Degradation prognosis based on a model of Gamma process mixture

Edith Grall-Maes, Pierre Beauseroy, and Antoine Grall

ICD-LM2S, Université de Technologie de Troyes,
UMR 6281 CNRS,Troyes, France
edith.grall@utt.fr
pierre.beauseroy@utt.fr
antoine.grall@utt.fr

ABSTRACT

A novel method is proposed to exploit jointly degradation measurements originating from a set of identical systems for making a degradation prognosis. The systems experience different degradation processes depending on operational conditions. The degradation processes are assumed to be Gamma processes. The aim is to cluster the degradation paths in classes corresponding to the different operational conditions in order to group properly the data for the estimation of degradation process parameters. A model of Gamma process mixture is considered and an expectation-minimization approach is proposed to estimate the unknown parameters. The feasibility of the method is shown on simulated cases. Prognosis results obtained with the proposed method are compared with results obtained with basic strategies (considering each system alone or all system together).

1. INTRODUCTION

To estimate the remaining useful lifetime (RUL) of a deteriorating system it is necessary to be able to model its deterioration in order to predict when the deterioration leads to a failure i.e. when it reaches a given threshold. To perform this RUL prognosis one generally relies on measurements of the degradation level and on a degradation model which is assumed to describe the degradation evolution in time (Si, Wang, Hu, & Zhou, 2011; Nystad, Gola, & Hulsund, 2012). For example, in the case of a metal pipe corrosion, the thickness provides a deterioration measure.

The Gamma process is widely used for degradation models when deterioration is monotonic and gradual (Van Noortwijk, 2009). This process is defined by a set of parameters, in particular the shape and scale parameters in the case of an homogeneous process. These parameters are usually unknown and must be estimated in order to perform prognosis. Obviously the reliability of prediction is directly related to the estimation precision. Most of the time, in operational conditions the amount of measurement is very limited. So when a set of similar systems is available, one can use the data coming from all the systems in the set to estimate the model parameters. The expected gain of using all measurements together is to improve the estimator precision (reduction of its variance for example).

By considering all systems as a single one while estimating the model parameters it is assumed that the degradation process model is the same for all systems. In most cases, the degradation process depends also on operating conditions that may be partially unknown. In the pipe corrosion example, the evolution of the pipe thickness depends on the used metal but it also depends on the characteristics of the fluid carried by the pipe (liquid/gaz, temperature, pressure...) and on the location (air/ground/underwater...) and on the environmental conditions of the pipe (temperature, humidity...).

In this paper, we consider that we have a limited amount of data from different systems. Each system has one operating condition among an unknown finite number. Then each system evolves in relation with its operating condition, which remains always the same. This is not a system which evolves in different classes corresponding to functioning modes, as in (Ramasso & Gouriveau, 2014).

We propose a method to cluster the observed systems in classes, corresponding to each operating condition. The degradation process is assumed to be ruled by a Gamma process model. The aim is to estimate the parameters of these Gamma processes in order to predict their RUL. In order to tackle the hypothesis of a number of operating conditions, a model of Gamma process mixture is introduced. An expectation-minimization algorithm is proposed to estimate the parameters of each process in the mixture model.

The problem is formalized in section 2. Then in section 3, the mixture model and the expectation-minimization algorithm (Ambroise & Govaert, 1998)(Hu & Sung, 2006) are presented to determine the clusters. In section 4, the considered prognosis is described and a criterion for comparing the prognosis values obtained according to different strategies of the Gamma process parameter estimation is proposed. In section 5, results on simulated data are presented and analyzed. A conclusion on the selection of process classes number and on future developments ends the paper.

2. PROBLEM FORMULATION

This section first gives a description of the problem and notations and ends with a brief recall about the Gamma process.

2.1. General Aim

The data we consider originates from N paths describing degradation process realizations. Each path p_n, $(n = 1, \ldots, N)$ is composed of $|p_n|$ observations $l_{n,i}$ with $i = 1..|p_n|$. The observation $l_{n,i}$ is characterized by the time instant $t_{n,i}$, and the deterioration level $x_{n,i} = x(t_{n,i}) \in \Omega_x \subset \mathcal{R}$. Then $p_n = \{l_{n,i} = (t_{n,i}, x_{n,i})\}_{i=1..|p_n|}$.

We suppose that the observation set can be divided into K unknown clusters C_k, $(k = 1, \ldots, K)$. In practise, each cluster would correspond to an operating condition. Each cluster represents a deterioration process characterized by some unknown parameter vector belonging to the parameter set $\Theta = \{\theta_k\}_{k=1}^{K}$. Besides, the latent membership, or cluster labels of paths are denoted by $\mathbf{z} = \{z_n\}_{n=1}^{N}$ where $z_n = k$ means that the path p_n belongs to the cluster k. A set of cluster labels defines a partition of all paths. Obviously, for a given path n, all observations $l_{n,i}$, $(i = 1, \ldots, \mid p_n \mid)$ belong to the same cluster.

The aim is to determine the latent cluster label for each observation and jointly the parameter vector of each process. Afterwards this parameter vector can be used to do prognosis. As an illustration, the prognosis we consider in this paper is the mean remaining useful lifetime, defined as the remaining time before reaching a given deterioration threshold which is the failure limit.

In this paper the chosen model for the deterioration processes is the Gamma process, parameterized by 3 parameters, a, b and u described in section **??**. The increments, given by $x(t_{n,i}) - x(t_{n,i-1})$ with $t_{n,0} = 0$ and $x(0) = 0$, are independent. Their density distribution in the cluster k depends on the time and on the parameter θ_k, and can be written as

$$f_k\left(x(t_{n,i}) - x(t_{n,i-1}) \mid t_{n,i}, t_{n,i-1}, \theta_k\right).$$

In the following, for simplicity we will use the notation $f_k(\Delta x_{n,i}|\theta_k)$. It has to be noticed that the density distributions of all the degradation increments are not the same because the increments are usually all different and/or the process may be not stationary.

The objective is to find out the unknown cluster labels $\{z_n\}_{n=1}^{N}$ and consequently the distribution parameter set $\theta = \{\theta_k\}_{k=1}^{K}$, such that paths in the same cluster originate from a process model with the same parameters.

The relevance of a partition described by $\mathbf{z}$ and a parameter set θ can be measured using the log-likelihood given by

$$l(\mathbf{z}, \theta) = \sum_{n=1}^{N} \sum_{i=1}^{|p_n|} log f_{z_n}(\Delta x_{n,i} \mid \theta_{z_n}) \tag{1}$$

2.2. Gamma process

Mathematically, the Gamma process is defined as follows: let $A(t)$ be a non-decreasing, right-continuous, real-valued function for $t \geq 0$, with $A(0) = 0$. The Gamma process with shape function $A(t)$ and scale parameter $b > 0$ is a continuous-time stochastic process $\{X(t), t \geq 0\}$ such that:

- $X(0) = 0$ with probability one;
- $\{X(t), t \geq 0\}$ is a stochastic process with independent increments;
- $X(t) - X(s)$ follows the Gamma distribution $\Gamma(A(t) - A(s), b)$ for $0 \leq s < t$

The definition of the Gamma process leads to two straightforward properties:

- $\{X(t),\ t \geq 0\}$ is a non-decreasing process.
- For all $t \geq 0$, the expectation value and the variance of $X(t)$ could be written as:

$$\mathrm{E}(X(t)) = \frac{A(t)}{b} \quad \mathrm{Var}(X(t)) = \frac{A(t)}{b^2} \tag{2}$$

In the degradation modeling framework, a non-homogeneous Gamma process defined by $A(t) = at^u$, $(a > 0, u > 0)$ is often considered. Thus the process is described by three parameters : a, b, and u. In this case, $X(t) - X(s)$ follows the Gamma distribution $\Gamma(a(t^u - s^u), b)$.

Two methods are often mentioned for the parameter estimation of the Gamma process: the moments estimation and the maximum likelihood estimation (Cinlar, Osman, & Bazant, 1977). The maximum likelihood estimator is asymptotically unbiased, which means the estimates converge to the true values as the number of observations increases as: $N \to \infty$. On the other hand, the moments approach leads to simpler formulae of the estimator. It is more straightforward to implement and the computation time is much reduced compared with the maximum likelihood method.

3. PROCESS CLUSTERING

The proposed approach is based on the computation of mixture models using the expectation-maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977). Beside, side information is considered according to (Shental, Bar-Hillel, Hertz, & Weinshall, 2003).

3.1. Related work

The EM algorithm is an iterative method that produces a set of parameters that locally maximizes the log-likelihood of a given sample, starting from an arbitrary set of parameters. It is often used to estimate the unknown parameters of a mixture model of K p.d.f. f_k given by:

$$f(x|\theta) = \sum_{k=1}^{K} \alpha_k f_k(x|\theta_k)$$

where α_k is the probability of class k and $f_k(x|\theta_k)$ the a posteriori probability of class k.

Furthermore the procedure we use is based on the work of Shental et al. (Shental et al., 2003) which describes an EM procedure for a Gaussian mixture model and for handling positive constraints, indicating that some observations arise from the same source. The data set is assumed to be a set of chunklets, and each chunklet is a set of points that originate from the same source. Alternating E steps and M steps leads to the estimation of the probability of each class, and the parameters (mean and variance) of each Gaussian class. The solution can be considered as a soft partition.

3.2. Proposed method

The problem we deal with, in comparison with the problem considered in (Shental et al., 2003), has to lead to a hard partition. Then we add an intermediate classification step between the E and M steps. Such a classification step has been introduced in the CEM algorithm (Celeux & Govaert, 1992, 1995) for hard classification problem using mixture models without constraints.

In comparison with the problem considered in (Shental et al., 2003), there is another main difference. The model is not a Gaussian mixture model and especially the degradation increments have different density distributions. The density distributions of all the degradation increments would be the same only in the case of homogeneous Gamma process, and of regularly sampled paths.

Thus, we have proposed an algorithm based on the mixture models for the problem of statistical process clustering with the two following properties. On the one hand, it takes into account that observations in a same path belong to a same class, and on the other hand it takes into account that a hard classification is searched.

The E step at iteration m consists in calculating an estimation of the *a posteriori* probability for each observation using the parameters $\theta^{(m-1)} = (a^{(m-1)}, b^{(m-1)}, u^{(m-1)})$. The posterior probability $c_{nk}^{(m)}$ at iteration m that the path n belongs to class k, given p_n and the parameter $\theta^{(m-1)}$ writes according to

$$\begin{aligned} c_{nk}^{(m)} &= p(z_n = k|p_n, \theta^{(m-1)}) \\ &= \frac{\alpha_k^{(m-1)} \prod_{i=1}^{|p_n|} f_k(\Delta x_{n,i}|\theta_k^{(m-1)})}{\sum_{r=1}^{K} \alpha_r^{(m-1)} \prod_{i=1}^{|p_n|} f_r(\Delta x_{n,i}|\theta_k^{(m-1)})} \end{aligned} \quad (3)$$

with

$$f_k(\Delta x_{n,i}|(a_k, b_k, u_k)) \sim \Gamma(a_k(t_{n,i}^{u_k} - t_{n,i-1}^{u_k}), b_k)$$

The expectation of the log-likelihood over all possible assignments which comply the given constraints is given by:

$$E(l(\mathbf{z}, \theta)) = \sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{i=1}^{|p_n|} log f(\Delta x_{n,i}|k, \theta_k) p(z_n = k|p_n, \theta_k^{(m-1)})$$

$$+ \sum_{k=1}^{K} \sum_{n=1}^{N} log \alpha_k p(z_n = k|p_n, \theta_k^{(m-1)})$$

The M step at iteration m consists in computing the parameters $\alpha_k^{(m)}$, and $\theta^{(m)}$ that maximize the expected log-likelihood found on the E step. The parameter $\alpha_k^{(m)}$ is given by

$$\alpha_k^{(m)} = \frac{1}{N} \sum_{n=1}^{N} (z_n^{(m)} = k) \quad (4)$$

and the parameters $(a^{(m)}, b^{(m)}, u^{(m)})$ are determined by maximization of the log-likelihood.

Then the algorithm is the following one.

- Initialize the parameter set $\theta^{(0)}$
- Repeat until $l(\mathbf{z}^{(m)}, \theta^{(m)}) - l(\mathbf{z}^{(m-1)}, \theta^{(m-1)}) < \epsilon$
  - compute $c_{nk}^{(m)}$ for each path n and each class k using relation (3)
  - determine the partition $\mathbf{z}^{(m)}$: choose $z_n^{(m)} = k$ corresponding to the largest value $c_{nk}^{(m)}$
  - determine the parameter vector $\theta^{(m)}$ that maximizes $l(\mathbf{z}^{(m)}, \theta^{(m)})$
  - compute the new value of the probability $\alpha_k^{(m)}$ for each class k using relation (4).

4. PROGNOSIS AND PERFORMANCE EVALUATION

4.1. Considered Prognosis

The considered prognosis is the remaining mean time until a threshold is reached. Let S be a threshold, p_n a path with its last observation $(t_{n,|p_n|}, x_{n,|p_n|})$, its class label z_n,

and a set of Gamma process parameters corresponding to the class label $\theta_{z_n} = (a_{z_n}, b_{z_n}, u_{z_n})$. If the last degradation level is smaller than the threshold, i.e. $x_{n,|p_n|} < S$, then it is possible to estimate the remaining mean time until the threshold is reached. This time is noted $T_{n,\theta_{z_n}}(S)$. Since the increment $\Delta X_{n,i}$ follows a Gamma distribution given by $\Gamma(a_{z_n}(t_{n,i}^{u_{z_n}} - t_{n,i-1}^{u_{z_n}}), b_{z_n})$, its mean is

$$E(\Delta X_{n,i}) = \frac{a_{z_n}}{b_{z_n}}(t_{n,i}^{u_{z_n}} - t_{n,i-1}^{u_{z_n}})$$

Then, taking $i = |p_n| + 1$, it leads to the value $T_{n,\theta_{z_n}}(S)$ given by:

$$T_{n,\theta_{z_n}}(S) = \left((S - x_{n,|p_n|})\frac{b_{z_n}}{a_{z_n}} + t_{n,|p_n|}^{u_{z_n}}\right)^{\frac{1}{u_{z_n}}} - t_{n,|p_n|}$$

For a given value of K, a set $\Theta = \{\theta_k\}_{k=1}^K$ and a set of class labels $\mathbf{z} = \{z_n\}_{n=1}^N$, the remaining mean time until a threshold S is reached can be computed for each path n.

4.2. Prognosis performance evaluation

In the case of simulated data, it is possible to compare the estimated prognosis result with the theoretical one. We have prefered to use the theoretical remaining useful time than a simulated value that we could obtain by running the path up to the failure threshold. The estimated prognosis result for path n, $T_{n,\widetilde{\theta}_{\widetilde{z_n}}}(S)$, is obtained using the estimated set of parameters $\widetilde{\Theta}$ and the estimated set of class labels $\widetilde{\mathbf{z}}$. The theoretical prognosis is noted $T_{n,\theta_{z_n}}(S)$.

A large number of metrics in the forecasting applications have been proposed, as accuracy and precision, which are classical metrics. The metrics we propose to use in this paper for assessing the prognosis is near to relative accuracy given in (Saxena, Celaya, Saha, Saha, & Goebel, 2010). It is a relative error criterion which allows to give the same importance to all classes. This is critical in our case because the precision depends on the class evolution. For a path n, and a threshold S we define the relative error $e_n(S)$ as:

$$e_n(S) = \frac{T_{n,\widetilde{\theta}_{\widetilde{z_n}}}(S) - T_{n,\theta_{z_n}}(S)}{T_{n,\theta_{z_n}}(S)} \quad (5)$$

Using all the paths for which the threshold is not reached for the last sample, i.e. $x_{n,|p_n|} < S$, it is possible to compute the mean of all the errors e_n to obtain $Ee(S)$ and to compute the standard deviation to obtain $Se(S)$

$$Ee(S) = \widehat{E}[\{e_n(S)\}_{n|x_{n,|p_n|}<S}] \quad (6)$$

$$Se(S)(S) = \widehat{\sigma}[\{e_n(S)\}_{n|x_{n,|p_n|}<S}] \quad (7)$$

The mean error should be equal to 0. The criterion which characterizes the performance of an approach is the standard deviation of the error.

Table 1. parameters - situations 1 and 2

Situation 1

	class 1	class 2	class 3	class 4
a	16.67	28.12	41.67	55.80
b	1.67	1.87	2.083	2.23
u	0.8	0.8	0.8	0.8

Situation 2

	class 1	class 2	class 3	class 4
a	16.67	24.5	33.75	43.21
b	1.67	1.75	1.87	1.96
u	0.8	0.8	0.8	0.8

5. RESULTS

Simulations have been done considering two situations. For both of them, there are 4 classes, each class with 6 paths, each path with 3 samples. The time increments are within an uniform distribution between 2 and 8. Parameters for both situations are given in table 1.

The mean theoretical evolution respectively for situation 1 and situation 2 is described in figure 1a and figure 3a. In situation 2, the classes are more similar than in situation 1: at each instant the mean values for 2 different classes are closer than in situation 1. However the standard deviations are the same for both situations.
Example of simulated data respectively corresponding to situation 1 and situation 2 are given in figure 2a and figure 4a.

The simulated data has been used to determine jointly the class of each path and the parameter set of each class, for different values of K (a priori number of classes). Simulations have been done for K between 1 and 7.
For the example of situation 1, the estimated class of each path for $K = 3, 4, 5$ is described in figures 2b, c, and d. The mean evolution of the degradation corresponding to the estimated parameters is given in figures 1b, c, and d. In the case of $K = 4$ classes, it is possible to determine the number of paths which are misclassified, since it corresponds to the theoretical number of classes. It can be seen than one path of class 1 ('+' red) is affected to class 2 ('*' green). All other paths are correctly classified.

Similar results for situation 2 are given in figure 4, for the estimated class of each path, and in figure 3, for the mean evolution of the degradation. In the case of $K = 4$ classes, 3 paths are misclassified: one path of class 1 is affected to class 2 and two paths of class 2 are affected to class 3.

The prognosis $T_{n,\widetilde{\theta}_{\widetilde{z_n}}}(S)$ has been determined in 9 cases of estimation of parameter $\widetilde{\theta}$:

- "semi-theoretical case" : the parameter set is estimated assuming the true class of each path is known;
- "path case" : a parameter set is estimated for each path using the 3 observations of the considered path;
- "estimated K-class case" (for $K = 1 \ldots 7$): the class of each path and the parameter set of each class are determine jointly.

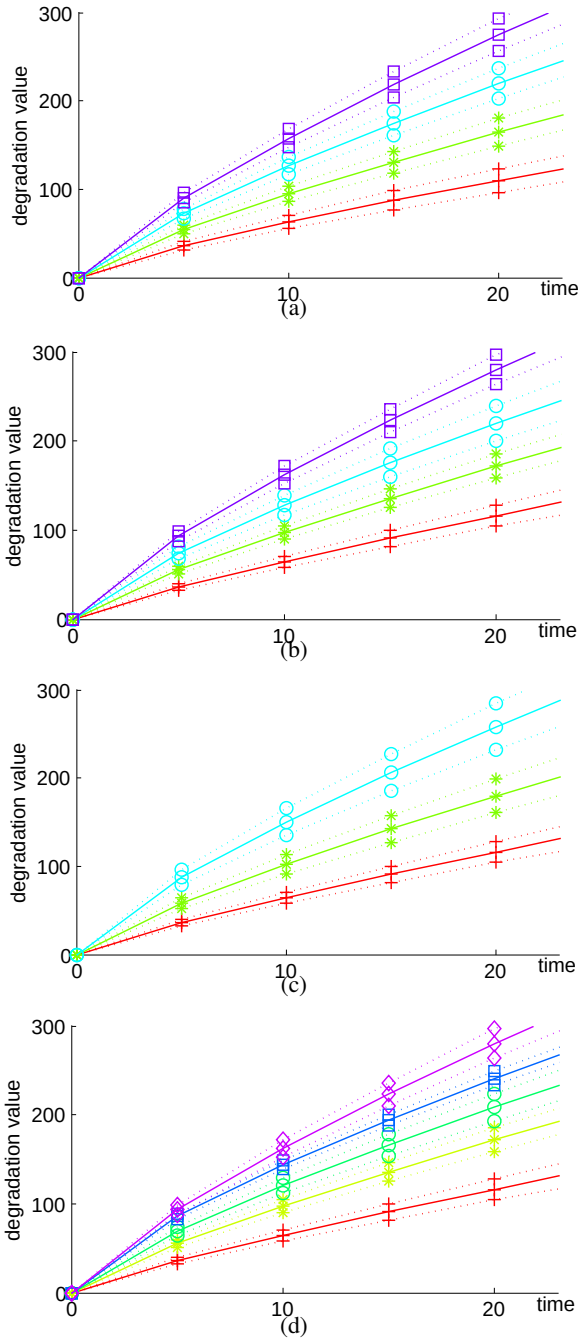

Figure 1. Evolution of the mean degradation value in situation 1 for (a) the theoretical parameters (b) the estimated parameters with 4 classes (c) the estimated parameters with 3 classes (d) the estimated parameters with 5 classes. Each color corresponds to an estimated class.

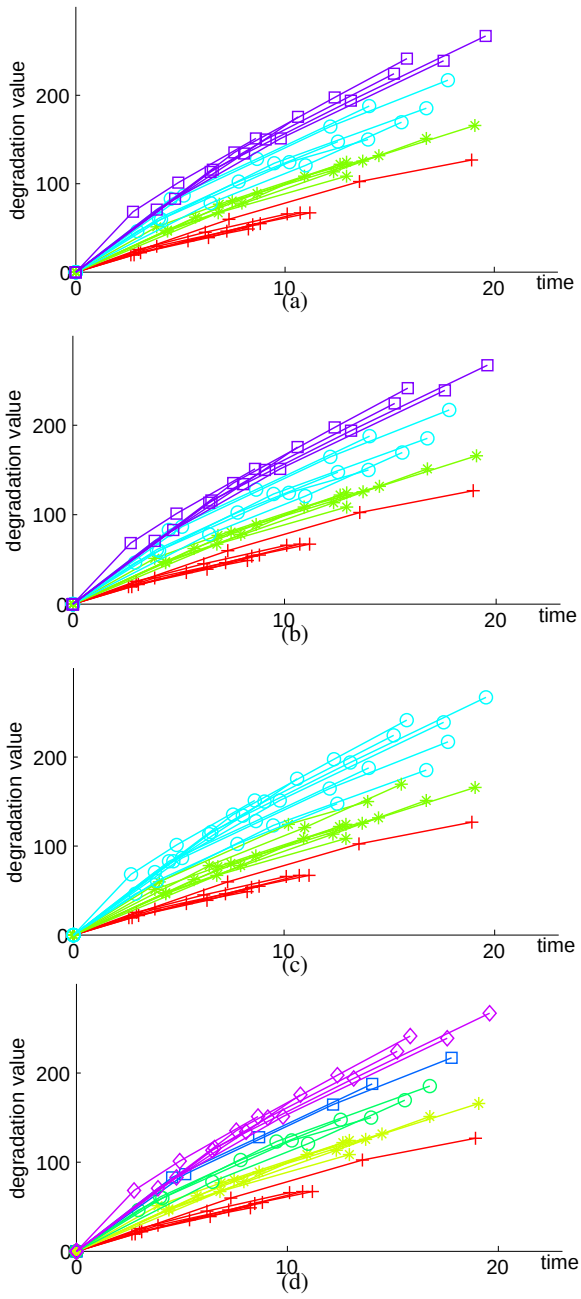

Figure 2. Example of simulated data in situation 2 for (a) the theoretical class (b) the estimated class for 4 classes (c) the estimated class with 3 classes (d) the estimated class with 5 classes. Each color corresponds to an estimated class.

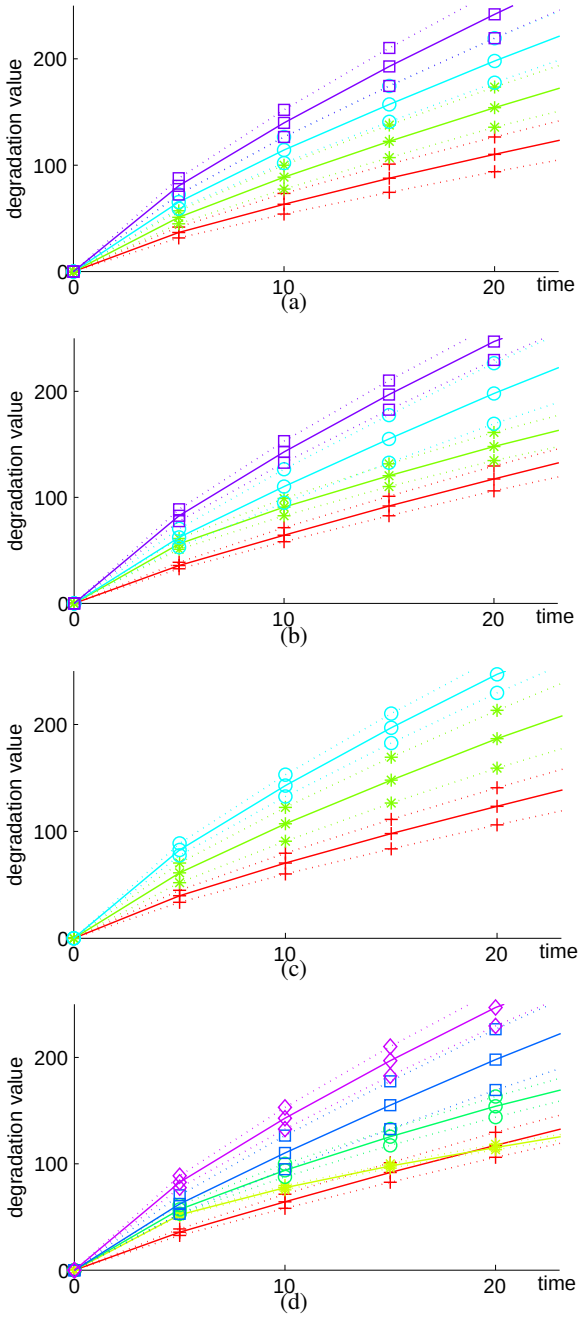

Figure 3. Evolution of the mean degradation value in situation 2 for (a) the theoretical parameters (b) the estimated parameters with 4 classes (c) the estimated parameters with 3 classes (d) the estimated parameters with 5 classes. Each color corresponds to an estimated class.

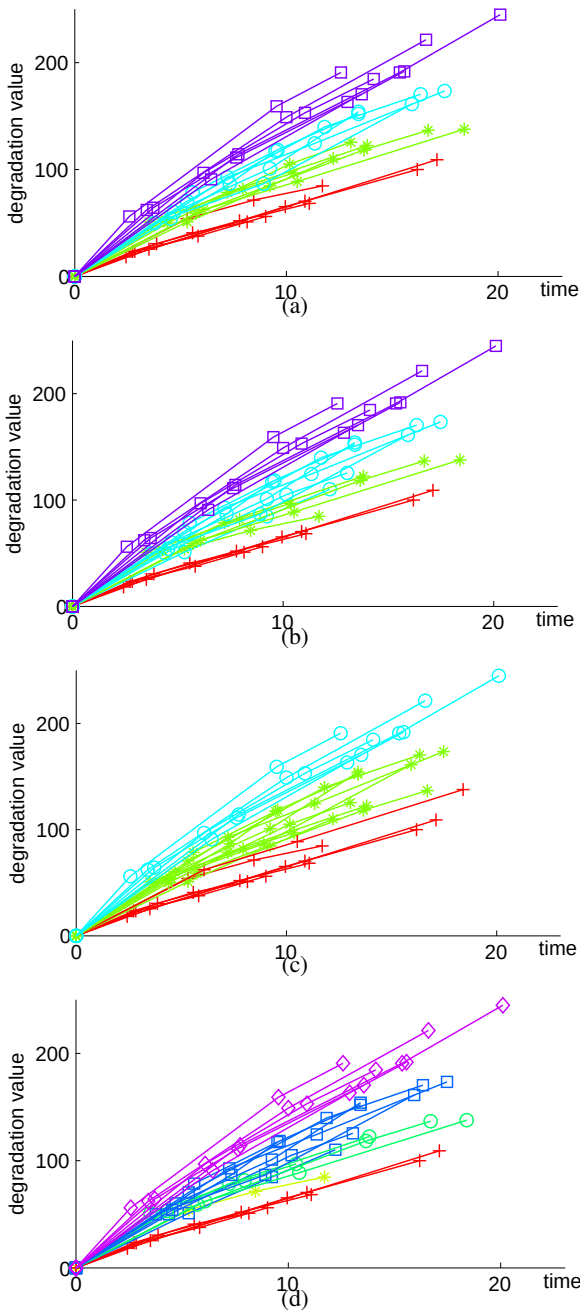

Figure 4. Example of simulated data in situation 2 for (a) the theoretical class (b) the estimated class for 4 classes (c) the estimated class with 3 classes (d) the estimated class with 5 classes. Each color corresponds to an estimated class.

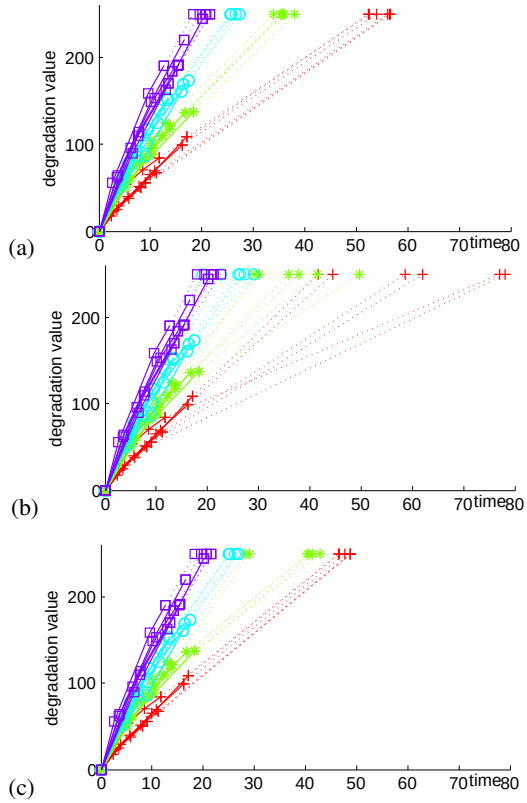

Figure 5. Example of prognosis for (a) the theoretical parameters (b) the estimated parameters for each path (c) the estimated parameters with 4 classes. Each color corresponds to a theoretical class.

The prognosis obtained in the "semi-theoretical case" leads to the minimum error which is reachable, for a given path set. The obtained error is due to the error of the parameters, arising from the estimation with a limited number of paths.

In figure 5, an example of prognosis for situation 2 with a threshold $S = 250$ shows "theoretical case", "path case" and "estimated 4-class case".

In the "path case" there exists a large variance and the bad estimation is due to the very low number of samples for each path. This is particularly visible when the degradation level at the last inspection time is far from the failure level i.e. when the time of prognosis is far from the failure time. In the case of class 1 (red '+') predicted mean failure times are in $[40, 80]$ instead of $[50, 60]$ for the "theoretical case". In the "estimated 4-class case", the impact of the misclassified paths appears clearly. One path of class 1 is affected to class 2 and two paths of class 2 are affected to class 3. Hence the estimated value for parameter u is smaller than its theoretical value for class 2 (green '*') and the green line on figure 3b is more curved than on figure 3a. As a consequence the estimated mean residual lifetime for class 2 is greater than the theoretical one.

The simulation has been repeated for 200 path sets and for

Table 2. Estimated mean $Ee(S)$ (relation 6) and estimated standard deviation $Se(S)$ (relation 7) for situation 1

(a)

threshold	200	300	350
semi-theor.	0.0057	0.0050	0.0055
1 path	0.0289	0.0276	0.0309
1 class	-0.1183	-0.0123	-0.0104
2-class	-0.0194	0.0002	0.0009
3-class	-0.0080	0.0031	0.0037
4-class	0.0067	0.0059	0.0064
5-class	0.0103	0.0089	0.0097
6-class	0.0126	0.0109	0.0119
7-class	0.0143	0.0128	0.0140

(b)

threshold	200	300	350
semi-theor.	0.0450	0.0520	0.0555
1 path	0.1629	0.1813	0.1954
1 class	0.2960	0.3462	0.3512
2-class	0.1954	0.1897	0.1924
3-class	0.1117	0.1193	0.1221
4-class	0.0554	0.0615	0.0650
5-class	0.0745	0.0821	0.0868
6-class	0.0859	0.0947	0.1004
7-class	0.0961	0.1062	0.1131

Table 3. (a) Estimated mean $Ee(S)$ (relation 6) and (b) estimated standard deviation $Se(S)$ (relation 7) for situation 2

(a)

threshold	150	200	250
semi-theor.	0.0073	0.0059	0.0054
1 path	0.0333	0.0280	0.0276
1 class	-0.1777	-0.0581	-0.0124
2-class	-0.0202	-0.0106	0.0003
3-class	-0.0010	0.0011	0.0059
4-class	0.0098	0.0080	0.0075
5-class	0.0138	0.0105	0.0099
6-class	0.0159	0.0120	0.0113
7-class	0.0186	0.0143	0.0136

(b)

threshold	150	200	250
semi-theor.	0.0413	0.0460	0.0500
1 path	0.1602	0.1649	0.1769
1 class	0.2232	0.2797	0.3007
2-class	0.1755	0.1651	0.1642
3-class	0.1090	0.1154	0.1169
4-class	0.0708	0.0744	0.0769
5-class	0.0821	0.0876	0.0913
6-class	0.0893	0.0950	0.0995
7-class	0.0984	0.1043	0.1099

three thresholds. The estimated prognosis values $T_{n,\widetilde{\theta}_{\widetilde{z_n}}}(S)$ have been compared with the theoretical value $T_{n,\theta_{z_n}}(S)$. The mean error $Ee(S)$ given by relation (6) and $Se(S)$ given by relation (7) have been computed for the 9 cases (described above) of estimated Gamma process parameters and for each threshold. The estimated mean error $Ee(S)$ and the estimated standard deviation $Se(S)$ for situation 1, obtained using 200 path sets, are given in tables 2(a) and (b). For situation 2 they are given in tables 3(a) and (b).

As expected, the estimated mean error is close to 0. For both situations, the worse result is obtained with the "estimated 1-class case". From estimated standard deviation point of view, the closest case to the "semi-theoretical case" is the "estimated 4-class case". It corresponds to the theoretical number of classes and to the expected result. As one could expect, results in situation 2 are worse than in situation 1 because the classes are more similar. Consequently the number of misclassified paths is larger than in situation 1 and the Gamma process parameters are estimated with a larger error. For both situations, when the number of classes is larger than the theoretical one, the impact is not very important. On the contrary when the number of classes is smaller than the theoretical one, some paths from different Gamma process are mixed and the parameters are not estimated correctly and consequently the prognosis error can be important.

6. CONCLUSION

In this paper, a method is proposed for making a degradation prognosis based on Gamma process model parameters that are estimated using degradation measurements on different systems. It is assumed that there are a number of operational conditions leading to different degradation processes. Estimating the Gamma process model parameters using only one system leads to poor results due to the limited number of samples. On the contrary, estimating the Gamma process model parameters considering only one Gamma process leads to poor results due to the mixture of systems with different degradation trends.

The proposed method consists in considering a mixture of Gamma process models. It allows to cluster the degradation paths in classes corresponding to the different degradation trends and to estimate the Gamma process parameters. It uses an expectation-minimization approach that takes into consideration that all measurements in a same path belong to the same class.

Simulations have been done and demonstrate the feasibility of the method. They have shown that grouping paths originating from the same process allows to really increase the prognosis performance in comparison with the two basic strategies (all paths in one class, one class per path). The best result has been obtained with the class number equal to the theoretical one ; however if the number of classes is sur-estimated the result evolves slowly. A method for choosing the number of classes, using the Bayesian information criterion (Kass & Raftery, 1995), is currently being studied.

REFERENCES

Ambroise, C., & Govaert, G. (1998). Convergence of an EM-type algorithm for spatial clustering. *Pattern Recognition Letters*, *19*, 919–927.

Celeux, G., & Govaert, G. (1992). A classification EM algorithm for clustering and two stochastic versions. *Computational statistics & Data analysis*, *14*(3), 315–332.

Celeux, G., & Govaert, G. (1995). Gaussian parcimonious clustering models. *Pattern Recognition*, *28*, 781–793.

Cinlar, E., Osman, E., & Bazant, Z. P. (1977). Stochastic process for extrapolating concrete creep. *Journal of the Engineering Mechanics Division*, *103*(6), 1069–1088.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society ,Series B*, *39*(1), 1–38.

Hu, T., & Sung, S. Y. (2006). A hybrid EM approach to spatial clustering. *Computational Statistics & Data Analysis*, *50*, 1188–1205.

Kass, R., & Raftery, A. (1995). Bayes factor. *J. Am. Statistical Association*, *90*, 733-795.

Nystad, B. H., Gola, G., & Hulsund, J. (2012, July 3-5). Lifetime models for remaining useful life estimation with randomly distributed failure thresholds. In *Proc. of first european conference of the prognostics and health management society 2012* (p. 141-147). Dresden, Germany.

Ramasso, E., & Gouriveau, R. (2014). Remaining useful life estimation by classification of predictions based on a neuro-fuzzy system and theory of belief functions. *Reliability, IEEE Transactions on*(99).

Saxena, A., Celaya, J., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for offline evaluation of prognostics performance. *International Journal of Prognostics and Health Management (IJPHM)*, *1*(1), 20.

Shental, N., Bar-Hillel, A., Hertz, T., & Weinshall, D. (2003, August 21-24). Computing gaussian mixture models with EM using side-information. In *Proc. of the 20th international conference on machine learning*. Washington DC, USA.

Si, X.-S., Wang, W., Hu, C.-H., & Zhou, D.-H. (2011). Remaining useful life estimation – a review on the statistical data driven approaches. *European Journal of Operational Research*, *213*(1), 1 - 14.

Van Noortwijk, J. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, *94*(1), 2–21.