



**HAL**  
open science

# Quantité minimale d'observations pour l'estimation de Modèles de Markov Cachés, dans le cadre d'une politique de maintenance préventive

Bernard Roblès, Manuel Avila, Florent Duculty, Pascal Vrignat, Stéphane  
Begot, Frédéric Kratz

► **To cite this version:**

Bernard Roblès, Manuel Avila, Florent Duculty, Pascal Vrignat, Stéphane Begot, et al.. Quantité minimale d'observations pour l'estimation de Modèles de Markov Cachés, dans le cadre d'une politique de maintenance préventive. 5 ièmes Journées Doctorales / Journées Nationales MACS, Jul 2013, Strasbourg, France. pp.1-6. hal-00844512

**HAL Id: hal-00844512**

**<https://hal.science/hal-00844512>**

Submitted on 15 Jul 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Quantité minimale d'observations pour l'estimation de Modèles de Markov Cachés, dans le cadre d'une politique de maintenance préventive

Bernard ROBLÈS<sup>1</sup>, Manuel AVILA<sup>1</sup>, Florent DUCULTY<sup>1</sup>, Pascal VRIGNAT<sup>1</sup>, Stéphane BÉGOT<sup>1</sup>,  
Frédéric KRATZ<sup>2</sup>

<sup>1</sup>Laboratoire PRISME

EA 4229 – Université d'Orléans

Château de La Source, Avenue du Parc Floral, BP 6749, 45067 Orléans cedex 2, France.

*Bernard.Robles@univ-orleans.fr, Manuel.Avila@univ-orleans.fr, Florent.Duculty@univ-orleans.fr,*

*Pascal.Vrignat@univ-orleans.fr, Stephane.Begot@univ-orleans.fr*

<sup>2</sup>Laboratoire PRISME,

EA 4229 – Université d'Orléans

ENSI, 88 boulevard Lahitolle 18020 Bourges Cedex, France.

*Frederic.Kratz@ensi-bourges.fr*

**Résumé**— Dans cet article, nous souhaitons évaluer la quantité minimale d'informations nécessaire pour l'estimation de modèles stochastiques, afin d'améliorer la conception d'une politique de maintenance préventive. Notre étude est basée sur des données fournies par un Modèle de Markov Caché (MMC) de synthèse. Ce modèle synthétique est destiné à produire des observations (ou « symboles ») utilisées dans le cadre d'une politique de maintenance préventive, ainsi qu'un indicateur de dégradation correspondant. Nous représentons ces événements chronologiques sous la forme de chaînes de Markov cachées d'ordre 1, également appelés « signatures ». La production de ces symboles est réalisée en utilisant des distributions uniformes ou normales. Nous mesurons ensuite la pertinence des paramètres du MMC à l'aide de l'entropie de Shannon. Les résultats montrent que l'on peut utiliser une quantité minimale d'informations pour décrire de manière optimale des modèles. Après une discussion sur l'utilisation d'une nouvelle « fenêtre glissante » de symboles utilisables dans un système de gestion de maintenance assistée par ordinateur, nous comparons l'ensemble des résultats avec des données réelles issues d'un processus continu.

**Mots-clés**— Modèle de Markov Caché, entropie de Shannon, sélection de modèles, maintenance préventive.

## I. INTRODUCTION

Dans cette étude, nous utilisons des MMC<sup>1</sup> pour la détection de panne, dans le cadre d'une politique de maintenance. Pour disposer d'une « vérité terrain », nous avons construit un *modèle synthétique* basé sur un *modèle de référence* [1]. Nous avons déjà montré dans [1] que ce modèle de référence est celui qui se rapproche le plus de la réalité. Il fournit également de bonnes prédictions dans la dégradation globale d'un processus. Ce *modèle synthétique* produit des observations ou symboles, couramment utilisés dans une GMAO<sup>2</sup> industrielle. Il produit également des indicateurs de dégradation : S1, S2, S3 et S4 représentés dans la Figure 1. Nous avons précédemment trouvé dans [2], une topologie optimale MMC pour une utilisation en maintenance industrielle (voir Figure 2).

1. Modèles de Markov Cachés

2. Gestion de la Maintenance Assistée par Ordinateur

De nombreux papiers du domaine utilisent le principe du *minimum Cross-Entropy* ([3] et [4]) pour l'optimisation des problèmes stochastiques. En effet, cette méthode introduite par [5], est généralement appliquée aux problèmes d'optimisation combinatoire (comme par exemple le problème du voyageur de commerce).

Dans ce papier, nous essayons de trouver la quantité minimale de données nécessaires pour estimer des modèles de manière optimale. Nous essayons ainsi de trouver une valeur maximale d'entropie ou une valeur de convergence. Cette valeur, si elle existe, nous donnera la taille minimale de la fenêtre des observations à utiliser pour l'estimation du modèle. Nous avons utilisé deux distributions différentes pour la production des symboles (distributions uniformes ou normales). Le critère de l'entropie de Shannon (un critère habituellement utilisé dans la sélection de modèle) nous a permis d'évaluer la pertinence des observations étudiées.

La structure de cette étude est la suivante : en Section II, nous décrivons succinctement les MMC, la méthode utilisée pour évaluer la pertinence des paramètres d'un modèle ainsi que le processus d'évaluation. En Section III, nous détaillons la conception du modèle synthétique. Les résultats sont donnés dans la Section IV et discutés en Section V. Pour terminer, nous présentons une application industrielle et comparons les résultats obtenus au § V-C.

Le but de cette étude est de trouver la quantité minimale d'informations nécessaire afin de garantir la pertinence des modèles. Nous souhaitons montrer que notre choix d'une taille minimale de données, pourrait permettre l'amélioration de la maintenance préventive et la prédiction de panne, dans un contexte réel.

## II. METHODOLOGIE

### A. Modèle de Markov Caché

Dans cette étude, nous avons choisi d'utiliser les MMC pour décrire des événements de maintenance industrielle.

Un Modèle de Markov Caché est composé d'un processus stochastique caché modélisé par une chaîne de Markov et un processus stochastiques observé dépendant des états du processus caché [6]. Selon Rabiner [7], Fox [8], c'est un automate à états cachés qui est constitué d'une variable non observable. Celle-ci représente l'état du système à modéliser. Seule la variable de sortie est observable. Cela nous permet d'avoir une séquence d'observations en sortie de l'automate (voir exemple d'une topologie Figure 1).

Un MMC est caractérisé par :

- le nombre d'états ;
- le nombre de symboles émis par état, les symboles correspondent à la sortie physique du système à modéliser ;
- la distribution de probabilité des transitions entre états ;
- la distribution de probabilité des symboles d'observations ;
- la distribution des probabilités de l'état initial.

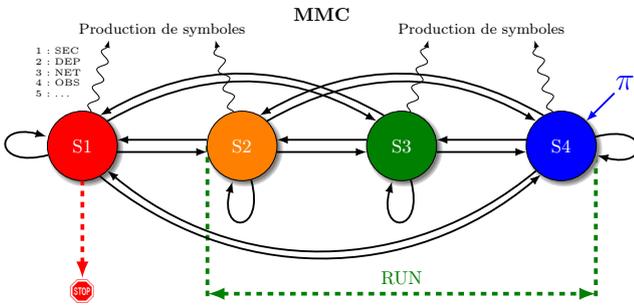


Fig. 1. Modèle de Markov Caché à quatre états discrets.

## B. Hypothèse de Markov

La prédiction de l'état futur n'est pas rendue plus précise par connaissance supplémentaire d'information a priori i.e. toute l'information utile pour l'estimation à l'instant  $t$  est contenue dans l'état présent du processus.

$$P(X_{n+1} = j | X_0, X_1, \dots, X_n = i) = P(X_{n+1} = j | X_n = i). \quad (1)$$

### B.1 Définitions formelles d'un Modèle de Markov Caché à observations discrètes

- Soit  $N$ , le nombre d'états cachés possibles et  $\mathbb{S} = \{s_1, s_2, \dots, s_N\}$ , l'ensemble des valeurs possibles de cette variable. On notera  $q_t$ , la valeur de cette variable à l'instant  $t$  ;
- Le processus ainsi modélisé, doit répondre à l'hypothèse Markovienne d'ordre 1 : l'état à un instant  $t$  ne dépend que de l'état à l'instant  $t - 1$  ;
- Soit  $T$ , le nombre total de symboles d'observations et nous notons  $O = \{o_1, o_2, \dots, o_T\}$ , la séquence d'observations du processus modélisé ;
- Soit  $A = \{a_{ij}\}$ , la distribution de probabilité de la transition d'état avec :

$$a_{ij} = P(q_{t+1} = s_j | q_t = s_i) \quad 1 \leq i, j \leq N, \quad (2)$$

- Soit  $B = \{b_j(k)\}$ , la distribution de probabilité des observations à l'état  $j$ , avec :

$$b_j(k) = P(O_t = o_k | q_t = s_j) \quad 1 \leq j \leq N \quad 1 \leq k \leq T, \quad (3)$$

avec  $O_t$ , la valeur de la variable d'observation à l'instant  $t$ .

- Soit  $\pi = \{\pi_i\}$ , la distribution des probabilités initiales, avec :

$$\pi = P(q_1 = s_i) \quad 1 \leq i \leq N. \quad (4)$$

- Le Modèle de Markov Caché sera noté :

$$\lambda = (A, B, \pi). \quad (5)$$

## B.2 Entropie de Shannon

Nous avons utilisé ce critère dans [2], pour évaluer la pertinence de plusieurs séquences d'observations (une séquence pouvant être assimilée à une signature). Nous avons ainsi pu trouver le modèle le plus pertinent. Dans cet article, nous avons utilisé l'entropie de Shannon sur des chaînes de Markov cachées d'ordre 1 pour les données simulées ainsi que pour des données issues d'un contexte réel (processus continu). L'entropie de Shannon, est une fonction mathématique qui permet de mesurer la quantité d'informations contenue dans une source d'informations. Cette source peut être un texte écrit dans une langue donnée, un signal électrique, un fichier informatique quelconque...

## B.3 Définition de l'entropie

L'entropie de Shannon est définie dans [9] comme suit :

$$H(S) = - \sum_{i=1}^n P_i \log P_i, \quad (6)$$

$P_i$  est la probabilité moyenne de voir apparaître le symbole  $i$  dans  $S$

## B.4 Principes de l'entropie maximale

Les deux principes de maximisation d'entropie [10] sont les suivants :

- principe d'assignation des probabilités à une distribution lorsque nous ne disposons pas d'informations complètes sur elle,
- de toutes les distributions de probabilité qui satisfont les contraintes, choisir celle qui présente l'entropie maximale au sens de Shannon [11].

Venkat [12] utilise ce 2<sup>ème</sup> principe pour la sélection de modèles, ainsi que [13] pour construire des modèles de plus en plus précis simplement en rajoutant de l'information. Notre démarche consiste à comparer la moyenne des entropies des différents modèles. La valeur d'entropie moyenne serait alors maximale pour les modèles les plus pertinents. Notre approche consiste à calculer l'évolution de l'entropie en fonction du nombre de symboles. Le but est de trouver un maximum d'entropie ou une valeur asymptotique d'entropie pour un nombre de symboles finis. Ces valeurs nous donneront ainsi une quantité minimale de données pour estimer correctement un modèle. Nous avons étudié ces calculs d'entropie sur les données du modèle de synthèse issues

des 2 distributions différentes. Nous avons ensuite analysé, d'après la même démarche, les données sur un cas concret.

La convergence ( $L$ ) est définie par la limite de la fonction d'entropie  $H$ , quand  $S$  tend vers  $+\infty$  :

$$\lim_{S \rightarrow +\infty} H(S) = L \quad (7)$$

### C. Processus d'évaluation

Nous avons réalisé un modèle de synthèse (voir § III-A) afin de produire des observations en adéquation avec de réelles observations d'une maintenance industrielle. Ces symboles sont obtenus en utilisant les distributions normales ou uniformes. Le modèle produit parallèlement des états cachés pour la topologie étudiée (S1 à S4).

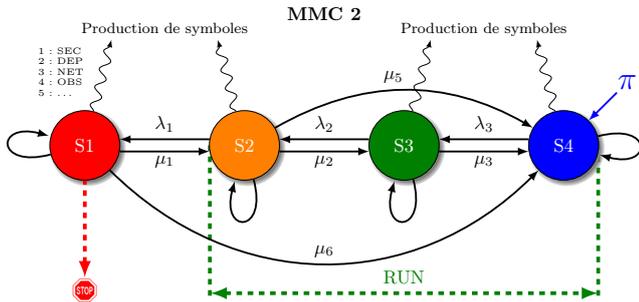


Fig. 2. Modèle de Markov Caché à 4 états. S1 à S4 représentent les niveaux de dégradations du processus.  $\lambda_i$  = taux de panne,  $\mu_i$  = taux de réparation.  $\pi$  est la matrice d'initialisation.

### D. Topologie MMC

Un MMC peut avoir une représentation graphique (Figure 2). La topologie choisie est en adéquation avec l'automate stochastique désiré. Les états représentent les niveaux de dégradation du processus (S1 à S4). Les symboles de l'automate représentent les observations du processus quelconque. {S4, S3, S2} sont les états correspondant à un fonctionnement continu du processus (« RUN »). {S1} est l'état « STOP », lorsque le processus est arrêté (dans notre cas, une situation en dépannage). Nous avons utilisé la topologie orientée modèle 2 (voir Figure 2), représentée par un automate à quatre états.

Un processus Markovien impose d'avoir un découpage temporel régulier. Un échantillonnage de l'information (1 Heure) nous a permis en effet d'utiliser une modélisation markovienne.

## III. SIMULATION INDUSTRIELLE

Les maintenances préventives sont depuis 1996, de plus en plus présentes dans l'industrie. Celles-ci ont permis d'avoir une meilleure maîtrise des arrêts de production non souhaités. Les agents de maintenance consignent leurs actions préventives ou non, (observations). Chaque agent peut rajouter un commentaire particulier sur son action de maintenance spécifique. Ces informations nourrissent une base de données généralement associée à une GMAO (voir exemple dans le Tableau I). Par exemple, les activités de maintenance deviennent des observations consignées « TEP, AU, DEP, ... » et caractérisent les activités

Noms	Date	Opé.	Cd	Ti	N°	Code
Dupond	11/01/2007	Huilage	VEP	20	1	9
Dupond	11/01/2007	Huilage	VEP	20	2	9
Dupond	12/01/2007	Huilage	TEP	30	3	5
Dupond	12/01/2007	Huilage	TEP	30	4	5
Dupond	13/01/2007	Cadenas	SEC	10	5	6
Dupond	13/01/2007	Cadenas	TEP	30	6	5
Dupond	13/01/2007	Cadenas	TEP	30	7	5
Dupond	16/01/2007	Huilage	DEP	90	8	1
Dupond	19/01/2007	Cadenas	AU	10	9	3
...	...	...	...	...	...	...

TABLE I

EXEMPLE DE CONSIGNATION DES ÉVÉNEMENTS.

de maintenance sur un processus réel. Pour pouvoir ensuite travailler avec les algorithmes choisis, ces observations sont ensuite codées avec des symboles subjectifs (9,5,1, ...) (voir Tableau II). Ces symboles définissent les actions de maintenance menées sur le processus. Par exemple, le symbole DEP correspond à un dépannage avec arrêt de la production. C'est un état critique qu'il faut minimiser. « MARCHÉ » et « ARRÊT », Tableau II, correspondent à l'état du processus, respectivement si la production est opérationnelle ou si la production est dans une situation d'arrêt non souhaité.

Etat du processus		
MARCHÉ		
ARRÊT		
N°Obs.	Symboles	Nature des interventions
1	DEP	(Dépannage / arrêt de la production)
2	RM	(Réglage Machine)
3	AU	(Autre)
4	OBS	(Observation)
5	TEP	(Travaux Entretien Préventif)
6	SEC	(Sécurité)
7	RAN	(Remise A Niveau / planifié)
8	NET	(Nettoyage Machine)
9	VEP	(Visite Entretien Préventif)
10	RAS	(pas d'intervention)

TABLE II

CODIFICATION SYMBOLIQUE DES INTERVENTIONS DE MAINTENANCE.

Ces différents états de maintenance sont simulés par notre modèle de synthèse, voir § III-A. Nous choisissons les paramètres des Modèles de Markov Cachés «  $\lambda_i$  » (taux de panne) et «  $\mu_i$  » (taux de réparation), voir Figure 2. Ces paramètres sont déterminés dans [1] afin de correspondre aux données issues d'une véritable GMAO industrielle (voir exemple Tableau I).

### A. Modèle de synthèse

La topologie d'un Modèle de Markov Caché dépend uniquement des éléments non nuls de la matrice de transition notée  $A = \{a_{ij}\}$ . Un tel modèle est dit à « connectivité totale » lorsque sa matrice de transition ne comporte aucun élément nul. Le modèle de synthèse génère des émissions de symboles en utilisant les deux distributions suivantes :

- La distribution uniforme : cette loi permet de modéliser des variables aléatoires uniformément réparties sur un intervalle, comme le montre la Figure 5. La variable aléatoire peut prendre  $n$  valeurs équiprobables possibles  $\{x_1, x_2, \dots, x_n\}$  définie par la probabilité sui-

vante :

$$P(x_i)_{i \in [1,n]} = \frac{1}{n}. \quad (8)$$

Cette loi est peu représentative d'un système réel, car tous les symboles ont le même poids.

- La distribution normale (Laplace – Gauss) : cette loi permet de modéliser de nombreuses répartitions de variables aléatoires. La loi normale est définie par la densité de probabilité  $\varphi : \mathbb{R} \rightarrow \mathbb{R}^+$ , d'espérance  $\mu$  et d'écart type  $\sigma$  :

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (9)$$

Cette loi peut être une bonne représentation d'un système réel, car les symboles ont des poids différents. La matrice  $A$  (probabilité de passage entre les états) est le résultat de calculs issus de l'algorithme d'apprentissage utilisé pour le modèle considéré (topologie désirée a priori). Dans notre situation, ces résultats sont fournis par le modèle de synthèse avec ses caractéristiques topologiques et reflétant une distribution des symboles suivant une loi normale (Figure 4).

Nous avons lancé une première simulation afin de produire 1000 observations. La matrice  $A$  est initialement définie dans [1] par :

$$A = \begin{pmatrix} 0.500 & 0.250 & 0 & 0.250 \\ 0.100 & 0.070 & 0.500 & 0.330 \\ 0 & 0.005 & 0.495 & 0.500 \\ 0 & 0 & 0.001 & 0.999 \end{pmatrix}. \quad (10)$$

Nous avons réalisé un modèle de synthèse sous MATLAB en utilisant le modèle de référence de [1]. Ces réglages numériques peuvent caractériser un contexte réel de dégradation et de réparation. Nous avons ainsi généré une séquence de 1000 symboles. Ces dernières correspondent aux observations que nous pouvons trouver dans une base de données issues d'une GMAO industrielle. Nous avons ainsi construit 12 séquences d'observations (une séquence peut être assimilée à une signature). Chaque séquence est ponctuée par une situation de dépannage (symbole DEP du Tableau II), où le processus est arrêté. Les symboles sont produits en utilisant la loi normale ou la loi uniforme.

Nous considérons ces séquences de symboles comme des chaînes de Markov cachées d'ordre 1 (voir Tableau III) pour modéliser les niveaux de dégradation d'un processus.

TEP	SEC	TEP	TEP	RAS	RAS	DEP	...
-----	-----	-----	-----	-----	-----	-----	-----

TABLE III

SÉQUENCE D'UN MESSAGE ISSUE DES DONNÉES DE MAINTENANCE.

Le processus d'évaluation de l'entropie utilise les spécifications suivantes (voir Figure 3) :

- 1000 2-uplets ( $Symb\_U, Etat\_U$ ), pour la distribution uniforme (Figure 5),
- 1000 2-uplets ( $Symb\_N, Etat\_N$ ), pour la distribution normale (Figure 4).

Chaque séquence se termine par un arrêt du processus (symbole DEP en rouge).

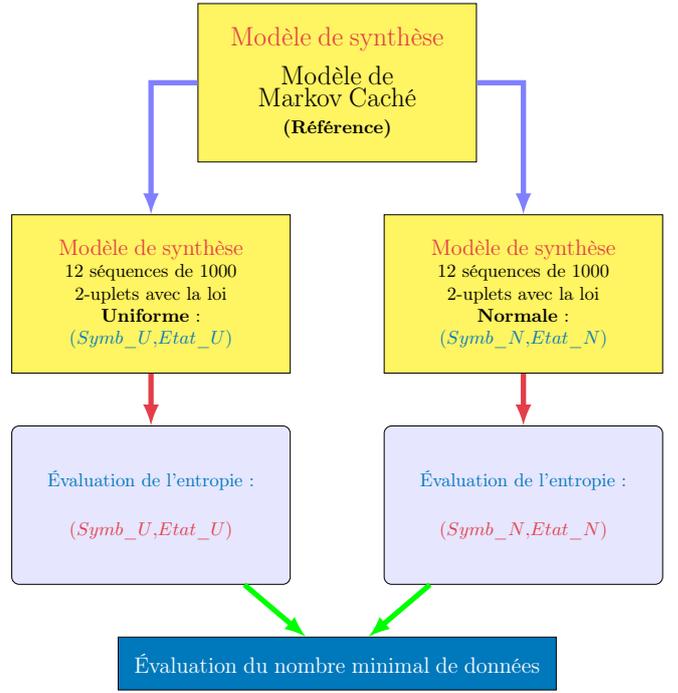


Fig. 3. Évaluation du nombre minimal de données : le modèle de synthèse génère 1000 couples ( $Symboles, Etats$ ) en utilisant les distributions normales ou uniformes ( $(Symb\_U, Etat\_U)$  et  $(Symb\_N, Etat\_N)$ ). Tous les 2-uplets sont répartis sur 12 séquences. Les 2-uplets sont utilisés dans le processus d'évaluation afin de déterminer le nombre minimal de données de modélisation, en utilisant l'entropie de Shannon.

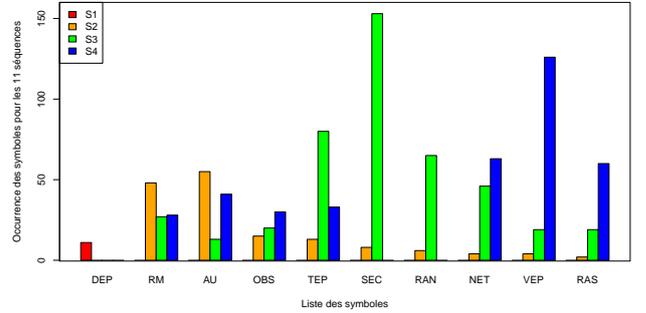


Fig. 4. Distribution normale des symboles produits par le MMC de référence. DEP représente l'état 1 du modèle, c'est un arrêt de production. Le symbole AU (AUtre) est le plus représentatif de l'état 2. (SECurité) est le plus représentatif de l'état 3 et VEP (Visite Entretien Préventif) pour l'état 4.

## B. Application industrielle

Les données récoltées pour cette étude sont issues d'une peseuse volumétrique sur une ligne de production de pain de mie. Tous les processus de l'usine étudiés sont liés entre eux de manière séquentielle. L'arrêt d'un processus engendre donc l'arrêt des éléments en aval. Ce processus de production est maintenu par la mise en place d'une politique de maintenance préventive. Pour ce faire, les agents de maintenance doivent consigner leurs actions ou observations dans une base de données centralisée (voir Tableau I). Un échantillonnage à la journée a été choisi conformément à la politique de maintenance interne. Si aucune donnée n'est saisie, nous insérons dans cette base de données le symbole

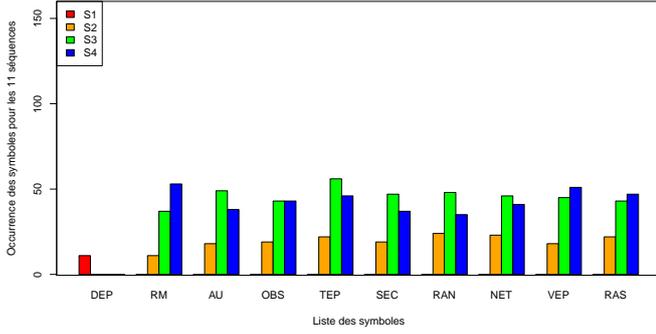


Fig. 5. Distribution uniforme des symboles produits par le MMC référence.

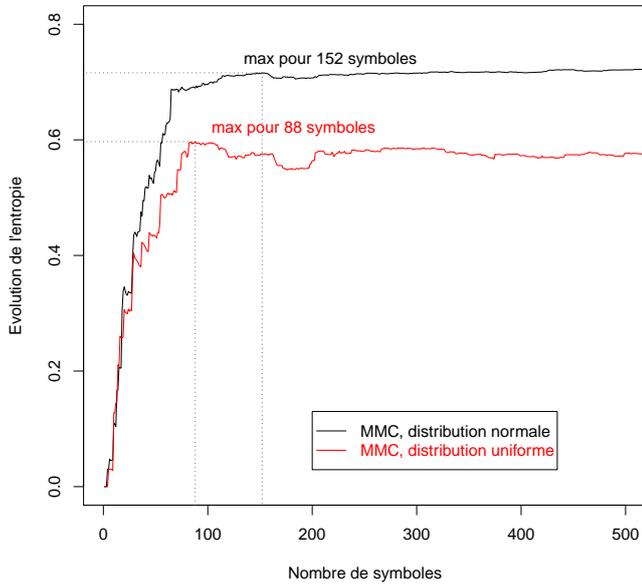


Fig. 6. Évolution de l'entropie de Shannon à partir des données issues du modèle de synthèse.

RAS (= tout va bien). Cet échantillonnage temporel régulier nous permet d'être dans une situation d'application de processus de Markov. Nous utilisons ensuite ces séquences de symboles (voir Tableau III) pour modéliser le niveau de dégradation du processus. Nous modélisons cette « signature » à l'aide de MMC. Les données récoltées sur 2 ans comprennent environ 2000 événements (voir Tableau I).

#### IV. RESULTATS

##### A. Évolution de l'entropie pour le modèle de synthèse

L'évolution de l'entropie de Shannon à partir des données issues du modèle de synthèse nous donne les résultats présentés sur la Figure 6. Nous observons un maximum d'entropie pour 152 symboles avec la distribution normale et 88 symboles avec la distribution uniforme. Nous voyons également que la courbe est asymptotique au-delà de 200 symboles.

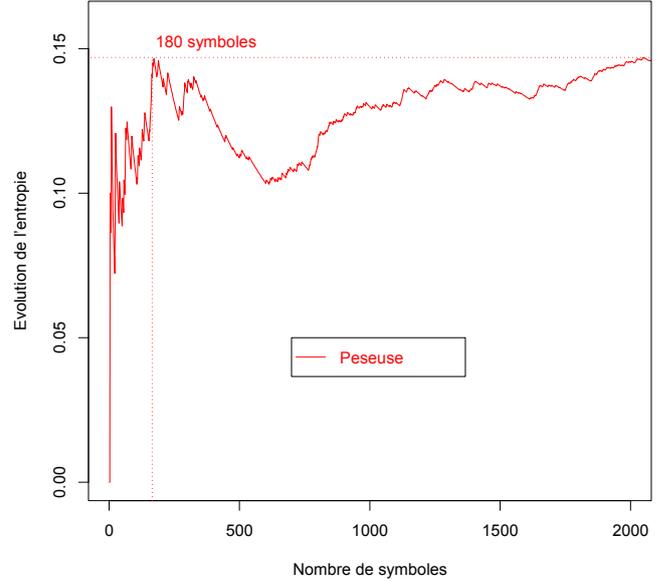


Fig. 7. Évolution de l'entropie pour l'application industrielle.

##### B. Évolution de l'entropie pour l'application industrielle

Les données industrielles, ont été collectées de 2005 à 2007. Les résultats de l'entropie de Shannon sont données Figure 7. La valeur maximale est atteinte pour 180 symboles. Cette valeur maximale de l'entropie correspond à la valeur asymptotique pour un nombre élevé de symboles.

#### V. DISCUSSION

Le but de cette étude était de trouver une quantité minimale de données pour estimer correctement un modèle.

##### A. Entropie maximale

Sans connaissance à priori, nous avons évalué différentes signatures par mesure de l'entropie de Shannon. Cette signature est considérée comme une chaîne de Markov cachée du 1<sup>er</sup> ordre. L'entropie de Shannon nous a permis précédemment [14] d'évaluer la pertinence de modèles. Nous avons trouvé dans cette étude un maximum d'entropie pour des données simulées issues de deux distributions différentes. La première valeur maximale est de 88 symboles pour la distribution uniforme et 152 symboles pour la distribution normale. Les résultats Figure 6 montrent aussi une valeur asymptotique de l'entropie, indiquant qu'il est possible d'utiliser un nombre limité de symboles dans la conceptualisation d'un modèle. Ce modèle ainsi établi avec un nombre fini de symboles, pourra être remis à jour en fonction des nouvelles activités de maintenance. Pour une maintenance prédictive, nous pouvons ainsi établir une « fenêtre glissante » contenant ce nombre de symboles trouvé précédemment.

##### B. Fenêtre glissante

La fenêtre glissante (voir Figure 8) contient le nombre minimal de symboles trouvé au travers des Figures 6 et 7, où l'entropie est maximale. Tous les symboles à l'intérieur de cette fenêtre permettraient de réaliser une estimation

des modèles régulièrement. Ces modèles dynamiques pourraient ainsi fournir une aide à la décision dans l'organisation quotidienne de la maintenance préventive. Cette quantité minimale de données pourra être implémentée dans des algorithmes d'apprentissage. Ceci permettra de réévaluer régulièrement le modèle. En effet, arrivé à l'instant « Maintenant » (Figure 8), nous ne connaissons pas la prochaine observation. Avec cette fenêtre glissante, nous serons en mesure d'estimer le prochain état caché. Le but principal de la fenêtre glissante est de fournir une mise à jour des modèles à l'instant souhaité.

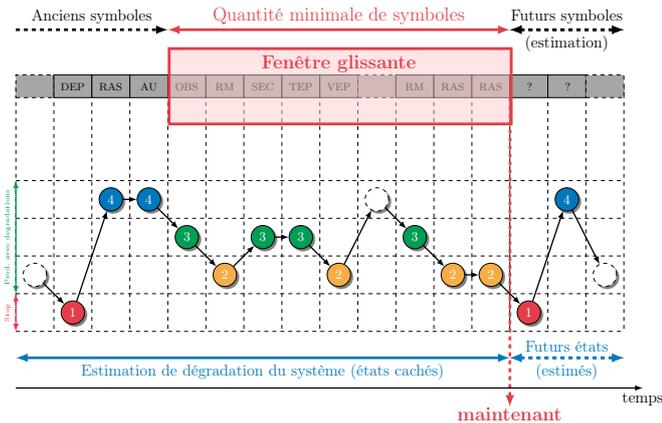


Fig. 8. La fenêtre glissante contient le nombre minimal de symboles qui donne une entropie maximale. Le but principal de la fenêtre glissante est de prévoir une mise à jour des modèles.

### C. Confrontation du modèle de synthèse avec l'application réelle

Nous comparons les données issues du modèle synthétique avec celles provenant d'une GMAO industrielle. La codification symbolique choisie étant identique dans les deux cas (MMC de synthèse et MMC situation réelle), nous avons une cohérence au niveau des observations du processus (voir Tableau II). Les données du modèle de synthèse sont du type  $Symb\_U$  pour la loi uniforme et  $Symb\_N$  pour la loi normale (voir Figure 3).

Dans l'étude du cas réel (voir contexte § III-B), nous avons trouvé une fenêtre de 180 symboles afin d'obtenir un maximum d'entropie. Comme nous pouvons le voir sur les Figures 6 et 7, la mesure de l'entropie (informations issues du contexte réel) suit le même comportement que l'entropie « synthétique ». Par ailleurs, toutes les valeurs du maximum d'entropie trouvées précédemment correspondent aux valeurs asymptotiques des évolutions des entropies respectives. Ainsi, ces valeurs d'entropie maximales nous donnent une idée sur le nombre minimal d'observations pour estimer le mieux possible les modèles. Ce nombre minimal de symboles nous donne des éléments pour choisir la taille minimale de la fenêtre glissante permettant d'optimiser la conceptualisation des modèles de maintenance préventive.

### D. Conclusion

Ainsi, l'étude réalisée montre que l'on peut trouver une quantité minimale de données pour estimer correctement un modèle. Il faudra valider ce nombre minimal d'observations avec de nouveaux tests sur le modèle de synthèse.

Ainsi, dans une future phase d'exploitation, on pourra donc réévaluer le modèle régulièrement. Dans ces conditions, l'expert pourra bénéficier d'informations lui permettant d'effectuer une programmation dynamique des interventions de maintenance puisque les actions programmées par la méthode vont modifier le comportement du système.

### RÉFÉRENCES

- [1] P. Vignat, M. Avila, F. Duculty, and F. Kratz, "Use of HMM for evaluation of maintenance activities," *IJAIS, International Journal of Adaptive and Innovative Systems*, Vol. 1, Nos. 3/4, pp. 216–232, 2010.
- [2] B. Roblès, M. Avila, F. Duculty, P. Vignat, and F. Kratz, "Statistical evaluation of Hidden Markov Models topologies, based on industrial synthetic model," in *14th IFAC Symposium on Information Control Problems in Manufacturing Information Control Problems in Manufacturing*, A. D. I. F. F. G. Boran-giu, Theodor ; Dolgui, Ed., vol. 14, no. 1, Bucharest, Roumanie, May 2012, pp. 1029–1034.
- [3] M. A. Shahid, "Cross entropy minimization for efficient estimation of sram failure rate." in *DATE*, W. Rosenstiel and L. Thiele, Eds. IEEE, 2012, pp. 230–235.
- [4] Z. Botev, D. Kroese, and T. Taimre, "Generalized cross-entropy methods with applications to rare-event simulation and optimization," *Simulation*, vol. 83, no. 11, pp. 785–806, Nov. 2007.
- [5] R. Y. Rubinstein, "Optimization of computer simulation models with rare events," *European Journal of Operations Research*, vol. 99, pp. 89–112, 1997.
- [6] S. Aupetit, N. Monmarché, and M. Slimane, "Hidden Markov models training using population based metaheuristics," in *Advances in Metaheuristics for Hard Optimization*, ser. Natural Computing Series, P. Siarry and Z. Michalewicz, Eds. Springer, 2007, pp. 415–438.
- [7] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceeding of the IEEE, 77(2) SIAM interdisciplinary journal*, pp. 257–286, 1989.
- [8] M. Fox, M. Ghallab, G. Infantes, and D. Long, "Robot introspection through learned hidden Markov models," *Artif. Intell.*, vol. 170, no. 2, pp. 59–113, 2006.
- [9] D. Hocker, L. Xiaohu, and S. S. Iyengar, "Shannon entropy based time-dependent deterministic sampling for efficient on-the-fly quantum dynamics and electronic structure," *J. Chem. Theory Comput.*, pp. 256–268, 2011.
- [10] A. Agouzal and T. Lafouge, "On the relation between the maximum entropy principle and the principle of least effort : The continuous case," *J. Informetrics*, vol. 2, no. 1, pp. 75–88, 2008.
- [11] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, 1948.
- [12] V. Chandrasekaran, J. K. Johnson, and A. S. Willsky, "Maximum entropy relaxation for graphical model selection given inconsistent statistics," *Laboratory for Information and Decision Systems, Massachusetts Institute of Technology Cambridge, MA 02139*, 2007.
- [13] M. Arminjon and D. Imbault, "Maximum entropy principle and texture formation," *Zeitschrift für angewandte Mathematik und Mechanik*, 80, Suppl. N°1, pp. 13–16, 2000.
- [14] B. Roblès, M. Avila, F. Duculty, P. Vignat, and F. Kratz, "Evaluation of relevance of stochastic parameters on Hidden Markov Models," in *Advances in Safety, Reliability and Risk Management, European Safety and Reliability Conference (ESREL)*, ISBN : 978-0-415-68379-1. Troyes, France : Taylor & Francis Group, Sep. 2011, p. 71.