

# Apprentissage partiellement supervisé dans les modèles de Markov cachés autoregressifs pour le monitoring et le pronostic de systèmes et structures mécaniques

## Partially supervised learning of ARHMM for structural health monitoring and prognostics

Emmanuel Ramasso

Vincent Placet

Institut FEMTO-ST, Univ. Bourgogne Franche-Comté, CNRS  
15B avenue des Montboucons, 25030 Besançon, cedex, France.

emmanuel.ramasso@femto-st.fr

vincent.placet@univ-fcomte.fr

### Résumé :

Une modification des mécanismes d'inférence et d'apprentissage dans les modèles de Markov cachés autorégressifs est proposée pour prendre en compte des connaissances a priori. Ces modèles sont particulièrement intéressants pour représenter statistiquement des séries temporelles. Le fait de pouvoir ajouter des connaissances a priori permet de l'utiliser dans de nombreuses applications. Deux applications sont présentées : une première sur le pronostic avec l'estimation du temps restant avant défaillance de turboréacteurs ; et une seconde sur l'analyse d'ondes élastiques transitoires hautes fréquences rencontrées dans les structures mécaniques.

### Mots-clés :

Monitoring de structures, pronostic, émissions acoustiques, séries temporelles

### Abstract:

A modification of the mechanisms of inference and learning in autoregressive hidden Markov models is proposed in order to take into account prior knowledge. These models are particularly interesting for statistical modelling of time series. Being able to add prior knowledge makes it possible to use it in many applications. Two applications are presented : a first on prognostics with the estimation of the remaining useful life before the failure of turbofan engines ; and a second on transient elastic wave analysis made of high frequencies content and encountered in mechanical structures.

### Keywords:

Structural Health Monitoring, Prognostics, Acoustic emission, Time-series

## 1 Introduction

Les modèles autorégressifs (AR) se sont révélés appropriés pour la représentation statistique de séries chronologiques dans divers domaines. En SHM par exemple (*SHM : Structural Health*

*Monitoring*), ce type de modèle est très souvent utilisé pour la détection de défaillances dans les structures mécaniques dans le domaine aéronautique [5], les engrenages [21], les portes d'ascenseur [22], les roulements à billes [18] ou encore les batteries [16]. [11] utilisent des modèles ARHMM, couplant un modèle AR avec un modèle de Markov caché (*HMM : Hidden Markov Model*), pour analyser des séries temporelles représentant des signes vitaux (multivariés) avec l'objectif de détecter le risque de mortalité dans des unités de soins de patients recevant des traitements particuliers. Les auteurs pointent l'intérêt de ce type de modèle (ARHMM) pour analyser des données issues de systèmes non stationnaires avec différents régimes de fonctionnement. Cette caractéristique est aussi exploitée par [1] où les auteurs sont intéressés de représenter statistiquement l'évolution temporelle des caractéristiques des vents à proximité d'éoliennes.

La contribution porte sur la modification du critère utilisé pour l'apprentissage des paramètres d'un ARHMM afin d'inclure des a priori sur la chaîne de Markov dans le but de rectifier la loi a posteriori. Ce type de modèle se situe dans la lignée de travaux initiés par Vannoorenberghé et Smets [19], Côme et al. [3] et

Denœux [4]. Pour les séries temporelles, les travaux proposés dans [15] se concentrent sur les HMM, c'est à dire un cas particulier du modèle présenté dans cet article. Le modèle ARHMM avec a priori sera appelé ARWHMM par la suite pour *AutoRegressive Weakly Hidden Markov Model*. Alors que l'ARHMM a été principalement utilisée pour la détection et le diagnostic, nous proposons dans cet article d'exploiter les ARWHMM pour le pronostic. Nous proposons par ailleurs d'exploiter ce modèle pour la détection d'anomalie ce qui n'a pas encore été envisagé dans la littérature. Les premiers résultats sont présentés dans la dernière section.

## 2 Apprentissage avec états incertains dans les ARWHMM

Une mesure au temps  $t$  fournie par un capteur est notée  $x_t$  et peut être la résultante d'une somme pondérée des mesures précédentes plus un terme d'erreur, où les poids sont définis conditionnellement à chaque état d'une chaîne de Markov :

$$x_t = - \sum_{\delta=1}^{\Delta} r_{\delta}(y_t) x_{t-\delta} + \varepsilon_t(y_t), \quad 1 \leq t \leq T \quad (1)$$

pour une séquence de  $T$  observations. Le terme de bruit

$$\varepsilon_t(y_t) \sim N(0, \Sigma_{y_t})$$

est supposé être une gaussienne de moyenne nulle et de variance  $\Sigma_{y_t}$  ajustée automatiquement pour chaque état caché dans la phase d'apprentissage.

Les coefficients AR pour l'état  $i$  sont notés  $r_{\delta}(y_t = i)$  où  $\delta = 1 \dots \Delta$  est le décalage temporel. L'ensemble des coefficients AR est donné par :

$$\mathbf{B}_i = \left( r_1(i), \dots, r_{\delta}(i), \dots, r_{\Delta}(i) \right) \quad (2)$$

La commutation entre états est régie par un processus stochastique prenant la forme d'une chaîne de Markov (Figure 1). Ce processus est

caractérisé par une matrice de transition  $\mathbf{A}$  dont un élément  $a_{ij} = p(y_t = j | y_{t-1} = i)$  représente la probabilité d'être dans l'état  $j$  à l'instant  $t$  étant donné que l'état était  $i$  à  $t - 1$ . La probabilité a priori de la chaîne est notée  $\mathbf{\Pi} = [\pi_1 \dots \pi_K]$ , où  $\pi_i$  est la probabilité de être dans l'état  $i$  à l'instant  $t = 1$ .

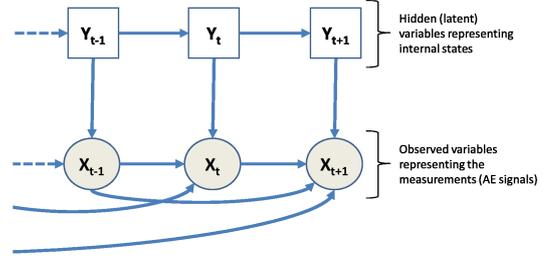


Figure 1 – Modèle graphique d'un ARHMM : les cercles représentent les variables continues observées, les carrés les variables à valeur entière. Tirée de [9].

Le problème est d'estimer les paramètres

$$\lambda = (\mathbf{A}, \mathbf{\Pi}, \mathbf{B}_i, \Sigma_i), \quad 1 \leq i \leq K \quad (3)$$

en présence d'informations incertaines et imprécises sur les variables cachées. Ces informations sont supposées prendre la forme d'un jeu de masses de croyances comme proposé dans E2M [4]. E2M permet de couvrir les cas non supervisé, supervisé, semi-supervisé et avec "labels bruités". Les a priori sont supposés être de la forme suivante :

$$\mathbf{W} = [\mathbf{w}_1; \dots; \mathbf{w}_t; \dots; \mathbf{w}_T]$$

avec  $\mathbf{w}_t = [w_t(1), \dots, w_t(i), \dots, w_t(K)]$  et  $w_t(i) \geq 0$  [8] :

- Si  $w_t(i) = 1$  pour un état  $i$  et  $w_t(j) = 0, j \neq i$  alors on retrouve le cas supervisé ;
- Si  $\forall t, \forall i, w_t(i) = 0$ , alors on retrouve le cas non supervisé.

Ces a priori peuvent prendre la forme de plausibilités pour faciliter leur interprétation.

Les paramètres de l'ARHMM sont optimisés

avec E2M où la fonction auxiliaire est

$$\begin{aligned} Q(\lambda, \lambda^{(q)}) &= \mathbb{E}_{\lambda^{(q)}} [\log(L(\lambda; \mathbf{Z})) | \mathbf{X}] \\ &= \sum_{\mathbf{Y}} p(\mathbf{Y} | \mathbf{X}, \lambda^{(q)}) \log L(\lambda; \mathbf{Z}) \end{aligned} \quad (4)$$

où  $(q)$  est une itération de E2M, et  $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ . La vraisemblance des données complètes est donnée par :

$$\begin{aligned} L(\lambda; \mathbf{Z}) &= p(y_1; \Pi) \left( \prod_{t=2}^T p(y_t | y_{t-1}; \mathbf{A}) \right) \times \\ &\quad \prod_{t=1}^T p(\mathbf{x}_t | y_t; \mathbf{r}_\Delta(y_t), \Sigma_{(y_t)}) \end{aligned} \quad (5)$$

Donc la fonction auxiliaire devient

$$Q(\lambda, \lambda^{(q)}) = \frac{\sum_{\mathbf{y}} w(\mathbf{y}) p(\mathbf{y} | \mathbf{x}, \lambda^{(q)}) \log L(\lambda; \mathbf{z})}{\mathbb{E}_\lambda [w(\mathbf{y})]}$$

On peut alors développer  $Q$ , dériver l'expression par rapport aux paramètres afin d'obtenir l'expression des nouveaux paramètres pour l'itération  $q + 1$ . Pour la chaîne de Markov :

$$\pi_i^{(q+1)} = \gamma_{1i}^{(q)} \quad (6a)$$

$$a_{ij}^{(q+1)} = \frac{\sum_{t=2}^T \xi_{t-1,t,i,j}^{(q)}}{\sum_{t=2}^T \sum_{l=1}^K \xi_{t-1,t,i,l}^{(q)}}, \quad (6b)$$

Les expressions de probabilités a posteriori  $\gamma$  et  $\xi$  en présence d'informations incertaines peuvent être retrouvées avec la même démarche que celle présentée dans [15] et en utilisant une version modifiée de l'algorithme *forward-backward*.

Pour le modèle d'observation, la covariance du bruit est donnée par [2] :

$$\begin{aligned} \Sigma_i^{(q+1)} &= \frac{1}{\sum_{t=1}^T \gamma_{ti}^{(q)}} \sum_{t=1}^T \gamma_{ti}^{(q)} \left[ \mathbf{x}_t + \sum_{\delta=1}^{\Delta} \mathbf{r}_\delta^{(q)}(i) \right. \\ &\quad \left. \mathbf{x}_{t-\delta} \right] \left[ \mathbf{x}_t + \sum_{\delta=1}^{\Delta} \mathbf{r}_\delta^{(q)}(i) \mathbf{x}_{t-\delta} \right]^T, \end{aligned}$$

et l'expression des coefficients  $\mathbf{B}_i^{(q+1)} = \left( \mathbf{r}_1^{(q+1)}(i), \dots, \mathbf{r}_\delta^{(q+1)}(i), \dots, \mathbf{r}_\Delta^{(q+1)}(i) \right)$  des

processus AR est donnée par

$$- \left[ \sum_{t=1}^T \gamma_{ti}^{(q)} \mathbf{x}_t \mathbf{u}_{t-1}^T \right] \left[ \sum_{t=1}^T \gamma_{ti}^{(q)} \mathbf{u}_{t-1} \mathbf{u}_{t-1}^T \right] \quad (7)$$

avec

$$\mathbf{u}_{t-1} = \left( \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{t-\Delta} \right)^T, \quad (8)$$

et la vraisemblance  $b_i(\mathbf{x}_t)$  (pour un état caché  $i$ ) est

$$b_i(\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_t + \sum_{\delta=1}^{\Delta} \mathbf{r}_\delta(i) \mathbf{x}_{t-\delta} | 0, \Sigma_i) \quad (9)$$

La passe *forward* permet d'évaluer la vraisemblance des données comme dans les modèles HMM standards [15] :

$$\alpha_{1i}^{(q)} = \pi_i^{(q)} w_{1i} b_i(\mathbf{x}_1), \quad (10a)$$

$$\alpha_{t,j}^{(q)} = b_j(\mathbf{x}_t) \underbrace{w_{tj}}_{a \text{ priori}} \sum_i \alpha_{t-1,i}^{(q)} a_{ij}^{(q)} \quad (10b)$$

avec

$$L(\lambda^{(q)}; \mathbf{X}, \mathbf{W}) = \sum_{i=1}^K \alpha_{Ti} \quad (11)$$

### 3 ARWHMM pour le pronostic

L'évaluation de l'état de santé d'un système en service peut être effectuée en déduisant l'état caché pour chaque observation. Pour cela on peut utiliser une passe *forward*. Cette passe peut d'ailleurs prendre en compte des a priori si ceux-ci sont disponibles pendant l'inférence, par exemple via des modèles physiques ou d'autres sources d'informations.

La durée de vie résiduelle peut être estimée de différentes manières. L'une d'entre elles consiste à entraîner le modèle pour déduire le temps restant à partir de données d'historique. Le temps restant est alors une cible pour la partie AR du modèle ARWHMM. C'est l'approche utilisée dans [9] que l'on appelle aussi "prédiction directe" [13]. Il suffit alors, en service, d'utiliser les données mesurées pour alimenter le modèle, qui déduit le temps restant.

Progressivement, on dispose de prédictions sur le temps restant, à partir desquelles on peut estimer une distribution que l'on peut mettre à jour à chaque mesure.

Lorsque les données sont relativement complexes, une approche consiste à construire un modèle par instance d'apprentissage et de réaliser l'inférence sur chaque modèle, suivi d'un tri des modèles en fonction de leur pertinence (qui peut être évaluée par l'estimation de la vraisemblance). Les modèles les plus pertinents sont ensuite combinés pour déduire le temps restant accompagné d'un intervalle de confiance. C'est le type d'approche utilisé dans [20] par exemple ou dans [12].

Une illustration est donnée ci-après. Elle porte sur des données simulées à partir d'un modèle de turboréacteurs développé à la NASA [6, 17]. Ce modèle prend en compte de nombreux paramètres et permet d'introduire des défaillances dans les différents sous-systèmes. Les jeux de données contiennent des mesures issues d'une vingtaine de capteurs (dont les mesures ont été bruitées) mais rien n'est spécifié concernant l'état initial du système ni le nombre ou les types de dégradation. Ces jeux de données ont été utilisés depuis 2008 par de nombreux chercheurs en pronostic (voir la synthèse dans [14]).

La figure 2 montre les indicateurs pour 4 jeux de données (avec différentes conditions de fonctionnement des turboréacteurs). Ces indicateurs synthétisent l'information des capteurs pour refléter l'état de santé. La méthode a été proposée dans [12], qui a donné des résultats de l'état de l'art sur plusieurs des jeux de données. Ces jeux montrent la variabilité sur le temps restant et sur l'état initial.

La figure 3 montre une estimation du temps restant (RUL : *Remaining Useful Life*) en fonction de la quantité d'a priori pour une des instances du jeu de données #1. Les a priori consistent à décomposer la dégradation en  $K$  paliers ( $K$  étant le nombre d'états dans les ARWHMM). Le cadre de discernement des a

priori représente donc ici des variables internes du turboréacteur à l'origine des observations. Cette figure montre que plus on augmente le nombre d'états, meilleur est l'estimation du temps restant. La contrepartie étant la capacité à généraliser sur de nombreux turboréacteurs, en particulier sur des cas complexes mêlant conditions de fonctionnement et plusieurs sources de défaillances. Dans [9], nous montrons par exemple que le modèle ARWHMM donne au moins d'aussi bons voire de meilleurs résultats que des réseaux de neurones. Les performances restent inférieures à [12] qui comporte très peu de paramètres lui permettant, de généraliser beaucoup mieux que la plupart des méthodes de la littérature.

L'intérêt des ARWHMM dans ce type d'application est en premier lieu dans la possibilité d'intégrer des a priori pendant l'apprentissage mais aussi et surtout pendant l'inférence (lorsque le système est en service). En effet, ceci permettra d'exploiter des modèles inverses issus de la physique pour, par exemple, améliorer le pronostic. L'autre intérêt est dans le fait qu'il s'agit d'un modèle avec un potentiel applicatif très large en monitoring.

## 4 ARWHMM pour la représentation d'ondes élastiques transitoires

Dans cette section, nous illustrons la méthode sur un autre type de données. Il s'agit de signaux transitoires générés par des endommagements dans des matériaux. Les endommagements libèrent en effet de l'énergie, dont une partie prend la forme d'ondes élastiques, se propageant à la surface du matériau et que l'on peut collecter par des capteurs piézo-électriques.

Les capteurs ont une empreinte importante sur les données qu'il est difficile de déconvoluer. De plus, les ondes, lors de la propagation, depuis l'endommagement jusqu'au capteur, subissent des distorsions importantes dues aux propriétés mécaniques du matériau, à sa

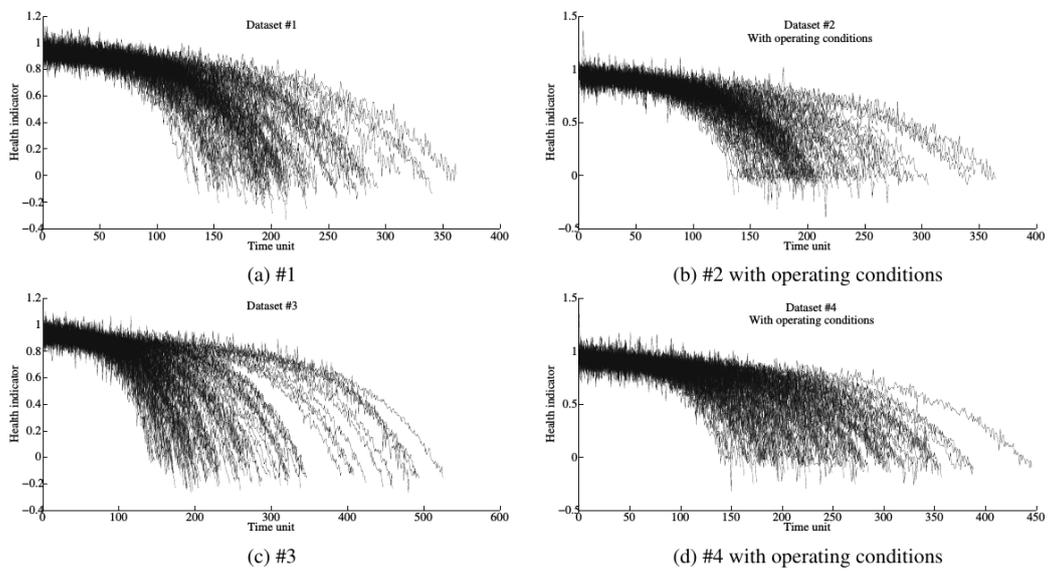


Figure 2 – Les jeux de données pour 4 conditions de fonctionnement.

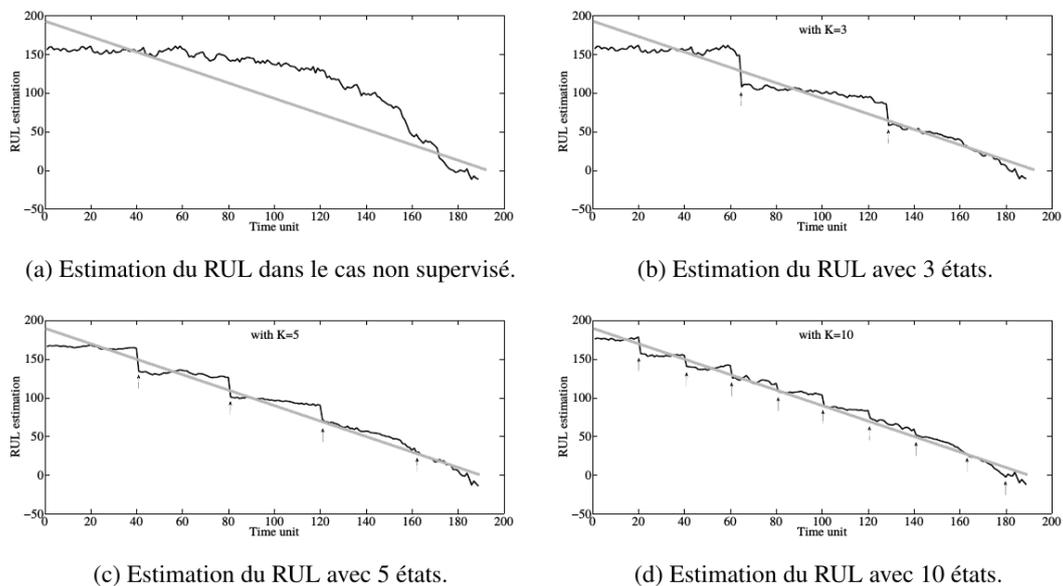


Figure 3 – Estimation du RUL pour différentes configurations de la structure latente dans le ARWHMM.

géométrie ou à l'endommagement cumulé. Il est donc difficile de déduire l'endommagement à partir de la simple mesure du signal sur le capteur piézo-électrique.

Une démarche synthétisée dans [13] et initiée à FEMTO-ST depuis 2012 dans le cadre du Labex "ACTION", consiste à mêler physique et statistique pour construire des modèles inverses plus précis, fiables et robustes pour le monitoring de structures. Le ARWHMM permet, comme vu précédemment, d'intégrer des connaissances a priori, possiblement issues de la physique. Il s'agit d'un modèle très intéressant pour le monitoring partiellement supervisé.

La figure 4 montre un signal transitoire, et son spectre, collecté sur une structure composite tubulaire utilisée dans des rotors à grande vitesse (extrait de [10] à partir de données issues de la thèse de [7]). Ce signal fait partie d'un ensemble de plusieurs centaines de milliers de signaux collectés pendant un essai de caractérisation. Le spectre du signal prédit à partir d'un ARWHMM (3 états, 10 prédécesseurs) est superposé sur la figure de droite. Sans chercher à optimiser le modèle, le ARWHMM prédit très bien l'évolution d'un signal transitoire complexe.

Etant donné que les sources de signaux ne sont encore pas bien connues, une étape fondamentale est de créer des groupes de signaux avec des caractéristiques similaires, puis d'analyser ces groupes a posteriori. Des méthodes de partitionnement sont pour cela utilisées. Elles sont appliquées, comme cela est fait en reconnaissance de formes, sur des descripteurs extraits à partir des signaux. Le ARWHMM est actuellement exploité pour représenter statistiquement les signaux transitoires afin de ne pas utiliser de descripteurs. Il n'y a en effet pas vraiment de consensus concernant le choix des descripteurs et ce qui amène parfois un biais d'interprétation. Le ARWHMM permet ainsi de représenter un signal transitoire à l'aide de seulement quelques paramètres.

L'intérêt est aussi de réduire la quantité d'information à stocker (les signaux sont généralement échantillonnés à plusieurs méga-Hertz) en ne conservant que les paramètres du modèle. Par exemple sur la figure 4, environ 50 paramètres sont utilisés pour représenter un signal de 1500 points quasiment parfaitement.

## 5 Conclusion

La modification apportée pour introduire des a priori dans des ARHMM repose sur l'algorithme E2M et permet de rectifier la distribution a posteriori sur les variables latentes au cours du temps, à la fois dans la phase d'apprentissage et de test. Le modèle résultant représente une solution intéressante pour différents problèmes liés au monitoring et pronostic de structures et systèmes mécaniques. Les travaux actuels portent sur le couplage entre reconnaissance de formes et physique de l'endommagement pour rendre ces modèles plus pertinents.

### Remerciements :

Ce travail a été financé par l'*EIPHI Graduate school* (contrat ANR-17-EURE-0002).

## Références

- [1] Pierre Ailliot and Valerie Monbet. Markov-switching autoregressive models for wind time series. *Environmental Modelling & Software*, 30 :92 – 101, 2012.
- [2] Joyce Chiang, Z Jane Wang, and Martin J McKeown. A hidden markov, multivariate autoregressive (hmm-mar) network framework for analysis of surface emg (semg) data. *IEEE Transactions on Signal Processing*, 56(8) :4069–4081, 2008.
- [3] E. Côme, L. Oukhellou, T. Denoeux, and P. Aknin. Learning from partially supervised data using mixture models and belief functions. *Pattern recognition*, 42(3) :334–348, 2009.
- [4] T. Denoeux. Maximum likelihood estimation from uncertain data in the belief function framework. *Knowledge and*

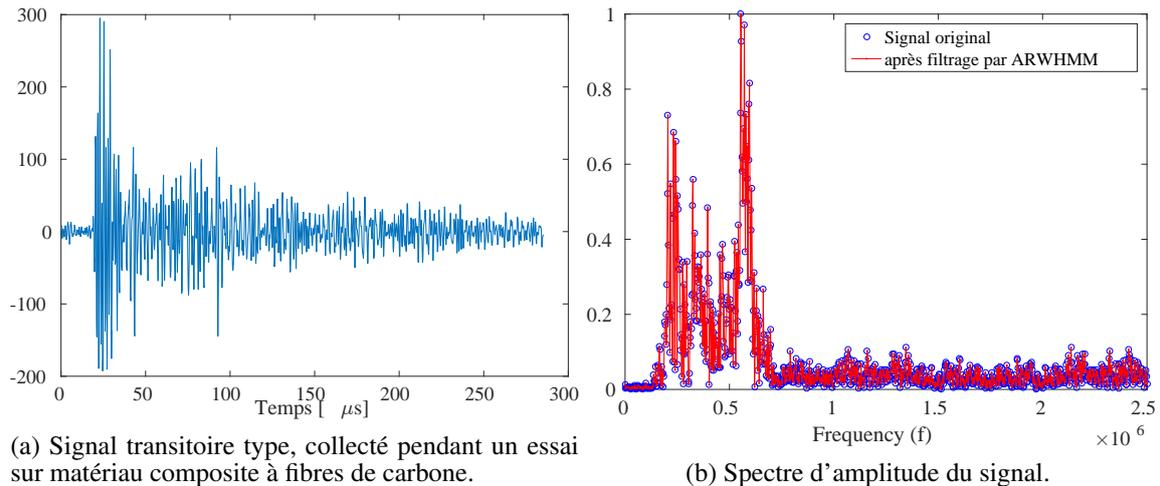


Figure 4 – Un signal type d’émission acoustique.

- Data Engineering, IEEE Transactions on*, 25(1) :119–130, 2013.
- [5] C.R. Farrar and K. Worden. *Structural Health Monitoring : A Machine Learning Perspective*. John Wiley & Sons, Ltd, 2013.
- [6] D.K. Frederick, J.A. DeCastro, and J.S. Litt. User’s guide for the commercial modular aero-propulsion system simulation (C-MAPSS). Technical report, National Aeronautics and Space Administration (NASA), Glenn Research Center, Cleveland, Ohio 44135, USA, 2007.
- [7] X. Gabrion. *Contributions à la caractérisation d’un matériau composite thermoplastique thermostable : Application à des structures cylindriques sous sollicitations multiaxiales*. PhD thesis, Univ. de Franche-Comté, 2014.
- [8] P. Juesas and E. Ramasso. Ascertainment-adjusted parameter estimation approach to improve robustness against misspecification of health monitoring methods. *Mechanical Systems and Signal Processing*, 81 :387–401, 2016.
- [9] Pablo Juesas, Emmanuel Ramasso, Sébastien Drujon, and Vincent Placet. On partially supervised learning and inference in dynamic bayesian networks for prognostics with uncertain factual evidence : Illustration with markov switching model. *Int. Journal on Prognostics and Health Management*, 7(51), 2016.
- [10] M. Kharrat, E. Ramasso, V. Placet, and M.L. Boubakar. A signal processing approach for enhanced acoustic emission data analysis in high activity systems : Application to organic matrix composites. *Mechanical Systems and Signal Processing*, 70 :1038–1055, 2016.
- [11] L.W.H. Lehman, S. Nemati, and R.G. Mark. Hemodynamic monitoring using switching autoregressive dynamics of multivariate vital sign time series. In *Computing in Cardiology*, Nice, France, 2015.
- [12] E. Ramasso. Investigating computational geometry for failure prognostics. *Int. Journal on Prognostics and Health Management*, 5(5) :1–18, 2014.
- [13] E. Ramasso. Méthodologies d’analyse de séries temporelles sous incertitudes aléatoires et épistémiques pour le suivi et le pronostic de l’état de systèmes et structures : De l’estimation d’une cinétique d’endommagement à son contrôle. Technical report, HDR, Univ. Bourgogne Franche Comté, 2016.
- [14] E. Ramasso and A. Saxena. Performance benchmarking and analysis of prognostic methods for CMAPSS datasets. *Internation-*

*tional Journal on Prognostics and Health Management*, 5(2) :1–15, 2014.

- [15] Emmanuel Ramasso and Thierry Denoeux. Making use of partial knowledge about hidden states in hmms : an approach based on belief functions. *IEEE Transactions on Fuzzy Systems*, 22(2) :395–405, 2014.
- [16] B. Saha, K. Goebel, and J. Christoffersen. Comparison of prognostic algorithms for estimating remaining useful life of batteries. In *Transactions of the Institute of Measurement and Control*, 2009.
- [17] A. Saxena, K. Goebel, D. Simon, and N. Eklund. Damage propagation modeling for aircraft engine run-to-failure simulation. In *International Conference on Prognostics and Health Management*, pages 1–9, Denver, CO, USA, 2008. IEEE.
- [18] S. Thanagasundram, S. Spurgeon, and F.S. Schlindwein. A fault detection tool using analysis from an autoregressive model pole trajectory. *Journal of Sound and Vibration*, 317 :975 – 993, 2008.
- [19] P. Vannoorenberghe and P. Smets. Partially supervised learning by a credal EM approach. In Lluís Godo, editor, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 3571 of *Lecture Notes in Computer Science*, pages 956–967. Springer Berlin Heidelberg, 2005.
- [20] T. Wang, J. Yu, D. Siegel, and J. Lee. A similarity-based prognostics approach for remaining useful life estimation of engineered systems. In *Int. Conf. on Prognostics and Health Management*, pages 1–6, 2008.
- [21] W. Wang and A.K. Wong. Autoregressive model-based gear fault diagnosis. *Journal of vibration and acoustics*, 124(2) :172–179, 2002.
- [22] Jihong Yan, Muammer Koç, and Jay Lee. A prognostic algorithm for machine

performance assessment and its application. *Production Planning & Control*, 15(8) :796–801, 2004.