Luke Bornn
CCS-6,
Statistical Sciences Group,
Los Alamos National Laboratory,
MS F600,
Los Alamos, NM 87545

Charles R. Farrar[1]
e-mail: farrar@lanl.gov

Gyuhae Park

Kevin Farinholt

The Engineering Institute,
Los Alamos National Laboratory,
MS T006,
Los Alamos, NM 87545

Structural Health Monitoring With Autoregressive Support Vector Machines

The use of statistical methods for anomaly detection has become of interest to researchers in many subject areas. Structural health monitoring in particular has benefited from the versatility of statistical damage-detection techniques. We propose modeling structural vibration sensor output data using nonlinear time-series models. We demonstrate the improved performance of these models over currently used linear models. Whereas existing methods typically use a single sensor's output for damage detection, we create a combined sensor analysis to maximize the efficiency of damage detection. From this combined analysis we may also identify the individual sensors that are most influenced by structural damage. [DOI: 10.1115/1.3025827]

1 Introduction

The extensive literature on structural health monitoring (SHM) has documented the critical importance of detecting damage in aerospace, civil, and mechanical engineering systems at the earliest possible time. For instance, airlines may be interested in maximizing the lifespan and reliability of their jet engines, or governmental authorities might like to monitor the condition of bridges and other civil infrastructures in an effort to develop cost-effective lifecycle maintenance strategies. These examples indicate that the ability to efficiently and accurately monitor all types of structural systems is crucial for both economic and life-safety issues. One such monitoring technique is vibration-based damage detection, which is based on the principal that damage in a structure, such as a loosened connection or crack, will alter the dynamic response of that structure. There has been much recent work in this area; in particular, Doebling et al. [1] and Sohn et al. [2] presented detailed reviews of vibration-based SHM. Because of random and systematic variability in experimentally measured dynamic response data, statistical approaches are necessary to ensure that changes in a structure's measured dynamic response are a result of damage and not caused by operational and environmental variabilities. Although much of the vibration-based SHM literature focuses on deterministic methods for identifying damage from changes in dynamic system response, we will focus on approaches that follow a statistical pattern recognition paradigm for SHM [3]. This paradigm consists of the four steps: (1) operational evaluation, (2) data acquisition, (3) feature extraction, and (4) statistical classification of features. The work presented herein focuses on steps (3) and (4) of this paradigm.

One approach for performing SHM is to fit a time-series predictive model such as an autoregressive (AR) model to each sensor output using data known to be acquired from the structure in its undamaged state. These models are then used to predict subsequent measured data, and the residuals (the difference between the model's prediction and the observed value) are the damage-sensitive feature that is used to check for anomalies. This process provides many estimates (one at each time step) of a single-dimension feature, which is advantageous for subsequent statistical classification. The logic behind this approach is that if the model fit to the undamaged sensor data no longer predicts the data subsequently obtained from the system (and hence the residuals are large and/or correlated), there has been some sort of change in the process underlying the generation of the data. This change is assumed to be caused by damage to the system. These linear time-series models have been used in such a damage-detection process that include applications to a wide range of structures and associated damage scenarios including cracking in concrete columns [4,5], loose connections in a bolted metallic frame structure [6], and damage to insulation on wiring [7]. However, the linear nature of this modeling approach limits the scope of application and the ability to accurately assess the condition of systems that exhibit nonlinearity in their undamaged state. In this paper, we demonstrate how support vector machines (SVMs) may be used to create a nonlinear time-series model that provides an alternative to these linear AR models.

Once a model has been chosen and the predictions from this model have been compared with the actual sensor data, there are several statistical methods for analyzing the resulting residuals. Sequential hypothesis tests, such as the sequential probability ratio test [6], may be used to test for changes in the residuals. Alternatively, statistical process control procedures, typically in the form of control charts, may be used to indicate abnormalities in the residuals [4]. In addition, sliding window approaches look at the features of successive subsets of data to detect anomalies [7]. For example, the sliding window approach of Ma and Perkins [8] looks at thresholds for the residuals such that the probability of an undamaged residual exceeding this threshold is 5%. A subset of n consecutive data points are then checked, and large values of the number g of points exceeding the threshold indicate damage, where g has a binomial distribution (i.e., $g \sim \text{bin}(n, 0.05)$).

To date, most of these time-series modeling approaches analyze data from one sensor at a time, and typically some sort of scheme is used to determine how many sensors need to indicate damage in order to trigger a system check [9]. As an alternative, in this paper we look at a statistically based method for combining multiple sensor output. From this combined output analysis, we can establish the existence of damage and also determine which sensors are contributing to the anomalous readings in an effort to locate the damage within the sensor network's spatial distribution. Previously Sohn et al. [5] used principal component analysis to com-

bine data from an array of sensors, but this study only examined these combined data in an effort to establish the existence of damage.

We first present a summary of the SVM approach to nonlinear time-series modeling. This procedure is illustrated on numerically generated data with artificial anomalies added to the baseline signal in an effort to simulate damage. This time-series modeling approach is then compared with linear AR models. Next the SVM method is coupled with a statistical analysis procedure that combines modeling results from multiple sensors in an effort to both establish the existence and the location of the damage. This procedure is applied to data from a laboratory test structure with damage that results in local nonlinear system response.

2 SVM-Based SHM

Existing methods for performing damage detection extract damage-sensitive features from data acquired on the undamaged system, and then use changes in those features as an indicator of damage. An AR model can be fit to the undamaged sensor output, and the residuals from predictions of subsequent data using this baseline model are then monitored for statistically significant changes that are assumed to be caused by damage. Specifically, an AR model with p autoregressive terms, AR(p), applied to sensor k may be written as

$$x_t^k = \sum_{j=1}^{p} \beta_j^k \cdot x_{t-j}^k + \varepsilon_t^k \tag{1}$$

where x_t^k is the representation of the measured signal at discrete times t from the kth sensor, β_j^k are the AR coefficients or model parameters, and ε_t^k is an unobservable noise term. Thus an AR model works by fitting a simple linear model to each point with the previous p observed points as dependent variables. Note that an n point time-series will yield $n-p$ equations that can be used to generate a least square estimate of the AR coefficients or the Yule–Walker method can be used to solve for the coefficients [10]. Autoregressive models work particularly well when modeling the response of linear, time-invariant systems. If the undamaged system is nonlinear, the AR process gives the best linear fit to the measured response, but there is no guarantee that this model will accurately predict responses obtained when the system is subjected to other inputs.

Because of the broad array of structural health monitoring problems, employing a linear model confines the scope of problems for which the AR methodology is appropriate. We thus seek to extend the fidelity of this general damage-detection approach by employing a nonlinear AR-type model based on SVMs, which have seen widespread use in machine learning and statistical classification fields. To simplify future development, we denote the vector $\{x_{t-p}^k, \ldots, x_{t-1}^k\}$ as $\mathbf{x}_{t-p:t-1}^k$. SVMs have many features that make them a more appropriate choice for SHM based on time-series analysis. With the right settings and the appropriate training, they are able to model any nonlinear relationship between the current time point x_t^k and the p previous time points $\mathbf{x}_{t-p:t-1}^k$, they are well suited for high-dimensional problems, and the methodology is easily generalized and is highly adaptable. Although SVMs have been used for SHM before [11–15], these approaches predominantly focus on one and two class SVMs, which are used for outlier detection and group classification, respectively. Our approach is unique in its combination of support vector regression, autoregressive techniques, and residual error analysis. Thus while earlier approaches look at classifying sections of the time-series response as damaged or undamaged directly (the dependent variable being a binary indicator), our methodology works by using support vector regression to model the raw time-series data, then subsequently predicting damage by monitoring the residuals of the model. We follow the development of SVMs for regression of Smola and Schölkopf [16] and Ma and Perkins [8].

First, assume we have data from a set of K sensors and we have measurements without damage for time $t=1, \ldots, t_0$ (i.e., if there is damage, it occurs after time t_0). Next we must decide the order p of our model. There are many methods for selecting p, such as partial autocorrelation or the Akaike information criterion (AIC), which are discussed in more detail in Ref. [4]. In general, we seek the lowest order model that captures the underlying physical process and hence will generalize to other data sets. As with linear AR modeling, we create the training set on which to build our SVM-based model by using each observation as the dependent variable and the previous p observations as independent variables. Our training samples are thus $\{(\mathbf{x}_{t-p:t-1}^k, x_t^k), \quad t=p+1, \ldots, t_0\}$.

Ideally we would like to find a function $\mathbf{f}$ such that $\mathbf{f}(\mathbf{x}_{t-p:t-1}^k) = x_t^k$ for all k and $t \le t_0$. However, the form of $\mathbf{f}$ is often restricted to the class of linear functions (as is the case for AR models).

$$\mathbf{f}(\mathbf{x}_{t-p:t-1}^k) = \langle w, \mathbf{x}_{t-p:t-1}^k \rangle \tag{2}$$

where $\langle , \rangle$ denotes the dot (or inner) product and w is a vector of model parameters. This restricted form makes perfect fit of the data impossible in most scenarios. As a result, we allow the prediction using $\mathbf{f}$ to have an error bounded by ε and find w under this constraint. With the recent advances in penalized regression methods such as ridge regression and lasso, the improved prediction performance of shrunken (or smoothed) models is now well-understood [17,18]. Thus in order to provide a model that maximizes prediction performance, we seek to incorporate shrinkage on the model parameters w. Such shrunken w may be found by minimizing the Euclidean norm subject to the error constraint ε, namely,

$$\text{minimize} \quad \tfrac{1}{2}\|w\|^2$$
$$\text{subject to} \quad \begin{cases} x_t^k - \langle w, \mathbf{x}_{t-p:t-1}^k \rangle \le \varepsilon \\ \langle w, \mathbf{x}_{t-p:t-1}^k \rangle - x_t^k \le \varepsilon \end{cases} \tag{3}$$

This model relies on the assumption that a linear model is able to fit the data to within precision ε. However, typically such a linear model does not exist, even for moderate settings of ε. As such, we introduce the slack variables ξ_t^+, ξ_t^- to allow for deviations beyond ε. The resulting formulation is

$$\text{minimize} \quad \tfrac{1}{2}\|w\|^2 + C \sum_{t=p+1}^{t_0} (\xi_t^+ + \xi_t^-)$$
$$\text{subject to} \quad \begin{cases} x_t^k - \langle w, \mathbf{x}_{t-p:t-1}^k \rangle \le \varepsilon + \xi_t^+ \\ \langle w, \mathbf{x}_{t-p:t-1}^k \rangle - x_t^k \le \varepsilon + \xi_t^- \end{cases} \tag{4}$$

The constant C controls the trade-off between giving small w and penalizing deviations larger than ε. In this form, we see that only points that lie outside of the bound ε have an effect on w. Figure 1 illustrates the process graphically.

Although this optimization problem is straightforward to carry out, the extension to nonlinearity is revealed by the dual formulation. We thus proceed by constructing a Lagrange function of the above by introducing a set of dual variables.

$$L := \frac{1}{2}\|w\|^2 + C \sum_{t=p+1}^{t_0} (\xi_t^+ + \xi_t^-) - \sum_{t=p+1}^{t_0} \alpha_t^+ (\varepsilon + \xi_t^+ - x_t^k + \langle w, \mathbf{x}_{t-p:t-1}^k \rangle)$$
$$- \sum_{t=p+1}^{t_0} \alpha_t^- (\varepsilon + \xi_t^- - x_t^k + \langle w, \mathbf{x}_{t-p:t-1}^k \rangle) - \sum_{t=p+1}^{t_0} (\eta_t^+ \xi_t^+ + \eta_t^- \xi_t^-) \tag{5}$$

where the dual variables α_t^+, α_t^-, η_t^+, and η_t^- are understood to be non-negative. It can be shown that this function has a saddle point at the optimal solution, and hence,
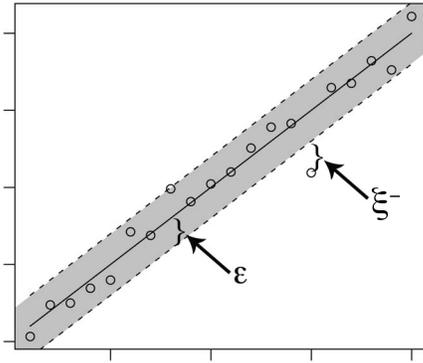
Fig. 1 Illustration of linear support vector regression fit

$$\frac{\partial L}{\partial w} = w - \sum_{t=p+1}^{t_0} (\alpha_t^+ - \alpha_t^-)\mathbf{x}_{t-p:t-1}^k = 0$$

$$\frac{\partial L}{\partial \xi_t^+} = C - \alpha_t^+ - \eta_t^+ = 0 \tag{6}$$

$$\frac{\partial L}{\partial \xi_t^-} = C - \alpha_t^- - \eta_t^- = 0$$

Substituting these saddle point constraints into L yields the following dual optimization problem:

$$\text{maximize} \begin{cases} -\frac{1}{2}\sum_{t,t'=p+1}^{t_0} (\alpha_t^+ - \alpha_t^-)(\alpha_{t'}^+ - \alpha_{t'}^-)\langle \mathbf{x}_{t-p:t-1}^k, \mathbf{x}_{t'-p:t'-1}^k \rangle \\ \\ -\varepsilon \sum_{t=p+1}^{t_0}(\alpha_t^+ + \alpha_t^-) + \sum_{t=p+1}^{t_0} x_t^k(\alpha_t^+ - \alpha_t^-) \end{cases}$$

$$\tag{7}$$

$$\text{subject to}\quad \begin{cases} x_t^k - \langle w, \mathbf{x}_{t-p:t-1}^k\rangle \le \varepsilon + \xi_t^+ \\ \langle w, \mathbf{x}_{t-p:t-1}^k\rangle - x_t^k \le \varepsilon + \xi_t^- \end{cases}$$

Notice that by the saddle point constraint $(w = \Sigma_{t=p+1}^{t_0}(\alpha_t^+ - \alpha_t^-)\mathbf{x}_{t-p:t-1}^k)$ we may write $\mathbf{f}$ as

$$\mathbf{f}(\mathbf{x}_{t-p:t-1}^k) = \sum_{t'=p+1}^{t_0} (\alpha_t^+ - \alpha_t^-)\langle \mathbf{x}_{t'-p:t'-1}^k, \mathbf{x}_{t-p:t-1}^k \rangle \tag{8}$$

In this way w may be viewed as a linear combination of the training points $\mathbf{x}_{t-p:t-1}^k$. Note also that in this formation both $\mathbf{f}$ and

the corresponding optimization can be described in terms of dot products between the data. In this way, we can transform the data using the function $\Phi: R^p \to F$ and compute the dot products in the transformed space. Such mappings allow us to extend beyond the linear framework presented above. Specifically, the mapping allows us to fit linear functions in F, which, when converted back to R^p, are nonlinear. A simplified example of this process is illustrated for a mapping $\Phi: R^2 \to R^3$, namely $\Phi(x,y) = (x^2, x\sqrt{y}, y)$, in Fig. 2. Here the data are generated using the relationship $y = x^2$.

To make use of this transformed space, we replace the dot product term with

$$\langle \Phi(\mathbf{x}_{t'-p:t'-1}^k), \Phi(\mathbf{x}_{t-p:t-1}^k)\rangle \tag{9}$$

If F is of high dimension, then the above dot product will be extremely expensive to compute. In some cases, however, there is a corresponding *kernel* that is simple to compute. For example, the kernel $k(\mathbf{x},\mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})^d$ corresponds to a map Φ into the space spanned by all products of exactly d dimensions in R^p. When $d,p=2$, for instance, we have

$$(\mathbf{x}\cdot\mathbf{y})^d = \left(\begin{pmatrix}x_1\\x_2\end{pmatrix}\cdot\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)^2 = \left(\begin{pmatrix}x_1^2\\\sqrt{2}x_1x_2\\x_2^2\end{pmatrix}\cdot\begin{pmatrix}y_1^2\\\sqrt{2}y_1y_2\\y_2^2\end{pmatrix}\right)$$

$$= (\Phi(\mathbf{x}), \Phi(\mathbf{y})) \tag{10}$$

defining $\Phi(\mathbf{x}) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$. More generally, it has been shown that every kernel that gives a positive matrix $(k(\mathbf{x},\mathbf{y}))_{ij}$ has a corresponding map $\Phi(\mathbf{x})$ [16]. One such family of kernels we focus on is radial basis function (RBF) kernels, which have the following form:

$$k(\mathbf{x},\mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2/(2\sigma^2)) \tag{11}$$

where σ^2 is the kernel variance. This parameter controls fit, with large values leading to smoother functions and small values leading to better fit. In practice, moderate values are preferred as a trade-off between model fit and prediction performance.

Whereas a traditional AR(p) model employs a linear model that is a function of the previous p time points, the SVM model looks at the previous p time points compared with all groups of p successive data points from the training sample. Specifically, the model has the form

$$\mathbf{f}(\mathbf{x}_{t-p:t-1}^k) = \sum_{j=p+1}^{t_0} \beta_j k(\mathbf{x}_{j-p:j-1}^k, \mathbf{x}_{t-p:t-1}^k) \tag{12}$$

Typically only a small fraction of the coefficients β_j are nonzero. The corresponding samples $\mathbf{x}_{j-p:j-1}^k$ are called support vectors of
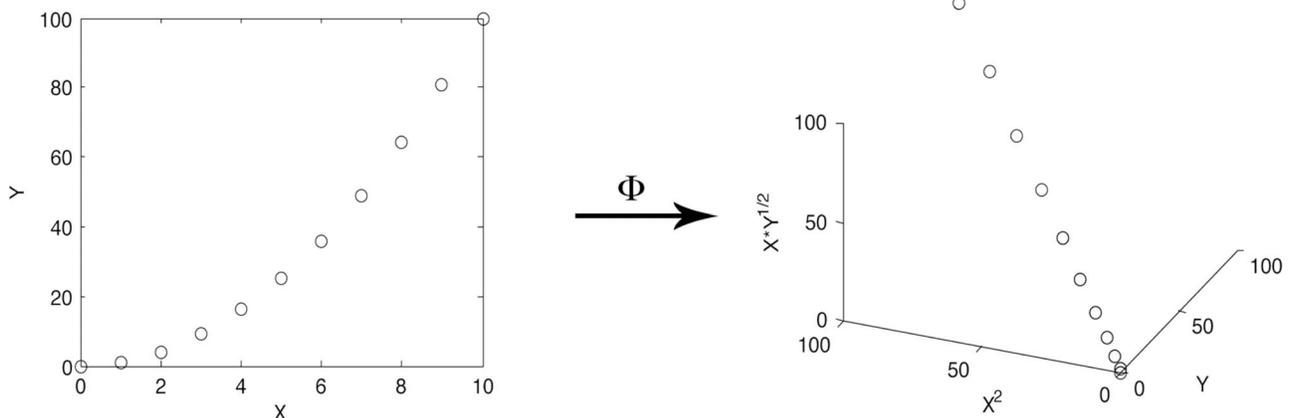

Fig. 2 Illustration of mapping to an alternate space to induce linearity
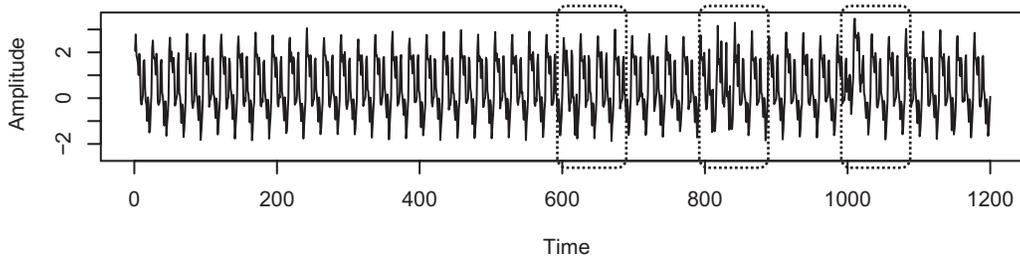
Fig. 3   Raw simulated data with highlighted artificial damage

the regression function because only these select samples are used in the formulation of the model. Once we have trained our model above, we use it to predict each future observation. We then take the residuals and use them as an indicator of structural change. For our purposes we employ a control chart to monitor if the system generating the data has changed. In this discussion the control chart is created by constructing 99% control lines that correspond to 99% confidence intervals for the residuals of the model fit to the undamaged data assuming the residuals are normally distributed. This normality assumption is further discussed in the experimental results below. These control lines are then extended through the remaining (potentially damaged) data, and damage is indicated when a statistically significant number of residuals, in this case more than 1%, lie outside these lines. Note that damage can also be indicated when the residuals no longer have a random distribution even though they may not lie outside the control lines.

RBF neural networks, which have the same form as Eq. (12), have previously been used to perform SHM [19]. However, fitting these networks requires much more user input such as selecting which β_j are nonzero as well as selecting the corresponding training points. In addition, the fitting of the neural network model is a rather complicated nonlinear optimization process relative to the simple quadratic optimization used in the support vector framework. Although the SVM models are more easily developed, Schölkopf et al. [20] demonstrated that SVMs still more accurately predict the data than the RBF neural networks despite their simplicity.

Example: Simulated damage. We now compare the performance of the SVM-based damage-detection method to a traditional AR model with coefficients estimated by the Yule–Walker method [10]. The data are generated as follows for discrete time points $t = 1, \ldots, 1200$:

$$x_t^1 = \sin^3(400\pi t/1200) + \sin^2(400\pi t/1200) + \sin(200\pi t/1200)$$

$$+ \sin(100\pi t/1200) + \Psi + \varepsilon \qquad (13)$$

where ε is Gaussian random noise with mean 0 and standard

deviation 0.1, and Ψ is a "damage" term. Three different damage cases are added to this time-series at various times as defined by

$$\Psi = \begin{cases} \varepsilon_1, & t = 600, \ldots, 650 \\ \frac{1}{2}\sin(1000\pi t/1200), & t = 800, \ldots, 850 \\ \varepsilon_2, & t = 1000, \ldots, 1050 \\ 0 & \text{otherwise} \end{cases} \qquad (14)$$

where ε_1 and ε_2 are Gaussian random noises with means 0 and 1, and standard deviations 0.5 and 0.2, respectively. Through the use of Ψ, we attempt to simulate several different types of damage to compare the models' performance handling each. This raw signal is plotted in Fig. 3 where it can be seen that the changes caused by the damage are somewhat subtle.

The order p for both models was set at 5, as determined from the autocorrelation plot in Fig. 4. This plot is the measure of correlation between successive time points for a given time lag. We see from the plot that after a lag of 5, the correlation is quite small, and hence little information is gained by including a longer past history p. The autocorrelation function is a standard method for determining model order for traditional AR models, and as such should maximize this method's performance, ensuring the SVM-based model is not afforded an advantage in this comparison.

The results of applying both the SVM model and a traditional AR model to the undamaged portion of the signal between time points 400 and 600 are shown in Fig. 5 where the signals predicted by these models are overlaid on the actual signal. A qualitative visual assessment of Fig. 5 shows that the SVM more accurately predicts this signal. A quantitative assessment is made by examining the distribution of the residual errors obtained with each model. The standard deviation of the residual errors from the SVM model is 0.26 while for the traditional AR it is 0.71, again indicating that the SVM is more accurately predicting the undamaged portion of this time-series.

In order for a model to excel at detecting damage, it must fit the undamaged data well (i.e., small and randomly distributed re-
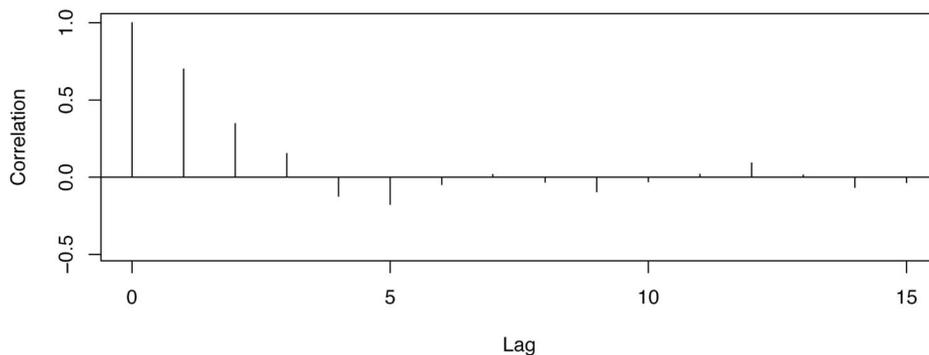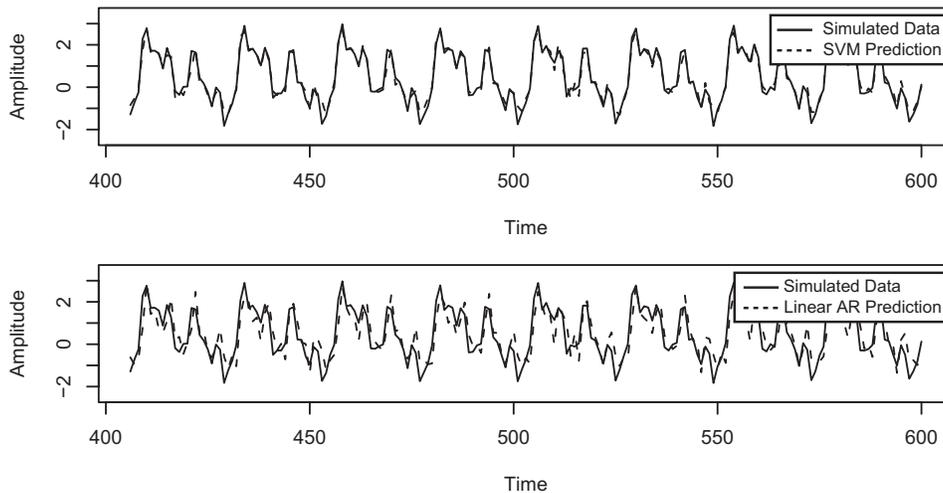

Fig. 4   Autocorrelation plot of simulated data

Fig. 5 SVM (top) and linear AR models (bottom) fit to subset of data

sidual errors) while fitting the damaged data poorly as identified by increased residual errors with possibly nonrandom distributions. In other words, the model must be sensitive to distributional changes in the data that result from damage. To quantify such changes, a control chart is developed based on the undamaged portion of the time-series to establish statistically based thresholds for the damage-detection process. As mentioned earlier, this control chart is calculated based on the fit to the undamaged data, specifically 99% confidence lines are drawn based on the undamaged residual error data and carried forward for comparison on the potentially damaged data. It is in this part of the process that the SVM's ability to more accurately represent the data enhances the damage-detection process. The 99% confidence lines for the SVM are much closer to the mean value of the residual errors and, hence, will more readily identify small perturbations to the underlying system that produce changes in the residual error distribution. In addition, the traditional AR model shows a trend in the residuals, indicating lack of model fit, even in the undamaged case. We see that during the times of damage the residuals for the SVM-based model exceed the control limits more than occurs with the residuals from the traditional AR model. In fact, the latter method would likely miss the damage between time points 1000 and 1050, where only one point exceeds the threshold versus over 10 for the SVM-based model. This result can be seen in Fig. 6.

Each method will perform differently for different types of damage; therefore, it is of interest to determine when each method will be successful in indicating damage. Because the traditional AR model fits a single model to the entire data, model fit will be very poor if the data is nonstationary (for instance, if the excitation is in the form of repeated impacts). Additionally, because the traditional AR model as presented above does not contain a moving average term, it will continue to fit when damage is in the form of a shift up or down in the raw time-series such as that which may result in strain reading when yielding occurs (demonstrated by the third damage scenario above). Conversely, the SVM-based method works by comparing each length of p data to all corresponding sets in the training set. Thus, if a similar sequence exists in the training set, we can expect the fit to be quite good. We see two scenarios in which the SVM-based method will perform poorly. First, if there is some initial damage in the "undamaged" scenario, and similar damage occurs in the testing set, the SVM model will likely fit this portion quite well. Second, if damage manifests itself in such a way that the time-series data are extremely similar to the undamaged time-series, the SVM methodology will be unable to detect it. However, we should emphasize that other methods, including the AR model, will suffer in such scenarios as well. As an attempted solution when the sensi-
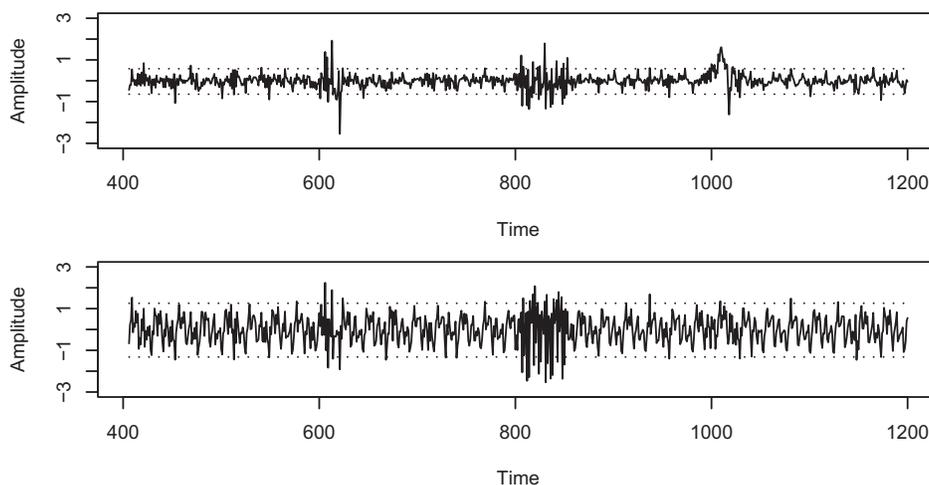
Fig. 6 Residuals from SVM (top) and linear AR models (bottom) applied to simulated data. The 99% control lines based on the residuals from the undamaged portion of the signal are shown with the horizontal lines.
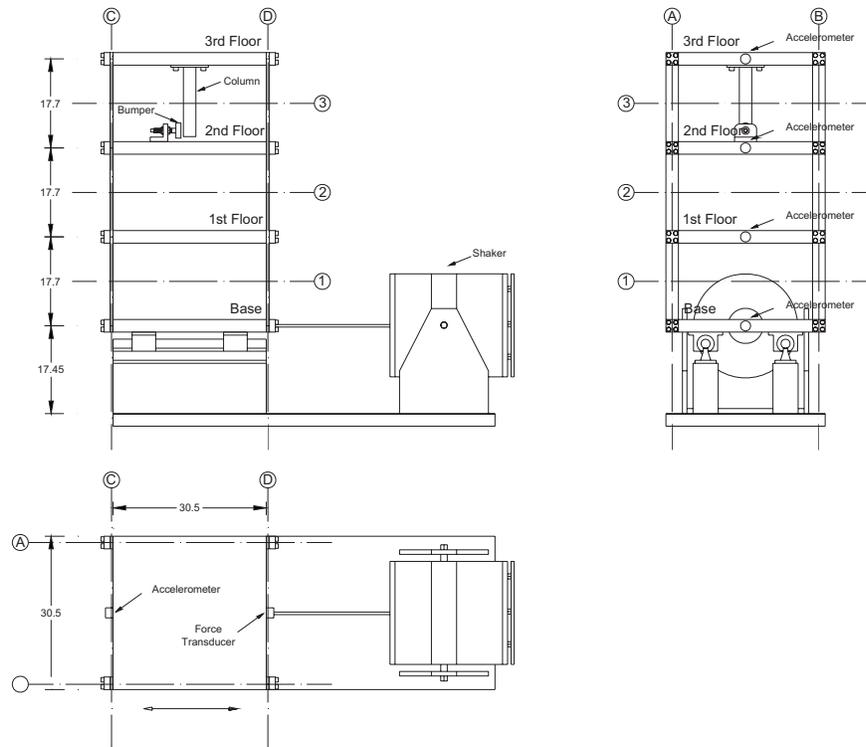
Fig. 7   Diagram of the experimental structure

tivity of the method to a given type of damage is unknown and simulation tests are impossible, various damage-detection methods could potentially be combined to boost the detection accuracy.

3   Joint Online SHM

In the undamaged state, a Gaussian distribution can often approximate the residuals from fitted models or control charts can be developed to invoke the central limit theorem and force some function of the residual errors to have a Gaussian distribution such as with an x-bar control chart [4]. If we have K sensors, each of whose residuals are Gaussian distributed, we would like a way of combining these residuals to come up with a damage-detection method that examines all K sensors. Noticing that the sum of K squared standard Gaussian random variables is distributed as a chi-squared random variable with K degrees of freedom, we square the residuals from each sensor (after they are normalized to have mean 0 and variance 1 based on the undamaged data) and add them together to create a new combined residual. These new combined residuals follow a chi-squared distribution, and hence we can make probabilistic statements about the residuals being typical or not (indicative of damage). Specifically, consider the combined residuals at some time point t.

$$\sum_{k=1}^{K} (r_t^k)^2 \tag{15}$$

where r_t^k is the normalized residual at time t for sensor k. Assuming the original residuals are Gaussian distributed, this random variable will have a chi-squared distribution with K degrees of freedom. Note that even when the original residuals are not approximately Gaussian, we may still employ a control chart on the combined residuals to give probabilistic statements regarding damage. For instance, when the residual errors from the fitted model have thicker tails than Gaussian, control charts must be employed to make probabilistic statements of the combined re-

sidual. However, as we will see in the following example, the residual errors are often very close to Gaussian.

In addition to these combined residuals allowing us to make statements regarding damage from multiple sensor output, they also provide us with a mechanism for determining which sensors are most influenced by the damage. This latter property is of particular importance for damage location. If this combined residual is large, and hence we determine that there is damage, we can look at the values $(r_t^k)^2$ for each sensor and from their magnitudes determine which sensors contributed the most to this large combined residual. If we detect damage over a range of values, we may average $(r_t^k)^2$ over this range for each sensor to determine how much each sensor is contributing to the anomalous reading.

Example: Experimental data. We look at joint online SHM using SVMs on experimental data from a structure designed to produce nonlinear response when it is "damaged." The structure is a three-story building (Fig. 7) consisting of aluminum columns and plates with bolted joints and a rigid base that is constrained to slide horizontally on two rails when excited by an electrodynamic shaker. Each floor is a $30.5 \times 30.5 \times 2.5$ cm^3 plate and is separated from adjacent floors by four $17.7 \times 2.5 \times 0.6$ cm^3 columns. To induce nonlinear behavior, a $15.0 \times 2.5 \times 2.5$ cm^3 column is suspended from the top floor and a bumper is placed on the second floor. The contact of this suspended column with the bumper results in nonlinear effects. The actual physical mechanism used to introduce the nonlinearity into the structure does not simulate a specific damage scenario. However, it does introduce response characteristics that would be observed in a structure where the damage results in an alternating stiffness state such as a crack opening and closing or the rattling of a loose connection. It is noted that this system is not scaled from a "real-world" prototype based on a rigorous similitude analysis.

The initial gap between the suspended column and the bumper is adjusted to simulate different levels of damage. In our test data we employ the case where the column is set 0.05 mm away from
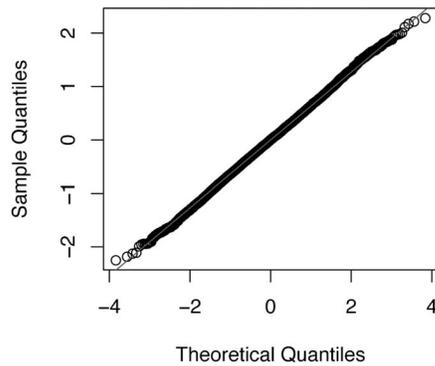
Fig. 8 Q-Q plot of residuals from SVM model

the bumper. The undamaged data is obtained when the bumper and suspended column do not contact each other. The structure is subjected to a random base excitation from the shaker in both its undamaged and damaged conditions. Accelerometers mounted on each floor record the response of the structure to these base excitations. A more detailed description of the test structure and the data obtained is available online.[2]

We first concatenate the undamaged data with the damaged data to demonstrate that the proposed methodology adequately detects the damage. The SVM time-series models are developed for each of the accelerometer measurements from the undamaged data as follows.

1. Select the number of time lags that will be used in the time-series models. In this case, eight time lags were used based on the AIC. Note that the number of time lags is analogous to the order of an AR model.
2. Select the parameters of the SVM model, including the kernel type and corresponding parameters as well as C and ε, which control model fit as described earlier. In our case, we used a Gaussian kernel with variance 1 and set $C=1$ and $\varepsilon=0.1$. We have found the methodology to be robust to choices of variance ranging over an order of magnitude. In addition, C could be increased to force fitting of extreme values, and ε could be lowered to enforce a closer fit to the training data.
3. Pass the data (arranged as dependent variable and previous p points as independent variables) to the optimization described by Eq. (7). In this case, we use the first 6000 undamaged points as training data. This step is handled by the wide variety of support vector machine software available covering multiple computing environments including MATLAB and R. In particular, we employ the LIBSVM library with accompanying MATLAB interface [21].
4. Once the SVM model is trained (i.e., the β_j in Eq. (12) are selected) in step 3, make predictions based on the new test data from the structure in its undamaged or damaged condition. Next, calculate the residual between the measured data and the output of the time-series prediction.
5. Square and add the residuals from each sensor as described by Eq. (15). Build a control chart for these combined residuals to detect damage (perhaps in conjunction with statistical tests such as a sliding window approach).

Note that steps 1–4 of this process are applied to each time-series recorded by the four accelerometers shown in Fig. 7.

First we will revisit the normality assumption that was made in constructing the control chart. Figure 8 shows the resulting Q-Q plot for the residuals from the SVM model fit to sensor four data

obtained with the structure in its undamaged state. The Q-Q plot compares the sample quantiles of the residuals to theoretical quantiles of a Gaussian distribution. We see in this figure that the sample quantiles fall very close to the theoretical line, and hence our residuals are approximately Gaussian.

Figure 9 shows the residual errors from the SVM fit to each of the accelerometer readings, respectively, and the corresponding 99% control limits that are based on the first 6000 points from the undamaged portion of each signal. There are 8192 undamaged points and 8192 damaged ones. Thus when we concatenate the data, the damage occurs at time point 8193 of 16,384. Figure 10 shows the density of the normalized residual errors from all the sensors that have been combined according to Eq. (15). We see that the distribution is very nearly chi-squared. In situations where the original residuals are not normal, this result will not be true, and hence probabilistic statements regarding the presence of damage must be made based on control charts.

Figure 11 shows the combined residuals as a function of time. The points in the upper portion of the plot show damage indication using the sliding window approach of Ma and Perkins [8], as described in the introduction and based on the 99% control lines. Specifically we use a window size of 6 which, when combined with the 99% control limit (dashed line in Fig. 11), detects damage whenever one or more of the six points in the window exceeds the control line (equivalent to binomial probability of 0.05). We see from Fig. 9 that sensors 3 and 4 are most influenced by damage. This result is expected as the bumper is mounted between these two sensors. In fact, if we look at the average values of $(r_t^k)^2$ (which are the individual squared residuals for sensor k) over the damaged section for each sensor, we see that the first two sensors have values 0.96 and 1.24, whereas the second two sensors have values 59.80 and 38.2, respectively. Thus from this numerical rating we can see that sensors 3 and 4 are most influenced by the damage, which agrees with the result shown in Fig. 9.

From this analysis it is evident that we can use the combined residuals to establish the presence of damage in a statistically rigorous manner and then examine the individual sensor residuals in an effort to locate the sensors most influenced by the damage. This latter information can be used to help locate the damage assuming that the damage is confined to a discrete location such as the formation of a crack in a welded connection. Further investigation is needed to assess how this procedure could be used to locate damage for more distributed damage such as that associated with corrosion.

4 Conclusion

Although the application of statistical techniques to structural health monitoring has been investigated in the past, these techniques have predominantly been limited to identifying damage-sensitive features derived from linear models fit to the output from individual sensors. As such, they are typically limited to identifying only that damage has occurred. In general, these methods are not able to identify which sensors are associated with the damage in an effort to locate the damage within the resolution of the sensor array.

To improve on this approach to damage detection, we have applied support vector machines to model sensor output time histories and have shown that such nonlinear regression models more accurately predict the time-series when compared with linear autoregressive models. Here the metric for this comparison is the residual errors between the measured response data and predictions of the time-series model.

The support vector machine autoregressive method is superior to traditional linear AR in both its ability to handle nonlinear dynamics as well as the structure of the model. Specifically, the support vector approach compares each new testing point to the entire training set whereas the traditional AR model finds a simple linear relationship to best describe the entire training set, which is then used on the testing data. For example, when dealing with
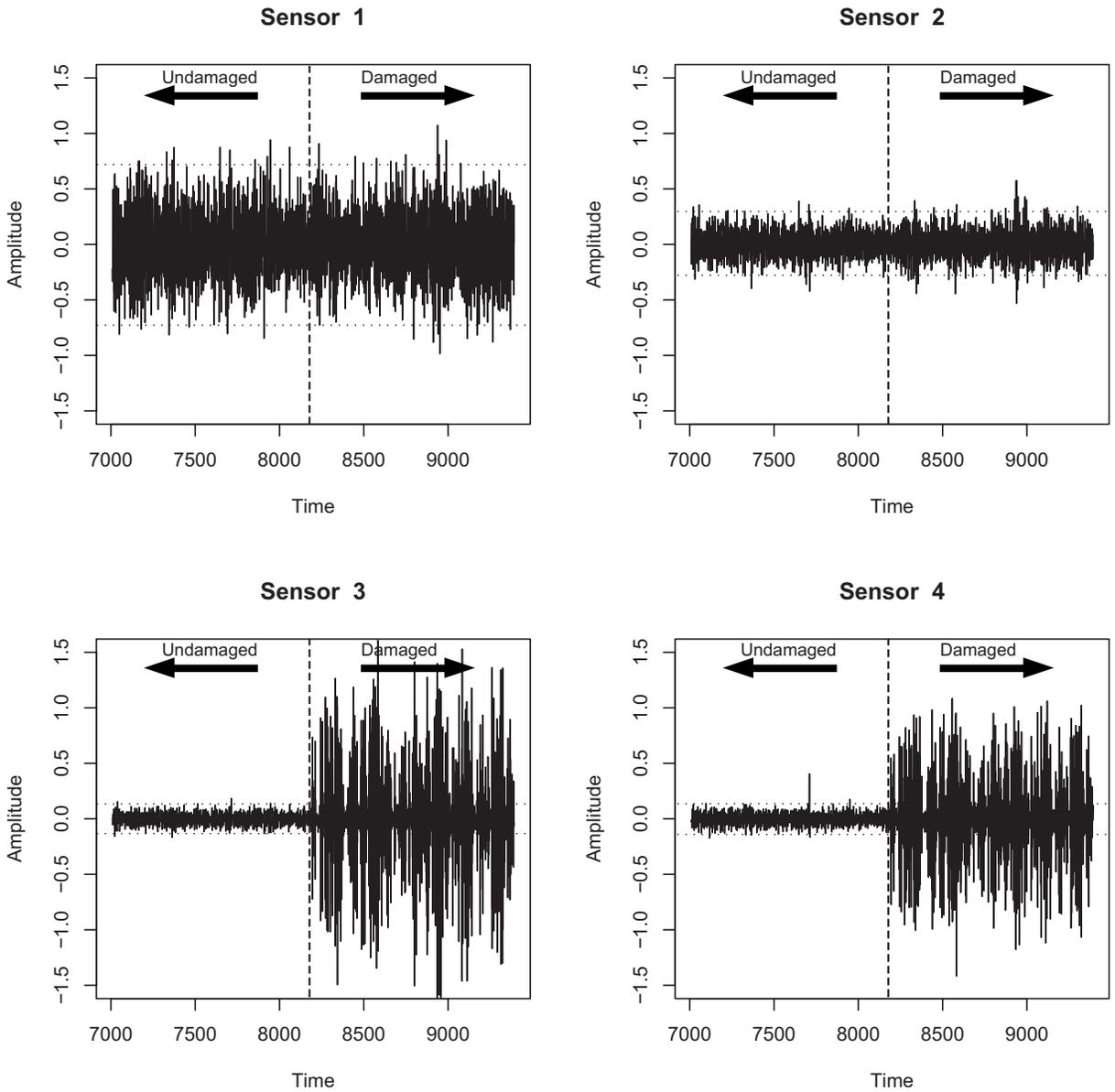
Fig. 9 Residuals from four sensors for $t=7000,\ldots,9384$. The horizontal lines are the 99% control lines.

transient impact data, the AR model will fail in trying to fit the entire time domain with a simple linear model. Whereas in the past RBF neural networks have been used to tackle this problem, these networks require significant user input and complex methods for fitting the model to the training data, and hence the simple support vector framework is preferred.

Furthermore, we have also shown how the residuals from the SVM prediction of each sensor time history may be combined in a statistically rigorous manner to provide probabilistic statements regarding the presence of damage as assessed from the amalgamation of all available sensors. In addition, this methodology allows us to pinpoint the sensors that are contributing most to the anomalous readings and therefore locate the damage within the sensor network's spatial resolution. The process was demonstrated on a test structure where damage was simulated by introducing an impact type of nonlinearity between the measured degrees of freedom. The authors acknowledge that the approach has only been demonstrated on a structure that was tested in a well-controlled
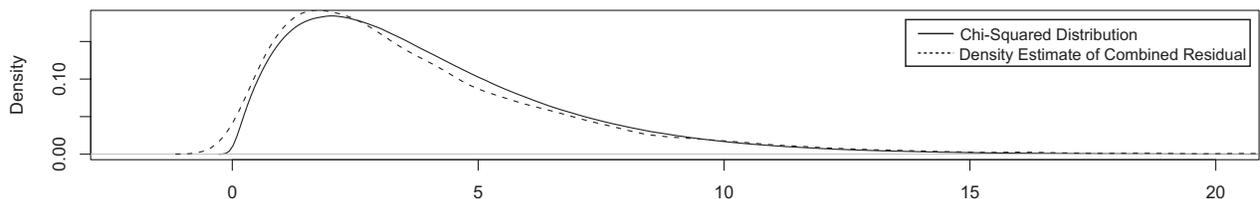
Fig. 10 Density estimate of combined residual (black) versus chi-squared distribution (dashed)
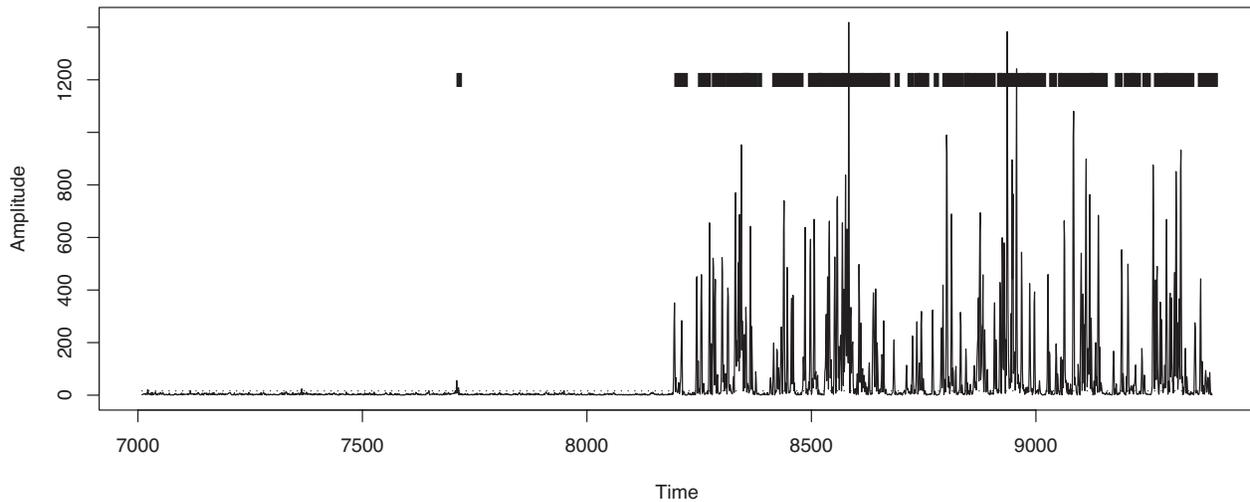
Fig. 11 Combined residuals from all four sensors. The 99% control line is shown as the dashed horizontal line. Sliding window damage indicators are indicated by the boxes across the top of the plot.

laboratory setting. This approach will have to be extended to structures subjected to real-world operational and environmental variabilities before it can be used in practice. However, the approach has the ability to adapt to such changes through the analysis of appropriate training data that span these conditions. Therefore, follow-on studies will focus on applying this approach to systems with operational and environmental variabilities as well as systems that exhibit nonlinear response in their undamaged state.

Acknowledgment

References

[1] Doebling, S., Farrar, C., Prime, M., and Shevitz, D., 1998, "A Review of Damage Identification Methods That Examine Changes in Dynamic Properties," Shock Vib. Dig., **30**, pp. 91–105.

[2] Sohn, H., Farrar, C. R., Hemez, F. M., Shunk, D. S., Stinemates, D. W., Nadler, B. R., and Czarnecki, J. J., 2004, "A Review of Structural Health Monitoring Literature From 1996–2001," Los Alamos National Laboratory, Report No. LA-13976-MS.

[3] Farrar, C. R., and Worden, K., 2007, "An Introduction to Structural Health Monitoring," Philos. Trans. R. Soc. London, Ser. A, **365**, pp. 303–315.

[4] Fugate, M., Sohn, H., and Farrar, C. R., 2001, "Vibration-Based Damage Detection Using Statistical Process Control," Mech. Syst. Signal Process., **15**, pp. 707–721.

[5] Sohn, H., Czarnecki, J., and Farrar, C. R., 2000, "Structural Health Monitoring Using Statistical Process Control," J. Struct. Eng., **126**, pp. 1356–1363.

[6] Allen, D., Sohn, H., Worden, K., and Farrar, C., 2002, "Utilizing the Sequential Probability Ratio Test for Building Joint Monitoring," Proc. SPIE, **4704**, pp. 1–11.

[7] Clark, G., 2008, "Cable Damage Detection Using Time Domain Reflectometry and Model-Based Algorithms," Lawrence Livermore National Laboratory, Document No. LLNL-CONF-402567.

[8] Ma, J., and Perkins, S., 2003, "Online Novelty Detection on Temporal Sequences," *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, DC, pp. 613–618.

[9] Herzog, J., Hanlin, J., Wegerich, S., and Wilks, A., 2005, "High Performance Condition Monitoring of Aircraft Engines," *Proceedings of GT 2005 ASME Turbo Expo*, Reno, NV, Paper No. GT2005-68485.

[10] Brockwell, P., and Davis, R., 1991, *Time Series: Theory and Methods*, Springer, New York.

[11] Worden, K., and Manson, G., 2007, "The Application of Machine Learning to Structural Health Monitoring," Philos. Trans. R. Soc. London, Ser. A, **365**, pp. 515–537.

[12] Shimada, M., Mita, A., and Feng, M. Q., 2006, "Damage Detection of Structures Using Support Vector Machines Under Various Boundary Conditions," Proc. SPIE, **6174**, pp. 61742K.

[13] Bulut, A., Singh, A. K., Shin, P., Fountain, T., Jasso, H., Yan, L., and Elgamal, A., 2005, "Real-Time Nondestructive Structural Health Monitoring Using Support Vector Machines and Wavelets," Proc. SPIE, **5770**, pp. 180–189.

[14] Worden, K., and Lane, A. J., 2001, "Damage Identification Using Support Vector Machines," Smart Mater. Struct., **10**, pp. 540–547.

[15] Chattopadhyay, A., Das, S., and Coelho, C. K., 2007, "Damage Diagnosis Using a Kernel-Based Method," Insight-Non-Destructive Testing and Condition Monitoring, **49**, pp. 451–458.

[16] Smola, A. J., and Schölkopf, B., 2004, "A Tutorial on Support Vector Regression," Stat. Comput., **14**, pp. 199–222.

[17] Copas, J. B., 1997, "Using Regression Models for Prediction: Shrinkage and Regression to the Mean," Stat. Methods Med. Res., **6**, pp. 167–183.

[18] Fu, W. J., 1998, "Penalized Regressions: The Bridge Versus the Lasso," J. Comput. Graph. Stat., **7**, pp. 397–416.

[19] Rytter, A., and Kirkegaard, P., 1997, "Vibration Based Inspection Using Neural Networks," Structural Damage Assessment Using Advanced Signal Processing Procedures, *Proceedings of DAMAS '97*, University of Sheffield, UK, pp. 97–108.

[20] Scholkopf, B., Sung, K. K., Burges, C. J. C., Girosi, F., Niyogi, P., Poggio, T., and Vapnik, V., 1997, "Comparing Support Vector Machines With Gaussian Kernels to Radial Basis Function Classifiers," IEEE Trans. Signal Process., **45**, pp. 2758–2765.

[21] Chang, C.-J., and Lin, C.-J., 2001, LIBSVM: A Library for Support Vector Machines, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.