An SHM Data-Driven Methodology for the Remaining Useful Life Prognosis of Aeronautical Subcomponents

Georgios Galanopoulos[1(✉)], Nick Eleftheroglou[1], Dimitrios Milanoski[1], Agnes Broer[2,3], Dimitrios Zarouchas[2,3], and Theodoros Loutas[1]

[1] Applied Mechanics Laboratory, Department of Mechanical Engineering and Aeronautics, University of Patras, 26504 Rio, Greece
gkgalanopoulos@gmail.com

[2] Structural Integrity and Composites Group, Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629HS Delft, The Netherlands

[3] Center of Excellence in Artificial Intelligence for structures, Aerospace Engineering Faculty, Delft University of Technology, Delft, The Netherlands

Abstract. Prognosis of the Remaining Useful Life (RUL) of a structure from Structural Health Monitoring data is the ultimate level in the SHM hierarchy. Reliable prognostics are key to a Condition Based Maintenance paradigm for aerospace systems and structures. In the present work, we propose a methodology for RUL prognosis of generic aeronautical elements i.e. single stringered composite panels subjected to compression/compression fatigue. Strain measurements are utilized in this direction via FBG sensors bonded to the stiffener feet. The strain data collected during the fatigue life are processed and used for the RUL prognosis. In order to accomplish this task, it is essential to produce Health Indicators (HIs) out of raw strain that can properly capture the degradation process. To create such HIs a new pre/post-processing technique is employed and a variety of different HIs are developed. The quality of the HIs can enhance the performance of the prognostic algorithms, hence a fusion methodology is proposed using genetic algorithms. The resulted fused HI is used for the RUL estimation of the SSCPs. Gaussian processes and Hidden Semi Markov Models are employed for RUL prognosis and their performance is compared. Despite the complexity the raw data we demonstrate the feasibility of successful RUL prognostics in a SHM-data driven approach.

Keywords: Structural Health Monitoring · RUL prognosis · Composite panels · Health Indicators · Fibber Bragg Gratings

1 Introduction

With the increasing use of composite materials in various safety critical industries, like automotive and aerospace, it is imperative to accurately monitor their

P. Rizzo and A. Milazzo (Eds.): EWSHM 2022, LNCE 253, pp. 244–253, 2023.
https://doi.org/10.1007/978-3-031-07254-3_24

structural behavior. Structural Health Monitoring (SHM) systems can be utilized for the consummation of that task. Intelligent diagnostics and prognostics using SHM is considered by some one of the most demanding task to achieve in a condition based maintenance scheme [3,13]. Of critical importance is Remaining useful life (RUL) prediction which is closely linked with SHM, since accurate SHM measurements are crucial for knowledge of the structure's degradation [9,18]. Prognostic methodologies are roughly classified into two major categories, model-based and data-driven [8,20]. Data-driven, which use stochastic modeling and machine learning (ML) methodologies, such as Neural Networks (NN), Gaussian process etc., to predict the End of Life (EoL) given historical data are more commonly used for complex structures, due to the difficulty of accurately modeling the degradation using physical equations.

There have been a number of studies employing data-driven methods for RUL prediction on composite structures. Eleftheroglou et al. [3] proposed a nonhomogeneous hidden semi Markov model (NHHSMM) for RUL prediction of openhole composite coupons subjected to constant amplitude tension-tension fatigue. Acoustic emission (AE) data were used as the input to train the NHHSMM. The predicted RUL displayed great results showing the potential of the framework for integration in different SHM datasets. An adaptive NHHSMM was developed in [4]. The ANHHSMM was able to adapt to unforeseen events such as midtest impacts, even though the training data did not contain such events. The ANHHSMM greatly outperformed the regular NHHSMM in such cases, demonstrating its capabilities. Wei et al. [22] also employed Markov chain models for fatigue life prediction of open-hole coupons. Infrared thermoelastic analysis and strain readings were used to train the model, which predicted the S-N curves with variability at a constant fatigue load. Rabiei et al. [16] proposed a dynamic Bayesian network (DBNN) for the RUL prediction of glass/epoxy specimens subjected to bending fatigue. Indirect damage measurements were used to estimate the damage state and train the DBNN, which predicted the damage state k steps ahead. Liu et al. [10] used Gaussian processes for non-linear regression to predict the RUL of composite specimens subjected to uni-axial and bi-axial fatigue. Real time strain gauge data were collected and used for the prediction. It was observed that the prediction results were more accurate the later the startpoint of the prediction.

An important aspect of data-driven prognostics are the degradation features. Such features are usually referred to as Health Indicators (HIs). Their quality affects the performance and accuracy of the prognostics methodologies [12]. The three main attributes that determine the quality of an HI are monotonicity, trendability and prognosability as stated by [1]. The higher these attributes the better the HI. HIs are divided into two categories [23], physical HI (pHIs), that results from direct measurements [5,7,14] and virtual HIs (vHIs), that are created using more sophisticated data processing on the direct measurements [5,11,19].

In this paper the HIs developed in [5] are fused together using Genetic Algorithms to create a HI with higher monotonicity and prognosability. The fused HI will be used by two machine learning models, i.e. Gaussian Process Regression and Non Homogeneous Hidden Semi Markov Model, to predict the RUL of single stringered composite panels.

2 Prognostic Models

2.1 Gaussian Process Regression

Gaussian Process Regression (GPR) is a probabilistic method for non-linear regression that estimates the posterior distribution by constraining the prior distribution to fit the training data. A GP is a collection of random variables with a joint Gaussian distribution, and are a function of $f(x)$ at $x = [x_1, x_2, \ldots, x_n]^T$. GP is completely specified [24] by its mean (Eq. 1) and covariance function (Eq. 2):

$$m(x) = E[f(x)] \tag{1}$$

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x'))] \tag{2}$$

Then the GP can be written as:

$$f(x) \sim GP(m(x), k(x, x')) \tag{3}$$

The mean function $m(x)$ is usually set to be zero. As it is noted in [24] different covariance functions yield different regression results, so this function should be considered carefully depending on the data. Assume a degradation history $H = [x_i, y_i]_{i=1}^{N}$, where x_i the input variables and $y_i = f(x_i) + \epsilon_i$ the noisy target variables, with ϵ_i is an i.i.d with 0 mean and σ_n^2 (ϵ_i $i.i.dN(0, \sigma_n^2)$). The joint distribution of observed target values $y = [y]_{i=1}^{N}$ and unobserved target values f^* at new input locations X^* can be denoted as:

$$\begin{bmatrix} y \\ f^* \end{bmatrix} \sim N(0, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K(X, X^*) \\ K(X^*, X) & K(X^*, X^*) \end{bmatrix}) \tag{4}$$

where I the identity matrix. The predictive (posterior) distribution for GPR, given the new inputs X^* and the historic input data X and targets y is defined by:

$$p(f^*|X, y, X^*) \sim N((\bar{f}^*), cov(f^*)) \tag{5}$$

$$\bar{f}^* = E[f^*|X, y, X^*] = K(X^*, X)[K(X, X) + \sigma_n^2 I]^{-1} y, \tag{6}$$

$$cov(f^*) = K(X^*, X^*) - K(X^*, X)[K(X, X) + \sigma_n^2 I]^{-1} K(X, X^*) \tag{7}$$

2.2 Non-Homogeneous Hidden Semi Markov Model (NHHSMM)

NHHSMM is a mathematical model that describes the association between a hidden stochastic degradation process, i.e. damage accumulation in composite materials, and an observed one which manifests via SHM data. The NHHSMM is actually a double stochastic process, where the hidden process is a finite Semi Markov chain and the observed process, conditioned on the hidden one.

To properly describe the bi-dimensional stochastic process, the model's parameters $\theta = \{\Gamma, B\}$ need to be estimated. These parameters characterize the degradation process (Γ) of the studied system via transition rate distributions between the hidden states, and the observation process (B) via an emission matrix that correlates hidden states and SHM data. The studied system, is assumed to start operation from a healthy state and during its life transits to states of higher degradation until it reaches its failure state.

The model's parameters θ are obtained via a Maximum Likelihood Estimator (MLE) θ^* of the model parameters θ through a procedure described in detail in [15]. The MLE algorithm leads to the maximization likelihood function $L(\theta, y(1:K))$, where $y(k)$ is the k-th degradation history, K is the number of available degradation histories and

$$L(\theta, y^{(1:K)}) = \prod_{k=1}^{K} Pr(y^{(k)})|\theta \Rightarrow \theta^* = \arg\max_{\theta}(\sum_{k=1}^{K} log(Pr(y^{(k)}|\theta)) \tag{8}$$

Setting initial values for Γ, B and solving the aforementioned optimization problem the parameter estimation process is obtained and prognostic-related measures can be defined and calculated. Regarding prognostics, the conditional reliability function, $R(t|y_{1:t_p}, L > t_p, M) = Pr(L > t|y_{1:t_p}, L > t_p, M)$, represents the probability that the composite material continues to operate after a time t, less than its lifetime L, further that the current time t_p given the SHM data $y_{1:t_p}$. In this study the mean and 95% confidence intervals of RUL are proposed as a prognostics measure. These measures were calculated via the cumulative distribution function (CDF) of RUL. The CDF of RUL is defined at any time point via the conditional reliability according to:

$$Pr(RUL_{t_p} \leq t|y_{1:t_p}, M) = 1 - R(t + t_p|y_{1:t_p}, M) \tag{9}$$

3 Case Study

3.1 Specimen Definition

Single stringered composite panels (SSCPs) were manufactured from IM7-8552 UD prepreg with $[45/-45/0/45/90/-45/0]_s$ for the skin and $[45/-45/0/45/-45]_s$ for the single T-shaped stringer. The total length of the panels is 300 mm, though only 240 mm are free, since 30 mm resin tabs were placed on the free edges to ensure uniform and proper load introduction [2,5]. FBG strain sensors were encased in two SMARTape^{TM} (provided by Smartec) [6] which were bonded at the stiffener's feet. A total of 10 FBGs were available (5 on each fiber), with a spacing of 20 mm, and were focused on measuring the strains at the middle section of the stiffeners' feet for an approximate area of 140 mm.

3.2 Test Definition

After determining the collapse load of the panels, which was on average 100 kN, two test campaigns were performed, with different loading scenarios. For the first, the SSCPs were subjected to constant amplitude compression-compression (C-C) fatigue. The second loading scenario involved variable amplitude C-C fatigue test, i.e. the load was applied in constant blocks and was arbitrarily increased after inspecting the extent of the damage using a phased array camera. An initial damage, either barely visible impact, or an artificial disbond, was introduced to the SSCPs before subjecting them to fatigue. A loading ratio R = 10 and a frequency f = 2 Hz were used during both test campaigns. Every 500 cycles the fatigue test was paused and quasistatic (QS) loadings were performed, during which the acquisition of the strains was made. In Tables 1 and 2 the detailed load sequences and cycles to failure are summarized.

Table 1. Constant amplitude fatigue coupon information

Spec label	Damage type	Max load	Cycles
CA-01	Impact 10 J	−65 kN	280,098
CA-02	Impact 10 J	−65 kN	144,969
CA-03	Impact 10 J	−65 kN	133,281
CA-04	Disbond 30 mm	−50 kN	100,000
		−65 kN	338,000
			438,000

3.3 Health Indicator Fusion

To improve the degradation features presented in [5] genetic algorithms were employed to combine these HIs and create an enhanced HI with improved monotonicity and prognosability [1]. The goal was to maximize the objective function:

$$F = Monotonicity + Prognosability \tag{10}$$

where Monotonicity and Prognosability are defined by Eq. (11) and Eq. (13) respectively:

$$Monotonicity = \frac{1}{N}\sum_{i=1}^{N} M_i \tag{11}$$

where,

$$M_i = \frac{(n_i^+)}{(n_i - 1)} + \frac{(n_i^-)}{(n_i - 1)}, i = 1, ..., N \tag{12}$$

Table 2. Variable amplitude fatigue coupon information

Spec label	Damage type	Max load	Cycles
VA-01	Impact 10 J	−40 kN	10,000
		−45 kN	80,000
		−50 kN	30,000
		−55 kN	70,000
		−60 kN	12,300
			202,300
VA-02	Impact 10 J	−40 kN	10,000
		−45 kN	80,000
		−50 kN	90,000
		−55 kN	63,000
			243,000
VA-03	Impact 10 J	−40 kN	10,000
		−45 kN	177,000
		−50 kN	30,000
			217,000
VA-05	Disbond 20 mm	−35 kN	10,000
		−39 kN	10,000
		−45 kN	10,000
		−50 kN	170,000
		−55 kN	85,000
		−60 kN	60,000
			354,000
VA-05	Impact 10 J	−40 kN	20,000
		−45 kN	75,000
		−50 kN	25,000
		−55 kN	62,000
		−60 kN	60,000
			242,000

and

$$Prognosability = exp(\frac{-std(HI_{fail})}{mean(|HI_{start} - HI_{fail}|)}) \tag{13}$$

Genetic algorithms (GAs) were selected for the optimization of the objective function, using the GPLAB toolbox [21] in Matlab. Three main GA parameters were investigated, concerning the selection of the population to create the next generation, the selection of individuals to produce children for the next generation and the survival of the current individuals to fill the population of the next generation. In total 27 parameters combinations were tested. A population of 150 and a generation (iteration) limit of 300 were arbitrarily selected after trial and error.

What ultimately guided our selection for the final fusion, was the fitness value, the simplicity and inputs of the fusion function. After selecting a model, the same GA optimization was run 50 more times to evaluate the repeatability. Though it was never managed to reproduce the same exact fusion function, similar functions with high monotonicity and prognosability were achieved. This was an anticipated result due to nature of the optimization problem and algorithm since there is no global maximum to be reached and it depends on the functions happening between the subsequent generations. For the training of the GA 3 specimens from each test campaign were used, leaving 3 specimens out to use for testing the applicability of the methodology. The resulted fusion function is shown in Eq. (14).

$$HI_{GA} = vHI_1(HI_4 - \frac{vHI_2 + 0.5HI_3}{vHI_2}) + 1 \tag{14}$$

where vHI_1, vHI_2, HI_3 and HI_4 are HIs developed in [5]. The resulted HI_{GA} is shown in Fig. 1. The average monotonicity and prognosability are 0.81 and 0.94 respectively.

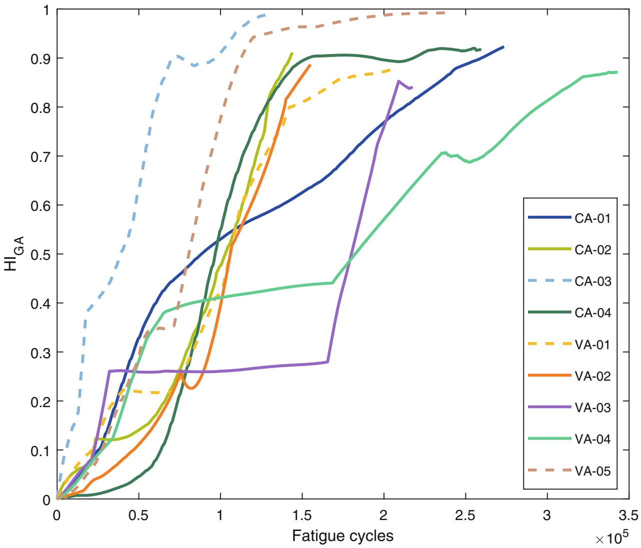
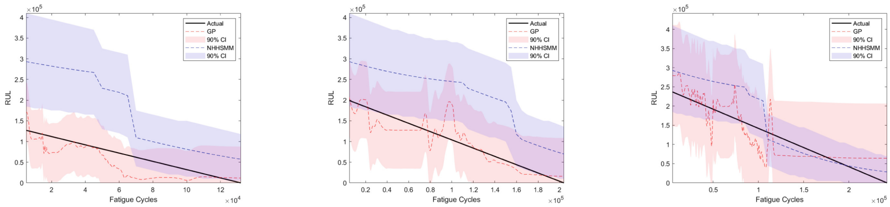


Fig. 1. HI_{GA} for all the composite panels. Solid lines represent specimens used for training, while dashed those used for testing

3.4 Remaining Useful Life Prognosis

For the GPR predictions a similarity based scheme was applied and only the 4 most similar specimens were used for the training of the GPR. The similarity was measured at the first 10000 cycles. The predicted RUL and the corresponding

90% prediction intervals are displayed in Fig. 2. Predictions for CA-03 using GPR is at first overestimating the RUL, though it remains always within the prediction intervals. As time progresses, the prediction is closing in to the true RUL. The NHHSMM predictions at first overestimates the RUL and near the EOL the prediction intervals start to include the true RUL. GPR for VA-01 is constantly close to the true RUL giving very good mean predictions, while the NHHSMM is constantly overestimating the true RUL. VA-05's predictions using GPR is at first close to the true RUL, however near the EoL the predicted RUL abruptly increases before it slowly decreases again. This is not an ideal behavior for a prediction since it increases the uncertainty near the end which is the opposite of what's desired. NHHSMM predictions follow a similar trend to those of the GPR's, following the overall trend of the true RUL. However compared to the GPR, the confidence intervals of the NHHSMM's are slightly narrower and hence provide less uncertainty in the predictions.



(a) CA-03 (b) VA-01 (c) VA-05

Fig. 2. RUL predictions for the three test specimens

RMSE and MAPE, two popular performance prognostic metrics are used to evaluate the predictions [17]. Table 3 summarizes these metrics for the three test set specimens. The overall best performance for both algorithms is achieved for VA-05 where NHHSMM provides slightly better predictions, while for CA-03 predictions are better using GPR. For VA-01 GPR shows the best overall prediction displaying the lowest metric values.

Table 3. Prediction RMSE and MAPE

Spec label	RMSE GPR (kcycles)	MAPE GPR	RMSE NHHSMM (kcycles)	MAPE NHHSMM
CA-03	43.0	33.9%	113.1	58.4%
VA-01	28.7	23.1%	107.3	54.2%
VA-05	50.5	29.5%	45.0	26.9%

4 Concluding Remarks

In this paper Remaining Useful Life prediction of single stringered composite panels is presented. The panels have been subjected in two different fatigue loading conditions and their fatigue life was monitored using FBG strain sensors. A novel Health Indicator was created using Genetic algorithms, fusing together Health Indicators extracted from strain measurements presented in previous work. The fused Health Indicator possessed highly desirable attributes such as high monotonicity and prognosability. Two machine learning models, i.e. Gaussian Process Regression and Non Homogeneous Hidden Semi Markov Model, were used to predict the RUL of three panels. Both methods showed good RUL predictions with comparable performance to each other demonstrating their ability of predicting RUL of more complex structures.

Funding. The research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high cost research equipment grant" (Project Number: 2573).

References

1. Baraldi, P., Bonfanti, G., Zio, E.: Differential evolution-based multi-objective optimization for the definition of a health indicator for fault diagnostics and prognostics. Mech. Syst. Signal Process. **102**, 382–400 (2018)
2. Broer, A., Galanopoulos, G., Benedictus, R., Loutas, T., Zarouchas, D.: Fusion-based damage diagnostics for stiffened composite panels. Struct. Health Monit. **1**(2), 613–639 (2021)
3. Eleftheroglou, N., Loutas, T.: Fatigue damage diagnostics and prognostics of composites utilizing structural health monitoring data and stochastic processes. Struct. Health Monit.: Int. J. **15**(4), 473–488 (2016). https://doi.org/10.1177/1475921716646579
4. Eleftheroglou, N., Zarouchas, D., Benedictus, R.: An adaptive probabilistic data-driven methodology for prognosis of the fatigue life of composite structures. Compos. Struct. **245**, 112386 (2020)
5. Galanopoulos, G., Milanoski, D., Broer, A., Zarouchas, D., Loutas, T.: Health monitoring of aerospace structures utilizing novel health indicators extracted from complex strain and acoustic emission data. Sensors **21**(17), 5701 (2021)
6. Glisic, B., Inaudi, D.: Development of method for in-service crack detection based on distributed fiber optic sensors. Struct. Health Monit.: Int. J. **11**(2), 161–171 (2011). https://doi.org/10.1177/1475921711414233
7. Lei, Y., Li, N., Lin, J.: A new method based on stochastic process models for machine remaining useful life prediction. IEEE Trans. Instrum. Meas. **65**(12), 2671–2684 (2016)
8. Li, Y., Peng, S., Li, Y., Jiang, W.: A review of condition-based maintenance: Its prognostic and operational aspects. Front. Eng. Manage. **7**(3), 323–334 (2020). https://doi.org/10.1007/s42524-020-0121-5
9. Ling, Y., Mahadevan, S.: Integration of structural health monitoring and fatigue damage prognosis. Mech. Syst. Signal Process. **28**, 89–104 (2012)

10. Liu, Y., Mohanty, S., Chattopadhyay, A.: Condition based structural health monitoring and prognosis of composite structures under uniaxial and biaxial loading. J. Nondestr. Eval. **29**(3), 181–188 (2010). https://doi.org/10.1007/s10921-010-0076-2
11. Loukopoulos, P., et al.: Reciprocating compressor prognostics of an instantaneous failure mode utilising temperature only measurements. Appl. Acoust. (2017). https://doi.org/10.1016/j.apacoust.2017.12.003
12. Loutas, T., Eleftheroglou, N., Georgoulas, G., Loukopoulos, P., Mba, D., Bennett, I.: Valve failure prognostics in reciprocating compressors utilizing temperature measurements, PCA-based data fusion, and probabilistic algorithms. IEEE Trans. Industr. Electron. **67**(6), 5022–5029 (2019)
13. Loutas, T., Eleftheroglou, N., Zarouchas, D.: A data-driven probabilistic framework towards the in-situ prognostics of fatigue life of composites based on acoustic emission data. Compos. Struct. **161**, 522–529 (2017). https://doi.org/10.1016/j.compstruct.2016.10.109
14. Milanoski, D.P., Loutas, T.H.: Strain-based health indicators for the structural health monitoring of stiffened composite panels. J. Intell. Mater. Syst. Struct. **32**(3), 255–266 (2020)
15. Moghaddass, R., Zuo, M.J.: An integrated framework for online diagnostic and prognostic health monitoring using a multistate deterioration process. Reliab. Eng. Syst. Saf. **124**, 92–104 (2014)
16. Rabiei, E., Droguett, E.L., Modarres, M.: Damage monitoring and prognostics in composites via dynamic Bayesian networks. In: 2017 Annual Reliability and Maintainability Symposium (RAMS), pp. 1–7. IEEE (2017)
17. Saxena, A., et al.: Metrics for evaluating performance of prognostic techniques. In: 2008 International Conference on Prognostics and Health Management, pp. 1–17. IEEE (2008)
18. Saxena, A., Goebel, K., Larrosa, C.C., Janapati, V., Roy, S., Chang, F.K.: Accelerated aging experiments for prognostics of damage growth in composite materials. Technical report (2011)
19. Shahid, N., Ghosh, A.: TrajecNets: online failure evolution analysis in 2D space. United Technologies Research Center, Penrose Wharf, Penrose Business Center, Cork, Ireland (2019)
20. Sikorska, J.Z., Hodkiewicz, M., Ma, L.: Prognostic modelling options for remaining useful life estimation by industry. Mech. Syst. Signal Process. **25**(5), 1803–1836 (2011). https://doi.org/10.1016/j.ymssp.2010.11.018
21. Silva, S., Almeida, J.: GPLAB-a genetic programming toolbox for MATLAB. In: Proceedings of the Nordic MATLAB Conference, pp. 273–278. Citeseer (2003)
22. Wei, B.S., Johnson, S., Haj-Ali, R.: A stochastic fatigue damage method for composite materials based on Markov chains and infrared thermography. Int. J. Fatigue **32**(2), 350–360 (2010)
23. Wen, P., Zhao, S., Chen, S., Li, Y.: A generalized remaining useful life prediction method for complex systems based on composite health indicator. Reliab. Eng. Syst. Saf. **205**, 107241 (2021)
24. Williams, C.K., Rasmussen, C.E.: Gaussian Processes for Machine Learning, vol. 2. MIT Press, Cambridge (2006)