civil structures because sensors are often placed in locations that are difficult and potentially dangerous to access, thus raising the cost of maintenance. Closely related to this is the issue of energy conservation, which Park et al. (2007) also addresses in detail in terms of the sensor's active/inactive cycle. Many researchers, such as Hackmann et al., Chebroly et al. (2008), Kim et al. (2007), Galbreath et al. (2003), and Stoianov et al. (2007), consider the active/inactive, or "sleep" cycle, in the context of their specific systems. Among acceleration sensing systems, Harms, et al. (2008) and Harms et al. (2009) address power consumption particularly aggressively.

A non-traditional hardware characteristic is mobility [Zhu et al.; Taylor et al. (2009)], involving a robotic device that travels to different parts of the structure. Another sensor feature is the ability to conduct self-assessment, checking its own health status. This feature has been demonstrated for piezoelectric sensors, as investigated by Park et al. (2006). Sensor location is one of the primary issues that we address in this paper, and is discussed in more detail below.

A number of software applications and systems have been proposed for use in conjunction with WSN for SHM. Of most interest are those that are embedded in the sensor (or network) itself, as opposed to those that intended only for use on a personal computer (PC). As this is a complicated topic, please note that the following is intended only to illustrate various software systems proposed, not to give a thorough review of work done in this area.

One of the most fundamental aspects of any embedded processing system is the operating system used. As mentioned before, TinyOS is a popular open-source system (TinyOS). It supports the use of a language called nesC, which is a dialect of C, meaning that any syntax that is acceptable in C is also acceptable in nesC (TinyOS). Dunkels et. al (2004) present an operating system called Contiki for WSN, which also uses C. One of its primary features is efficient in-service software updating capability. Farooq and Kunz (2011) give an in-depth review of operating systems for WSN, including the two mentioned above. However, one that they do not discuss Simple Operating System (SOS), presented by Han et al. (2005), whose design focuses on the updating capability similar to Contiki.

A popular type of embedded software for SHM is the agent-based one. This approach was proposed in detail by Ruiz-Sandoval (2004). Agent software is “a computer system, situated in some environment, that is capable of flexible autonomous action in order to meet its design objectives,” (Jennings et al. 1998). Zhao et al. (2009), Chen and Liu (2011), and Smarsly et al. (2011) also address agent-based systems.

Allen (2004) proposed a system, in which the embedded software could be changed by a remote user. Thus, the type of processing, or the parameter of an algorithm, could be changed without the need to physically uninstall the wireless module. This system is integrated with non-embedded MATLAB-based software called DIAMOND II, developed at Los Alamos National Laboratory (LANL).

Paek et al. (2010) discuss the Tenet architecture. Tenet runs on the TinyOS platform and is based on and intended for two-tiered systems, as used in Paek et al. (2006).

A.2 NETWORK AND COMMUNICATION CHARACTERISTICS

One of the most important characteristics of a WSN in terms of communication is what data is being communicated. In a distributed processing system, the data analysis takes place at the node itself and the results are communicated back to the base station. In a central processing system, the raw data is communicated to and then processed by the base station. Many systems lie on the continuum between these extremes. Another characteristic, network topology, refers to the configuration of the nodes with respect to the communication connections that they have to each other. Three major topologies are utilized in WSN for SHM. The star topology is composed of a single central node, to which all other nodes communicate directly. Whelan and Janoyan (2010) used a star topology to test a 137-ft. long bridge. Second, the multi-tier topology has multiple clusters of nodes, each with a star topology, whose central nodes are connected. Ruiz-Sandoval (2004) proposes a multi-tier system that is established by the central node. Paek et al. (2006) uses a two-tier topology with 5 central nodes, which they call “Masters” and 20 nodes which they call “motes.” Third, the peer-to-peer topology is used when identical nodes pass information to each other for ultimate receipt by a central node. Pakzad et al. (2008) used a peer-to-peer topology for an experimental implementation on the 1280-m main span of the Golden Gate Bridge. This nomenclature is borrowed/adapted from Lynch and Loh (2006). However, some researchers [including Paek et al. (2006) and Pakzad et al. (2008)] refer to a linear topology, which is, effectively, one of the above topologies with the sensors arranged in (spatially) linear manner.

The communication protocol refers to the set of routines by which nodes pass information to each other within the specific topology. IEEE standards establish the basic components for these protocols, but leave room for custom protocol arrangements. For example, Whelan and Janoyan (2009) used the Tmote Sky platform, which had a Chipcon CC2420 transceiver, operating under the IEEE 802.15.4 standard, but they developed their own sophisticated communication protocol and achieved excellent transmission success rate

One common communication system that uses the peer-to-peer topology is called ZigBee, which is popular for many WSN applications, including SHM. Zigbee is based on the IEEE 802.15.4 protocol (Zigbee). WSN proposed by Harms et al. (2009) and Bocca et al. (2009) use ZigBee. It is possible to use the 802.15.4 standard without ZigBee (Whelan and Janoyan 2009).

The IEEE 802.15.4 standard is part of the IEEE standards from 802.11 through 802.16 governing wireless networks (Taylor et al. 2009), include WiFi networks and the Bluetooth communication system. Taylor et al. (2009) in their mobile-agent node system, use the 802.11 standard for communication between a base station and the mobile-agent, and Zigbee for communication between the mobile agent and the sensors. Paek et al. (2006) use 802.15.4 standard for their motes and 802.11 for the Masters, since the 802.15.4 is much more energy-efficient (Lynch and Loh 2006).

A significant consideration in network design is the local transmission range of the nodes. This consideration becomes more important when applying WSN to very large (bridge) structures. Paek et al.

(2006) note that their 802.11 standard radio has a range of over 120 ft. when equipped with an antenna. Kim et al. (2007) report up to 175-ft. range between motes using MICAz motes [which are IEEE 802.15.4 transceivers (Hill et al. 2004)] with bi-directional antennas. However, Kim et al. (2004) also report that another group of MICAz motes had as low as a 50-ft. transmission range.

While communication between sensors within a structure is important, remote communication with decision makers is also important, especially if the structure is in imminent danger of collapse. Harms et al. (2009) address this issue by equipping the base station of their system with GSM (Global System for Mobile Communications, or Groupe Spécial Mobile) connection capability, allowing it to send emails and text messages to appropriate personnel. Sridhar et al. (2009) also explored remote communication via the GSM network. Two-way communication is discussed in the context of software in the main body of this paper.

APPENDIX B. EXPERIMENTAL SHM SYSTEMS

APPENDIX B. EXPERIMENTAL SHM SYSTEMS

The particular application of interest for this paper is that of bridges. A number of complete systems have been proposed and tested for SHM of bridges. We have mentioned many of them above in our outline of SHM system characteristics. A few of these warrant particular attention, and thus we will discuss them in more detail, with specific attention to their architectures.

Paek et al. (2006) present a new architecture developed around software the TENET architecture. This system utilizes a tiered topology. Specifically, the Intel Stargate platform was used for the “Master” (upper-tier) nodes and the MICAz platform was used for the motes (lower tier nodes). Master nodes have IEEE 802.11x standard radios, and motes have IEEE 802.15.4 standard radios, which allow the motes to consume less energy. The Masters have a 5-dBi antenna and were placed approximately 120 ft. apart. Each master had a cluster of four motes spaced 30 ft. apart. The system had no remote communication capability, requiring manual retrieval of data from the Masters.

Mechitov et al. (2006) propose certain improvements on previous systems. They use a self-healing peer-to-peer topology, optimized on the basis of path length and secondarily on link quality. The Mica2 wireless module and embedded 433 MHz RF transceiver are used for communication. For sensing, a Silicon Designs model 1221 is used for adequate A-to-D precision. The Mica2 can use TinyOS, though the authors made no mention of it.

Chebrolu et al. (2008) developed a wireless sensor network system for railway bridge monitoring in India. An ad hoc network was used, with a single master node for each span of the bridge. Routing was accomplished by a “neighbor discovery phase,” in which each node establishes a connection with its neighbors, followed by a “tree construction phase,” in which the routing tree is built, starting from the master node, based on quality of connections. Each of these master nodes deposits information directly on the passing train. Tmote Sky wireless modules, with 8 MHz processors running TinyOS 2.0, are used, along with IEEE 802.15.4 radios, and triaxial MEMS-based ADXL accelerometers.

Pakzad et al. (2008) discuss implementation of an SHM system on the Golden Gate Bridge. A linear topology was used, which required multihop communication. The communication protocol was managed by certain routines in TinyOS operating system. The wireless module used was a MicaZ which has 512 kB of flash memory, 8 MHz clock speed processor, and a 2.4 GHz Chipcon CC2420 transceiver (IEEE 802.15.4 standard), to which the designers added a bi-directional antenna. Dual accelerometers are used: ADXL202 for high-amplitude signals; and Silicon Designs 1221L for low-amplitude.

Whelan and Janoyan (2009) discuss development and testing of a system, called the “Wireless Sensor Solution” (WSS) on a small beam bridge in New York State. The network is set up as a star system with each node communicating directly with the base station. Communication takes place via 2.4GHz Chipcon CC2420 transceivers and a complex protocol system, which includes acknowledgement from the central node to the peripheral ones. WSS uses the Moteiv Tmote Sky wireless module. The dual-axis accelerometer is a LIS2L02AL made by STMicroelectronics. TinyOS 1.x was used, but was supplemented with additional software written by the authors for hardware interfacing. The WSS underwent further testing; see Whelan and Janoyan (2010), Whelan et al. (2010) and Gangone et al. (2011).

Taylor et al. (2009) propose a mobile sensing system. The topology of the network is tiered, where many sensors connect to central computing/processing stations, multiple computing/processing stations connect to power/communication stations, and each of the power/communication stations connect to a mobile unit, when it arrives, via a very short-range wireless connection. Both data and energy are transmitted through this network. Impedance data are gathered by special sensor nodes.

Harms et al. (2008, 2009a, 2009b) present the development of a wireless sensor platform called “SmartBrick.” It is intended to be more energy-efficient than its predecessors, with a five-year life projected in Harms et al. (2009b). Two types of units are used: base stations; and sensor nodes. As their names imply, the sensor nodes gather data and transmit it to the base station for processing via the Zigbee system. The base stations process the data on a 16-bit microcontroller unit (MCU). The units are very

similar except that the base station includes a modem that allows communication via email, text message, or FTP. Using the FTP, a web interface has been developed for remote visualization of data. In addition to vibration sensing via a triaxial MEMS (microelectromechanical system) accelerometer, the unit also senses temperature and tilt. The SmartBrick has its own (crude) analysis software.

Bilbao et al. (2011) present a system based on custom-built wireless modules, with different designs for nodes and the base station. Connected in a peer-to-peer topology, they use the Digimesh communication system, which is based on the IEEE 802.15.4 standard, along with XBee communication hardware. For sensing, an ST Microelectronics triaxial accelerometer is used, and processing takes place on a Texas Instruments microcontroller. Similar to the concept discussed by Allen (2004), the researchers plan to develop this system to be capable of remote communication and control [via TCP/IP through a GPRS (General Packet Radio Service) device], and thus is easily adapted to new software.

APPENDIX C. TECHNICAL DETAILS

APPENDIX C. TECHNICAL DETAILS

C.1 MIGRATING THE ALGORITHM

Before ORNL's failure forewarning algorithm could be implemented in a new complex structure sensor environment, one preliminary task was to ensure that the FORTRAN code would compile and run with open-source GNU compilers. The code, written by Lee Hively (ORNL), had originally been written and compiled with the Compaq Visual FORTRAN compiler on the Windows XP operating system. Because of the volume of input data, an auxiliary code had been written to convert the many files of

ASCII data into one large binary file for faster reading and computation. However, the Compaq compiler's statements for opening and reading binary files contained a non-standard tag* that was not available with the GNU gfortran compiler.

The code's open statements were modified to use standard tags to read the binary file (FORM='UNFORMATTED'). However, the code was still producing I/O errors when attempting to read from the binary file. Rewriting the binary file with a version of the auxiliary code that had been recompiled with gfortran also resulted in the same runtime errors. It was subsequently discovered, through trial and error, that minor syntactic changes were creating bugs which hampered the code in its current running environment [the dynamic-link library (DLL) Cygwin 1.7.9-1].† Once those bugs were removed, the program correctly read the data and delivered the same output produced with the Compaq compiler under Windows. The next step was to create a virtual machine to run the new application in a portable environment. Ubuntu 10.10 was selected as the operating system.

VMWare Player 3.1.4 was installed to run Ubuntu 10.10 on the same machine that had been used to compile and run the code with the DLL. With a few minor exceptions (e.g., case sensitivity in the data file paths, and formatting in the read statements), the final version of the code which had compiled and run successfully with Cygwin, also ran successfully under Ubuntu 10.10. Once the code was reading and processing the example data correctly in the target OS, the code could then be executed with any system capable of running the virtual machine environment.

C.2 UTILIZING THE VIRTUAL MACHINE

To demonstrate the scalability of the failure forewarning algorithm, the research team needed access to a cluster of processors where multiple instantiations of the forewarning algorithm could be run simultaneously, as well as the computing and storage capacity for an increasing volume of input data. For guaranteed compatibility in the code's compilation and execution, a virtual machine was created to provide a consistent running environment for the code, regardless of the physical hardware available—whether in laboratory conditions, or within the central processor of a wireless sensor on a bridge pier. Since a virtual machine, or software version of a physical computer, can run completely independently from the host, a virtual machine created on one computer can run as an isolated operating system on another computer. By using a virtual machine, the algorithm code could be tested on different hardware without the chance of compiler errors and inconsistencies. The decision to use a virtual machine to test the FORTRAN code would also prove effective in testing the scalability and efficiency of the failure forewarning algorithm.

*FORM='BINARY' in the OPEN statement

† In retrospect, some of the compilation and runtime errors might have been avoided by eliminating the intermediate DLL step, and initially implementing the code with the GNU compiler on a virtual machine.

By uploading the Ubuntu 10.10 virtual machine files onto an ORNL cluster computer, the virtual machine and the code could be run with various configurations. Specifically, the scalability of the code could be tested by running the virtual machine with different numbers of processors dedicated to the codes execution. The idea being tested was twofold. The first test would demonstrate whether this algorithm and its implementation in FORTRAN had been written in such a way that would result in linear scalability: as the number of processors running the code was increased, the code output produced in the same execution time would increase linearly. The second test would demonstrate that regardless of the amount of data being analyzed, the program would still maintain its scalability.

The first test was performed as follows. First, a single processor would be dedicated to the virtual machine. A preliminary run of the code would be executed, and the number of parameter sets computed would be recorded. The virtual machine would then be reconfigured to have access to successively more processors (1,2,3,4, etc.), and the process would be repeated. The code would be executed in this way until a clear trend emerged between the cumulative processing time across all CPUs and the amount of output. The second test would demonstrate the second part of scalability (the volume of data) by following essentially the same process, except this time, by increasing the amount of data given to the virtual machine along with the RAM. If the output produced still increased linearly with the amount of total processing time, regardless of the addition of multiple channels of data, then scalability would be effectively demonstrated.

Four clones of the same virtual machine, each with identical versions of the algorithm code and data input file were installed in the cluster. Each virtual machine had access to the equivalent of one 64 bit CPU with 256 GB of RAM. In addition, the codes were set to open a file to retrieve the MAC address of the machine running the code. Each virtual machine had a different MAC address and thus stored a different number, which was used as a seed in the random number generator to initiate a unique data training process in each virtual machine. A bash script was written to begin the training process on a specified number of processors simultaneously. In this way, the data training process could use multiple processors at once to cover more ground in the search for optimal phase space reconstruction parameters. After this, the second test was performed by setting the code in each virtual machine to process additional channels of input data. The reasoning behind demonstrating scalability in bridge monitoring applications can be found in the Methods section, *Demonstrating Algorithm Scalability*. See the Results section for the outcome of both tests.

APPENDIX D. TESTING THE ALGORITHM

APPENDIX D. TESTING THE ALGORITHM

D.1 STATISTICAL VALIDATION METHODOLOGY FOR VIBRATION-BASED STRUCTURAL HEALTH MONITORING

The last decade in SHM research has seen increased attention devoted to statistical validation of damage detection methods. Some of the consistent advocates of verification and validation techniques in SHM are Doebling, Farrar, and Hemez, whose collaborations with each other and other researchers since the year 2000 have emphasized the traditional lack of statistical validation in SHM and its importance in future research (Doebling et al. 2000; Farrar et al. 2001; Hemez et al. 2005; Farrar and Lievan 2007). These researchers' use of statistical validation deals mostly with improving the accuracy of laboratory- and computer-based structural models in imitating experimental vibration data. However, Doebling et al.'s (2000) definition of validation as "the evaluation of the accuracy of a computational prediction with respect to experimental data" is applicable to determining the reliability of a data-based algorithm to detect and predict real-life damage response.

Statistical validation of the results of a particular damage detection method is not to be confused with the use of statistical techniques within the algorithm itself. In vibration-based SHM analysis, the basis of damage detection is usually a statistical comparison of test data to some threshold determined from the baseline data, whether the damage feature is a statistical measure or correlation calculated from the vibrational data itself (as in traditional SHM methods), a measure of the amount of nonlinearity in the data (as in Adams 2002), or the degree of error in the model's predicted response (as in Sohn and Farrar 2001 and Nichols et al. 2003). In the first case, the role of statistical validation is to evaluate model reliability by testing the distributions of simulated vibrational response as compared to experimental data (when the original baseline and test data were simulated from a model) or by testing the effectiveness of the damage detection and decision tools with new or subsequent real-life data (when the analysis is data-based). The use of "conventional statistical methods" (CSM) to identify and quantify damage based on statistical measures (mean, standard deviation, skewness, and kurtosis) of the accelerometer data itself, is typically only effective in analyzing linear processes (Protopopescu and Hively 2005).

When using linear modal or frequency response models to simulate or analyze vibrations, nonlinearities are considered a hindrance to the analysis. However, nonlinearities are in fact an inherent characteristic of damage growth and detection (Adams 2002; Protopopescu and Hively 2005), and can therefore serve as a tool in SHM. Consequently, more recent SHM research has begun to account for nonlinearities in vibrational data (Adams 2002). Here, the damage feature (and the role of statistical testing and correlations) can become either the amount of nonlinearity in the system, or the ratio of linear to nonlinear behavior (as in Adams 2002). As has been previously pointed out (Protopopescu and Hively 2005), algorithms which identify and quantify nonlinearities provide a good contrast between linear and nonlinear systems, but do not characterize it by effectively measuring changes between nonlinear systems. As mentioned before, the weakness of "traditional nonlinear measures" (TNM) in characterizing nonlinear processes has to do with the fact that TNMs are based upon differences of averages, whereas the PSDMs used in the current approach are based upon sums of the absolute values and squares of differences.

Phase-space time-serial analysis presents an effective alternative to linear methods and to the analyses which simply identify or quantify nonlinearity. In phase-space analyses, the damage features can be produced from performing analyses of attractor geometries (Torkamani 2010), or (as in the current method) calculating statistical measures from the distribution on the phase-space attractor, itself, rather than from the vibrational data. This is a relatively new avenue in SHM where some researchers, such as Nichols, Todd, and Trendafilova, have begun to venture.

Statistical analysis plays more than one role in vibration-based SHM. First, statistical measures are an inherent part of damage detection algorithms and decision-making tools—be they linear or nonlinear,

physics-based or data-based models. Whether the damage feature is a statistical measure or correlation calculated from either the vibrational data (as in traditional SHM methods) or on the phase-space attractor (as in this analysis); a measure of the amount of nonlinearity in the data (e.g., Adams 2002); or the degree of error in the model's predicted response (e.g., Sohn and Farrar 2001 and Nichols et al. 2003); the basis of damage detection is usually a statistical comparison of test data to some threshold determined from the baseline data. However, the use of statistical techniques within the damage detection algorithm, itself, is not to be confused with statistical validation of the results of a particular SHM method. After the preliminary data have been analyzed and some decision-making algorithm is in place, confidence in the damage detection method can be strengthened with statistical validation of the algorithm: 1) reliability when faced with subsequent data of the type that it is likely to encounter; and 2) applicability to the real-life phenomena which it was designed to analyze.

The reliability of a particular damage detection method can be measured with a variety of techniques, not limited to: examining and comparing Type I and Type II error rates (false positives and false negatives in damage detection); and quantifying uncertainty with indicators such as confidence intervals, regression correlation coefficients, or a series of hypothesis tests. Statistical validation showing applicability to real phenomena must be carried out differently for the various types of SHM models. For example, the growing use of statistical methods to test structural simulation models involves comparing the distributions of the model-generated vibrational data against the distributions of real, experimentally-collected data. Such models have no shortage of simulated data of any quality and frequency, but they have the burden of demonstrating that their methods' effectiveness in detecting damage in computerized models will translate to effectiveness in detecting damage in a real structure. For a data-based SHM algorithm, however, the algorithm and decision-making technique have been developed based on real data, so the task becomes gauging the likelihood that the indications they give are not somehow a predestined result of the data that were used to develop the algorithm. Phase-space analysis requires high-quality, high-frequency data in order to train the algorithm to a structure's specific vibrational signature. Most researchers have no such priorities when collecting accelerometer data. Even more confounding: data from real-life structure damage can only be collected once. The somewhat circular problem for data-based time-serial analysts, then, becomes statistically testing the accuracy of the damage detection algorithm, when the data available for algorithm testing are the same data that were used for training it.

ORNL's forewarning algorithm has achieved high success rates in predicting events in many applications. In a 2005 paper, the algorithm was used to analyze EEG data to predict seizures. The algorithm had a sensitivity of $39/40$ [$TP/(TP + FN)$], and a specificity of $17/20$ [$TN/(TN + FP)$]. However, even after such results, it was clear that the quality of the algorithm as a failure forewarning system was still contingent upon the nature of the forewarning times the algorithm was producing. Thus, it was necessary to show that the distribution of forewarning times did not represent a cumulative distribution of random, independent events. In this case, the algorithm would be no more likely to predict an event at any point in the time series before its occurrence, and the length of forewarning times would essentially be a matter of chance. In the end, the data did deviate significantly from the exponential distribution, and the null hypothesis was rejected. More details on the test, and how it was selected and implemented, are given below.

The failure forewarning mode of the algorithm is not utilized in this paper; however, in order to build upon previous applications and show the methodology for future applications in SHM, a brief statistical validation of previous failure forewarning results is performed and its results are summarized in the section below. The data analyzed are forewarning times of seizure events produced from the same EEG data used in (Hively et al. 2005). The purpose of the statistical validation is to demonstrate whether the forewarnings produced by the algorithm are simply governed by chance, or if they constitute some consistent prediction of epileptic events. Only the datasets which contain correctly-predicted seizure events are considered (39 out of the 40 samples containing a seizure event).

Two hypothesis tests are performed on the failure forewarning times. The first is an Anderson-Darling test of uniformity. If the probability of receiving forewarning were uniformly likely throughout

the pre-seizure period—that is, the occurrence of failure forewarning were independent of time before the seizure (any timestep had a $1/T$ chance of containing the forewarning signal, where T is the length, in s, of the pre-seizure period)—then the distribution of forewarning times would follow a uniform cumulative distribution function. Therefore, if the null hypothesis of this first test was rejected, then it would demonstrate strong evidence against the scenario that the 39 event forewarnings represent random, independent events occurring any randomly-selected time within the T s. However, knowing that only the first occurrence of forewarning was recorded in the 39 datasets, and therefore the fact that it was first makes it conditional upon all timesteps preceding it not containing a forewarning signal, the forewarning times are subjected to a second hypothesis test, this time against an exponential distribution. The exponential distribution corresponds to the distribution of random, independent events in time, occurring “without replacement” (that is, each event has an equal probability of occurring in any discrete, instantaneous timestep, but the probability of receiving forewarning at each subsequent timestep is a decreasing function). If this hypothesis test is rejected, this would demonstrate strong evidence against the scenario that the 39 correctly-predicted event forewarnings were distributed in such a way that is likely to have happened by chance.

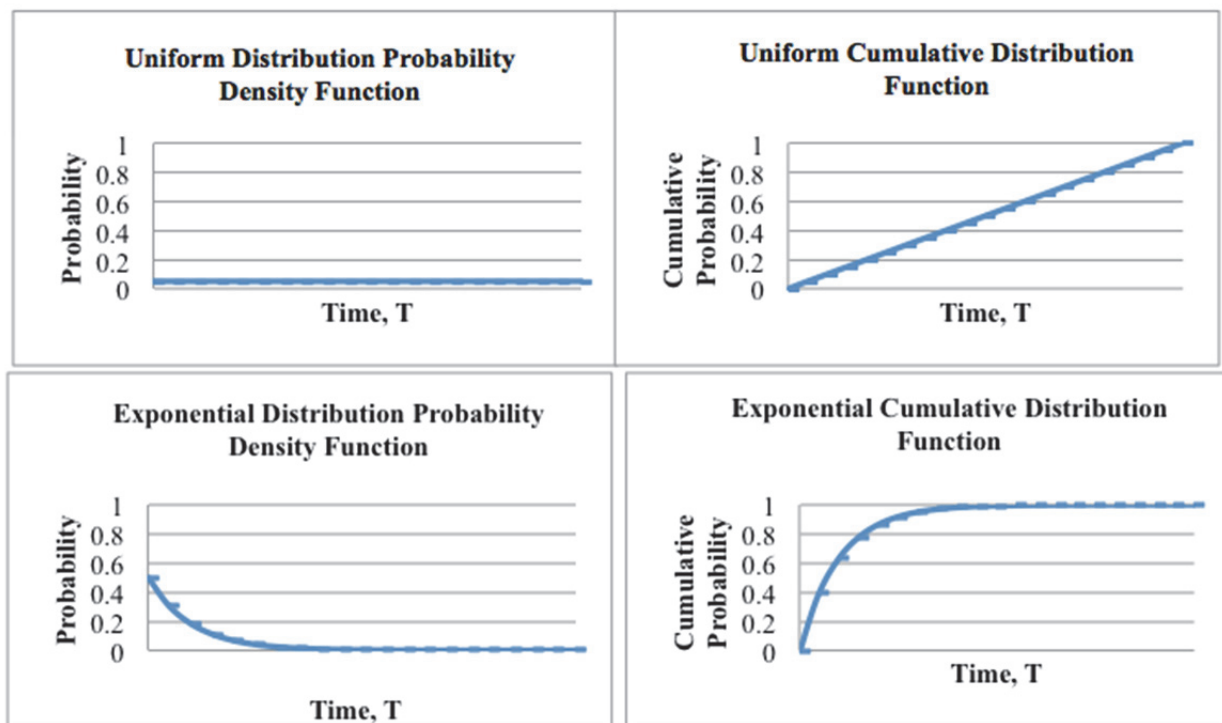


Fig. D.1. Probability Density Function and Cumulative Probability Function Shapes. *Top left:* The shape of the probability density function of a uniform distribution; *Top right:* The shape of a uniform cumulative distribution function; *Bottom left:* The shape of the probability density function of an exponential distribution; *Bottom right:* The shape of an exponential cumulative distribution function.

D.2 SELECTING A STATISTICAL TEST

There are several statistical tests which can be used to test data against a selected distribution. With a sample size of over 35, as was available with the EEG data, there were multiple tests which could have produced reasonably reliable results. The most common in the relevant literature and the most straight-

forward tests to implement, were investigated for their testing power. Below is a summary of the ultimate decision to implement the Anderson-Darling test.

Although perhaps the simplest and most common method for testing various distributions, the Chi-squared test is often shunned by researchers due to its potential decrease in accuracy when bin frequencies (when data are divided into ranges or ‘bins’) fall below a minimum number, usually accepted to be five. However, some research argues that Chi-squared tests with bin frequencies of less than five are not significantly less accurate than those with frequencies greater than five (Roscoe 1971; Miller 1975). According to this research, test reliability depends more on total sample size than on bin frequency, with minimum bin frequency being more a rule of thumb than empirical fact (Roscoe 1971). Like all of these tests, the accuracy of Chi-squared depends on which distribution is being tested, and it is more powerful for some of these data shapes—even with low bin frequencies—than many other tests (Miller 1975).

The Kolmogorov-Smirnov (or KS test) is a widely-known alternative to the Chi-squared test, and is a commonly-accepted method for avoiding the low bin frequency problem in Chi-squared tests (Roscoe 1971; NIST 2011). This acceptance is because KS uses the cumulative distribution function (CDF) instead of the probability distribution function (PDF), and therefore eliminates the need for bins altogether. The KS p-value tables are identical, regardless of the distribution being tested, which allows for simpler acquisition of the appropriate critical values, but sacrifices some testing accuracy (NIST 2011; ERI 2011). With some distributions, KS does not perform any better than Chi-squared (Roscoe 1971), and it is in fact one of the lowest-performing tests when sample sizes are lower (below around 30) (Miller 1975; Marhuenda 2005).

Some other well-known goodness-of-fits tests for uniformity include Anderson-Darling (AD), Shapiro-Wilk, and Watson (to name a few), which tend to be more accurate than the KS or Chi-squared tests (Cheng and Spiring 1987) (especially with smaller sample sizes and when testing for certain distributions) (Miller 1975; Marhuenda 2005). The power of each test also varies with the shape of the data. Out of these, the most sensitive test when testing centrally-concentrated data for uniformity is Watson (Miller 1975; Marhuenda 2005). The most consistently sensitive test for most other shapes is Anderson-Darling (Miller 1975).

The Anderson-Darling test also uses CDF to calculate test statistics, and is known to be more sensitive to variance in the tails of the data than the KS test (NIST 2011). With the standard AD test, a good model for p-value with respect to sample size has not been found, and consequently AD tables tend to use experimentally-determined mean values (one for each percentile, independent of sample size) (Rahman 2006). Therefore, once a distribution and alpha level are selected, only one single value is needed to test for significance. Although these critical values are based on means, they are produced from very close values, and they are more accurate for having been calculated multiple times, and with separate tables generated specifically for each distribution type.

With higher sample sizes, any one of these tests performs reasonably well (Miller 1975). However, AD was chosen because it is a well-known test with consistently high power for most data shapes (Miller 1975; Marhuenda 2005). Especially appropriate for these data, AD is designed to give more weight to tail data than KS (NIST 2011). Since the cumulative shape of the data is particularly non-linear in the right tail (Fig. D.2), the AD test for exponential distribution was an appropriate choice.

D.3 TESTING AND RESULTS

An auxiliary code to perform the AD test for exponential distribution was written in FORTRAN. It was to then be included as a subroutine in the larger algorithm. Fig. D.2 shows the distribution of the EEG data, compared to the theoretical exponential distribution. Clearly, as the rejection of the null hypothesis suggests, it is unlikely that the data come from an exponential distribution. With an AD statistic of 4.7089, the p-value is less than 0.001. This is fairly conclusive evidence that the forewarning

times are not governed by chance. Test output is provided below.

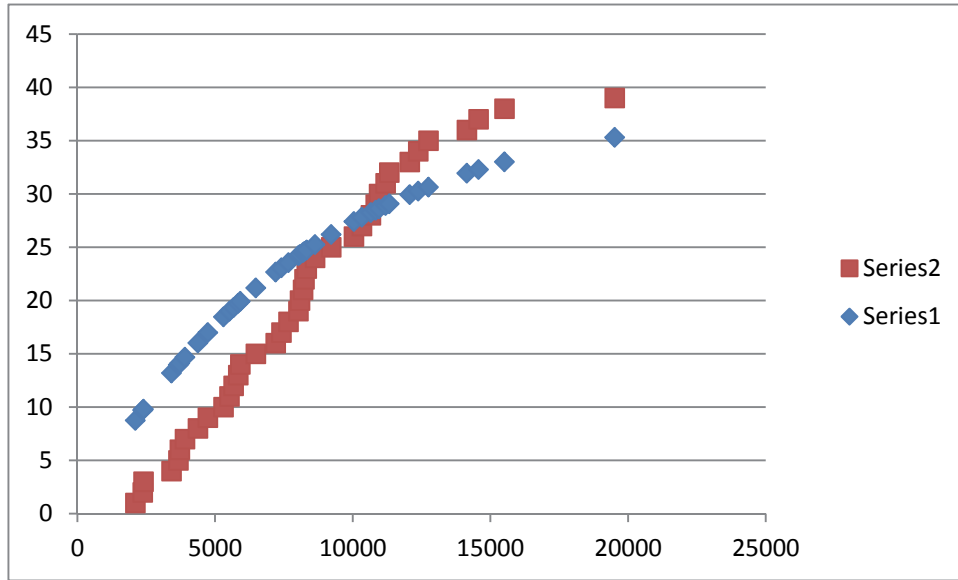


Fig. D.2. Cumulative frequency of actual data (series 2) vs theoretical exponential distribution (series 1).

(Series 1)

Sample size: 39

The number of true negatives is: 17

The number of false positives is: 3

The number of false negatives is: 1

H0: The data follow an exponential distribution

H1: The data do not follow an exponential distribution

The Anderson-Darling test statistic is: 4.70893. The corresponding probability value for accepting the null hypothesis is $p < 0.001$, which provides strong evidence to reject H0.

0.1000 0.6370

0.0500 0.7570

0.0250 0.8770

0.0100 1.0380

APPENDIX E. SHM RESEARCHERS CONTACTED FOR DATA

APPENDIX E. SHM RESEARCHERS CONTACTED FOR DATA

Contact Name	Email address	Institution	Status
Dr. Kerop Janoyan (and formerly Dr. Matthew Whelan, who is now at UNC-Charlotte)	kerop@clarkson.edu	Clarkson University	After contact by email and phone conversation, they wanted some documentation, so we sent a couple reports but have not heard back
Dr. Erin Bell	erin.bell@unh.edu	University of NH	Responded, had no applicable data
Dr. Ziyad Duron	zee_duron@hmc.edu	Harvey Mudd College	No response
Dr. Rakesh Gupta	rakesh.gupta@oregonstate.edu milo.clauson@oregonstate.edu	Oregon State University	Corresponded with Dr. Gupta and Milo Clauson (lab technician), but awaiting response
Mr. Vicente Suarez	Vicente.J.Suarez@nasa.gov	NASA Structural Dynamics Laboratory	No response
Dr. Peter Avitabile	Peter_Avitabile@uml.edu	University of Massachusetts Lowell	Responded, had no applicable data
Dr. Michael Todd	mdtodd@ucsd.edu	University of California San Diego	Responded, said that he might have something, but never got back
Dr. John DeWolf	dewolf@engr.uconn.edu	University of CT	Responded, but the data he had was not applicable
Dr. Ricardo Medina	ricardo.a.medina@unh.edu	University of NH	Responded, said he might have ideas, but never got back
Dr. Jim Beck	jimbeck@caltech.edu	CalTech	Referred us to NEES website : http://stage.neeshub.org/dataview/spreadsheet/scmbp
Dr. Bo Chen	bochen@mtu.edu	Michigan Tech	No response

Dr. Jason Ingham	j.ingham@auckland.ac.nz	University of Auckland	Referred me to Lucas Hogan (PhD student), who was unable to give us data for intellectual property reasons (publishing of original research before dissertation)
Dr. Chenyang Lu	lu@WUSTL.EDU	Washington University in St. Louis	Referred me to Shirley Dyke at Purdue (below)
Shirley Dyke	sdyke@purdue.edu	Purdue University	Referred us to database: https://nees.org/resources/1687
Dr. Antonino Morassi	antonino.morassi@uniud.it	University of Udine, Italy	Responded and said that only the frequency response functions had been kept, not the time serial data
Mark Rumsey and Joshua Paquette	marumse@sandia.gov and japaque@sandia.gov	Sandia National Laboratory	No response