On leveraging network-wide information from hotspot sensor networks using Multi-output Gaussian Process regression models

Ahmad Amer
PhD Candidate
Center for Mobility with Vertical Lift (MOVE)
Rensselaer Polytechnic Institute, Troy, NY, USA

Fotis Kopsaftopoulos
Assistant Professor
Center for Mobility with Vertical Lift (MOVE)
Rensselaer Polytechnic Institute, Troy, NY, USA

ABSTRACT

With the needs for full structural state awareness and health monitoring as well as emerging challenges of Urban Air Mobility (UAV) and Future Vertical Lift (FVL), Health and Usage Monitoring systems (HUMS) need to be more accurate, robust and reliable than ever before. In active-sensing guided-wave networks in particular, conventional Damage Index (DI)-based approaches have been the industry standard for decades because of their computational simplicity and ability to do the damage detection and quantification tasks. However, under specific circumstances, like for specific actuator-sensor paths within a network or due to varying operational conditions, DIs can suffer from various drawbacks that make them prone to inaccurate and/or ineffective damage quantification. This study builds on previous work by the authors where DIs were used to train single-output Gaussian Process regression models (SOGPRMs) for robust damage quantification, and the accuracy limit of SOGPRMs was shown to depend on the evolution of the chosen DI formulation with damage size. In this study, multi-output GPRMs (MOGPRMs) are used instead in order to leverage information about damage size from multiple actuator-sensor path DI values. It is shown that the proposed approach can overcome the different shortcomings of DI evolution with damage size in the different path by capturing the correlation between the DI evolution for different paths. The proposed framework is applied for an Al coupon with simulated damage, and the damage size quantification results are compared with those of SOGPRMs. It is shown that the information fusion approach exhibited by MOGPRMs gives more accurate damage size estimations compared to SOGPRMs.

INTRODUCTION

Novel Urban Air Mobility (UAM) and Future Vertical Lift (FVL) aircraft configurations are shaping the future technologies through boasting research in many fields. One of the most relevant of such fields to the Health and Usage Monitoring Systems (HUMS) community is structural/state awareness (Refs. 1–4), where information provided by sensors on the aircraft can be used to detect and identify varying states and/or damage events. In this context, the effectiveness of HUMS today depends heavily on the availability of information-rich sensor data from multiple sensors across an aircraft. In particular, sensor networks designed to monitor structural "hot spots" can provide critical information on damage development at an early stage (Refs. 5–7) – information that is conventionally leveraged in the industry today by using health/damage indicators (herein denoted as DIs) (Refs. 8–10). One of the major challenges facing such endeavors is the variation in the "amount" of damage-relevant information carried by the different signal paths within a given sensor network for the purposes of damage detection and quantification. One cause of such a phenomenon is the signal paths intersecting or not intersecting damage (Ref. 11), which in turn might lead to the DIs of the latter paths evolving in a way uncorrelated to damage (Ref. 12), making those paths not very useful in damage quantification (Refs. 13, 14). Another reason for the said variation is the phenomenon of saturation (Ref. 15), which might occur depending on the relative location of a signal path from damage. This phenomenon may also limit the amount of information a given signal path provides beyond a specific damage size. These causes, along with the oftentimes non-linear and complex structural response (Refs. 16–18) cause such variations in the information that can be extracted from each of the different signal paths within a sensor network.

When it comes to piezoelectric sensor networks, using multi-path data sets for damage detection and localization is a standard procedure (Refs. 19–24), where feature extraction metrics, oftentimes probabilistic, are used for detecting and accurately localizing damage. However, the problem of accurate damage quantification has been mostly tackled throughout the literature within the framework of single actuator-sensor paths (Refs. 17, 25–28), where the signal path with the largest change in DI values, for instance, would be used for analysis. That being said, there have been limited efforts on combining information from different paths and/or SHM

Presented at the VFS International 77th Annual Forum & Technology Display, May 10–14, 2021.

techniques towards the aim of damage detection and/or quantification (Ref. 29). For instance, observing uncertainties in each individual path DI (which originate from varying environmental/operational conditions), Jin *et al.* (Ref. 12) proposed an arithmetic fusion algorithm where DIs in the time and the frequency domains, based on energy and amplitude, are all summed over all the actuator-sensor path signals coming from a steel plate in order to "visualize" fatigue crack growth. Also, Derriso *et al.* (Ref. 30) examined combining information form guided-wave and acoustic emission signals for detecting fatigue cracks in Aluminum dogbone samples. Overall, the techniques fusing information from, or estimating the correlation between different actuator-sensor path signals in he literature are either not probabilistic (making them prone to uncertainties), or do not tackle accurate damage quantification.

Indeed, even when it comes to damage detection and quantification using single actuator-sensor path signals in active-sensing guided-wave SHM, the most promising of such endeavors are those involving the use of advanced modeling techniques for damage detection and quantification (Refs. 19, 31–35), which are used to overcome the drawbacks of conventional DIs. In this context, in recent work, the authors proposed frameworks based on machine learning techniques for damage detection and quantification, in which single-output Gaussian Process Regression & Classification Models (GPR/CMs) were used (Refs. 14,36), which included the prior selection of damage-intersecting paths based on statistical path selection algorithms (Refs. 11, 14). In addition, the idea of leveraging information from different signal paths for damage quantification was also tackled by the authors (Ref. 13), in which multiple damage-intersecting, or mixed paths were used to train single-output GPRMs for more accurate damage quantification models. However, in the latter study, there were no prior assumptions on the possible correlation between the different signal paths; that is, the correlation of the information carried by multiple signal paths was not considered.

The **aim** of the present study is the application of multi-output GPRMs (MOGPRMs) trained using data from multiple signal paths are used in order to leverage the many signals coming from an active-sensing, guided-wave sensor network monitoring any given "hot spot", whilst capturing the correlation between the different actuator-sensor paths. The proposed framework does not only promise more accurate damage detection and quantification processes compared to models that use single-path signals, but also promises a more efficient learning process by capturing the hidden correlation between the different signals. Trained using readily-available state-of-the-art DI results, MOGPRMs are applied to experimental results from an aluminum coupon in the context of damage detection and quantification, and then compared with single-output GPRM (SOGPRM) results in order to highlight the effect of capturing the said correlation between signals on the damage state and size prediction accuracy of GPRMs.

BACKGROUND

Reference Damage Index

The state-of-the-art DI formulation used in this study for training GPRMs was adopted from the work of Janapati *et al.* (Ref. 17) owing to the exhibited high sensitivity to damage features compared to sensitivity to material, sensor, and adhesive properties. Given a healthy signal ($x_o[t]$) and a signal coming from an unknown state ($x_u[t]$), the selected DI can be formulated as follows:

$$N_u = \frac{x_u[t]}{\sqrt{\sum_{t=1}^{n} x_u^2[t]}}, \tag{1a}$$

$$N_0 = \frac{\sum_{t=1}^{n} (x_o[t] \cdot N_u)}{x_o[t] \cdot \sum_{t=1}^{n} x_o^2[t]} \tag{1b}$$

$$DI = \sum_{t=1}^{n} (N_u - N_o) \tag{1c}$$

where N_u and N_o are normalized unknown- and baseline-state signals, respectively.

Single-output Gaussian Process Regression Models (SOGPRMs)

GPRMs are kernel-based linear regression formulations that have the ability to model non-linear relationships between observations and inputs. In this study, GPRMs are used to model the evolution of the Z statistic with frequency and/or damage size as will be shown later. Given a training data set $\mathcal{D}$ containing n inputs-observation pairs $\{(\mathbf{x}_i \in \mathbb{R}^D, y_i \in \mathbb{R},\ i = 1,2,3,\ldots,n\}$, a GPR model can be formulated as follows:

$$y = f(\mathbf{x}) + \epsilon \tag{2}$$

such that a GP prior with mean $m(\mathbf{x})$ and kernel $k(\mathbf{x},\mathbf{x}')$ is placed on the latent function $f(\mathbf{x})$, and an independent, identically-distributed (*iid*), zero-mean Gaussian prior with variance σ_n^2 is placed on the noise term ϵ as follows:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x},\mathbf{x}')), \quad \epsilon \sim iid\, \mathcal{N}(0,\sigma_n^2) \tag{3}$$

In this study, $m(\mathbf{x})$ is set to zero, and the squared exponential kernel function is used for f:

$$k(\mathbf{x},\mathbf{x}') = \sigma_0^2 \exp(-\frac{1}{2}(\mathbf{x}-\mathbf{x}')^T \Lambda^{-1}(\mathbf{x}-\mathbf{x}')) \tag{4}$$

where σ_0^2 is the output signal power (variance), and Λ is a diagonal matrix of the characteristic length scales of each dimension (D, i.e each covariate) in the input data set. There will be a separate length scale for every covariate in the data. Thus, for a single-input (i.e. $D = 1$), the entries along the diagonal of Λ^{-1} will be identical.

Training GPRM training involves optimizing the hyperparameters ($\theta \equiv \sigma_0^2, \Lambda, \sigma_n^2$) *via* Type II Maximum Likelihood (Ref. 37, Chapter 5, pp. 109). In this method, the marginal likelihood (evidence) of the training observations (outputs) is maximized with respect to the hyperparameters. For computational reasons, its negative log is minimized instead as follows:

$$\widehat{\theta} = \arg\min_{\theta}\{-\log p(\mathbf{y}|X,\theta)\} \quad (5a)$$

$$-\log p(\mathbf{y}|X,\theta) = -\log \mathcal{N}(\mathbf{y}|\mathbf{0}, K_{XX}+\sigma_n^2\mathbb{I}) \quad (5b)$$

$$= -\frac{1}{2}\mathbf{y}^T(K_{XX}+\sigma_n^2\mathbb{I})^{-1}\mathbf{y} - \frac{1}{2}\log|K_{XX}+\sigma_n^2\mathbb{I}| - \frac{n}{2}\log 2\pi \quad (5c)$$

Prediction One of the most powerful results of the assumptions in GPRMs is that a joint Gaussian distribution can be assumed between the training observations $\mathbf{y}$, and the test observation(s) (to be predicted) at the set of test inputs ($\mathbf{x}_*$) as follows:

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} = \mathcal{N}\left[\mathbf{0}, \begin{matrix} K_{XX}+\sigma_n^2\mathbb{I} & \mathbf{k}_{X\mathbf{x}_*} \\ \mathbf{k}_{\mathbf{x}_*X} & k_{\mathbf{x}_*\mathbf{x}_*}+\sigma_n^2\mathbb{I} \end{matrix}\right] \quad (6)$$

In equation 6, K_{AB} is used as a shorthand for $K(A,B)$, and $\mathbb{I}$ is the identity matrix. The predictive distribution over the prediction $\mathbf{y}_*$ can then be defined from the properties of multivariate Gaussian distributions (Ref. 38) as follows:

$$p(y_*|\mathbf{x}_*, X, \mathbf{y}) = \mathcal{N}(\mathbb{E}\{y_*\}, \mathbb{V}\{y_*\}) \quad (7)$$

such that

$$\mathbb{E}\{y_*\} = \mathbf{k}_{\mathbf{x}_*X}(K_{XX}+\sigma_n^2\mathbb{I})^{-1} \quad (8)$$

$$\mathbb{V}\{y_*\} = k_{\mathbf{x}_*\mathbf{x}_*} - \mathbf{k}_{\mathbf{x}_*X}(K_{XX}+\sigma_n^2\mathbb{I})^{-1}\mathbf{k}_{X\mathbf{x}_*} + \sigma_n^2) \quad (9)$$

Multi-output Gaussian Process Regression Models (MOGPRMs)

The method used in this study for training, and predicting using, MOGPRMs is the multivariate GPR algorithm presented in (Ref. 39). The details on the algorithm operation when it comes to training and prediction are presented therein; a summary of that presentation will be described next. Briefly, the algorithm is based on the matrix-variate Gaussian distribution, for which the probability density function can be presented as follows:

$$p(X|M,\Sigma,\Omega) = (2\pi)^{-\frac{dn}{2}}\det(\Sigma)^{-\frac{d}{2}}\det(\Omega)^{-\frac{n}{2}} \times \mathrm{etr}\left(-\frac{1}{2}\Omega^{-1}(X-M)^T\Sigma^{-1}(X-M)\right) \quad (10)$$

In the equation above, $X \in \mathbb{R}^{n\times d}$, $M \in \mathbb{R}^{n\times d}$, $\Sigma \in \mathbb{R}^{n\times n}$, and $\Omega \in \mathbb{R}^{d\times d}$ denote the random matrix, matrix mean, column covariance and row covariance matrices, respectively. Also, $\det(\cdot)$ and $\mathrm{etr}(\cdot)$ denote the determinant and the exponential of a matrix trace, respectively. A random matrix following a matrix-variate Gaussian distribution can be written as:

$$X \sim \mathcal{MN}(M,\Sigma,\Omega) \quad (11)$$

In the Multivariate GPR algorithm, the regression representation can be posed as follows: with each input, there lies d outputs, and thus the data set can be given by $\{\mathbf{x}_i \in \mathbb{R}^D, \mathbf{y}_i \in \mathbb{R}^{1\times d},\ i = 1,2,3,\ldots,n\}$, such that D is the dimension of each input point. The MOGPRM can then be formulated as:

$$\mathbf{y}_i = \mathbf{f}(x_i), \qquad \mathbf{f} \sim \mathcal{MGP}(\mathbf{u}(\mathbf{x}), k'_{\mathbf{x},\mathbf{x}'}, \Omega) \quad (12)$$

such that $\mathcal{MGP}$ denotes a multivariate Gaussian Process, $\mathbf{u}$ is the vector mean of the process (which is set typically to $\mathbf{0}$ as done with the SOGPRMs), and k' is the used kernel (herein also the squared exponential kernel described in equation 4 plus an added noise term, which replaces the missing noise term in the MOGPRM formulation in equation 12, that is:

$$k'(\mathbf{x}_i,\mathbf{x}_j) = k(\mathbf{x}_i,\mathbf{x}_j) + \delta_{ij}\sigma_n^2 \quad (13)$$

where δ_{ij} is the Kronecker delta.

Training just like with SOGPRM, minimizing the negative log likelihood function was chosen in (Ref. 39) as the hyperparameter optimization method. However, due to the presence of the extra hyperparameters in Ω, the negative log marginal likelihood would look as follows:

$$\frac{nd}{2}\ln(2\pi) + \frac{d}{2}\ln(\det(K')) + \frac{n}{2}\ln\det(\Omega) + \frac{1}{2}\mathrm{tr}((K')^{-1}Y\Omega^{-1}Y^T) \quad (14)$$

In this equation, Y is the full matrix of training outputs ($Y \in \mathbb{R}^{n\times d}$), and regular gradient-based numerical optimization techniques can be used just like with SOGPRMs.

Prediction Similar to SOGPRMs, prediction of each new set of outputs $\mathbf{f}_*$ given incoming test inputs X_* can be achieved as follows:

$$p(\mathbf{f}_*|X,Y,X_*) = \mathcal{MN}(\mathbb{E}\{\mathbf{f}_*\}, \hat{\Sigma}, \hat{\Omega}) \quad (15)$$

such that

$$\mathbb{E}\{\mathbf{f}_*\} = K'^T_{X_*X}K'^{-1}_{XX}Y \quad (16)$$

$$\mathbb{V}\{\mathbf{f}_*\} = \hat{\Sigma}\otimes\hat{\Omega} = [K'_{X_*X_*} - K'^T_{X_*X}K'^{-1}_{XX}K'_{X_*X}]\otimes\Omega \quad (17)$$

such that $\mathbb{E}\{\mathbf{f}_*\}$ and $\mathbb{V}\{\mathbf{f}_*\}$ are the vectors of predictive means and variances for the test points X_*. Note that, in this context, $\mathbf{f}_*$ is equivalent to $\mathbf{y}_*$ in the SOGPRM context.

Damage Size Quantification

In the case presented herein, the test observations at which prediction should take place are the GPRM targets/outputs, not the inputs i.e. a test DI (y_* - target) would be available, and the damage size (x_* - input) would be estimated by the GPRM. This reverse prediction process of predicting the damage size from a test DI point can be applied by estimating the probability that a test target value y_* belongs to a specific damage size x. The process of estimating this input state probability in

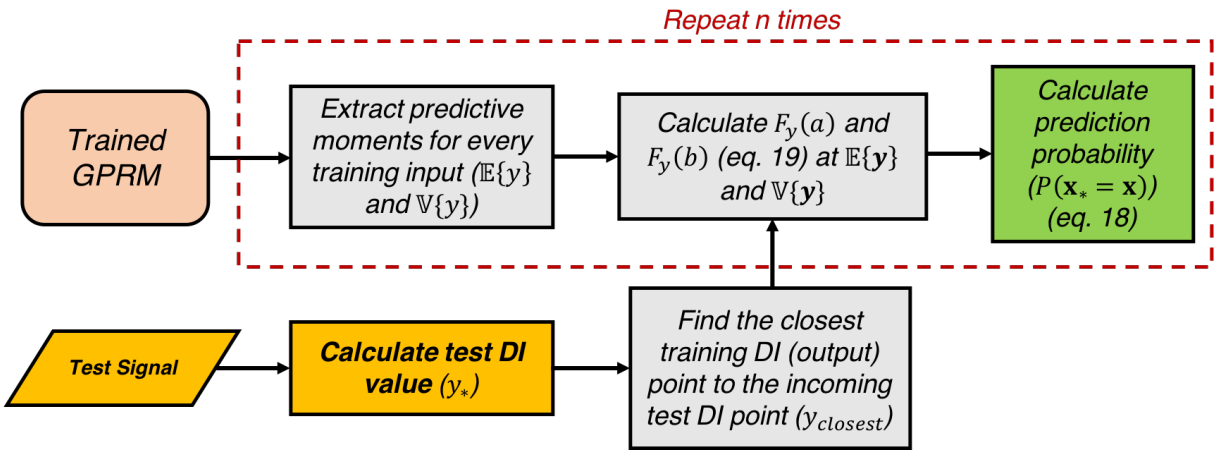

Fig. 1. A schematic flow diagram showing the steps taken in this study to calculate the damage size prediction probability.

schematically outlines in Figure 1. As shown, this probability ($P(\mathbf{x}_* = \mathbf{x})$) can be estimated from the Cumulative Distribution Function of the targets ($F_y(\cdot)$) as follows:

$$P(\mathbf{x}_* = \mathbf{x}) = F_y(b; E\{y\}, \mathbb{V}\{y\}) - F_X(a; E\{y\}, \mathbb{V}\{y\}) \quad (18)$$

such that,

$$F_y(s;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{t=-\infty}^{s} e^{\frac{-(t-\mu)^2}{2\sigma^2}} \quad (19a)$$

$$a = y_* - 2\sqrt{\mathbb{V}\{y_{closest}\}} \quad (19b)$$

$$b = y_* + 2\sqrt{\mathbb{V}\{y_{closest}\}} \quad (19c)$$

such that $\mathbb{V}\{y_{closest}\}$ is the GPRM predictive variance of the closest training DI value ($y_{closest}$) to the value of the incoming test DI point (y_*), while $\mathbb{E}\{y\}$ and $\mathbb{V}\{y\}$ are the GPRM predictive mean and variance at the training input x, respectively. This probability is calculated for every damage size in the training data, and information from the uncertainty in the GPRM (predictive variance) corresponding to the closest training DI value to the test one is utilized for properly estimating damage size probability. This damage size quantification framework can be easily expanded to MOGPRMs by replacing the predictive moments of SOGPRMs by those of MOGPRMs.

EXPERIMENTAL SETUP

A 152.4×304.8 mm (6×12 in) 6061 Aluminum coupon (2.36 mm/0.093 in thick) was used to demonstrate the proposed approach in this study. Six piezoelectric sensors type PZT-5A (Acellent Technologies, Inc) were attached to the coupon using Hysol EA 9394 adhesive. For curing the adhesive, the coupon was put under vacuum for 24 hours at room temperature. The coupon was then mounted onto a tensile testing machine (Instron, Inc). This allowed for the application of multiple static loading conditions in order to simulate real-life situations. 1-4 three-gram weights were attached onto the

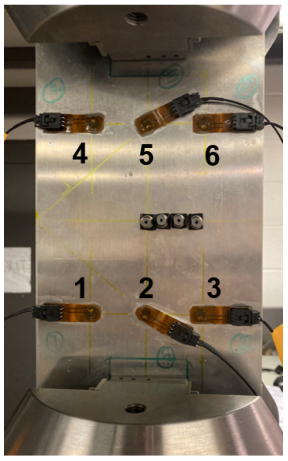

Fig. 2. The test case used in this study with four 3-gm weights simulating damage (largest damage size) shown here with the testing machine's grips.

surface of the plate during each loading state in the manner shown in Figure 2. For brevity, data sets from a static load state of 10 kN only were presented here.

An actuation of 5-peak tone bursts (5-cycle Hamming-filtered sine waves) with an amplitude of 90 V peak to peak and 250 kHz center frequency was generated at each sensor and the response signals of the sensors on the other side of the plate were collected. Twenty response signals per structural case were collected at each sensor (sampling rate of 24 MHz) using a ScanGenie III data acquisition system (Acellent Technologies, Inc). Acquired signals were imported into Matlab.[1] for further analysis.

For the implementation of the GPRMs, DI values from different numbers of paths were chosen and used to simultaneously

[1]Matlab version R2020a, GPRM training and prediction: the different functions within the GPML package available at http://www.gaussianprocess.org/gpml/code/matlab/doc/index.html, as well as the Multivariate Gaussian Process Regression (MV-GPR) package available at http://www.github.com/Magica-Chen/gptp_multi_output.com

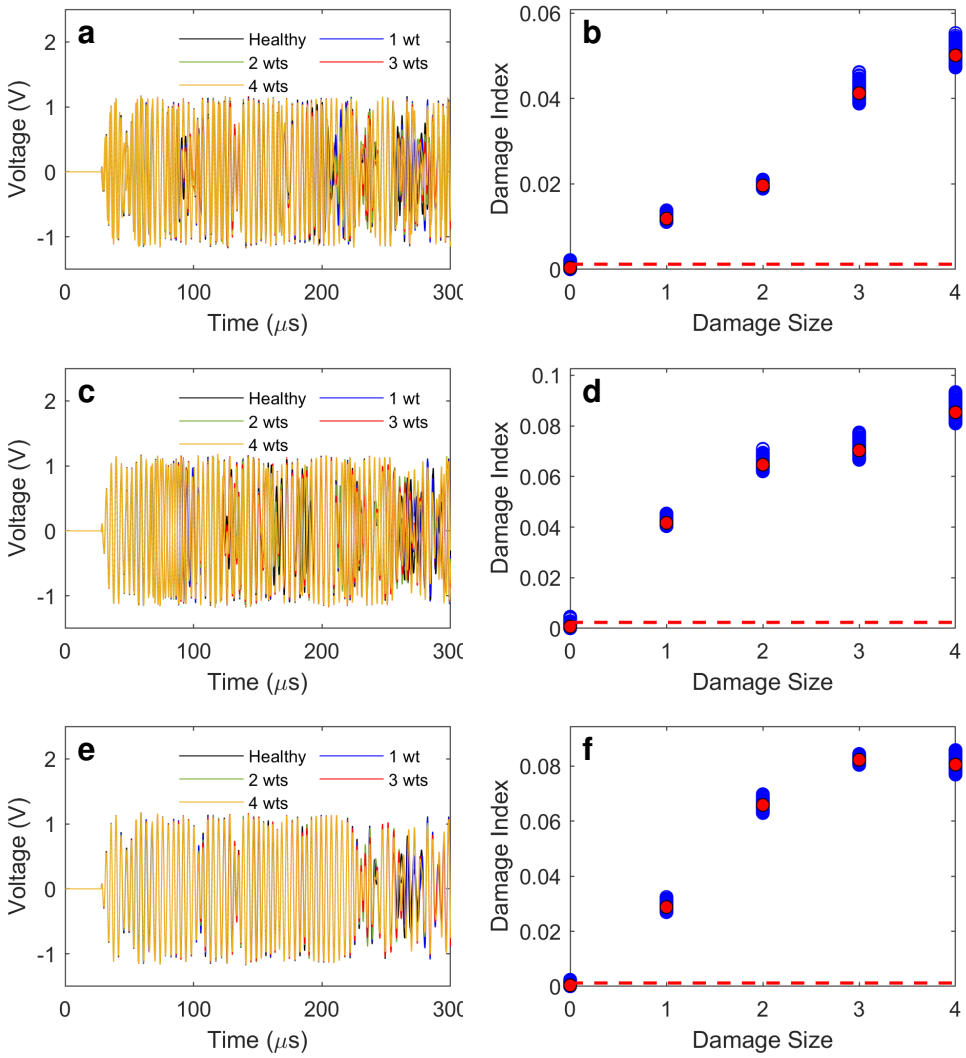

Fig. 3. Response signals and DI values for multiple damage sizes in the Al coupon at a load of 10 kN: (a) signal from path 1-4; (b) DI values for path 1-4; (c) signal from path 5-1; (b) DI values for path 5-1; (e) signal from path 6-2; (f) DI values for path 6-2. The red circles in the DI plots, (b), (d) and (f), correspond to the sample DI mean for each damage size.

train MOGPRMs as well as independently train SOGPRMs, i.e. one SOGPRM per path (e.g. for three paths, one MOGPRM and three SOGPRMs are trained). Then, the prediction results produced using the framework outlined in Figure 1 are compared across MO and SOGPRMs for each path in order to show the potential improvements of capturing cross-path data correlation training MOGPRMs for multiple paths simultaneously.

RESULTS AND DISCUSSION

MOGPRMs *versus* SOGPRMs

Each actuator-sensor signal path within a sensor network would generally result in a different DI evolution with damage size, with some showing uniform evolution, and some paths exhibiting DI saturation. Also, with the abundance of varying operational and/or environmental states, the amount of variation in the DI values would differ across different paths and/or damage sizes. The premise of the work presented herein is to leverage whatever information about damage size that is embedded within the signals coming from each path. This proposed information-fusion framework should then be able to provide more accurate and robust damage size estimations since it would be capable of making up for any non-uniformity in DI evolution, which could otherwise throw off single-path quantification models. Figure 3 shows the signals and corresponding DI plots for three different paths in the Al coupon under different damage states. The three paths presented here are a combination of damage-intersecting and damage-non-intersecting paths. They also exhibit somewhat different DI evolution, with path 1-4 showing varying levels of noise across damage sizes, path 5-1 showing some saturation

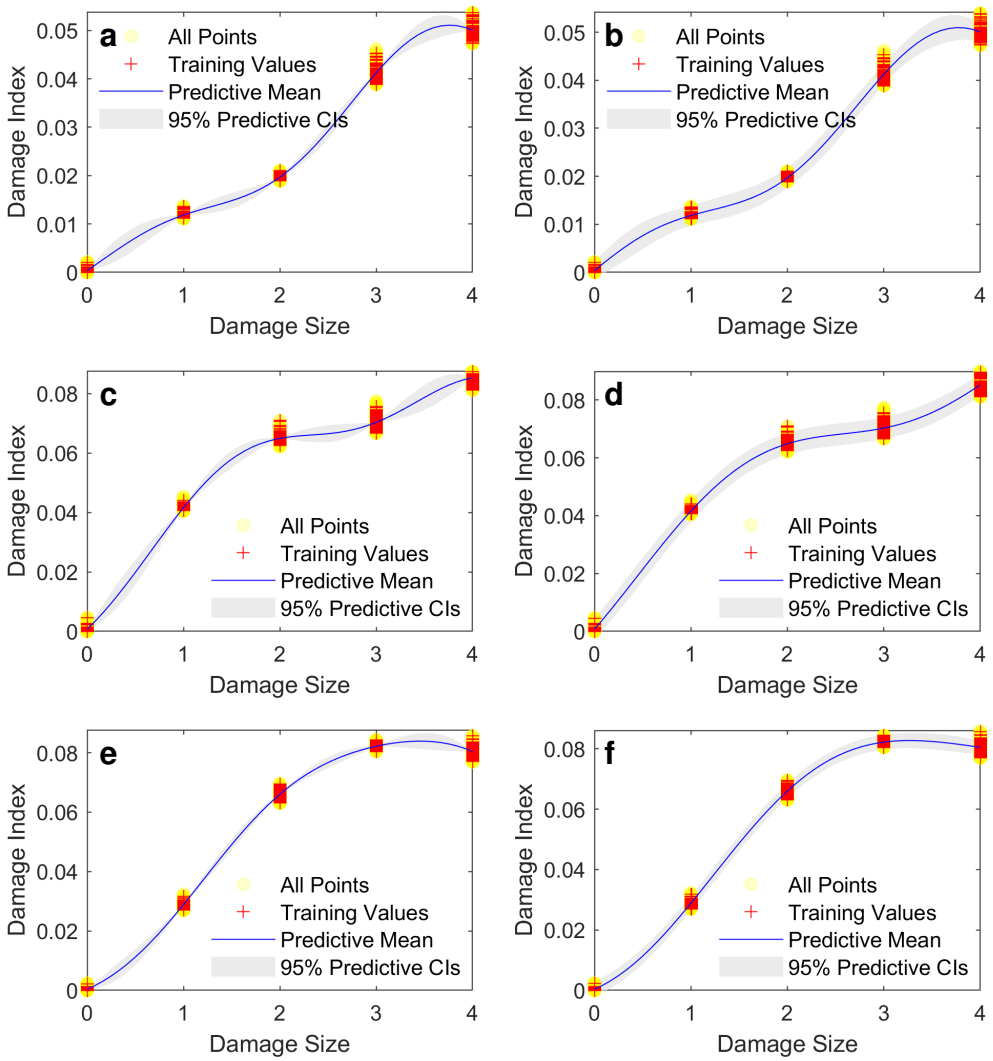

Fig. 4. GPRM predictive mean and confidence intervals (CIs) for the Al coupon based on the path-specific DI values from MO (three paths) and SOGPRMs: (a) MOGPRM predictive mean and CIs for path 1-4; (b) SOGPRM predictive mean and CIs for path 1-4; (c) MOGPRM predictive mean and CIs for path 5-1; (d) SOGPRM predictive mean and CIs for path 5-1; (e) MOGPRM predictive mean and CIs for path 6-2; (f) SOGPRM predictive mean and CIs for path 6-2.

between the DI values when 2 and 3 weights were attached, and path 6-2 showing saturation when 3 and 4 weights were attached to the coupon. As will be shown later, these non-uniformity features in the DI values can potentially throw off SOGPRMs trained for damage quantification using only any of these paths.

The DI values presented in the different panels of Figure 3 were separated into training and testing data sets for the purposes of training each SOGPRM (three total) as well as one MOGPRM. Table 1 summarizes the training and testing procedures of the models presented here, and the test data-based mean squared error was used as the criteria for comparing the trained models.

Figure 4 panels a, c, and e show the resulting MOGPRM predictive means and confidence bounds for each of the three paths presented in Figure 3. Panels b, d, and f of the same Figure show the corresponding SOGPRM predictive moments for the three paths. It is worth mentioning here that, again, the MOGPRM results come form a single trained MOGPRM, while the SOGPRM results come from three different models. As shown, the MOGPRMs can nicely capture the DI evolution for the three studied paths just like SOGPRMs. However, As shown in panels a, c, and e, the MOGPRM tends to underestimate the variance in the data as clear in the narrow confidence bounds compared to SOGPRMs. This underestimation may be avoided by proper initialization of the MOGPRM hy-

Table 1. Summary of DI-trained GPRM* information† for three paths in the Al coupon with simulated damage.

Signal	MSE		Training Time (s)		Prediction Time (s)	
Path	SO	MO	SO	MO	SO	MO
1-4	1.15E-6	8.97E-4	4.81	58.09	0.076	0.044
5-1	3.04E-6	3.6E-3	3.94		0.076	
6-2	1.45E-6	3.8E-3	4.04		0.070	

*500 DI data points from each path (out of 2000) were used for training.
†Numbers approximated to the last quoted decimal place.

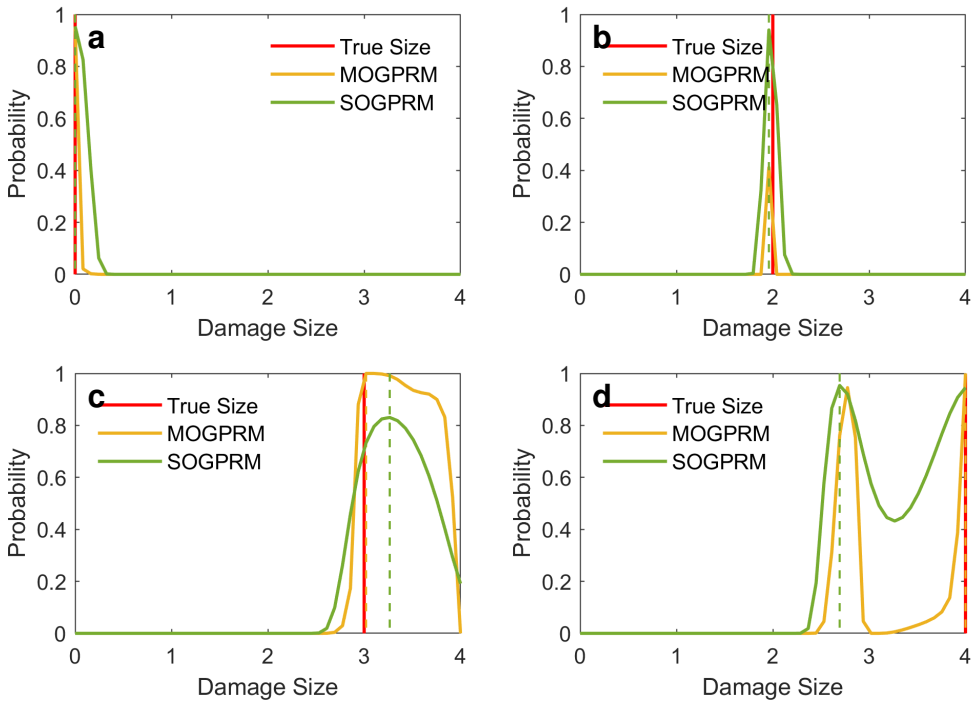

Fig. 5. Indicative damage size prediction results from MO (three paths) and SOGPRMs for path 6-2 in the Al coupon: (a) prediction probabilities for the healthy case; (b) prediction probabilities for two attached weights; (c) prediction probabilities for three attached weights; (d) prediction probabilities for four attached weights. Dashed vertical lines indicate the maximum probability corresponding (by color) to each model.

perparameters. This being said, as will be shown next, this phenomenon did not affect the accuracy of damage size estimation in this study.

In order to compare the damage quantification performance of the MOGPRM and SOGPRM approaches, damage size probabilities of test DI data sets (data not used in training) were estimated for path 6-2 and are shown in Figure 5 for the healthy and three damage states. As shown, although both models show accurate damage size estimation in the healthy state and in the case of two attached weights as indicated by the maximum probabilities (dashed vertical lines), the MOGPRM surpasses the SOGPRM when it comes to damage size estimation in the cases of three and four attached weights as shown in panels c and d, respectively. This enhancement exhibited by the MOGPRMs can be attributed to the MOGPRMs capturing the correlation between and fusing the information from the different path DI values at those two damage states, and consequently producing a more accurate prediction in each case. It is important to note here that the results shown in Figure 5 only come from four test DI points. In order to evaluate the overall performances of each model type, summary results of predictions from all DI test points need to be examined.

Figure 6 shows the summary results from the trained MOGPRM and SOGPRMs for all three paths shown in Figure 3. As shown, the trained MOGPRM results in generally sharper predictions for the different DI values with less outliers, as exhibited by the case of four attached weights in path 1-4 (panels a and b), the cases of 2 and 3 attached weights in path 5-1 (panels c and d), and the case of three attached weights in path 6-2 (panels e and f). In the latter path, for the case of four attached weights, it can be observed in panel (f) that the trained SOGPRM for this path cannot accurately differentiate between this case and the case of three attached weights, with the median and percentile predictions all lying around three weights, owing to the overlap in their DI values (see Figure 3f). However, the trained MOGPRM shows a sharper three-weight state prediction, as well as expands the percentiles of the 4-weight case to encompass the true state (4 weights) as shown in panel (e). These summary results show the superiority of MOGPRMs over SOGPRMs, as well as the capability of the former to fuse damage size information from multiple-path DI data sets thus making up for any shortcomings in single-path DI sets.

Effect of Number of Paths in MOGPRMs

With the superiority of MOGPRMs concluded, the question of whether adding information from more paths (i.e. adding DI sets from more paths in the MOGPRM training process) would yield better results was tackled next. Studies were done on training MOGPRMs using 5, 7, and 9 paths. Table 2 shows the paths from which the training DI sets were extracted for each of the four MOGPRMs. Figure 7 panels a, b, c and d show summary prediction results for path 5-1 coming from using DI sets from 3, 5, 7, and 9 paths, respectively, in training MOGPRMs. As shown, although the 3-path model already exhibited sharper prediction results compared to the trained SOGPRM (as shown in Figure 6c), especially around the saturation area (damage sizes of 2 and 3), the 5-path model exhibits yet sharper predictions than either. These results show that adding more paths in the learning process

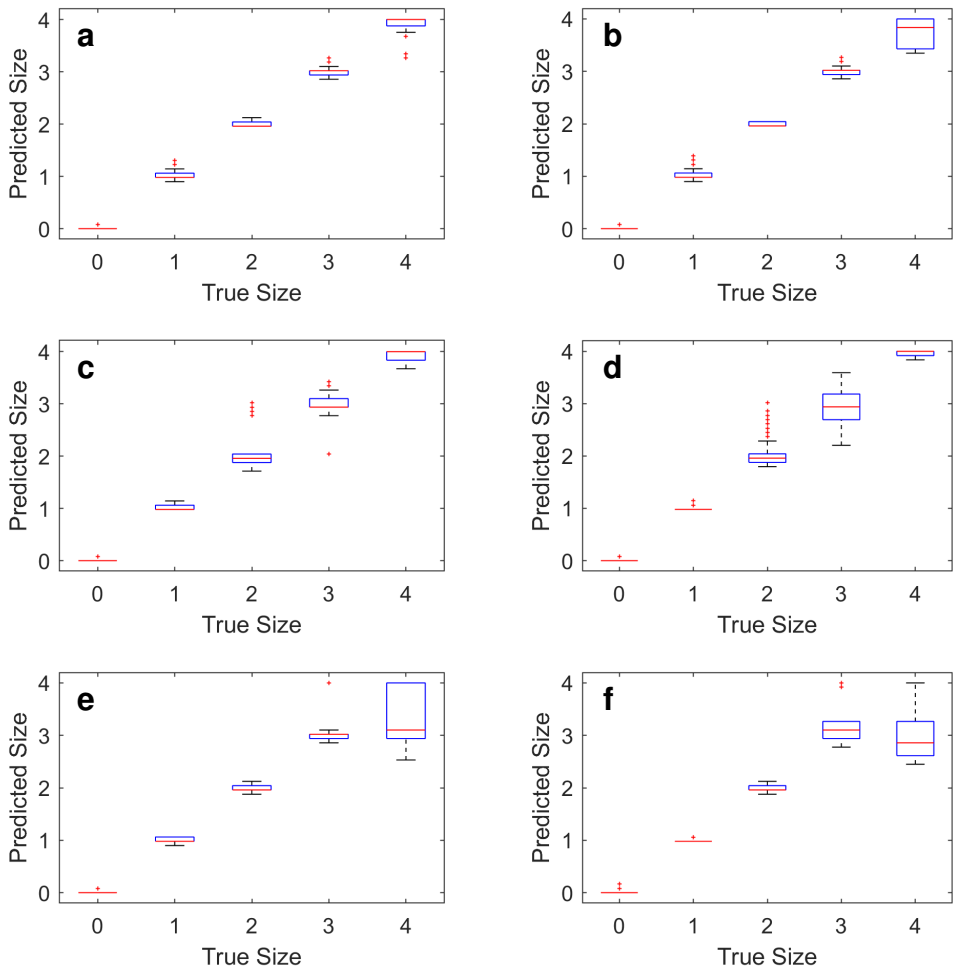

Fig. 6. True/predicted damage size boxplots for the Al coupon: (a) 3-path MOGPRM state predictions for path 1-4; (b) SOGPRM state predictions for path 1-4; (c) 3-path MOGPRM state predictions for path 5-1; (d) SOGPRM state predictions for path 5-1; (e) 3-path MOGPRM state predictions for path 6-2; (f) SOGPRM state predictions for path 6-2.

Table 2. Actuator-sensor signal paths used in the different MOGPRMs presented herein.

Number of included paths	1-4	5-1	6-3	4-2	6-2	4-1	4-3	2-6	1-6
Three	○	○			○				
Five	○	○	○	○	○				
Seven	○	○	○	○	○	○	○		
Nine	○	○	○	○	○	○	○	○	○

can lead to more accurate and robust damage size estimations. Switching onto the results for 7 paths (panel c), the results seem to be almost the same as for the 5-path model, with a few more outliers for the 2-weight damage state, and slightly broader prediction spectrum for the 4-weight case. Looking onto the results of the MOGPRM trained using 9 paths (panel d), it can be concluded that adding more paths might as well decrease accuracy, as indicated by the more dispersed predictions around the 2 and 3-weight cases. This deterioration in damage quantification accuracy can be attributed to the inclusion, in the training process, of DI sets from specific paths that exhibit a high degree of saturation around these two weights. This observation dictates that training MOGPRMs needs to be done with care with respect to the DI evolution of the paths being included in the training process: if many of the DI sets involved in the training process do not evolve nicely with damage size, this will lead to less accurate damage predictions.

In order to "visualize" the changes in damage size prediction probabilities as more paths are involved in the training of MOGPRMs, the prediction probabilities for two indicative test DI points were plotted for path 5-1 (Figure 8). As shown in Figure 8a, the increase in prediction accuracy by adding DI data sets from more paths to the training process can be clearly seen from the uniform narrowing of the prediction probability as more paths are added. This is attributed again to the

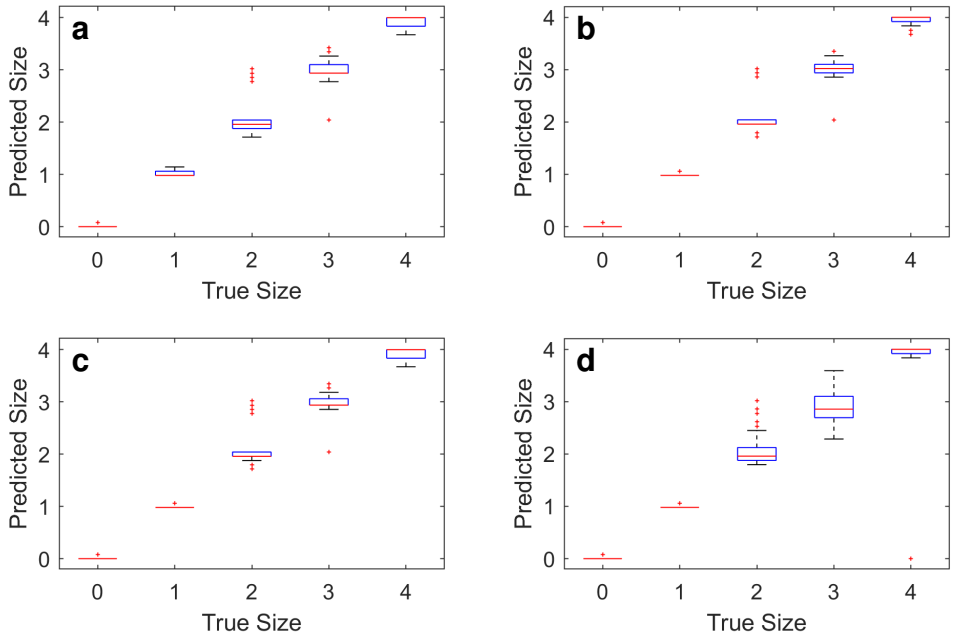

Fig. 7. MOGPRM true/predicted damage size boxplots for the Al coupon for path 5-1: (a) 3-path MOGPRM state predictions; (b) 5-path MOGPRM state predictions; (c) 7-path MOGPRM state predictions; (d) 9-path MOGPRM state predictions.

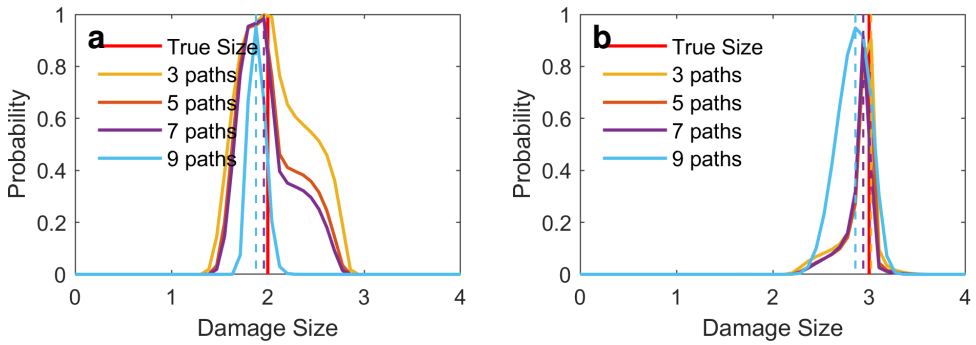

Fig. 8. Indicative damage size prediction results for path 5-1 from multiple MOGPRMs trained with DI sets from a different number of paths in the Al coupon: (a) prediction probabilities for two attached weights; (b) prediction probabilities for three attached weights. Dashed vertical lines indicate the maximum probability corresponding (by color) to each MOGPRM.

evolution of the DI sets used in training each MOGPRM. Observing the evolution of the prediction probabilities for the damage size of three attached weights (Figure 8b), the same overall trend can be observed up till the 7-path model, albeit with a very slight enhancement in the predicted damage size for the three models. However, the 9-path MOGPRM shows the broadest prediction probability, as expected from the results shown in Figure 7d. This is again attributed to the nature of the added training data sets.

Figures 9 and 10 show the corresponding analysis for path 1-4. As shown in Figure 9, a very similar trend in prediction accuracy evolves for path 1-4 compared to path 5-1. Also, as can be seen with the predictions shown in Figure 10, which are coming from indicative test DI points, the 9-path models shows a slightly less sharp (panel a) or lower (panel b) prediction probability, while the three other models show similar results, with the probabilities generally becoming better as more paths are included into the training process. Two conclusions can be made from this analysis. Firstly, adding DI sets from more paths has to be implemented considering the evolution of the DI with damage size. Secondly, in the case the "right" data sets are added to the MOGPRM training process, there seems to be some damage sizes for which there is a tangible enhancement in quantification accuracy, whilst there is barely any enhancement for other damage sizes. This latter conclusion might be related to the dispersion (or noise) in the calculated DI values for each specific damage size. Overall, these results shed light on the importance of selecting the proper DI data sets (or proper paths) when training MOGPRMs where the evolution of the DI values with damage size, as well as their dispersion for each, need to be considered.

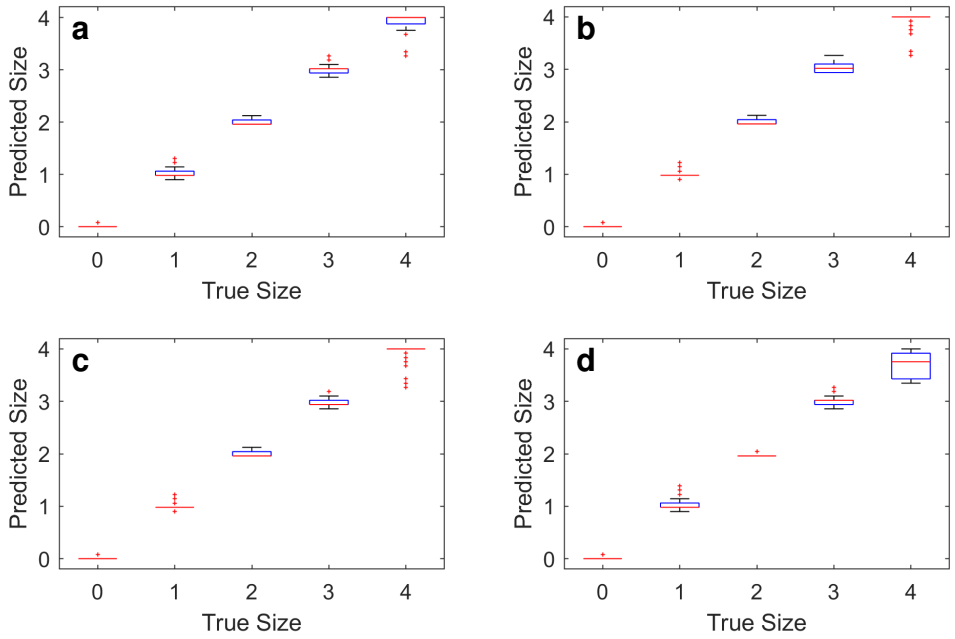

Fig. 9. MOGPRM true/predicted damage size boxplots for the Al coupon for path 1-4: (a) 3-path MOGPRM state predictions; (b) 5-path MOGPRM state predictions; (c) 7-path MOGPRM state predictions; (d) 9-path MOGPRM state predictions.

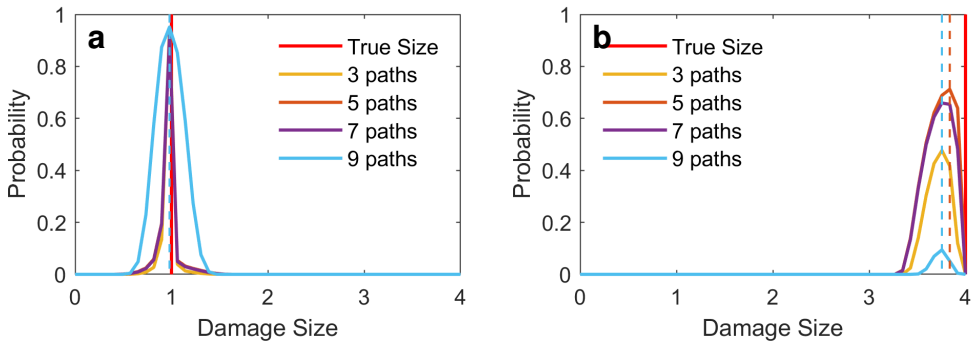

Fig. 10. Indicative damage size prediction results for path 1-4 from multiple MOGPRMs trained with DI sets from a different number of paths in the Al coupon: (a) prediction probabilities for one attached weight; (b) prediction probabilities for two attached weights. Dashed vertical lines indicate the maximum probability corresponding (by color) to each MOGPRM.

CONCLUSIONS

In this study, an information fusion approach was proposed for accurate and robust damage quantification within active-sensing, guided-wave SHM networks. Multi-output Gaussian Process regression models (MOGPRMs) were trained using DI values coming from different actuator-sensor paths in order to leverage damage size information from each path-based DI data set, while making up for the shortcomings of following damage evolution that might be encountered with some paths. The proposed framework was applied onto an Al coupon with multiple simulated damage states (attached weights), and the damage size prediction results were presented and compared to those of independently training one single-output GPRM (SOGPRM) for each path. It was shown that MOGPRMs exhibit more accurate and robust damage size estimation as shown by the sharper predictions and reduced number of outliers for each path. In addition, it was also shown that adding DI data sets from more paths in the MOGPRM training process has the potential of yielding more accurate (sharper) damage size predictions as long as the added DI data sets from the different paths do not exhibit too much distortion in the evolution with damage size or too much noise for each damage size. Overall, the approach proposed herein exhibits efficient handling of multi-path data sets and shows promise for probabilistic damage quantification within active-sensing, guided-wave SHM.

Author contact: Ahmad Amer, amera2@rpi.edu; Fotis Kopsaftopoulos, kopsaf@rpi.edu

ACKNOWLEDGMENTS

This work is carried out at the Rensselaer Polytechnic Institute under the Army/Navy/NASA Vertical Lift Research

Center of Excellence (VLRCOE) Program, grant number W911W61120012, with Dr. Mahendra Bhagwat and Dr. William Lewis as Technical Monitors.

REFERENCES

[1]Ahmed, S., Amer, A., Varela, C., and Kopsaftopoulos, F. P., "Data Driven State Awareness for Fly-by-Feel Aerial Vehicles via Adaptive Time Series and Gaussian Process Regression Models," Proceedings of the InfoSymbiotics/DDDAS2020 Conference, 2020.

[2]Kopsaftopoulos, F., Nardari, R., Li, Y.-H., and Chang, F.-K., "A stochastic global identification framework for aerospace structures operating under varying flight states," *Mechanical Systems and Signal Processing*, Vol. 98, 2018, pp. 425–447.
doi: 10.1016/j.ymssp.2017.05.001

[3]Dutta, A., McKay, M., Kopsaftopoulos, F., and Gandhi, F., "Rotor Fault Detection and Identification on a Hexacopter Based on Statistical Time Series Methods," Vertical Flight Society 75th Annual Forum, Philadelphia, PA, May 2019.

[4]Dutta, A., McKay, M., Kopsaftopoulos, F., and Gandhi, F., "Statistical Time Series Methods for Multicopter Fault Detection and Identification," Vertical Flight Society International Powered Lift Conference, San Jose, CA, Jan 2020.

[5]Yadav, S. K., Chung, H., Kopsaftopoulos, F. P., and Chang, F.-K., "Damage Quantification of Active Sensing Acousto-ultrasound-based SHM Based on a Multi-path Unit-cell Approach," Proceedings of the 11th International Workshop on Structural Health Monitoring (IWSHM 2017), September 2017.

[6]Mitra, M. and Gopalakrishnan, S., "Guided wave based structural health monitoring: A review," *Smart Materials and Structures*, Vol. 25, (5), MAY 2016.

[7]Diamanti, K. and Soutis, C., "Structural health monitoring techniques for aircraft composite structures," *Progress in Aerospace Sciences*, Vol. 46, (8), NOV 2010, pp. 342–352.

[8]Davis, M., Bouquillon, B., Smith, M., Allred, C., Sarjeant, R., Loverich, J., and Bordick, N., "Rotor load and health monitoring sensor technology," American Helicopter Society 71st Annual Forum Proceedings, May May 2015.

[9]LeFevre, B., Davis, M., Marr, C., Rusak, D., and Johnson, C., "Integrated Hybrid Structural Management System (IHSMS): usage and loads monitoring," American Helicopter Society 73st Annual Forum & Technology Display Proceedings,, May 2017.

[10]Schenck, E., Davis, M., Bond, R., Meyer, J., and Rusak, D., "Integrated Hybrid Structural Management System (IHSMS) – aircraft impact monitoring," American Helicopter Society 73rd Annual Forum Proceedings, May 2017.

[11]Amer, A. and Kopsaftopoulos, F. P., "Probabilistic active sensing acousto-ultrasound SHM based on non-parametric stochastic representations," Proceedings of the Vertical Flight Society 75th Annual Forum & Technology Display, May 2019.

[12]Jin, H., Yan, J., Li, W., and Qing, X., "Monitoring of fatigue crack propagation by damage index of ultrasonic guided waves calculated by various acoustic features," *Applied Sciences*, Vol. 9, 2019, pp. 4254.

[13]Amer, A. and Kopsaftopoulos, F. P., "Probabilistic Damage Quantification *via* the Integration of Non-parametric Time-series and Gaussian Process Regression Models," Proceedings of the 12th International Workshop on Structural Health Monitoring (IWSHM 2019), September 2019.

[14]Amer, A. and Kopsaftopoulos, F. P., "Towards Unified Probabilistic Rotorcraft Damage Detection and Quantification via Non-parametric Time Series and Gaussian Process Regression Models," Proceedings of the Vertical Flight Society 76th Annual Forum & Technology Display, October 2020.

[15]Castro, E., Moreno-Garcia, P., and Gallego, A., "Damage Detection in CFRP Plates Using Spectral Entropy," *Shock and Vibration*, 2014, pp. 1–8.

[16]Wilson, C. L., Lonkar, K., Roy, S., Kopsaftopoulos, F., and Chang, F.-K., "Structural Health Monitoring of Composites," *Comprehensive Composite Materials II*, edited by P. W. R. Beaumont and C. H. Zweben, Elsevier Ltd., 2018, pp. 382–407.

[17]Janapati, V., Kopsaftopoulos, F., Li, F., Lee, S., and Chang, F.-K., "Damage detection sensitivity characterization of acousto-ultrasound-based structural health monitoring techniques," *Structural Health Monitoring*, Vol. 15, (2), 2016, pp. 143–161.

[18]Ahmed, S. and Kopsaftopoulos, F. P., "Uncertainty quantification of guided waves propagation for active sensing structural health monitoring," Proceedings of the Vertical Flight Society 75th Annual Forum & Technology Display, May 2019.

[19]Zhao, J., Gao, H. D., Chang, G. F., Ayhan, B., Yan, F., Kwan, C., and Rose, J. L., "Active health monitoring of an aircraft wing with embedded piezoelectric sensor/actuator network: I. Defect detection, localization and growth monitoring," *Smart Materials and Structures*, Vol. 16, (4), 2007, pp. 1208–1217.

[20]Memmolo, V., Ricci, F., Boffa, N. D., Maio, L., and Monaco, E., "Structural Health Monitoring in Composites Based on Probabilistic Reconstruction Techniques," *Procedia Engineering*, Vol. 167, 2016, pp. 48–55.

[21]Flynn, E. B., Todd, M. D., Wilcox, P. D., Drinkwater, B. W., Croxford, A. J., and Kessler, S., "Maximum-likelihood estimation of damage location in guided-wave structural health monitoring," *Proceedings of The Royal Society A, Burlington, VT*, Vol. 467, (2133), 2011, pp. 2575–2596.

[22]Todd, M. D., Flynn, E. B., Wilcox, P. D., Drinkwater, B. W., Croxford, A. J., and Kessler, S., "Ultrasonic wave-based defect localization using probabilistic modeling," American Institute of Physics Conference Proceedings, May 2012.

[23]Clarke, T. and Cawley, P., "Enhancing the defect localization capability of a guided wave SHM system applied to a complex structure," *Structural Health Monitoring*, Vol. 10, (3), 2011, pp. 247–259.

[24]Sharif Khodaei, Z. and Aliabadi, M. H., "A Multi-Level Decision Fusion Strategy for Condition Based Maintenance of Composite Structures." *Materials*, Vol. 9, (9), September 2016.

[25]Xu, B., Zhang, T., Song, G., and Gu, H., "Active interface debonding detection of a concrete-filled steel tube with piezoelectric technologies using wavelet packet analysis," *Mechanical Systems and Signal Processing*, Vol. 36, 2013, pp. 7–17.

[26]Ihn, J. and Chang, F.-K., "Pitch-catch active sensing methods in structural health monitoring for aircraft structures," *Structural Health Monitoring*, Vol. 7, (1), 2008, pp. 5–19.

[27]Giurgiutiu, V., "Piezoelectric Wafer Active Sensors for Structural Health Monitoring of Composite Structures Using Tuned Guided Waves," *Journal of Engineering Materials and Technology*, Vol. 133, (4), 2011, pp. 041012.

[28]Nasrollahi, A., Deng, W., Ma, Z., and Rizzo, P., "Multimodal structural health monitoring based on active and passive sensing," *Structural Health Monitoring*, Vol. 17, (2), 2018, pp. 395–409.

[29]Kralovec, C. and Schagerl, M., "Review of Structural Health Monitoring Methods Regarding a Multi-Sensor Approach for Damage Assessment of Metal and Composite Structures." *Sensors (Basel, Switzerland)*, Vol. 20, (3), February 2020.

[30]Derriso, M. M., Little, I., John, E., Vehorn, K. A., Davies, M. J., and DeSimio, M. P., "Crack detection using combinations of acoustic emission and guided wave signals from bonded piezoelectric transducers," Technical report, AIR FORCE RESEARCH LAB WRIGHT-PATTERSON AFB OH AIR VEHICLES DIRECTORATE, 2011.

[31]Su, Z. and Ye, L., "Lamb wave-based quantitative identification of delamination in CF/EP composite structures using artificial neural algorithm," *Composite Structures*, Vol. 66, 2004, pp. 627–637.

[32]Song, G., Gu, H., and Mo, Y.-L., "Smart aggregates: multi-functional sensors for concrete structure —a tutorial and a review," *Smart Materials and Structures*, Vol. 17, 2008, pp. 033001.

[33]Tibaduiza, D. A., Mujica, L. E., Rodellar, J., and Güemes, A., "Structural damage detection using principal component analysis and damage indices," *Journal of Intelligent Material Systems and Structures*, Vol. 27, (2), 2016, pp. 233–248.

[34]He, J., Ran, Y., Liu, B., Yang, J., and Guan, X., "A Lamb wave based fatigue crack length estimation method using finite element simulations," The 9th International Symposium on NDT in Aerospace, Xiamen, China, November 2017.

[35]Peng, T., Saxena, A., Goebel, K., Xiang, Y., Sankararaman, S., and Liu, Y., "A novel Bayesian imaging method for probabilistic delamination detection of composite materials," *Smart Materials and Structures*, Vol. 22, 2013, pp. 125019–125028.

[36]Amer, A., Ahmed, A., and Kopsaftopoulos, F. P., "Active-sensing Structural Health Monitoring via Statistical Learning: An Experimental Study under Varying Damage and Loading States," Special Volume, 10th European Workshop on Structural Health Monitoring (EWSHM 2022), 2020.

[37]C. E. Rasmussen and C. K. I. Williams, editors, *Gaussian Processes for Machine Learning*, MIT Press, 2006.

[38]Rogers, T., Gardner, P., Dervilis, N., Worden, K., A.E., M., Papatheou, E., and Cross, E., "Probabilistic modelling of wind turbine power curves with application of heteroscedastic Gaussian process regression," *Renewable Energy*, Vol. 148, 2020, pp. 1124–1136.

[39]Chen, Z., Wang, B., and Gorban, A. N., "Multivariate Gaussian and Student-t process regression for multi-output prediction," *Neural Computing and Applications*, Vol. 32, (8), 2020, pp. 3005–3028.