Classification des défauts mécaniques des robinets en utilisant les données de consommation électrique

ANGELA DIGULESCU[1,2], CORNEL IOANA[1], CYRIL PLANTON[3], FRANCK HARS[3]

[1] Gipsa-LAB, Institute National Polytechnique de Grenoble
11 Rue des Mathématiques, 38400 Saint-Martin-d'Hères, France
[2] Académie Technique Militaire « Ferdinand I »
39-49 Boulevard George Cosbuc, 050141 Bucarest, Roumanie
[3] EDF - Division Technique Générale (DTG)
21 Avenue de l'Europe, 38100 Grenoble, France
[1]angela.digulescu@grenoble-inp.fr, cornel.ioana@grenoble-inp.fr,
[3]cyril.planton@edf.fr, franck.hars@edf.fr

Résumé - L'objectif de ce papier est de proposer et de valider une méthodologie d'analyse de données de consommation électrique pour l'alimentation de plusieurs types de robinets afin de pouvoir classifier les différents défauts. L'intérêt opérationnel est de pouvoir réaliser, grâce à cette classification, une détection préventive des défauts tout en utilisant des signaux bien plus facilement accessibles que des mesures mécaniques *in situ*. Cette analyse s'appuie sur une technique mixte décomposition en modes empiriques et représentation en diagramme de phase. Les paramètres extraits avec cette technique permettent une classification des différents types de défauts, réalisée à partir d'une base de données réelle composée des signaux de consommation électrique.

Abstract – The aim of this paper is to propose and validate a methodology for the analysis of the power consumption for the supply of several types of valves in order to classify different faults. The operational interest is to obtain the capability of performing a preventive detection of these faults, using this classification based on the signals easier to access than the mechanic measures on site. This analysis is based on a hydride technique composed by the empirical mode decomposition and the phase diagram representation. The featured extracted by this technique allow the classification of several faults, as illustrated from data issued from real experiments.

1 Introduction

Dans la maintenance conditionnelle sur la robinetterie, l'analyse de données de consommation électrique pour l'alimentation de plusieurs types de robinets est plus facilement accessible que la réalisation des mesures mécaniques. Cette dernière permet de classifier les différents défauts et statuer ainsi sur le bon fonctionnement de ces robinets. Une classification des défauts basée sur les données électriques peut mettre en place une détection préventive des défauts, souhaitable pour la maintenance du système tout en assurant une accessibilité aisée car les signaux de consommation sont bien plus facilement accessibles que les organes mécaniques de robinets.

Ce papier présente quelques méthodes de caractérisation pour chacun des types de défauts, visant à extraire les différents paramètres discriminants pour le type de défaut. Nous avons testé un ensemble des techniques différentes et leur pertinence par rapport aux règles mécaniques de fonctionnement (et, de fait, la pertinence par rapport aux défauts). Ces méthodes ont été validées sur une base de données composée des signaux correspondants à plusieurs régimes de fonctionnement, en présence des défauts et après la réparation (situation normale de fonctionnement).

De façon générale, un signal d'alimentation d'un robinet électrique se compose de deux régimes : le régime transitoire, correspondant aux parties initiales de l'ouverture et de la fermeture des robinets ; le régime permanent, défini par les manœuvres stationnaires entre les ouvertures/fermetures. Les méthodes de caractérisation seront donc définies par rapport à ces deux régimes.

Le papier est organisé comme il suit : la 2eme section présente toutes les méthodes de caractérisation des signaux acquis, la 3eme section discute les résultats et la section finale présente les conclusions.

2 Méthodes de caractérisation

La Figure 1 présente l'architecture générale de traitement de données organisée comme suit :

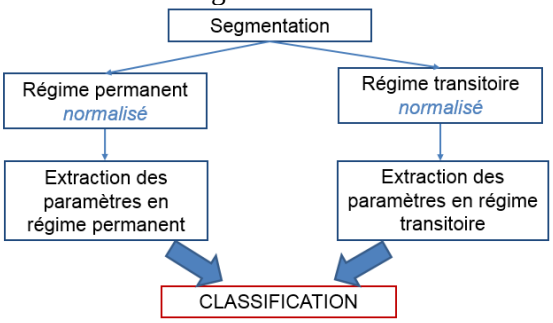


Figure 1: Organisation globale du traitement des données

2.1 La segmentation

La première étape, la segmentation, permet de séparer les parties transitoires des signaux de consommation électrique du celui permanent et ce de façon automatique. Pour ce faire, nous utilisons le spectrogramme du signal, calculé avec l'expression suivante :

$$S(t,f)=\sum_i x(i)h(t-i)e^{-j2\pi f\cdot i} \quad (1)$$

où x est le signal analysé et h est la fenêtre d'analyse de durée D. Dans notre cas, la durée de la fenêtre d'analyse est de 128 échantillons et nous avons également utilisé un recouvrement de 50%.

Une fois le spectrogramme calculé nous déterminons le marginal en temps, obtenu par la sommation de toutes les fréquences, pour chaque instant de temps :

$$X(t)=\sum_f |S(t,f)| \quad (2)$$

A partir du marginal en temps, nous détectons les instants d'apparition des parties utiles des signaux analysés telles que le marginal en temps est supérieur au seuil α que nous fixons par rapport à l'énergie moyenne du spectrogramme :

$$\{D_k\}=\sup\{X(t)\geq\alpha\} \quad (3)$$

A partir de ces instants de temps, le régime permanent (*x_perm*) et les deux régimes transitoires de début (*x_transit_start*) et de fin (*x_transit_stop*), qui peuvent être d'ouverture ou de fermeture, sont définis ainsi :

$$\begin{aligned} x_perm(t)&=x(t), t\in[D_k+10\cdot T_r, D_{k+1}-10\cdot T_r]\\ x_transit_start&=x(t), t\in[D_k, D_k+10\cdot T_r]\\ x_transit_stop&=x(t), t\in[D_k-10\cdot T_r, D_{k+1}] \end{aligned} \quad (4)$$

où T_r est la période du signal.

La figure 2 présente, pour une fermeture d'un robinet, les trois régimes segmentés.

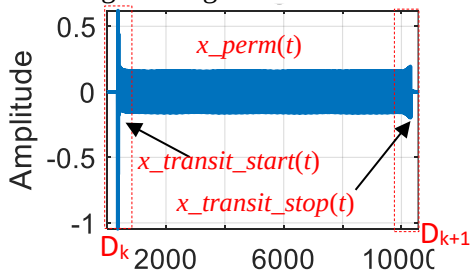


Figure 2: Illustration des signaux correspondants au régime permanent et aux deux régimes transitoires

Les paragraphes suivants présenteront *les méthodes d'analyse adaptées aux deux régimes.*

2.2 Analyse du régime permanent

La première analyse consiste à calculer le spectre du signal, par la *DFT* (Discrete Fourier Transform) que nous nous exprimons en échelle logarithmique :

$$\begin{aligned} X(f)&=10\log_{10}\left|\sum_{n=1}^{N} x_perm[n]\exp\left(-j\frac{2\pi kn}{N}\right)\right|^2 [dB]\\ f&=\frac{k}{N}\frac{F_e}{2}, k=0,1,...,N-1 \end{aligned} \quad (5)$$

où F_e est la fréquence d'échantillonnage (1 kHz, dans notre étude) et N est le nombre d'échantillons du signal *x_perm.*

Le calcul du spectre du signal en régime permanent permet d'estimer précisément les fréquences et de les ***comparer avec les fréquences théoriques*** (associées aux fonctionnements mécaniques). Le niveau moyen du spectre nous permet également d'identifier le premier type de défaut, ***le graissage***. Pour ceci, nous évaluons le plancher moyen du spectre et s'il augmente avec une valeur de 8%, nous pouvons conclure la présence d'un problème de graissage. La valeur de 8% a été déterminée à travers des tests sur les signaux de la base de données. Physiquement, cette augmentation s'explique par l'effort supplémentaire requis similaire à un bruit blanc pour manœuvrer en absence du graissage car la friction sera plus élevée dans ce cas.

La règle d'identification du défaut de graissage est définie par la formule suivante :

$$\begin{aligned} &Si\ E_{moy}=\int_0^{50Hz} X_{actual}(f)df \geq 1.08\%\cdot\int_0^{50Hz} X_{prec}(f)df\\ &Alors\ Défaut\ Graissage \end{aligned} \quad (6)$$

où X_{prec} représente le spectre du signal en régime permanent dans la fenêtre d'analyse précédente alors que X_{actual} représente le spectre du signal dans la fenêtre d'analyse courante et $50\,Hz$ est la fréquence fondamentale du courant..

Un autre paramètre important pour l'analyse du signal en régime permanent est ***l'enveloppe du signal***, soit la modulation à faible fréquence qui modifie l'amplitude de la composante sinusoïdale principale du signal d'alimentation électrique. A partir du signal en régime permanent *x_perm* l'enveloppe est extraite à l'aide de la transformée de Hilbert qui est utilisée pour déterminer, à partir du signal réel, la composante en quadrature du signal x :

$$\mathrm{H}x(t)=x(t)\otimes\frac{1}{\pi t}=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(t)}{t-\tau}d\tau \quad (7)$$

où $\otimes$ est l'opérateur de convolution. L'enveloppe est définie avec l'expression mathématique :

$$Env(t)=|\mathrm{H}x(t)| \quad (8)$$

La présence d'une ***forme harmonique*** de l'enveloppe peut témoigner d'un défaut visible en régime permanent. C'est pourquoi notre méthodologie d'analyse du signal permanent passe par la détection de la forme harmonique de l'enveloppe qui se fera dans le domaine spectrale. Pour ceci, le spectre de l'enveloppe est calculé et comparé, dans une bande B_{env} = 0.5-20 Hz, avec un seuil de -25 dB, déterminé expérimentalement sur l'ensemble des signaux permanents ayant une enveloppe d'amplitude :

$$\begin{aligned} \hat{E}nv(f)&=10\log_{10}\left|\sum_{n=1}^{N} Env[n]\exp(-j\cdot 2\pi\cdot f\cdot n)\right|^2\\ Si\int_{B_{env}}&\hat{E}nv(f)df\geq -35\,dB\Rightarrow Env.harmonique \end{aligned} \quad (9)$$

La présence de la modulation d'amplitude harmonique pendant le régime permanent peut être due à un **défaut de centrage arcade ou à un défaut de tige**, c'est pourquoi une analyse plus approfondie est nécessaire pour pouvoir classifier ces types de défauts.

Nous analysons cette variation à travers l'évolution ***de la variance au cours du temps***. Pour ceci, l'enveloppe

est segmentée en fenêtres de N_w=200 échantillons et pour chaque segment, la variance est calculée :

$$e_k(t) = Env(t) \cdot h_D(t - t_k)$$
$$D = 200; t_k = k \cdot D \quad (10)$$
$$V_k = \frac{\sum_{i=1}^{N} \left(e_k[i] - \frac{1}{N} \sum_{i=1}^{N} e_k[i] \right)^2}{N}$$

où $h_D(\cdot)$ est la fenêtre rectangulaire de largeur D.

L'ensemble des valeurs locales de la variance $\mathbf{V}=[V_1,\ldots,V_k]^T$ sera utilisé pour réaliser une régression linéaire, définie par :

$$\mathbf{Y} = b + c \cdot \mathbf{V}$$
$$avec\ b, c\ estimés\ par \min_{b,c} (\mathbf{Y} - b - c \cdot \mathbf{V}) \quad (11)$$

La pente de la courbe de régression **Y** permet de distinguer les défauts de centrage arcade et les défauts de tige. Pour ceci, la pente de la courbe de régression linéaire est comparée au seuil de 0.005 (calculé sur des fenêtres de régression linéaires de 5000 échantillons) déterminé expérimentalement, pour assurer un taux de bonne classification proche de 100% :

$$Si\ c > 0.005 \Rightarrow Défaut\ \text{Centrage Arcade}$$
$$\text{Si } non \Rightarrow Défaut\ \text{Tige} \quad (12)$$

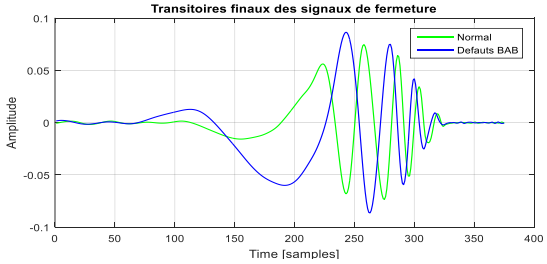
2.3 Analyse du régime transitoire

Nous avons analysé plusieurs classes de techniques de caractérisation des parties transitoires du signal d'alimentation [1]-[4], correspondantes aux fermetures/ ouvertures des robinets. Nous avons démontré l'intérêt pour la modélisation des signaux transitoires *x_transit_start*(*t*) et *x_transit_stop*(*t*) par *Empirical Mode Decomposition (EMD)* :

$$x_transit_start(t) = \sum_{k=1}^{N_{mod}} IMF_k^{start}(t) + \text{Residual}$$
$$x_transit_stop(t) = \sum_{k=1}^{N_{mod}} IMF_k^{stop}(t) + \text{Residual} \quad (13)$$

où IMF_k sont les fonctions modales intrinsèques d'ordre k ($k=1,\ldots,N_{mod}$; N_{mod} est le nombre de modes considérés dans la décomposition ; 6 dans notre étude) et **Residual** est le résidu après la décomposition sur N_{mod} modes. Pour la simplicité des notations, nous allons ignorer l'indice du type de la manœuvre – *start ou stop*.

Dans nos différentes analyses, nous avons trouvé que le mode 1 (IMF_1) porte généralement des informations pertinentes pour la discrimination entre le fonctionnement normal et celui en présence de différents types de défauts (Figure 3).



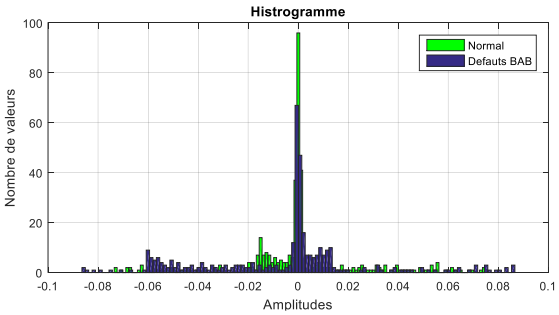


Figure 3 : Exemple d'IMF avant l'intervention (présence du défaut) et après l'intervention (considéré comme fonctionnement normal). Les deux signaux correspondent aux transitoires finaux d'une fermeture, en situation normale et en présence d'un défaut BAB (boite à butées)

Le cas normal indique une variation plus régulière avec une concentration de l'histogramme sur quelques valeurs dominantes alors qu'en présence du bruit, la concentration de l'histogramme s'étale sur plusieurs valeurs. Afin de quantifier cette observation, que nous avons pu faire sur plusieurs types de défauts, nous proposons l'utilisation de ***l'entropie*** qui est une mesure du degré de désordre dans un vecteur de données. Par définition, l'entropie est calculée, pour un signal $x(t)$, avec la formule suivante :

$$Ent\{x(t)\} = -\sum_{k=1}^{N_v} p_k \log_2 p_k \quad (14)$$

où p_k est la probabilité pour que la valeur du signal x soit x_k, avec k entre 1 et N_v le nombre total des valeurs distinctes du signal x. Les probabilités p_k sont calculées à l'aide de l'histogramme qui quantifie les nombres de fois que le signal passe par les valeurs distinctes d'amplitude sur le nombre total d'échantillons du signal, N.

Les valeurs de l'entropie du IMF 1 seront utilisées à la fois pour identifier les éventuels défauts mais également, comme on le verra par la suite, pour classifier les différents types de défauts.

Mais la variation de l'entropie, bien qu'elle donne une idée du degré d'ordre qui accompagne la dynamique d'un signal, ne peut pas à elle seule permettre la distinction entre les différents types de défauts. C'est pourquoi nous nous sommes intéressés en parallèle à une approche qui permet de quantifier *la dynamique des transitoires*.

2.4 Analyse signal complet

Il s'agit de la représentation des signaux en ***diagrammes de phase*** (phase diagramme analysis). Pour un signal décrit par ses échantillons $x(t) = \{x[k]\}, k = 1,\ldots,N$, le diagramme de phase *PhD* à deux dimensions aura pour définition :

$$PhD(i,k) = [x[k], x[k + Lag]]^T \quad (15)$$

où $i=1, 2$ et Lag est le retard entre les échantillons formant un vecteur bidimensionnel. Dans ce papier le Lag a le valeur 4.

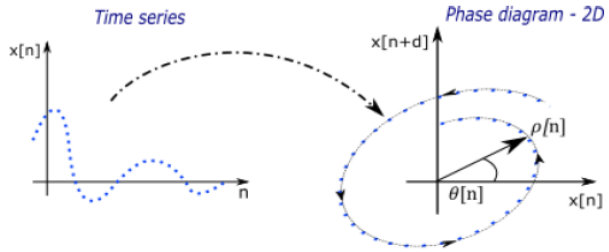


Figure 4: Le principe de l'analyse en diagramme de phases

A partir du diagramme de phase, les coordonnées polaires de la trajectoire, définies par le rayon *Rho*[*k*] et l'angle par rapport à l'axe *x*[*k*], $\theta[k]$, sont calculées par :

$$\rho[k] = \sqrt{PhD(1,k)^2 + PhD(2,k)^2}$$
$$\theta[k] = \arctan\frac{PhD(2,k)}{PhD(1,k)} \quad (16)$$

Avec ces éléments nous évaluons la ***surface de l'ellipse*** décrite par le vecteur *v* pendant une rotation complète, soit une période du signal transitoire.

$$Area[k] = \sum_{i=k}^{k+N_k-1} \frac{Rho[i]\cdot Rho[i+1]}{2}\cdot\sin(\theta[i+1]-\theta[i]) \quad (17)$$

sous la contrainte que $\sum_{i=k}^{k+N_k}\theta[i] = \pi$.

Ce paramètre sera ensuite utilisé pour évaluer la ***dynamique de l'angle*** de la trajectoire avec la formule :

$$Angle = \sum_{i=L}^{N-1}\frac{Area[i+1]-Area[i]}{i\cdot(N-L)} \quad (18)$$

où *N* est le nombre d'échantillons et *L* représente l'indice temporel pour lequel la surface devient plus grande que la surface correspondante au centre du signal :

$$L = \arg_k\left\{Area[k] \geq 1.05\cdot Area\left[\frac{N}{2}\right]\right\} \quad (19)$$

Avec l'ensemble des paramètres d'entropie du mode 1 en régime transitoire d'ouverture-fermeture (start/stop) et les angles de la variation du diagramme de phase sur le signal complet, nous définissons le vecteur des descripteurs du régime transitoire.

3 Résultats

Les paramètres estimés pour les deux régimes permettent l'identification des différents types de défauts. Pour une visualisation rapide et intuitive, nous mettons dans un diagramme 3D les paramètres de classification (l'entropie du début et fin du transitoire et l'angle décrit par l'aire du diagramme des phases), constitué par les clusters des données déterminés par l'analyse de la totalité de la base de données.

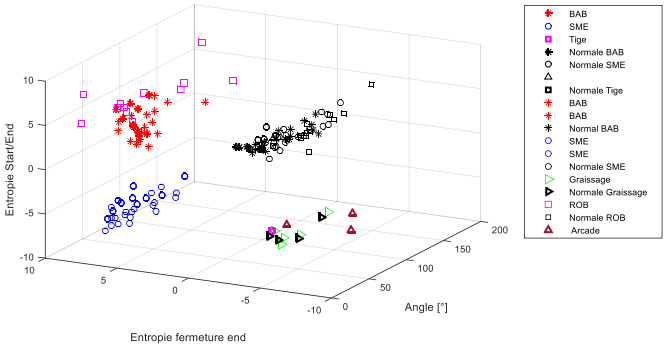


Figure 5: Visualisation des résultats de classification des défauts « visibles » en régime transitoire avec trois paramètres de classification : entropies de l'IMF 1 pour le start et, respectivement, stop ; la dynamique de l'angle de l'analyse en diagramme de phase

La classe normale a été constituée par les paramètres issus de l'analyse des signaux ***après*** l'intervention alors que chacune des classes a été constituée par les vecteurs d'analyse déterminés par le traitement des signaux avant l'intervention –Figure 5.

Le clustering final, issu de cette analyse avec les seuils et paramètres considérés dans la section 2, constitue l'étape importante du processus *Machine Learning* qui permettra, en conditions opérationnelles, d'identifier et de classifier chaque défaut rencontré.

4 Conclusions

Les travaux de recherche menés dans le cadre de ce projet ont permis de définir plusieurs algorithmes d'analyse des signaux de consommation électrique pour identifier et classifier les différents types de défauts mécaniques, ce qui est l'objectif central de ces travaux. Nous avons démontré que, pour répondre complétement à cet objectif, une analyse distincte en régimes permanent et transitoire est nécessaire et, pour s'affranchir de tout modèle réducteur, nous avons concentré nos analyses autour des méthodes *data driven* : EMD et analyse en diagramme de phase.

Les différents paramètres extraits ont permis d'identifier et de classer les différents défauts et la validation des critères a été effectuée sur l'ensemble de la base de données. Les perspectives de ce projet envisagent la définition d'une approche bayésienne de classification automatique à partir du clustering actuel et en rajoutant des résultats futurs pour conforter le diagnostic.

Bibliographie

[1] J. Zbilut, C. Webber Jr., "Embeddings and delays as derived from quantification of recurrence plots", Physical Letters A , 171, 199-203,1992.

[2] N. Marwan, S. Schinkel, J. Kurts, "Recurrence plots 25 years later – Gaining confidence in dynamic transitions", Europhysics Letters, 101, 20007, 2013.

[3] A. Digulescu, I. Murgan, C. Ioana, I. Candel, A. Serbanescu, "Applications of Transient Signal Analysis using the concept of recurrence plot analysis", Recurrence Plots and Their Quantifications: Expanding Horizons, Springer Proceedings in Physics, 2016.

[4] P. Flandrin, « Explorations in Time-Frequency Analysis », Cambridge University Press, 2018