KSCE Journal of Civil Engineering (2018) 22(3):987-992

DOI 10.1007/s12205-018-1290-1

Probabilistic Learning for Decision-making on Civil Infrastructures

pISSN 1226-7988, eISSN 1976-3808
www.springer.com/12205

A Bayesian Learning Method for Structural Damage Assessment of Phase I IASC-ASCE Benchmark Problem

Chang Kook Oh* and James L. Beck**

Received September 15, 2017/Revised 1st: December 8, 2017/Accepted December 15, 2017/Published Online

Abstract

Rapid progress in the field of sensor technology leads to acquisition of massive amounts of measured data from structures being monitored. The data, however, contains inevitable measurement errors which often cause quantitative damage assessment to be ill-conditioned. The Bayesian learning method is well known to provide effective ways to alleviate the ill-conditioning through the prior term for regularization and to provide meaningful probabilistic results for reliable decision-making at the same time. In this study, the Bayesian learning method, based on the Bayesian regression approach using the automatic relevance determination prior, is presented to achieve more effective regularization as well as probabilistic prediction and it is expanded to provide vector outputs for monitoring of a Phase I IASC-ASCE simulated benchmark problem. The proposed method successfully estimates damage locations as well as its severities and give considerable promise for structural damage assessment.

Keywords: *Bayesian learning method, vector outputs, automatic relevance determination prior, damage assessment, Phase I IASC-ASCE benchmark problem*

1. Introduction

Recent advancements in sensor technology such as development of new sensor units, wireless sensors, sensor networks, and so on, enable collection of huge amounts of raw data from structures under monitoring. The collected data contain valuable information on structural damage. Structural Health Monitoring (SHM) similar to human health monitoring intends to perform damage assessment, e.g., identify, locate and estimate damage, within the monitored structures by using the measured data or damage-sensitive features extracted from the measurements (Oh and Sohn, 2009; Oh *et al.*, 2009).

Sophisticated deterministic data processing methods developed in the field of machine learning, for example, Neural Networks, Support Vector Machine (SVM), and so on, have been applied to SHM. Recently-developed SVM, in particular, is widely known as a powerful algorithm for data classification and/or regression by solving a convex optimization problem to successfully produce a global minimum rather than local ones, as well as alleviating over-fitting through model complexity control (Burges, 1998; Vapnik, 1998; Schölkopf and Smola, 2002). Deterministic methods such as SVM, however, can not effectively deal with inherent uncertainties caused by inevitable measurement errors in the data.

Probabilistic learning methods, however, can provide appropriate probabilistic predictions and efficiently handle all of the involved uncertainties that may make inverse problems ill-conditioned. These advantages motivated application of the Bayesian learning method, especially with the incorporation of an Automatic Relevance Determination (ARD) prior, to SHM. The adopted Bayesian method with the ARD prior and the same kernel basis functions as SVM is called the Relevance Vector Machine in the machine learning field and it is able to overcome disadvantages of SVM, for example, (1) incapability to deal with uncertainties, (2) waste of data to decide a parameter to define the objective functions, (3) extensive memory requirements in large-scale tasks and so on (Tipping, 2000 and 2001; Tipping and Faul, 2003).

In this study, the Bayesian learning method based on a regression approach with the ARD prior is investigated for the purpose of SHM. The original Bayesian learning method with the ARD prior is able to provide only a scalar output (Tipping, 2001; Oh *et al*., 2008), but the method adopted in this study is expanded to produce vector outputs such as multivariate damage locations and severities (Thayananthan, 2005; Oh, 2008).

This paper is organized as follows. In Section 2, the mathematical formulation for the proposed method is presented. Monitoring results of the benchmark structure are provided in Section 3 and conclusions are given in Section 4.

*Associate Professor, School of Civil and Environmental Engineering, Kookmin University, Seoul 02707, Korea (Corresponding Author, E-mail: ockoogi@kookmin.ac.kr)
**Professor, Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, U.S.A. (E-mail: jimbeck@caltech.edu)

2. A Bayesian Learning Method with the ARD prior for Vector Outputs

The procedure to apply the proposed Bayesian method to SHM is to (1) extract damage-sensitive features from measured signals, (2) train the algorithm using the prepared training dataset, and (3) predict structural damage from newly-obtained data which were not utilized in the training phase. In this section, the training and predicting procedure of the proposed Bayesian method to produce vector outputs is presented.

2.1 Training Phase

Training data consist of input vectors $\underline{x} \in R^L$ along with their corresponding output vectors $\underline{y} \in R^M$. Input vectors are damage-sensitive features extracted and pre-processed from measured signals and output vectors represent locations and/or severities of damage. The training dataset $\mathcal{D}_N = \{(\underline{x}_i, \underline{y}_i): i = 1, \ldots, N\}$ is employed to learn a regression function $\underline{f}(\underline{x}\,|\,\underline{\theta})$ to relate input and output vectors during the training or learning phase. The regression function chosen in this study is formulated using the Gaussian kernel centered at $\underline{x}_i$, i.e., $k(\underline{x}, \underline{x}_i) = \exp\left(-\frac{\| \underline{x} - \underline{x}_i \|^2}{w}\right)$, which is also frequently selected for SVM:

$$f_m(\underline{x}\,|\,\underline{\theta}) = \theta_{m0} + \sum_{i=1}^{N} \theta_{mi} k(\underline{x}, \underline{x}_i) = \underline{\tau}(\underline{x})^T \underline{\theta}_m \tag{1}$$

where w is called a width of the Gaussian kernel, $\underline{\theta} = \left[\underline{\theta}_1^T, \ldots, \underline{\theta}_M^T\right]^T \in R^{(N+1)M}$, $\underline{\theta}_m = \left[\theta_{m0}, \theta_{m1}, \ldots, \theta_{mN}\right]^T \in R^{N+1}$ for $m = 1, \ldots, M$, and $\underline{\tau}(\underline{x}) = \left[1, k(\underline{x}, \underline{x}_1), \ldots, k(\underline{x}, \underline{x}_N)\right]^T \in R^{N+1}$.

An uncertain prediction error $\underline{\varepsilon} \sim \mathcal{N}(\underline{0}, \Omega)$ where $\Omega = \text{diag}(\sigma_1^2, \ldots, \sigma_M^2)$ to account for the fact that no model gives perfect predictions is included in a probability model (Beck, 2010):

$$\underline{y} = \underline{f}(\underline{x}\,|\,\underline{\theta}) + \underline{\varepsilon} \tag{2}$$

Then, Bayes' Theorem is used to estimate the most probable value of parameter $\underline{\theta}$ given the data:

$$p(\underline{\theta}\,|\,\mathcal{D}_N, \underline{\alpha}, \underline{\sigma}^2) = \frac{p(\mathcal{D}_N\,|\,\underline{\theta}, \underline{\sigma}^2)\, p(\underline{\theta}\,|\,\underline{\alpha})}{p(\mathcal{D}_N\,|\,\underline{\alpha}, \underline{\sigma}^2)} \tag{3}$$

where $\underline{\sigma}^2 = \left[\sigma_1^2, \ldots, \sigma_M^2\right]^T$ and $\underline{\alpha} = \left[\alpha_0, \ldots, \alpha_N\right]^T \in R^{N+1}$ called hyperparameters controls the distribution of the prior for $\underline{\theta}$.

The likelihood for independent predictions of the data $\underline{y}_i$ becomes a product of Gaussians:

$$p(\mathcal{D}_N\,|\,\underline{\theta}, \underline{\sigma}^2) = \prod_{m=1}^{M} \mathcal{N}(\underline{\nu}_m\,|\,\mathbf{\Phi}\underline{\theta}_m, \sigma_m^2 \mathbf{I}) \tag{4}$$

where $\underline{\nu}_m = [(\underline{y}_1)_m, \ldots, (\underline{y}_N)_m]^T$ and $\mathbf{\Phi} = \left[\underline{\tau}(\underline{x}_1), \ldots, \underline{\tau}(\underline{x}_N)\right]^T \in R^{N\times(N+1)}$. The prior distribution over $\underline{\theta}$ controlled by hyperparameter $\underline{\alpha}$ is:

$$p(\underline{\theta}\,|\,\underline{\alpha}) = \prod_{m=1}^{M} \mathcal{N}(\underline{\theta}_m\,|\,\underline{0}, \mathbf{A}^{-1}) \tag{5}$$

where matrix $\mathbf{A} = \text{diag}(\alpha_0, \ldots, \alpha_N)$. The resulting posterior PDF using Bayes' Theorem in Eq. (3) is:

$$p(\underline{\theta}\,|\,\mathcal{D}_N, \underline{\alpha}, \underline{\sigma}^2) \propto \prod_{m=1}^{M} \mathcal{N}(\underline{\theta}_m\,|\,\hat{\underline{\theta}}_m, \Sigma_m) \tag{6}$$

where $\hat{\underline{\theta}}_m = \sigma_m^{-2} \Sigma_m \mathbf{\Phi}^T \underline{\nu}_m$ and $\Sigma_m = \left(\sigma_m^{-2} \mathbf{\Phi}^T \mathbf{\Phi} + \mathbf{A}\right)^{-1}$ are the most probable values a posteriori of $\underline{\theta}_m$ and its covariance matrix, respectively.

The most probable values of hyperparameters $\hat{\underline{\alpha}}$ and variances $\hat{\underline{\sigma}}^2$ are estimated using Bayesian model class selection based on the evidence and training dataset $\mathcal{D}_N$. The selected most probable model class maximizes the log evidence $\mathcal{L}(\underline{\alpha}, \underline{\sigma}^2)$ for a uniform prior PDF (Beck and Yuen, 2004; Tipping and Faul, 2003):

$$\begin{aligned}
\mathcal{L}(\underline{\alpha}, \underline{\sigma}^2) &= \ln p(\mathcal{D}_N\,|\,\underline{\alpha}, \underline{\sigma}^2) \\
&= -\frac{1}{2} \sum_{m=1}^{M} \left[N \ln 2\pi + \ln |\mathbf{C}_m| + \underline{\nu}_m^T \mathbf{C}_m^{-1} \underline{\nu}_m\right] \\
&= \mathcal{L}(\underline{\alpha}_{-i}, \underline{\sigma}^2) + \sum_{m=1}^{M} \left[\ln \alpha_i - \ln(\alpha_i + S_{mi}) + \frac{Q_{mi}^2}{\alpha_i + S_{mi}}\right] \\
&= \mathcal{L}(\underline{\alpha}_{-i}, \underline{\sigma}^2) + l(\alpha_i, \underline{\sigma}^2)
\end{aligned} \tag{7}$$

where $\mathcal{L}(\underline{\alpha}_{-i}, \underline{\sigma}^2)$ is the evidence excluding terms related with $\underline{\tau}(\underline{x}_i)$, $\mathbf{C}_m = \sigma_m^2 \mathbf{I} + \mathbf{\Phi} \mathbf{A}^{-1} \mathbf{\Phi}^T$, $S_{mi} = \frac{\alpha_i s_{mi}}{\alpha_i - s_{mi}}$, $Q_{mi} = \frac{\alpha_i q_{mi}}{\alpha_i - s_{mi}}$, $s_{mi} = \sigma_m^{-2} \underline{\tau}(\underline{x}_i)^T \underline{\tau}(\underline{x}_i) - \sigma_m^{-4} \underline{\tau}(\underline{x}_i)^T \mathbf{\Phi} \Sigma_m \mathbf{\Phi}^T \underline{\tau}(\underline{x}_i)$ and $q_{mi} = \sigma_m^{-2} \underline{\tau}(\underline{x}_i)^T \underline{\nu}_m - \sigma_m^{-4} \underline{\tau}(\underline{x}_i)^T \mathbf{\Phi} \Sigma_m \mathbf{\Phi}^T \underline{\nu}_m$.

Therefore, maximization of the log evidence $\mathcal{L}(\underline{\alpha}, \underline{\sigma}^2)$ can be achieved by iterative maximization of $l(\underline{\alpha}_i, \underline{\sigma}^2)$ with respect to α_i and σ_m^2 (Tipping, 2001). In practice, most of the α_is become infinite, which makes the corresponding θ_{mi} zero and so prunes the associated kernel terms from Eq. (1) (Tipping, 2001; Thayananthan, 2005).

2.2 Prediction Phase

Prediction of output vector $\tilde{\underline{y}}$ for a new input vector $\tilde{\underline{x}}$ is given by the Theorem of Total Probability from the robust posterior probability based on the most probable model class determined in the training phase:

$$p(\tilde{\underline{y}}\,|\,\tilde{\underline{x}}, \mathcal{D}_N, \hat{\underline{\alpha}}, \hat{\underline{\sigma}}^2) = \mathcal{N}(\tilde{\underline{y}}\,|\,\underline{y}_*, \mathbf{\Omega}_*) \tag{8}$$

where $\underline{y}_* = [y_{1*}, \ldots, y_{M*}] \in \mathrm{R}^M$, $y_{m*} = \hat{\underline{\theta}}_m^T \underline{\tau}(\tilde{\underline{x}})$, $\mathbf{\Omega}_* = \text{diag}(\sigma_{1*}^2, \ldots, \sigma_{M*}^2) \in \mathrm{R}^{M\times M}$, $\sigma_{m*}^2 = \hat{\sigma}_m^2 + \underline{\tau}(\tilde{\underline{x}})^T \hat{\Sigma}_m \underline{\tau}(\tilde{\underline{x}})$ and $\hat{\Sigma}_m = \left(\hat{\sigma}_m^{-2} \mathbf{\Phi}^T \mathbf{\Phi} + \mathbf{A}\right)^{-1}$.

3. IASC-ASCE Structural Health Monitoring Benchmarks for Damage Assessment

3.1 IASC-ASCE Benchmark Structure and Identification of Modal Parameters

The primary purpose of the IASC-ASCE benchmarks is to offer a common platform for numerous researchers to apply various SHM methods to an identical structure and to compare their performance. The benchmarks comprise two phases, e.g., Phase I and Phase II, each with simulated and experimental benchmarks. In this study, Phase I simulated data generated from an existing 4-story, 2-bay × 2-bay, steel braced frame as shown in

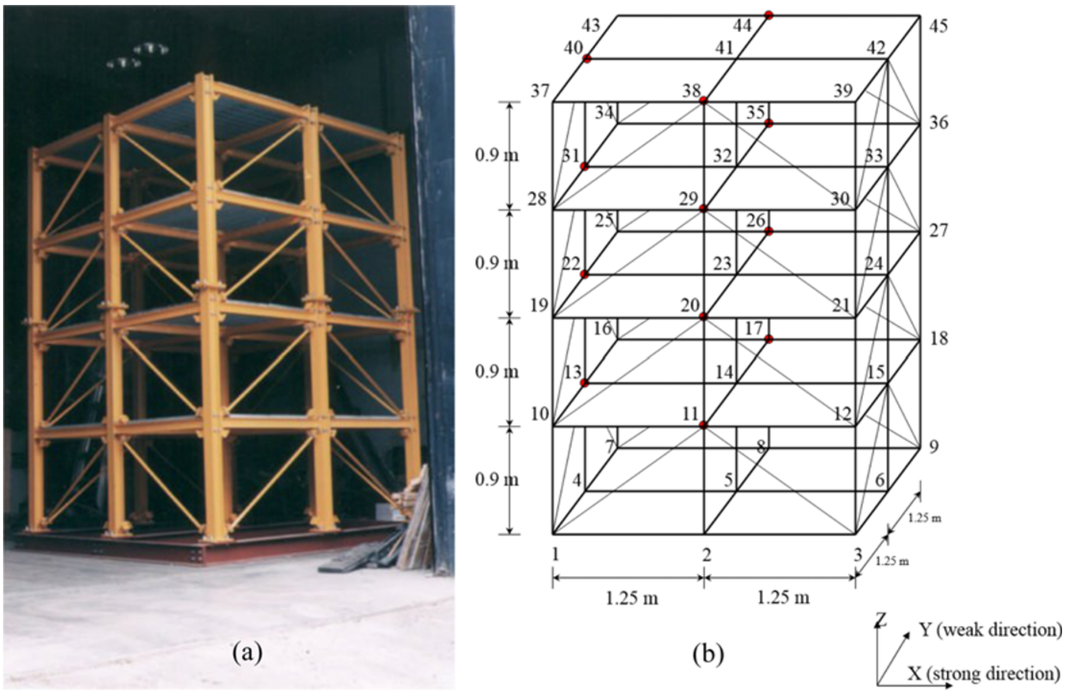

Fig. 1. (a) Steel Frame Scaled Model Structure used for Benchmark, (b) Its Diagram

Fig. 1(a) are employed for prediction of structural damage (Johnson *et al*., 2000; Johnson *et al*., 2004).

Prediction datasets are pre-processed from the acceleration time histories simulated in the strong (X) and weak (Y) directions and measured at the locations marked with red dots in Fig. 1(b). Two kinds of Finite Element (FE) models of 12 and 120 degrees of freedom (DOF) are employed to generate the time histories for pre-defined damage cases and damage patterns listed in Table 1 (Johnson *et al.*, 2004). Mode shapes with associated frequencies are extracted by using MODE-ID, i.e., modal identification program (Beck, 1996) and prepared for prediction.

Training datasets, however, are generated from simpler models than the benchmarks for all damage cases, for example, lumped-mass shear building models with reduced DOFs such as 4-DOF and 12-DOF for damage cases 1-3 and 4-5, respectively, in order to reflect modeling errors. The proposed Bayesian regression approach is tested on the simpler damage cases 1-3 first and then extended to more realistic damage cases 4-5.

Table 1. Damage Cases and Patterns for Prediction using the Benchmark Model

Description	Case				
	1	2	3	4	5
12 DOF Model	O		O	O	
120 DOF Model		O			O
Symmetric Mass	O	O	O		
Asymmetric Mass				O	O
Ambient Excitation	O	O			
Shaker on Roof			O	O	O
Damage Patterns: Remove Followings					
1) All Braces in the 1st Story	O	O	O	O	O
2) All Braces in the 1st & 3rd Stories	O	O	O	O	O
3) One Brace in the 1st Story				O	O
4) One Brace in the 1st & 3rd Stories				O	O
5) 4) & Loosen. Floor Beam at the 1st Level					O
6) 2/3 Stiffness in One Brace at the 1st Story					O

3.2 Damage Cases 1-3

3.2.1 Training Phase

Training datasets for damage cases 1-3 are simulated from a simplified (compared with the benchmark) 4-story 4-DOF lumped mass shear building model. Note that data for prediction are generated from 12-DOF and 120-DOF FE models for damage cases 1, 3 and 2, respectively. A lumped mass matrix with entities of $\mathbf{M} = \mathrm{diag}\{3242, 2652, 2652, 1809\}$ kg is utilized (Yuen *et al.*, 2004) and damage is imposed as a reduction of inter-story stiffness scaled by parameter θ_i of undamaged stiffness $k_i^u = 68.1$ MN/m:

$$k_i^{pd} = \theta_i k_i^u$$

where k_i^{pd} is the possibly damaged i^{th} story stiffness.

Changes of the 1st and 2nd mode shape vectors and corresponding modal frequencies are selected for the input vector $\underline{x}$, because modal parameter changes are known to be more insensitive to modeling error than the parameter values (Lam *et al.*, 2006; Oh, 2008) :

$$\underline{x}_i = [\Delta\underline{\phi}_1 \;\; \Delta f_1 \;\; \Delta\underline{\phi}_2 \;\; \Delta f_2]_i \text{ with the component of } \underline{\phi}_j \text{ at roof} = 1 \tag{9}$$

where $\Delta\underline{\phi}_j$ and Δf_j are changes of the j^{th} mode shape and modal

Table 2. Stiffness Loss Predictions for Damage Cases 1-3

	Damage Case	Damage Pattern	Story 1	Story 2	Story 3	Story 4
Target	1, 2, 3	1	**0.71**	0.00	0.00	0.00
		2	**0.71**	0.00	**0.71**	0.00
Prediction	1	1	**0.74**	0.00	0.03	0.04
		2	**0.71**	0.00	**0.69**	0.05
	2	1	**0.75**	0.00	0.01	0.00
		2	**0.72**	0.02	**0.76**	0.01
	3	1	**0.73**	0.04	0.03	0.05
		2	**0.74**	0.03	**0.71**	0.03

frequency between undamaged and possibly damaged structures, respectively.

Along with the input vector $\underline{x}$, output vector $\underline{y} \in R^4$ represents 9 different levels of stiffness losses, i.e., 0%, 10%, 20%, …, 80% losses, to the undamaged inter-story stiffness for each story. For example, $\underline{y} = [0.20 \;\; 0 \;\; 0.40 \;\; 0]^T$ means 20% and 40% stiffness losses in the 1st and 3rd stories, respectively.

Training dataset $\mathcal{D}_N$ comprising 47 feature vectors is generated via simple eigen-analysis of the mass and stiffness matrices based on the restriction that simultaneous damage is assumed to occur at two different locations at most.

3.2.2 Prediction Phase

Input vectors for prediction, i.e., $\underline{\tilde{x}}$ of the same form as Eq. (9), are simulated from damaged structures with 71% stiffness reduction at the 1st and 1st,3rd stories for damage patterns 1 and 2, respectively, as shown in Table 2. Note that prediction is carried out with respect to $\underline{\tilde{x}}$ whose damage severities are different from those of the training data. Time histories with the addition of 10% noise to signal are generated from the benchmark structure with a damping coefficient of 0.01 and a time step of 0.004 sec. Stationary response for 20 sec after the initial 10 sec transient response was employed to extract modal parameters using MODE-ID (Beck, 1996). As listed in Table 2, prediction results $\underline{\tilde{y}}$ in terms of presence, location and severity of damage agree with target values, i.e., 71% stiffness losses at the 1st and 1st, 3rd stories, with considerable accuracy.

3.3 Damage Cases 4 and 5

For damage cases 4-5, the number of possible damage scenarios to take into consideration increases enormously, since damage can be imposed within faces of the structure rather than just in stories as damage cases 1-3. For example, if 4 separate damage levels of 20%, 40%, 60%, and 80% are imposed simultaneously at 4 different locations among the entire 16 possible positions, then the total number of possible damage scenarios becomes $\sum_{i=0}^{4} {}_{16}C_i 4^i = 503{,}745$ resulting in $\mathbf{\Phi} \in R^{503745 \times 503745}$, where C represents combinatorial factor.

To overcome those aforementioned computational difficulties, a two-step approach is carried out (Yuen and Lam, 2006). In the first step, only damage locations are identified, and then the severities of damage are estimated in the next step. In this study, a threshold value of damage index to decide the presence of damage is set to be 0.6, which is ad-hoc to a certain extent. However, the estimated damage severities in the second step could rectify false results in the previous step, if any.

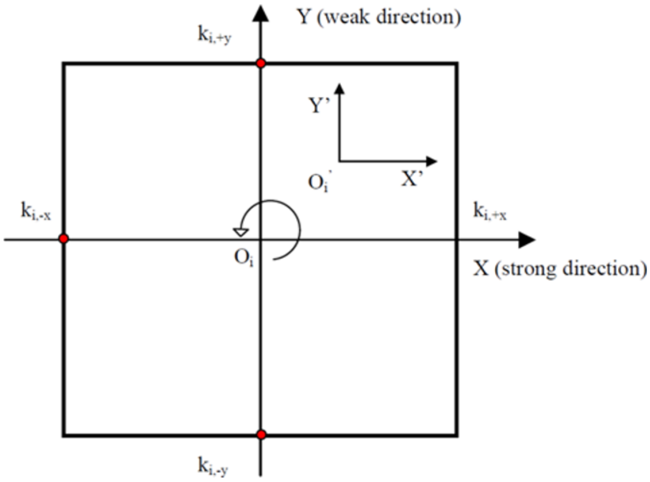

Fig. 2. Floor Plan for Benchmark Structure

3.3.1 Training Phase

Training data of damage cases 4 and 5 are simulated from a 3-dimensional, 12-DOF, lumped mass shear building model to locate damage in the faces of the building. Note that as for damage cases 1-3, a simplified analytical model is utilized to consider modeling errors. Damage is imposed using stiffness loss parameters θ_{ij} for the i^{th} story ($i = 1,...,4$) and face j ($j = +x,-x,+y,-y$) (see Fig. 2):

$$k_{ij}^{pd} = \theta_{ij} k_{ij}^{u} \tag{10}$$

where k_{ij}^{u} are stiffness of the undamaged shear building model, e.g., $k_{i,+x}^{u} = k_{i,-x}^{u} = 34.0MN/m$ and $k_{i,+y}^{u} = k_{i,-y}^{u} = 53.5MN/m$.

The local stiffness matrices can be computed and transformed with respect to geometric center (Yuen *et al.*, 2004) and assembled to generate mode shape vectors and corresponding frequencies at locations marked with red dots in $-x, +y, -y$ faces in Fig. 2. Extracted modal parameters are pre-processed to construct input vectors, e.g., a damage signature defined as $\underline{x}_i = [\Delta\underline{\phi}_1/\Delta f_1 \;\; \Delta\underline{\phi}_2/\Delta f_2]_i$ for the first step and damage-sensitive feature vectors in Eq. (9) for the second step. Training output vectors for the first and second steps are damage index $\in R^8$ to locate damage, whose components have values of either 0 or 1 for undamaged or damaged states, respectively, and damage severities expressed as stiffness reduction factor θ_{ij} in Eq. (10).

To construct the damage signatures for the purpose of damage identification (Step 1), only a single damage level, i.e., 50% stiffness reduction, is imposed on the stiffness in the strong and weak directions. Note that the stiffness of both elements facing each other, e.g., elements in +Y and –Y faces, are reduced together when the damage is assigned. To estimate damage severity (Step 2), two levels of damage, i.e., 30% and 70% of stiffness losses, are imposed on the structural elements identified

Table 3. Damage Indices to Identify Damage Locations for Damage Cases 4-5

Damage Case	Damage Pattern	Story 1		Story 2		Story 3		Story 4	
		$\theta_{i,y}$	$\theta_{i,x}$	$\theta_{i,y}$	$\theta_{i,x}$	$\theta_{i,y}$	$\theta_{i,x}$	$\theta_{i,y}$	$\theta_{i,x}$
Target - Damage Cases 4 & 5	1	**1.00**	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00
	2	**1.00**	**1.00**	0.00	0.00	**1.00**	**1.00**	0.00	0.00
	3	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00
	4	0.00	**1.00**	0.00	0.00	**1.00**	0.00	0.00	0.00
	5	0.00	**1.00**	0.00	0.00	**1.00**	0.00	0.00	0.00
	6	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00
Prediction - Damage Case 4	1	**1.48**	**2.05**	0.19	0.49	0.23	0.44	0.20	0.34
	2	**1.16**	**1.53**	-0.12	0.43	**0.76**	**0.85**	0.29	0.32
	3	-1.94	**0.95**	0.42	0.37	**1.93**	0.39	**2.59**	0.07
	4	0.06	**1.24**	0.57	**0.95**	**1.23**	0.59	0.35	0.28
Prediction - Damage Case 5	1	**1.82**	**2.26**	0.19	0.43	0.11	0.44	-0.02	0.32
	2	**0.97**	**1.42**	0.42	0.33	**0.79**	**0.90**	0.07	0.26
	3	**3.60**	**0.90**	-1.29	0.43	**1.06**	0.48	-2.59	2.62
	4	-0.22	**1.27**	0.19	0.56	**1.63**	0.57	0.08	0.35
	5	-0.23	**1.30**	0.21	0.56	**1.63**	0.56	0.08	0.36
	6	**5.10**	**0.79**	**3.09**	-0.19	**1.38**	**0.71**	-7.01	-0.85

Table 4. Stiffness Reduction Factors to Estimate Damage Severities for Damage Cases 4-5

Damage Case	Damage Pattern	Story 1				Story 3			
		$\theta_{i,-y}$	$\theta_{i,+x}$	$\theta_{i,+y}$	$\theta_{i,-x}$	$\theta_{i,-y}$	$\theta_{i,+x}$	$\theta_{i,+y}$	$\theta_{i,-x}$
Target - Damage Cases 4 & 5	1	**0.45**	**0.71**	**0.45**	**0.71**	0.00	0.00	0.00	0.00
	2	**0.45**	**0.71**	**0.45**	**0.71**	**0.45**	**0.71**	**0.45**	**0.71**
	3	0.00	**0.36**	0.00	0.00	0.00	0.00	0.00	0.00
	4	0.00	**0.36**	0.00	0.00	**0.23**	0.00	0.00	0.00
	5	0.00	**0.36**	0.00	0.00	**0.23**	0.00	0.00	0.00
	6	0.00	**0.23**	0.00	0.00	0.00	0.00	0.00	0.00
Prediction - Damage Case 4	1	**0.43**	**0.70**	**0.44**	**0.72**	0.00	0.00	0.00	0.00
	2	**0.42**	**0.69**	**0.43**	**0.71**	**0.46**	**0.72**	**0.42**	**0.69**
	3	0.00	**0.37**	0.00	0.00	0.03	0.00	0.02	0.00
	4	0.00	**0.30**	0.00	0.02	**0.18**	0.00	0.03	0.00
Prediction - Damage Case 5	1	**0.44**	**0.71**	**0.43**	**0.73**	0.00	0.00	0.00	0.00
	2	**0.43**	**0.71**	**0.42**	**0.70**	**0.46**	**0.70**	**0.43**	**0.73**
	3	0.03	**0.41**	0.02	0.07	0.00	0.00	0.02	0.00
	4	0.00	**0.36**	0.00	0.00	**0.22**	0.00	0.02	0.00
	5	0.00	**0.37**	0.00	0.00	**0.22**	0.00	0.07	0.00
	6	0.00	**0.21**	0.01	0.03	0.04	0.00	0.02	0.00

*Maximum value of the estimated stiffness reduction factor for story 2 and 4 (not listed in Table 4) is 0.09. Note that target value for the corresponding location is 0.00.

in the first step in order to generate the training dataset.

3.3.2 Prediction Phase – Step 1: Identification of Damage Locations

Full models of the benchmarks with 12 and 120 DOF for damage cases 4 and 5, respectively, are used to generate acceleration time histories for prediction. The added noise level, damping coefficient, and time step size are identical with those used for damage cases 1-3, that is 10%, 0.01, and 0.004 sec, respectively. MODE-ID (Beck, 1996) is utilized to extract modal parameters as before.

Step 1 prediction results listed in Table 3 demonstrate that all of the actually damaged elements can be successfully identified by the proposed method. Some of the damage indices that are larger than the previously set threshold of 0.6, even though no damage is imposed, are corrected for in the next step. Note that $\theta_{i,y} = \theta_{i,+y} = \theta_{i,-y}$ and $\theta_{i,x} = \theta_{i,+x} = \theta_{i,-x}$ only in the first step (damage assigned in each face is estimated in the second step as shown in Table 4). Underlined and bold values in Table 3 stand for suspected damage locations based on the predicted damage index values.

3.3.3 Prediction Phase – Step 2: Assessment of Damage Severities

After identifying the potential damage locations, another Bayesian learning algorithm is trained to predict damage severities at suspected damage locations. As before, input vectors in Eq. (9) are pre-processed from acceleration time histories of the benchmarks with the assigned levels of damage given by the target values in Table 4.

Prediction results listed in Table 4 demonstrate that the false identification of damage in the previous step is corrected in the present step and the severities of damaged elements in each face of the 1st and 3rd stories are estimated with a high degree of accuracy as well. The predicted maximum damage severity of the 2nd and 4th stories (not listed in Table 4) is 0.09 (Oh, 2008), which could be reasonably judged to be undamaged. Note that the effect of loosening the floor beam at the 1st level in damage pattern 4 and 5 of damage case 5 is negligible, as addressed by Yuen *et al.* (2004).

Based on the results so far achieved, it can be concluded that the proposed two step Bayesian learning method is a promising tool for SHM to successfully locate damage as well as to estimate its severities.

4. Conclusions

In this paper, a Bayesian learning method with the ARD prior and kernel basis functions is introduced and expanded to provide vector outputs in order to investigate its potential capability for damage assessment in structural health monitoring. The proposed Bayesian method can (1) explicitly handle all of the involved uncertainties from measurement errors, (2) provide meaningful probabilistic decision-making based on the associated confidence level, (3) control the trade-off between the data fit and model complexity in an automatic manner, and (4) prune irrelevant kernels enabling rapid prediction when new data are obtained.

The proposed method is applied to the Phase I IASC-ASCE benchmark problem to investigate its feasibility for damage assessment. The procedure for SHM can be summarized as (1) training data are simulated from simplified structures that can reflect modeling errors, e.g., lumped mass models with reduced DOF, and pre-processed to construct feature vectors as in Eq. (9) or damage signatures, (2) the proposed algorithm is trained using the prepared training data, (3) data for prediction are obtained from the monitored structure, and (4) damage locations and its severities are estimated using the prediction data.

Based on the prediction results obtained in this study and listed in Table 2 ~ Table 4, the proposed Bayesian learning method with vector outputs demonstrates its ability to accurately identify, locate, and estimate damage for the benchmark problems and it can also be a promising and powerful real time tool for structural health monitoring using huge amounts of collected data with intrinsic errors.

References

Beck, J. L. (1996). "System identification methods applied to measured seismic response." *Proceedings of the Eleventh World Conference on Earthquake Engineering*, Acapulco, Mexico.

Beck, J. L. and Yuen, K. V. (2004). "Model selection using response measurements: A bayesian probabilistic approach." *Journal of Engineering Mechanics*, Vol. 130, pp. 192-203, DOI: 10.1061/(ASCE)0733-9399(2004)130:2(192)..

Beck, J. L. (2010). "Bayesian System Identification based on Probability Logic." *Structural Control and Health Monitoring*, Vol. 17, pp. 825-847, DOI: 10.1002/stc.424.

Burges, C. J. C. (1998). "A tutorial on support vector machines for pattern recognition." *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, pp. 121-167, DOI: 10.1023/A:1009715923555.

Ching, J. and Beck, J. L. (2004). "Bayesian analysis of the Phase II IASC-ASCE structural health monitoring experimental benchmark data." *Journal of Engineering Mechanics*, Vol. 130, No. 10, pp. 1233-1244, DOI: 10.1061/(ASCE)0733-9399(2004)130:10(1233).

Johnson, E. A., Lam, H. F., Katafygiotis, L., and Beck, J. L. (2000). "A benchmark problem for structural health monitoring and damage detection." *Proceedings of the 14th ASCE Engineering Mechanics Conference*, Austin, Texas.

Johnson, E. A., Lam, H. F., Katafygiotis, L., and Beck, J. L. (2004). "Phase I IASC - ASCE Structural health monitoring benchmark problem using simulated data." *Journal of Engineering Mechanics*, Vol. 130, No. 3, pp. 3-15, DOI: 10.1061/(ASCE)0733-9399(2004)130:1(3).

Lam, H. F., Yuen, K. V., and Beck, J. L. (2006). "Structural health monitoring via measured ritz vectors utilizing artificial neural networks." *Computer-Aided Civil and Infrastructure Engineering*, Vol. 21, No. 4, pp. 232-241, DOI: 10.1111/j.1467-8667.2006.00431.x.

Oh, C. K. (2008). *Bayesian learning for earthquake engineering applications and structural health monitoring*, Ph. D. Thesis, California Institute of Technology, Pasadena, U.S.A.

Oh, C. K. and Sohn, H. (2009). "Damage diagnosis under environmental and operational variations using unsupervised support vector machine." *Journal of Sound and Vibration*, Vol. 325, pp. 224-239, DOI: 10.1061/(ASCE)0733-9399(2008)134:12(1013).

Oh, C. K., Beck, J. L., and Yamada, M. (2008). "Bayesian learning using automatic relevance determination prior with an application to earthquake early warning." *Journal of Engineering Mechanics*, Vol. 134, No. 12, pp. 1013-1020, DOI: 10.1061/(ASCE)0733-9399(2008)134:12(1013).

Oh, C. K., Sohn, H., and Bae, I. (2009). "Statistical novelty detection within the Yeongjong suspension bridge under environmental and operational variations." *Smart Materials and Structures*, Vol. 18, pp. 1-9, DOI: 10.1088/0964-1726/18/12/125022.

Schölkopf, B. and Smola, A. J. (2002). *Probability, Reliability and Statistical Methods in Engineering Design*., Wiley, New York.

Thayananthan, A. (2005). *Template-based Pose Estimation and Tracking of 3D Hand Motion.* Ph.D. Thesis, University of Cambridge.

Tipping, M. E. (2000). "The Relevance Vector Machine." *Advances in Neural Information Processing Systems 12*, pp. 652-658.

Tipping, M. E. (2001). "Sparse bayesian learning and the relevance vector machine." *Journal of Machine Learning Research*, Vol. 1, pp. 211-244, DOI: 10.1162/15324430152748236.

Tipping, M. E., and Faul, A. C. (2003). "Fast marginal likelihood maximisation for sparse bayesian models." *Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics*, Key West, FL.

Vapnik, V. N. (1998). *Statistical Learning Theory.* Wiley, New York.

Yuen, K. V. and Lam, H. F. (2006). "On the complexity of artificial neural networks for smart structure monitoring." *Engineering Structures*, Vol. 27, No. 7, pp. 977-984, DOI: 10.1016/j.engstruct.2005.11.002.

Yuen, K. V., Au, S. K., and Beck, J. L. (2004). "Two-stage structural health monitoring approach for Phase I Benchmark studies." *Journal of Engineering Mechanics*, Vol. 130, pp. 16-33, DOI: 10.1061/(ASCE)0733-9399(2004)130:1(16).