

DETECTION DE DEFAUTS ET SEGMENTATION  
DE SIGNAUX DE CONTROLE NON DESTRUCTIF

C. VILLEMUR\* - F. CASTANIE\* - D. GARREAU\*\*

\*ENSEEIHNT/GAPSE 2, Rue Camichel 31071 TOULOUSE CEDEX

\*\*EDF/Direction Etudes et Recherches 6, Quai Watier 78400 CHATOU

RESUME

Le présent article traite de l'application de méthodes paramétriques de segmentation à des signaux de contrôle non-destructif (C.N.D.).

Cette étape se situe en amont d'une chaîne de traitement intégralement paramétrique dont l'aboutissement consiste en une classification automatique des signaux. Nous avons comparé les performances de deux algorithmes — franchissements d'un seuil par l'erreur de prédiction linéaire (E.P.L.) issue d'un modèle autorégressif (AR); mesure d'un rapport de vraisemblance généralisée — en termes de fausses alarmes, de non-détections, de retard à la détection, de complexité opératoire sur une base de données incluant les divers types de défauts à détecter dans les signaux C.N.D. Ces deux algorithmes, notamment le premier, ont déjà été utilisés avec succès pour le traitement de signaux dans des domaines aussi divers que les télémesures spatiales ou l'électromyographie.

SUMMARY

This paper deals with an application of segmentation parametric methods on non-destructive evaluation signals.

This stage takes place in the beginning of an all-parametric processing procedure the final aim of which consists in an automatic classification of signals. We have made a comparison of two algorithms — level crossings of the linear prediction error (LPE) obtained through an autoregressive (AR) modelling; measure of a generalized likelihood ratio (GLR) — in terms of false alarms, non detections, detection delay, computational cost on a data base containing the different kinds of defects to be detected. These two algorithms, the first one essentially, have shown to be very successful in the processing of signals in domains such as spacial telemetry or electromyography.

I - INTRODUCTION

L'implantation de méthodes paramétriques pour le traitement des signaux courants de Foucault\* (C.D.F.) utilisés pour le contrôle des tubes de générateur de vapeur (G.V.) de centrales nucléaires a été entreprise dans l'optique de venir en aide aux experts humains seuls chargés pour l'instant du dépouillement et de l'analyse des données.

Le diagnostic, dont le caractère subjectif dû au facteur humain est très fortement marqué, est en outre altéré par plusieurs phénomènes venant perturber les mesures. L'utilisation de l'informatique est d'autre part rendue obligatoire de par la masse considérable de données à traiter et à archiver lors de chaque inspection. En implantant les méthodes de modélisation paramétrique du signal, on réalise non seulement une compression importante des données — c'est l'aspect purement modélisation — mais en outre on peut les utiliser à des fins de segmentation du signal. D'autres méthodes de détection non-paramétriques auraient pu être testées mais elles ne permettaient pas de passer à l'étape suivante: la classification automatique [9].

II - CONTROLE PAR COURANTS DE FOUCAULT

Le matériau à contrôler — les tubes de G.V. — 'baigne' dans un champ magnétique créé par une sonde montée en mode différentiel alimentée par un courant alternatif. L'apparition d'une hétérogénéité dimensionnelle ou structurelle perturbe le trajet des courants induits — les C.D.F. — dans le matériau conducteur. Le but du contrôle consiste donc à

rechercher dans la réponse recueillie — la variation d'impédance du bobinage de la sonde (deux composantes X et Y en quadrature: on doit donc traiter un signal complexe) — la présence des dites hétérogénéités qui temporellement prennent la forme de transitoires très courts pour la quasi-totalité des défauts rencontrés et de courbes de Lissajou dans le plan de phase. Ces dernières constituent le principal outil d'investigation des experts C.D.F. Le problème principal consiste à isoler un défaut du signal enregistré sachant que ce dernier peut être 'noyé' par diverses perturbations que nous allons détailler.

II-1 - Les problèmes d'interprétation

Ces éléments perturbateurs ont plusieurs origines; nous citerons les suivants:

— le bruit de fond dû au laminage des tubes, davantage présent sur la voie X. Ce bruit qui présente un caractère 'coloré' a déjà été traité par des méthodes telles que le corrélo-filtre.

— les plaques-supports, les fixations aux extrémités du tube, les différences de structure magnétique, les entrées-sorties du tube.

- les chocs éventuels lors du montage.
- les variations de vitesse de la sonde.

II-2 - Objectifs et avantages d'un traitement paramétrique

L'interprétation visuelle des courbes de Lissajou varie d'un groupe d'experts à un autre, des techniques reposant sur le traitement paramétrique du signal et l'informatique peuvent remédier à ce problème. Elles sont en effet reproductibles, implantables sur microprocesseurs et offrent donc la possibilité d'inspections en temps réel sur site, d'autre part elles peuvent délivrer des informations non accessibles à l'opérateur humain ou ayant échappées à son investigation.

\* Ce travail a été accompli dans le cadre d'une convention INPT/EDF et fait partie de la thèse de doctorat de C. VILLEMUR, membre du GRECO/SARTA.



### III - DESCRIPTION DES ALGORITHMES DE SEGMENTATION

Le signal observé  $x(n)$  s'écrit soit sous la forme:  $x(n)=b(n)$  en l'absence de tout signal perturbateur — défauts, signaux provenant des structures métalliques de support, etc... — soit sous la forme:  $x(n)=s(n)+b(n)$  quand apparaît l'objet à détecter ( $b(n)$  représente le bruit supposé additif décrit au paragraphe II.2).

Nous nous proposons de détecter en ligne les changements de caractéristiques spectrales du signal à l'aide d'un modèle AR : l'apparition d'une non-stationnarité  $s(n)$  se caractérise en effet par une modification brusque des paramètres du modèle.

Les algorithmes de segmentation doivent satisfaire aux conditions suivantes :

— faible taux de fausses alarmes (F.A.) et de non-détections (N.D.).

— faible retard à la détection.

— détection 'symétrique': résultats similaires si l'on traite la non-stationnarité dans le sens 'avant' et dans le sens 'rétrograde'

Deux algorithmes, que nous désignerons par  $n^1$  et  $n^2$  par la suite, ont été mis en oeuvre:

— étude des franchissements d'un seuil par l'erreur de prédiction linéaire (EPL) issue d'un modèle AR ( voir remarque ci-dessous); analyse d'un mot de dépassement [1] [2] [3]

— algorithme de APPEL-BRANDT: calcul d'un rapport de vraisemblance généralisée [4] [6] [8].

Ces deux algorithmes opèrent un filtrage séquentiel du signal et utilisent une fenêtre dans le premier cas, trois dans le second cas.

#### III-1 - Franchissements d'un seuil par l'E.P.L.

Si nous considérons le modèle AR d'ordre  $p$  d'un signal échantillonné complexe  $x(n)$ :

$$x(n) = -\sum_{k=1}^p a_k x(n-k) + b(n) \quad (3.1)$$

l'erreur de prédiction linéaire est obtenue en filtrant le signal  $x(n)$  à l'aide du filtre blanchissant  $A(z)$ . Les propriétés de franchissement d'un seuil par l'EPL — méthode déjà abondamment décrite dans la littérature [1] [2] [3] — permettent de conclure à l'apparition d'une non-stationnarité quand l'EPL séjourne au-dessus d'un seuil  $d$  pendant  $L$  échantillons.

Le seuil  $d$  est défini comme suit à partir de la probabilité de fausse-alarme  $\alpha$ :

$$(P(|e_n| > d = k \cdot \sigma_e))^{L-1} \ll \alpha \quad (3.2)$$

$\hat{\sigma}_e^2$ : puissance de l'erreur en sortie de l'algorithme de modélisation AR.

Si la détection se révèle positive, la puissance de l'erreur  $\hat{\sigma}_e^2$  et les paramètres du modèle sont réinitialisés. La nouvelle fenêtre d'estimation débutera quelques échantillons après la non-stationnarité détectée.

On définit le mot de dépassement  $D_L(n)$  [1] [2] de la façon suivante : soit une variable binaire  $I(k)$

$$\begin{cases} I(k)=1 & \text{si } e_k > d \\ I(k)=0 & \text{sinon} \end{cases}$$

$$D_L(n) = \{I(n), I(n+1), \dots, I(n+L-1)\} \quad (3.3)$$

Il existe certainement un mot de dépassement pour chaque type de non-stationnarité rencontrée. En pratique, sur les signaux C.D.F. à notre disposition, nous nous limiterons, et ce pour des raisons de robustesse des algorithmes, à détecter quelques dépassements consécutifs ( $L=4$ ). Cependant nous avons pu mettre en évidence que la détection d'une non-stationnarité particulière pouvait se faire à l'aide d'une signature spécifique [10]. On pourrait donc opérer ainsi une pré-classification.

Nous pouvons énoncer une stratégie de détection très simple: 'Si un dépassement du seuil dure au

moins  $L$  échantillons consécutifs, il y a présence d'une non-stationnarité. Il ne faut pas tenir compte des dépassements du seuil dans les autres cas.

#### Remarque

L'algorithme AR utilisé est de type covariance, formulé en complexe, et a été développé par L. MARPLE [7].

#### III-2 Mesure d'un rapport de vraisemblance

Cet algorithme [6] [8] utilise un rapport de vraisemblance généralisée. A l'instant  $n$ , on opère une modélisation sur une fenêtre à mémoire croissante [FC] de longueur  $(1, n_0)$  et sur une fenêtre glissante [FG] de longueur fixe  $(n_0, n)$  que nous noterons LW. Nous estimons en outre un modèle AR sur l'ensemble des observations  $(1, n)$ . Nous avons donc à notre disposition trois énergies résiduelles:  $e_{rFC}$ ,  $e_{rFG}$  et  $e_{rLW}$  en posant  $[T] = [FC+FG]$ .

Les signaux étant supposés gaussiens, on calcule une mesure  $d(n)$  qui dans ce cas est assimilable à un test de vraisemblance généralisée:

$$d(n) = 1/n \cdot \text{Log}(e_{rLW}/n) - 1/n_0 \cdot \text{Log}(e_{rFC}/n_0) - 1/LW \cdot \text{Log}(e_{rFG}/LW) \quad (3.4)$$

Cette distance est distribuée comme une loi du  $\chi^2$  à  $p$  degrés de liberté,  $p$  désignant toujours l'ordre du modèle AR. Nous pouvons adapter ensuite une procédure de détection analogue à celle de l'EPL: compter  $L$  dépassements consécutifs par la distance décrite ci-dessus, si tel est le cas on signale la présence d'une non-stationnarité et on réinitialise l'algorithme. Ce critère est bien sûr totalement empirique car les propriétés théoriques de franchissement du seuil ne sont pas connues ( ce qui expliquera les cas de simulations où le positionnement du seuil se fait d'une manière totalement erratique). L'influence de l'ordre de modélisation  $p$  et de la longueur de fenêtre LW est bien sûr primordial. Un ordre faible ( $\omega 4,6$ ) est en général préférable à un ordre élevé, d'autre part une fenêtre trop courte autorise une trop grande sensibilité aux bruits de fond.

Cet algorithme qui a fait l'objet d'applications sur des signaux E.E.G. [8] est moins souple de mise en oeuvre que le précédent ( intervention de trois fenêtres), il est aussi plus coûteux en temps de calcul ( voir paragraphe IV).

#### Remarque

L'algorithme de modélisation est identique à celui cité en référence au paragraphe III.1.

### IV - SIMULATIONS

Sur une quinzaine de signaux représentant les classes de défauts les plus couramment rencontrées — ce sont des défauts, dits 'étalons', usinés sur des tubes neufs et répertoriés sur une base de données [11] —, nous évaluons les performances de chaque algorithme.

Les figures IV.1 à IV.4 illustrent un cas où la non-stationnarité est apparente sur les deux voies en même temps — avec toutefois un bruit plus important sur Y que sur X — et où les 2 algorithmes, testés dans des conditions identiques ( ordre  $p=4$ ), fournissent une détection (  $L=4$ ) quasiment pour le même échantillon temporel (  $r=225$ ). En ce qui concerne l'algorithme  $n^2$ , la fenêtre glissante LW a une longueur de 30 points ( les algorithmes ne sont pas réinitialisés après la détection).

Par contre sur les figures IV.5 à IV.8, si l'algorithme  $n^1$  ( les paramètres de réglage  $p, L, LW$  ayant gardé les mêmes valeurs que précédemment) permet de conclure à une détection sur un signal où le transitoire révélateur d'un défaut n'est visible que sur la voie Y, l'algorithme  $n^2$  s'avère inefficace: le niveau élevé du seuil interdit toute détection. Ceci est fréquemment le cas pour cet algorithme où le positionnement du seuil découle de critères très empiriques d'où le comportement parfois erratique du test ( \*voir tableau ci-dessous)

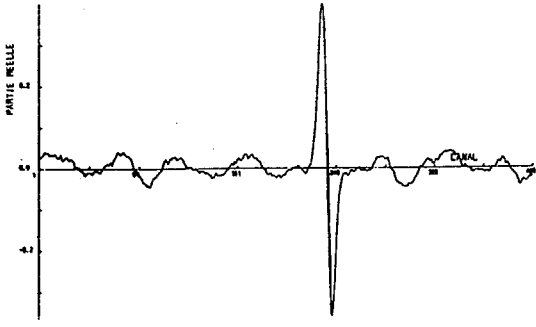


fig IV.1

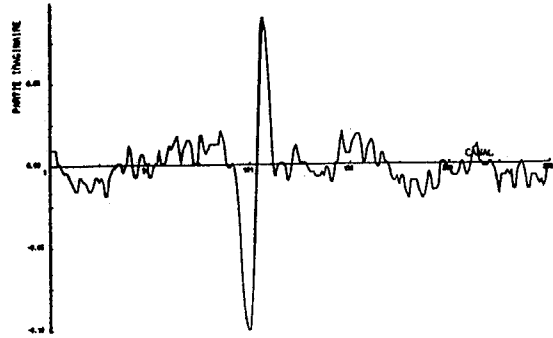


fig IV.6

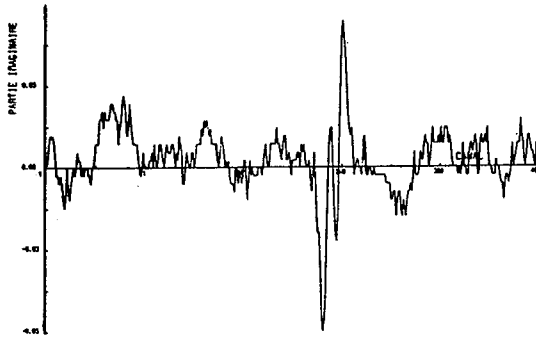


fig IV.2

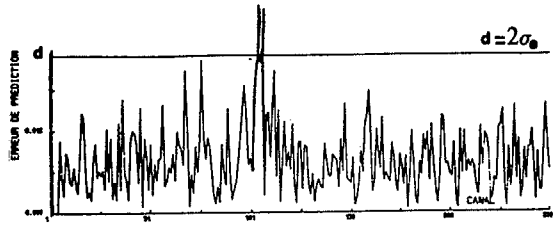


fig IV.7

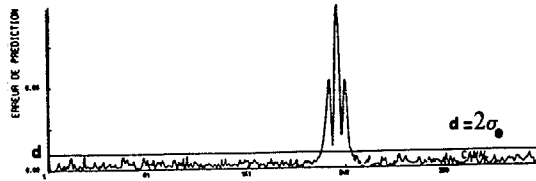


fig IV.3

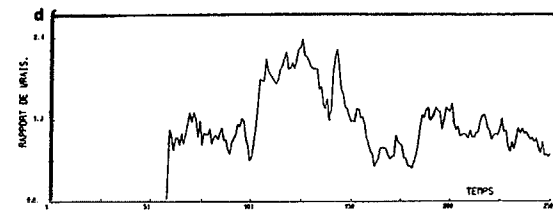


fig IV.8

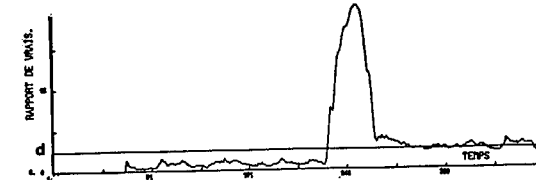


fig IV.4

**légendes**

- IV.1: partie réelle sur 400 points.
- IV.2: partie imaginaire id
- IV.3: algorithme n°1: EPL.
- IV.4: algorithme n°2: APPEL-BRANDT.

**légendes**

- IV.5: partie réelle sur 250 points.
- IV.6: partie imaginaire id
- IV.7: algorithme n°1: EPL.
- IV.8: algorithme n°2: APPEL-BRANDT.

**Remarque**

Dans le cas n°1, on pourrait facilement diminuer la variance de l'erreur en augmentant l'ordre de modélisation ( $p=14$ ). Pour un tel ordre, les temps de calcul, dans le cas n°2, deviennent prohibitifs.

Si nous dressons un tableau comparatif des fausses alarmes (FA) et des non-détectées (ND) comptabilisées sur l'ensemble des simulations pour chacun des algorithmes:

	EPL	APPEL
FA	0	3
ND	2	7*

p=4

	EPL	APPEL
FA	0	--- **
ND	0	--- **

p=14

\*\* : problèmes numériques.

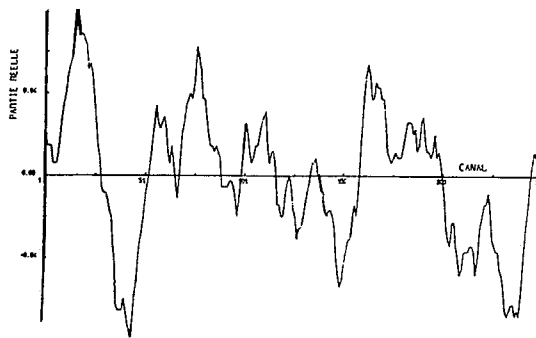


fig IV.5

En ce qui concerne les retards à la détection, si l'algorithme n°1 réagit quasi instantanément, il n'en est pas de même pour le n°2 qui présente parfois une inertie proportionnelle à la longueur de fenêtre glissante LW. L'algorithme n°1 se révèle plus robuste en présence d'un bruit de fond élevé.

D'autre part, d'un strict point de vue calculatoire, l'avantage de l'algorithme n°1 — modélisation en  $o(p^2)$  plus filtrage à  $p$  coefficients — est évident par rapport à l'algorithme n°2 — trois modélisations en  $o(p^2)$  plus mesure d'une distance —



## V - CONCLUSION

L'algorithme de APPEL-BRANDT dont la formulation théorique est très 'séduisante' ne semble pas approprié pour le traitement des signaux CND si l'on en juge par ses performances en FA/ND et retard. Par contre, l'analyse des franchissements d'un seuil par l'EPL fournit un algorithme particulièrement efficace et facilement implantable sur microprocesseur. Sa robustesse, alliée à sa facilité de mise en oeuvre et à sa faible complexité opératoire, le rend très adapté à un traitement temps réel sur site qui est l'objectif de cette étude. Cet algorithme prend place dans un ensemble de traitement paramétrique parfaitement cohérent car la deuxième phase — la reconnaissance, qui fait l'objet d'une publication à paraître [9] — utilisera les mêmes outils de modélisation que ceux développés ici.

## Bibliographie

- [1] P.SOULE - F.CASTANIE: 'Méthode de détection de non-stationnarités par analyse des propriétés de franchissement d'un seuil par l'erreur de prédiction linéaire', GAPSE/ENSEEIH, GRETSI 1983, pp 899-903.
- [2] E.DAYMIER -F. CASTANIE: 'Analyse d'une méthode de détection des sauts de moyenne de variance', GAPSE/ENSEEIH, GRETSI 1985.
- [3] P. SOULE: 'Application des méthodes de modélisation paramétrique à l'analyse et à la segmentation des signaux de télémesures spatiales', Thèse de 3ième cycle I.N.P. Toulouse, 1983.
- [4] J.F. AGNEL: 'Elimination de fouillis en radar et détection séquentielle de cibles: une approche par la reconnaissance des formes et le filtrage en treillis', Thèse de docteur ingénieur E.N.S.T. Paris, Fév 1985.
- [5] M. BASSEVILLE - A. BENVENISTE: 'Détection séquentielle de changements brusques des caractéristiques spectrales d'un signal numérique', IRISA, Publication interne, n°161, Mars 1982.
- [6] U. APPEL - A. BRANDT: 'Performance comparison of two sequential algorithms using growing reference window' dans Proc. of the 6th International Conference on Analysis and Optimization of Systems, INRIA, Nice, Juin 1984.
- [7] L. MARPLE: 'A new autoregressive spectrum analysis algorithm' dans Trans. on ASSP, vol 28, n°4, 1980.
- [8] A. BRANDT: 'Detecting and estimating jumps using ladder algorithms and likelihood ratio tests', ICASSP 1983, pp 1017-1020.
- [9] C. VILLEMUR et al.: 'Modélisation paramétrique et classification automatique de signaux de forme transitoire', GAPSE/ENSEEIH, à paraître GRETSI 87.
- [10] C.VILLEMUR: rapports de contrat GAPSE/ENSEEIH-EDF Etudes et Recherches, 1985-87.
- [11] A. ROUSSILLO-B.GEORGEL: 'Une ébauche de banque de signaux pour la reconnaissance automatique de défauts par courants de Foucault', Note technique externe (1ère et 2ième partie), EDF/Chatou, Juin 1986.
- [12] A. KADE-B.GEORGEL-D.MOYSE: 'Deux méthodes nouvelles de paramétrisation de signaux: application à la classification de défauts de tubes de générateurs de vapeur', rapport d'études, EDF/Chatou SDM, Sept. 1986.