



**HAL**  
open science

## Classification contextuelle pour système d'aide à la décision pour machines-outils

Zhiqiang Wang, Mathieu Ritou, Catherine M. da Cunha, Benoit Furet

► **To cite this version:**

Zhiqiang Wang, Mathieu Ritou, Catherine M. da Cunha, Benoit Furet. Classification contextuelle pour système d'aide à la décision pour machines-outils. Colloque National S-mart/AIP-PRIMECA, Apr 2019, Les Karellis, France. hal-02100713

**HAL Id: hal-02100713**

**<https://hal.science/hal-02100713>**

Submitted on 16 Apr 2019

**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.

# Classification contextuelle pour système d'aide à la décision pour machines-outils

Zhiqiang Wang

Laboratoire des Sciences du Numérique de Nantes  
(LS2N, UMR CNRS 6004), Université de Nantes  
2 av. J Rouxel, 44475 Carquefou – France  
[Zhiqiang.wang@univ-nantes.fr](mailto:Zhiqiang.wang@univ-nantes.fr)

Catherine Da Cunha

Laboratoire des Sciences du Numérique de Nantes  
(LS2N, UMR CNRS 6004), Ecole Centrale de Nantes  
1 rue de la Noë, 44321 Nantes – France  
[Catherine.da-cunha@ec-nantes.fr](mailto:Catherine.da-cunha@ec-nantes.fr)

Mathieu Ritou

Laboratoire des Sciences du Numérique de Nantes  
(LS2N, UMR CNRS 6004), Université de Nantes  
2 av. J Rouxel, 44475 Carquefou – France  
[Mathieu.ritou@univ-nantes.fr](mailto:Mathieu.ritou@univ-nantes.fr)

Benoît Furet

Laboratoire des Sciences du Numérique de Nantes  
(LS2N, UMR CNRS 6004), Université de Nantes  
2 av. J Rouxel, 44475 Carquefou – France  
[Benoit.furet@univ-nantes.fr](mailto:Benoit.furet@univ-nantes.fr)

**Résumé** — Dans le contexte général de l'Industrie 4.0, une entreprise de fabrication moderne dispose de nombreuses données numériques qui pourraient être utilisées pour rendre les machines-outils plus intelligentes et faciliter la prise de décision en matière de gestion opérationnelle. L'une des premières étapes de l'approche d'exploration de données est la sélection précise de données pertinentes. Pour ce faire, les données brutes doivent être classées dans différents groupes de contextes. Cet article présente une étude comparative d'algorithmes d'apprentissage automatique non-supervisé pour la classification contextuelle ; qui est utile pour un système d'aide à la décision pour machines-outils. Les vérifications par fouilles manuelles montrent que la méthode GMM permet d'obtenir de bons résultats de classification contextuelle, contrairement à celle des K-means.

**Mots-clés** — UGV, Industrie 4.0, apprentissage non-supervisé.

## I. INTRODUCTION

L'Usinage à Grande Vitesse (UGV) a fortement augmenté les vitesses de coupe par rapport à l'usinage conventionnel. Cependant, en UGV, l'opérateur ne peut pas correctement détecter les incidents lorsqu'il gère plusieurs machines dans une cellule de production. De plus, il n'a pas le temps d'arrêter la machine à temps. Par conséquent, un système de surveillance est nécessaire pour protéger la machine-outil et les pièces à usiner (surtout les pièces aéronautiques qui sont à forte valeur ajoutée). En outre, les pièces en alliage d'aluminium sont usinées dans la masse et peuvent mesurer plus de 10m de long, elles ont des formes complexes et des épaisseurs finales très faibles.

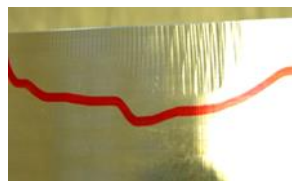


Fig.1 Exemple de bris d'outil (à gauche) et de broutement (à droite) [4]

## A. Etat de l'art sur les systèmes de surveillance de machine-outil

Des articles reprenant l'état de l'art de la surveillance de l'usinage listent l'ensemble des défauts mesurables et étudiés [1]. Les principaux sont, le broutement, le bris d'outil, l'usure outil (Figure 1). Quintana et al. [2] présentent un état de l'art sur le broutement et les méthodes existantes pour le détecter et l'éviter. Zhou et al. [3] présentent une revue des méthodes utilisées pour la surveillance de l'usure d'outil dans les processus de fraisage, y compris les capteurs, l'extraction de caractéristiques et les modèles de surveillance. Godreau et al. [4] présentent le lien entre la qualité de la surface et le broutement. Dans le contexte de l'Industrie 4.0, des données numériques abondantes sont disponibles dans une entreprise de fabrication moderne et pourraient être utilisées pour rendre les machines-outils plus intelligentes et pour aider les prises de décisions opérationnelles. Lenz et al. [5] proposent une approche holistique de l'analyse des données des machines-outils afin de combiner les tâches et de regrouper les objectifs d'analyse entre différents départements dans l'entreprise. Morgan et al. [6] présentent la conception et le développement d'un système de surveillance de processus cyber-physiques pour les machines-outils. Liu et al. [7] proposent une méthode de développement systématique pour Machines-outils 4.0. Afin d'extraire les données des bases de données de manière plus efficace et plus pertinente, des algorithmes d'apprentissage automatiques sont développés.

## B. Etat de l'art sur les algorithmes d'apprentissage automatique

L'apprentissage automatique (Machine Learning) est un champ d'étude de l'intelligence artificielle qui se base sur des approches statistiques et qui peut notamment être appliqué aux signaux de machines-outils. Il est catégorisé par deux modes d'apprentissage : apprentissage supervisé et apprentissage non-supervisé. L'apprentissage supervisé apprend à classer à partir d'échantillons de sortie déjà étiquetés. Il a pour but de faire des prédictions correctes sur des données non présentes dans l'ensemble d'apprentissage. Au contraire, l'apprentissage non-supervisé apprend à classer des données non étiquetées. De

nombreuses industries étudient et utilisent l'apprentissage automatique pour améliorer leurs processus actuels. Kim et al. [8] listent et résument les contributions en usinage à l'aide d'algorithmes d'apprentissage automatique. Il existe différents types d'apprentissage supervisé : les réseaux de neurones, les machines à vecteurs de support, la méthode des plus proches voisins, l'arbre de décision etc. Denkena et al. [9] présentent une méthode qui utilise les machines à vecteurs de support pour modéliser les données de processus obtenues afin d'optimiser automatiquement les machines-outils et prédire la durée de vie d'outil. Krishnakumar et al. [10] comparent les réseaux de neurones et les arbres de décision pour la surveillance de bris d'outil lors de l'usinage de titane en utilisant les signaux de vibration. Gouarir et al. [11] présentent une méthode de prédiction de l'usure des outils en cours de processus basée sur les réseaux de neurones à convolution en utilisant les signaux d'efforts. Dans l'industrie, il n'est généralement pas possible de collecter des données de sortie déjà étiquetées. Par conséquent, les méthodes d'apprentissage non-supervisé sont plus adaptées pour les applications industrielles.

Il existe différents types d'apprentissage non-supervisé : K-means, classification hiérarchique (ascendante ou descendante) [12], des estimations de densité de distribution (ex. modèle de mélange gaussien - GMM). Bhangue et al. [13] construisent un modèle de prédiction d'énergie pour usiner une pièce en appliquant un algorithme de régression gaussienne. Godreau et al. [14] calculent les seuils de critères en appliquant la méthode de l'estimateur du maximum de vraisemblance (Maximum Likelihood Estimation - MLE). K-means est une méthode de segmentation de données. Étant donné des points et un entier  $k$ , le problème est de diviser les points en  $k$  clusters, de façon à minimiser un calcul de distances. Son avantage est que sa complexité est linéaire ce qui le rend applicable à de grandes volumes de données. Tandis que son inconvénient est que les clusters dépendent de l'initialisation, de la distance choisie ainsi que de l'équilibre entre les classes. Alors que GMM peut correctement classifier les données avec des classes déséquilibrées.

Des recherches sont faites en utilisant ces algorithmes directement sur les données brutes pour rendre les machines-outils plus intelligentes. Cependant, peu de recherches se concentrent sur les prétraitements de données pour identifier le contexte, et permettre une sélection très fine des données d'usinage. Cet article présente une étude comparative d'algorithmes d'apprentissage automatique non-supervisé pour la classification contextuelle ; qui est utile pour un système d'aide à la décision pour machines-outils. Les méthodes sont appliquées à une base de données d'usinage collectée dans l'industrie. Des vérifications par fouilles manuelles sont ensuite effectuées pour analyser les résultats de classification.

## II. DEMARCHE D'ANALYSE DES DONNEES

### A. Cadre de l'étude

Les travaux présentés dans cet article participent au projet ANR SmartEmma qui vise à développer des machines-outils intelligentes et connectées pour l'UGV. La Figure 2 présente le principe d'exploitation des données enregistrées. Un dispositif, appelé Emmatools, collecte les données mesurées pendant l'usinage et les stocke dans une base de données. Les méthodes d'apprentissage automatique (Machine Learning) vont être

développées pour d'une part effectuer des classifications contextuelles, et d'autre part agréger les données. Notons qu'1Go de données est collecté par jour et par machine-outil. Avant l'exploration de données, il est souhaitable de faire une classification contextuelle. En effet, une meilleure sélection de données est préférable pour un calcul plus pertinent d'indicateurs de performance (KPI), réduisant le bruit. Enfin, des KPI et des méthodes d'analyse des données seront définis pour améliorer l'efficacité du processus UGV. Les KPI ainsi évalués deviennent un support d'aide à la décision pour le pilotage de l'entreprise.

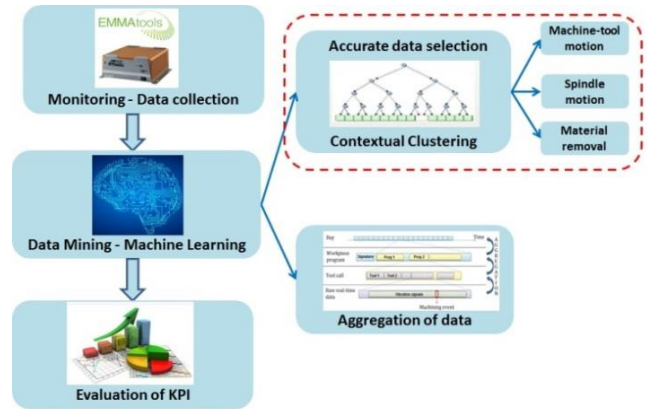


Fig.2 Processus de fouille des données d'usinage

### B. Données collectées

Emmatools est un système de surveillance et de collecte de données en usinage développé par le laboratoire LS2N pendant le projet FUI UsinAE [15]. Il est installé sur des machines-outils dans des usines de constructeurs aéronautiques.

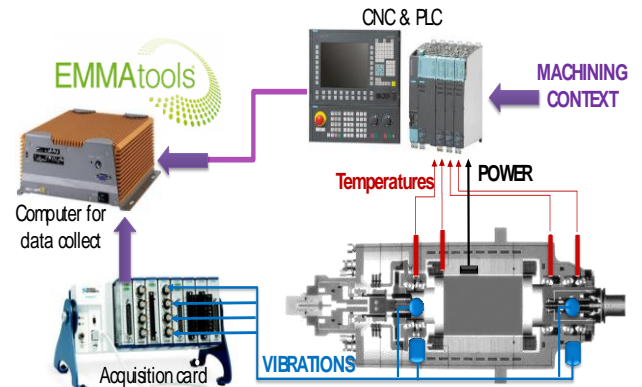


Fig.3 Schéma de principe de l'Emmatools

La Figure 3 présente le schéma de principe de l'Emmatools. Il collecte les données chaque 0,1s. Les données sont issues de deux sources d'information : la commande numérique et les capteurs ajoutés. La commande numérique fournit des informations de contexte ainsi que certaines données de capteurs déjà présents. Les données de contexte correspondent aux informations telles que l'outil en broche, le programme en cours... Les données de capteurs déjà présents comprennent la vitesse de rotation de la broche, les différentes vitesses d'axe, la puissance instantanée de la broche... Afin de compléter ces mesures, 4 accéléromètres sont intégrés à la broche (radialement à ses paliers).

		Vitesse d'avance de la machine-outil Vf							
		Vf=0		Vf>0					
				$\Delta Vf > T_{\Delta Vf}$			$\Delta Vf \leq T_{\Delta Vf}$		
				$Arms \leq T_{Arms}$ & $P \leq P_{vide}$	$Arms > T_{Arms}$ ou $P > P_{vide}$		$Arms \leq T_{Arms}$ & $P \leq P_{vide}$	$Arms > T_{Arms}$ ou $P > P_{vide}$	
$\Delta P \leq T_{\Delta P}$	$\Delta P > T_{\Delta P}$	$\Delta P \leq T_{\Delta P}$	$\Delta P > T_{\Delta P}$						
Rotation de la broche N	N=0	Broche : A Machine : A Usinage : N	Broche : Arrêtée (A) Machine-outil : Vitesse Variante (VV) Usinage : Non (N)			Broche : Arrêtée (A) Machine-outil : Vitesse Constante (VC) Usinage : Non (N)			
	N>0	$\Delta N > T_{\Delta N}$	Broche : VV Machine : A Usinage : N	Broche : VV Machine : VV Usinage : Non (N)	Broche : VV Machine : VV Usinage : eng. Constant (UC)	Broche : VV Machine : VV Usinage : eng. Variant (UV)	Broche : VV Machine : VC Usinage : Non (N)	Broche : VV Machine : VC Usinage : eng. Constant (UC)	Broche : VV Machine : VC Usinage : eng. Variant (UV)
		$\Delta N \leq T_{\Delta N}$	Broche : VC Machine : A Usinage : N	Broche : VC Machine : VV Usinage : N	Broche : VC Machine : VV Usinage : UC	Broche : VC Machine : VV Usinage : UV	Broche : VC Machine : VC Usinage : N	Broche : VC Machine : VC Usinage : UC	Broche : VC Machine : VC Usinage : UV

Tableau 1. Classification des différents états de la machine-outil

### C. Classification contextuelle

La classification contextuelle est nécessaire pour un calcul plus pertinent d'indicateurs (KPI). Il est intéressant de savoir :

Si la broche est arrêtée ou si elle tourne (à vitesse constante ou variante),

Si la machine-outil est arrêtée ou si elle avance (à vitesse constante ou variante)

Si l'outil n'usine pas ou s'il usine (avec un engagement de l'outil constant ou variant).

En combinant ces trois informations contextuelles, il y a 17 états potentiels pour la machine-outil. Le Tableau 1 présente ces 17 états avec leurs conditions.

Soit une variable X (ex. Vf ou N). La classification va être effectuée à partir des variations de X, notée  $\Delta X$  (ex.  $\Delta Vf$  ou  $\Delta N$ ) qui est défini comme suit :

$$\text{éq. (1): } \Delta X = \frac{X_n - X_{n-1}}{\Delta t}$$

$X_n$  est la mesure à l'enregistrement  $n$ ,  $X_{n-1}$  est celle de l'enregistrement précédant et  $\Delta t$  est la période d'échantillonnage des données ( $=0,1s$ ). L'objectif est de calculer un seuil  $T_{\Delta X}$  (ex.  $T_{\Delta Vf}$ ,  $T_{\Delta N}$ ) par des algorithmes d'apprentissage automatique pour permettre la classification.

Il y a trois clusters pour la vitesse d'avance de la machine-outil Vf (m/min): arrêtée (Vf=0); avance à vitesse constante (Vf>0 et  $\Delta Vf \leq T_{\Delta Vf}$ ); avance à vitesse variante (Vf>0 et  $\Delta Vf > T_{\Delta Vf}$ ); ainsi que trois clusters pour la vitesse de rotation de la broche N (tr/min) : arrêtée (N=0); tourne à vitesse constante (N>0 et  $\Delta N \leq T_{\Delta N}$ ); tourne à vitesse variante (N>0 et  $\Delta N > T_{\Delta N}$ ).

Il est également possible de savoir si l'outil usine ou pas. Pour cela il est proposé d'effectuer une classification suivant les vibrations en usinage, par le critère Arms (valeur efficace d'accélération, root mean square, en  $m/s^2$ ) et la puissance P

(kW) quand la vitesse d'avance Vf n'est pas nulle. En usinage, il est admis que si l'outil usine, sa vibration (Arms) est plus grande que si l'outil n'usine pas, il est aussi admis que la puissance (P) va dépasser la puissance de la rotation à vide de la broche ( $P_{vide}$  quand l'outil n'usine pas). C'est-à-dire, si  $P > P_{vide}$  ou  $Arms > T_{Arms}$ , l'outil usine; si  $P \leq P_{vide}$  et  $Arms \leq T_{Arms}$ , l'outil n'usine pas. L'objectif est de trouver le seuil  $T_{Arms}$  ( $m/s^2$ ) et la puissance à vide  $P_{vide}$  qui permet de déterminer les clusters « l'outil usine » et « l'outil n'usine pas ».

Il est aussi possible de savoir si l'outil usine avec un engagement de l'outil constant ou variant. Comme précédemment, l'objectif est de trouver le seuil de classification  $T_{\Delta P}$  (kW/s).

Chaque machine a des caractéristiques différentes, il faut donc effectuer un apprentissage initial pour chaque machine. Les données d'une journée d'usinage sont utilisées pour cela.

### III. METHODE DE CLASSIFICATION CONTEXTUELLE

Nous comparons ici des méthodes d'apprentissage non-supervisé sur les données venant d'une usine de production de pièces de structure aéronautique du groupe Mécachrome. Les 2 méthodes qui ont été testées sont K-means et GMM.

#### A. Classification du mouvement de la machine

Les performances des algorithmes de K-means et GMM ont été testées pour la détermination de seuils de classification contextuelle, basée sur la vitesse d'avance Vf (m/min). Le contexte des enregistrements pour lesquels Vf=0 est considéré comme connu : « machine-outil arrêtée ».

#### Utilisation de K-mean

K-means est appliqué aux variations de vitesse  $\Delta Vf$  afin de les classer en deux clusters, qui conduisent ensuite à trois états de la machine, définis Tableau 1. Le seuil  $T_{\Delta Vf}=466,2$  mm/s<sup>2</sup> est trouvé. La Figure 4 représente le résultat de cette classification appliquée à  $\Delta Vf$  en fonction du temps sur 17h. On peut visualiser les 3 états.

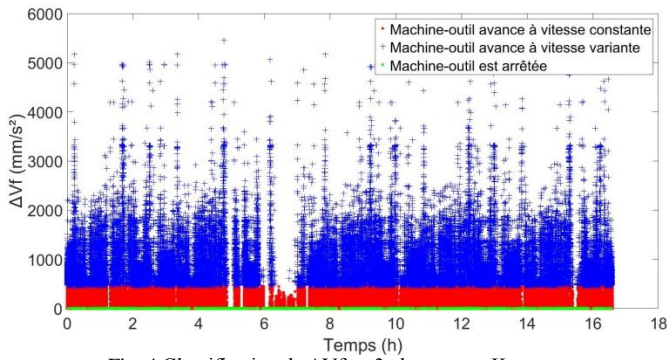


Fig. 4 Classification de  $\Delta Vf$  en 3 clusters par K-means

Le cluster « machine arrêtée » est en vert, le cluster « machine-outil avançant à vitesse constante » est en rouge et le cluster « machine-outil avançant à vitesse variable » est en couleur bleu. La performance de K-means peut être vérifiée par fouille manuelle sur Vf, en traçant les 3 clusters à partir du seuil résultant de l'apprentissage. La Figure 5 illustre les classifications obtenues sur une période de 60s.

On trouve qu'il a des erreurs de classification pour des points identifiés comme à vitesse constante alors que la machine accélère. Nous pouvons en déduire que le seuil  $T_{\Delta Vf}=466,2$  mm/s<sup>2</sup> obtenu par K-means est trop grand.

Nous vérifions donc ici que K-means ne donne pas de bons résultats sur les données avec des classes déséquilibrées [16]. Ce qui est le cas dans les données  $\Delta Vf$ .

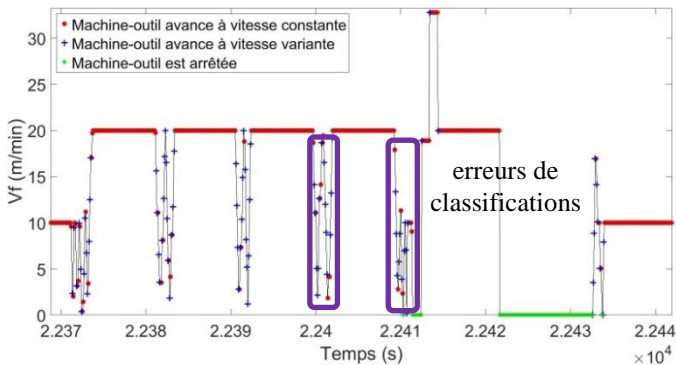


Fig.5 Classification de Vf en 3 clusters par K-means (sur 60s)

#### Utilisation de GMM

Un modèle de mélange gaussien (Gaussian Mixture Model, GMM) est une fonction de densité de probabilité paramétrique représentée par la somme pondérée des densités de composantes gaussiennes [17]. Les paramètres de GMM sont estimés à partir des données d'apprentissage à l'aide de l'algorithme d'espérance-maximisation (EM).

GMM est appliqué sur les  $\Delta Vf$  afin de les classer en deux clusters, qui conduisent ensuite à trois états de la machine (cf. Tableau 1). Selon la densité de probabilité de  $\Delta Vf$ , il y a 37.6% des données où  $\Delta Vf=0$  (le grand pic central). Donc, on enlève tous les  $\Delta Vf=0$  avant d'appliquer GMM parce que les  $\Delta Vf=0$  sont déjà labellisés comme 'Vf constante'. Cette distribution de  $\Delta Vf$  peut être modélisée par 2 gaussiennes (Y1 et Y2), la distribution Y3 est la somme de Y1 et Y2, voir Figure 6 :

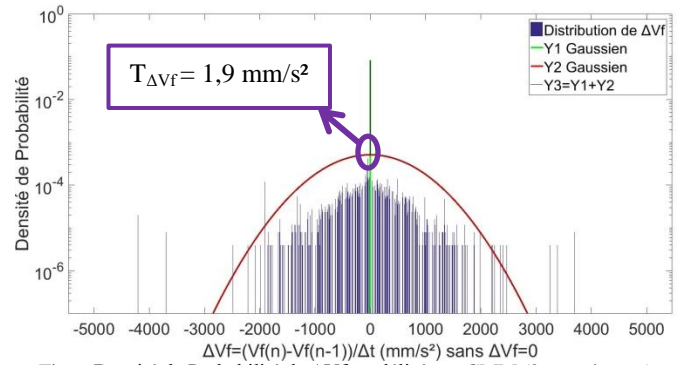


Fig. 6 Densité de Probabilité de  $\Delta Vf$  modélisé par GMM (2 gaussiennes)

Selon GMM, le seuil  $T_{\Delta Vf}$  est obtenu comme l'intersection des deux gaussiennes ( $T_{\Delta Vf}=1,9$  mm/s<sup>2</sup>) afin de minimiser les erreurs de diagnostic (faux positifs ou faux négatifs). Après les fouilles de données manuelles, aucune erreur n'est observée sauf les points de 'front montant' et les points de 'front descendant'. Tous les points de 'front montant' et les points de 'font descendant' sont affectés dans le cluster « Machine-outil avançant à vitesse constante ». Les résultats de classification sont présentés sur la Figure 7 (sur le même jeu de données que pour la Figure 5).

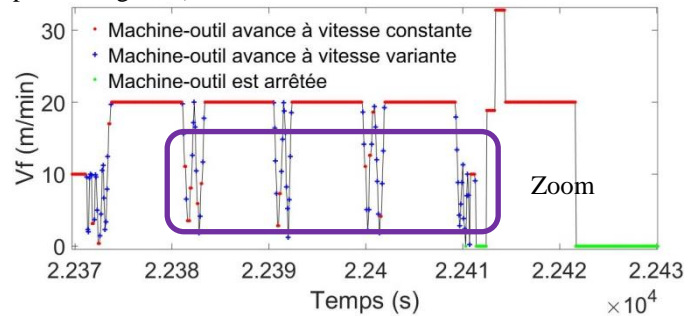


Fig. 7 Classification de Vf en 3 clusters par GMM (sur 60s)

Il y a des points identifiés comme « machine-outil avançant à vitesse constante » alors que la vitesse d'avance semble varier. Après zoom, la vitesse d'avance est bien constante dans les zones, comme on peut le constater sur la Figure 8.

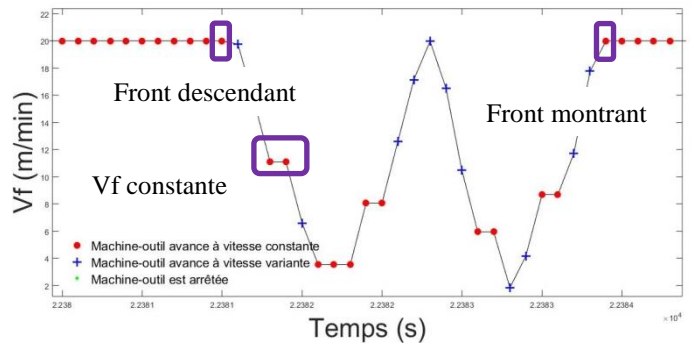


Fig. 8 Classification de Vf en 3 clusters par GMM (zoom).

La classification de Vf est donc bien faite selon GMM.

#### B. Classification de la rotation de la broche

Le contexte des enregistrements pour lesquels  $N=0$  est considéré comme connu : « broche arrêtée ».

On enlève tous les  $\Delta N=0$  avant d'appliquer GMM car les  $\Delta N=0$  sont déjà labellisés comme « broche tournant vitesse constante ». La distribution de  $\Delta N$  (sans  $\Delta N=0$ ) peut être modélisée par 2 gaussiennes : un gaussien autour de zéro (Y1) ; un autre gaussien Y2 sont les données « broche tournant à

vitesse variable », Y3 est la somme des 2 gaussiennes (Y3=Y1+Y2), voir la Figure 9.

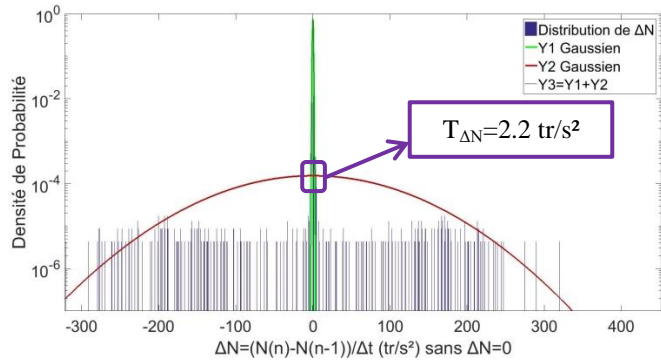


Fig.9 Densité de Probabilité de ΔN (sans 0) modélisé par GMM (2 gaussiennes)

Selon GMM, le seuil  $T_{\Delta N}$  est obtenu comme l'intersection des deux gaussiennes ( $T_{\Delta N}=2,2 \text{ tr/s}^2$ ) afin de minimiser les erreurs de diagnostic. Tous les points de 'front montant' et les points de 'font descendant' sont mis dans le cluster « broche tournant à vitesse constante » comme dans la classification de la vitesse d'avance Vf. La Figure 10 représente N en fonction du temps pendant 17h avec un seuil  $T_{\Delta N}=2,2 \text{ tr/s}^2$ .

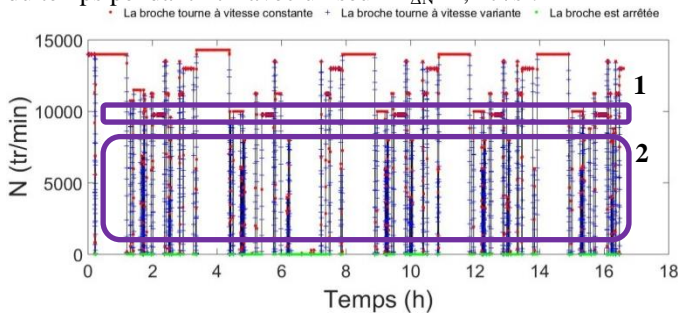


Fig. 10 Classification de N en 3 clusters selon  $T_{\Delta N}=2,2 \text{ tr/s}^2$

Dans la zone 1, il y a des points identifiés comme « broche tournant à vitesse variable » alors que la vitesse de la broche semble constante. Après zoom, la vitesse de la broche est bien variable dans les zones, voir Figure 11 pour un zoom sur la zone 1.

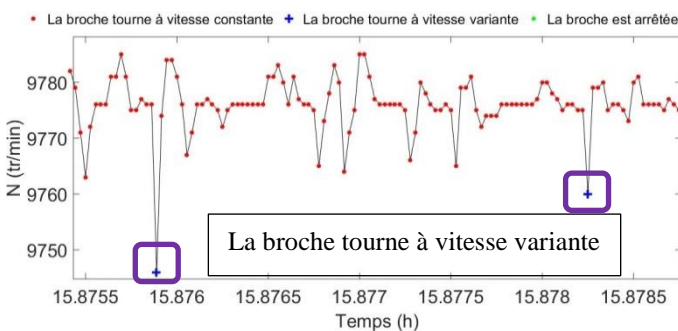


Fig. 11 Classification de N en 3 clusters selon  $T_{\Delta N}=2,2 \text{ tr/s}^2$  en zoom la zone 1

Dans la zone 2 il y a des points identifiés comme « broche tournant à vitesse constante » alors que la vitesse de la broche semble variable. Après zoom, la vitesse de la broche est bien constante dans la zone 2.

### C. Classification de l'engagement de l'outil

Ici, la méthode de GMM est appliquée sur les enregistrements pour lesquels « la broche tourne à vitesse constante » et  $V_f > 0$  (cf. Tableau 1). Le seuil pour le cluster « broche tournant à

vitesse constante » est obtenu par la classification contextuelle sur N :  $T_{\Delta N}=2,2 \text{ tr/s}^2$  (cf. section III.B).

La distribution de Arms dans ce cluster-là, peut être modélisée par 3 gaussiennes : une gaussienne représente que « l'outil n'usine pas » (Y1 vert), une autre représente que « l'outil usine » (Y3 bleu), la 3<sup>ème</sup> gaussienne (Y2 rouge) représente de fortes vibrations d'usinage, la distribution de Y4 est la somme des 3 gaussiennes ( $Y4=Y1+Y2+Y3$ ). Le seuil  $T_{Arms}$  se trouve dans l'intersection entre les deux gaussiennes Y1 et Y3  $T_{Arms}=6,24 \text{ m/s}^2$ , afin de minimiser les erreurs de classification (faux positifs ou faux négatifs, cf. Figure 12).

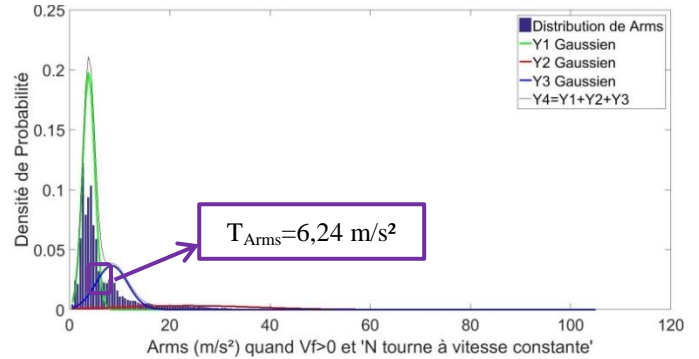


Figure 12 Densité de Probabilité de Arms modélisé par GMM (3 gaussiennes)

La puissance quand l'outil tourne à vide  $P_{vide}$  ne varie pas beaucoup. Il est plus rapide à l'obtenir par fouilles manuelles  $P_{vide} = 1,6 \text{ kW}$ . Donc, dans les données pour qui « la broche tourne à vitesse constante » et  $V_f > 0$  (cf. Tableau 1), on définit que si  $Arms > T_{Arms}$  ou  $P > P_{vide}$ , la classification est attribuée à « l'outil usine ». Dans le cluster « l'outil usine », il est possible de savoir si l'outil usine avec un engagement de l'outil constant ou variant (cf. Tableau 1). Pour vérifier la classification, la N (tr/min), la Vf (m/min), l'Arms (m/s²), ainsi que la puissance P (kW) classifié en 3 clusters vont être tracés sur une même figure. Plusieurs zones sont zoomées et vérifiées. Figure 13, un exemple est présenté.

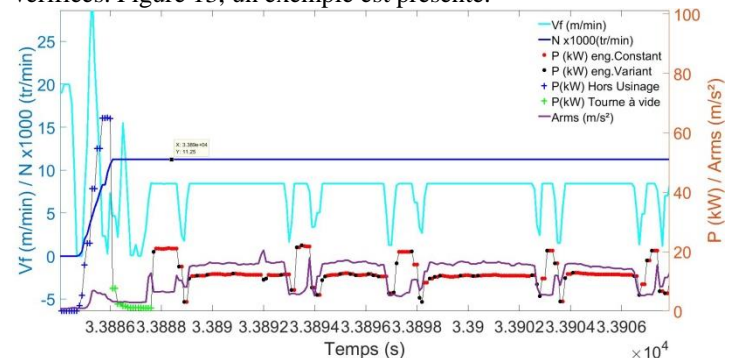


Figure 13 N, Vf, P et Arms en 3 clusters

Vf est en ligne cyan, N est en ligne bleu, Arms est en ligne violette. Les classifications « l'outil n'usine pas » sont représentées par des signes « + » de couleur bleue sur la courbe de puissance P, « l'outil usine avec engagement constant de l'outil » par des points rouges, « l'outil usine avec engagement variant de l'outil » par des points noirs. Lorsque la broche a atteint sa vitesse de consigne (11 250 tr/min dans Fig. 13), l'outil n'usine pas tout de suite. Il y a une sous-classe appelée « broche tournant à vide ». Cette classe est représentée en vert. Elle est définie ainsi : de l'atteinte de la vitesse consigne N jusqu'à l'instant où la puissance dépasse la puissance à vide  $P_{vide}$  (qui représente le début de l'usinage). Le but est de supprimer des vibrations parasites liées aux

démarrages de la broche, qui engendreraient sinon des erreurs de classification. La sous-classe «*broche tournant à vide* » est aussi affectée au cluster «*outil n'usine pas* ». En résumé, 3 clusters de la classification de l'engagement de l'outil sont définis. «*L'outil usine* » quand  $Arms > T_{Arms}$  ou  $P > P_{vide}$  en excluant la sous-classe «*broche tournant à vide* » ; il est constitué de deux clusters : «*l'outil usine avec un engagement constant de l'outil* » et «*l'outil usine avec un engagement variant de l'outil* » ; enfin, «*l'outil n'usine pas* » quand  $Arms \leq T_{Arms}$  et  $P \leq P_{vide}$  en incluant la sous-classe «*broche tournant à vide* ».

#### IV. CONCLUSIONS ET PERSPECTIVES

L'une des premières étapes de fouille de données est la sélection précise de données pertinentes. Pour ce faire, les données brutes doivent être classées dans différents groupes de contextes. Cet article présente tout d'abord les 17 classes potentielles en usinage. Puis, deux algorithmes d'apprentissage automatique non-supervisés (K-means et GMM) sont testés sur des données venant d'une entreprise d'usinage aéronautique. Les résultats de classification sont analysés par fouilles manuelles et conduisent aux conclusions suivantes :

1) K-means n'est pas pertinent pour les données brutes en raison du déséquilibre de ses classes, mais GMM fonctionne bien dans ce cas-là.

2) Un seuil de classification pour les 3 clusters «*broche arrêtée*», «*broche tournant à vitesse constante*» et «*broche tournant à vitesse variante*» a été déterminé par GMM et conduit à des classifications correctes.

3) Un seuil de classification pour les 3 clusters : «*Machine-outil arrêtée*», «*Machine-outil avançant à vitesse constante*» et «*Machine-outil avançant à vitesse variante*» a été déterminé par GMM et conduit également à des classifications correctes.

4) Un seuil de classification pour les 3 clusters : «*l'outil n'usine pas*», «*l'outil usine avec un engagement constant de l'outil* » et «*l'outil usine avec un engagement variant de l'outil* » a été déterminé par GMM et conduit également à des classifications correctes.

Ces valeurs sont spécifiques à la machine-outil étudiée mais la démarche de Machine Learning est répliquable.

Dans une prochaine étape, les données de simulation de trajectoires vont être confrontées aux classifications «*l'outil usine / n'usine pas* » obtenues par la méthode de GMM.

#### V. RÉFÉRENCES

- [1] Teti, R., Jemielniak, K., O'Donnell, G., & Dornfeld, D. (2010). Advanced monitoring of machining operations. *CIRP Annals-Manufacturing Technology*, 59(2), 717-739.
- [2] Quintana, G., & Ciurana, J. (2011). Chatter in machining processes: A review. *International Journal of Machine Tools and Manufacture*, 51(5), 363-376.
- [3] Zhou, Y., & Xue, W. (2018). Review of tool condition monitoring methods in milling processes. *The International Journal of Advanced Manufacturing Technology*, 1-15.
- [4] Godreau, V. (2017). Extraction des connaissances à partir des données de la surveillance de l'usinage, thèse de l'Université de Nantes.
- [5] Lenz, J., Wuest, T., & Westkämper, E. (2018). Holistic approach to machine tool data analytics. *Journal of Manufacturing Systems*.
- [6] Morgan, J., & O'Donnell, G. E. (2018). Cyber physical process monitoring systems. *Journal of Intelligent Manufacturing*, 29(6), 1317-1328.
- [7] Liu, C., Vengayil, H., Zhong, R. Y., & Xu, X. (2018). A systematic development method for cyber-physical machine tools. *Journal of Manufacturing Systems*, 48,13-24.
- [8] Kim, D.-H., Kim, T. J., Wang, X., Kim, M., Quan, Y.-J., Oh, J. Min S.-H., Kim, H., Bhandari, B., Yang, I. & Ahn, S.-H. (2018). Smart Machining Process Using Machine Learning: A Review and Perspective on Machining Industry. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 5(4), 555-568.
- [9] Denkena, B., Dittrich, M. A., & Uhlich, F. (2016). Self-optimizing cutting process using learning process models. *Procedia Technology*, 26, 221-226.
- [10] Krishnakumar, P., Rameshkumar, K., & Ramachandran, K. I. (2015). Tool Wear Condition Prediction Using Vibration Signals in High Speed Machining (HSM) of Titanium (Ti-6Al-4 V) Alloy. *Procedia Computer Science*, 50, 270-275.
- [11] Gouarir, A., Martínez-Arellano, G., Terrazas, G., Benardos, P., & Ratchev, S. (2018). In-process Tool Wear Prediction System Based on Machine Learning Techniques and Force Analysis. *Procedia CIRP*, 77, 501-504.
- [12] Jacques Julien. Fouille de données / Data mining Université Lumière Lyon 2, 2001.
- [13] Bhinge, R., Biswas, N., Dornfeld, D., Park, J., Law, K. H., Helu, M., & Rachuri, S. (2014). An intelligent machine monitoring system for energy prediction using a Gaussian Process regression. 2014 IEEE International Conference on Big Data, 978-986.
- [14] Godreau, V., Mathieu, R., Etienne, C. et al. (2018). Continuous improvement of HSM process by data mining. *Journal of Intelligent Manufacturing*, p. 1-8.
- [15] De Castelbajac, C., Ritou, M., Laporte, S., & Furet, B. (2014). Monitoring of distributed defects on HSM spindle bearings. *Applied Acoustics*, 77, 159-168.
- [16] Tufféry, Stéphane. Data mining et statistique décisionnelle: l'intelligence des données. Editions Technip, 2012.
- [17] Reynolds, D. (2015). Gaussian mixture models. *Encyclopedia of biometrics*, 827-832.