A data-driven method for predicting structural degradation using a piezoceramic array

Kyle R Mulligan¹, Chunsheng Yang², Nicolas Quaegebeur¹, and Patrice Masson¹

¹ GAUS, Dept. of Mechanical Engineering, Université de Sherbrooke, Sherbrooke, QC, J1K 2R1, CANADA. Kyle.Mulligan@USherbrooke.ca Nicolas.Quaegebeur@USherbrooke.ca Patrice.Masson@USherbrooke.ca

² Information and Communications Technology, National Research Council of Canada, Ottawa, ON, K1A 0R6, CANADA. Chunsheng.Yang@nrc-cnrc.gc.ca

ABSTRACT

There is a growing use of carbon fiber reinforced polymers (CFRPs) in modern airframes with still a limited understanding of the in-service behavioral characteristics of these structures. Structural Health Monitoring (SHM) technologies that use surface-bonded piezoceramic (PZT) transducers to generate and measure guided waves within these structures have demonstrated promising damage detection and localization results and potential for data gathering in data-driven damage prognosis. This paper investigates the development of a data-driven SHM based damage prognosis system for estimating remaining useful life (RUL) of CFRP coupons following damage initiation. A robust and realistic laboratory data gathering methodology is introduced as a building block for evaluating the feasibility of data-driven damage prognosis for in-service aerospace structures. Data are gathered using a PZT-based SHM system. Using the gathered raw guided wave signals, a number of time and frequency domain features are first extracted which are derived from existing damage imaging and detection algorithms. Then, using various combinations of the feature sets as inputs to generic data mining algorithms, the paper presents estimates of the predicted RUL against actual damage diameter progression.

1. INTRODUCTION

Modern operation of aircraft generates vast amounts of 'operational data' from on-board aircraft systems and 'maintenance data' from offline maintenance procedures. Prognostic model development maximizes the potential uses of these two data sources with the goal of discovering 'a priori' knowledge of component failures and abnormal system behavior (Zaluski, Létourneau, Bird, & Yang, 2010). Such prognostic tools would be invaluable to end-users of these aircraft in allowing the development of opportunistic and preventive maintenance procedures designed to cut costs and increase safety. Prognostic modelling would close the loop in the development of adequate health management systems for aircraft which are built with the combination of diagnostic, prognostic, and repair management technologies (Létourneau et al., 2005).

In any data-driven prognostic model development approach, data are gathered using a sensory and maintenance network. In the aerospace domain, data analysis of structural damage is tracked off-line by experienced engineers through nondestructive inspection (Roemer, Ge, & Liberson, 2005). Data are therefore collected when a problem is detected and no information is gathered leading up to the event. Without a means to gather data prior to a problem, data-driven prognostic model development is impossible. The use of Structural Health Monitoring (SHM) strategies based on guided wave propagation for injecting and receiving guided waves, using lead-zirconate-titanate (PZT) piezoceramic transducers bonded to aerospace structures for damage detection, is however increasing. This allows data gathered regularly during in-service operation of these aircraft to be used in the development of data-driven prognostic models and eventual decision support systems. Currently, however, SHM systems are not being readily used on in-service aircraft fleets (commercial or military) because the technology has not yet transitioned from research to practice (Chattopadhyay et al., 2012). Efforts to develop health management systems that include damage prognosis are limited because gathered data from SHM systems for real applications are not yet available.

An increasing importance in industry for decision support systems and condition based monitoring (CBM) through on-

Kyle R Mulligan et.al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



Figure 1. Overview of the data-driven methodology

board SHM comes from a growing use of composite Carbon Fiber Reinforced Polymers (CFRPs) in both military and civilian aircraft fabrication and production. As these aircraft are commencing in-service use, there is limited historical data available for understanding their in-service behavioral characteristics (Soutis, 2005). Damage in aerospace composites under in-service loading is mainly due to fatigue from stresses that vary depending on mission loading profiles. Operational impacts also occur which are less predictable. For CFRP structures, fatigue and impact damages lead to: fiber breakage, matrix cracking, delamination, and interface debonding (O'Brien, 2001; Rhymer, Kim, & Roach, 2012). The result of the initiation of such damage with further use of the aircraft could have detrimental effects on the overall strength of the structure. A desire for the aircraft manufacturer, owner, and user is for the development of on-board SHM systems for fault detection and eventual failure prediction of the remaining useful life (RUL) of these components to enhance opportunistic and preventive maintenance procedures with anticipated impacts on availability, safety, and cost (Zaluski et al., 2010).

With no commonly available operational data to build datasets for data-driven model development, much research has focused on using Material Testing Systems (MTS) in laboratory environments for data gathering and on analytical and Finite Element (FE) models of composite structural behavior (Liu, Mohanty, & Chattopadhyay, 2009). In laboratory environments, RUL estimations have been presented that use features created from data gathered using MTS fatigue systems in combination with acoustic emission or PZT-based SHM systems (Saxena, Goebel, Larrosa, Janapati, & Roy, 2011). Promising results in predicting remaining fatigue loading cycles have been put into the types of loading pat-

terns that should be used to mimic real applications (Liu et al., 2009; Larrosa & Chang, 2011). One must consider what loading patterns actual aircraft endure in order to develop realistic prognostic models. Furthermore, parameters measured from the MTS system cannot be used as features in the prognostic model development process because these features are not available in real applications and vary by composite structure type.

Many analytical models are available (Hashin, 1980; Chang & Chang, 1987; Roebuck, Gorley, & McCartney, 1989; Choi, 1990; Hou, Petrinic, Ruiz, & Hallett, 2000; Nairn, 2000; O'Brien, 2001; Beaumont, Diamant, & Shercliff, 2006) for understanding damage propagation and the transition from matrix micro-cracks to delamination. In practice, validation of structural components using such models is time consuming and complex and many assumptions for simplification are made when estimating fatigue (Liu et al., 2009). FE models exist to evaluate impact damage size and depth in aluminum sandwich structures (Nguyen, Jacombs, Thomson, Hachenberg, & Scott, 2005) but none, to the authors' knowledge, have been validated and used for RUL prediction. SHM allows a straight-forward data-driven implementation for damage prognosis which in turn provides a practical solution to difficulties faced in physics-based modelling (M. Chiachio, Chiachio, Saxena, Rus, & Goebel, 2013; J. Chiachio, Chiachio, Saxena, Rus, & Goebel, 2013).

To address these issues, this work presents a novel datadriven approach for estimating RUL for CFRP aerospace grade coupons exposed to low energy drop-weight impacts. The approach relies on a PZT-based SHM systems instrumented onto CFRP coupons in a pitch and catch configuration. Data are acquired from guided wave signals generated and received in the coupons and later pre-processed to define a feature set. Observations in the feature set are divided



Figure 2. Schematic (a) and picture (b) of the experimental setup of a CFRP aerospace grade composite coupon. The picture shows the coupon inside the drop-weight impact system.

into training and testing datasets where several generic supervised learning algorithms are applied to the data to estimate RUL. The laboratory data is gathered using a methodology designed as a foundation to build upon for simulating real world operating conditions. The purpose of this first step is to evaluate and demonstrate the feasibility of assessing structural damage using damage prognosis.

2. DATA-DRIVEN PROGNOSTIC MODEL DEVELOPMENT METHODOLOGY

There is an emerging use of data-driven based damage prognosis in the aerospace domain with the onset of an increasing use of composite structural components. In order to demonstrate the feasibility of damage prognosis for RUL prediction in aerospace grade composite structures following impact events a thorough, robust, and realistic laboratory data gathering procedure must be considered. This methodology consists of three major steps: data generation and processing, feature generation, and modelling and evaluation following operational data gathering. The defined data-driven based methodology can be extrapolated into experiments for developing prognostic models to predict damage degradation or RUL in aerospace structures. This section, together with Fig. 1, describes the details for each of the different stages in the methodology along with observed challenges.

2.1. Data generation and processing

In the present project, destructive testing using a drop-weight impact system is used to investigate damage prognosis for estimating RUL. By defining a critical damage size, the RUL is predicted in remaining number of impacts (RNI) from the critical impact point which is estimated for the critical damage size ($\simeq 7.1$ mm). Although somewhat arbitrary in this work, the critical damage size is estimated from nondestructive inspection measurements following a series of impacts over the entire sample size of CFRP coupons used in this work. Important considerations for the methodology must be made because following destructive impacts, data cannot be re-collected. Due to the high cost of destructive testing in aerospace structures, CFRP aerospace grade coupons instead of real aerospace structures are used for operational data gathering. The maintenance database however cannot be created until established maintenance procedures are defined and developed for such CFRPs. Structural fabrication and instrumentation and damage initiation procedures are first described in this section followed by a description for data gathering and processing.

2.1.1. Structural fabrication and instrumentation

A CFRP aerospace grade sheet is fabricated using the facilities at the National Research Council of Canada (NRC) Institute for Aerospace Research (IAR) in Ottawa, Ontario. The sheet is fabricated using pre-preg fibre layers autoclave cured inside a vacuum bag while keeping the tolerance on the ply orientation for the fabrication process below 0.5°. Following autoclave curing, the CFRP sheet is cut into 9 individual coupons of size 10.16 cm by 15.24 cm using a diamond saw. The composite layup and mechanical properties of the coupons are presented in Table 1.

Following fabrication, instrumentation of the CFRP coupons is performed. Two 5 mm diameter with 0.5 mm thickness PZT transducers (Physik Instrumente®) are bonded 10 cm apart in a pitch and catch configuration to the surface of the CFRP coupons using epoxy. The properties of the PZT transducers and the epoxy adhesive are outlined in Table 2. Damping tape at the plate edges is used to prevent edge reflections of guided wave signals.

Plies	Dimensions	Orientation	Thickness	Eq. Young's	Eq. Poisson's	
16	10.16cm x 15.24cm	[0.90]»	(mm) 2.45+0.01	74.5+8.6	0.33 ± 0.01	

Item	Material	Thickness	Young's	Poisson's	Density
		(mm)	modulus (GPa)	ratio	(kg/m ³)
Plain weave fibers (uni-dir)	ACG MTM45-1 GA045	0.15	61.5	0.3	1180
PZT	PZT 5A	0.5	65.0	0.31	7750
Bonding layer	Ероху	0.1	3.5	0.3	1100

Table 1. CFRP coupon properties

Table 2. Properties of the CFRP layers, PZT transducer, and bonding layer.

2.1.2. Damage initiation

Following instrumentation, coupons are placed individually into a drop-weight impact system for damage initiation. Two steel shims are placed on the upper and lower sides of the coupons and clamped using 5 quick clamps onto the impact system frame. Data acquisition described in the next section, is repeated for each coupon to define the operational database. The experimental setup of the plate is presented in Fig. 2.

ASTM standard 'D7136/D7136M-05' (ASTM, 2007) provides guidelines for impacting CFRPs using a drop-weight impact system such that damage including dents/depressions, splits/cracks, combined splits/delaminations, and combined large cracks with fibre breakage can be reproduced. For the thickness of the coupons used in this work, the standard dictates that an impact energy of 16.7J will induce all of the described failure modes. In other words, following an impact of this magnitude, the coupon would be deemed damaged and would require repair or replacement. For successful development of data-driven data mining prognostic models, the objective is to measure and produce data features that show progressive changes far away from a failure point. This would allow more time for fleet managers to schedule and perform opportunistic maintenance. Therefore, impacting the composite coupons using the standardized energy would not produce any useful data as the coupons would be too severely damaged (Guida, Marulo, Meo, & Russo, 2012) and there would be no justification for damage prognosis.

In reality, sudden impacts do occur which can only be detected and assessed but a more realistic application for damage prognosis in the aerospace domain would be for prediction of RUL after the onset of lower energy impacts from sources such as accidental tool drops or runway debris impacts from take-off and landing (2-3 J, 3-5 m/s) (Tomblin et al., August 1999). These impacts cause barely visible impact damage (BVID) which commonly occurs from constant maintenance and operation of the aircraft. Accumulations of BVID in similar proximities on the airframe surface can eventually lead to severe damage (Rhymer et al., 2012). Therefore, in this work, damage is applied in the form of 5J impacts to each composite coupon to replicate BVID. Actual damage size and depth is measured following each impact using a non-destructive phased array ultrasonic testing tool (Olympus OMNIScan MX2) (gold standard) (Olympus, 2013). These types of damage investigated herein are excellent candidates for RUL estimation as they are pertinent and common.

2.1.3. Data gathering

Guided wave PZT based SHM systems are being explored in current research and show promise in damage detection, localization, and material characterization for aerospace structures and are ideal candidates for in-situ data generation and gathering (Giurgiutiu & Bao, 2004; Quaegebeur, Masson, Langlois-Demers, & Micheau, 2010; Saxena et al., 2011; Ostiguy, Quaegebeur, Mulligan, Masson, & Elkoun, 2012). Before implementing such a system for structural assessment, key considerations must be made for transducer signal generation and acquisition parameters. In this case, the acquired data must be useful for features to be developed for later use in damage prognosis. For signal generation using guided wave propagation, a burst function must be selected with an appropriate amplitude, frequency, number of cycles, and duration (Staszewski, Mahzan, & Traynor, 2009). Features may be developed by investigating time, frequency, dispersive, and correlation aspects of the acquired signals following generation. Some of these features may be sensitive to different frequency ranges which vary by material. Therefore, in this study, considerations made for parameter selection are based on damage analysis and material characterization techniques (Giurgiutiu & Bao, 2004; Quaegebeur et al., 2010; Ostiguy et al., 2012).

For guided wave pitch and catch measurements a broadband frequency range is covered with fine steps between each frequency such that a diverse amount of data are gathered (Quaegebeur, Masson, Micheau, & Mrad, 2012) in order to develop new features not found in literature. The subband generation technique (Quaegebeur et al., 2012) is used to determine the transfer function between the transmitting and receiving transducers. To do this, an impulse excitation signal which is decomposed over 11 sub-bands over a frequency range below 1 MHz is transmitted using a generating PZT and received by the acquisition PZT. The signals are amplified using a UA-8400 amplification system (Produitson). Baseline guided wave measurements are taken before the impact sequence such that pristine and damage sig-



Figure 3. Example of a time domain signal without damage (a) and with damage (c). The difference between the undamaged and damaged time signals is shown in (e). The amplitude part of the frequency domain transfer function representation of the time domain signals without damage (b) and with damage (d). The difference between the undamaged and damaged frequency domain signals is shown in (f). All signals are normalized to the baseline.

nals can be later compared. Post-processing of the measured impulse response is performed using windowing and reconstruction filters to determine the transfer function. The input signal voltage burst is generated using an HP 33120A generator with a sampling frequency of 15 MHz. The output of the measurement circuit is acquired using a high impedance National Instruments PCI-5105 12-bit DAQ board configured through a custom LabVIEW interface. The generated signals are recorded at a fixed sampling frequency of 6 MHz and averaged 1000 times in order to increase SNR and low-pass filtered at 1.5 MHz.

2.1.4. Data pre-processing

Data-driven damage prognosis models require as input a dataset which is composed of instances of vectors of attribute values. The attributes and their respective values are extracted from the operational and maintenance databases obtained for specific applications. In the case of composite structures in aircraft, historical databases are not yet available. To build a historical database in this case, considerations must be made for selecting an appropriate SHM system that can provide *insitu* and regular monitoring of the aircraft structure during future applications for in-service operation.

In order to use supervised learning algorithms, data must be pre-processed. This involves the addition of problem identification and index attributes to the gathered operational data. Problems separated by problem identification numbers are created by applying the data gathering procedure to independent aerospace structures. With an attribute for identifying each problem, the data are split into training and testing datasets. For each of the training problems, an index attribute is added to each instance for supervised learning. The index attribute associates a number to each instance within each problem in sequential order where the failure point is generally labeled '-1' and the furthest point from the problem is '-N', where N is the total number of instances within a problem. In model training, the index attribute is used to train the predictive model for each individual observation. In model testing, testing data is inputted into the trained models and the index attribute for each observation is estimated. The predicted index for each instance in the testing dataset is compared to the actual index by replacing the index attribute for each instance. Depending on the quality and structure of the gathered data, data labelling can be performed at this stage or later, following feature generation.

2.2. Feature generation

Features are generated from raw data measurements using data transformations to improve the initial, as measured, representation (Zaluski et al., 2010). This is done by augmenting the initial representation with new features created using methods from process physics, signal processing, time-domain and frequency analysis, wave dispersion and correlation, and constructive induction. Different and more complex features can then be extrapolated from these findings. Once a feature set is defined, attribute evaluation tools (Hall, 2000;



Figure 4. Method for extracting the one sample instance of the dispersion feature from a measured pitch and catch signal.

Kira & Rendell, 1992) and domain knowledge can be used to remove redundant or irrelevant features to optimize model development and reduce computation time.

Guided wave measurements are used to build a feature set with instances for each of the 9 coupons. A total of 151 parameters are extracted based on time and frequency domain signals. An example of guided wave signals obtained using sub-band generation are provided in Fig. 3, with (c and d) and without (a and b) coupon damage showing the time and frequency domain representations. The difference between the undamaged and damage signal is shown in Fig. 3 (e) for the time domain and frequency domain in Fig. 3 (f). From the difference between the signals alone, slight differences can be observed. Therefore, both domains are considered in the feature extraction process. Once features are extracted, they are assembled into a database with the addition of problem identification and index attributes. Based on the problem identification attribute the database is separated into training and testing datasets. Each type of features generated for the two domains are described in the following and summarized in Tab. 3.

2.2.1. Time domain features

In this section, features extracted from time domain representation of signals are described. This includes a feature that uses the root mean square (RMS) of the time domain signal, a feature that exploits wave dispersion commonly found in SHM, and finally a wave correlation feature.

Root Mean Square feature

The RMS of reconstructed broadband signals initially generated and measured in a pitch and catch configuration where a hole is increasingly grown in the path between the two transducers has shown to increase linearly in (Quaegebeur et al., 2012) for experiments on aluminum plates. Although numerous impact damages will not perforate the CFRP plates as a hole, a dent with increasing diameter will be produced. As the penetration depth increases, generation and reflections of guided wave modes should become more prominent and effects on the measured signal RMS should also be observed. With a significance reported in the signal RMS, the mean, median, standard deviation, variance, minimum, and maximum are also calculated and plotted against the respective impact number. A total of 7 parameters are therefore extracted in this step.

Wave dispersion feature

Wave dispersion commonly occurs in SHM when generating and measuring guided wave bursts. Depending on material properties, when frequencies that compose a generated wave burst do not propagate with the same velocity, the measured wave burst becomes stretched in time. This phenomena depends on material characteristics. Generally most algorithms used in damage imaging are implemented in conditions such that dispersion is avoided. These algorithms depend on the Time of Flight (ToF) of a series of measured wave bursts which is not easily measured accurately in the presence of dispersion. In this work, an attempt to exploit dispersion as a feature is presented. The approach for calculating the dispersion feature presented in Fig. 4 is similar to that of obtaining the Minimum Resolvable Distance (MRD) parameter presented in (Wilcox, Lowe, & Cawley, 2001).

Using the impulse response obtained from post-processing measured signals using the sub-band generation technique (Quaegebeur et al., 2012), output signals for frequencies between 25 kHz - 800 kHz at steps of 25 kHz (32 fre-



Figure 5. Method for extracting the one sample instance of the correlation feature from one range of the frequency domain transfer function.

quencies) are generated using synthetically constructed wave bursts. The maximum amplitude over the entire signal is then selected using a peak detection algorithm. Two peaks before and after the maximum peak are determined with amplitudes that are less that 20 dB of the maximum peak. The purpose of this is to isolate the first reception of the propagated wave bursts. The time difference between the outer peak is then calculated. The time difference is plotted for each frequency against the respective impact number for a total of 32 features.

Wave correlation feature

Extraction of the wave correlation feature set from time domain signals is presented in Fig. 5. Time domain signals are first bandpass filtered according to 5 frequency ranges namely: low frequency (10 kHz - 200 kHz) and below PZT resonance frequency (200 kHz - 400 kHz) constituting the low frequency range (Fig. 5), PZT resonance frequency (400 kHz - 600 kHz), and above PZT resonance frequency (600 kHz - 800 kHz) and high frequency (800 kHz - 1010 kHz) constituting the high frequency range (Fig. 5). The maximum cross-correlation of each frequency range is calculated with respect to the baseline acquired prior to destructive testing. A buffer is used to store the maximum correlation result obtained by cross-correlating all impacts with the baseline for each frequency range. The RMS, mean, median, standard deviation, variance, maximum, and minimum of all maximum correlation values for each frequency range is calculated for a total of 35 features.

2.2.2. Frequency domain features

In this section, features extracted from the frequency domain representation of signals are described. This includes a feature that uses the Power Spectral Density (PSD) of the frequency domain signal and features that exploit the amplitude and phase of the transfer function for guided wave signals propagating in the coupons.

Power Spectral Density feature

The PSD is investigated based on features used in (Larrosa & Chang, 2011). The PSD is calculated from the reconstructed broadband signals initially generated and measured in a pitch and catch configuration. Then, the RMS, mean, median, standard deviation, variance, minimum, and maximum of the PSD are calculated and plotted against the respective impact number. Another 7 parameters are therefore extracted in this step.

Transfer function feature

The transfer function in the frequency domain can be extracted using the sub-band generation technique when used in pitch and catch (Quaegebeur et al., 2012). This provides an amplitude and phase relationship for frequencies below 1 MHz. Based on observation and results reported in (Mulligan, Quaegebeur, Ostiguy, Masson, & Létourneau, 2013), five frequency ranges of the transfer function are isolated for feature extraction based on transducer resonance frequency: low frequency (< 200 kHz), below PZT resonance frequency (200 kHz - 400 kHz), PZT resonance frequency (400 kHz - 600 kHz), above PZT resonance frequency (600 kHz - 800 kHz), and high frequency (> 800 kHz). For each frequency range the RMS, mean, median, standard deviation, variance, minimum, and maximum of the amplitude and phase are calculated and plotted against the respective impact number for a total of 70 parameters.

A summary of the total number of time and frequency domain parameters is shown in Tab. 3. For prognostic model development, a total of 151 parameters are used which are generated solely from guided wave signals.

Domain	Feature	Total
Time	RMS	7
	Wave dispersion	32
	Wave correlation	35
Frequency	PSD	7
	Transfer function	70
	151	

Table 3. Summary of parameters extracted from the raw guided wave signals used in the prognostic model development process.

2.3. Modelling and evaluation

Following data gathering and processing, and feature generation, any supervised learning algorithm can be applied to create prognostic models. Prognostic models are created and evaluated using training and testing datasets respectively. Often an iterative process by which features are added and removed is used to optimize model accuracy and reduce the feature set to remove redundant or irrelevant features. By comparing the results of each model, statistical analysis of model performance is conducted by comparing the predicted index to the actual index. In this case the index corresponds to the remaining number of impacts (RNI). Main statistics including: mean, standard deviation, and mean square error (MSE) of the error between the predicted and actual RNI are also calculated to evaluate the robustness of each model.

With 151 total features, initial reduction of the feature set is performed using the attribute selection tool in the Weka suite of machine learning algorithms to reduce the size of the feature set for model training and testing (Witten & Frank, 2005). The attribute selector requires the addition of a class variable which is labelled as '1' for observations close to the failure point and '0' otherwise. Therefore, 7 impacts closest to the failure point are selected to be labelled as the damage size becomes severe at this point (depth = 0.45-0.55 mm, diameter = 4-7 mm). A wrapper subset attribute evaluator using a 'Naive Bayes' classifier is selected with the search functions: best first, genetic algorithm, greedy stepwise, rank search, random search, and exhaustive search. These algorithms provide a starting point to define a good feature set for model training and testing. Domain knowledge is also used to add and remove features that are neglected or suggested from the attribute selectors. A number of new input feature sets can therefore be defined at this stage.

After the feature sets are assembled into a database, 5 of the 9 problems (plates) are used for model training and the other 4 are used for testing. Then, a Leave One Batch Out (LOBO) cross-validation technique is applied using the best feature set where each problem is chosen individually to test the

model trained by the other problems (8 training, 1 testing) in a round-robin fashion. LOBO is used to assess the variability between the gathered datasets (training and testing) processed by each model and to test model performance when smaller datasets are used (Zaluski et al., 2010). Four generic regression based models are used in this work: sequential minimal optimization (SMO) support vector machine, multiple perceptron artificial neural network (ANN), linear regression (LR), and least mean square (LMS). Based on the results of the attribute selection step, feature subsets can be extracted from the global feature set database and used for training and testing each model. In order to iterate the data-driven model prognostic model development process to obtain accurate prognostic models using feature subsets, an automated system named EBM3 (Environment for Building Models for Machinery Maintenance) is used and presented in (Zaluski et al., 2010). In addition to speeding up execution, the EBM3 system allows the research team to maximize reuse of software components and experimental methodologies between applications (Zaluski et al., 2010). Over 200 experiments in EBM3 are performed using the two split configuration and various feature subsets. The goal of each experiment is to obtain the best prediction results while minimizing the required number of features. In this case, good estimation results are obtained using only 11 features of the original 151. The consistency for model training and testing of the feature subsets yielding the best model is later evaluated using LOBO.

3. PROGNOSTIC MODEL EXPERIMENTAL RESULTS

3.1. Model performance

When selecting and separating training and testing datasets, one assumes that the same model evaluation results can be obtained for any combination. In other words, any combination of problem (plate) pairs used for training and testing datasets should provide the same results. In reality, some datasets are more robust than others in that similar tendencies are found within the data. If data leading up to a problem (plate) event behaved in a similar manner in every case, structural RNI prediction would be easy. Cross-validation provides a sense of the similarities between each dataset and tests model performance when smaller datasets are used. The cross-validation results are provided in remaining number of impacts in Tab. 4 for each regression estimation where N is the number of instances within a 25 impact range including baseline measurements (N_{max} =26).

Model evaluation results outlined in Tab. 4 show the ANN model having the lowest average variability between datasets (designated by low standard deviations). An average error of approximately 6 - 7 impacts is observed with this model and an average MSE of 63 with 25.1 instances within the 26 impact range (including baseline). The SMO, linear regression, and LMS models show larger variability, error, and MSE

Dataset	SMO			ANN		LR			LMS			
	Err±std	MSE	Ν	Err±std	MSE	Ν	Err±std	MSE	Ν	Err±std	MSE	Ν
1	5.4±2.6	35	26	4.6±3.0	29	26	5.4±2.6	35	26	8.0±14.9	86	26
2	5.4±3.3	40	26	6.3±3.6	52	26	4.6±3.2	32	26	5.7±6.3	70	26
3	7.0±4.1	66	26	6.4±3.8	55	26	7.9±4.5	82	26	8.0±6.1	100	26
4	3.2±2.8	18	26	5.6±3.8	45	19	$3.2{\pm}2.2$	15	22	$2.9{\pm}2.6$	15	15
4	3.2±2.8	18	26	5.6±3.8	45	19	$3.2{\pm}2.2$	15	22	$2.9{\pm}2.6$	15	15
5	15.4±35.6	1457	25	11.2±6.8	169	26	9.6±23.2	608	25	41.5±96.1	10611	26
6	20.1±47.2	2543	26	4.8±4.2	40	25	49.7±141.0	1581	26	134.8±393.8	166811	24
7	5.2±3.3	37	26	4.9±3.4	35	26	4.7±3.1	31	26	5.3±3.6	40	25
8	6.5±3.9	57	26	7.4±5.1	81	26	6.5±4.1	58	26	6.6±7.6	130	26
9	6.7±3.6	57	26	6.9±4.1	64	26	6.6±4.2	60	25	6.6±5.8	76	23
Avg.	8.3±11.8	479	25.9	6.5±4.2	63	25.1	10.9±9.8	278	25.3	24.4±59.6	19771	24.1

Table 4. Model evaluation for RNI estimates on test data for the SMO, ANN, LR, and LMS regression algorithms using LOBO.

overall. In all cases where large MSE is found, the final two observations closest to the failure point have large estimation errors. This indicates that the SMO, LR, and LMS models may perform optimally at estimating the RNI farther away from the failure point. In a next step, this may be useful when performing data clustering (Chen, Han, & Yu, 1996).

3.2. Model results

In order to associate a number of impacts to a critical damage size (failure threshold), the non-destructive phased array ultrasonic testing tool is used following each destructive impact to measure damage diameter versus impact number. Fig. 6 presents an experimental damage size versus number of impacts curve obtained by taking an average of all nondestructive phase array ultrasonic measurements for the training and testing coupons. A failure threshold is defined at impact number 25 corresponding to a damage diameter of $\simeq 7.1$ mm. In a real application, this threshold point is defined from analytical and experimental tests on the material. Model estimates using the testing dataset in remaining number of impacts are associated with a corresponding damage size based on the experimental damage size measurements for each testing plate and plotted in Fig. 6 with the average experimental damage size versus impact number measurements. From the figure, for an impact number lower than 5, the linear regression through the damage size estimation points suggests that damage exists prior to detection by the ultrasonic testing tool. Overall however, both the experimental damage size measurements and damage size estimation points increase linearly towards the failure threshold as the impact number increases. The damage size estimation model however is underestimating damage size shown by the slope of the regression fit to the damage size estimation points which increases at a slower rate than that of the experimental damage size measurements. This is caused by estimation errors between 15 and 20 impacts as features extracted from the guided wave signals must suffer from a reduced sensitivity in this range.

The model results for estimated RNI plotted against actual RNI are presented in Fig. 7. The ANN algorithm is evaluated using the testing dataset composed of a subset of 11 features from the original 151 including: time domain signal mean, low PZT frequency range mean, low PZT resonance frequency range median, PZT resonance frequency range median and standard deviation, high PZT resonance frequency range maximum, high PZT resonance frequency phase mean and variance, variance of the PSD, and the low frequency range correlation variance. The model is capable of estimating RNI within a 30 impact range for all instances in the training and testing datasets. In Fig. 7 the slope of the linear regression fit for points within a 25 impact window is 0.51 with an intercept of 9.6. A slope close to unity with an intercept at 0 in training would represent an ideal model. If the model is then tested using a testing dataset demonstrating similar tendencies found in the training dataset, the prediction should match the actual parameter of interest (in this case RNIs) for all instances. In reality however, it is difficult to create testing datasets without variation from the training dataset in the data-driven approach. The slope of the linear regression fitted to the model results using the testing dataset is 5% lower than in the training case. This means that the model is over estimating the RNIs and a failure is detected by the model sooner than it actually occurs. In a real application, this would increase costs as maintenance procedures would be initiated before being necessary. In other words, structural components would be over maintained (Zaluski et al., 2010; Létourneau et al., 2005). Looking at the points close to the failure point, the model does however succeed at estimating the RNIs within range. This indicates that with more problems (plates), a larger number of points could converge in this area. This would potentially increase the slope of the linear regression fit while reducing the intercept across the ordinate. Performance assessment based on model error indicates a RNI estimation error for training of 3.7±3.2 and 8.4 ± 6.7 for testing and a MSE value of 24 for training and of 114 for testing.

Combining previous results leads to a more useful representation of RNI versus damage diameter, presented in Fig. 8. RUL in remaining number of impacts (RNI) is obtained by subtracting each estimated diameter for each impact number shown in Fig. 6 from the failure threshold ($\simeq 7.1$ mm) and



Figure 6. Estimated damage diameter versus impact number compared to experimental measurements made using the nondestructive phased array ultrasonic testing tool following each destructive impact.



Figure 7. Estimated RNI versus actual RNI using results from the ANN model evaluated using the testing dataset.

plotting RNI (as units for RUL) versus damage diameter. To summarize, in a real implementation of the predictive model:

- 1. Each observation containing a feature set obtained using guided wave data is evaluated by the model.
- 2. The result (in RNIs) is associated to a corresponding damage diameter using an average of the experimental damage size measurements gathered after each impact for each plate.
- 3. Using a pre-defined failure threshold, the estimated damage size diameter is subtracted and RNI is plotted versus damage size.
- 4. By constructing a plot of RNIs versus damage diameter, RUL can be predicted by tracing a linear regression through the points. Note that in order to plot a linear regression, at least two points are required. As the number of points increases, the prediction accuracy also increases.

4. DISCUSSION

To summarize, using a data-driven approach for estimating RNI following the initiation of impact damage in CFRP aerospace structures, desirable model performance is



Figure 8. Predicted RUL in remaining number of impacts (RNI) versus damage diameter compared to experimental measurements made using the non-destructive phased array ultrasonic testing tool following each impact.

achieved based on operational data gathered in a laboratory environment. Key considerations are presented in this paper for defining a robust and realistic laboratory data acquisition and damage initiation methodology. This is an important foundation to build upon such that new features relevant to damage prognosis can be explored before higher scale attempts are made on real aircraft. Evidently guided wave measurements using SHM technologies while used in-service are subject to variations induced by temperature (Konstantinidis, Wilcox, & Drinkwater, 2007; Croxford, Moll, Wilcox, & Michael, 2010), hydrostatic pressure (Croxford et al., 2010), impact damages themselves among other factors. In fact, damage has been shown in (Mulligan, Ostiguy, Masson, Elkoun, & Quaegebeur, 2011; Mulligan, Masson, Létourneau, & Quaegebeur, 2011; Mulligan, Quaegebeur, Masson, & Létourneau, 2012; Mulligan, Quaegebeur, Masson, Brault, & Yang, 2013) to significantly influence guided wave signals by degradation to the bonding layer between the transducer and the structural surface. Compensation strategies for temperature have been proposed in (Croxford et al., 2010). A signal correction factor (SCF) has been proposed in (Mulligan, Quaegebeur, Masson, et al., 2013) to compensate for bonding layer degradation based on transducer electrical admittance measurements demonstrated in (Park, Park, Yun, & Farrar, 2009) as a potential bonding layer degradation assessment metric. With a robust foundation for a data gathering methodology by which data are acquired with consistent damages that mimic potential in-service damages, investigations into these variations are now possible by building on the current methodology.

Four types of features including time and frequency domain,

wave dispersion and correlation features are presented in this work. Although the best results used only a combination of time domain, frequency domain, and correlation features, those results including wave dispersion features were not far off with slopes and intercepts of the predicted versus actual RNI within 5% while correctly estimating remaining impact numbers within range for most instances. Variations between the estimation, prediction, and measurement results are likely related to variability between the data gathered from the testing plates which is indicated in the cross-validation. These uncertainties in part can be attributed to bonding layer degradation of the transducers due to the initiation of impacts. Damage initiation using drop-weight impacts has been shown to induce transducer bonding layer degradation (Mulligan, Quaegebeur, Ostiguy, et al., 2013; Mulligan, Quaegebeur, Masson, et al., 2013). Bonding layer degradation in any amount influences the amplitude and phase of guided wave bursts measured in a pitch and catch configuration. Changes to the amplitude and phase of a guided wave burst can drastically influence wave amplitude, phase, dispersion and correlation. Variability is introduced through comparison of these features with baseline measurements made prior to damage initiation which are not influenced by impact damage.

Another cause for low model performance could be that data clustering is not used in this work. Clustering is commonly used in data-driven damage prognosis where instances of the feature set can be grouped into similar categories. In this work, based on the loading pattern, instances could be separated into 4 sections based on damage severity. Clustering has demonstrated that some features are more sensitive in some categories more than others (Chen et al., 1996). Perhaps, the wave dispersion features could perform better in assessing certain severities of damage than others. Furthermore, model evaluation has shown that some models may be more appropriate for estimating RUL further from the critical damage size. Using clustering, these models could only be applied to such observations.

With these considerations in mind, a number of questions remain to be answered in this domain. However, promising results demonstrating that impact damage is predictable are shown in this work. To overcome the aleatory uncertainty in the training process, the addition of more and different problems, features, clustering, and data correction techniques, is required.

5. CONCLUSION

In this work, a novel data-driven approach for estimating RNI and damage size for CFRP aerospace grade coupons exposed to drop-weight impact damage is presented. The approach relies on a PZT based SHM system instrumented onto the CFRP coupons in a pitch and catch configuration for data acquisition. Time and frequency domain data are acquired for training and testing several generic regression based machine learning models on several CFRP coupons over a specified loading pattern. The laboratory data is gathered using a methodology designed as a foundation to build upon for simulating real world operating conditions. The purpose of this first step is to evaluate the feasibility of assessing structural damage using damage prognosis.

In future work, signal correction presented in (Mulligan, Quaegebeur, Masson, et al., 2013) will be applied to the data gathered data to compensate for bonding layer degradation induced by the impact damage. This will assure that only structural damage is being assessed not to include damage to the PZT SHM system. The acquisition of data for new problems will reduce the variability observed in the cross-validation step of this work while hopefully improve results of the predicted versus actual RUL estimating resulting adding confidence for eventual testing on in-service aircraft.

The ultimate goal of building upon this work is to determine if there exists a feature set capable of revealing damage tendencies prior to detection using a non-destructive phased array ultrasonic testing system and to determine if data mining and machine learning combined with SHM provide equivalent or more accurate measures of damage compared to ultrasonic systems.

ACKNOWLEDGEMENTS

This study has been conducted with the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the National Research Council of Canada (NRC). The authors would like to thank Sylvain Létourneau and Marvin Zaluski from the NRC for providing guidance in this project.

REFERENCES

- ASTM. (2007). Standard test method for measuring the damage resistance of a fibre-reinforced polymer matrix composite to a drop-weight impact event (Tech. Rep. Nos. D7136/D7136M-07). American Society for Testing and Materials.
- Beaumont, P. W. R., Diamant, R. A., & Shercliff, H. R. (2006). Failure processes in composite materials: getting physical. *Journal of Material Science*, 41(20), 6526–6546.
- Chang, F., & Chang, K. (1987). A progressive damage model for laminated composites containing stress concentrations. *Journal of Composite Materials*, 21(9), 834– 855.
- Chattopadhyay, A., Papandreou-Suppapola, A., Kim, B. K., Kovvali, N., Farrar, C. R., Triplett, M. H., & Derriso, M. M. (2012). Workshop on transitioning structural health monitoring technology to military platforms (Tech. Rep.). DTIC.
- Chen, M.-S., Han, J., & Yu, P. S. (1996). Data mining: an overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 866–883.
- Chiachio, J., Chiachio, M., Saxena, A., Rus, G., & Goebel, K. (2013). Based prognostic framework to predict fatigue damage evolution in composites. In *Proceedings of the annueal conference of the prognostics and health management society, new orleans, la, usa, 14-17 october.*
- Chiachio, M., Chiachio, J., Saxena, A., Rus, G., & Goebel, K. (2013). Fatigue damage prognosis in frp composites by combining multi- scale degradation fault modes in an uncertainty Bayesian framework. In *Proceedings* of the 9th international workshop on structural health monitoring, stanford, ca, usa, 10-12 september.
- Choi, H. (1990). *Damage in graphite/epoxy laminated composites due to low-velocity impact* (Unpublished doctoral dissertation). Stanford University.
- Croxford, A. J., Moll, J., Wilcox, P. D., & Michael, J. E. (2010). Efficient temperature compensation strategies for guided wave structural health monitoring. *Ultrasonics*, *50*(4-5), 517–528.
- Giurgiutiu, V., & Bao, J. J. (2004). Embedded-ultrasonics structural radar for in situ structural health monitoring of thin-wall structures. *Structural Health Monitoring*, 3(2), 121–140.
- Guida, M., Marulo, F., Meo, M., & Russo, S. (2012). Certification by birdstrike analysis on c27j fullscale ribless composite leading edge,. International Journal of Impact Engineering, doi:

10.1016/j.ijimpeng.2012.10.002.

- Hall, M. (2000). Correlation-based feature selection for discrete and numeric class machine learning. In *Proceedings of the 17th international conference on machine learning*.
- Hashin, Z. (1980). Failure criteria for unidirectional fiber composites. *Journal of Applied Mechanics*, 32(1), 1921–1944.
- Hou, J. P., Petrinic, N., Ruiz, C., & Hallett, S. R. (2000). Prediction of impact damage in composite plates. *Composites Science and Technology*, 60(2), 228–273.
- Kira, K., & Rendell, L. (1992). A practical approach to feature selection. In *Proceedings of the 9th international conference on machine learning*.
- Konstantinidis, G., Wilcox, P. D., & Drinkwater, B. W. (2007). An investigation into the temperature stability of a guided wave structural health monitoring system using permanently attached sensors. *IEEE Sensors Journal*, 7(5), 905–912.
- Larrosa, C., & Chang, F.-K. (2011). Characterization of matrix micro-cracking in composite laminates using builtin piezo-electric sensors. In *Annual conference of the prognostics and health management society, montreal*, *qc, canada, 25-29 september.*
- Létourneau, S., Yang, C., Drummond, C., Scarlett, E., Valdés, J., & Zaluski, M. (2005). A domain independent data mining methodology for prognostics. In Essential technologies for successful prognostics: proceedings of the 59th meeting of the society for machinery failure prevention technology, virginia beach, virginia, 18-21 april.
- Liu, Y., Mohanty, S., & Chattopadhyay, A. (2009). A gaussian process based prognostics framework for composite structures. In *Proceedings of spie, san diego, ca, usa, 8-12 march.*
- Mulligan, K. R., Masson, P., Létourneau, S., & Quaegebeur, N. (2011). An approach to compensate for the degradation of the monitoring system in damage detection. In *Proceedings of the Canadian Institute for NDE*.
- Mulligan, K. R., Ostiguy, P.-C., Masson, P., Elkoun, S., & Quaegebeur, N. (2011). Assessment of PZT transducer bonding techniques under drop-weight impact loading in composites. In *Proceedings of spie*.
- Mulligan, K. R., Quaegebeur, N., Masson, P., Brault, L.-P., & Yang, C. (2013). Compensation of piezoceramic bonding layer degradation for structural health monitoring. *Structural Health Monitoring, doi:* 10.1177/1475921713500516.
- Mulligan, K. R., Quaegebeur, N., Masson, P., & Létourneau, S. (2012). Correction of data gathered by degraded transducers for damage prognosis in composite structures. In Annual conference of the prognostics health management society.
- Mulligan, K. R., Quaegebeur, N., Ostiguy, P.-C., Masson, P.,

& Létourneau, S. (2013). Comparison of metrics to monitor and compensate for piezoceramic degradation in structural health monitoring. *Structural Health Monitoring*, *12*(2), 153–168.

- Nairn, J. A. (Ed.). (2000). *Matrix microcracking in composites*. Elsevier.
- Nguyen, M. Q., Jacombs, S. S., Thomson, R. S., Hachenberg, D., & Scott, M. L. (2005). Simulation of impact on sandwich structures. *Composite Structures*, 67, 217– 227.
- O'Brien, T. K. (Ed.). (2001). Fracture mechanics of composite delamination - asm handbook - composites. ASM International.
- Olympus. (2013). *Ultrasonic flaw detectors-omniscan mx2* (Tech. Rep.). Author.
- Ostiguy, P.-C., Quaegebeur, N., Mulligan, K. R., Masson, P., & Elkoun, S. (2012). In-situ mechanical characterization of isotropic structures using guided wave propagation. *Smart Materials and Structures*, 21(1), 1-9.
- Park, S., Park, G., Yun, C.-B., & Farrar, C. R. (2009). Sensor self-diagnosis using a modified impedance model for active sensing-based structural health monitoring. *Structural Health Monitoring*, 8(1), 71–82.
- Quaegebeur, N., Masson, P., Langlois-Demers, D., & Micheau, P. (2010). Dispersion-based imaging for structural health monitoring using sparse and compact arrays. *Smart Materials and Structures*, 20(1), 1–12.
- Quaegebeur, N., Masson, P., Micheau, P., & Mrad, N. (2012). Broadband generation of ultrasonic guided waves using sub-band decomposition. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 59*(5), 928–938.
- Rhymer, J., Kim, H., & Roach, D. (2012). The damage resistance of quasi-isotropic carbon-epoxy composite tape laminates impacted by high velocity ice. *Composites: Part A*, 43(7), 1134–1144.
- Roebuck, B., Gorley, T. A. E., & McCartney, L. N. (1989). Mechanical property test procedures for metal matrix composites. *Materials Science and Technology*, 5(2), 105–117.
- Roemer, M. J., Ge, J., & Liberson, A. (2005). Autonomous impact damage detection and isolation prediction for aerospace structures. In *Proceedings of ieee aerospace conference, big sky, mt, usa, 5-12 march.*
- Saxena, A., Goebel, K., Larrosa, C. C., Janapati, V., & Roy, S. (2011). Accelerated aging experiments for prognostics of damage growth in composite materials. In *Proceedings of the 8th internationa workshop on structural health monitoring, stanford, ca, usa, 13-15 sept.*
- Soutis, C. (2005). Fibre reinforced composites in aircraft construction. *Progress in Aerospace Sciences*, 41(2), 143–151.
- Staszewski, W. J., Mahzan, S., & Traynor, R. (2009). Health monitoring of aerospace composite structures - active

and passive approach. *Composites Science and Technology*, 69(11–12), 1678–1685.

- Tomblin, J., Lacy, T., Smith, B., Hooper, S., Vizzini, A., & Lee, S. (August 1999). *Review of damage tolerance for composite sandwich airframe structures* (Tech. Rep. No. DOT/FAA/AR-99/49). Federal Aviation Administration, Office of Aviation Research, Washington, DC.
- Wilcox, P., Lowe, M., & Cawley, P. (2001). The effect if dispersion on long-range inspection using ultrasonic guided waves. NDT&E international, 34(1), 1–9.
- Witten, I. H., & Frank, E. (Eds.). (2005). *Data mining: Practical machine learning tools and techniques, 2nd edition*. Morgan Kaufmann, San Francisco.
- Zaluski, M., Létourneau, S., Bird, J., & Yang, C. (2010). Developing data mining-based prognostic models for cf-18 aircraft. In *Proceedings of asme turbo expo 2010: Power for land, sea and air, glascow, uk, 14-18 june.*

BIOGRAPHIES

Kyle R Mulligan obtained his Bachelor's degree in Systems and Computer Engineering at Carleton University in Ottawa in 2007. He then continued to obtain his Master's degree in Biomedical Engineering in 2009 at Carleton University. Currently he is pursuing a Doctorate degree in Mechanical Engineering at the Université de Sherbrooke. He is an active volunteer at local high schools to promote the awareness and importance of Engineering in today's society. His research interests include: medical imaging techniques, biomechanics, nanorobotics, and structural/prognostic health management.

Chunsheng Yang is a Senior Research Officer at the National Research Council of Canada. He is interested in data mining, reasoning technologies such as case-based reasoning, rule-based reasoning and hybrid reasoning, multi-agent systems, and distributed computing. He received an Hons. B.Sc. in Electronic Engineering from Harbin Engineering University, China, an M.Sc. in computer science from Shanghai Jiao Tong University, China, and a Ph.D. from National Hiroshima University, Japan. He worked with Fujitsu Inc., Japan, as a Senior Engineer and engaged on the development of ATM Network Management Systems. He was an Assistant Professor at Shanghai Jiao Tong University from 1986 to 1990 working on Hypercube Distributed Computer Systems. He was a Program Co-Chair for the 17th International Conference on Industry and Engineering Applications of Artificial Intelligence and Expert Systems. Dr. Yang is a guest editor for the International Journal of Applied Intelligence. He has served Program Committees for many conferences and institutions, and has been a reviewer for many conferences, journals, and organizations, including Applied Intelligence, NSERC, IEEE Trans., ACM KDD, PAKDD, AA-MAS, IEA/AIE. Dr. Yang is a senior IEEE member.

Nicolas Quaegebeurobtained his bachelor's degree in engineering physics from Ecole Nationale de Techniques avancées (EnsTa ParisTech, France) and his master's degree from ATIAM (IRCAM, University of Paris and Telecom ParisTech) with specialization in acoustics, signal processing, and informatics applied to music in 2004. He obtained his Ph.D. degree on nonlinear vibrations and acoustic radiation of electrodynamic transducers from Ecole Polytechnique (Paris-Tech, France) in 2007. He joined the GAUS group at the Université de Sherbrooke for a postdoctoral fellowship on active vibration control in 2008 and on structural health monitoring using guided waves in 2009, before becoming a research assistant in the same group in 2011. His research interests include nonlinear vibrations, advanced signal processing, active vibration control, piezoelectric transducers, guided wave propagation modelling, and structural health monitoring.

Patrice Masson received his bachelor's degree in engineering physics and his master's degree in mechanical engineering from École Polytechnique de Montréal, Canada, in 1989 and 1991, respectively. He obtained his Ph.D. degree in mechanical engineering from the Université de Sherbrooke, Canada, in 1997. In 2000, he became a faculty member in the Mechanical Engineering department of the Université de Sherbrooke, where he is now the director of the Acoustics and Vibration Group of the Université de Sherbrooke (GaUs). He has contributed to establish the new curriculum and is teaching the mechatronics courses. Prof. Masson leads or co-leads multi-institutional projects within AUTO21 and the CRIAO. His research interests include structural health monitoring, signal processing, smart materials and structures, and active noise and vibration control. Prof. Masson is a member of the acoustical society of america (ASA), the society of automotive Engineers (SAE), the International society for optical Engineering (SPIE) and the ordre des Ingénieurs du Québec (OIG).