

# Vers une analyse de la dynamique des paramètres physiologiques en Unité de Soins Intensifs

S. Sharshar\*, M-C. Chambrin\*

\* EA 2689, INSERM IFR 114 ITM, Bât. Vancostenobel, CH&U de Lille, Place de Verdun  
F-59037 Lille Cedex  
samir.sharshar@mac.com  
chambrin@lille.inserm.fr

**Résumé.** Le monitoring des paramètres physiologiques en milieu de Réanimation génère un flux abondant et continu de données. L'extraction et la synthèse des informations sont une étape obligatoire pour tout système d'aide à la décision. L'Analyse de la Tendance Linéaire par Morceaux (A.T.L.M.) est une méthode originale d'analyse de la dynamique qui utilise deux niveaux d'interprétation. Le premier niveau, monoparamétrique, exprime la variation de chaque paramètre en quatre classes (constant, diminue, augmente, transitoire) selon la méthode décrite dans (Calvelo et al. 2001). Le second niveau, multiparamétrique, définit le comportement du système par une variable continue ; l'introduction de la connaissance du domaine est alors nécessaire pour différencier des états. La combinaison des deux niveaux d'informations permet l'élaboration de scénarios. Le propos de ce document est d'exposer la méthode sur des signaux simulés présentant des modifications pouvant s'observer en clinique et d'introduire les perspectives offertes par son implémentation en ligne.

## 1 Introduction

De façon générale, l'approche de l'aide à la décision peut être exprimée en termes de transformation de paramètres en données puis en information, enfin en connaissance (Coiera 1994, Mora et al. 1993). Le passage d'un niveau d'abstraction au suivant implique de mettre en œuvre des moyens de transcodage pour effectuer le passage d'un système de représentation à un autre. Cette chaîne d'abstraction (Calvelo 1999) peut être schématisée en un processus qui comprend :

1. L'acquisition des paramètres,
2. Le traitement univarié des paramètres,
3. Le traitement multivarié des paramètres,
4. L'intégration des connaissances du domaine,
5. La présentation des différentes informations au personnel soignant.

La connaissance du domaine peut intervenir à chacune des étapes 1 à 3 mais sans pour autant être nécessaire. L'étape 1 fait appel aux procédés classiques du traitement de signal. Elle est, pour les paramètres étudiés, déjà intégrée aux moniteurs. L'A.T.L.M. ou Analyse de la Tendance Linéaire par Morceaux s'applique aux étapes 2 et 3 de ce processus d'abstraction en n'introduisant l'intervention de l'expertise qu'en bout de chaîne.

L'A.T.L.M. procède en différentes étapes détaillées dans la section 2. Pour illustrer les résultats de la méthode, nous avons choisi 6 paramètres simulés  $a$  à  $f$  reportés sur la figure 1.

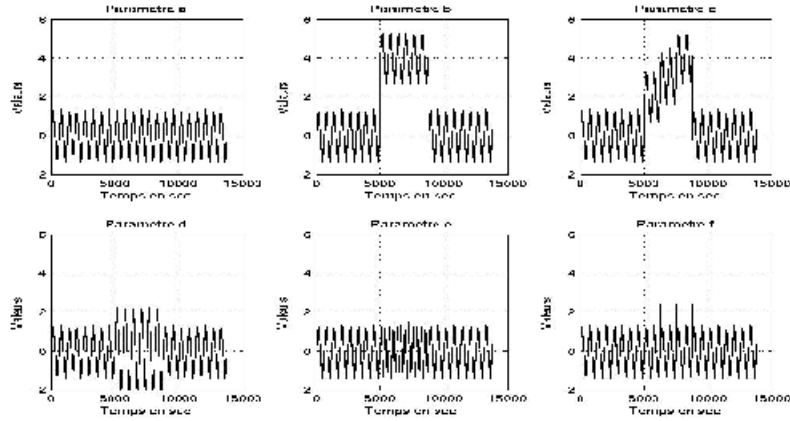


FIG. 1 – Paramètres tests

Le paramètre *a* est construit à partir d'une sinusoïde bruitée et quantifiée. Les paramètres *b* à *f* sont construits à partir du paramètre *a* en introduisant des modifications entre 5028 et 8799 sec.

## 2 Méthodologie

### 2.1 Échelles caractéristiques et transformation numérique → symbolique

L'A.T.L.M. est basée sur l'extraction d'une tendance locale. Cette tendance est calculée selon la méthode décrite dans (Calvelo et al. 2000). Brièvement, le traitement consiste à calculer la tendance par régression linéaire sur une fenêtre dont la taille est calculée en fonction de la dynamique propre de chaque paramètre et appelée *échelle* ou *temps caractéristique*  $\tau$ . Nous avons retenu deux critères pour la détermination de l'échelle caractéristique : le premier correspond au principe de la recherche de la meilleure approximation de la dérivée ( $\tau_f$ ), le second à l'échelle qui fournit un modèle global linéaire par morceau ( $\tau_z$ ).

À l'échelle caractéristique, à chaque instant on dispose de la valeur du paramètre, de sa *tendance* caractérisée par la pente de la droite de régression ( $r$ ). Le calcul de l'écart-type fournit un indicateur de la *stabilité* locale du paramètre. C'est à partir du partitionnement du plan formé par *tendance* et *stabilité* que les notions de tendance qualitative (augmentation, diminution) et de stabilité (stable, instable) sont obtenues. Ce traitement réalise une transformation numérique → symbolique des données. Différentes modalités de partitionnement ont été testées sur des données issues de protocoles cliniques (Chambrin et al. 2003) afin de déterminer les critères les mieux adaptés aux données traitées (valeurs fixes égales à -0.7 et +0.7 pour la tendance et un seuil fixé au 95° percentile de la distribution des écarts-type pour la stabilité). Les symboles issus de ces deux traitements ont été combinés de façon à caractériser le signal selon 4 classes, les *AgClasses* : constant (C), transitoire (T), augmentation (A) et diminution (D). L'application de la méthode sur les données simulées est illustrée par la figure 2.

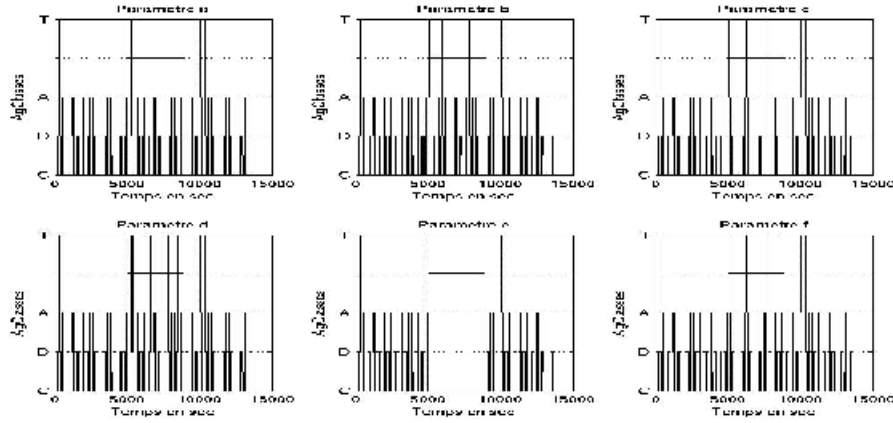


FIG. 2 - Résultats de la symbolisation en AgClasses. C=constant, D=diminution, A=augmentation, T=transitoire. La période correspondant aux modifications est repérée par une ligne horizontale. La composante sinusoïdale du paramètre est rendue par la succession des D et A ; sur le paramètre a subsistent quelques T liés au bruit ; les périodes de modifications des paramètres b et c sont rendues par des T et D correspondant à la brutale décroissance observée à 8799 sec ; le changement d'amplitude du paramètre d est repéré par des T ; le changement de fréquence est repéré par la disparition de la composante D et A au profit de C les valeurs aberrantes de f sont repérées par des T.

## 2.2 Calcul des index de gaussianité Ig

L'Ig se propose d'apprécier le caractère infra ou supra gaussien de la distribution des coefficients de régression  $r$  établis en 2.1. Nous cherchons par l'établissement de cet index à repérer les *changements de répartition* des tendances locales. Nous avons retenu le coefficient d'aplatissement ou kurtosis comme représentatif de la forme de la distribution des  $r$ . Son calcul est réalisé sur une fenêtre glissante de longueur fixe par pas de 1 et, sur toute la longueur de l'observation pour chaque paramètre et à chaque échelle caractéristique  $\tau_f$  et  $\tau_z$ . La figure 3 illustre le résultat obtenu sur les paramètres  $a$  à  $f$  à  $\tau_f$ .

## 2.3 Indicateur $\theta$ de comportement global du système

Cet indicateur est construit à partir de l'analyse en composantes principales des valeurs centrées réduites des Ig définis en 2.2. Cette analyse est réalisée sur une fenêtre glissante de longueur fixe par pas de 1. La valeur moyenne ( $\Delta$ ) des valeurs projetées sur la première composante principale est calculée et positionnée par rapport à l'intervalle  $[-1 ; +1]$ . On définit ensuite une variable  $\delta$  qui prend la valeur 0 lorsque  $\Delta$  est dans l'intervalle  $[-1 ; +1]$  et la valeur 1 lorsque  $\Delta$  est en dehors de cet intervalle. L'indicateur  $\theta$  est défini par :

$$(1) \Theta = \frac{1}{t_2 - t_1} \sum_{t=t_1}^{t_2} \delta(t) \text{ avec } t_2 - t_1 \text{ correspondant à la taille de la fenêtre}$$

La figure 4 illustre le résultat du calcul de  $\theta$  en prenant en compte l'ensemble des paramètres  $a$  à  $f$ . L'introduction de la connaissance permet à ce stade de déterminer une ou plusieurs valeurs seuils sur  $\theta$  afin de différencier des états. Sur les paramètres tests, nous

avons défini 2 états : état de base et état modifié sur la période entre 5028 et 8799 sec, de la même façon que l'expertise clinique permettrait de déterminer des périodes d'états stables ou non. Dans notre exemple, l'établissement d'une courbe ROC (*Receiver Operating Curve*) fixe le seuil à une valeur de 0.34. (Aire sous la courbe = 0.99, Sp = 98%, Se = 99%).

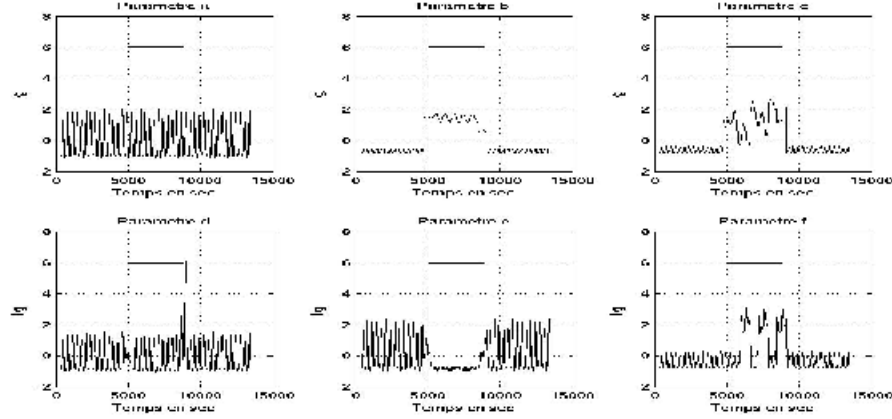


FIG. 3 - Valeurs centrées réduites de l'index  $I_g$  à  $\tau_j$  des paramètres a à f. La période correspondant aux modifications du paramètre a de base est repérée par une ligne horizontale. On note une excellente différenciation des états pour les paramètres b, c, e, f. Le changement d'amplitude (paramètre d) est simplement repéré au début et à la fin.

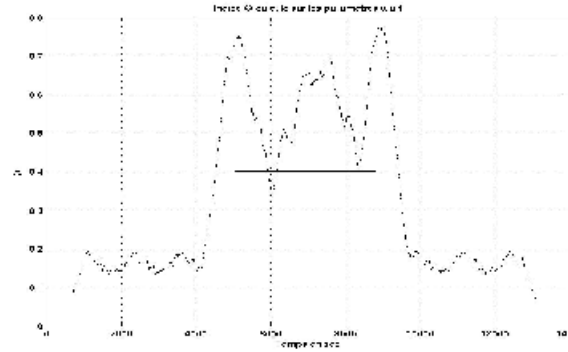


FIG. 4 – Indicateur  $\theta$  calculé pour l'ensemble des paramètres a à f. La période de modification du paramètre a de base est repérée par une ligne horizontale.

## 2.4 Construction de mots et de scénarios

$\theta$  permet de définir différents états pour chacun desquels une *interprétation* des résultats de la symbolisation par les *AgClasses* est envisageable. Cette interprétation passe par l'élaboration de *mots* qui sont à la base des *scénarios*.

Qu'est ce qu'un mot ? Nous définissons le mot comme l'assemblage des valeurs des *AgClasses* de  $i$  paramètres à l'instant  $t$ . Chaque mot est donc constitué d'un « alphabet » de 4

lettres : C, D, A, T. Un mot défini sur  $i$  paramètres prendra à l'instant  $t$  un « sens » parmi 4 puissance  $i$  combinaisons possibles. Nous cherchons ensuite à définir les mots *pertinents* d'un état du système (tel que défini par  $\theta$ ). Un simple histogramme de fréquence des mots dans un état est insuffisant. De la même façon que dans un texte un mot très fréquent peut être considéré comme significatif, il ne peut être qu'un simple mot de liaison. À l'inverse un mot peu présent peut contenir tout le sens du texte. Nous proposons une transformation de la fréquence  $w$  du mot dans l'état par la *word frequency* :

$$(2) wf = 1 + \log(w)$$

Nous cherchons ensuite une *caractérisation* de l'état par les mots. La succession d'états définis par  $\theta$  peut, pour poursuivre l'analogie avec l'analyse textuelle, être considérée comme une succession de *parties* d'un document. Par un raisonnement analogue à celui employé pour définir la pertinence par  $wf$ , on calcule pour chacun des états et pour chaque mot la *word relative frequency*

$$(3) wrf = wf \cdot \log\left(\frac{n}{df}\right)$$

où  $n$  = nombre total d'états définis par  $\theta$ ,  $df$  le nombre d'états dans lequel le mot est présent.

Qu'est ce qu'un scénario ? Un scénario est un plan qui porte en abscisse le temps et en ordonnée les mots de l'état par ordre décroissant de  $wrf$ . Le scénario relatif à un état est construit à partir de ses mots caractéristiques et de leur instant d'occurrence. Il sera utilisé comme base d'apprentissage d'évènements rencontrés en clinique. La figure 5 illustre les résultats du calcul de  $wrf$  sur les paramètres  $b$ ,  $c$  et  $d$  et donne un exemple de scénario.

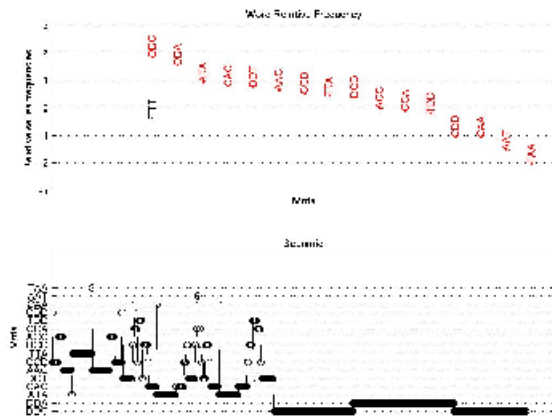


FIG. 5 - Illustration de la construction d'un scénario. En haut résultats du calcul de  $wrf$  : en noir, état « stable » entourant la période de modification ; en rouge : état « instable » correspondant à la période de modification. En bas, scénario construit sur l'état instable.

### 3 Conclusion et perspectives

## A.T.L.M. vers une analyse de la dynamique des paramètres physiologiques

L'A.T.L.M. est un outil cohérent d'étude de la dynamique. Il permet par une symbolisation de la variation des données et la définition d'états d'accéder à un niveau d'abstraction pouvant être interfacé avec d'autres systèmes de représentation de la connaissance.

Les paramètres tests choisis, même s'ils sont caricaturaux, présentent des modifications proches de celles pouvant être rencontrées en pratique. Les résultats obtenus par l'utilisation de l'A.T.L.M. sur des paramètres «réels» restent à valider par une implémentation «en ligne» en milieu de Réanimation. Dans ce contexte, l'expertise clinique devra être utilisée pour déterminer sur un jeu de données deux états correspondant à la notion d'état clinique stable ou non afin d'évaluer la valeur seuil de  $\theta$ . La construction du scénario proposée en 2.4 sera utilisée comme base pour l'apprentissage d'événements rencontrés en clinique.

## Références

- Calvelo D. (1999), Apprentissage de modèles de la dynamique pour l'aide à la décision en monitoring clinique. Thèse d'université. De Lille1. Décembre 1999.
- Calvelo D., Chambrin M.C., Pomorski D. et Ravaux P. (2000), Towards symbolization using data-driven extraction of local trends for ICU monitoring, *Artif Intell Med.*, 19(3), pp 203-223, 2000.
- Calvelo D., Chambrin M.C. et Pomorski D. (2001), From local trend extraction to symbolization of time-series, *Intelligent Data Analysis*, 5(1), pp 41-57, 2001.
- Chambrin M.C., Bijaoui E., Carry P.Y., Charbonnier S., Dojat M., Garbay C., Mangalaboyi J., Perdrix J.P., Ravaux P., Sharshar S., Silvent A.S. et Vilhelm C. (2003), OSSCAR : A collaborative project for intelligent monitoring in intensive care and anesthesia, Workshop on "Computers in Anaesthesia and Intensive Care: Knowledge-Based Information Management" . Nine European Conference on Artificial Intelligence in Medicine (AIME), Protaras, Chypre 18-22 October 2003.
- Coeira E. (1994), Monitoring in Anesthesia and Intensive Care, chapter Automated signal interpretation, pp 32-42, W.B. Saunders Co Ltd, 1994.
- Mora F.A., Passariello G., Carrault G., Le Pichon J.P.(1993), Intelligent patient monitoring and management systems : a review. *IEEE EMB Magazine*, 12(4), pp 23-33, 1993.

## Summary

In Intensive Care Units, clinical information systems generate a large amount of data often observed with high sampling frequencies. To help the physician in his medical decision process, decision support system should provide the proper extraction and interpretation of the information contained in such massive data sets. Windowed Linear Trend Analysis is an original method for analysing dynamics of physiological parameters. Two levels of interpretation are provided: the first at monoparametric level into 4 classes (constant, increasing, decreasing, transient) using the method described in (Calvelo et al., 2001) and the second at the multiparametric level into two or more states depending on domain knowledge. The combination of these 2 levels allows the elaboration of scenarios. This paper describes the whole methodology applied

A.T.L.M. vers une analyse de la dynamique des paramètres physiologiques

on a set of simulated parameters, and introduces some perspectives about its on line implementation.