

Robust structural health monitoring under environmental and operational uncertainty with switching state-space autoregressive models

Structural Health Monitoring

1–19

© The Author(s) 2018

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1475921718757721

journals.sagepub.com/home/shm



Anthony Liu¹, Lazhi Wang¹, Luke Bornn¹ and Charles Farrar²

Abstract

Existing methods for structural health monitoring are limited due to their sensitivity to changes in environmental and operational conditions, which can obscure the indications of damage by introducing nonlinearities and other types of noise into the structural response. In this article, we introduce a novel approach using state-space probability models to infer the conditions underlying each time step, allowing the definition of a damage metric robust to environmental and operational variation. We define algorithms for training and prediction, describe how the algorithm can be applied in both the presence and absence of measurements for external conditions, and demonstrate the method's performance on data acquired from a laboratory structure that simulates the effects of damage and environmental and operational variation on bridges.

Keywords

Structural health monitoring, vector autoregressive models, switching state-space models, environmental and operational variability, EOV uncertainty

Introduction

A central challenge in structural health monitoring (SHM) is distinguishing the effects of damage on structural behavior from the effects of environmental and operational variation (EOV). Real-world structures are exposed to constantly changing conditions, and therefore, methods robust to the effects of EOV must be established in order for SHM to be practical. This article introduces a state-space approach novel to SHM that uses knowledge of the structure's behavior under different conditions to infer the conditions affecting the structure and apply a model that is appropriate given those conditions, thereby allowing the effects of damage to be isolated from the effects of EOV.

SHM can be framed as a novelty detection problem in which a model for healthy structural behavior is established and new observations are classified as *healthy* or *damaged* depending on whether or not they continue to follow that model.¹ An important component of novelty detection is therefore the selection of features that allow for discrimination between healthy and damaged structures. It is common in SHM to construct these features from a structure's dynamic properties

(e.g. modal properties), and this is known as the vibration-based approach to SHM.^{2–4} Autoregressive (AR) time-series models provide a robust way to capture the dynamic properties of structures, and their residuals can be incorporated into a control chart framework as features for damage detection.^{5,6} The AR model is useful not only in the univariate case but can also be applied to damage detection in multivariate time series, such as those generated by sensor networks.⁷

The dynamic properties of structures are known to be sensitive to variations in environmental and operational conditions.^{8–10} The task of engineering damage-sensitive features robust to EOV is known as *data normalization*. One approach to data normalization is to preprocess the data and identify combinations of

¹Department of Statistics, Harvard University, Cambridge, MA, USA

²Engineering Institute, Los Alamos National Laboratory, Los Alamos, NM, USA

Corresponding author:

Anthony Liu, Department of Statistics, Harvard University, Science Center 400 Suite, One Oxford Street, Cambridge, MA 02138, USA.
Email: aliu14@post.harvard.edu

features invariant under EOV. Figueiredo et al.¹¹ and Cross et al.¹² demonstrate how methods such as principal components analysis (PCA), nonlinear PCA (auto-encoder neural networks), and cointegration can be applied to exactly this purpose. An alternative to learning commonalities across different conditions is to explicitly model the individual behaviors of the structure under different EOV. Farrar et al.¹³ describe an approach in which a “reference table” of models is constructed to enumerate normal behavior under a range of environmental and operational conditions; the model residuals of the reference model that most closely resemble the new observations are then used to evaluate the health of the structure. When measurements of the EOV are available, directly modeling the structure’s dynamic properties as a function of the conditions can be effective, as demonstrated by Worden et al.,¹⁴ who use a decision tree regression model to explicitly model how the dynamic properties of a bridge change with the external temperature.

This article introduces a state-space model called the switching vector autoregressive (SVAR) model as a novel approach that extends this latter class of data normalization methods. The SVAR model divides the full range of EOV into distinct regimes of behavior, learning from each discrete state an AR model that describes that state’s expected behavior. This allows the model to effectively normalize out the effects of EOV and focus solely on damage, similar to the methods of Farrar et al.¹³ and Worden et al.¹⁴ However, in contrast to these other methods, the SVAR model also provides a probabilistic framework that allows us to infer the state of the underlying system, making it practical for situations where measurements of the environmental and operational variables affecting the structure are unavailable. In addition to outlining basic inference methods, we also detail inference methods for when measurements of state variables are partially available, as well as simplifications to inference that can be applied when the state space is large. Our methodology is evaluated upon experimental data collected across multiple observation windows with distinct environmental and operational conditions. We also include a simulation study to demonstrate that although the SVAR model assumes discrete states, its probabilistic framework allows it to effectively model observations even when EOV are changing continuously, due to the SVAR model’s ability to combine the behaviors learned from each constituent model when making new estimates, weighting the influence of each state by its likelihood. A similar state-space approach to SHM can be found in Avendaño-Valencia and Fassois,¹⁵ though without the transition dynamics modeled by the SVAR approach.

We first describe the general AR approach to SHM before introducing the SVAR model, detailing the algorithms it requires for inference, prediction, and damage detection. We then present a simulated example to illustrate how the method is applied before using data from a laboratory test structure with simulated EOV to demonstrate the advantages it offers over non-switching AR models, the nonlinear autoregressive support vector machine (AR-SVM) model, and minor PCA.

Background

The approach developed in this article has its foundations in the vibration-based SHM approach of Fugate et al.,⁶ who demonstrated how statistical process control could be applied to SHM by constructing control charts from the residuals of AR models. In the following section, we describe this methodology in full, describing inference, prediction, and damage-detection methods for the univariate AR model and the multivariate vector autoregressive (VAR) model in the context of SHM.

AR models

Let $\{x_1, x_2, \dots, x_T\}$ be a univariate zero-mean time series, representing the output of a single sensor. An AR model of order p is a linear model

$$x_t = \sum_{i=1}^p a_i x_{t-i} + \varepsilon_t$$

for $t \geq p+1$, where the errors $\{\varepsilon_{p+1}, \dots, \varepsilon_T\}$ are i.i.d. random variables with a $\mathcal{N}(0, \sigma^2)$ distribution, and a_1, \dots, a_p are the *AR coefficients*. These parameters can be estimated using either maximum likelihood or the Yule–Walker method.¹⁶ Note that AR models assume white noise, i.e., uncorrelated errors. In the presence of colored noise, it may make sense to similarly model the AR structure of the noise terms, in which case the class of autoregressive-moving-average (ARMA) models may be more applicable.¹⁶

Determining model order. It is important to carefully select the order of AR models so as to prevent overfitting. Several of the most common metrics for AR model selection are the partial autocorrelation function (PACF), model likelihood, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC).¹⁶ The PACF looks at the autocorrelation of the time series at different lags, and tries to select the simplest model that includes all significantly correlated lags. The model likelihood examines the closest fit achievable at each order, and tries to identify the point

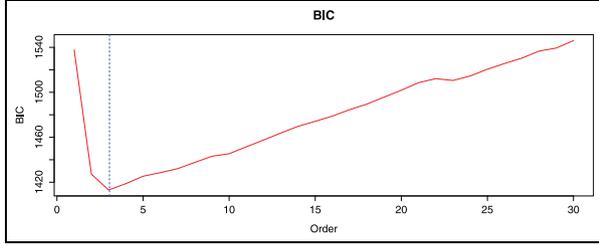


Figure 1. Plot of the BIC as a function of model order for data generated from an AR model ($p = 3$) fit with maximum likelihood: $x_t = 0.4x_{t-1} + 0.3x_{t-2} + 0.15x_{t-3} + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, 1)$. The dashed vertical line indicates where BIC reaches its minimum value, which is at $p = 3$.

after which additional parameters have only a marginal effect on the fit, usually identified as a plateau in the model likelihood as a function of order. The AIC and BIC are penalized versions of the model likelihood that approximate the predictive error of the model; the optimal order under these metrics is therefore the model that minimizes the AIC or BIC.^{17,18}

While these metrics do not always agree with one another on the appropriate model order, they reach similar enough conclusions that the choice of metric does not significantly affect the performance of AR damage-detection methods.¹⁹ In this article, we use the BIC method to select model order because it is less subjective than the PACF or model likelihood, and tends toward lower order models than AIC. The BIC is defined as

$$\text{BIC}(p) = -2 \ln L(p) + k_p \ln N$$

where $L(p)$ is the model likelihood of an AR model of order p , k_p is the number of parameters of a model of order p , and N is the number of data points observed.¹⁸ A plot demonstrating how BIC changes as a function of model order is presented in Figure 1.

Damage detection. We detect damage by constructing control charts from the residuals of an AR model fit to the sensor output.⁶ Under the control chart framework, damage detection is equivalent to identifying observations that are outliers to the process; these outliers are defined as points exceeding thresholds set to quantiles of the feature's distribution, called *control lines*.²⁰ A typical control chart, known as a 99% control chart, has control lines set at the 0.5% and 99.5% quantiles of the residual's distribution; these thresholds can be adjusted for a specific SHM problem based on the relative tolerance for false-positive versus false-negative indications of damage.

Vector AR models

Let $\{\mathbf{x}_1, \dots, \mathbf{x}_T\}$ be a zero-mean multivariate time series representing the output of a sensor network, where each \mathbf{x}_t is a $m \times 1$ vector of observations (x_{t1}, \dots, x_{tm}) , with x_{ti} representing the observation from the i th sensor at time t , and with T representing the total number of time steps and m the total number of sensors. A VAR model of order p is then defined as

$$\mathbf{x}_t = \sum_{i=1}^p A_i \mathbf{x}_{t-i} + \varepsilon_t$$

where the errors $\{\varepsilon_{p+1}, \dots, \varepsilon_T\}$, $\varepsilon_t \in \mathbb{R}$, are i.i.d. random variables with a $\mathcal{MVN}(\mathbf{0}, Q)$ distribution, $\mathbf{0}$ is a zero-vector of dimension $m \times 1$, and Q is a $m \times m$ covariance matrix. The AR coefficients A_1, \dots, A_p are $m \times m$ matrices, where $A_i[j, k]$ describes the dependence of \mathbf{x}_{ij} on $\mathbf{x}_{(t-i)k}$; note that if A_1, \dots, A_p and Q are all diagonal matrices, then the m sensors are independent from one another. The methods used for inference and model selection for VAR models are the same as those used for AR models. Examples of VAR models in the SHM literature include Bodeux and Golinval,²¹ De Stefano et al.,²² Bornn et al.,²³ and Dzunic et al.²⁴

VAR models are useful not only for damage detection but also damage localization, as the AR matrices encode the dependence relationships between each sensor.^{23,24}

Damage detection. As with univariate AR models, the residuals of VAR models can be used to construct control charts. However, since the residuals are now multivariate, they are first transformed to a univariate statistic by standardizing the variance of each dimension and computing the sum of squares

$$z_t = \varepsilon_t^T Q^{-1} \varepsilon_t$$

Since ε_t is assumed multivariate normal, z_t is distributed according to a χ_m^2 distribution.²⁵ The likelihood of each observation can also be used as a damage statistic.

Methods

The SVAR model introduced in this article is an example of a state-space model. Here, the signal evolves according to one of several possible discrete regimes of behavior between which the system alternates. Each component state is associated with a model that defines the behavior expected under that state; in the case of the SVAR model, each component state is modeled by a VAR model. This framework finds an intuitive application in SHM under EOV, as we can group similar sets

of environmental and operational conditions into discrete states, and then infer the likely value of future observations by considering the behavior of the model under each of these states. If the observed behavior of the system then still deviates from expectation, we can have the confidence that the observed novelty is due to factors other than a benign change in the environmental and operational conditions. Furthermore, while the state space of an SVAR model is finite and discrete, provided that the training states well represent the range of possible variation in behavior, a trained SVAR model can also successfully be applied to signals whose behavior is changing continuously, as demonstrated in the simulation study below.

A similar method was proposed by Farrar et al.¹³ using a “reference table” that specified different models for different sets of environmental and operational conditions. New observations would then be compared to each of the models in the reference table; the residual from the best fitting model would then be used as the feature for damage detection. More recently, Worden et al.¹⁴ developed a method using treed Gaussian processes to model directly how the dynamic properties of structures vary with the environmental and operational conditions. The general concept is similar to that of the state-space model, i.e., to normalize out the effects of EOV by modeling the way dynamic properties depend on environmental and operational conditions, though that method differs from this article in that it is continuous and requires direct measurement of external conditions. Worden et al.¹⁴ model the relationship between temperature and the fundamental frequencies of the structure by learning a decision tree that partitions the range of temperatures into regimes of similar behavior, each of which is then fit with a separate Gaussian process. This partitioning of the space is necessary as the relationship between temperature and the dynamic properties can differ greatly in different ranges, in particular above and below the freezing point.

The methodology described below builds on these methods by providing a probabilistic framework that specifies not only how observations are generated under each state but also how the states transition between one another, leading to predictions more robust than those provided by the simple reference table approach. In addition, the probabilistic structure provides a way to infer the state underlying new observations, allowing the model to provide accurate predictions even without complete knowledge of the current environmental and operational conditions, an important capability as direct measurement of the sources of EOV is not always available.

Avendaño-Valencia and Fassois¹⁵ have also proposed a mixture model approach that shares elements of the methodology introduced in this article, in

particular by separately modeling the vibration response of the structure under different states. A key motivation of their work was also in addressing the challenges of SHM in contexts where direct measurement of EOV was not possible. Our approach to damage detection mainly differs in two areas: (1) Avendaño-Valencia and Fassois use a more flexible class of models for each component state than the VAR models used in this article (they use linear parameter varying autoregressive (LPV-AR) and functional series time-dependent autoregressive (FS-TAR) models), and (2) their model represents all healthy behavior under a single state, as opposed to dividing healthy behavior into multiple states as in the SVAR approach. Further work combining the two approaches would be interesting, especially to see whether using both more complex models and a richer state space would lead to better damage-detection capabilities than either individually.

SVAR models

Let $\{\mathbf{x}_1, \dots, \mathbf{x}_T\}$ be a zero-mean multivariate time series representing the output of a sensor network and $\{s_1, \dots, s_T\}$ be a hidden univariate discrete-valued time series representing the environmental and operational conditions, or *state*, of the structure at each time point, where $\mathbf{x}_t \in \mathbb{R}^m$ and $s_t \in \{1, \dots, K\}$. Each state $1, \dots, K$ corresponds to a regime of environmental and operational conditions, where the full range of possible EOV has been discretized into a finite number of bins. An SVAR model is defined as a process where (1) the state variables follow a Markov transition process and (2) each observation \mathbf{x}_t is generated according to the AR model corresponding to the current state s_t . In other words, an SVAR model is a stochastic process generated as

$$s_{t+1} \sim \text{Categorical}(Z[s_t, 1], \dots, Z[s_t, k])$$

$$\mathbf{x}_{t+1} | s_{t+1} \sim \mathcal{MVN} \left(\sum_{i=1}^{p^{(s_{t+1})}} A_i^{(s_{t+1})} \mathbf{x}_{t+1-i}, Q^{(s_{t+1})} \right)$$

where $Z \in \mathbb{R}^{k \times k}$ is a transition matrix with $Z[i, j] = P(s_t = j | s_{t-1} = i)$ and the matrices $A_1^{(k)}, \dots, A_{p^{(k)}}^{(k)}, Q^{(k)}$ are the parameters of the VAR model for state k . We can also write the model in the form

$$\mathbf{x}_t = \sum_{i=1}^{p^{(s_t)}} A_i^{(s_t)} \mathbf{x}_{t-i} + \varepsilon_t$$

where $\{\varepsilon_{p+1}, \dots, \varepsilon_T\}$ are i.i.d random variables with $\varepsilon_t \sim \mathcal{MVN}(\mathbf{0}, Q^{(s_t)})$. A graphical representation of an SVAR model of order 2 is presented in Figure 2. The horizontal lines linking the state variables indicate that

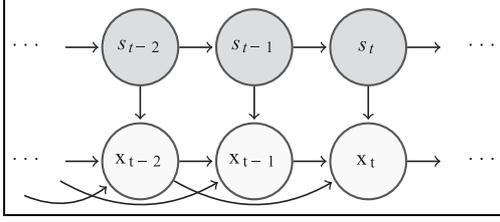


Figure 2. A graphical representation of a switching VAR model of order 2. Each observation depends on both the corresponding hidden state as well as the prior two observations; each state only depends on the previous state.

each state variable depends only on the value of the preceding state variable, representing the Markov transition process; note that if the state dependency structure is believed to be more complex, we can construct an extended space by including the previous 2, 3, ..., and so on state variables while maintaining the Markov framework. The dependencies of each x_t are more complex, as x_t depends not only on the previous two time observations but also the current state s_t . A comprehensive reference for inference and prediction with SVAR models can be found in Murphy.²⁶

A useful way to extend the SVAR model is to model the state of the system as the composite of several different *factors*; in the context of SHM, the factors would represent components of EOV such as temperature and live load (e.g. traffic loading on a bridge). More formally, in a factorized SVAR model, we describe the state as a multivariate variable $s_t = (s_{t1}, \dots, s_{tL}) \in \mathbb{R}^L$, where s_{tL} is the observation of the l th factor at time t ; the total number of states is then $K = \prod_{l=1}^L K_l$, where K_l is the number of levels of the l th factor. This provides a more intuitive representation of the state as the Cartesian product of the individual components of EOV. With the assumption that the states transition independently, the transition matrix can then be decomposed as

$$\begin{aligned} Z[i, j] &= P(s_t = \mathbf{j} | s_{t-1} = \mathbf{i}) \\ &= \prod_{l=1}^L P(s_{tl} = j_l | s_{(t-1)l} = i_l) \\ &= \prod_{l=1}^L Z_l[i_l, j_l] \end{aligned}$$

where $Z_l \in \mathbb{R}^{K_l \times K_l}$ is the transition matrix of the l th factor. This is convenient as it allows the transition process of each factor to be considered separately, and reduces the number of parameters in the model. For an example of a factorized state-space model in practice, see Quinn et al.,²⁷ who apply a factorized state-space model to the domain of patient health monitoring.

Inference. If the training data are *labeled* (i.e. there are measurements of the environmental and operational

conditions), then the SVAR model parameters can be estimated by isolating the observations corresponding to each state and using the inference methods for VAR models described earlier. The transition matrix Z can be estimated by computing the empirical transition probabilities

$$\hat{Z}[i, j] = \frac{n_{i,j}}{n_i}$$

where $n_{i,j}$ is the number of times state i transitions to state j , and n_i is the number of times state i appears; if the factorized representation is used, the factor transition matrices can be analogously estimated using the counts for each factor level.

If the training data are *unlabeled* (i.e. the environmental and operational conditions are unknown), then the model parameters can be learned with *expectation-maximization* (EM), an iterative algorithm that converges on a local maximum of the likelihood.²⁶ The first step of the EM algorithm is to randomly initialize the SVAR model parameters $A_i^{(k)}$, $Q^{(k)}$, Z to a set of values $\theta^{(0)}$. The E-step then uses the initial estimates $\theta^{(0)}$ to compute the state probabilities $\pi_{t,k} = P(s_t = k | \mathbf{x}_{1:T}, \theta^{(0)})$ using the forward-backward algorithm (see the following section). The $\pi_{t,k}$ are then fixed in the M-step so that the maximum likelihood estimates of each of the parameters can be computed

$$\begin{aligned} A_i^{(k)} &= \left(\sum_{t=2}^T \pi_{t,k} R_{t,t-1} \right) \left(\sum_{t=2}^T \pi_{t,k} R_{t-1} \right)^{-1} \\ Q^{(k)} &= \left(\frac{1}{\sum_{t=2}^T W_t^i} \right) \left(\sum_{t=2}^T \pi_{t,k} R_t \perp A_i \sum_{t=2}^T \pi_{t,k} R_{t,t-1} \right) \\ Z[i, j] &= \frac{\sum_{t=2}^T P(s_{t-1} = i, s_t = j | \mathbf{x}_{1:T})}{\sum_{t=1}^{T-1} \pi_{t,i}} \end{aligned}$$

where $R_t = \mathbf{x}_t \cdot \mathbf{x}_t$ and $R_{t,t-1} = \mathbf{x}_t \cdot \mathbf{x}_{t-1}$; if a factorized representation is used, the factor transition matrices can be analogously estimated using the computed probabilities for each factor level. Note that the number of states must be fixed before applying the EM algorithm; model selection criteria like BIC can be applied to select the number of states to use in the model.

Prediction. The expected value of \mathbf{x}_t under the SVAR model is a weighted average of the expected value of \mathbf{x}_t under each state-specific model

$$\begin{aligned} \hat{\mathbf{x}}_t &= \mathbb{E}[\mathbf{x}_t] \\ &= \sum_{k=1}^K \pi_{t,k} \mathbb{E}[\mathbf{x}_t | s_t = k] \\ &= \sum_{k=1}^K \pi_{t,k} \hat{\mathbf{x}}_t^{(k)} \\ &= \sum_{k=1}^K \pi_{t,k} \left[\sum_{i=1}^{p^{(k)}} A_i^{(k)} \mathbf{x}_{t-i} \right] \end{aligned}$$

The residuals of the SVAR model are similarly a weighted sum of the state-specific residuals

$$\begin{aligned}\hat{\varepsilon}_t &= \mathbf{x}_t - \hat{\mathbf{x}}_t \\ &= \mathbf{x}_t - \sum_{j=1}^K \pi_{t,j} \hat{\mathbf{x}}_t^{(j)} \\ &= \sum_{j=1}^K \pi_{t,j} \left[\mathbf{x}_t - \hat{\mathbf{x}}_t^{(j)} \right] \\ &= \sum_{j=1}^K \pi_{t,j} \varepsilon_t^{(j)}\end{aligned}$$

(the equivalence of the second and third steps follows because $\sum_j \pi_{t,j} = 1$ for each time t).

Often, the environmental and operational conditions underlying new observations are not known; when this is the case, the probabilities $\pi_{t,k} = P(s_t = k | \mathbf{x}_{1:T})$ can be estimated using the *forward-backward algorithm*. The forward-backward algorithm makes two passes over the data: (1) recursively computing the partial-data probabilities $P(s_t = k | \mathbf{x}_{1:t}, \theta)$, which are then used for (2) computing the full-data probabilities $\pi_{t,k} = P(s_t = k | \mathbf{x}_{1:T}, \theta)$. Specifically, we first iterate forwards in time starting at $t = 1$ and compute

$$P(s_t = j | \mathbf{x}_t, \mathbf{x}_{1:t-1}) = \frac{1}{c_t} L_t(j) \sum_{i=1}^K Z[i,j] P(s_{t-1} = i | \mathbf{x}_{1:t-1})$$

for each $t \in [2, T]$, where c_t is a normalizing constant and $L_t(j)$ is the likelihood of the observation at time t under state j ; the probabilities $P(s_1 = j)$ are initialized as $P(s_1 = j) = (1/c_1) L_1(j)$. The backward pass is then used to compute the full-data probabilities by iterating backwards from $t = T$ to $t = 1$

$$P(s_t = j | \mathbf{x}_{1:T}) = \sum_{j'=1}^K \frac{P(s_t = j | \mathbf{x}_{1:t}) Z[j,j']}{P(s_{t+1} = j' | \mathbf{x}_{1:t})} P(s_{t+1} = j' | \mathbf{x}_{1:T})$$

Damage detection. As with the VAR model, a univariate statistic for damage detection with an approximately chi-square distribution can be constructed from the model residuals. Note that the SVAR residuals are distributed

$$\varepsilon_t = \sum_{j=1}^K \pi_{t,j} \varepsilon_t^{(j)} \sim \mathcal{MVN} \left(0, \sum_{j=1}^K \pi_{t,j}^2 Q^{(j)} \right)$$

and therefore standardized residuals can be constructed as

$$\mathbf{z}_t = \varepsilon_t^T \left(\pi_{t,1}^2 Q^{(1)} + \cdots + \pi_{t,K}^2 Q^{(K)} \right)^{-1} \varepsilon_t$$

The likelihood of each prediction can also be used as a damage statistic.

Computational complexity of prediction. The computational complexity of prediction is dominated by the forward-backward algorithm, which is an $O(TK^2)$ algorithm; this implies that prediction is linear in the number of observations, but quadratic in the number of states. However, in most real-world long-term monitoring cases, the number of different environmental and operational conditions will be dwarfed by the length of the time series, so the linear factor will dominate. The linear cost could be further reduced by using sliding windows, truncating the computation from the length of the whole time series to only the size of the window.

Simulation study

The following example will illustrate how to fit an SVAR model, demonstrate its ability to infer the environmental and operational conditions of a system, and show its advantages over non-switching VAR models. In addition, the simulation study will demonstrate how the switching model can use a discrete state space to effectively model systems with continuously evolving environmental and operational conditions by interpolating between the models for known states. Consider a time series $\{\mathbf{x}_1, \dots, \mathbf{x}_{900}\}$ where $\mathbf{x}_t \in \mathbb{R}^2$ with three different states

$$\begin{aligned}\text{state 1 : } A_1^{(1)} &= \begin{pmatrix} 0.5 & 0 \\ 0 & -0.5 \end{pmatrix}; Q^{(1)} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix} \\ \text{state 2 : } A_1^{(2)} &= \begin{pmatrix} 0.3 & 0 \\ 0 & -0.3 \end{pmatrix}; Q^{(2)} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix} \\ \text{state 3 : } A_1^{(3)} &= \begin{pmatrix} -0.5 & 0 \\ 0 & 0.5 \end{pmatrix}; Q^{(3)} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}\end{aligned}$$

Let the observations \mathbf{x}_t be generated as

$$\mathbf{x}_t = A_1^{(s_t)} \mathbf{x}_{t-1} + \varepsilon_t; \quad \varepsilon_t \sim \mathcal{MVN}(\mathbf{0}, Q^{(s_t)})$$

where the hidden state variables s_t are

$$s_t = \begin{cases} 1, & t \in [1, 300] \\ 2, & t \in [301, 600] \\ 3, & t \in [601, 900] \end{cases}$$

This will serve as the training data.

We then define a set of damaged states with the same AR coefficients, but whose errors are decoupled. This is intended to reflect subtle forms of damage that may not significantly change the readings from individual sensors, but rather affect the correlation between different regions of the structure²³

$$\begin{aligned} \text{state 4 : } A_1^{(4)} &= \begin{pmatrix} 0.5 & 0 \\ 0 & -0.5 \end{pmatrix}; Q^{(4)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \text{state 5 : } A_1^{(5)} &= \begin{pmatrix} 0.3 & 0 \\ 0 & -0.3 \end{pmatrix}; Q^{(5)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \text{state 6 : } A_1^{(6)} &= \begin{pmatrix} -0.5 & 0 \\ 0 & 0.5 \end{pmatrix}; Q^{(6)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Now that we've defined a series of healthy and damaged states, we construct a test time series to compare the performance of the SVAR model to that of a baseline using a VAR model. In order to simulate gradually evolving environmental and operational conditions, we construct the test time series such that it gradually shifts between the six states. Specifically, the test time series has observations $\{\mathbf{w}_1, \dots, \mathbf{w}_{500}\}$ with $\mathbf{w}_t \in \mathbb{R}^2$, generated as

$$\mathbf{w}_t = \left((1 - \alpha_t)A_1^{(i_t)} + \alpha_t A_1^{(j_t)} \right) \mathbf{w}_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim \mathcal{MVN}(\mathbf{0}, ((1 - \alpha_t)Q^{(i_t)} + \alpha_t Q^{(j_t)}))$

$$(i_t, j_t) = \begin{cases} (1, 2), & t \in [1, 100] \\ (2, 3), & t \in [101, 200] \\ (3, 4), & t \in [201, 300] \\ (4, 5), & t \in [301, 400] \\ (5, 6), & t \in [401, 500] \end{cases}$$

and $\alpha_t = 0.01(t \bmod 100)$ so that α_t linearly interpolates between each pair of states i_t and j_t .

Next, we train the SVAR and VAR models and evaluate their ability to detect damage. The train and test data are plotted in Figure 3. Since the training data are labeled, we can train the SVAR model by splitting the data along each state and learning the VAR parameters individually using the Yule-Walker method. Since this small simulated example lacks true transitions, we set the transition matrix Z to

$$Z = \begin{pmatrix} 0.98 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 \\ 0.01 & 0.01 & 0.98 \end{pmatrix}$$

We then fit the model to the test data and construct control charts from the standardized chi-square statistics with control lines set at the 99% quantile of the χ_2^2 distribution; the results are plotted in Figure 4. Note that the SVAR model has a stronger response to the simulated damage than the VAR model, as can be noted not only by the number of SVAR damage statistics that exceed the control line but also by the greater overall magnitude of the squared residuals for the SVAR model on the damaged states. The receiver operating characteristic (ROC) curves for the SVAR and VAR approaches are plotted in Figure 5 along with their area under curve (AUC) scores, which reinforce the stronger performance of the SVAR method. Examining the state probabilities, also in Figure 4, we observe that the state probabilities inferred for the

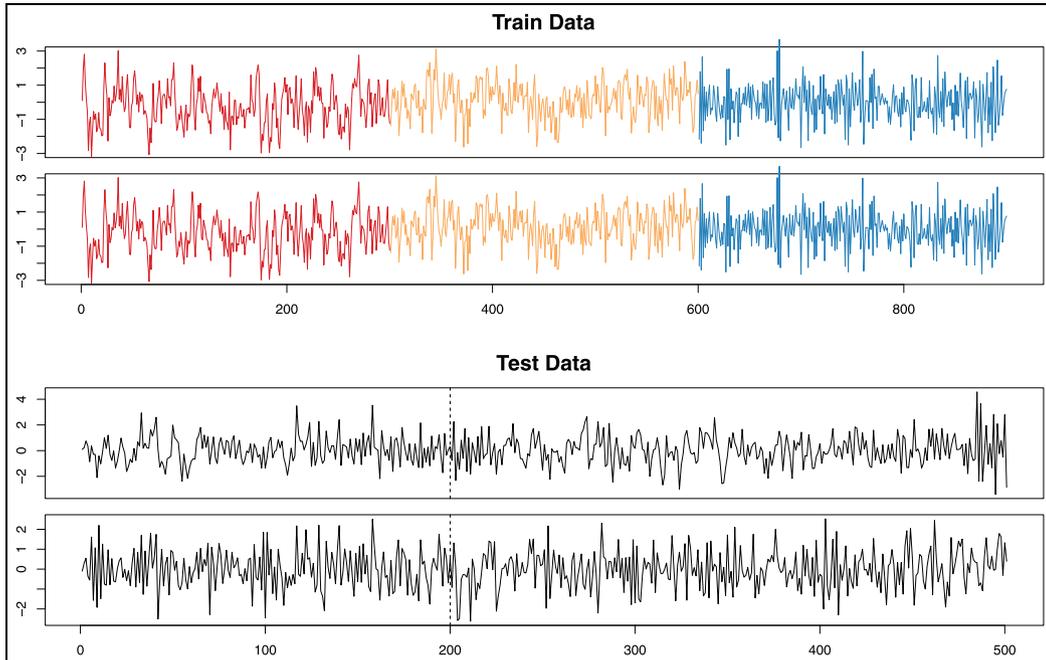


Figure 3. The train and test data for the simulated example. Each color in the train data represents a different state. For the test data, the dashed line indicates the point at which the first damage state begins to be introduced, for example, where state 3 begins to transition to state 4.

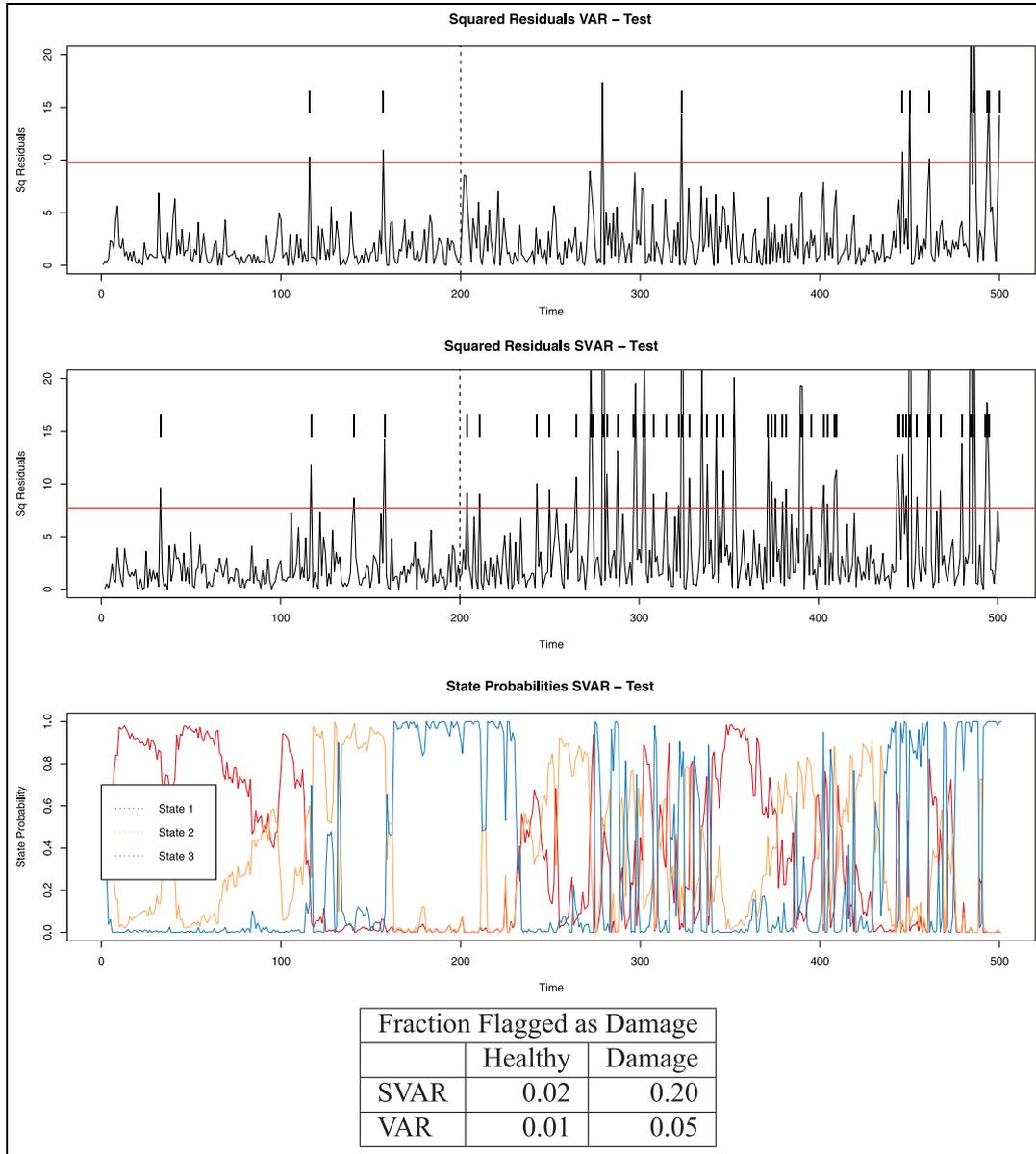


Figure 4. The 99% control charts for the residuals of the SVAR and VAR model in the simulated example, estimated state probabilities for the SVAR model, and summary table enumerating the fraction of the time damage are flagged for in the healthy and damaged data (e.g. the false- and true-positive rates, respectively). We define the “damaged” data as the time steps $t \in [201, 500]$ in the test data, as time $t = 201$ is when the first damaged state is introduced; in the residual plots, this time step is indicated by the dashed line. Both the residual plots and the table support that the SVAR model has a stronger response to damage than the VAR model at a similar false-positive rate, and that the SVAR model’s state probabilities correctly track the changing states even when the decoupling is introduced.

healthy states track correctly with the gradual changes in each state the test data, and that the state probabilities inferred for the damaged states mirror those of the undamaged states that share the same AR coefficients. This implies that, as desired, the SVAR model was able to correctly match up the damaged states in the test data to the closest healthy states and from there recognize that the signal was no longer behaving as expected.

Partial knowledge of environmental and operational conditions

We have now discussed inference and prediction methods for the SVAR model under either complete knowledge or complete lack of knowledge of the environmental and operational conditions; however, knowledge of the states is more often only partially incomplete. For example, some factors may be actively

measured while others are unobserved, as in the case of a vehicle whose temperature and speed are always monitored, but whose payload is not. As another example, there may be noisy measurements that imply a probability distribution across states, or the availability of measurements for particular factors may change over time.

Let $\{\mathbf{d}_1^{(l)}, \dots, \mathbf{d}_T^{(l)}\}$, where $\mathbf{d}_t^{(l)} = (d_{t,1}^{(l)}, \dots, d_{t,K}^{(l)})$ and $\sum_j d_{t,j}^{(l)} = 1$, be a multivariate time series representing our external knowledge of the factor l at each time t ; these reflect how our prior beliefs of the factor levels change over time. Note that complete knowledge is indicated by setting one of the $d_{t,k}^{(l)} = 1$ and the rest to 0, and complete lack of knowledge is indicated by uniformly setting $d_{t,k}^{(l)} = 1/K_l$ for each k .

Now let us reconsider the problem of computing the state probabilities. First, the forward pass becomes

$$P(s_t = j | \mathbf{x}_t, \mathbf{x}_{1:t-1}, \mathbf{d}_{1:t}) = \frac{1}{c_t} L_t(j) \sum_{i=1}^K [P(s_t = j | s_{t-1} = i, \mathbf{d}_{1:t}) P(s_{t-1} \mathbf{x}_{1:t-1}, \mathbf{d}_{1:t-1})]$$

Then, the backward pass becomes

$$P(s_t = j | \mathbf{x}_{1:T}, \mathbf{d}_{1:T}) = \sum_{j'=1}^K \left[\frac{P(s_t = j | \mathbf{x}_{1:t}, \mathbf{d}_{1:t}) P(s_{t+1} = j' | s_t = j, \mathbf{d}_{1:t}) P(s_{t+1} = j' | \mathbf{x}_{1:T}, \mathbf{d}_{1:T})}{P(s_{t+1} = j' | \mathbf{x}_{1:t}, \mathbf{d}_{1:t})} \right]$$

The key difference is that the transition probabilities are now dependent on the external knowledge

$$\begin{aligned} P(s_t = j | s_{t-1} = i, \mathbf{d}_{1:t}) &= \prod_{l=1}^L P(s_{t,l} = j | s_{t-1,l} = i, \mathbf{d}_{1:t}^{(l)}) \\ &= \frac{1}{c_t} \prod_{l=1}^L \mathbf{d}_t^{(l)} Z^{(l)}[i, j] \\ &= \frac{1}{c_t} \left(\prod_{l=1}^L \mathbf{d}_{t,j}^{(l)} \right) Z[i, j] \end{aligned}$$

where c_t is a normalization constant. We can characterize the influence of the external knowledge on the transition probabilities as a time-dependent transition matrix Z_t , such that

$$Z_t[i, j] = \frac{1}{c_t} \left(\prod_{l=1}^L \mathbf{d}_{t,j}^{(l)} \right) Z[i, j]$$

This new time-dependent transition matrix can then be inserted into the original forward-backward equations to compute the new externally dependent state probabilities. Note that in the case of complete knowledge of the states, the time-dependent transition matrix becomes a delta function, while in the no-knowledge case, the

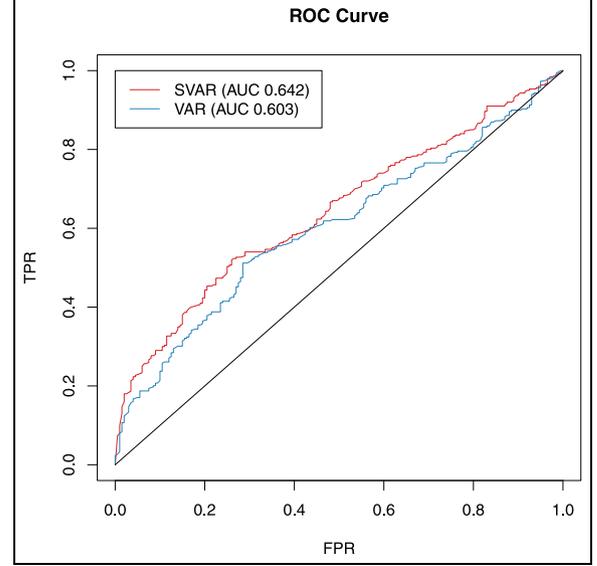


Figure 5. The ROC curves for the SVAR and VAR methods, comparing the false- and true-positive rates achieved by the two approaches. Here, by “false-positive,” we mean a time step where the residuals exceeded the threshold yet the underlying true state was healthy, and by “true-positive,” we mean a time step where the residuals exceeded the threshold and the underlying true state was damaged.

time-dependent transition matrix is equivalent to the original transition matrix, reflecting the desired impact of external knowledge on prediction and inference.

Learning the SVAR parameters under partial knowledge can be accomplished with the EM algorithm, the only change being that the E-step should incorporate the new time-dependent transition matrix into the computation of the state probabilities.

Related methods

In the following section, we elaborate upon some of the methods against which the SVAR model will be compared in the experimental study.

AR-SVM

Bornn et al.^{25,28} introduce the AR-SVM model to SHM, demonstrating how it can be used to model systems with a nonlinear dynamic response. The AR-SVM model is an application of the methods of support vector regression (SVR) to time series. In SVR, the objective is to find a vector w such that all responses y_i lie within a margin ε of the fitted values $f(x_i) = \langle w, x_i \rangle + b$, such that $\|w\|^2$ is as small as possible. The optimization and constraints are

$$\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\eta_t^+ + \eta_t^-)$$

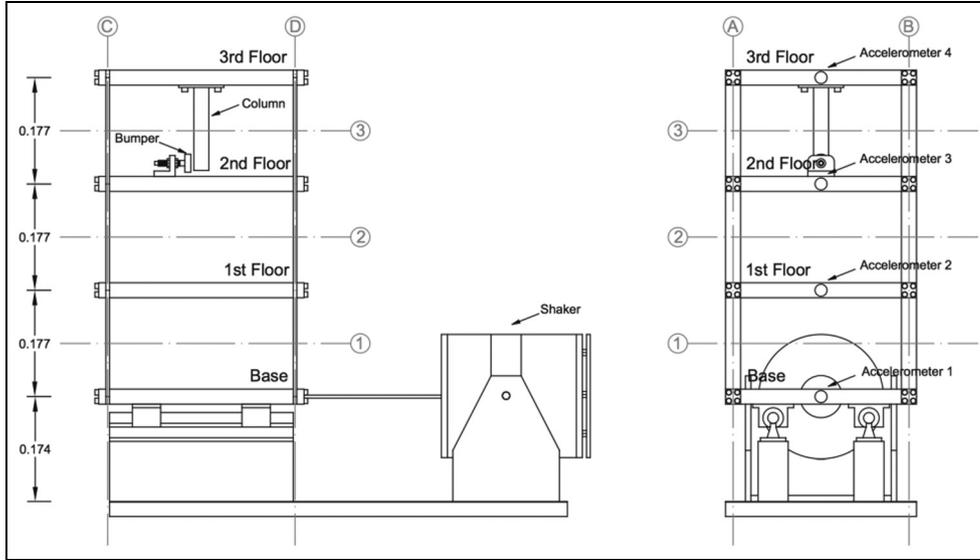


Figure 6. Diagram of the experimental structure.

$$\text{subject to } \begin{cases} y_i - \langle w, \mathbf{x}_i \rangle \leq \varepsilon + \eta_i^+ \\ \langle w, \mathbf{x}_i \rangle - y_i \leq \varepsilon + \eta_i^- \end{cases}$$

where the slack variables η_i^+ and η_i^- allow for inexact fits to the margin ε , and C is a constant controlling the trade-off between the size of w and the allowance of deviations beyond ε . When applied to time series, the (y_i, \mathbf{x}_i) correspond to the pairs $(\mathbf{x}_t, \mathbf{x}_{t-p:t-1})$, where p is the order of the model. Bornn et al.²⁵ apply the AR-SVM to a sensor network by fitting a separate AR-SVM model to each sensor. The residuals were then standardized to construct approximately χ_m^2 -distributed statistics that could be used in control charts for damage detection.

Minor PCA

Cross et al.¹² demonstrate several methods for identifying linear combinations of features robust to EOV, most prominently minor PCA and cointegration. In the following study, we focus on minor PCA, as cointegration requires larger data sets than available in this experiment.

PCA is a statistical tool that identifies a basis spanning the data whose first component has the greatest variance possible across the data, and whose each subsequent component has the greatest variance possible across the subspace of the data orthogonal to the preceding components. Therefore, by transforming the data along the last few principal components, new features can be constructed that have lower variance across the different EOV; this method is also known as minor PCA.

Our implementation of minor PCA follows the methods of Cross et al.¹² First, in order to construct the features, the training and test sets were divided into

100-point samples (e.g. such that there were 160 total samples in the training set and 40 in each test set). Then, the Fourier transform of each sample was computed and spectral lines 16 through 30 extracted to construct the data matrix.

Since minor PCA is only developed in a univariate context in Cross et al.,¹² we constrain ourselves to using the data obtained from the accelerometer on the third floor, the sensor exhibiting the most sensitivity to damage. We then apply PCA to the 160×15 data matrix extracted from the training data and compute the five least variable components; the training and test sets are then transformed along these components. The distance of the features corresponding to each chunk in the test set from the mean of the training features is computed to identify damage, using Mahalanobis squared distance as a measure of *discordance*

$$D(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}})^T S^{-1} (\mathbf{x} - \bar{\mathbf{x}})$$

where $\bar{\mathbf{x}}$ is the mean of the (transformed) features of the training data and S is the sample covariance matrix. A threshold for damage is determined using a Monte Carlo method and any point in the test set exceeding the distance threshold is flagged for damage.¹²

Experiment

Test structure

The performance of each model was evaluated on data from a laboratory test structure, depicted in Figure 6. The structure is a three-story building made of four $30.5 \times 30.5 \times 2.5$ cm aluminum plates held up by four

Table 1. A description of the different states present in the data set. Gap refers to the distance between the bumper and column used to simulate damage; mass (location) refers to the size of the weight and where the weight was placed. Stiffness reduction refers to halving the thickness of the listed column. Omissions imply default conditions: no gap, no mass, normal stiffness.

Dataset				
Index	Gap	Mass (base)	Mass (first floor)	Stiffness
1	–	–	–	–
2	–	1.2 kg	–	–
3	–	–	1.2 kg	–
4	–	–	–	1BD
5	–	–	–	1AD and 1BD
6	–	–	–	2BD
7	–	–	–	2AD and 2BD
8	–	–	–	3AD
9	–	–	–	3AD and 3BD
10	0.20 mm	–	–	–
11	0.15 mm	–	–	–
12	0.13 mm	–	–	–
13	0.10 mm	–	–	–
14	0.05 mm	–	–	–
15	0.20 mm	1.2 kg	–	–
16	0.20 mm	–	1.2 kg	–
17	0.10 mm	–	1.2 kg	–

17.7 × 2.5 × 0.6 cm aluminum columns each. The building is attached at its base to a shaker that agitates the structure horizontally. Four accelerometers with sensitivities of 1000 mV/g were mounted at the center of each plate to record the response of the structure to the excitation of the shaker. Damage was simulated through the impact of a 15.0 × 2.5 × 2.5 cm column suspended from the ceiling of the top floor with a bumper on the floor which could be adjusted to modify the extent of impact. Different environmental and operational conditions were simulated by adding weights to various floors and by adjusting the stiffness (equivalently, the thickness) of various columns, as described in Figueiredo and Flynn.²⁹ Each simulated time series is composed of 8192 data points sampled at 3.125 ms intervals, and the structure was excited by random excitation in the range of 20–150 Hz. The different environmental and operational states simulated in the data set are summarized in Table 1.

This test structure is designed to be a laboratory-scale way to capture the effects of environmental and operational variability that one of the co-authors directly measured during field tests of two in situ highway bridges.^{9,30} The varying mass is intended to simulate the changes in the bridges' dynamic response properties caused by the mass loading of vehicles, and the changes in stiffness represented by varying the column thicknesses are intended to simulate the effects of day–night temperature differentials that were observed in the bridge study reported in Farrar et al.⁹ The changes in mass and column thicknesses are designed

to produce approximately the same changes in the first mode frequency of the test structure as observed on the in situ bridges resulting from varying traffic loads or thermal conditions (both of which produced an approximately 5% change in the first mode frequency of the respective structures). The actual bridge data were not used because traffic was not allowed on the I-40 bridge once damage was introduced and damage was not introduced into the Alamosa Canyon Bridge. Although the structure only offers an idealization of these environmental and operational effects, it provides a reasonable approximation to the variability associated with the actual measured dynamic response of in situ structures coupled with simulated damage scenarios. Furthermore, this structure lets researchers simulate many combinations of idealized sources of variability and damage, data which are difficult to obtain from in situ structures.

Randomization study

A randomization study was used to evaluate and compare the performance of four different damage detection and data normalization methods on the data obtained from the laboratory structure: (1) the VAR model, (2) the AR-SVM model, (3) the SVAR model, and (4) minor PCA.

Taking the perspective of SHM as a novelty detection problem, there are two different types of novelty that a data normalization algorithm can encounter: damage in the structure, or an unfamiliar set of

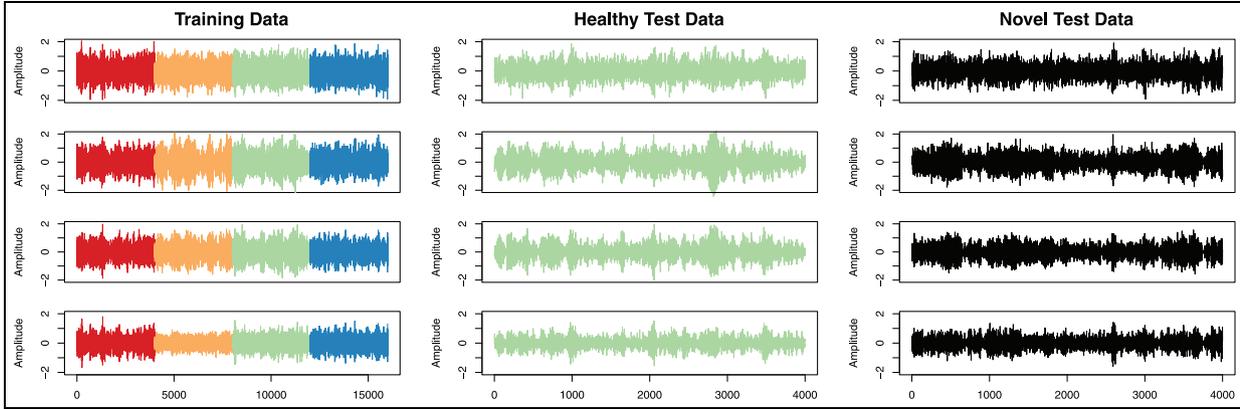


Figure 7. A plot of the time series from a sample iteration of the randomization study; each row represents a different sensor. The training data are composed of samples from states 1, 9, 8, and 6, in that order; the healthy test data are a different sample from state 8; the novel test data are a sample from state 10. The different colors in the training data sample demarcate the different states.

environmental and operational conditions. In order to study these two sources of novelty separately, we conducted two variations of the randomization study, one in which the models were evaluated upon a damaged state (e.g. one of states 10–17) with EOV matched to the states in the training set, and one in which the novel states were healthy states (e.g. one of states 1–9) not present in the training set. In this way, we ensured that in each study, only one type of novelty was present. We henceforth refer to these two variations of the randomization study as the *matched* study and *novel EOV* study, respectively. The states selected for train and test in each iteration of the each study are detailed in Tables 2 and 3 in Appendix 1.

In each iteration of the randomization study, four random healthy states and one random novel state were selected. A 16,000-point training set was then constructed by concatenating 4000-point segments of each selected healthy state. After fitting the models to the training set, the models were applied to two different test sets: (1) a 4000-point segment from a second realization of one of the selected healthy states and (2) a 4000-point segment from a novel state. The fraction of points flagged as novel by each model on each test set was then computed, providing a measure of the false- and true-positive rates of each algorithm. An example of the training and test data constructed for the model is plotted in Figure 7.

We now present a more detailed illustration of how the experiments were conducted, using the experiment described by the first row of Table 2 as an example. For this iteration, the signal for training the model was constructed by concatenating 4000-point segments from realizations of states 1, 7, 6, and 9, in that order (specifically using time steps $t \in [2001, 6000]$ from the realization of each state). Then, the healthy segment of

the test signal was constructed by extracting time steps $t \in [2001, 6000]$ from a second realization of one of the training states—in this example, state 1. Since this was an iteration of the *matched* study, the novel test state was randomly selected from the subset of damaged states whose environmental and operational condition matched at least one of the healthy states, and in this example, state 10 was selected, whose conditions match those of state 1 (for the *novel EOV* study, the novel segment would have been selected from the complement of the healthy states used in training, for example, one of states 2–5 or 8). The novel segment of the test signal was then constructed by extracting time steps $t \in [2001, 6000]$ from a realization of state 10, and the complete test data constructed by concatenating the novel segment to the healthy segment. Each method was then trained and evaluated accordingly on the train and test data.

An important note is that the damage-detection algorithms for the SVAR, VAR, and AR-SVM methods were slightly modified for the randomization study. In this study, we used the average damage statistic across 100 time-step segments rather than the damage statistic at each time step. This provided the SVAR, VAR, and AR-SVM methods a coarser resolution, and in particular allowed for correct comparison between those methods with the PCA method, which required chunking the data into 100 time step segments for accurate spectral analysis. The damage statistics used are still the ones outlined in the methodology section above: marginal likelihood for the SVAR and VAR methods, the chi-square damage statistic for the AR-SVM method, and discordance for the PCA method.

Results

The true-positive rates for each study (using 99% control lines) are summarized in Figure 8, and ROC curves

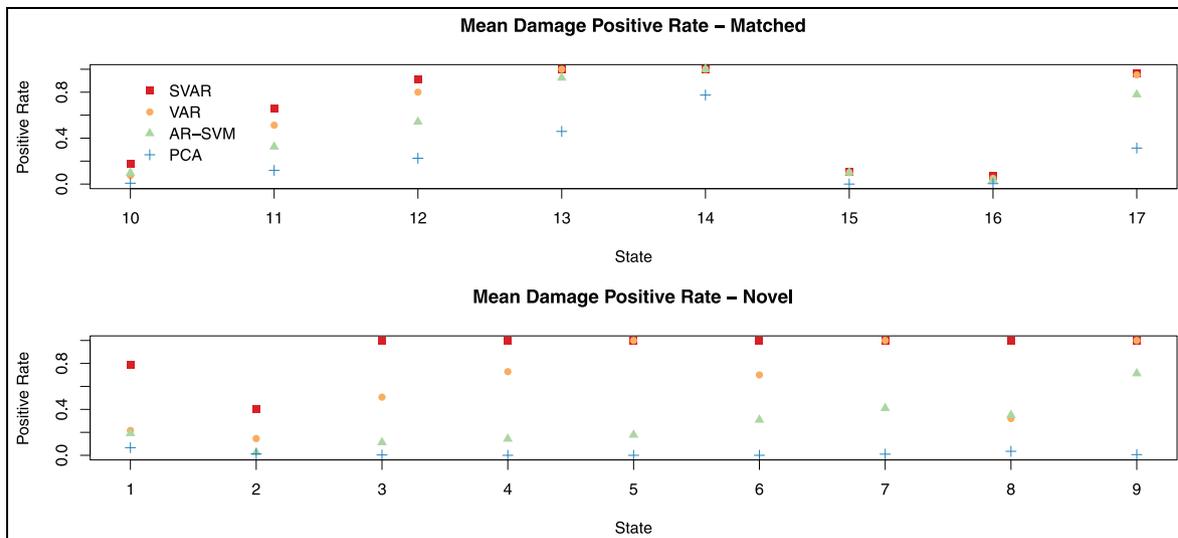


Figure 8. Mean true-positive rate for states in the matched study (above) and novel EOV study (below), i.e., the percentage of time steps flagged in the novel test case in each study. Note that sometimes when the true-positive rate is near 100%, as in states 5, 7, 9, 13, 14, and 15, the SVAR marker is occluded by the VAR marker.

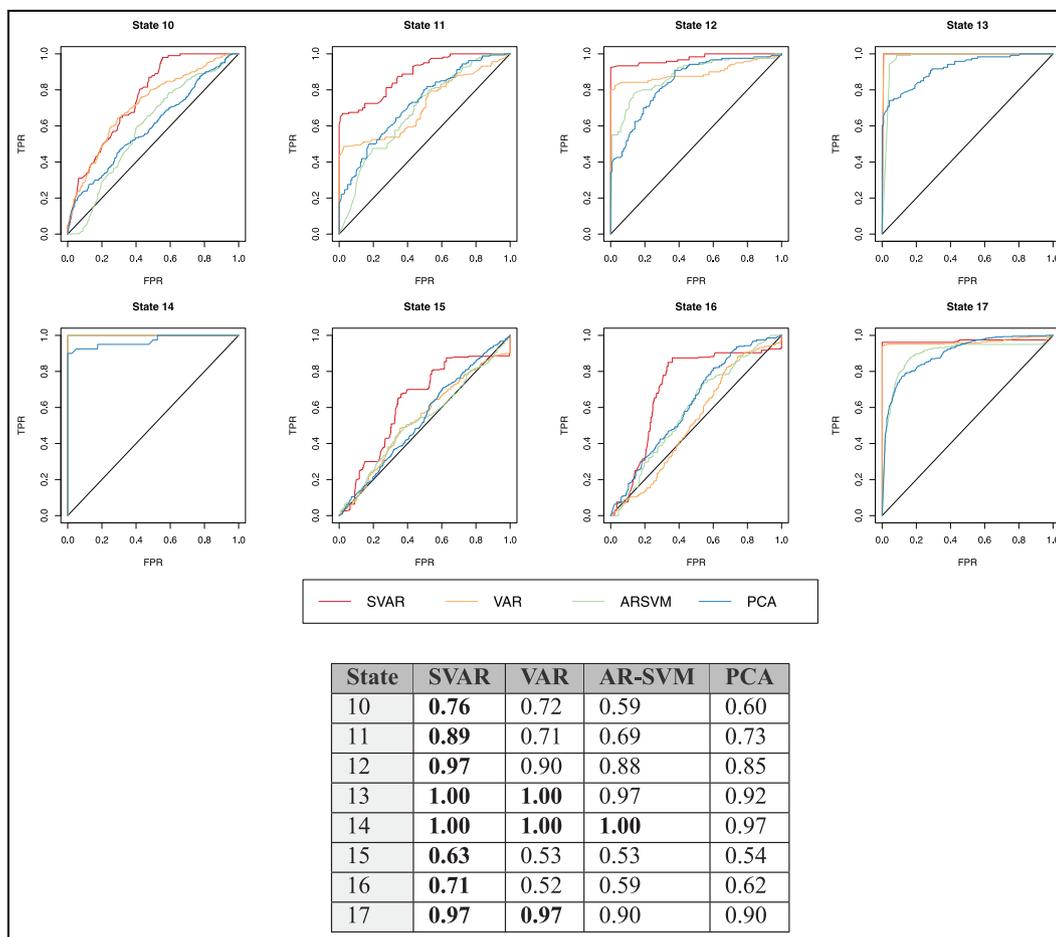


Figure 9. Matched study: Comparison of ROC curves obtained with the SVAR, VAR, AR-SVM, and minor PCA methods in the matched study. In each case, the models are trained on healthy states, then tested on a replication of one of the training states and a novel, damaged state whose environmental and operational conditions match one of the training states.

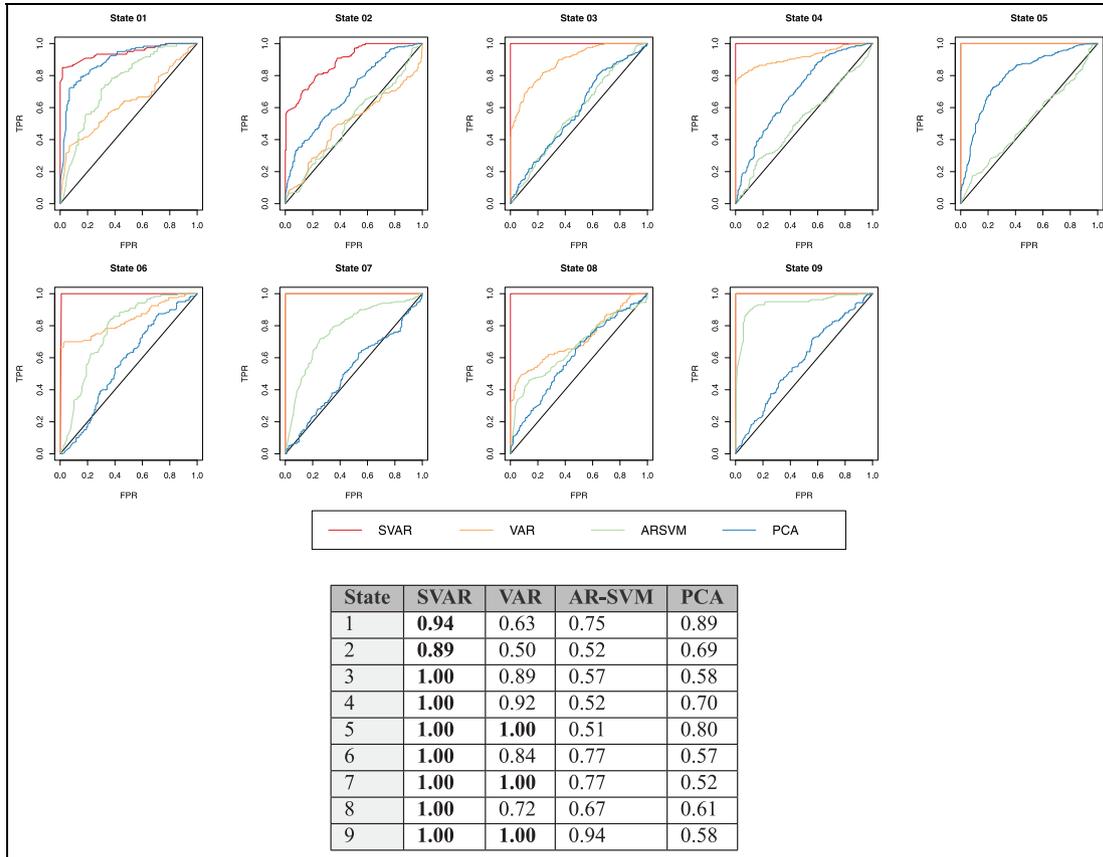


Figure 10. Novel EOJ study: Comparison of ROC curves obtained with the SVAR, VAR, AR-SVM, and minor PCA methods in the novel study. In each case, the models are trained on healthy states, then tested on a replication of one of the training states and a novel, undamaged state with distinct environmental and operational conditions.

and ROC-AUC are presented in Figures 9 and 10. There are detailed summaries of each iteration of the randomization study in Tables 2 and 3 of Appendix 1.

Matched study. In the *matched* study, the SVAR method consistently yields the highest AUC and true-positive rates, though with other methods performing similarly at the higher damage levels. All methods tend to perform better as the severity of damage increases.

The high performance of the VAR and SVAR methods suggests that AR methods in general are well suited for identifying damage of this type. The left-heavy shape of many of the ROC curves of the AR models suggests that they yielded large damage statistics whenever there was an impact, but otherwise fit the data closely. This is as expected, since the behavior of the structure in these damaged states only deviates at the impacts. Relatedly, the PCA method's slightly worse performance may be attributable to it having less resolution by operating in the frequency domain, and therefore not reacting as severely as the AR methods.

The state probability estimates of the SVAR model for state 10 in one of the *matched* study iterations are plotted in Figure 11; note that the SVAR method correctly identifies the EOJ underlying state 10 to be associated with state 1. The ability to identify the underlying state with great fidelity suggests that the gain in model performance of the SVAR over the VAR method can be explained by the fact that the SVAR method learns a separate model for each state it encounters in training. In contrast to other methods, which must fit a single model across multiple states, this allows the SVAR model to maintain "stricter expectations" for the behaviors it observes, increasing its relative sensitivity to novelty.

Novel EOJ study. In the *novel EOJ* study, the SVAR model again yields the highest AUC and true-positive rates. Note that the SVAR model has higher AUC scores in this study than in the *matched* study; this may be attributable to the sparsity of impacts in the damaged states, for example, while the novelty in the

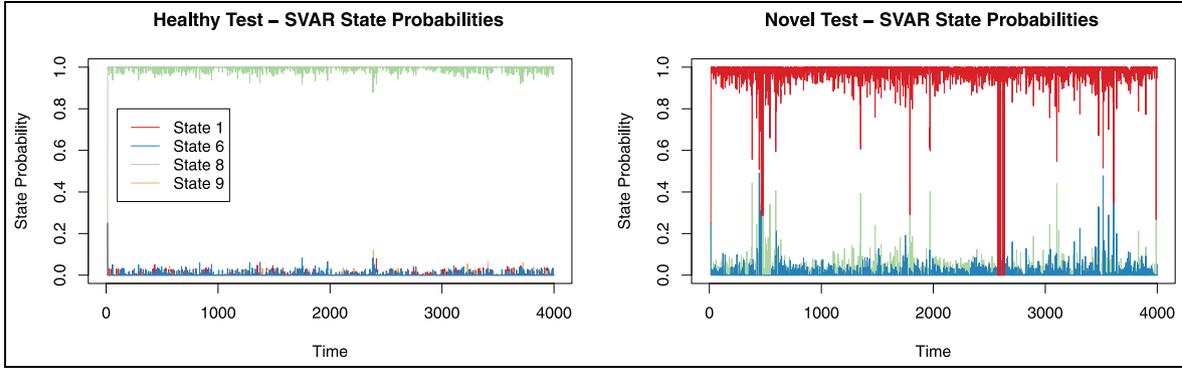


Figure 11. The probabilities of each state estimated by the SVAR model at each time step on an iteration of the matched randomization study. The training states are 1, 9, 8, and 7; the healthy test state is state 8; the novel test state is state 10. This is the same iteration from the matched randomization study test data as in Figure 7.

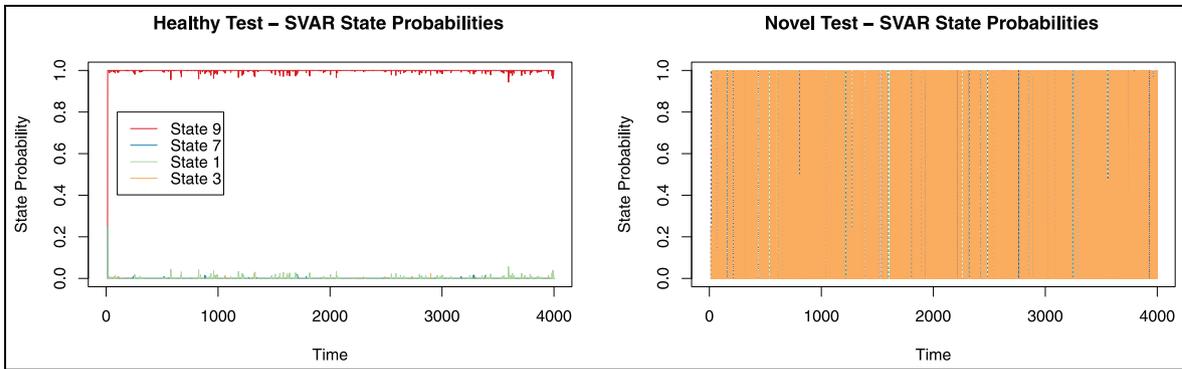


Figure 12. The probabilities of each state estimated by the SVAR model at each time step on an iteration of the novel EOV randomization study. The training states are 9, 3, 1, and 7; the healthy test state is state 9; the novel test state is state 8.

matched study is intermittent, the novelty in *novel EOV* study is continuous, perhaps leading to a higher rate of true positives.

The same “stricter expectations” explanation may explain the SVAR model’s performance in the *novel EOV* study, as the specificity with which the SVAR model learns to recognize each state means that new conditions that do not resemble any of the states upon which the model was trained will be identified as novel. In Figure 12, we provide an example of the SVAR model estimating state probabilities for an iteration of the *novel EOV* study; note how on the novel test case the model does not settle on a confident belief in any of the four training states, in contrast to the example shown in the *matched* study.

One way to interpret these results is that there is a limit to the robustness of the ability of the SVAR model to average between the behavior of its known states to approximate the behavior of new states, and therefore that care must be taken in arranging comprehensive training data for the SVAR approach. In contrast, since the other methods have learned to expect greater

variance in the behavior of healthy structures, they are less sensitive to the appearance of a new state.

Conclusion

This article proposes a novel method for vibration-based damage detection robust to EOV that unites the AR novelty detection framework with switching state-space models. In contrast with other data normalization methods, the SVAR and other state-space models utilize the full information provided by the training data, learning the behavior of each regime of environmental and operational conditions separately rather than only summarizing the common trends across them. As demonstrated by the experimental results, this leads to greater sensitivity to damage without raising the false-positive rate. In addition, the state-space model framework provides a way to infer the underlying environmental and operational conditions, making it practical for situations where there is only partial knowledge of the EOV in the training or test data.

An important simplification in this article was the choice to model the EOVS with discrete states, an assumption enabled by the dataset but not reflected by many real-world sources of EOVS. While the model is still effective in the continuous case because of its ability to probabilistically average across different states, defining effective methods for discretizing continuous processes so that the resulting SVAR model learns a representative set of models warrants further study. The discretization could be defined manually or, noting that defining discrete bins is similar to the unlabeled case, learned using a method such as EM. It may be possible to hybridize the discrete state-space framework described here with the continuous models used by Worden et al.¹⁴ and, in particular, to develop a model where some components are discrete and others are handled continuously.

Another interesting direction would be to incorporate more complex model components, such as the LPV-AR and FS-TAR models used in Avendaño-Valencia and Fassois,¹⁵ within the state-space approach, as that may allow for better capturing of the behavior of each individual state. It would also be of interest to apply the state-space model to the problem of damage identification, i.e., not only detecting the presence of damage but also classifying what type of damage had occurred.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this research was supported in part by U.S. National Science Foundation (Grant No. 1461435), by DARPA under (Grant No. FA8750-14-2-0117), by ARO under (Grant No. W911NF-15-1-0172), and by NSERC.

References

- Worden K. Structural fault detection using a novelty measure. *J Sound Vib* 1997; 201(1): 85–101.
- Doebbling S, Farrar C, Prime M, et al. A review of damage identification methods that examine changes in dynamic properties. *Shock Vib Digest* 1998; 30: 91–105.
- Carden EP and Fanning P. Vibration based condition monitoring: a review. *Struct Health Monit* 2004; 3(4): 355–377.
- Fan W and Qiao P. Vibration-based damage identification methods: a review and comparative study. *Struct Health Monit* 2011; 10(1): 83–111.
- Pandit SM and Wu S. *Time series and system analysis with applications*. New York: Wiley, 1983.
- Fugate ML, Sohn H and Farrar CR. Vibration-based damage detection using statistical process control. *Mech Syst Signal Pr* 2001; 15: 707–721.
- Mosavi A, Dickey D, Seracino R, et al. Identifying damage locations under ambient vibrations utilizing vector autoregressive models and Mahalanobis distances. *Mech Syst Signal Pr* 2012; 26: 254–267.
- Weijtens W, Verbelen T, De Sitter G, et al. Data normalization for foundation SHM of an offshore wind turbine: a real-life case study. In: *Proceedings of the 7th European workshop on structural health monitoring*, Nantes, 8–11 July 2014, pp. 788–795.
- Farrar CR, Cornwell PJ, Doebling SW, et al. *Structural health monitoring studies of the Alamosa Canyon and I-40 bridges*. Technical report LA-13635-MS, July 2000. Los Alamos, NM: Los Alamos National Laboratory.
- Alampalli S. Influence of in-service environment on modal parameters. In: *Proceedings of the 16th international modal analysis conference*, Santa Barbara, CA, 2–5 February 1998, p. 111.
- Figueiredo E, Park G, Farrar CR, et al. Machine learning algorithms for damage detection under operational and environmental variability. *Struct Health Monit* 2011; 10: 559–572.
- Cross EJ, Manson G, Worden K, et al. Features for damage detection with insensitivity to environmental and operational variations. *P Roy Soc A: Math Phy* 2012; 468: 4098–4122.
- Farrar CR, Sohn H and Worden K. *Data normalization: a key for structural health monitoring*. Technical report LA-UR-01-4212, 1 January 2001. Los Alamos, NM: Los Alamos National Laboratory.
- Worden K, Cross E and Brownjohn J. Switching response surface models for structural health monitoring of bridges. In: Koziel S and Leifsson L (eds) *Surrogate-based modeling and optimization*. New York: Springer Science + Business Media, 2013, pp. 337–358.
- Avendaño-Valencia LD and Fassois SD. Damage/fault diagnosis in an operating wind turbine under uncertainty via a vibration response Gaussian mixture random coefficient model based framework. *Mech Syst Signal Pr* 2017; 91: 326–353.
- Cryer JD and Chan K. *Time series analysis: with applications in R*. 2nd ed. New York: Springer, 2008.
- Akaike H. A new look at the statistical model identification. *IEEE T Automat Contr* 1974; 19(6): 716–723.
- Schwarz G. Estimating the dimension of a model. *Ann Stat* 1978; 6(2): 461–464.
- Figueiredo E, Figueiras J, Park G, et al. Influence of the autoregressive model order on damage detection. *Comput-Aided Civ Inf* 2011; 26: 225–238.
- Montgomery D. *Introduction to statistical quality control*. New York: John Wiley & Sons, Inc., 2012.
- Bodeux J and Golinval J. Modal identification and damage detection using the data-driven stochastic subspace and ARMAV methods. *Mech Syst Signal Pr* 2003; 17(1): 83–89.

22. De Stefano A, Sabia D and Sabia L. Structural identification using VARMA models from noisy dynamic response under unknown random excitation, structural damage assessment using advanced signal processing procedures. In: *Proceedings of DAMAS 97*, Sheffield, 30 June–2 July 1997, pp. 419–429.
23. Bornn L, Farrar CR, Higdon D, et al. Modeling and diagnosis of structural systems through sparse dynamic graphical models. *Mech Syst Signal Pr* 2016; 74: 133–143.
24. Dzunic Z, Chen JG, Mobahi H, et al. A Bayesian state-space approach for damage detection and classification. *Mech Syst Signal Pr* 2017; 96: 239–259.
25. Bornn L, Farrar CR, Park G, et al. Structural health monitoring with autoregressive support vector machines. *J Vib Acoust* 2009; 131: 021004.
26. Murphy K. Switching Kalman filters. Technical report, 21 August 1998. The University of British Columbia, Vancouver, BC, Canada, 1998.
27. Quinn JA, Williams CK and McIntosh N. Factorial switching linear dynamical systems applied to physiological condition monitoring. *IEEE T Pattern Anal* 2009; 31: 1537–1551.
28. Bornn L, Farrar CR and Park G. Damage detection in initially nonlinear systems. *Int J Eng Sci* 2010; 48: 909–920.
29. Figueiredo E and Flynn E. Three-story building structure to detect nonlinear effects. Technical report, Los Alamos National Laboratory, Los Alamos, NM, 2009.
30. Farrar CR, Baker WE, Bell TM, et al. *Dynamic characterization and damage detection in the I-40 bridge over the Rio Grande*. Technical report LA-12767-MS, 1 June 1994. Los Alamos, NM: Los Alamos National Laboratory.

Appendix I

Table 2. Detailed results of the false- and true-positive rates of each iteration of the matched randomization study, as well as the states featured in each iteration.

Matched EOv study										
Training states	Healthy state	Novel state	SVAR		VAR		SVM		Minor PCA	
			Healthy	Novel	Healthy	Novel	Healthy	Novel	Healthy	Novel
1, 7, 6, 9	1	10	0.13	0.25	0.03	0.15	0.25	0.13	0.08	0.20
3, 4, 9, 1	9	12	0.00	0.93	0.00	0.83	0.03	0.55	0.00	0.35
3, 6, 4, 8	3	17	0.00	0.95	0.00	0.95	0.08	0.78	0.03	0.48
1, 7, 6, 9	7	11	0.00	0.68	0.00	0.53	0.05	0.43	0.00	0.25
2, 4, 6, 9	9	15	0.00	0.10	0.00	0.10	0.10	0.08	0.03	0.05
1, 8, 9, 3	3	10	0.00	0.10	0.05	0.05	0.08	0.03	0.03	0.20
1, 8, 6, 7	7	13	0.00	1.00	0.00	1.00	0.00	0.90	0.00	0.65
2, 6, 3, 8	2	16	0.00	0.08	0.00	0.05	0.03	0.03	0.10	0.05
2, 6, 9, 1	2	15	0.00	0.10	0.00	0.10	0.03	0.05	0.00	0.00
3, 2, 8, 9	8	15	0.03	0.13	0.00	0.10	0.18	0.08	0.08	0.03
2, 1, 7, 5	1	10	0.08	0.35	0.00	0.05	0.25	0.18	0.10	0.13
2, 3, 8, 1	2	15	0.00	0.10	0.00	0.10	0.00	0.03	0.00	0.00
2, 3, 7, 1	1	17	0.13	0.98	0.00	0.95	0.25	0.75	0.08	0.43
2, 5, 9, 7	9	15	0.00	0.13	0.05	0.10	0.18	0.15	0.03	0.08
3, 9, 4, 2	3	17	0.00	0.95	0.00	0.95	0.10	0.80	0.05	0.53
3, 5, 7, 4	7	17	0.00	0.95	0.03	0.95	0.08	0.78	0.03	0.70
1, 9, 8, 6	8	10	0.03	0.10	0.00	0.05	0.08	0.05	0.00	0.13
3, 7, 9, 5	7	17	0.00	0.98	0.03	0.95	0.08	0.78	0.03	0.70
1, 3, 2, 6	6	16	0.08	0.08	0.00	0.05	0.10	0.03	0.03	0.00
1, 4, 5, 8	8	13	0.03	1.00	0.03	1.00	0.08	0.90	0.03	0.65
2, 6, 3, 4	4	15	0.00	0.10	0.03	0.10	0.05	0.15	0.00	0.05
1, 6, 2, 7	1	11	0.03	0.68	0.00	0.53	0.25	0.38	0.05	0.20
3, 7, 4, 5	4	16	0.00	0.05	0.00	0.05	0.03	0.03	0.05	0.10
3, 8, 9, 7	3	16	0.00	0.08	0.03	0.05	0.10	0.05	0.05	0.10
1, 7, 6, 3	3	10	0.00	0.25	0.08	0.10	0.10	0.18	0.03	0.20
1, 6, 2, 4	1	10	0.00	0.05	0.00	0.05	0.15	0.05	0.03	0.08
2, 9, 1, 6	6	10	0.08	0.10	0.00	0.15	0.05	0.10	0.00	0.05
1, 6, 2, 3	1	16	0.03	0.08	0.00	0.05	0.28	0.03	0.08	0.00
2, 4, 3, 5	4	17	0.00	0.95	0.00	0.95	0.03	0.80	0.00	0.60
1, 4, 2, 7	2	12	0.00	0.90	0.00	0.83	0.00	0.53	0.00	0.30
2, 5, 8, 3	5	17	0.03	0.98	0.05	0.95	0.05	0.80	0.00	0.53
2, 7, 3, 9	9	15	0.00	0.13	0.05	0.10	0.15	0.13	0.00	0.05
3, 5, 8, 2	8	15	0.03	0.13	0.03	0.10	0.20	0.13	0.08	0.13
2, 1, 3, 4	1	11	0.00	0.60	0.00	0.55	0.18	0.23	0.08	0.23
3, 9, 8, 2	8	17	0.03	0.98	0.00	0.95	0.18	0.80	0.05	0.48
3, 8, 5, 9	5	16	0.03	0.08	0.08	0.05	0.05	0.08	0.00	0.15
1, 3, 5, 6	5	13	0.03	1.00	0.05	1.00	0.05	0.98	0.00	0.63
1, 5, 8, 6	5	11	0.03	0.68	0.08	0.45	0.05	0.28	0.00	0.25
2, 8, 7, 4	7	15	0.00	0.10	0.00	0.10	0.05	0.08	0.00	0.08
1, 8, 9, 4	8	10	0.00	0.05	0.00	0.05	0.08	0.05	0.00	0.15
1, 7, 3, 9	1	10	0.13	0.25	0.00	0.05	0.28	0.13	0.10	0.23
1, 4, 2, 6	4	12	0.00	0.90	0.03	0.75	0.00	0.55	0.00	0.30
2, 6, 9, 7	2	15	0.00	0.10	0.00	0.10	0.03	0.18	0.10	0.05
3, 4, 8, 7	8	17	0.03	0.95	0.00	0.95	0.10	0.73	0.05	0.55
2, 9, 7, 4	2	15	0.00	0.10	0.00	0.10	0.03	0.08	0.15	0.05
2, 1, 6, 7	2	14	0.00	1.00	0.00	1.00	0.03	1.00	0.00	0.90
1, 9, 2, 5	2	10	0.00	0.30	0.00	0.05	0.03	0.10	0.03	0.18
3, 6, 4, 7	4	17	0.00	0.95	0.03	0.95	0.05	0.78	0.00	0.60
2, 8, 7, 3	8	16	0.03	0.08	0.03	0.05	0.15	0.03	0.05	0.08
2, 8, 5, 3	3	17	0.00	0.98	0.00	0.95	0.10	0.80	0.05	0.58

EOV: environmental and operational variation; SVAR: switching vector autoregressive; VAR: vector autoregressive; SVM: support vector machine; PCA: principal components analysis.

Table 3. Detailed results of the false- and true-positive rates in each iteration of the novel EOV randomization study, as well as the states featured in each iteration.

Novel EOV study											
Training states	Healthy state	Novel state	SVAR		VAR		SVM		Minor PCA		
			Healthy	Novel	Healthy	Novel	Healthy	Novel	Healthy	Novel	
5, 4, 1, 6	1	3	0.03	1.00	0.08	0.25	0.15	0.03	0.15	0.03	
8, 2, 4, 9	2	3	0.00	1.00	0.00	0.53	0.03	0.15	0.05	0.03	
9, 5, 7, 1	7	3	0.00	1.00	0.00	0.15	0.05	0.13	0.03	0.10	
8, 7, 3, 9	7	2	0.00	1.00	0.03	0.45	0.08	0.03	0.03	0.18	
9, 3, 1, 7	9	8	0.00	1.00	0.03	0.05	0.10	0.38	0.00	0.05	
8, 2, 7, 4	8	5	0.03	1.00	0.00	1.00	0.08	0.15	0.08	0.00	
1, 2, 7, 6	6	5	0.08	1.00	0.00	1.00	0.08	0.40	0.03	0.03	
8, 3, 6, 7	8	1	0.03	1.00	0.03	0.35	0.15	0.15	0.08	0.38	
9, 6, 5, 4	4	8	0.00	1.00	0.00	0.15	0.05	0.33	0.00	0.10	
1, 3, 7, 9	7	2	0.00	0.33	0.00	0.00	0.08	0.05	0.00	0.03	
6, 1, 9, 7	6	4	0.05	1.00	0.00	1.00	0.05	0.20	0.00	0.00	
7, 6, 8, 9	7	3	0.00	1.00	0.00	1.00	0.08	0.18	0.00	0.05	
4, 6, 9, 8	4	1	0.00	1.00	0.05	0.13	0.00	0.18	0.00	0.18	
9, 8, 4, 3	8	5	0.00	1.00	0.00	1.00	0.10	0.15	0.08	0.00	
6, 4, 2, 1	6	5	0.00	1.00	0.00	1.00	0.05	0.13	0.03	0.00	
6, 4, 7, 3	4	2	0.00	1.00	0.03	0.43	0.05	0.08	0.00	0.18	
6, 5, 9, 7	6	8	0.08	1.00	0.00	0.38	0.18	0.45	0.10	0.10	
2, 7, 3, 8	7	4	0.00	1.00	0.00	0.98	0.08	0.10	0.00	0.00	
8, 1, 5, 6	8	2	0.03	0.05	0.03	0.00	0.08	0.00	0.03	0.00	
2, 5, 7, 6	6	4	0.08	1.00	0.00	0.10	0.13	0.13	0.08	0.00	
5, 1, 8, 4	8	7	0.03	1.00	0.03	1.00	0.08	0.43	0.03	0.08	
2, 3, 7, 9	2	5	0.03	1.00	0.00	1.00	0.03	0.20	0.08	0.00	
7, 3, 6, 1	6	8	0.08	1.00	0.00	1.00	0.08	0.48	0.00	0.20	
6, 1, 2, 3	1	7	0.03	1.00	0.00	1.00	0.28	0.35	0.03	0.00	
2, 4, 6, 8	2	3	0.00	1.00	0.00	0.58	0.03	0.10	0.03	0.08	
6, 4, 1, 9	4	2	0.00	0.03	0.03	0.00	0.00	0.00	0.00	0.00	
3, 2, 6, 8	2	4	0.00	1.00	0.00	0.98	0.03	0.10	0.03	0.00	
7, 1, 3, 2	1	6	0.13	1.00	0.00	0.10	0.25	0.15	0.10	0.00	
4, 1, 7, 9	1	3	0.03	1.00	0.00	0.35	0.18	0.13	0.10	0.05	
2, 9, 6, 4	9	7	0.00	1.00	0.00	1.00	0.10	0.35	0.03	0.00	
6, 1, 2, 8	1	4	0.00	1.00	0.00	1.00	0.13	0.10	0.00	0.00	
1, 9, 5, 3	5	6	0.03	1.00	0.08	1.00	0.05	0.23	0.00	0.03	
3, 2, 5, 1	3	4	0.03	1.00	0.00	0.05	0.10	0.13	0.03	0.00	
4, 9, 1, 8	8	7	0.00	1.00	0.00	1.00	0.08	0.35	0.00	0.03	
1, 4, 2, 7	2	9	0.00	1.00	0.00	1.00	0.00	0.60	0.03	0.10	
3, 7, 6, 2	3	4	0.00	1.00	0.05	1.00	0.10	0.25	0.00	0.00	
4, 9, 6, 1	9	8	0.00	1.00	0.00	0.03	0.03	0.13	0.00	0.03	
4, 7, 6, 3	3	9	0.00	1.00	0.00	1.00	0.10	0.80	0.03	0.13	
9, 1, 8, 6	1	3	0.05	1.00	0.00	1.00	0.18	0.08	0.10	0.03	
4, 3, 7, 2	7	9	0.00	1.00	0.00	1.00	0.08	0.80	0.00	0.08	
8, 2, 9, 5	2	6	0.00	1.00	0.00	1.00	0.05	0.55	0.18	0.15	
7, 4, 2, 9	2	5	0.00	1.00	0.00	1.00	0.03	0.15	0.15	0.00	
8, 2, 1, 6	8	5	0.03	1.00	0.03	1.00	0.08	0.15	0.00	0.00	
5, 6, 7, 2	2	9	0.00	1.00	0.00	1.00	0.03	0.65	0.18	0.00	
6, 1, 5, 9	6	3	0.05	1.00	0.00	0.20	0.05	0.13	0.00	0.03	
4, 2, 8, 1	1	5	0.00	1.00	0.00	1.00	0.18	0.13	0.10	0.00	
3, 2, 9, 1	1	5	0.08	1.00	0.00	1.00	0.20	0.15	0.05	0.00	
4, 1, 8, 3	4	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	
3, 5, 4, 9	4	7	0.00	1.00	0.00	1.00	0.03	0.58	0.00	0.10	
4, 2, 5, 3	5	1	0.03	0.38	0.08	0.18	0.05	0.25	0.00	0.25	

EOV: environmental and operational variation; SVAR: switching vector autoregressive; VAR: vector autoregressive; SVM: support vector machine; PCA: principal components analysis.