

Extreme Prognostics for Remaining Useful Life Analysis of Composite Structures

Nick Eleftheroglou¹, Dimitrios Zarouchas² and Rinze Benedictus³

^{1,2,3}*Structural Integrity & Composites Group, Aerospace Engineering Faculty, Delft University of Technology, 2629 HS Delft, the Netherlands*

N.Eleftheroglou@tudelft.nl

D.Zarouchas@tudelft.nl

R.Benedictus@tudelft.nl

ABSTRACT

The procedure of fatigue damage accumulation in composite structures is still unknown and depends on several parameters such as type and frequency of loading, stacking sequence and material properties. Additionally, the nonhomogeneous and anisotropic nature of composites result to a stochastic activation of the different failure mechanisms and make the estimation of remaining useful life (RUL) very complex but interesting task. Data driven probabilistic methodologies have found increasing use the last decade and provide a platform for reliable estimations of RUL utilizing condition monitoring (CM) data. However, the fatigue life of a specific composite structure has a quite significant scatter, with specimens that either underperform or outperform. These specimens are often referred as outliers and the estimation of their RUL is challenging. This study proposes a new RUL probabilistic model, the Extreme Non-Homogenous Hidden Semi Markov Model (ENHHSMM) which is an extension of the Non-Homogenous Hidden Semi Markov Model (NHHSMM). The ENHHSMM uses dynamic diagnostic measures, which are estimated based on the training and testing CM data and adapts dynamically the trained parameters of the NHHSMM. The available CM data are acoustic emission data recorded throughout fatigue testing of open-hole carbon–epoxy specimens. RUL estimations from the ENHHSMM and NHHSMM are compared. The ENHHSMM is concluded as the preferable option since it provides more accurate outlier prognostics.

1. INTRODUCTION

Enhancing safety, availability, and reliability of engineering systems through prognostics has gained momentum over the last decade. Research on prognostics is often correlated with

Nick Eleftheroglou 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.

estimating the remaining useful life (RUL) of the system utilizing condition monitoring (CM) data, where the RUL estimation implies to find its probability density function (pdf). In general, similar engineering systems, which serve under the same conditions, tend to fail at different times during their life-span. Composite structures are a profound example because their lifetime, especially when they are subjected to service loads such as fatigue, deviate remarkable. The main reason for this deviation is the process of damage initiation and propagation, known as fatigue damage accumulation.

The fatigue damage accumulation depends on several parameters such as the type of material and the lay-up, loading frequency and sequence, the form of the fatigue cycle and in combination with the multi-phase nature of composites and the variation of defects, different failure mechanisms are stochastically activated. Thus fatigue damage analysis, and consequently the prognostics, becomes very complex and at the same time challenging task. The fatigue life of a specific composite structure has a quite significant scatter, with extreme cases the left and right outlier. The estimation of their RUL is challenging because the training process of the selected probabilistic model doesn't take into account CM data that are associated with the outlier performance. Therefore, probabilistic methodologies that will account for the extreme cases, are required in prognostics field.

Methods based on stochastic filtering (Orchard & Vachtsevanos, 2009), multi-stage degradation models (Rabiner, 1989) and covariate hazard models (Lu & Liu, 2014) are common methodologies, which can take lifetime scattering into account (Si, Zhang & Hu, 2017). Since the phenomenon of damage accumulation of composites structures is stochastically correlated with CM data, multi-stage degradation models, such as Markov models (MMs), are the preferable approach in order to estimate the RUL of composite structures. MMs have been utilized as early as the 80s (Bogdanoff & Kozin, 1985) but the Markovian

assumption, i.e. future degradation state depends on the current degradation state, is not generally enough in order to describe efficiently the damage accumulation phenomenon in engineering systems. Driven of that drawback, Hidden Markov models (HMMs) have been introduced by Rabiner (1989). HMM is a multistate structure where each state is hidden correlated with the damage accumulation phenomenon. The main disadvantage in that case is the assumption of an exponentially sojourn time distribution for each of the hidden states, which is not always valid. Hidden Semi Markov models (HSMMs) relax this assumption allowing the unconstrained selection of the sojourn time distributions (Peng & Dong, 2011). Both in HMMs and HSMMs, there is a limitation regarding the state transition which is independent on the age of the engineering system or the sojourn time in the current hidden state. In order to take into account this limitation Moghaddass and Zuo (2014^a) extended the HSMM approach developing the Non-Homogenous Hidden Semi Markov model (NHHSMM). According to this model, the state transition depends on the current hidden state, the sojourn time of the current hidden state and the total age of the studied system. However, a limitation of all the aforementioned models, i.e. MMs, HMMs, HSMMs and NHHSMMs, is the lack of adaptation regarding the estimated model's parameters, while the engineering system i.e. composite structure, is operating. To our knowledge, such data driven probabilistic methodologies, which can dynamically adapt the estimated parameters have not been developed for prognostic tasks yet.

Therefore, there is a need for developing methodologies that are able to dynamically adapt the model's parameters based on the testing CM data in the sense that outlier aspects can be reflected. The contribution made in this paper is to propose a new RUL probabilistic model, the Extreme Non-Homogenous Hidden Semi Markov model (ENHHSMM) which is an extension of the NHHSMM. The ENHHSMM, which is described in Section 2, uses the available diagnostic measures and adapts dynamically the trained parameters of the NHHSMM. The remainder of this paper is organized as follows: in Section 3, the case study analysis is presented and finally the paper is concluded in Section 4.

2. EXTREME NHHSMM

As already mentioned in introduction there is a need for developing dynamic probabilistic models which will be able to adapt the estimated model's parameters using the testing CM data. Figure 1 summarizes the RUL prediction methodology, which consists of two parts; the training and testing process. The training process contains the training CM data and the stochastic model while the testing process uses the training process' output, the extracted testing CM data and dynamic diagnostic measures.

The aim of this probabilistic model, which called ENHHSMM, is to be able to estimate more accurately the RUL of an outlier based on the adapted parameters and the available testing CM data.

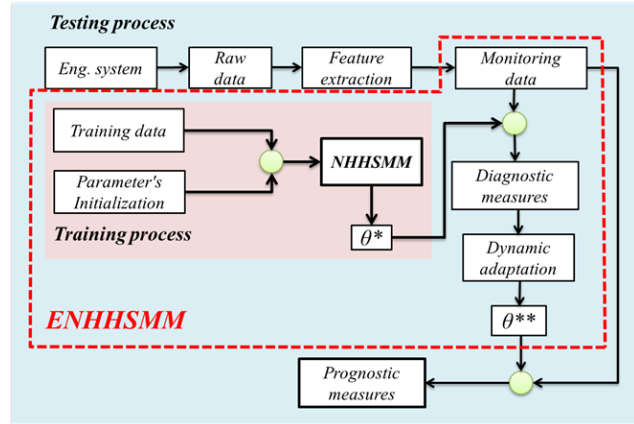


Figure 1. Flowchart of the RUL prediction methodology.

2.1. NHHSMM

This subsection briefly reviews the fundamentals of the NHHSMM. The interested reader can refer to Moghaddass and Zuo (2014) and Eleftheroglou and Loutas (2016) for a more detailed description. The NHHSMM consists of a bi-dimensional stochastic process. The first process forms a finite Semi Markov chain, which is not directly observed, and the second process, conditioned on the first one, forms a sequence of independent random CM data variables. In order to describe the aforementioned bi-dimensional stochastic process the model's parameters θ have to be estimated.

The parameter estimation consists of the initialization and training procedure. The purpose of the initialization procedure is to identify a set of parameters ζ , with high computational efficiency, which will associate the damage accumulation phenomenon and the available CM data. The purpose of the training procedure is to estimate the parameters $\theta = \{\Gamma, \mathbf{B}\}$. Γ parameters characterize the transition rate distribution between the hidden states (degradation process), while \mathbf{B} parameters deal with the correlation between the hidden states and CM data (observation process). This correlation is represented in a nonparametric and discrete form via a matrix called emission matrix. The purpose of the training procedure is to estimate the parameters θ based on the selected parameters ζ . The complete model \mathbf{M} is defined when ζ and θ are known, $\mathbf{M} = \{\zeta, \theta\}$.

The initialization procedure is obtained by defining the following parameters:

- Number of hidden states (N). N refers to the number of discrete levels of degradation. However, hidden states are not quantitatively but just

qualitatively correlated with the damage accumulation phenomenon. The main assumption in this paper is that the system under study starts to operate on its perfect functioning state, hidden state 1, until its total failure i.e. state N. The final state N is not hidden but self-announcing and always corresponds to the failure state. As a result, the last observation of the available CM data should be unique dictating a common failure threshold in the available CM data.

- Transition between the hidden states (Ω). This parameter defines the connectivity between the N selected hidden states and it can be soft (gradual transition to neighbour hidden state), hard (sudden transition from any hidden state to failure state N) and multistep (transition to an intermediate state between the current hidden state and the failure state). Figure 2 illustrates the three possible types of transition. The fatigue damage accumulation, which always increases during the life of the structure, dictates that the available transitions are left to right only.
- Transition rate function (λ). This parameter is the main describer of the degradation process since each transition is going to follow this λ transition rate function. The transition process can depend on the involved hidden states (Markovian property), the sojourn time of the current hidden state, the total operation time (aging) and any other combination between the aforementioned parameters. The most commonly used distributions for the λ function are the Weibull, Gaussian, Exponential and Gamma failure rates. In this study the Weibull failure rate is used since it is the most generic one.
- Discrete CM indicator space ($\mathbf{Z}=\{z_1, z_2, \dots, z_V\}$). The selection of this parameter is crucial for the observation process since emission matrix has N (number of hidden states) rows and V (number of discrete condition monitoring values) columns. The entry in the element (i,j) of the emission matrix represents the probability that z_j CM value is observed when the system is in hidden state i.

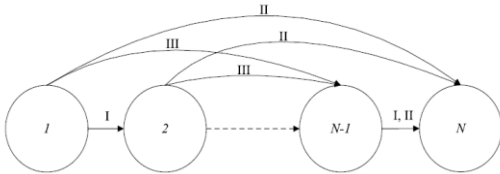


Figure 2. Soft (I), hard (II) and multistep (III) types of transition.

To summarize, the initialization topology can be denoted as $\zeta=\{N, \Omega, \lambda, V\}$. With regards to the training procedure, parameters $\theta=\{\Gamma, \mathbf{B}\}$ are obtained via the maximum likelihood estimation. Moghaddass and Zuo (2014) proposed a method for defining the Maximum Likelihood Estimator (MLE) θ^* of the model parameter θ which leads to maximize the likelihood function $L(\theta, \mathbf{y}^{(1:K)})$, where $\mathbf{y}^{(k)}$ is the k-th degradation history, K is the number of available degradation histories and

$$L(\theta, \mathbf{y}^{(1:K)}) = \prod_{k=1}^K \Pr(\mathbf{y}^{(k)} | \theta) \xrightarrow{L'=\log(L)} L'(\theta, \mathbf{y}^{(1:K)}) = \sum_{k=1}^K \log(\Pr(\mathbf{y}^{(k)} | \theta)) \Rightarrow \theta^* = \arg \max_{\theta} \left(\sum_{k=1}^K \log(\Pr(\mathbf{y}^{(k)} | \theta)) \right) \quad (1)$$

setting initial values for Γ , \mathbf{B} and solving the aforementioned optimization problem, the parameter estimation process is obtained and diagnostics and prognostics can be estimated.

2.2. Diagnostics

Finding a monotonic degradation measure, which at least reflects qualitatively the damage accumulation has always been an interesting and challenging topic in a real time CM applications (Shen et al., 2012). Therefore, a reasonable measure to monitor the overall health status of an engineering system is the diagnostic measure Most Likely State (MLS) (Moghaddass & Zuo, 2014^b), which can be determined via Eq. (2).

$$\text{MLS}(t|y_{1:t}, \mathbf{M}^*) = \arg \max_i \Pr(Q_t = i | y_{1:t}, \mathbf{M}^*) \quad (2)$$

This measure maximizes the probability $\Pr(Q_t = i | y_{1:t}, \mathbf{M}^*)$ of being at the hidden state i at the time point t given the CM data up to time t. With $\mathbf{M}^*=\{\zeta, \theta^*\}$ a specific model topology is denoted.

2.3. Dynamic adaptation process

In addition to the NHHSM's assumptions, presented by Moghaddass and Zuo (2014^a), extra assumptions should be considered for the development of the ENHSM:

- The emission matrix is independent on the time since it correlates CM values and hidden states. As a result, we assume that the emission matrix parameters remain the same during the dynamic adaptation process, that means $\mathbf{B}^{**} = \mathbf{B}^*$.
- The scale and shape parameters of the Weibull failure rate distribution describe the degradation process Γ . The shape parameter can be interpreted

as a value that indicates when the failure rate remains constant, decreases or increases over time. On the other hand, the scale parameter shifts the distribution along the abscissa scale. Assuming that all specimens have the same volume of damage at the end of their fatigue life, the scale Weibull parameter adapts only enabling the sojourn time of each hidden state to shift in time. In order to quantify this shift the aforementioned dynamic diagnostic measure MLS is used. During the testing process the MLS is estimated enabling observation of the transition time from the current hidden state i to any new hidden state j . Therefore, the sojourn time of the i hidden state can be defined ($\text{mean}\Gamma_{i,j}^{**}$). However, the PDF of sojourn time, at hidden state i , is estimated based on the NHHSM's Γ^* parameters ($\text{mean}\Gamma_{i,j}^*$) and a comparison between these two sojourn times ($\text{mean}\Gamma_{i,j}^{**}$, $\text{mean}\Gamma_{i,j}^*$) is achieved. Since the target of the ENHSM is to estimate more accurately the RUL of the testing system the scale Γ^* parameters are dynamically adapted in order to have as mean sojourn time the value which the MLS has defined ($\text{mean}\Gamma_{i,j}^{**}$). This adaptation is determined via introducing the Eq. (3) (Deng & Jiang, 2017).

$$\text{Scale}_{\Gamma_i, j^{**}} = \frac{\text{mean}\Gamma_{i, j^{**}}}{\text{Gamma}(1 + 1/\text{Shape}_{\Gamma_i, j^*})} \quad (3)$$

- The ratios between the training and testing sojourn times of hidden state i and $i+1$ should be constant. To demonstrate this last assumption, which dynamical updates the sojourn times of the future hidden states based on the current and past hidden states' sojourn time adaptation, the following flowcharts and pseudo code are presented.

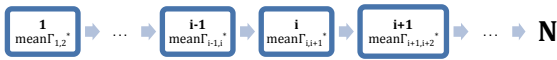


Figure 3. Sojourn times per hidden state based on the NHHSM Γ^* parameters.



Figure 4. Sojourn times per hidden state based on the MLS diagnostic measure when the engineering system just transited from the hidden state i to $i+1$.

The following pseudo code dynamically adapts the sojourn time of each (past, present and future) hidden state when the

engineering system just transited from the present hidden state i to $i+1$.

```

For s=1 to N
    If s<i+1 then
        meanGamma_{s,s+1}^{**} = T_{s,s+1}
        RF_{s,s+1} = \frac{\text{mean}\Gamma_{s,s+1}^{**}}{\text{mean}\Gamma_{s,s+1}^*}
    Else
        Rescaling_Factor = mean(RF)
        meanGamma_{s,s+1}^{**} = Rescaling_Factor * \text{mean}\Gamma_{s,s+1}^*
    End If
End For
    
```

Based on the aforementioned three assumptions the dynamic adaptation process, which is the key element of the ENHSM, is introduced in the present work and receives as inputs the extracted CM data and the estimated model's parameters θ^* . By utilizing the available CM data and θ^* parameters the MLS diagnostic measure is calculated. Afterwards the dynamic adaptation process can be implemented. The rationale behind the adaptation is according to Figure 5.

2.4. Prognostics

Prognostic measures can be defined based on the θ^{**} parameters and the testing CM data. In other words, conditional to the testing CM data and the complete extreme model $\mathbf{M}^{**} = \{\zeta, \theta^{**}\}$, prognostics tries to estimate the probability of being in hidden states $1, \dots, N-1$ at a specific time points in future i.e. the conditional reliability function. Conditional reliability function, $R(t|y_{1:t_p}, L > t_p, \mathbf{M}^{**}) = \Pr(L > t|y_{1:t_p}, L > t_p, \mathbf{M}^{**})$, represents the probability that the studied system continues to operate after a time t , less than life-time L ($L > t$), further than the current time t_p given that the system has not failed yet ($L > t_p$), the testing CM data $y_{1:t_p}$ and the complete model \mathbf{M}^{**} . In this study the mean and confidence intervals of RUL are proposed as prognostic measures. These measures were calculated via the cumulative distribution function (CDF) of RUL (Moghaddass & Zuo, 2014^b). The CDF of RUL is defined at any time point via the conditional reliability according to the following equation:

$$\Pr(\text{RUL}_{t_p} \leq t|y_{1:t_p}, \mathbf{M}^{**}) = 1 - R(t + t_p | y_{1:t_p}, \mathbf{M}^{**}) \quad (4)$$

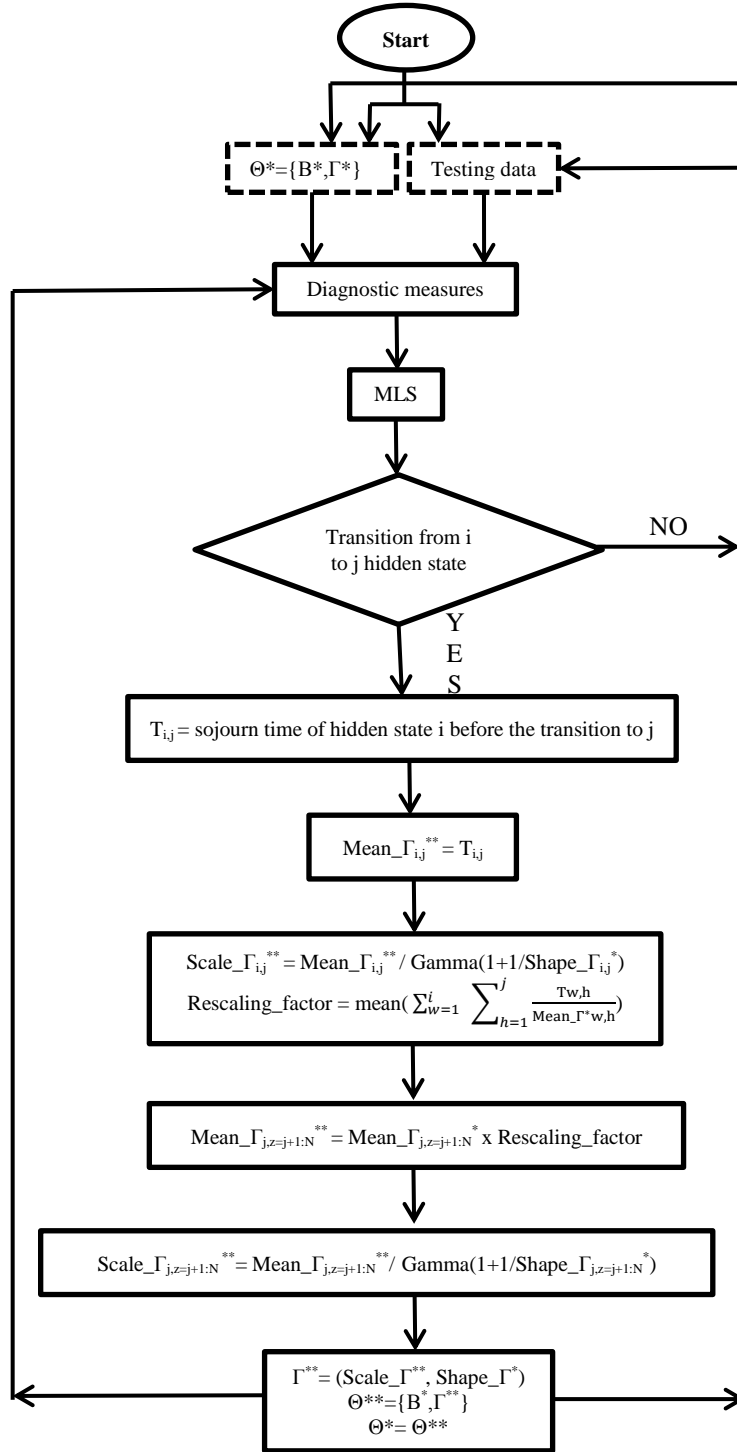


Figure 5. Dynamic adaptation process flowchart.

3. CASE-STUDY

To demonstrate the RUL estimations of the ENHHSMM, a case study on composite structures is considered where open-hole carbon/epoxy specimens were subjected to fatigue loading until failure.

3.1. Experimental campaign

A laminate with $[0/\pm 45/90]_{2s}$ lay-up was manufactured using the autoclave process and seven specimens with the following geometrical details: dimensions [300mm x 30mm] and a central hole of 6mm diameter were tested at 90% of the static tensile strength with $R=0$ and $f=10$ Hz.

An acoustic emission (AE) system was used in order to perform the AE measurements. $1/A$ (1/amplitude) calculated cumulatively in periodic windows of 500 cycles for all the tested specimens. $1/A$ is chosen since it produced more monotonic observation sequences than other conventional AE features e.g. energy, hits, RA (rise time/amplitude), as it was observed in (Eleftheroglou & Loutas, 2016). The AE degradation histories for seven specimens are shown in Figure 6 with left outlier's (specimen05) degradation history highlighted. The reader can refer to Eleftheroglou et al. (2016) for a more detailed description of the experimental campaign.

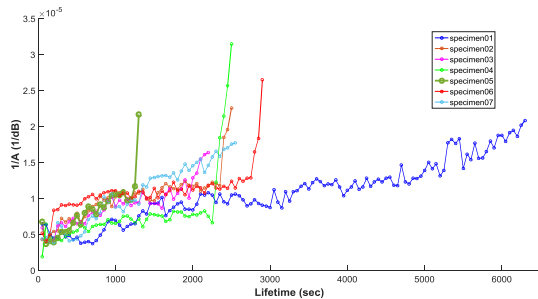


Figure 6. Acoustic Emission degradation histories.

3.2. Extreme NHHSM Implementation

Initially, the procedure of damage accumulation in composite structures under fatigue loading is modelled via the NHHSM and $\theta^* = \{B^*, \Gamma^*\}$ parameters were determined via the Maximum Likelihood Estimation procedure Eq. (1). In Figure 7 the dashed lines depicts the NHHSM estimated Γ^* parameters.

Regarding the initialization procedure the Bayesian Information Criterion (BIC) was employed to estimate the optimum number of hidden states which was found at $N=4$. Actually, the number of degradation states N suggesting that the damage accumulation process of the studied composite structure can be approximated as a four-state procedure. Figure 8 presents the estimations of the aforementioned diagnostic measure MLS as calculated from Eq. (2) at each time point during the fatigue test of specimen05. Figure 8

reflects that specimen05 is an outlier since the sojourn time of the hidden state 1 based on MLS is just 600 sec and based on the NHHSM is 840.6 sec (Figure 5), similar results were obtained for the sojourn time of the hidden state 2 since MLS sojourn time is 600 sec and NHHSM sojourn time is 1123.4 sec. Utilizing the NHHSM estimated parameters θ^* , the testing CM data and the MLS estimations the ENHHSMM can be defined and dynamically adapt the parameters θ^* to θ^{**} , following the process which described in subsection 2.3. In Figure 7 the outcome of the ENHHSMM is presented with solid lines.

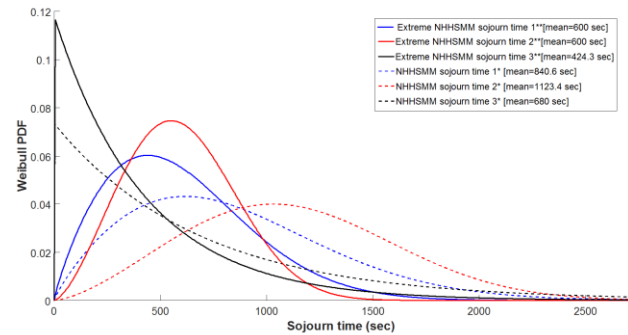


Figure 7. Sojourn time Weibull distributions utilizing the Γ^* and Γ^{**} parameters.

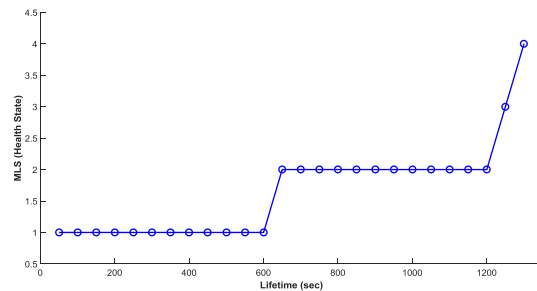


Figure 8. MLS diagnostic measure of specimen05.

Based on Figure 7 the ENHHSMM Weibull pdfs are shifted to the left as it was desired since Specimen05 is the left outlier of this case study. In this direction the ENHHSMM RUL estimations are expected to be more accurate, comparing with the NHHSM RUL estimations since the mean sojourn time values are getting shorter.

3.3. Remaining Useful Life Estimations

Specimen05 is actually the left outlier in the sense that fails earlier than all the specimens in its training set. The degradation history of Specimen05 has not been included in the training procedure. Consequently, the minimum failure time of this training data set is 2200 sec while the Specimen05's failure time is 1300 sec, see Figure 6. Figure 9 presents the RUL estimations of the NHHSM and the ENHHSMM regarding that specimen.

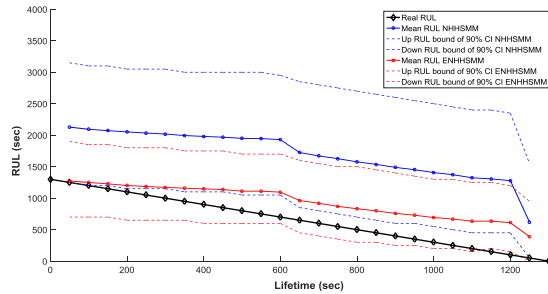


Figure 9. RUL estimations of the left outlier specimen05.

However, it is not possible to detect an outlier before it reaches the end of its lifetime. Therefore, in terms of robustness the proposed model should be also applied to non-outlier cases. Based on Figure 6, specimen03 is an interesting case since its degradation rate is similar to left outlier's rate but its failure time is almost double i.e. 2200 sec. Figure 10 presents the RUL estimations of the NHHSSM and the ENHSSM of specimen03.

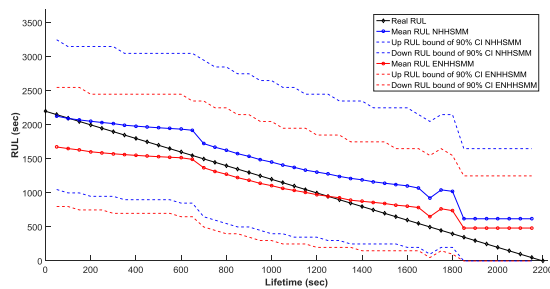


Figure 10. RUL estimations of specimen03.

Based on Figure 9 and 10 the ENHSSM provides better outlier and non-outlier prognostics since the mean ENHSSM RUL estimations are able to approach more satisfactorily the real RUL estimations than the NHHSSM. Additionally, the confidence intervals of the ENHSSM contain the real RUL curve during almost the whole lifetime of specimen05/specimen03 and their distance is shorter than the classic model in both of the two aforementioned specimens. Furthermore regarding the outlier, the initial mean 'extreme' RUL estimations almost overlaps to the real RUL and at the same time the 'classic' RUL estimations are not accurate, 150% overestimation. Therefore, the ENHSSM can identify early enough an outlier and adapt the RUL estimations in an efficient and accurate way.

4. CONCLUSIONS

In this paper, a new dynamic probabilistic model, which can estimate more accurately the RUL of an outlier was developed. The efficiency and the correctness of the ENHSSM were evaluated during experimental campaign of open-hole carbon/epoxy specimens. The specimens were subjected to constant amplitude fatigue loading up to failure

and AE technique were employed in order to provide the required CM data. In conclusion, the results demonstrate that the ENHSSM provides better prognostics than the NHHSSM. As a result, adapting the model's parameters using the available diagnostic measures has the potential to estimate the RUL of extreme and non-extreme cases more efficiently.

REFERENCES

- Bogdanoff, J. L. & Kozin, F. (1985). *Probabilistic models of cumulative damage*. New York, USA: Wiley-Interscience
- Deng, B. & Jiang D. (2017). Determination of the Weibull parameters from the mean value and the coefficient of variation of the measured strength for brittle ceramics. *Journal of Advanced Ceramics*. vol. 6, pp. 149–156. doi: 10.1007/s40145-017-0227-3
- Eleftheroglou, N. & Loutas, T. (2016). Fatigue damage diagnostics and prognostics of composites utilizing structural health monitoring data and stochastic processes. *Structural Health Monitoring*. vol. 15, pp. 473-488. doi:10.1177/1475921716646579
- Eleftheroglou, N., Zarouchas, D.S. & Loutas, T.H. (2016). In-situ fatigue damage assessment of carbon-fibre reinforced polymer structures using advanced experimental techniques. *Proceedings of 17th European Conference on Composite Materials*. June 26-30, Munich, Germany. doi: 10.13140/RG.2.1.4750.5529
- Lu, XF. & Liu, M., (2014). Hazard rate function in dynamic environment. *Reliability Engineering and System Safety*. Vol. 130, pp. 50–60. <https://doi.org/10.1016/j.res.2014.04.020>
- Moghaddass, R. & Zuo M. J. (2014^a). Multistate degradation and supervised estimation methods for a condition-monitored device. *IIE Transactions*. vol. 46, pp. 131–148. doi: 10.1080/0740817X.2013.770188
- Moghaddass, R. & Zuo, M. J. (2014^b). An integrated framework for online diagnostic and prognostic health monitoring using a multistate deterioration process. *Reliability Engineering and System Safety*. vol.124, pp. 92–104. <https://doi.org/10.1016/j.res.2013.11.006>
- Orchard, M. & Vachtsevanos, G. (2009). A particle-filtering approach for on-line fault diagnosis and failure prognosis. *Transactions of the Institute of Measurement and Control*. vol. 31, pp. 221–246. doi:10.1177/0142331208092026
- Peng, Y. & Dong, M. (2011). A prognosis method using age dependent hidden semi-Markov model for equipment health prediction. *Mechanical Systems and Signal Processing*. vol. 25, pp. 237–252. doi:10.1016/j.ymsp.2010.04.002
- Rabiner, L., (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of IEEE*. vol. 77, pp. 257–285. doi: 10.1109/5.18626

- Shen, Z., He, Z., Chen, X., Sun, C. & Liu, Z. (2012). A monotonic degradation assessment index of rolling bearings using fuzzy support vector data description and running time. *Sensors*. vol. 12, pp. 10109–35. doi: 10.3390/s120810109
- Si, X. S., Zhang, Z. X. & Hu, C. H. (2017). *Data-Driven Remaining Useful Life Prognosis Techniques*. Beijing, China: Springer Series in Reliability Engineering