TABLE DES MATIÈRES

Page

LISTE DES TABLEAUX

Page

LISTE DES FIGURES

Page

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

AC	Accuracy
AG	Algorithme Génétique
ANN	Artificial Neural Network
BP-NN	Back Propagation-Neural Network
CPP	Compound Poisson Process
DMSMS	Diminishing Manufacturing Sources and Materials Shortages
FA	Forêt Aléatoire
GBM	Gradient Boosting Machine
GLM	Generalized Linear Model
GA-RF	Genetic Algorithm-Random Forest
GA	Genetic Algorithm
HMM	Hidden Markov Model
LTB	Last Time Buy
MVS	Machine à Vecteurs de Support
ML	Machine Learning
NN	Neural Network
OOB	Out Of Bag
PAM	Partitioning Around Medoid
PSO	Partical Swarm Optimization
PLS	Partical least Square
PSO-RF	Partical Swarm Optimization-Random Forest
RF	Random Forest

ROC	Receiver Operator characteristic
RNA	Réseaux de Neurones Artificiels
SVM	Support Vector Machine
SE	Sensitivity
SP	Specificity
STD	Standard Deviation
YEOL	Years to End of Life

LISTE DES SYMBOLES ET UNITÉS DE MESURE

$argmin$	Approche discriminative
$fitness\ (x)$	Fonction du coût
P	Probabilité
k	Nombre de classes (clusters)
μ_k	Centroïde de chaque groupe
d	Distance
λ_i	Intensité des ordres d'arrivé
t	Durée de mesure
x_i	Phase de cycle de vie

INTRODUCTION

0.1 Contexte et problématique de la recherche

Aujourd'hui, des milliers de composants électroniques deviennent obsolètes, ce qui est engendré principalement par l'avancement technologique (O'Dowd, 2010). L'obsolescence est le statut donné à une pièce lorsqu'elle n'est plus disponible ou fabriqué par le fournisseur, ce qui cause des retards et des coûts énormes. Ce genre d'obsolescence fait référence au type DMSMS (F. J. Romero Rojo, Roy, Shehab, & Cheruvu, 2012; F. R. Rojo, Roy, & Shehab, 2010). La figure 0.1 résume ce problème.

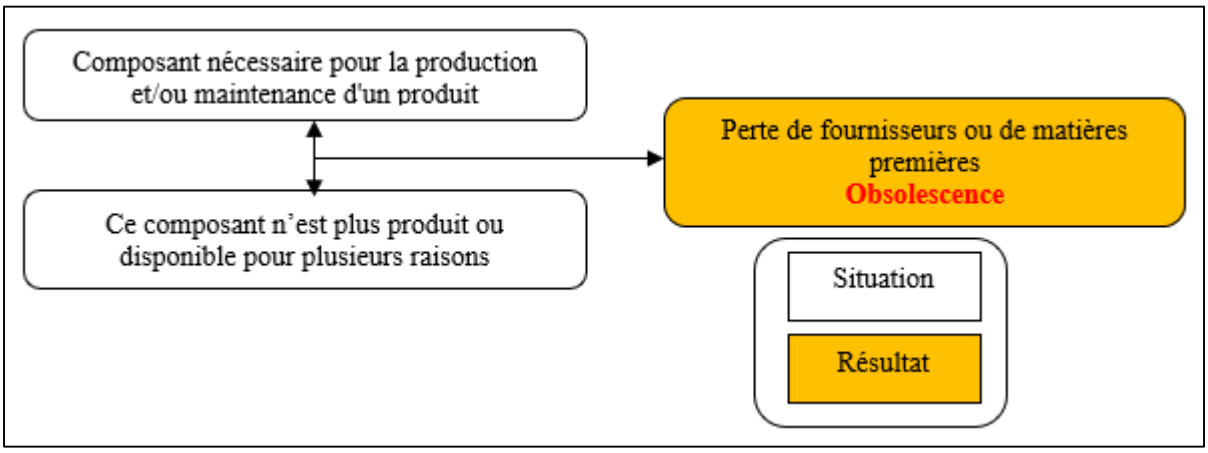

Figure 0.1 Apparition de l'obsolescence adaptée de Bartels et al (2012)

L'arrêt de production peut avoir de nombreuses causes, la croissance rapide de la technologie électronique est la cause la plus fréquente (Mellal, 2020). L'obsolescence des composants électroniques est depuis longtemps perçue comme un défi majeur dans la conception des systèmes à longue durée de vie. Les composants électroniques présentent souvent des problèmes majeurs de durabilité des systèmes en raison de leur court cycle de vie. Les problèmes d'obsolescence apparaissent généralement dans les systèmes dont le cycle de vie est plus long que celui de leurs composants, tels que l'automobile, l'avionique, etc. (Xiaozhou, Thornberg, & Olsson, 2014). Ce conflit rend le maintien de ces équipements très coûteux.

La croissance de l'industrie électronique a entraîné de profonds changements dans les composants électroniques. De nouvelles technologies sont introduites sur le marché à un rythme croissant. Gordon Moore, a stipulé que la densité des semiconducteurs double environ tous les 18 mois (Voller & Porté-Agel, 2002). En effet, cette évolution rapide de la technologie présente l'un des principaux facteurs d'augmentation du taux d'obsolescence (Bartels, Ermel, Sandborn, & Pecht, 2012). Aujourd'hui, le court cycle de vie et le manque de prévision représentent un défi pour plusieurs entreprises qui doivent prendre en considération le risque d'obsolescence, en particulier dans le secteur électronique, qui représente l'un des secteurs les plus dynamiques de l'économie mondiale (P. Sandborn, Prabhakar, & Ahmad, 2011; Solomon, Sandborn, & Pecht, 2000).

La nature de la prévision de l'obsolescence est réactive, définie par la résolution du problème une fois qu'il se produit. Actuellement, la plupart des entreprises ne disposent pas de méthodes permettant de prédire efficacement l'obsolescence, et sont donc obligées de compter sur des stratégies réactives. Malheureusement, les stratégies réactives sont souvent plus coûteuses que les stratégies proactives. Elles nécessitent des ressources supplémentaires (temps et matériel) pour être résolues et peuvent contribuer à retarder davantage la satisfaction du client. En revanche, les stratégies proactives permettent aux entreprises de disposer de plus de temps pour planifier et réagir avec une approche efficace et fiable (P. Sandborn, 2013). De ce fait, la prévision de l'obsolescence apparaît comme l'une des approches les plus efficaces pour mieux gérer l'obsolescence. Grâce aux méthodes de prévisions, les entreprises peuvent assurer le support des pièces en service et atténuer tout impact négatif en identifiant les pièces susceptibles de devenir obsolètes. De plus, la prévision permet aux concepteurs et ingénieurs de gérer plus efficacement la conception des produits à longue durée de vie en fonction du cycle de vie prévu des pièces.

Afin de faire face à ces défis, plusieurs études ont été menées pour créer des modèles capables de prévoir efficacement l'obsolescence. Des méthodes statistiques telles que la régression, les séries chronologiques et la méthode gaussienne ont déjà été utilisées dans plusieurs travaux (Gao, Liu, & Wang, 2011; Jungmok & Namhun, 2017; P. Sandborn, 2017). Également, des

méthodes de prévision du risque ont été proposées (C. L. Josias, 2009; F. R. Rojo, Roy, & Kelly, 2012; van Jaarsveld & Dekker, 2011a). Dans la littérature, la majorité des travaux sont basés sur la prévision d'obsolescence d'un seul type de composant. Toutefois, il existe des milliers de composants qui affectent la vie quotidienne et, par conséquent, il est nécessaire de modifier ces méthodes afin d'améliorer la prévision sur une grande échelle de composants.
À cet égard, l'apprentissage machine apparaît comme une approche prometteuse de prévision (D. Wu, Jennings, Terpenny, & Kumara, 2016; Zurada, 1992). C'est une approche d'analyse de données qui automatise la reconnaissance des modèles de données sans la manipulation humaine et capable de prévoir un large échantillon de composants électroniques avec un haut degré de précision. Au cours des dernières décennies, l'apprentissage automatique a attiré l'attention des chercheurs dans plusieurs disciplines et a été appliqué dans de nombreux domaines (X. Wu et al., 2008).

0.2 Objectifs de la thèse

Cette thèse a pour objectif de développer et d'optimiser des modèles de prévision du risque et de cycle de vie de l'obsolescence des composants électroniques. Pour atteindre ce but, plusieurs objectifs spécifiques ont été assignés à l'étude. Ces objectifs ont été identifiés après une évaluation des lacunes dans certaines approches existantes, pour ensuite proposer de nouvelles méthodes de prévision.
La question principale de cette recherche est la suivante :
« Quels sont les modèles de prévision les mieux adaptés pour prévoir le risque et cycle de vie d'obsolescence sur une grande échelle de composants électroniques, et comment évaluer la performance de ces modèles? »
Pour répondre à la question de recherche, trois objectifs spécifiques sont alors visés :

1) Développer des modèles basés sur l'exploration de données (apprentissage supervisé et non supervisé) pour prévoir le risque d'obsolescence sur un grand échantillon de composants électroniques.
2) Introduire des algorithmes d'optimisation métaheuristique pour améliorer la performance des algorithmes d'apprentissage machine, étant donné que la précision de

la classification peut être réduite en raison des caractéristiques non pertinentes et redondantes de l'ensemble de données.

3) Développer un modèle pour prévoir le cycle de vie d'obsolescence en introduisant le modèle de Markov et le processus de poisson composé.

La démarche de la recherche proposée est guidée par une approche de la science de conception décrite dans le chapitre 2.

0.3 Motivation et contribution

Le marché d'aujourd'hui est devenu très concurrentiel. Avec la quatrième révolution de l'industrie et l'énorme vague de génération de données des entreprises ont incité une adaptation étroite aux entreprises et aux ingénieurs de conception face aux changements des préférences et des exigences des clients. L'apprentissage automatique a beaucoup capté l'attention dans plusieurs domaines, permettant aux industriels d'extraire des connaissances à partir de données à grande échelle pour mieux gérer l'obsolescence. Les méthodes présentées dans cette thèse peuvent apporter un nouvel éclairage sur les problèmes d'obsolescence face à la situation actuelle d'un marché hautement concurrentiel.

Cette thèse est motivée par un besoin industriel. En effet, ces travaux de recherche s'inscrivent dans un cadre d'un projet industriel en collaboration avec une entreprise aéronautique.
Partant de ce besoin industriel par rapport aux problèmes d'obsolescence et pour répondre aux nouvelles exigences des clients et à la mutation du marché face à l'innovation accrue, l'entreprise partenaire cherche les moyens permettant de minimiser les risques et de mieux prévoir l'émergence de l'obsolescence afin de réduire son impact potentiel sur la chaine de production.

L'entreprise partenaire est confrontée fréquemment à des problèmes d'obsolescence. Comme elle ne possède aucun outil pour gérer efficacement l'obsolescence, elle est obligée de recourir à des méthodes réactives. Ce problème est devenu un enjeu majeur dans leur gestion de cycle de vie des systèmes complexes (hélicoptères), et a provoqué des ennuis néfastes comme le remplacement prématuré et imprévu des sous-systèmes et les investissements énormes

auxquels ils devront faire face. À la vue de ce manque, l'entreprise a cherché à implanter un outil de prévision qui va l'aider à réduire les coûts supplémentaires engendrés par l'obsolescence et à devenir proactive pour accroître sa part de marché. Les travaux de recherche et le problème industriel ont permis de cibler le contexte de la recherche, ce qui a amené à définir la problématique et à bien cerner la méthodologie.

La principale contribution de cette thèse est d'introduire de nouvelles approches permettant de mieux prévoir l'obsolescence des composants électroniques. La recherche repose sur trois contributions spécifiques décrites plus amplement dans le chapitre 2.
La première contribution est liée à l'introduction de l'apprentissage machine supervisé et les algorithmes d'optimisation pour prévoir le risque d'obsolescence. Tandis que la deuxième contribution se concentre sur l'apprentissage machine non supervisé lié au risque d'obsolescence. La troisième contribution, quant à elle, introduit une méthode améliorée basée sur les limites d'une approche existante pour prévoir le cycle de vie d'obsolescence.

0.4 Organisation de la thèse

La thèse est organisée en six chapitres. Le chapitre 1 présente une synthèse de l'état de l'art de diverses notions théoriques associées au contexte de la recherche. Le chapitre 2 introduit la démarche méthodologique. Ensuite, trois articles de revue sont présentés dans les chapitres 3, 4 et 5 qui répondent aux objectifs de la thèse cités précédemment. Le chapitre 6 est consacré à une discussion générale et résume les contributions de cette thèse, les limites et les travaux futurs. Enfin, la thèse est achevée par une conclusion.

CHAPITRE 1

REVUE DE LITTÉRATURE

Ce chapitre vise à élaborer un examen critique de la littérature et offre une formulation détaillée de diverses notions théoriques ainsi que les problématiques associées au contexte de la recherche.

Le chapitre inclut un résumé qui décrit les limites et les faiblesses des différentes approches utilisées dans la littérature. L'examen est divisé en trois sections associées à chaque contexte de recherche, comme illustrées dans la Figure 1.1.

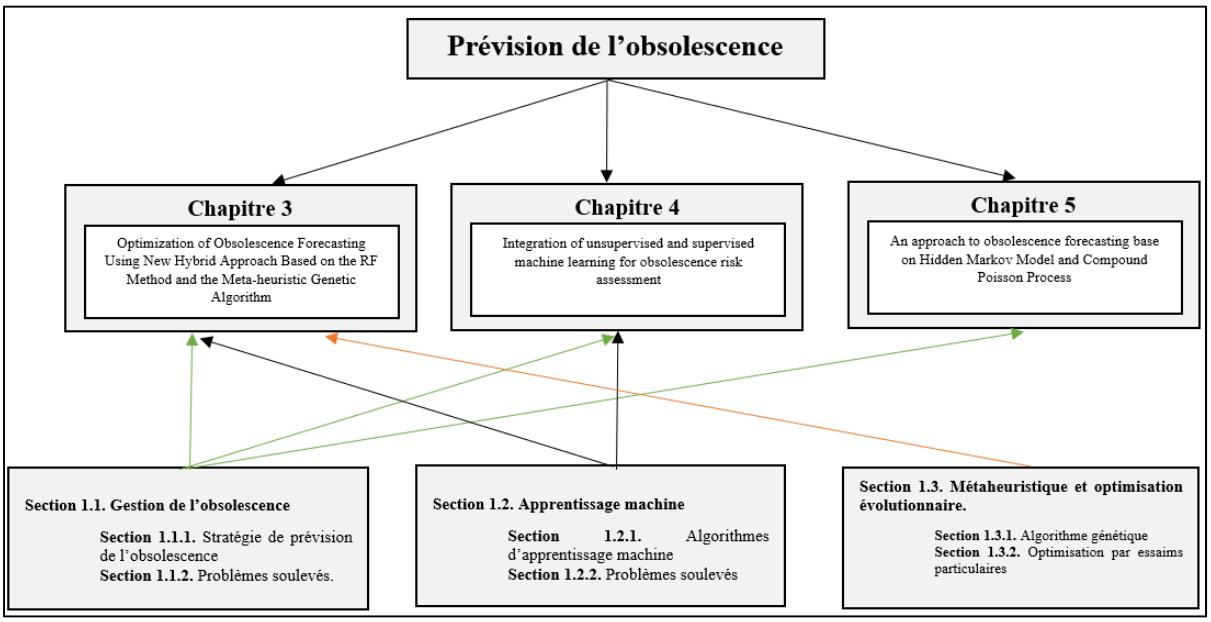

Figure 1.1 Structure et portée des sujets abordés dans la littérature

1.1 Gestion de l'obsolescence

L'objectif de la gestion de l'obsolescence est de minimiser les coûts supplémentaires et l'impact négatif de l'obsolescence. La norme IEC 62402 de la gestion d'obsolescence fournit des orientations et des lignes directrices aux entreprises pour les aider à planifier un processus

efficace pour gérer l'obsolescence et d'assurer le support tout au long du cycle de vie des pièces (Adetunji, Bischoff, & Willy, 2018; Zaabar, Beauregard, & Paquet, 2018).

Trois niveaux de gestion d'obsolescence ont été abordés dans la littérature : la gestion réactive, proactive et stratégique (Bartels, Ermel, Sandborn, et al., 2012; P. Sandborn, 2013). En pratique, la plupart des entreprises ne disposent pas de méthodes proactives pour gérer l'obsolescence et sont donc obligées de compter sur des stratégies réactives à court terme. Les approches réactives les plus connues et applicables par les entreprises à court terme sont : « le stock existant » : c'est un stock détenu par la chaîne d'approvisionnement qui contient les composants de rechange originaux. Le stock existant doit être suffisant pour assurer l'approvisionnement tout au long du cycle de vie restant du système (Bartels, Ermel, Pecht, & Sandborn, 2012a; Livingston, 2000; Pince & Dekker, 2011; F. R. Rojo et al., 2010; Z. Shi & Liu, 2020).
« Le dernier achat » (en anglais : last time buy) consiste à acheter une quantité suffisante pour supporter la production jusqu'à la reconception (C. Jennings & Terpenny, 2015). La troisième approche réactive est la « substitution d'un composant » (en anglais alternate part), consiste à remplacer un composant ou un équipement par un autre avec la même performance, les mêmes dimensions et les mêmes caractéristiques mécaniques. Néanmoins, ces approches ne sont que temporaires et peuvent engendrer des coûts importants non prévus si l'entreprise manque de moyens pour se procurer les pièces requises.

La gestion proactive, soit la méthodologie à long terme, repose sur la surveillance proactive des informations relatives au cycle de vie des pièces afin d'éviter les arrêts de production. La prévision de l'obsolescence, qui est le centre d'attention de cette recherche, s'avère essentielle au niveau de la gestion proactive.

L'approche stratégique, quant à elle, constitue une combinaison de la gestion réactive et proactive pour minimiser le coût du cycle de vie. La prévision d'obsolescence peut être utilisée comme une entrée dans un niveau de gestion stratégique.

1.1.1 Stratégie de prévision de l’obsolescence

La prévision est une estimation probabiliste concernant un événement. Les prévisions fournissent aux décideurs des informations qui peuvent être utilisées pour gérer, de manière proactive, l'écart entre ce à quoi pourrait ressembler l'avenir et ce à quoi nous voulons qu'il ressemble (Parvin & Beruvides, 2017).

La prévision de l'obsolescence est une approche de gestion proactive qui permet aux fabricants des systèmes d’identifier l'obsolescence des composants en avance. La prévision d'obsolescence est basée sur des méthodes qualitatives et quantitatives, c’est-à-dire soit les données sont disponibles et c'est possible de les collecter, soit aucune donnée n'est disponible et, par conséquent, la méthode qualitative devient la meilleure approche de prévision. La méthode Delphi est l’une des approches qualitatives qui ont été appliquées pour estimer le risque d’obsolescence (F. R. Rojo et al., 2012). Cette méthode est basée principalement sur des questionnaires et des enquêtes faites auprès des experts pour identifier les facteurs qui ont un impact sur l’obsolescence.
Les méthodes quantitatives de prévision d’obsolescence auxquelles la thèse est consacrée sont divisées en deux types, la prévision du risque et la prévision de cycle de vie. Par ailleurs, il existe deux démarches de prévision : la prévision à long terme (un an ou plus) qui permet une gestion proactive basée sur la planification du cycle de vie pour soutenir un système, et la prévision à court terme qui peut être observée à partir de la chaîne d'approvisionnement. Les prévisions à court terme peuvent impliquer la réduction du nombre de sources, la réduction des stocks et l’augmentation des prix.

1.1.1.1 Prévision du risque d’obsolescence

Notions et concepts du risque

La gestion du risque permet à une organisation d'évaluer tous les risques liés à la chaîne logistique et de contrôler les mesures d'atténuation de manière structurée. Pour cela, les entreprises doivent adopter une approche proactive pour identifier le risque en avance. Dans le

pratique, le risque industriel est généralement associé aux notions tels que, un obstacle potentiel, une vulnérabilité d'un produit, une probabilité d'occurrence d'un évènement, etc. De ce fait, ISO31000 a été élaboré dans le but d'atteindre et de gérer efficacement le risque. Cette norme établit un ensemble d'exigences et de lignes directrices aux entreprises pour assurer l'amélioration continue du service. Selon cette norme, le risque est défini comme la probabilité d'occurrence d'un évènement et sa conséquence.

Prévision du risque d'obsolescence

Les méthodes de prévision du risque d'obsolescence sont utilisées pour estimer la probabilité qu'une pièce devienne obsolète et son impact sur le système, en se basant sur une matrice de criticité. Ce concept est équivalent à celui défini par la norme ISO31000. Dans le cas de prévision d'obsolescence, une analyse du risque doit être effectuée dans le but d'obtenir un aperçu sur la disponibilité future des pièces critiques, généralement basé sur des facteurs et des indices d'obsolescence.

Deux principales approches qui utilisent des facteurs pour identifier le niveau du risque (faible, moyen ou élevé) ont été adoptées et largement discutées dans la littérature (F. R. Rojo et al., 2012). Dans ce contexte, Rojo a développé une méthodologie pour prévoir le risque d'obsolescence en se basant sur des indicateurs d'obsolescence tels que l'année de fin de vie, le nombre de sources disponibles, le taux de consommation de la pièce et la disponibilité du stock. En effet, ces facteurs ont été identifiés à partir d'une enquête auprès d'une équipe d'experts industriels. Cette approche a permis d'élaborer un plan d'évaluation du risque pour chaque composant identifié.

L'approche développée par Josias (2009), quant à elle, vise à créer un indice de risque pour mesurer le risque d'obsolescence. Parmi les paramètres identifiés grâce à cette technique, on retrouve : le nombre de sources disponibles pour chaque pièce et le niveau du risque de l'entreprise. Ces paramètres sont mesurés sur une échelle de risque de zéro à trois (C. L. Josias, 2009). Pour calculer l'indice de risque, un poids est associé à chaque paramètre, ce poids est modulable selon chaque entreprise. L'indice de risque est calculé comme suit :

$$Indice\ de\ risque = w_1 \left(\left(\frac{\alpha-1}{2}\right)\right) + w_2 \left(\left(\frac{\beta-1}{2}\right)\right) + w_3 \left(\left(\frac{\gamma-1}{2}\right)\right) + w_4 \left(\left(\frac{\delta-1}{2}\right)\right) \quad (1.1)$$

Où :

w_i : Moyenne pondérée des facteurs $\alpha, \beta, \gamma\ et\ \delta$ pour i= 1,2,3,4
α: Part de marché de l'entreprise (1= faible, 2=moyen et 3=élevé)
β : Nombre de fournisseurs disponibles (1= faible, 2=moyen et 3=élevé)
γ: Phase de cycle de vie (1= faible, 2=moyen et 3=élevé)
δ: Niveau du risque de la compagnie (1= faible, 2=moyen et 3=élevé)

Par ailleurs, Jaarsveld a introduit une autre méthode pour estimer le risque d'obsolescence (van Jaarsveld & Dekker, 2011a). Dans son modèle basé sur la chaîne de Markov, Jaarsveld (2011) utilise deux états de transition x_1 et x_0. Selon ces deux états, x_1 présente l'état où la demande est élevée, et x_0 présente l'état absorbant où la demande est morte. La transition d'un état à un autre se base sur la variation de la demande qui suit un processus de poisson composé. Cependant, cette méthode ne peut pas prévoir très loin dans le futur, car elle n'est pas capable de prévoir l'ensemble du cycle de vie puisqu'elle ne possède qu'une seule transition entre deux états; de telle façon que l'estimation de l'obsolescence n'est pas précise. Les limites et les améliorations portées à cette méthode sont discutées au chapitre 2.

Enfin, Zolghadri (2018) a introduit la théorie de Bayes pour prédire la probabilité d'obsolescence en se basant sur deux états : obsolète et non obsolète, afin d'évaluer sa propagation et son impact sur le système (Zolghadri, Addouche, Boissie, & Richard, 2018).

1.1.1.2 Prévision du cycle de vie d'obsolescence

La prévision du cycle de vie fait référence à un processus pour prévoir la durée pendant laquelle le produit sera en production et estimer la date à laquelle il deviendra obsolète. Ces dates permettent aux industriels de choisir la meilleure approche proactive pour gérer l'obsolescence avant qu'elle ne se produise. Il existe dans la littérature de nombreux modèles de prévision qui peuvent être utilisés. Les méthodes courantes incluent : la régression, les méthodes par séries chronologiques, les moyennes mobiles et le lissage exponentiel. Toutefois, la méthode choisie dépend généralement des caractéristiques de la pièce et des données disponibles.

Les méthodes de prévision du cycle de vie sont développées en se basant sur la période pendant laquelle le produit est disponible sur le marché et que le client peut se le procurer. En effet, la majorité des produits suivent cinq étapes durant leur cycle de vie : introduction, croissance, maturité, déclin, et obsolescence (C.-M. Huang, Romero, Osterman, Das, & Pecht, 2019; Solomon et al., 2000). La phase d'introduction se réfère à la technologie introduite dans le marché pour la vente, et tout ce qui survient avant la phase d'introduction représente la période de développement du produit. Une fois que les produits sont introduits et réussissent à être vendus, après une certaine période les concurrents à faibles coûts entrent dans le marché et ainsi le volume de la demande commence à diminuer. De ce fait, de nombreuses entreprises cessent la production, ce qui rend le produit sans support et obsolète. La Figure 1.2 présente la transition de vie d'un état à un autre, où l'axe des abscisses représente les différentes phases de cycle de vie, et l'axe des ordonnées représente la variation de la demande à travers chaque phase (Solomon et al., 2000).

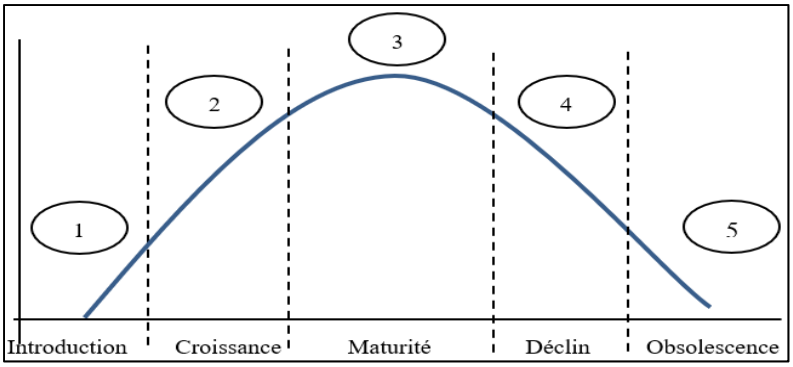

Figure 1.2 Modèle d'une courbe de cycle de vie d'un produit

Solomon (2000) a été le premier à introduire la méthode de prévision de la date d'obsolescence à partir de la courbe de cycle de vie. Sa méthodologie est basée sur les indicateurs d'entrée suivants : la part du marché, le risque du marché et le groupe technologique du composant. Chaque phase de cycle de vie est calculée à partir de la moyenne (μ) et de l'écart type (σ). La détermination de la zone d'obsolescence été définie comme : ($\mu + 2.5\sigma - p$, $\mu + 3.5\sigma - p$).

Où p est la date actuelle, μ est la moyenne de la courbe de cycle de vie calculée à partir des données de ventes, et σ représente la différence entre la moyenne et la date actuelle.

Depuis les dernières décennies, l'exploration des données (en anglais : Data mining) a beaucoup été utilisée. Une première version de l'exploration des données a été introduite par Josias et al. (2004) (C. Josias, Terpenny, & McLean, 2004). Ils ont développé un modèle de régression qui intègre l'interaction entre les attributs d'un processeur afin de déterminer les caractéristiques technologiques qui influencent l'obsolescence du processeur. En revanche, le modèle n'a pas été utilisé pour prévoir la date d'obsolescence.
Par ailleurs, Gao et al. (2011) ont suivi la même approche que Josias et ses collaborateurs. Ils ont développé un modèle de régression pour prévoir la tendance d'obsolescence d'un composant électronique, nommée FPGA. Le modèle réagit sur la corrélation entre la date d'obsolescence, la date d'introduction et les attributs liés au composant. Un coefficient est assigné à chaque indice basé sur la signification de corrélation (Gao et al., 2011).

Dans la littérature, il existe deux approches d'exploration de données. La première approche est utilisée pour les composants qui ont un pilote paramétrique évolutif clair (en anglais : *Evolutionary parametric driver*), c'est-à-dire les composants technologiques dont les caractéristiques évoluent au fil du temps (P. Sandborn, 2017; P. A. Sandborn, Mauro, & Knox, 2007; Solomon et al., 2000). La prévision du cycle de vie d'obsolescence est obtenue à partir des données de ventes en utilisant la distribution gaussienne (Zheng, Nelson, Terpenny, & Sandborn, 2012). Dans le même contexte, Jungmok et Namhun (2017) ont développé une approche évolutive pour prévoir le cycle de vie d'obsolescence basé sur la modélisation par les séries chronologiques (en anglais : time series modelling) (Jungmok & Namhun, 2017). Cette méthode a été comparée avec l'approche de Sandborn (2007) et il a été prouvé qu'elle peut mieux prévoir le cycle de vie, étant donné que la méthode de Sandborn exige que le pilote paramétrique évolutif soit connu pour effectuer la prévision, alors que cette méthode ne l'exige pas.

D’un autre côté, plusieurs composants électroniques n’ont pas de pilote paramétrique évolutif clair, d’où le fait que leur évolution dans le temps n’est pas connue. Dans ce contexte, une approche introduite par Sandborn, Prabhakar et al. (2011) consiste à développer un algorithme pour prévoir la date d’obsolescence basée sur la durée d’approvisionnement du composant, donnée par la relation suivante : $L_p = D_0 - D_i$

Où, D_0 est la durée pendant laquelle la pièce était disponible à l'approvisionnement auprès de son fabricant d'origine, alors que D_i est la date d’introduction de la pièce. Évidemment, si la durée d’approvisionnement peut être prévue et la date d’introduction connue, alors la date d’obsolescence peut être prévue à partir de cette équation. Cependant, cette approche nécessite non seulement un vaste échantillon de données d'entrées, mais peut également générer de graves erreurs en utilisant une seule variable explicative (date de l’introduction) (P. Sandborn et al., 2011).

1.1.2 Problèmes soulevés et pistes de recherche

La première partie de la revue de littérature a permis d’évaluer les approches existantes pour la prévision de risque et de cycle de vie d’obsolescence. En effet, les méthodes classiques telles que la régression linéaire, la distribution gaussienne et les séries chronologiques ont couramment été utilisées pour prévoir le cycle de vie d’obsolescence. Par ailleurs, des indicateurs clés ont été identifiés et utilisés pour prévoir le risque d’obsolescence.

Néanmoins, la majorité de ces méthodes sont basées sur la prévision d’obsolescence d'un seul type de composant. De plus, elles sont principalement basées sur des connaissances humaines pour estimer l'obsolescence, ce qui cause des erreurs de prévision. Conséquemment, il est nécessaire de modifier ces méthodes par une méthode plus efficace qui doit tenir compte des défauts et limites des approches existantes afin d’améliorer la prévision de milliers de fin de vie des composants qui affectent la vie quotidienne.

De ce fait, une piste de recherche qui paraît particulièrement pertinente consiste à introduire l’apprentissage machine appliqué à l’obsolescence, capable de prévoir un large échantillon de composants électroniques avec un haut degré de précision sans la manipulation humaine.

Contrairement aux approches existantes qui nécessitent des données de vente et l'intervention humaine, l'apprentissage machine ne l'exige pas. L'introduction d'une nouvelle approche de prévision de l'obsolescence est nécessaire en raison du manque d'évolutivité et de précision dans la majorité des méthodes actuelles.

L'autre piste de recherche s'intéresse particulièrement aux chaînes de Markov. En effet, dans la littérature, une seule étude a évoqué l'intégration des chaînes de Markov pour prévoir l'obsolescence (van Jaarsveld & Dekker, 2011a). L'une des limites majeures de cette étude réside dans le fait qu'elle n'est pas capable de prévoir l'ensemble du cycle de vie. Compte tenu de cette limite, une autre piste de recherche prometteuse conduira à la recherche d'une nouvelle approche pour améliorer la méthode actuelle.

1.2 Apprentissage machine

L'apprentissage machine est un champ d'études de l'intelligence artificielle basé sur la modélisation statistique. C'est une méthode d'analyse des données à grande dimension qui automatise la reconnaissance des modèles de données sans la manipulation humaine. L'objectif de l'apprentissage automatique est de construire des systèmes informatiques qui s'améliorent automatiquement avec des informations supplémentaires.

Au cours des dernières décennies, l'apprentissage machine a attiré l'attention de nombreux chercheurs dans diverses disciplines et a été appliqué dans de nombreux domaines. Sa performance a été justifiée dans la littérature et a gagné en popularité dans de nombreux domaines d'application, puisqu'il peut traiter de grands ensembles de données avec de nombreuses variables.

Dix algorithmes ont été identifiés dans IEEE (X. Wu et al., 2008) tels que adaboost*, SVM, k-means, decision tree, naïve bayes, random forest, neural network, etc.* comme les meilleurs prédicteurs en data mining. Ces algorithmes offrent une approche prometteuse pour la prévision. D'ailleurs, ils ont montré une bonne précision de prévision dans plusieurs travaux

(Hassan, Khosravi, & Jaafar, 2013; Hippert & Taylor, 2010; Homchalee & Sessomboon, 2013; Jilani, Amjad, Jaafar, & Hassan, 2012; Yun, Ping, & Li, 2010).

Il existe deux types d'apprentissage, soit supervisé et non supervisé. En l'absence d'une variable à expliquer (étiquette), l'apprentissage est dit non supervisé. Lors de l'apprentissage supervisé, la réponse désirée du modèle est spécifiée par l'utilisateur. L'apprentissage supervisé construit des modèles de prévision par rapport aux données qui ont des étiquettes connues et prédit des étiquettes pour de nouvelles données inconnues. Dans le cas d'obsolescence, la variable de décision est définie comme catégorielle (obsolète/non obsolète) ou numérique représentée par une date d'obsolescence.

Quel que soit le domaine d'application, les étapes d'apprentissage supervisé sont toujours les mêmes. La première étape consiste à extraire les données, généralement à partir des bases de données disponibles. La deuxième étape est l'exploration des données, qui consiste à bien choisir les caractéristiques et les spécifications des données dont le modèle aura besoin. Une fois que l'échantillon est prêt, il sera divisé en deux groupes : 2/3 pour l'entraînement et 1/3 pour le test et la validation. L'ensemble des apprentissages sera introduit dans les algorithmes pour créer le modèle de prévision. Par la suite, les paramètres du modèle comme le taux et la durée d'apprentissage, le nombre de variables, etc. seront estimés. Les paramètres sont optimisés en fonction de la méthode d'estimation de l'erreur obtenue. L'exploitation du modèle choisi peut être réalisée sur de nouvelles données.

Les étapes de l'apprentissage non supervisé sont différentes de celles de l'apprentissage supervisé. En effet, l'apprentissage non supervisé ou le clustering, consiste à séparer les données en plusieurs classes k. Chaque classe contient un ensemble de données similaires.

L'avancement technologique de l'intelligence artificielle au niveau des systèmes fait de l'apprentissage machine une méthode parfaite pour la prévision de l'obsolescence, puisque des milliers de composants deviennent obsolètes chaque jour, ce qui permet de récupérer plus

d'informations sur ces composants à partir desquelles les systèmes s'améliorent pour apprendre et fournir des résultats plus précis.

Les principaux défis de l'apprentissage machine consistent à identifier les algorithmes les mieux adaptés pour la prévision de l'obsolescence et aussi de savoir quelles sont les caractéristiques qui forment le modèle. Récemment, l'apprentissage machine a reçu une attention particulière pour la prévision de l'obsolescence. Plusieurs auteurs ont utilisé une méthode basée sur l'exploration des données en créant des algorithmes d'apprentissage supervisé pour prévoir le risque et la date d'obsolescence (Yosra Grichi, Yvan Beauregard, & Thien-My Dao, 2018; Y Grichi, Beauregard, & Dao, 2017; C. Jennings, Wu, & Terpenny, 2016a). Dans le cas de prévision de la date d'obsolescence, la régression est utilisée pour prévoir une valeur numérique. Pour prévoir le risque, la classification est utilisée pour prédire le statut, soit actif ou obsolète. La Figure 1.3 présente le schéma fonctionnel d'un algorithme basé sur l'apprentissage supervisé.

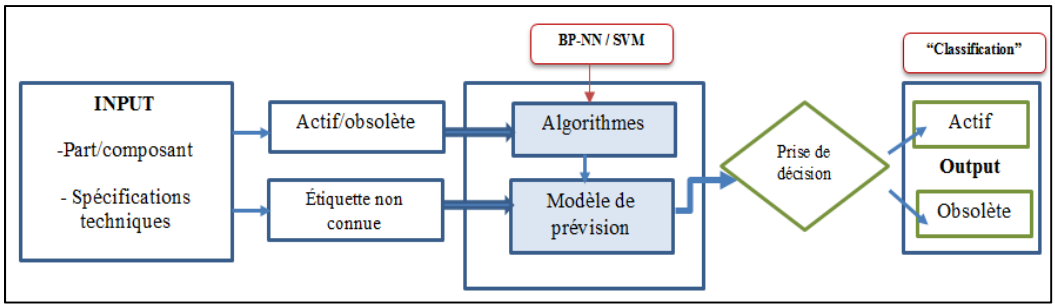

Figure 1.3 Classification de l'obsolescence adaptée de C. Jennings et al. (2016a)

Les résultats des deux articles mentionnés précédemment ont montré que la forêt aléatoire est le meilleur algorithme de prévision de l'obsolescence, comparativement aux autres algorithmes.

1.2.1 Description qualitative des principaux algorithmes d'apprentissage machine

Les phrases suivantes présentent une description de différents algorithmes d'apprentissage machine. Ces algorithmes sont classés comme les meilleurs algorithmes en data mining selon X. Wu et al. (2008), leur performance a été justifiée dans la littérature (X. Wu et al., 2008).

Les Réseaux de neurones artificiels: constituent un type d'apprentissage qui offre une approche alternative prometteuse caractérisée par plusieurs avantages, comme la modélisation d'un grand ensemble de données. Le modèle des RNA peut être généralisé. En effet, après avoir entrainer les données qui leur sont présentées, les RNA peuvent souvent déduire la partie invisible de la population même si les données de l'échantillon contiennent des données bruyantes. Comme toute autre approche, les RNA présentent des inconvénients (Guoqiang, Patuwo, & Hu, 1998). Le surapprentissage (en anglais : overfitting) est l'une des principales limites dans l'apprentissage et qui sont causées par l'entraînement trop long ou avec beaucoup de nœuds cachés.

Les machines à vecteurs de support (SVM), s'appellent aussi séparateurs à vaste marge, sont devenues largement utilisées dans de nombreuses applications. Elles exigent peu de données pour l'entraînement et offrent une bonne performance de classification. Cependant, il existe quelques limites liées à cette approche. En fait, même si les SVM représentent l'un des meilleurs classificateurs, un meilleur hyperplan peut admettre quelques erreurs sur les données d'entraînement.

L'arbre de décision (en anglais : *decision tree*) est utilisé comme un modèle prédictif. Il est utilisé notamment en *data mining*. Dans ces structures arborescentes, les feuilles représentent les valeurs de la variable cible et les lignes de branchement correspondent à des combinaisons de variables d'entrée qui conduisent à ces valeurs. Dans l'analyse des décisions, un arbre de décision peut être utilisé pour représenter explicitement les décisions prises et les processus qui y conduisent (Rokach & Maimon, 2007).

L'algorithme forêt aléatoire (RF) a été introduit par Breiman (2001). Il représente une intégration de plusieurs prédicteurs d'arbres où chaque arbre dépend séparément des valeurs d'un vecteur aléatoire. La précision de la classification augmente considérablement lorsque le groupe d'arbres est agrandi. La forêt aléatoire consiste à faire tourner de nombreux arbres de décision qui sont construits de manière aléatoire, puis à les générer. L'échantillonnage bootstrap (OOB: Out-of-Bag sampling) est utilisé en association avec RF pour avoir une meilleure estimation de la distribution de l'ensemble de données d'origine. En effet, le bootstrap signifie la sélection au hasard d'un sous-ensemble de données pour chaque arbre plutôt que d'utiliser toutes les données pour construire les arbres. En matière de statistiques, si les arbres ne sont pas corrélés, cela réduit la variance prévue. Le principal avantage des forêts aléatoires est leur résistance aux variances et aux biais (Breiman, 2001). Les forêts aléatoires sont utilisées en régression et classification. Dans le cas d'une régression, la variable dépendante à prédire est continue, l'erreur dans cette prévision est appelée erreur Out Of Bag (OOB) et est calculée comme suit (Breiman, 2001) :

$$\text{Error OOB}_s = \frac{1}{n} \sum_{k=1}^{n} (y_k - \widehat{y_k})^2 \qquad (1.2)$$

Pour la classification, la variable dépendante prédite est catégorielle. En effet, une variable catégorielle avec N modalités est décomposée en un tableau disjonctif (avec N-1 variables) selon un schéma de codage 0-1. La forêt aléatoire est construite en échantillonnant arbitrairement le sous-ensemble d'entités ainsi que le sous-ensemble d'apprentissage pour chaque système. Le vote majoritaire fournit la prédiction finale.

K-means est un algorithme d'apprentissage machine non supervisé, qui se base sur la partition des données en plusieurs groupes k (d'où vient le nom de k-means). Le nombre de groupes k est le seul paramètre dans le modèle k-means, qui doit être spécifié en avance par l'utilisateur. Les étapes de base pour le regroupement de k-means sont les suivantes : 1) Sélectionner le nombre de clusters k. 2) Sélectionner au hasard des points de données distincts qui représentent le nombre de catégories. 3) Attribuer chaque point de données à son centroïde (cluster k) le plus proche en mesurant la distance entre le premier point de données et les points de classes

initiaux. 4) Calculer la moyenne de chaque cluster et remplacer le centre d'origine par le nouveau centre de position pour chaque cluster. L'algorithme converge lorsque les points affectés ne changent plus. Autrement, il faut répéter les étapes précédentes jusqu'à ce qu'aucun point de données ne change de position. La performance du clustering peut être évaluée en ajoutant la variation au sein de chaque cluster. La distance entre les points et les classes k est mesurée en fonction d'une distance numérique. À cet égard, on trouve la distance de Manhatten et la distance de Minkowski (Singh, Yadav, & Rana, 2013). La distance euclidienne, quant à elle, présente la mesure de similarité la plus courante utilisée pour les cas de clustering, donnée par l'équation suivante (X. Wu et al., 2008):

$$\sum_{i=1}^{n} \left(argmin_j \, \|x_i - \, \mu_k\|^2 \right) \tag{1.3}$$

Le concept de base de k-means repose sur des calculs mathématiques (la moyenne) en utilisant une distance numérique (généralement la distance euclidienne). Cependant, il existe plusieurs circonstances où les variables sont catégorielles. Dans ce cas, k-modes peut être adopté et la distance est calculée en fonction de la distance de Hamming (Z. Huang, 1998). Par ailleurs, des ensembles de données mixtes se produisent fréquemment dans de nombreux domaines. En effet, Les données mixtes comprennent des caractéristiques numériques et catégorielles simultanément. Dans ce cas, la distance peut se calculer par Gauss Kernel, la distance de Gower, ou encore de combiner la distance euclidienne avec la distance de Hamming (Ahmad & Khan, 2019; Ren, Liu, Wang, & Pan, 2016).

1.2.2 Problèmes soulevés

Comme toutes les approches de prévision, l'apprentissage machine présente des problèmes qui peuvent compromettre la validité des estimations. Premièrement, les données doivent être assez fiables et à jour. Un autre problème est lié à la sélection des variables qui peuvent biaiser la performance du modèle. En effet, il est difficile de capturer les variables correctes qui influencent l'évolution technologique. De ce fait, la performance de classification peut être réduite en raison des caractéristiques non pertinentes et redondantes de l'ensemble de données.

En réalité, pas toutes les variables sont utiles pour l'apprentissage. Par conséquent, le bon choix des variables utilisées pour l'apprentissage peut maximiser la performance des algorithmes (Macas et al., 2014).

Dans ce cadre, des travaux ont été menés sur la sélection des caractéristiques (en anglais : feature selection), dont les algorithmes d'optimisation métaheuristiques qui ont largement été discutés dans la littérature et qui ont mené à des améliorations considérables au niveau de la classification (Dianati, Song, & Treiber, 2002; Yang & Honavar, 1998a).

1.3 Optimisation métaheuristique

Contrairement aux méthodes exactes classiques, comme la programmation linéaire ou dynamique, les métaheuristiques sont des méthodes de résolution approchées. Un algorithme d'optimisation métaheuristique vise à résoudre un problème d'optimisation difficile. Il est adopté lorsque le recours aux méthodes classiques n'est plus efficace. Les problèmes d'optimisation dans cette thèse sont associés à l'intelligence artificielle; notamment à l'apprentissage machine qui a pour objectif de comprendre et analyser la structure de données de grosses tailles et de les intégrer dans des systèmes, pour ensuite optimiser ou prédire quelque chose à partir des modèles construits. En effet, les algorithmes métaheuristiques apparaissent dans les modèles de l'apprentissage machine comme des algorithmes de recherche qui visent à apprendre les caractéristiques du modèle pour ensuite donner les meilleures solutions à un problème bien déterminé, et ce, lorsque l'algorithme atteint l'optimum global.

Les algorithmes métaheuristiques ont largement été utilisés dans le domaine de l'apprentissage machine et ont apporté des meilleurs résultats (K.-Y. Chen & Wang, 2007; Du, Liu, Yu, & Yan, 2017; Gonzalez, Padilha, & Barone, 2015; Pan, Xue, Zhang, & Zhao, 2011).

Il existe deux types de métaheuristiques : 1- À solution unique, aussi connue sous le nom de méthode de trajectoire, tels que la recherche tabou (Glover, 1986), le recuit simulé (Kirkpatrick, Gelatt, & Vecchi, 1983), la recherche locale, la recherche à voisinage variable (VNS) (Mladenović & Hansen, 1997) et la recherche exhaustive. Ces algorithmes sont amplement utilisés dans les problèmes d'optimisation, le plus connu est le problème de

recherche des plus proches voisins. 2- À population de solutions, dispose de plusieurs points appelée population, le plus connue est l'algorithme génétique (AG). De plus, on trouve l'Essaims particulaires (PSO), les colonies de fourmis, etc. Ces algorithmes évolutifs s'inspirent de la nature biologique et du mécanisme d'évolution de la nature.

1.3.1 Comparaison et choix des algorithmes de l'étude

L'avantage principal de l'AG et le PSO réside du fait qu'ils sont des algorithmes évolutifs indépendants et capables de s'adapter à n'importe quel problème. L'algorithme génétique dispose d'une qualité spéciale par rapport aux autres techniques d'optimisation et c'est l'exploration de l'espace de recherche en se basant sur des paramètres aléatoires. Par ailleurs, le PSO dispose de plusieurs avantages comme la flexibilité et la robustesse, étant donné qu'il détient ses caractéristiques à partir des phénomènes naturels.

Le choix de l'algorithme adéquat dépend de chaque problème à étudier. Le critère de choix de l'algorithme dans cette recherche a été basé principalement sur la complexité en temps (temps nécessaire que l'algorithme prend jusqu'à sa convergence), et la complexité en espace mesurée en fonction de ses entrées (capacité de mémoire).

La complexité en temps est définie comme le temps de convergence d'atteinte de l'optimum global. Il est généralement associé à la taille de la population et à la dimension du problème (le nombre d'itérations, etc.). Le temps total est la somme de toutes ces mesures. En effet, les algorithmes comme le recuit simulé ou la recherche exhaustive prennent beaucoup de temps pour se converger. La recherche exhaustive par exemple, consiste à essayer toutes les solutions possibles. Cependant, chaque fois que la dimension du problème augmente, la fonction de complexité en temps augmente en conséquence. Tandis que le PSO commence à partir d'une position aléatoire et tente de s'évoluer peu à peu pour améliorer l'ensemble de solutions (la population). Cet avantage lui permet de gagner du temps pour se converger beaucoup plus rapidement à l'optimum global (Lobo, Goldberg, & Pelikan, 2000; Zhang, Zou, & Shen, 2018). Par rapport à la complexité en espace (la mémorisation), comparé à la recherche exhaustive, le

PSO et le AG montre un meilleur résultat puisqu'il dispose d'une grande robustesse face à la dimension du problème et même au mauvais choix des paramètres.
À l'appui de ces critères, l'algorithme génétique et l'optimisation par essaims particulaires ont été choisis.

1.3.2 Algorithme génétique

Holland (1975) a été la première personne à introduire l'algorithme génétique canonique (CGA). Les algorithmes génétiques (AG) sont des algorithmes d'optimisation basés sur des techniques dérivées de la génétique et des mécanismes d'évolution naturelle. Bien que les AG puissent être utilisés dans la sélection et la classification des fonctionnalités, ils peuvent également résoudre des problèmes non linéaires et trouver la meilleure solution à travers certaines opérations pour obtenir un ensemble de paramètres de conception souhaités. Ces opérations sont définies comme suit : la sélection consiste à choisir les individus de la population actuelle qui survivront et se reproduiront. Cette opération est basée sur la valeur de la fonction « fitness » qui évalue les solutions. Le croisement (en anglais : crossover) est défini comme des individus qui sont divisés, au hasard, en paires. Les chromosomes sont ensuite copiés et recombinés pour former deux descendants avec les caractéristiques des deux parents. La mutation présente une modification aléatoire de la valeur d'un allèle dans un chromosome. Cet opérateur évite une convergence prématurée vers des optima locaux. Il est appliqué avec une probabilité fixe, P_m (Holland, 1975). Les AG peuvent s'adapter à n'importe quel espace de recherche. Du fait de leur nature modulaire, ces algorithmes évolutifs sont presque totalement indépendants du problème à optimiser. Cependant, il est nécessaire d'effectuer plusieurs expériences pour ajuster des paramètres de l'algorithme tels que la taille de la population, la probabilité de croisement et de mutation ainsi que le nombre de générations.
Le principe de fonctionnement de l'algorithme est comme suit : la première phase consiste à construire une population initiale au hasard. Après cela, les AG sélectionnent les meilleurs individus (parents) de la génération actuelle à chaque étape pour créer les enfants de la génération suivante. La sélection des individus est basée sur les résultats de la fonction fitness de chaque individu. Après cela, le chromosome de la prochaine génération est sélectionné en

choisissant les éléments les plus adaptés de la nouvelle génération. L'opération suivante consiste à choisir un point aléatoire et à échanger les codes (bits) des paires de chromosomes et à mettre en œuvre une mutation, en remplaçant l'élément des chromosomes actuels par un nouvel élément généré. Cette étape est connue sous le nom de croisement et de mutation. La phase de reproduction d'une nouvelle population est construite à partir des individus sélectionnés selon des opérateurs de croisement et de mutation. L'algorithme s'arrête une fois que le nombre d'itérations requis ou le temps d'exécution est atteint. La solution actuelle est adoptée lorsque le test est validé; sinon, les étapes sont reprises à nouveau.

1.3.3 Optimisation par essaims particulaires

L'approche par essaims particulaires, appelée aussi intelligence collective (*PSO*) a été développée par Eberhart et Kennedy en 1995(Kennedy, 2010; Kennedy & Eberhart, 1995). Dans un contexte d'optimisation, plusieurs chercheurs ont appliqué le *PSO* en data mining (Du et al., 2017; Jinglin, Yayun, Yanan, & Weilan, 2017; Lin, Ying, Chen, & Lee, 2008; Xiaodan, 2017). Le processus de fonctionnement de la PSO est comme suit : l'étape 1 consiste à initialiser au hasard une population. L'étape 2 est consacrée à la mesure de la forme physique de chaque particule dans la population. La 3^{e} étape consiste à mettre à jour la vitesse et la position de chaque particule en recherchant les meilleures performances pour chaque particule (optimum local). Si la forme physique actuelle est meilleure (*pbest*) que la forme physique précédente, le *pbest* précédent est remplacé par le *pbest* actuel. Le processus de sélection continue jusqu'à ce que le processus converge. L'arrêt de l'algorithme est lié à la satisfaction du critère de terminaison.

1.4 Problèmes soulevés et pistes de recherche

Les méthodes d'optimisation métaheuristiques ont largement été discutées dans la littérature. Les résultats satisfaisants ont favorisé l'exploration de nouvelles pistes pour les intégrer dans ce domaine de recherche. Pour cela, une piste de recherche est formée, vise à intégrer des algorithmes d'optimisation métaheuristiques dans les algorithmes de l'apprentissage machine afin d'améliorer la précision des modèles de prévision. Les algorithmes d'optimisation

présentés dans cette thèse ont pour fonction de sélectionner les variables appropriées et d'optimiser les paramètres des algorithmes de l'apprentissage machine.

Avec autant d'avantages, il existe cependant une limite avec les approches métaheuristiques qu'il faut mentionner. Ces algorithmes sont performants, mais derrière cette performance, le bon choix des paramètres. En effet, les paramètres de l'algorithme doivent être ajustés adéquatement et cela peut être obtenu après plusieurs expérimentations et avec l'expertise de l'utilisateur qui lui permettra de faire un choix.

1.5 Conclusion

La revue de littérature a permis de fournir un contexte bien défini de l'analyse prédictive. Deux types de modélisation statistique sont analysés pour la prévision de l'obsolescence : la modélisation statistique classique et l'introduction de l'apprentissage machine. La plupart des approches de prévision de l'obsolescence étudiées dans la littérature sont basées sur des modèles classiques. Toutefois, cette recherche a révélé d'autres méthodes de prévision qui semblent être performantes, précises et stables. En résumé, une méthode efficace de prévision de l'obsolescence devrait prendre en considération les facteurs qui influencent la technologie; elle devrait aussi être capable de prévoir un grand nombre de données et garantir une meilleure précision de prévision. Le travail dans cette thèse est une contribution de la recherche qui est dans cette voie.

CHAPITRE 2

DÉMARCHE MÉTHODOLOGIQUE

Ce chapitre vise à préciser le choix et la démarche méthodologique associés au contexte de recherche. Trois principales sections sont discutées. La première détaille le choix de la stratégie de recherche suivie, la deuxième présente le cheminement méthodologique détaillé et le lien entre les articles, et la dernière présente les articles de revue.

2.1 Choix de la démarche méthodologique

La formulation du problème de recherche se définit par le choix du thème de recherche à étudier. En effet, le thème de cette recherche a été motivé principalement par un besoin industriel; ce besoin a permis d'identifier les lacunes existantes dans la littérature et de formuler la problématique de recherche. D'autre part, des sondages et des communications avec l'entreprise partenaire ont contribué à bien cibler le cadre conceptuel de la recherche.
Chaque recherche suit une démarche méthodologique. Dans cette étude, la méthodologie suivie est celle dictée par la science de conception qui se base principalement sur l'analyse des lacunes dans les méthodes de prévision d'obsolescence actuelles afin de les améliorer d'une manière plus efficace qu'auparavant.

La démarche de la science de conception est une méthode de recherche scientifique qui consiste à fournir un nouvel outil pour résoudre soit un problème non résolu (par exemple industriel), soit à améliorer une recherche antécédente. Les outils de conception peuvent être un modèle, une méthode ou un concept (Fortin & Gagnon, 2010; Hevner, March, Park, & Ram, 2004). De ce fait, l'approche de la science de la conception semble justifiée.

2.2 Stratégie de recherche

L'objectif principal de la thèse vise à développer et à améliorer des modèles de prévision du risque et de cycle de vie de l'obsolescence des composants électroniques. La démarche

méthodologique suit les mêmes étapes que la science de conception : identification du problème, sélection de solutions, développement et démonstration de la solution choisie et, finalement, l'interprétation.

La Figure 2.1 présente la stratégie de recherche. La première étape correspond à l'étape préliminaire de la méthodologie de la science de la conception et c'est la connaissance du problème. Un examen de littérature est proposé à cet égard, qui consiste à identifier la problématique de recherche. Cette étape vise à analyser les lacunes dans certaines approches existantes pour ensuite définir le problème et les requis avant l'étape de conception. Les étapes suivantes sont consacrées à la conception et au développement des outils de prévision. Chaque étape présente aussi les outils utilisés pour chacun des concepts.

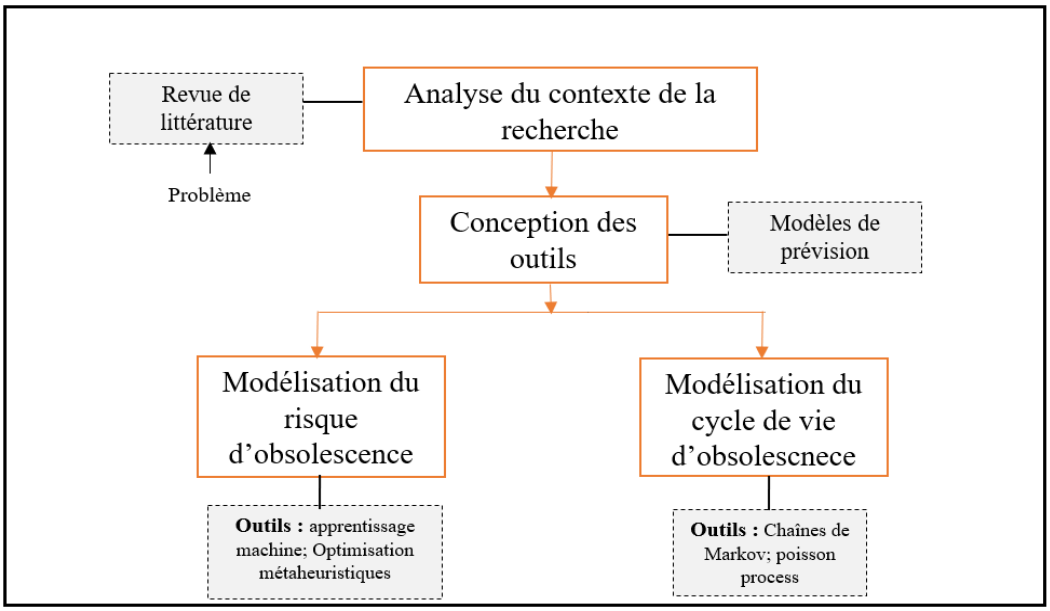

Figure 2.1 Démarche méthodologique

La recherche présentée dans ce travail a permis d'élaborer trois articles de revue qui sont inclus intégralement dans les chapitres 3 à 5. En effet, la recherche s'articule sur deux axes principaux : 1) la prévision du risque de l'obsolescence 2) la prévision du cycle de vie de l'obsolescence. Le premier axe de recherche introduit l'apprentissage machine appliqué à l'obsolescence des composants électroniques. Deux modèles basés sur l'exploration de données, soit l'apprentissage supervisé et non supervisé sont proposés dans la démarche

méthodologique afin de prévoir le risque de l'obsolescence. La modélisation introduit également les algorithmes d'optimisation métaheuristiques pour améliorer la performance du modèle. Le deuxième axe de recherche s'intéresse à la prévision du cycle de vie et à la date d'obsolescence, en intégrant les chaînes de Markov et le processus de poisson composé.

2.3 Démarche conceptuelle des outils de prévision d'obsolescence

2.3.1 Justification du choix des méthodes de modélisation

Le choix des méthodes de conception dans ce travail est justifié en se reposant sur deux atouts importants, la littérature et le besoin industriel. Partant par le besoin industriel, et comme discuté auparavant, l'entreprise partenaire avait des problèmes d'obsolescence. Des communications et des entretiens ont révélé la faiblesse et le manque de stratégies efficaces pour gérer l'obsolescence. En revanche, ces entretiens ont favorisé la compréhension et à positionner leurs besoins et leurs attentes. Les parties prenantes de ce projet ont exprimé leur besoin d'un outil pour la prévision de l'obsolescence. Cependant, la portée et la forme de ce modèle n'étaient pas claires de leur côté. Pour déterminer la méthode adéquate, une revue critique de la littérature est amenée à cet égard, cette dernière a révélée de nombreuses méthodes de modélisation de risque et de cycle de vie d'obsolescence. Toutefois, la méthode choisie dépend généralement des caractéristiques de la pièce et des données disponibles dans l'entreprise. Dans la littérature, la majorité des travaux sont basés sur la prévision d'obsolescence d'un seul type de composant. Néanmoins, il existe des milliers de fin de vie des composants électroniques, qui affectent par conséquent les systèmes complexes. De ce fait, l'apprentissage machine est adopté pour prévoir un large échantillon de données. D'autre part, le choix de la méthode basée sur les chaînes de Markov est justifié à partir des limites dans les approches existantes dans la littérature. Cette méthode apparaît comme une meilleure approche pour prévoir le cycle de vie, étant donné qu'elle prend en considération toutes les étapes du cycle de vie. D'après Jaarsveld et Dekker (2011), une bonne méthode de prévision d'obsolescence doit tenir compte de tous les états du cycle de vie dont lesquels la demande peut augmenter ou baisser. Cette méthode peut améliorer considérablement la précision de la prévision de la date d'obsolescence. D'un autre côté, la chaîne de Markov dispose d'une qualité

importante et c'est la propriété puissante de Markov, implique que les propriétés des variables aléatoires liées à l'avenir ne dépendent que de l'information de l'état présent, et non de l'information provenant des états passés. Ces méthodes sont donc proposées pour répondre au besoin industriel et à contribuer au développement scientifique.

2.3.2 Modélisation du risque d'obsolescence basé sur l'apprentissage supervisé

La démarche du travail commence par introduire les articles de conférence qui ont contribué à rédiger les trois articles de journaux. Des articles de conférence sont ajoutés aux annexes I à III, dans le but de supporter davantage le sujet de recherche et le cheminement méthodologique. Ils contribuent aussi à développer les réponses aux objectifs spécifiques pour aider à développer les approches proposées.

2.3.2.1 Étude préliminaire : comparaison des algorithmes

La première phase de la conception consiste à mener une comparaison entre les algorithmes de réseaux de neurones artificiels, Adaboost, machine à vecteurs de support, forêt aléatoire et arbre de décision qui sont capables de prévoir le risque d'obsolescence en se basant sur l'apprentissage supervisé afin d'estimer le meilleur prédicteur. Cette étude a été réalisée et documentée dans un article de conférence présenté dans l'annexe I, intitulé *« Obsolescence forecasting strategy of technological components - A comparative study of algorithms »* et présenté à la conférence *« 7th IESM Conference, October 11 – 13, 2017, Saarbrücken, Germany ».* En effet, le choix des algorithmes a été basé sur une analyse de la littérature pour distinguer les meilleurs algorithmes de prévision.

La démarche méthodologique suit les étapes suivantes : 1) la collection des données. Les données sont collectées à partir d'une base de données disponible qui fournit des informations sur les caractéristiques techniques des composants électroniques ainsi que leurs statuts (en production ou obsolète), plus de 700 composants ont été collectés. 2) Par la suite, les données sont divisées en deux groupes aléatoires, soit 2/3 pour l'entraînement et 1/3 pour le test et la validation. L'échantillon est introduit ensuite dans l'algorithme pour construire le modèle de

prévision à partir de l'entraînement de l'échantillon. 3) La sortie du modèle est présentée par une matrice de confusion donnée par le Tableau 2.1.

Tableau 2.1 Matrice de confusion tiré de Grichi et al. (2017)

Algorithmes	Actuel	Obsolète	En production	Erreur
BP-NN	**Obsolète**	12	9	21%
	En production	16	60	
SVM	**Obsolète**	9	12	22%
	En production	9	67	
AdaBoost	**Obsolète**	53	26	25%
	En production	32	110	
RF	**Obsolète**	10	11	19%
	En production	7	69	
Decision tree	**Obsolète**	53	26	24%
	En production	27	115	

Les résultats ont montré que RF a surpassé les autres algorithmes avec la meilleure précision. Il est à noter que les algorithmes montrent une amélioration de la précision lorsque la taille de l'échantillon augmente à l'entrainement.

2.3.2.2 Prévision d'obsolescence basée sur l'algorithme forêt aléatoire

La deuxième phase de la conception est consacrée à développer un modèle de régression basé sur la forêt aléatoire pour prévoir la tendance de l'obsolescence. Le modèle réagi sur la corrélation entre la date d'introduction et d'autres caractéristiques techniques. Cette étude a été réalisée et documentée dans un article de conférence, intitulé *« A random forest method for obsolescence forecasting »* et présenté à la conférence *« IEEE International Conference on Industrial Engineering and Engineering Management (IEEM) 2017, Singapore ».*

La Figure 2.2 présente le fonctionnement de la forêt aléatoire pour construire le modèle de prévision.

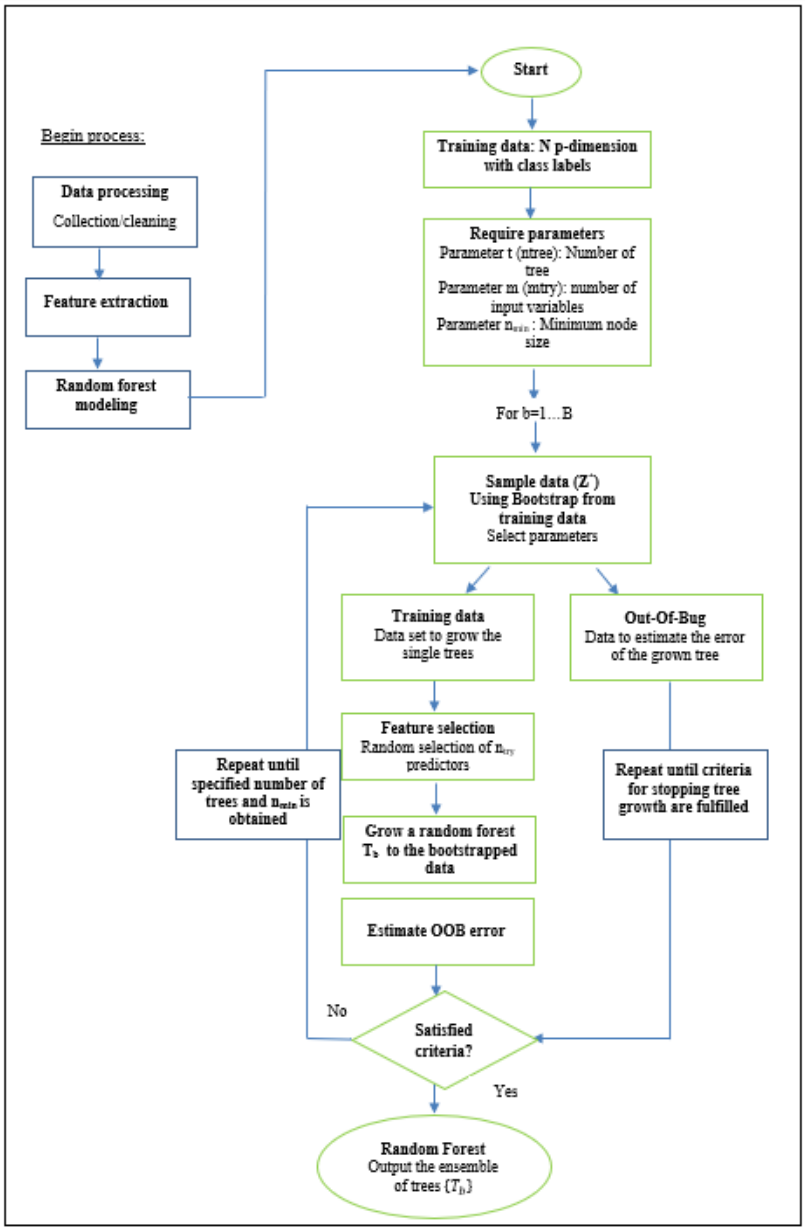

Figure 2.2 Mécanisme de l'algorithme forêt aléatoire

Tirée de Grichi et al. (2017)

L'échantillon est divisé en deux sous-ensembles. Le premier comprend généralement 70% des données et il est utilisé pour entrainer les données, c'est ce qu'on appelle l'ensemble d'apprentissages. Le deuxième sous-ensemble comprend 30% des données et il est utilisé pour le test et la validation du modèle. L'ensemble d'apprentissages est introduit dans le RF pour créer le modèle prédictif. Ensuite, les paramètres du modèle sont estimés. Pour chaque nœud

terminal, ces étapes seront répétées jusqu'à ce que le nombre d'arbres soit spécifié et que la taille minimale du nœud soit obtenue. L'étape finale est l'estimation de l'erreur OOB pour le modèle et finalement la sortie est présentée comme un ensemble d'arbres (T_b).
La sélection des paramètres constitue une étape importante dans la conception de l'algorithme. Pour le modèle du forêt aléatoire dans la prévision de l'obsolescence, trois paramètres sont requis. En effet, pour chaque nœud le tirage des variables se fait sans remplacement et uniformément parmi toutes les p variables explicatives (chaque variable a une probabilité 1/p à être choisie). Le nombre m qui représente le nombre de variables tirées aléatoirement à chaque nœud, est fixé au début de la construction du forêt et est identique pour tous les arbres.

2.3.2.3 Optimisation de la prévision de l'obsolescence

À partir des résultats de la comparaison entre les algorithmes de l'apprentissage machine, l'étape suivante de la méthodologie est consacrée à l'amélioration de la performance de l'algorithme forêt aléatoire. En effet, la littérature et l'étude expérimentale ont révélé quelques limites par rapport à la prévision avec l'apprentissage machine, et ce, au niveau de la sélection des variables. La méthode vise à intégrer un algorithme métaheuristique dans le modèle de forêt aléatoire pour optimiser les paramètres et les variables appropriées. Le modèle est donné par la Figure 2.3.

La construction du GA-RF suit plusieurs étapes. Premièrement, les paramètres et le modèle de la forêt aléatoire sont construits. Ensuite l'algorithme génétique et les paramètres associés sont intégrés pour construire le modèle GA-RF. Une étude expérimentale est réalisée basée sur les données collectées auparavant pour la comparaison des algorithmes. Pour examiner la faisabilité de cette approche de même que les résultats, cette partie présente une comparaison entre GA-RF, RF, GBM, ainsi que la régression logistique pas à pas, sous forme d'une matrice de confusion.

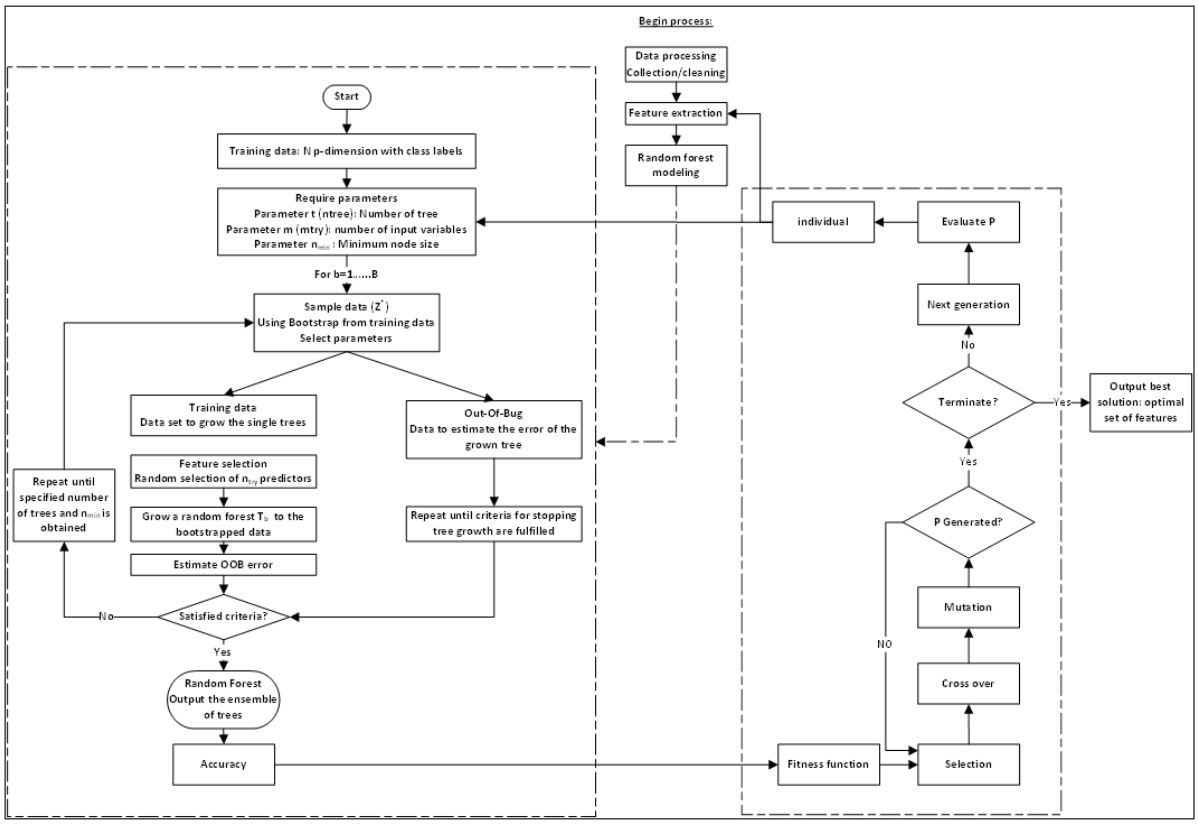

Figure 2.3 Architecture du modèle GA-RF tirée de Grichi et al. (2018)

2.3.3 Modélisation du risque d'obsolescence basé sur l'apprentissage non-supervisé

La deuxième partie de la modélisation du risque d'obsolescence vise à introduire une approche de combinaison entre l'apprentissage non-supervisé (clustering) et l'apprentissage supervisé. La méthodologie est basée sur deux phases. La première est définie comme un problème de classification, discutée dans les sections précédentes. Cependant, la deuxième phase est consacrée à introduire l'apprentissage machine non-supervisé pour former des groupes avec k-prototype. L'objectif d'intégrer la technique de clustering consiste à regrouper les composants non-obsolètes dans différentes catégories de niveau de risque basé sur le modèle prédictif développé dans la première phase.

Étant donné que k-prototype divise uniquement l'ensemble de données en plusieurs groupes, le modèle prédictif aide à définir les groupes comme étant à risque faible, moyen ou élevé. En effet, ce modèle réagit sur les données qui sont classées comme non obsolètes. À partir de cela, un calcul de probabilité est élaboré pour chaque composant prédit comme non obsolète (ce qui veut dire = 1). Pour considérer un composant actif avec une probabilité élevée, la valeur doit être proche de 1. Par exemple, si la probabilité d'un groupe est inférieure ou égale à 0.5, cela veut dire que le composant a une probabilité faible d'être en production et le risque de devenir obsolète est élevé. En revanche, si la probabilité est supérieure à 0.5, (la valeur est proche de 1), cela signifie que la probabilité d'être actif est élevée, et le risque de tomber obsolète est faible. À partir de ces probabilités les classes peuvent être définies.

La raison du choix de cette méthode est que certains composants n'ont pas de pilote paramétrique évolutif clair (P. Sandborn et al., 2011) , et par conséquent, la prévision de leur obsolescence n'est pas aussi simple. Cette approche aidera à évaluer le niveau de risque basé sur des techniques de regroupement pour aider les gestionnaires à choisir les meilleures approches de stratégies de mitigation.

2.3.4 Développement d'un modèle de prévision du cycle de vie d'obsolescence

Le deuxième axe de recherche se concentre sur la modélisation d'une approche de prévision du cycle de vie. La méthodologie utilisée est celle dictée par la science de la conception qui vise à améliorer une technique ou une approche existante. La méthode proposée considère les cinq étapes de la courbe de vie d'un produit pour prévoir le cycle de vie jusqu'à l'obsolescence. Les cinq états sont présentés comme suit :

État 5: présente la phase d'introduction où un nouveau produit arrive sur le marché (la demande est lente mais en augmentation);

État 4: présente la phase de croissance, la demande commence à augmenter;

État 3: présente la phase de maturité, la demande est élevée et les prix sont plus bas;

État 2: présente la phase de déclin, baisse de la demande et les concurrents commencent à introduire un nouveau produit;

État 1: présente le dernier état du cycle de vie où la demande est morte.

La chaîne de Markov et le processus de poisson composé sont proposés à cet égard. Dans chaque état, la demande suit un processus de poisson homogène à intensité constante. La probabilité de transition est estimée avec la distribution de poisson avec un taux de demande variable. La transition entre les états peut sauter plusieurs étapes. Par exemple, la transition de l'état 5 à 1 peut être provoquée par l'intégration d'un concurrent dans le marché. Elle peut également être causée par plusieurs événements futurs indépendants, tels que l'introduction de plusieurs produits concurrents qui peuvent entraîner une baisse de la demande. La probabilité de transition entre les états du cycle de vie est présentée dans la Figure 2.4.

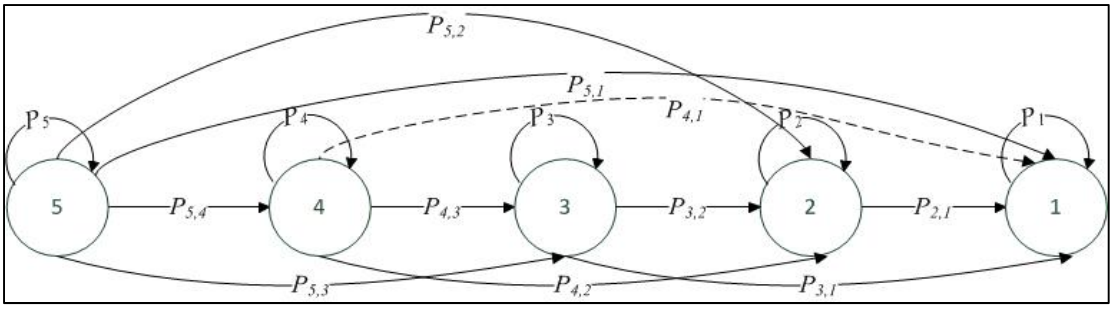

Figure 2.4 Probabilité de transition entre plusieurs états tirée de Grichi et al. (2019)

La méthode proposée permet de prévoir la date à laquelle le composant sera obsolète, en tenant compte de l'ensemble de la courbe de cycle de vie estimé à partir du taux d'arrivé des commandes. Les hypothèses retenues de l'étude sont les suivantes :

- Il y a cinq états de Markov avec seulement une transition vers l'avant, dans lequel le dernier est considéré comme l'état absorbant.
- Les transitions entre les états se produisent de façon aléatoire dans le temps, et la durée dans un état particulier suit une distribution exponentielle.
- Dans chaque état, la demande suit un processus de poisson homogène basé sur un processus de comptage de Markov (*Markov counting process*) à intensité constante λ_i.
- La gestion du stock est considérée comme suit : Si la demande augmente, il peut y avoir des problèmes de disponibilité temporaires. Si la demande diminue, les stocks sont

temporairement élevés. Cependant, si la demande chute, les stocks deviennent obsolètes.

2.4 Lien entre les objectifs spécifiques de la thèse et les articles de revue

. L'article 1, présenté dans le chapitre 3, a répondu au premier objectif spécifique de la thèse par le développement d'une méthode qui consiste à prendre en considération la prévision du risque de l'obsolescence basé sur l'apprentissage machine supervisé et l'optimisation en intégrant les algorithmes d'optimisation métaheuristiques. D'un autre côté, l'apprentissage non supervisé est introduit pour répondre au risque d'obsolescence discuté dans le deuxième article et présenté dans le chapitre 4. L'article 3, présenté au chapitre 5, a répondu au troisième objectif qui concerne la prévision du cycle de vie de l'obsolescence.

Tableau 2.2 Lien entre les objectifs de la thèse et les articles de revue

Objectif de la recherche	Développement et optimisation des modèles de prévision de l'obsolescence des composants électroniques		
Objectifs spécifiques	Développer et optimiser un modèle basé sur l'apprentissage machine supervisé pour prévoir le risque d'obsolescence	Développer un modèle basé sur l'apprentissage machine non-supervisé pour prévoir le risque d'obsolescence	Développer un modèle pour prévoir le cycle de vie d'obsolescence
Outils utilisés	Forêt aléatoire; algorithme génétique;	*k-means cluster; PAM*	Chaînes de Markov; processus de poisson composé.
Article répondant à chaque objectif	**Chapitre 3** *"Optimization of Obsolescence Forecasting Using New Hybrid Approach Based on the RF Method and the Meta-heuristic Genetic Algorithm"*	**Chapitre 4** *"Integration of unsupervised and supervised machine learning for obsolescence risk assessment"*	**Chapitre 5** *"An approach to obsolescence forecasting based on Hidden Markov Model and Compound Poisson Process"*

2.5 Présentation des articles de journal

2.5.1 Article de revue nº 1: Optimisation de la prévision de l'obsolescence en utilisant une approche basée sur la forêt aléatoire et l'algorithme génétique.

Le chapitre 3 présente un article de revue scientifique, publié en août 2018 dans la revue *« American journal of management »* et intitulé *« Optimization of obsolescence forecasting using new hybrid approach based on the RF method and the meta-heuristic genetic algorithm ».*
Cet article propose une nouvelle approche nommée *GA-RF*, qui contribue à améliorer la précision de la classification de la forêt aléatoire, et ce, après une étude préliminaire basée sur une comparaison, et qui a pour objectif de choisir le meilleur algorithme de prévision.
Cet article propose une intégration d'un algorithme d'optimisation métaheuristique dans le modèle de la forêt aléatoire. En effet, l'algorithme génétique (GA) choisi, a pour fonction de rechercher les paramètres optimaux et d'optimiser la sélection des caractéristiques appropriés pour construire la forêt aléatoire (RF) afin d'améliorer sa classification.

La première partie de la méthodologie est consacrée à la modélisation de RF, les paramètres et les variables sont initiés, les étapes de l'apprentissage sont élaborées. La deuxième partie, quant à elle est, consacrée à la modélisation de l'algorithme génétique, ensuite l'algorithme génétique est intégré dans le RF pour construire le modèle après avoir initialisé les données d'entrées. La sortie présente le modèle de forêt aléatoire pour la classification de l'obsolescence. Une étude complémentaire à cet article vise à intégrer un autre algorithme d'optimisation connu sous le nom « optimisation par essaims particulaires » (en anglais : *partical swarm optimisation*), présenté dans l'annexe III et qui vise à comparer les deux modèles GA-RF avec PSO-RF.

Cet article contribue au domaine des connaissances par l'introduction de l'apprentissage machine et améliore la prévision de l'obsolescence par l'intégration d'une approche métaheuristique, en proposant une approche innovante de combinaison entre la forêt aléatoire et l'algorithme génétique. La performance est calculée en termes de mesures statistiques d'un test de classification binaire. Ces mesures de performance sont calculées comme suit :

$$Précision\ (AC) = \frac{TP+TN}{TP+TN+FP+FN} = 93.3\% \ ; \ \ Erreur = 1 - AC = 6.7\%$$

$$Sensibilité\ (SE) = \frac{TP}{TP+FN} = 90.4\% \ ; \ \ Spécificité\ (SP) = \frac{TN}{TN+FP} = 95.4\%$$

Où *TP, TN, FP* and *FN* sont définis respectivement comme vrais positifs (*true positive*), vrais négatifs (*true negative*), faux positifs (*false positive*) et faux négatifs and (*false negative*). Ces mesures sont distinguées à partir de la matrice de confusion.

2.5.2 Article de revue nº 2: Intégration de l'apprentissage automatique supervisé et non supervisé pour évaluer le risque d'obsolescence.

Le chapitre 4 présente un article de revue scientifique, soumis en juin 2020 dans la revue *« IEEE Transactions on Components, Packaging and Manufacturing Technology » et intitulé « Integration of unsupervised and supervised machine learning for obsolescence risk assessment».* Cet article propose une approche d'intégration combinée de l'apprentissage non supervisé avec l'apprentissage supervisé pour évaluer le risque d'obsolescence. La contribution de cette étude vise à introduire une méthode efficace pour la prévision de l'obsolescence basée sur deux apprentissages machine. La différence entre ce travail et d'autres travaux antérieurs est que ces travaux de recherche impliquent l'apprentissage supervisé et l'apprentissage non supervisé simultanément. Le risque d'obsolescence est alors estimé sur deux étapes différentes: la classification et la catégorisation du niveau de risque. Au meilleur de notre connaissance, l'apprentissage non supervisé n'a jamais été appliqué pour la prévision de l'obsolescence.

L'article est divisé en deux phases principales. La première phase évoque un problème de classification de l'apprentissage machine supervisé. Le but est de développer un modèle efficace capable de classifier l'obsolescence d'un grand nombre de composants (obsolètes ou en production). Le modèle et les résultats ont été tirés du premier article de revue. En revanche, la deuxième phase de l'article est consacrée à regrouper un ensemble de composants dans différentes classes de niveau de risque basé sur l'apprentissage non supervisé basé sur la l'algorithme k-prototype. L'évaluation du risque dans chaque groupe est basée sur le modèle prédictif obtenu à partir de la première phase. En effet, le modèle prédictif est intégré comme

une entrée dans le modèle d'apprentissage non supervisé afin d'obtenir des probabilités. Ces probabilités sont appliquées aux données prédites comme non obsolètes (=1), et à partir de ces probabilités les classes de risque peuvent être définies.

2.5.3 Article de revue nº 3: Prévision de l'obsolescence basée sur le modèle de Markov et le processus de Poisson composé

Le chapitre 5 présente un article de revue scientifique, publié en décembre 2019 dans la revue *« International Journal of Industrial Engineering and Operations Management »* et intitulé *« An approach to obsolescence forecasting based on Hidden Markov Model and Compound Poisson Process » (Yosra Grichi, Beauregard, & Dao, 2019).* L'article proposé évoque le manque de méthodes efficaces de prévision dans la littérature et s'inspire des limites dans certaines approches pour proposer une nouvelle méthode qui peut mieux prévoir le cycle de vie d'un produit. L'approche proposée aidera les entreprises à améliorer les prévisions d'obsolescence et à réduire son impact sur la chaîne d'approvisionnement. La première partie de l'article est consacrée à une étude critique et une comparaison avec des études antérieures pour identifier les lacunes afin de justifier la contribution de la nouvelle approche. La deuxième partie de l'article est consacrée au développement du modèle de Markov. La méthode introduit la théorie de Markov pour estimer les paramètres du modèle. Afin d'estimer la probabilité de transition, cinq états de Markov sont utilisés. L'état transitoire est direct et la durée dans un état particulier suit une distribution exponentielle. Cependant, la transition d'un état à l'autre peut se produire en cas de changement dans l'ordre de la demande. Finalement, la troisième partie de l'article est consacrée à la validation et la discussion des résultats.

Cet article contribue au domaine des connaissances en démontrant d'abord les limites d'une approche de prévision existante afin de permettre de définir une méthode puissante et innovante basée sur la chaîne de Markov qui considère tous les états du cycle de vie d'un composant jusqu'à l'obsolescence. Contrairement à la méthode développée par Jaarsveld (2011), cette méthode prévoit la possibilité de changements de la demande à travers les différents états du cycle de vie.

CHAPITRE 3

OPTIMIZATION OF OBSOLESCENCE FORECASTING USING NEW HYBRID APPROACH BASED ON RANDOM FOREST METHOD AND METHAHEURISTIC GENETIC ALGORITHM

Yosra Grichi, Yvan Beauregard, Thien-My Dao

Département de génie mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

Article publié dans « American journal of management », aout 2018

3.1 Abstract

Obsolescence is highly complex problems due to the influence of many factors such as technological advancement. However, prediction of obsolescence appears to be one of the most efficient solutions. This paper proposes a novel approach known as GA-RF for obsolescence forecasting. Genetic algorithm (GA) searches for optimal parameters and feature selection to construct a random forest (RF) in order to improve the classification of RF. To examine the feasibility of this approach, this paper presents a comparison between GA-RF, RF, Stepwise logistic regression, and stochastic gradient boosting. Experimental results show that GA-RF outperformed the other methods with 93.3% of accuracy, 90.4% of sensitivity and 95.4% of specificity.

3.2 Introduction

Obsolescence is a major problem caused by the increasing development of technologies. Various researchers define obsolescence in various ways (F. R. Rojo et al., 2010). A component becomes obsolete when the technology used to manufacture it is no longer available, supported or produced by the supplier (F J Romero Rojo, Baguley, Shaikh, Roy, & Kelly, 2012; Shen & Willems, 2014). In other words, obsolescence can be characterized by the loss of a supplier or raw materials (P. Sandborn, 2013; P. S. a. P. Sandborn, 2002).

The negative effects of technological obsolescence on the production performances have been studied in the literature and represent a major challenge in the long term. Technological obsolescence causes problems in the supply chain and in the management of electronic systems (P. Sandborn et al., 2011). Rapid technological progress is one of the factors that increase the rate of obsolescence. When one hears about technological obsolescence, the first thing that comes to mind is electronics. The electronic industry has emerged as the fastest growing sector and has spread widely around the world. As defined by Moore's law, the rapid evolution of electronic components continues to grow, which stipulates that semiconductor density doubles approximately every 18 months (Homchalee & Sessomboon, 2013; P. Sandborn, 2008; Tomczykowski, 2003). This evolution creates new electronic components every year with short lifetimes. In the USA, the industry has grown at a rapid rate since the 1990s.

New technologies are introduced in the market at increasing rates. Today, the short technological lifecycle and the lack of forecasting represent a challenge for several companies that need to take into account the risk of obsolescence. However, obsolescence forecasting appears to be one of the best solutions in the management obsolescence as it assists manufacturers to identify part obsolescence. Through obsolescence forecasting, companies can ensure support for parts in service. There are two types of obsolescence forecasting: long-term forecasting (1 year or longer), which allows a proactive management and life cycle planning to support a system, and short-term forecasting, which can be observed from the supply chain. Short-term forecasting may involve reducing the number of sources, reducing inventories, and increasing the price.

To overcome the problems caused by obsolescence, many studies have been conducted to create models that can effectively forecast obsolescence. Statistical methods such as regression, Partial least square regression (PLS), logistical regression and Gaussian method have been previously employed in many works (Gao et al., 2011; Jungmok & Namhun, 2017; P. Sandborn et al., 2011; Solomon et al., 2000). However, to conduct life cycle forecasting for thousands of components that affect day-to-day lives, there is a need to modify these methods. In this context, there are precise and stable forecasting approaches whose performance has been justified in the literature. In fact, results show that these approaches could be used for

nonlinear predictions with high accuracy and without human manipulation (Zurada, 1992). In the past few decades, machine learning has attracted the attention of many researchers in various disciplines and has been applied in many areas. According to (X. Wu et al., 2008) among the top ten algorithms identified by the Institute of Electrical and Electronics Engineers (IEEE) are AdaBoost, SVM, K-Means, decision tree, and Naïve Bayes. These algorithms have been proven to be good predictors. In addition, applying machine learning in obsolescence forecasting has received a lot of attention in the last two years (C. Jennings et al., 2016a).

In previous work, benchmarking studies have shown that the RF performs the best among current classification techniques and numerous experiments have shown that RF has a high degree of satisfactory classifications accuracies (Y Grichi et al., 2017; C. Jennings et al., 2016a). On the other hand, features selection has an important effect on the classification accuracy, because not all of them are useful for classification. Therefore, selecting the best features can achieve a better performance.

To improve the accuracy of machine learning algorithms for obsolescence forecasting, this study presents a novel random forest approach based on genetic algorithm to perform features selection and parameters optimization of random forest algorithm based on supervised learning. This paper also examines the feasibility of applying this proposed method by comparing it to other algorithms which are random forest, stepwise logistic regression and stochastic gradient boosting. To the best of the researcher's knowledge, this is the first time the RF algorithm based on genetic algorithm is applied to predict obsolescence of technological components.

3.3 Literature review

3.3.1 Obsolescence forecasting

Forecasting obsolescence is reactive in nature and is based on the resolution of the problem once noticed. The most classical reactive approaches are last LTB and exiting stock (F. R. Rojo

et al., 2010). However, there are two types of forecasting methods, namely forecasting of the obsolescence risk and forecasting of the obsolescence date (life cycle forecasting). Obsolescence risk is used to predict the probability that a component still in production (C. L. Josias, 2009; F. J. Romero Rojo et al., 2012; F. R. Rojo et al., 2012; van Jaarsveld & Dekker, 2011a). A few researchers focus on the prediction of the risk of obsolescence. In this context, Rojo conducted a Delphi study to analyze the risk of obsolescence. They developed a risk index using some indicators, which are years to end of life, the number of sources available, and the consumption rate versus availability of the stock. Another approach developed by Josias aims to create a risk index by measuring the manufacturers' market share, number of manufacturers, life cycle stage, and company's risk level. Finally, Jaasveld used product demand data history to estimate the risk of obsolescence. Alternatively, (Solomon et al., 2000) was the first to introduce the life cycle forecasting method. In his paper, Solomon conducted a study to predict the obsolescence date from the life cycle curve, which included six stages: introduction, growth, maturity, saturation, decline, and obsolescence. Another method based on data mining was developed by (P. A. Sandborn et al., 2007). The obsolescence date was obtained by applying Gaussian method. Moreover, other researchers have introduced regression analysis to predict the date of obsolescence (Gao et al., 2011). Last, (Y Grichi et al., 2017; C. Jennings et al., 2016a) have used data-driven method by create machine learning algorithms to forecast the obsolescence.

Today, the short life cycle of technology and the lack of forecasting represent a major challenge for the electronics industry. In fact, rapid technological progress leads to the production of technological components with very short lifetimes (Ward & Sohns, 2011). However, with this advancement of technology, obsolescence should be forecasted more accurately. Data-driven methods appear to be the most efficient solutions, we can find for example neural network (Guoqiang et al., 1998). In fact, these networks are capable of assimilating complex relationships between several variables, neural networks show good stability, better accuracy, and the ability to predict and manage a large data (Memmedli & Ozdemir, 2011). Furthermore, we find the support vector machine (SVM) algorithm. This algorithm was applied in several areas such as forecasting and optimization, SVM is considered one of the most precise and robust approaches within all the recognized algorithms. Only a few examples are required to

understand the model, and it does not depend on the number of dimensions (X. Wu et al., 2008).

3.3.2 Random forest

Introduced by (Breiman, 2001), Random forests are an integration of tree predictors where every tree depends on the values of a random vector separately. A similar distribution applies for all the trees in the forest. The tree classifier of a forest has a generalization error which relies on the strong correlation between all trees in the forest. Classification accuracy increases significantly when the group of trees is enlarged. A primary example is bagging, where to raise every tree, an arbitrary selection (without replacement) is done from the set examples. Another example is random split selection where arbitrarily, the split is selected from among the K best splits at every single node. A random forest algorithm consists of rotating many decision trees that are randomly constructed and then generating them. Bootstrap sampling (OOB: Out-Of-Bag sampling) is used in RF to have a better estimate of the distribution of the original dataset. Indeed, bootstrapping means randomly selecting a subset of the data for each tree rather than using all the data to build the trees. In statistical terms, if the trees are uncorrelated, this reduces the forecast variance. The main advantage of random forests is their resistance to variances and biases.

The random forest algorithm is used in the regression case to predict a continuous dependent variable, and in the classification case to predict a categorical dependent variable. For the regression type, a random forest consists of a set of simple prediction trees; each can produce a numerical response when presented with a subset of explanatory variables or predictors. The error in this forecast is called Out of Bag (OOB) error.

For the classification type, a categorical variable with N modalities is broken down into a disjunctive array (with N-1 variables) according to a 0-1 coding scheme. Thus, a categorical variable with N modalities can be considered as a set of N-1 variables, of which only one will assume the value 1 for a given observation. In fact, the ability to make predictions on a random subset of predictive variables is one of the strengths of the Random Forest module, which makes it particularly well suited to processing data sets with extremely high predictive

variables. This random feature selection encourages systems diversity, and by the end, it enhances classification performance. The random forest is constructed by sampling arbitrarily the features subset as well as the training subset regarding every system. The majority vote provides the final prediction. Finally, the random forest attains a favorable and vigorous performance with various applications (Cheng, Chan, & Qiu, 2012; J. Friedman, Hastie, & Tibshirani, 2001).

3.3.3 Genetic algorithm

(Holland, 1975) was the first person to introduce the canonical genetic algorithm (CGA). Genetic algorithms are optimization algorithms based on techniques derived from genetics and natural evolution mechanisms. While GA can be used in features selection and classification, it can also solve nonlinear problems and find the best solution through some operations to obtain a set of desired design parameters. These operations are defined as follows:

Selection involves choosing the individuals of the current population that will survive and reproduce. This operation is based on the value of the fitness function that evaluates the solutions. Crossover is defined as individuals that are randomly divided into pairs. The chromosomes are then copied and recombined to form two offspring with characteristics from both parents. Mutation presents a random modification of the value of an allele in a chromosome. This operator avoids premature convergence towards local optima. It is applied with a fixed probability, Pm.

GAs can adapt to any search space. Because of their modular nature, these evolutionary algorithms are almost completely independent of the problem to be optimized. They require a measure of the quality of the solution and a definition of space by a code and operators that allow it to browse effectively. With these advantages and the power of this algorithm, it is necessary to take into account the specificities of the problem to design or choose the code and adequate operators. It is necessary to perform several experiments to adjust such parameters of the algorithm as the size of the population, the probability of crossing and mutation, the generation number, and the replacement technique among others.

A simple GA works with a set of candidate solutions called a population, in which each chromosome (individual) represents a possible solution to the given problem. The individuals are obtained through a fitness function. The probability of having duplicates for each gene was defined by Holland as:

$$PS_i = \frac{f_i}{\sum_{i=1}^{s} f_i}; \text{ with } f_i \text{ presents the fitness} \quad (3.1)$$

The first phase of the algorithm involves building an initial population randomly. After that, the GA selects at each step the best individuals (parents) in the current generation to create the children of the next generation. The selection of individuals is done based on the results of every individual's fitness function. After that, the next generation's chromosome is selected by choosing the most adapted elements of the new generation. The next operation involves choosing a random point and exchanging the codes (bits) of chromosome pairs and implementing mutation by replacing the element of the current chromosomes with a new generated element. This step is known as crossover and mutation. The reproduction phase of a new population is constructed from the selected individuals though crossover and mutation operators. The algorithm stops once the required number of iterations or the execution time is achieved. The current solution is taken when this test is checked; otherwise, the steps are repeated all over again.

3.4 Proposed framework

Throughout this research, the literature has showed how to evaluate the most promising approaches to obsolescence prediction based on selection criteria such as the capacity and stability of the algorithms. Based on these criteria, some of machine learning algorithms have been the subject of this research. This work presents a random forest algorithm based on genetic algorithm for obsolescence forecasting. The RF classifier is chosen because it represents a successful ensemble learning algorithm which has been proven as the best

predictor for forecasting obsolescence risk based on a previous comparative study with a high degree of accuracy.

The main objective is to apply genetic algorithm to enhance the classification performance of random forest. The framework of predicting obsolescence of technological components using RF based on GA is illustrated in Figure 3.1.

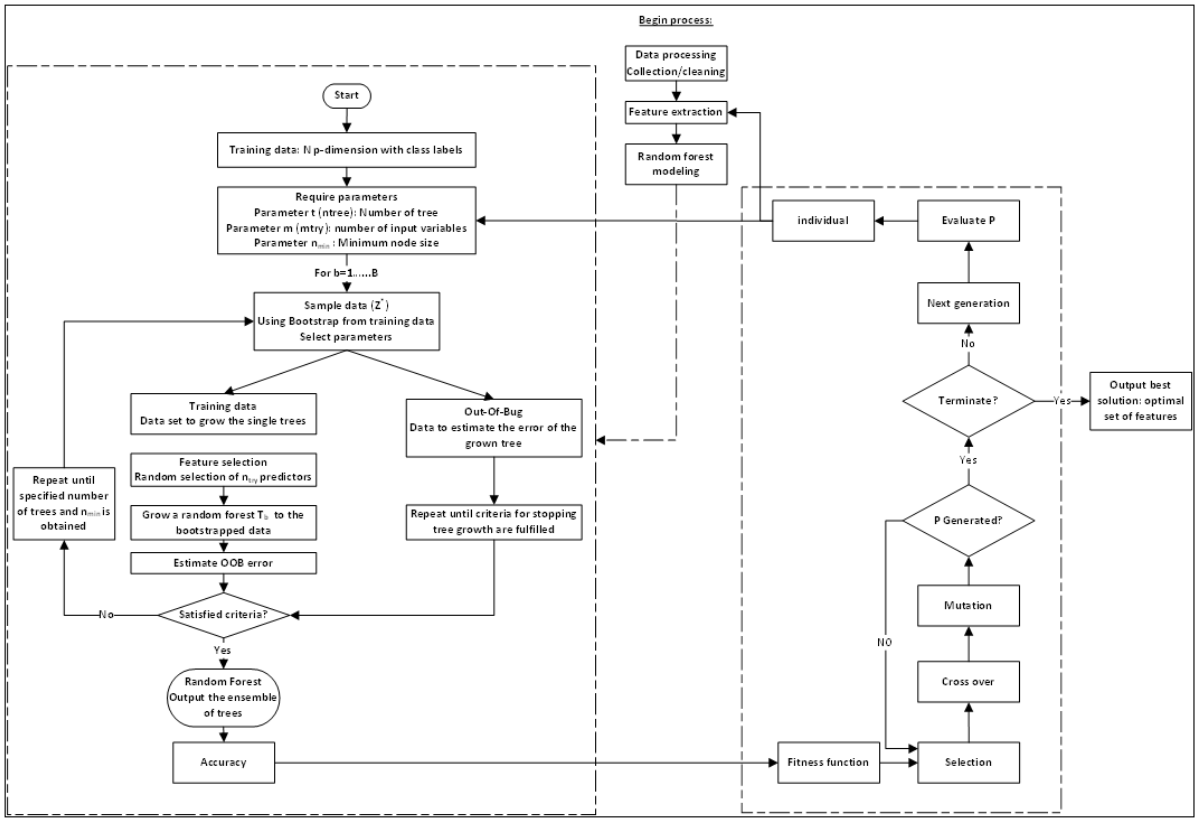

Figure 3.1 A flow chart of the proposed framework

3.4.1 Modeling of random forest

The bootstrap adopted in this research is based on Breiman work. Each tree is constructed using a different bootstrap sample from the data sets, each having different subsets of attributes and parameters. Every final node gives a classification, and the forest chooses the classification by the majority vote. The first step of RF algorithm is the extraction of the data from an available database. Then, the data is extrapolated by choosing the attributes and specifications

required by the model. After that, the data is randomly subdivided into two subsets. The first subset consists of 70% of the data and is used to form the models; this is called the learning set. The second subset consists of 30% of the data, which is used to evaluate and validate the model. Once the data is collected and the features are extracted, a sample data that uses a bootstrap that grows from the training data is presented. Then, the learning set is introduced into the RF to create the predictive model, whose parameters are also estimated. For each terminal node of the tree, these steps are repeated until the specified number of trees is reached and the minimum node size is obtained. Next, the OOB error for the model is estimated. Finally, the output is represented as an ensemble of trees Tb. To make prediction at a new point x, let $\widehat{c_b}(x)$ be the class predicting of the b^{th} random forest tree, then $\widehat{C_{rf}^{b}} = majority\ vote\{\widehat{c_b}(x)\}$. The flow chart of random forest is shown in Figure 3.2.

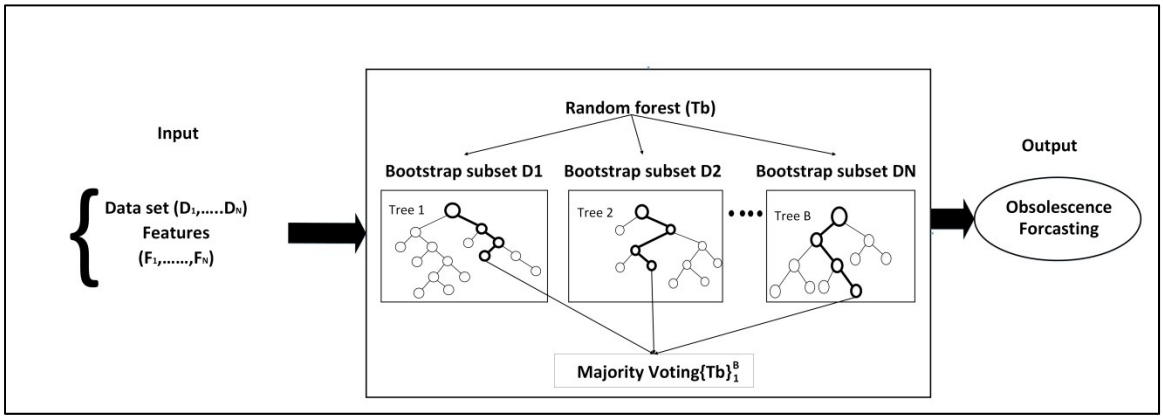

Figure 3.2 A flowchart of Random forest

For the RF model, three parameters are required. At each node, the variables are drawn uniformly and without replacement among all the p explanatory variables (each variable has a probability $1/p$ to be chosen). The number m ($m \leq p$) is fixed at the beginning of the forest construction and is identical for all trees. By default, the number of trees (n_{tree}) is 500, and m is equal to $p^{1/2}$ as suggested by (Breiman, 2001). Here, p presents the total number of variables/features while m is the number of variables in each split.

3.4.2 Modeling of Genetic algorithm

The main objective of applying GA is to enhance the classification performance of random forests for obsolescence forecasting. However, evolutionary algorithms such as genetic algorithm give a promising approach to multi-criteria optimization problems (Yang & Honavar, 1998b). The individual in GA is coded with bit strings. First, the population is initialized. After that, the evaluation of the fitness function is conducted where the classification accuracy of the random forest is used to evaluate the fitness of individuals.

$$Fitness(x) = accuracy\ (x) \tag{3.2}$$

The individual with the highest fitness is selected as the results of the RF algorithm when the evolution is over. Each individual (chromosome) in the population represents a solution to the feature subset problem. Although the proposed genetic algorithm proceeds with the next generation, the operation of genetic includes selection, crossover and mutation. In addition, the crossover operation is done by exchanging the bit strings of a pair of selected individuals at random positions to generate two new offspring, while each bit of the chromosome is treated as a gene. Mutation operation involves replacing one bit with a randomly generated one in a chromosome. This operation avoids premature convergence towards local optima (Figure 3.3).

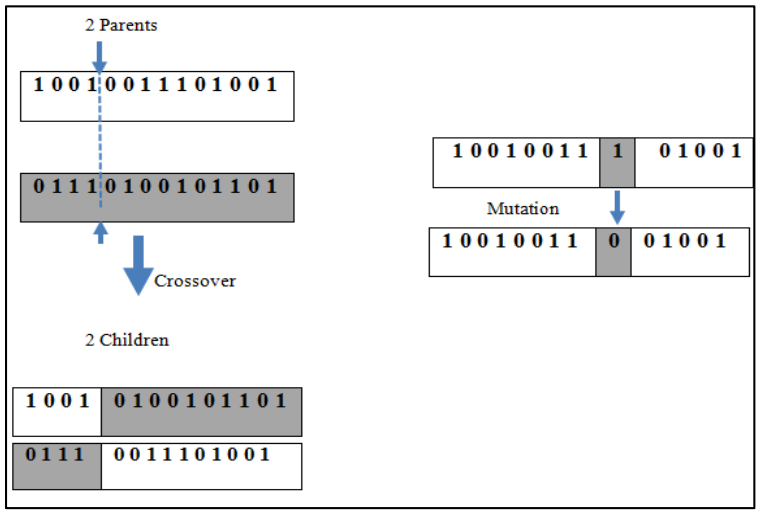

Figure 3.3 Crossover and mutation operations

Let F be the total number of the ensemble features available to choose. Each feature in the GA represents a binary bit (0 or 1), where 1 represents a selected feature while 0 represents a feature that is not selected. Also, let K be the total number of classifiers. Each tree constructed contains two parameters to be optimized, which are m (the number of variables in each split) and n_{min} (the minimum node size). Figure 3.4 presents a sample individual composed of two classifiers and a dataset with 10 features (e.g. for the first tree, the features 2, 7, and 9 are assigned).

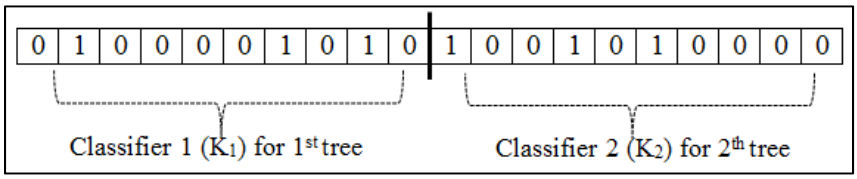

Figure 3.4 An example of selection features

3.4.3 Integration of GA-RF

As previously discussed, the method of getting the decision function (ensemble of trees) uses GA to evolve. In this phase, the decision functions can be evolved using GA following the flowchart shown in Figure 3.1.

The steps of GA-RF algorithm are as follows: First, the input is initialized. For the RF, the number of trees (defined by the users), the training data set, and the feature set are initialized. For the GA, the number of generations, population size, and the probability of crossover, mutation and selection, are initialized. The output presents the random forest model for obsolescence classification. The second step involves grouping the features and the data set by drawing bootstrap samples from the original data. In step 3, a tree is evolved for feature set and RF parameters combined using GA. Thereafter, a population is initialized. After population initialization, fitness evaluation for each individual is carried out. The fitness of the individual is defined as the accuracy of the classification result. For every individual, a classification label can be obtained when a data record comes. When all the training data records are collected, the accuracy of the individual can be obtained by considering all the classification results. After fitness evaluation, GA conducts other genetic operations and produces a new population by selecting the best individual in each generation using the elitism

strategy. Elitism consists to Copy one or more of the best chromosomes in the new generation to not lose the best solution.

3.5 Numerical case study

3.5.1 Data processing

A total of 1000 cellphones characteristics were used in this study, obtained from (Connor et al. 2016). The data provided information such as the brand, date of introduction, and the status (production/end-of-life) of different cellphones. Also, the data contained many technical specifications such as the screen size, GPS, and keyboard. While some data used in this paper are categorical variables (yes or no), others are continuous. These predictive variables (13 features) present the RF input. The target variable (dependent variable) is the obsolescence risk. About 70% of the data (700 cellphones) is used for training while 30% (300 cellphones) is used for testing. To obtain a better estimation of classification accuracy, a 5-fold cross-validation method was applied.

3.5.2 Experimental results

The R programming is used to develop the model (the code is available on request from the main author). The characteristic of GA and RF (based on R output) are as follows:
Maximum generations: 20, Population per generation: 50, Crossover probability: 0.8, Mutation probability: 0.1, Elitism: 0, Number of trees: 500.
The most frequently used metrics to describe these types of results are accuracy and error rate. The results are described in terms of accuracy, Error rate, sensitivity, specificity, and Cohen's KAPPA given by the following equations (C. Jennings, Wu, & Terpenny, 2017)

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.3}$$

$$Error\ rate = 1 - AC \tag{3.4}$$

$$SE = \frac{TP}{TP + FN} \tag{3.5}$$

$$SP = \frac{TN}{TN + FP} \tag{3.6}$$

$$K = \frac{\Pr(a) - \Pr(e)}{1 - Pr(e)} \tag{3.7}$$

TP, TN, FP, and *FN* were defined as true positive, true negative, false positive and false negative respectively. *Pr(a)* presents the probability of success of classification (accuracy) and *Pr(e)* presents the probability of success due to chance.

The accuracy of GA-RF was initially present by a confusion matrix (Table 3.1) for training and testing set.

Tableau 3.1 Confusion matrix of GA-RF algorithm

		Predict	
GA-RF	**Actual**	Available	Discontinued
Training	Available	**404**	6
	Obsolete	4	**286**
Testing	Available	**166**	12
	Obsolete	8	**113**

To validate the results, the GA-RF algorithm is benchmarked with the RF algorithm, Stochastic Gradient Boosting, as well as the Stepwise GLM (stepwise logistic regression) model, which selected the best logistic regression using the stepwise method and AIC as the criterion for selecting the best model. Parameters of the GA-RF and the RF algorithms were optimized on the training sample using 5-fold cross validation. The same 5 folds were used in both procedures for consistency. The comparisons of the algorithms are shown in Table 3.2 and 3.3.

Tableau 3.2 The accuracy measures comparison (testing sample)

Accuracy measure	RF	GA-RF	Stepwise GLM	Stochastic Gradient Boosting
Accuracy	0.913	**0.933**	0.906	0.923
No Information Rate	0.582	0.582	0.582	0.582
Kappa	0.821	**0.862**	0.807	0.841
Sensitivity	0.880	**0.904**	0.880	0.896
Specificity	0.937	**0.954**	0.925	0.943
Error rate	0.087	**0.067**	0.094	0.077
Balanced Accuracy	0.908	**0.929**	0.903	0.919

Tableau 3.3 The accuracy measures comparison (training sample)

Accuracy measure	RF	GA-RF	Stepwise GLM	Stochastic Gradient Boosting
Accuracy	0.947	**0.986**	0.914	0.936
No Information Rate	0.583	0.583	0.583	0.583
Kappa	0.891	**0.971**	0.824	0.867
Sensitivity	0.914	**0.980**	0.901	0.904
Specificity	0.971	**0.990**	0.924	0.958
Error rate	0.053	**0.014**	0.086	0.064
Balanced Accuracy	0.943	**0.985**	0.912	0.931

The accuracy of all algorithms was significantly higher than no information rate, which equals to 0.582 (proportion of discontinued products in the validation sample). Random Forest with feature selection based on the genetic algorithm outperformed stepwise logistic regression, the random forest algorithm and gradient boosting, the performance of which was already outstanding. All commonly used accuracy measures are higher for the GA-RF approach. The error rate was 6.7% for GA-RF, 8.7% for RF, 9.4% for GLM, while the GBM error was 7.7%.

Figure 3.5 presents a chart comparing the training and testing of each model. For the training set, GA-RF outperforms RF by 3.9%, GLM by 7.2% and GBM by 5%. For the testing set, GA-RF also outperforms GA, GLM and GBM by 2%, 2.7% and 1% respectively.

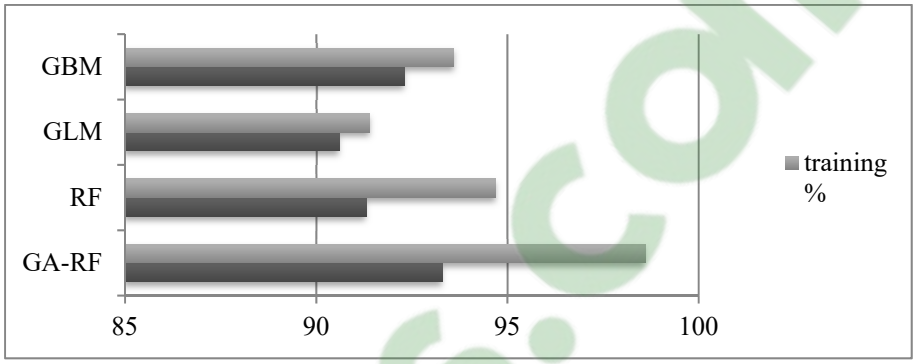

Figure 3.5 Accuracy of algorithms

Figure 3.6 presents a ROC curve. The curve compares the false positive rate to the true positive rate. We can see from the above curve that the proposed method presented the highest accuracy of classification with a small difference compared to the other models. However, the accuracy, specificity and sensitivity in the table 1 and 2 are more impressive.

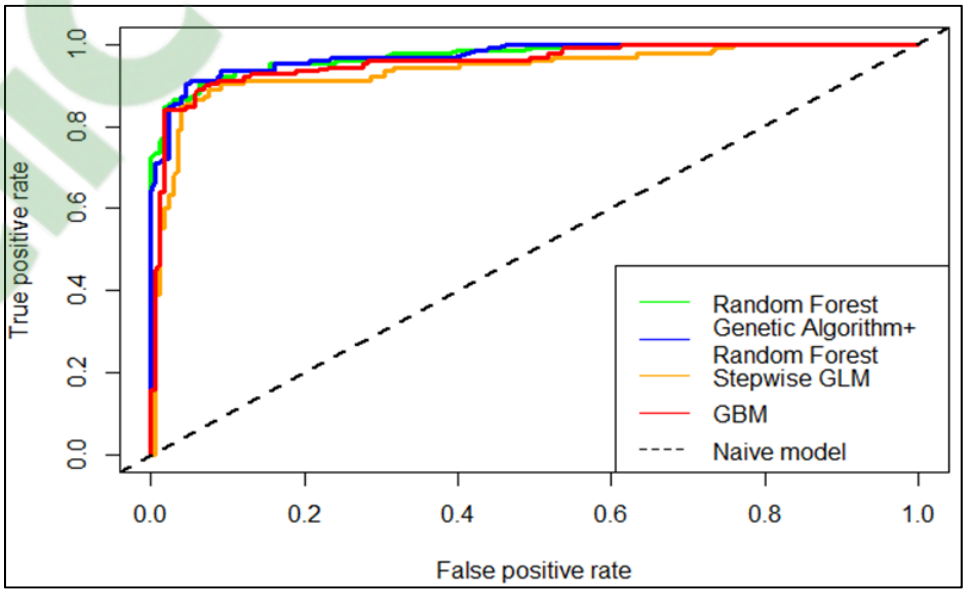

Figure 3.6 Receiver Operator Characteristic curve

3.6 Conclusion

Research on obsolescence is growing due to the serious challenges that complex systems face. Basically, obsolescence problems are often sudden and not planned, causing delays and extra cost. This research presents a method that can be used to estimate the risk of obsolescence using machine learning with a high degree of accuracy. In fact, random forest method is one of a kind machine learning method and appears as the best predictor for forecasting obsolescence risk.

This paper presents a novel approach called GA-RF algorithms, which combines a random forest and genetic algorithm for predicting obsolescence. GA is used to improve the classification of random forest by select an optimal set of features and parameters optimization for classification. The optimal set of features was adopted for the training and testing of the model to obtain the optimal outcomes in classification. In order to validate this approach, GA-RF was compared to random forest algorithm, Stepwise logistic regression and stochastic gradient boosting. GA-RF was found to be more efficient with a high degree of accuracy by 93.3% of accuracy, 90.4% of sensitivity and 95.4% of specificity after the five-fold cross-validation. Therefore, the GA-RF with feature selection subset has better performance than the other methods.

Furthermore, this model can be helpful for manufacturers and researchers that interest to apply evolutionary algorithm in predictive machine learning models. However, the exploitation of this model can be carried out on new data.

For future work, a larger data can give more accurate results. Other optimization algorithms can also be used, such as Particle Swarm Optimization (PSO), which is widely used for optimization, characterized by its robustness and efficiency. In fact, no works have been done before using this approach. Therefore, the researcher intends to include it in the future work by comparing the performance of PSO-RF and GA-RF.

CHAPITRE 4

INTEGRATION OF UNSUPERVISED AND SUPERVISED MACHINE LEARNING FOR OBSOLESCENCE RISK ASSESSMENT

Yosra Grichi, Yvan Beauregard, Thien-My Dao

Département de génie mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

Article soumis à la revue « IEEE Transactions on Components, Packaging and Manufacturing Technology », Juin 2020

4.1 Abstract

Component obsolescence has become an important issue that cannot be ignored. Today, thousands of electronic components are becoming obsolete driven by many factors. These factors include but are not limited to competitive market pressure, technological advancement. Companies face significant challenges due to various obsolescence factors involved when they use systems with a long-life cycle. Obsolescence risk forecasting has gained attention in enterprises to sustain market and ensure good services. This study presents an approach for obsolescence risk assessment where supervised and unsupervised learning techniques are combined. The task is distributed into two levels. The first level is defined as a classification problem tackled with a supervised learning approach to forecast obsolescence risk. In the second level, unsupervised machine learning technique is proposed to partition data into different classes (clusters) that describe risk level based on the predictive model obtained in the first level. The supervised learning model is used in unsupervised learning model to obtain predicted probabilities for each point predicted as 1 (non-obsolete). The study also compares and evaluates the quality of k-means with Partitioning Around Medoids method. The experimental results show that PAM provides better performance using Dunn index.

Keywords:

Obsolescence, risk assessment, unsupervised machine learning, k-means clustering

4.2 Introduction

New technologies have revolutionized the electronics industry due to the rapid advancement of electronic components. This evolution creates new electronic components with a short life cycle. Obsolescence generally occurs when the technology is either old or no longer produced by the supplier, which causes huge delays and costs. Obsolescence emerges in systems that have a longer life cycle than their components, such as automotive, avionics, military, etc. This phenomenon appeared in the aeronautical sector in the 1990s. Prior to this, long-life systems did not face this problem because their components were customized for clients.

The advancement in the electronics industry brings new challenges in the market. This growth of the electronics industry has spurred dramatic changes in the electronic parts with new technologies being introduced in the market with increasing rate (Gravier & Swartz, 2009). Today, the short life cycle of technology and the lack of forecasting represent a challenge for several companies.

Currently, most companies do not have methods to effectively predict obsolescence risk, and therefore they are forced to rely on reactive strategies. Unfortunately, reactive strategies are often more expensive than proactive strategies. For that reason, obsolescence risk forecasting appears as an essential step in the proactive level to manage obsolescence effectively.
A series of studies have primarily focused on the development of obsolescence risk assessment techniques (C. L. Josias, 2009; F. R. Rojo et al., 2012). The purpose of the studies is to identify the critical parts that need to be assessed and to put a key priority in order to choose the optimal mitigation strategies. However, these approaches are based on manual procedures.
With the development of technology, thousands of parts are becoming obsolete every day. Hence, classical approaches are not fruitful to forecast a large number of parts.

Integration of machine learning (ML) for part obsolescence has become a hotspot. Indeed, classification performance is becoming more and more essential to forecast obsolescence risk of thousands of parts.

This study aims to support obsolescence risk assessment by integrating unsupervised ML techniques with supervised ML classifiers. The research is designed to develop predictive models based on supervised learning. In addition to this, this model is further used by unsupervised ML models that can group together items into different classes of risk levels. Clusters are identified based on the correlation and similarity between features. The risk level category is identified based on the average probability P(Y=1|x) of the classification model for data predicted as 'active' (1) in the first phase. K-means clustering algorithm is used in this study. In fact, k-means is one of the ten top algorithms in data mining (X. Wu et al., 2008). In addition, k-means is one of the best algorithms that is commonly used in the clustering process.

The contribution of this study is to introduce an effective method for obsolescence forecasting based on integration of both supervised and unsupervised ML. This approach will help to predict the risk level of upcoming data prior to taking a decision. In order to test the quality of clusters, this work provides a comparison between k-means with PAM algorithm.

4.3 Literature review

Obsolescence can have various impacts on industries. These impacts are caused by functional, technological or logistical obsolescence. Several obsolescence factors were identified in the literature which can be classified as technical and market factors (Gravier & Swartz, 2009; C. L. Josias, 2009; F. R. Rojo et al., 2012).

The effective obsolescence management requires three mitigation strategies. These strategies can be defined as reactive, proactive, and strategic (Bartels, Ermel, Sandborn, et al., 2012). Reactive management involves immediate decision. Possible solutions in this level include existing stock which represents the least expensive approach (F. R. Rojo et al., 2010). It is a stock held by the supply chain that contains the original spare components. The existing stock should be sufficient to ensure supply throughout the remaining life cycle of the system. Last Time Buy is another common approach frequently used by companies when a supplier informs the end of life of a product. Therefore, the manufacturer decides to purchase a good enough quantity to support production until redesign. In addition to this, another well known approach

is described as an alternate part, consisting of replacing the part with another one with the same performance, dimensions and mechanical characteristics.

Proactive management approach is based on the proactive monitoring of life cycle information on parts in order to prevent the obsolescence risk. The paper focuses on obsolescence forecasting, which is part of proactive management.

Finally, strategic management represents long-term strategic planning for part obsolescence in order to satisfy the system's requirements. In other words, it is a combination of the proactive and reactive approaches to minimize the cost of the life cycle. Part obsolescence forecasting can be the input for the strategic management level.

4.3.1 Obsolescence risk forecasting

Obsolescence forecasting can be distributed into two groups, obsolescence risk forecasting and life cycle forecasting. Life cycle forecasting uses regression to predict a numeric value of when the part will become obsolete. However, obsolescence risk forecasting is used to estimate the probability that a part will become obsolete based on a scale risk.

Two main approaches have been adopted and widely discussed in the literature which use key factors to identify the risk level (C. L. Josias, 2009; F. R. Rojo et al., 2012). Rojo developed the first approach to predict the obsolescence risk based on obsolescence indicators which are: year of end of life, number of sources available, consumption rate and stock availability. Similarly, Josias developed the second approach to create a risk index to measure the obsolescence risk through a risk scale level. Four obsolescence factors ($\alpha, \beta, \gamma\ et\ \delta$) were identified to calculate the risk index. These parameters are measured on a risk scale from zero to three. To calculate the risk index, a weight is associated with each parameter.

Jaarsveld has also introduced another method using Markov chain to estimate the risk of obsolescence (van Jaarsveld & Dekker, 2011a). Jaarsveld proposed a Markov chain model in which he used two transition states x_1 and x_0, where x_1 showed high demand, and x_0 presented the absorbing state in which the demand is dead. The transition from one state to another is based on the variation in demand followed by a compound poisson process. (Zolghadri et al.,

2018) introduced Bayes' theory to predict the probability of obsolescence based on two states: obsolete and non-obsolete to assess its impact on the system.

4.3.2 Machine learning

Machine learning has attracted the attention of many researchers in various disciplines and has been applied in many areas. Machine learning problems can be classified as either supervised or unsupervised learning. Supervised learning approach is used when the labels (the output variable is known) are used to train the algorithm. Unlikely supervised machine learning, unsupervised ML algorithms are applied on unlabeled data.

Recently, supervised machine learning has been applied for obsolescence forecasting. (Yosra Grichi, Yvan Beauregard, et al., 2018; Y Grichi et al., 2017; Yosra Grichi, Dao, & Beauregard, 2018; C. Jennings et al., 2016a) have developed an approach based on comparison of different ML algorithms such as artificial neural network (ANN), AdaBoost, support vector machine (SVM), random forest (RF) and decision tree to predict obsolescence risk. The problem is solved as a classification problem (obsolete or active). On the other hand, the obsolescence life cycle was applied to predict the obsolescence date based on regression approach.

Yet, unsupervised ML has not been applied for obsolescence forecasting. Unsupervised ML is referring to clustering algorithms and consists of having only input data (not labeled data). Unsupervised learning includes two categories of algorithms: clustering and association algorithms. Clustering involves separating or dividing a dataset into several groups, so that datasets belonging to the same groups come together and define a particular cluster (group). However, association consists in discovering relationships between variables in large databases. Unsupervised ML algorithms use data mining techniques to cluster instead of giving predictions directly. Therefore, these unsupervised ML algorithms are often used as complementary tools to supervised ones. In this work, we propose an unsupervised clustering using k-means algorithm.

K-means algorithm is an unsupervised learning method to partition a dataset into k-clusters, in which the variation within a class is low and between class is high. This is achieved by

clustering the points that are close to each other in terms of some predetermined metric/distance function (Asgharbeygi & Maleki, 2008; X. Wu et al., 2008). The k-means has been discovered by several researchers across different disciplines (H. P. Friedman & Rubin, 1967). The basic steps for k-means clustering are: 1) select number of clusters k. If the number of clusters is not known in advance, elbow method can be used which searches for the optimal k cluster. 2) initiate centroids by randomly selecting k points from the data without replacement. 3) assign each data point to its nearest centroid (cluster center) by measuring the distance between the 1st and the three-initial centroid using Euclidean distance. 4) calculate the mean of each cluster and replace the original center with the new position center for each cluster. The algorithm converges when the assignments no longer change. Otherwise, the previous steps should be repeated until no data point changes its position.

The algorithm operates on a set of d-dimensional vectors: $D_n = x_i \{x_1, x_2, \dots, x_n\}$ where x_i is the data point allocated to a k class. Each class is represented by a prototype μ_k, and each iteration needs n × k comparisons, which determines the time complexity of one iteration. The number of iterations required for convergence varies depend on $\{n\}$.

4.4 Research methodology

This study emphasized on integrating clustering algorithms with predictive outcomes of supervised ML to forecast the obsolescence risk. The proposed method is presented in Figure 4.1. The process of the methodology is based on two phases: Phase 1 is devoted to forecast obsolescence risk using classification algorithms (obsolete/ non obsolete). Phase 2 is dedicated to integrating k-means clustering with information coming from the results of supervised learning (phase 1). The development process is conducted in several steps. First, individual supervised ML algorithms such as ANN, SVM, RF decision tree, etc. are developed and compared. Additionally, features reduction technique is applied using meta-heuristic optimization approach to improve the classification accuracy and to accelerate the clustering process. The methodology process detail of phase 1 is given through this reference (Yosra Grichi, Thien-My Dao, et al., 2018). Once the predictive model is ready, the next phase is dedicated to the clustering. The input of the clustering model contains three elements: the

initialization of the data processing (x), number of clusters and the predictive model (f) obtained from phase 1, given by the following function:

$$Cluster = f(data(x), model, k = 3) \quad (4.1)$$

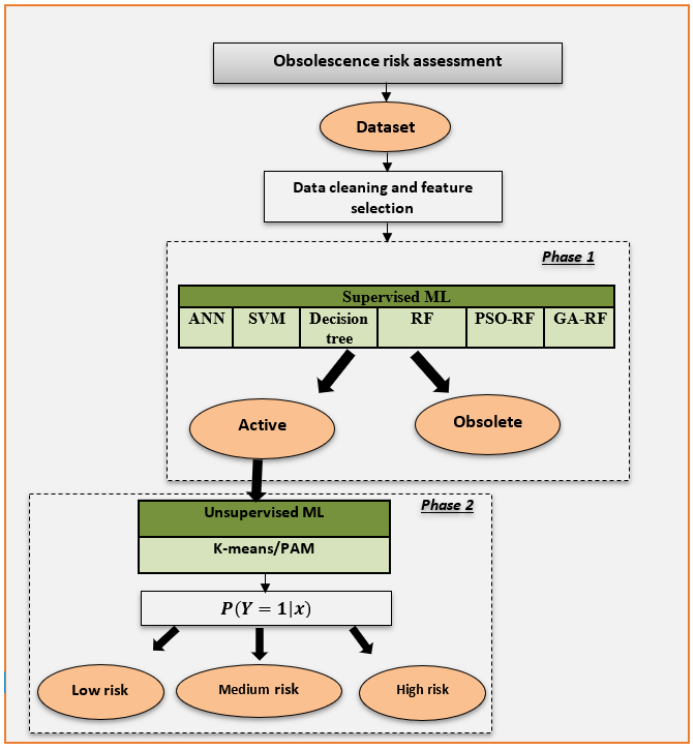

Figure 4.1 Research flow

The methodology process of phase 2 is as follows: **1-** specify the number of clusters. In fact, there are several methods that can estimate the optimal number of clusters, the most known approach is the elbow method. In this study, a comparison between different methods using the majority vote is developed. Elbow method and Silhouette width are adopted, which they can be applied for mixed data. **2-** get predictions from the phase 1 model (f) by choosing the best predictive model. The model from the phase 1 is used to get the predicted probabilities $P(Y = 1|\mathrm{x})$ for each data point predicted as active. **3-** present the development of k-means clustering based on subset (x) for the cases where predicted output is 1. K-means is initially chosen. Then, it has been found that it can be applied only for numeric variables. As the dataset contain mixed data, k-prototype is adopted. Cluster centers of k-prototype are represented by

mean values for numeric variables and mode values for categorical variables. The reason of choosing k-prototype given its development is simple, efficient and easy to interpret the results. K-prototype works in a similar way as k-means but with different distance function. The k-prototype algorithm integrates k-means and k-modes algorithms to cluster objects described by numeric and categorical variables (Z. Huang, 1998). The distance function of k-prototype is given by Euclidean and hamming distance, presented as follows:

$$d(\tilde{x}, \tilde{y}) = \sum_{i=1}^{m} (x_i - y_i)^2 + \lambda \sum_{i=m+1}^{n} \emptyset(x_i, y_i) \quad (4.2)$$

Where, consider the first m variables to be numeric variables, and the remaining (n-m) to be categorical variables. $\emptyset(x_i, y_i)$=0 if and only if $x_i = y_i$, else $\emptyset(x_i, y_i)$=1.

Clusters are formed using k-prototype. In fact, k-prototype split the dataset into groups based on general spread of data over variables. Given that k-prototype only separate the dataset into multiple groups, the predictive model helps to define the groups as low, medium, or high risk. Indeed, this model reacts to data which is classified as non-obsolete. In this study, we attempt to put the emphasis on how probable for a cluster to fall in the case of Y=1. A probability calculation is developed for each component predicted as in production (active); this probability gives between two data points the more that has a high probability to be active (which means equal to 1). The partition in groups is controlled using a risk index developed based on likelihood of obsolescence occurring. In fact, the risk is controlled as following: if the risk probability for a specific cluster is lower or equal to 0.5, therefore, the component has a high probability of becoming obsolete. Hence, this cluster is categorized as high. If the risk probability of cluster is between 0.5 and 0.8, the cluster is categorized as medium. However, if the risk probability of a cluster is greater than 0.8 (the value is close to 1), the risk is low. Based on these probabilities, the clusters can be defined.

Let M_1, M_2, M_3 denotes the index of the points that belong to the three clusters. The group-name of cluster will be identified based on the following average probability for data belong to 1:

$$P(Y = 1|\mathrm{x} \in M_i) = \frac{1}{m_i} \sum_{\mathrm{i} \in M_i} P(Y = 1|\mathrm{x} \in x_i); \mathrm{i}=1,2,3 \quad (4.3)$$

Based on the average probability, the clusters can be ranked into high, medium or low. Finally, return the output of final clusters. Once the cluster model is finished, a new data point can be assigned to the cluster with the closest centroid. Similar features for each group will be identified which affect the obsolescence risk based on the average probability.

4.5 Experimental results

Two datasets are used in this work. The first dataset provides information about cell phones collected from the online forum known as GSM Arena (https://www.gsmarena.com). Thirteen attributes are used as predictive variables. The second dataset is assembled from micron technology website (https://www.micron.com). The data contains more than 700 components with a known label, including technical specifications. The detailed information of datasets is presented in Table 4.1.

Tableau 4.1 Description of datasets in this work

Dataset	Total	Obsolete/Active	Features
Cellphones	1000	417/583	13
Memory technology	734	161/573	11

Given that the information was extracted from different sites, the data may have lack of information and missing value. Consequently, cleaning and formatting data is done to make accurate predictions. In addition to this, feature optimization is applied in the phase 1 and integrated with random forest algorithm to improve the results. The experimental results for

the phase 1 are presented from this reference (Yosra Grichi, Thien-My Dao, et al., 2018). Experimental results show that PSO-RF outperformed GA-RF with 96% of accuracy.

R studio programming is adopted to develop k-prototype cluster. K-prototype algorithm is chosen because the datasets contain mixed data. The input of the cluster contains initial data, PSO-RF model and the number of clusters. The number of clusters is an essential parameter in partitioning clustering that should be specified in advance by the user. This considered as a weak point in clustering since the optimal number of clusters is subjective which depends on the method used to measure the similarities between clusters and the parameters used for partitioning. To determine the optimal number of clusters, three different methods are adopted and widely used in partitioning clustering (Ahmad & Khan, 2019), which will be applied in order to decide the best number of clusters using the majority vote. These methods are, Elbow method, Silhouette method and Gap statistic method. Elbow method helps to determine the optimal value of clusters that have the smallest value of inertia (sum of squared distances of samples to their closest cluster center). Silhouette method is used to calculate the average silhouette width of different k clusters. Since a high average silhouette width indicate a good clustering, then the optimal number of k is the one that maximize that average (Rousseeuw & Kaufman, 1990). Gap statistic is another approach that can be used in order to determine the number of clusters. In fact, this method compares the variation in intra-cluster. During the simulation process, it was found that Gap statistic cannot be applied to categorical variables. In fact, during finding the Gap statistics, there is a simulation involved drawing sample from uniform distribution between the range of each of the x variables which can be done only for numeric variables. For this reason, a comparison is made between within-sum of square and silhouette width for k-prototype approach. According to the observations in Figure 4.2, it can be concluded that the optimal number of clusters is 3 based on k-prototype clustering algorithm. The optimal number of clusters is chosen from the point in which the variance is no longer reduced significantly. Based on the elbow curve, the variance curve between k=1 to 3 is more significantly than the variance from k=4. Therefore, the optimal value of k is 3.

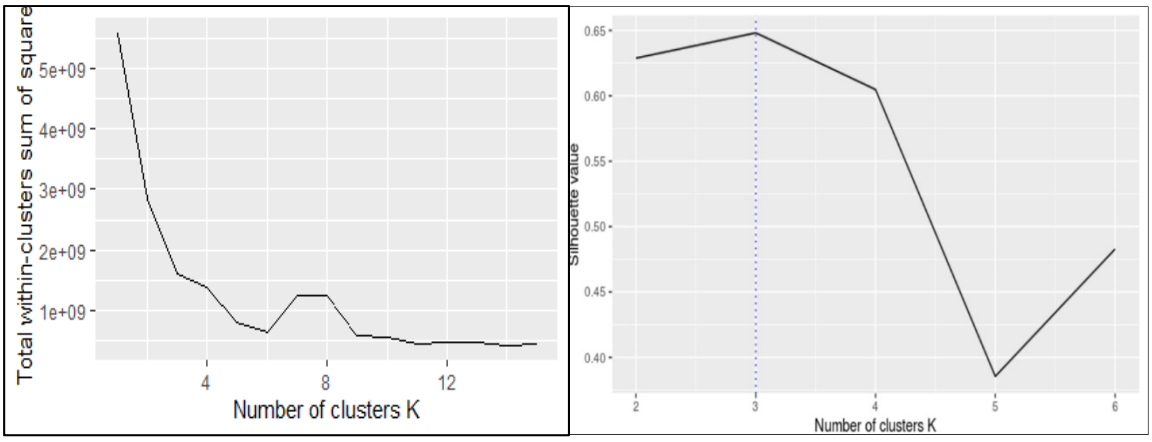

Figure 4.2 Elbow and Silhouette plot to determine the optimal value of k

The results of k-prototype approach within clusters centroids are given by table 4.4 and 4.5. Table 4.2 and Table 4.3 present the similarity between data points within a cluster and what differs from other clusters.

Tableau 4.2 Datapoint similarity within each cluster centroid for cells dataset

Cluster	Headphone	MP3	Monophonic	Wav	GPS	Email	Resolution	Screen size	Weight
1	Yes	Yes	No	Yes	Yes	Yes	3054.7	4.1	130.5
2	Yes	Yes	No	No	No	No	1131.8	2.4	437.1
3	Yes	Yes	No	Yes	Yes	Yes	620.9	6.3	618.5

Tableau 4.3 Datapoint similarity within each cluster centroid for memory dataset

Cluster	density	depth	width	voltage	pin.count	clock rate	Cycle time	Op.temp
1	4000	752	8	1.233	84	1044.666	0.938	0C to +95C
2	555	61.125	11.25	2.737	59.625	209.25	5.8125	-40C to +85C
3	8000	2000	4	1.2	78	1067	0.938	0C to +95C

Based on the average probability $P(Y = 1|\mathrm{x})$ of the classification model, the results are given by table 4.4.

Tableau 4.4 Clusters names based on average probability using k-prototype (cell phones dataset)

Cluster	Average probability	Group name	# of observations
2	0.50	Group_1	160
1	0.769	Group_2	251
3	0.97	Group_3	585

Tableau 4.5 Clusters names based on average probability using k-prototype (memory technology dataset)

Cluster	Average probability	Group name	# of observations
1	0.48	Group_1	43
2	0.787	Group_2	112
3	0.94	Group_3	235

As previously discussed, the risk probability lower or equal to 0.5 indicate a high probability of becoming obsolete. Therefore, Group_1 will be assumed as a high risk in which the group of parts should receive a particular attention. Mitigation strategies decisions should be set up to manage them proactively in order to reduce the probability to be obsoletes.

4.5.1 Comparison between k-prototype and PAM

To validate the performance of the clustering technique, Partitioning Around Medoids (PAM) is chosen to be benchmarked with k-prototype. PAM appears in the literature as an efficient algorithm for mixed data, since it has been compared with different algorithms and has shown a good result. It's simple and fast. The experimental results produced a cluster accuracy compared to other partitioning algorithms (Budiaji & Leisch, 2019).

Clusters are defined based on Gower's distance given by the following equation (Park & Jun, 2009; Van der Laan, Pollard, & Bryan, 2003):

$$d(\tilde{x}, \tilde{y}) = \sum_{i=1}^{m} \frac{|x_i - y_i|}{R_i} + \sum_{i=m+1}^{n} \emptyset(x_i, y_i) \qquad (4.4)$$

Where: consider the first m variables to be numeric variables and (n-m) to be categorical variables. R_i denotes the range of i[th] feature. $\emptyset(x_i, y_i)$=0 if and only if $x_i = y_i$ else $\emptyset(x_i, y_i)$=1
Based on Gower's distance, the cost function of PAM is given as follows:

$$c = \sum_{c_i} \sum_{p_i \in c_i} d(p_i, \bar{c}_i) \tag{4.5}$$

The steps start by initializing k random points of n data points as medoids. Then, associate each data point to the closest medoid by using Gower distance method for mixed data. For each data point (n) which is not a medoid m, swap m and n and associate each data point to the closest medoid, then recalculate the cost until the cost decreases. If the total cost is higher than the previous step, undo the swap.

4.5.2 Dunn index calculation

Quality of clustering is an important stage in cluster evaluation. Ideally, clusters should minimize intra-cluster variation and maximize the distance between clusters (Handl, Knowles, & Kell, 2005). After analysis and comparison between different index calculation in order to choose the best suitable index, Dunn index is chosen in this work as it can provide a good cluster from nonlinear combinations of compactness and separation. In fact, it was found that Davies-Bouldin index uses Minkowski distance, which is not applicable for mixed data.
Dunn index is given by the following equation (Ansari, Azeem, Ahmed, & Babu, 2015):

$$Dunn = \frac{\min\limits_{1 \le i \le j \le q} d(C_i, C_j)}{\max\limits_{1 \le k \le q} diam(C_k)} \tag{4.6}$$

The dis-similarity between two clusters C_i and C_j is given by:

$$d(C_i, C_j) = \min_{x \in C_{i, y \in C_j}} d(x, y) \tag{4.7}$$

The diameter of a cluster is given by:

$$diam(C_k) = \max_{x,y \in C_k} d(x,y) \tag{4.8}$$

Figure 4.3 illustrates the performance of k-prototype and PAM based on both datasets. Horizontal axis shows the optimal number of clusters compared to the vertical axis which represents Dunn index value.

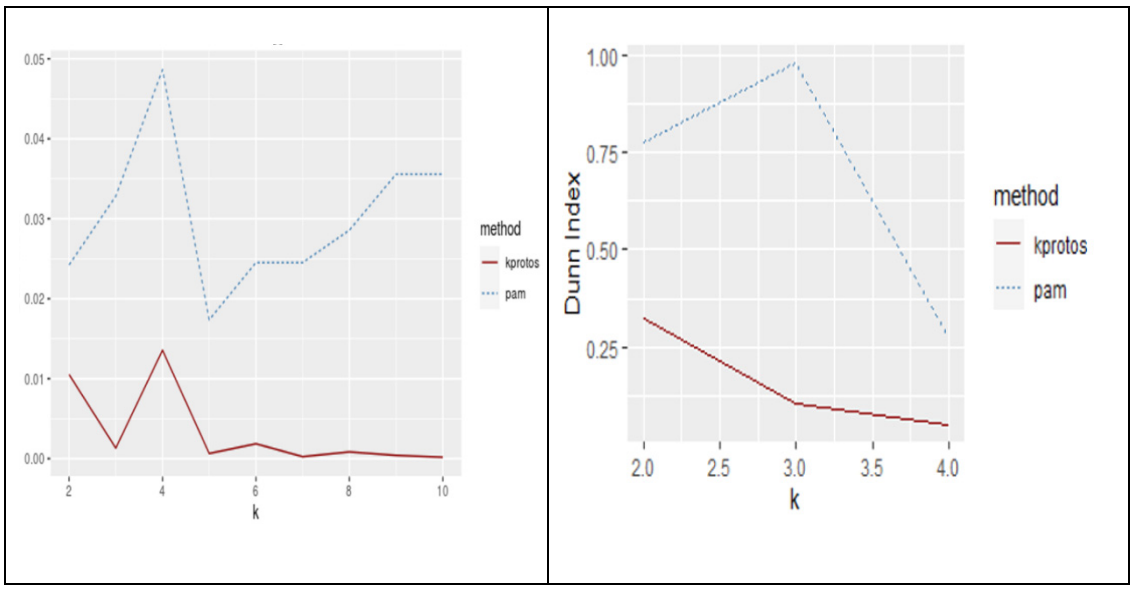

Figure 4.3 Comparison between k-prototype and PAM based on Dunn index

PAM shows the highest Dunn index for all numbers of clusters. This indicates that PAM is significantly better than k-prototype and can produce more compact clusters that are well separated from each other.

4.6 Discussion

In the previous sections, an integration method of two machine learning was applied using two datasets. The first part was interested on applying supervised learning classification. Then, a clustering process was applied to classify a set of components which are still in production,

into different groups of risk level based on risk probability. During the clustering process, two methods were applied and compared in order to choose the optimal value of k clusters. These methods are: Elbow method and Silhouette width. Both methods give the same number of clusters. Afterwards, k-prototype algorithm was adopted to develop the clustering model and compared to PAM using Dunn index to evaluate the clustering quality. Experimental result shows that PAM is better than k-prototype in term of clustering quality.

Although the proposed method shows a good potential in obsolescence risk assessment, there are, however, certain limits which must be taken into consideration for future work. The proposed approach of clustering was mainly based on the probability to rank the clusters into different risk levels based on the results of the supervised learning model. In fact, other criteria can be considered, for example, a risk index can be created based on obsolescence factors and correlations between them. Similar test to CL Josias (2009) can be invested in this regard. The risk index may help to improve the validity of clustering. Additionally, the proposed method can be tested using other available datasets of electronic components. Furthermore, the cluster quality can be evaluated with other methods to validate the results. After experimental tests, it was found that k-prototype has some limitations in terms of hamming distance. In fact, hamming distance is a good representative for categorical values which can be represented as 0 or 1. However, a clustering errors may occur for other categorical variables. It would be interested to test other partitioning algorithms such as the extended algorithm w-k-prototype developed by Huang et al. (J. Z. Huang, Ng, Rong, & Li, 2005), and compared to the k-prototype suggested in this study (Ahmad & Khan, 2019).

4.7 Conclusion

The rapid growth of electronics industry has produced components having a short life cycle. Obsolescence risk forecasting become a crucial activity in proactive management in order to support activities in sectors that sustain long life systems. To manage thousands of potentially obsolete components effectively, machine learning technique would be the best approach to assess obsolescence risk. The study investigated a technique in which two different types of

traditional machine learning algorithms (supervised and unsupervised learning) are concatenated. Supervised ML is used to calculate obsolescence risk based on classification algorithms. On the other hand, unsupervised ML is applied to classify data into a scale of risk level based on probability. To categorize the data into three different risk levels clusters based on information derived from a predictive model of the supervised ML, k-prototype clustering algorithm is proposed. Two different datasets are used in the experimental setup to test the hypothesis. To evaluate the best clustering approach in obsolescence risk assessment, a benchmark PAM approach is used. Finally, experimental results show that PAM outperformed k-prototype in terms of Dunn's validity index. The study also suggested that obsolescence risk assessment based on clustering methods could be used to help choose the best mitigation strategies approaches.

CHAPITRE 5

AN APPROACH TO OBSOLESCENCE FORECASTING BASED ON HIDDEN MARKOV MODEL AND COMPOUND POISSON PROCESS

Yosra Grichi, Yvan Beauregard, Thien-My Dao

Département de génie mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

Article publié dans « International Journal of Industrial Engineering and Operations Management (IJIEOM) », Décembre 2019

5.1 Abstract

The popularity of electronic devices has sparked research to implement components that can achieve better performance and scalability. However, companies face significant challenges when they use systems with a long-life cycle, such as in avionics, which leads to obsolescence problems. Obsolescence can be driven by many factors, primary among which could be the rapid development of technologies that lead to a short life cycle of parts. Moreover, obsolescence problems can prove costly in terms of intermittent stock availability and unmet demand. Therefore, obsolescence forecasting appears to be one of the most efficient solutions. This paper presents a review of gaps in the actual approaches and proposes a method that can better forecast the product life cycle. The proposed approach will help companies to improve obsolescence forecasting and reduce its impact in the supply chain.

The method introduces a stochastic approach to estimate the obsolescence life cycle through simulation of demand data using Markov chain and homogeneous compound Poisson process. This approach uses multiple states of the life cycle curve based on the change in demand rate and introduces hidden Markov theory to estimate the model parameters. Numerical results are provided to validate the proposed method. To examine the accuracy of this approach, the standard deviation of obsolescence time is calculated. The results showed that the life cycle curves of parts can be predicted with high accuracy.

Keywords

Obsolescence; Life Cycle Forecasting; Markov Chain, Hidden Markov Model; Compound Poisson Process.

5.2 Introduction

Parts obsolescence has long been perceived as a significant challenge in designing and sustaining long-lived systems. This phenomenon is referred to as diminishing manufacturing sources and material shortages (DMSMS), and electronic parts frequently entail serious problems of sustainability of the whole systems, due to their short life cycles. Obsolescence issues arise in systems that have a longer life cycle than their components, such as automotive, avionics, military, etc. These sectors have had to face several challenges on this account.

Obsolescence is a major problem caused by the rapid evolution of technologies. Researchers have defined obsolescence in various ways. (F. R. Rojo et al., 2010) define obsolescence as a part becoming obsolete when the technology used to manufacture it is no longer available, supported or produced by the supplier (Shen & Willems, 2014). In other words, obsolescence can be characterized by the loss of suppliers or raw materials (P. Sandborn, 2013), causing delays and extra cost. The negative effects of technological obsolescence on production performance have been studied in the literature and represent a major challenge in the long term. Technological obsolescence causes problems in the supply chain and in the management of electronic systems (P. Sandborn et al., 2011). In other words, the incessant progress of technology is one of the factors that increase the rates of obsolescence. Electronic components are the parts most affected by technological obsolescence. In fact, the electronics industry has emerged as the fastest growing sector and has spread worldwide. Based on Moore's law, the evolution of technologies mainly comprised of electronic parts will continue to grow rapidly, and it is estimated that in the same way, semiconductor density can double every year (Homchalee & Sessomboon, 2013; P. Sandborn, 2008). Consequently, many of the electronic parts that constitute a product have individual life cycles which are significantly shorter than the life cycle of the product.

Growth of the electronics industry has spurred dramatic changes in the electronic parts, with new technologies being introduced in the market at an increasing rate. Today, short technological life cycles and the lack of forecasting represent a challenge for several companies, especially for the electronics industry which represents one of the most dynamic sectors of the world economy (Solomon et al., 2000). From this perspective, obsolescence forecasting appears to be one of the most efficient solutions to managing obsolescence, as it assists manufacturers in identifying obsolete parts. Through obsolescence forecasting, companies can ensure support for parts in service and mitigate any negative impact by identifying parts that are likely to become obsolete. Moreover, obsolescence forecasting enables engineers to more effectively manage the introduction and ongoing use of long-field-life products based on the projected life cycle of the parts. Obsolescence prediction methodology is a critical element within risk-informed parts selection and management processes.

To overcome the problems caused by obsolescence, several studies have been conducted to create models that can effectively forecast obsolescence. Statistical methods such as regression, partial least square regression (PLS), time series and Gaussian method have been previously employed in several works (Gao et al., 2011; Jungmok & Namhun, 2017). However, there is a need to modify these methods to facilitate life cycle forecasting for thousands of components that affect our day-to-day lives. In this context, there are precise and stable forecasting approaches whose performance has been justified in the literature. Moreover, results have shown that these approaches could be used for nonlinear predictions with high accuracy and without human manipulation (Zurada, 1992). In the past few decades, machine learning has attracted the attention of many researchers in various disciplines and has been applied in many areas. According to (X. Wu et al., 2008), among the top ten algorithms identified by the Institute of Electrical and Electronics Engineers (IEEE) are AdaBoost, SVM, K-Means, decision tree, and naïve Bayes. These algorithms have been proven to be good predictors. In addition, applying machine learning in forecasting obsolescence risk and life cycle has received a lot of attention in the last three years (Y Grichi et al., 2017; C. Jennings et al., 2016a). Moreover, Jaarsveld introduced a forecasting obsolescence model (van Jaarsveld

& Dekker, 2011a) that uses Markov chain with only two states, where the transition between states is characterized by a decrease in demand order.

5.2.1 Scope and contribution

In previous work, benchmarking studies have shown an approach to forecast obsolescence risk using Markov chain with a two-state model, with only one state in which the part is moving and in the other state the demand for the part has dropped dead. This last state is assumed to be an absorbing state (van Jaarsveld & Dekker, 2011a). However, this approach is limited and not adequate to forecast the whole life cycle, since it considers only one transition between two states, so that the estimation of obsolescence is not accurate. In general, the life cycle of a product goes through multiple stages as presented in Figure 5.1, which are: introduction–growth–maturity–decline– obsolescence. It is important to consider the whole of the life cycle in order to improve the accuracy of forecasting.

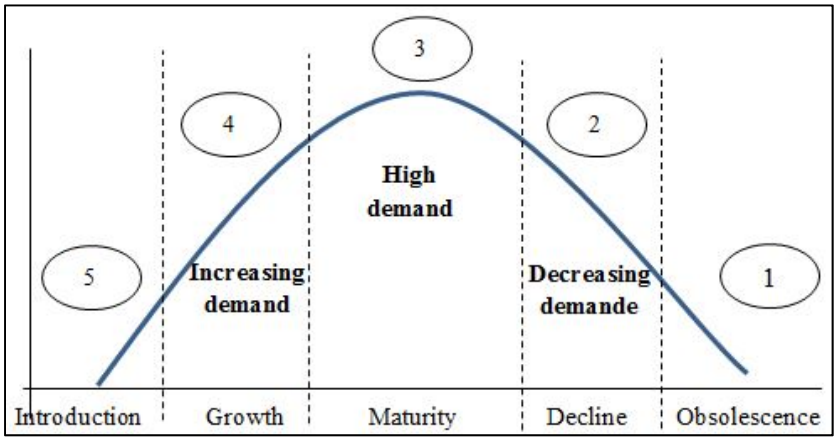

Figure 5.1 Life cycle stages adapted from Solomon et al. (2000)

This paper addresses a new approach to forecasting the life cycle of electronic parts obsolescence. Markov chain performed with compound Poisson process (CPP) is proposed to build a predictive model for the parts' life cycle. The proposed method considers five stages of a part's life cycle to forecast the obsolescence risk time.

In fact, multiple states will greatly complicate the estimation of the model parameters; therefore, the theory of hidden Markov models is used to estimate the model parameters. To the best of our knowledge, methods to estimate the life cycle using Markov chain and compound Poisson process are not available in the literature.

Table 5.1 shows the difference between the proposed method (Markov chain performed with compound Poisson process and homogeneous hidden Markov theory to estimate model parameters) and Jaarsveld's model (simple Markov model with only two states).

Tableau 5.1 Jaarsveld's method vs. proposed method

	Jaarsveld's method	**Proposed method**
Model	Markov chain with two states	Markov chain with multiple states
Advantages	- Method very simple and easy to implement - Estimation of parameters is simple - Simple optimization of model	- Powerful method - Provides for a possibility of sudden changes in the order rate through the states. - The model uses multiples states with a positive demand rate as well as negative demand.
Disadvantages	- The method described in the above model provides for decreases in demand but not increases. In practice, the parts go through several stages of an entire life cycle, with possibility of positive demand rate.	- The estimation of parameters using HMM is quite complicated.

In this paper, a Markov driven homogeneous Poisson process is introduced based on specific features. These key features are as follows:

- There are five Markov states with only forward transition, in which the last one is considered as an absorbing state.
- The transitions between different states occur randomly in time, and the duration in one state follows an exponential distribution.
- In each state, the demand follows a homogeneous Poisson counting process with constant intensity λ_i., i.e., constant rate of arrival of orders.
- Transitions occur when there is a sudden change in the intensity of order arrivals.
- The sequence of observed states s_i can be characterized by the corresponding λs_i, and they are in descending order, $\lambda s_1 > \lambda s_2 > \ldots, > \lambda s_n$

5.3 Potential obsolescence forecasting strategies

The nature of obsolescence forecasting so far has been the management of the problem once it occurs. In practice, most firms do not have effective methods for predicting obsolescence and therefore are forced into over reliance on reactive strategies. The most famous reactive approach used by companies is last time buy (LTB) and existing stock (F. R. Rojo et al., 2010). However, these strategies are only temporary and can fail if the organization runs out of ways to procure the required parts, and it can be very costly.

Obsolescence forecasting can be divided into long-term and short-term methodologies: anything lasting more than one year is named as long-term forecasting, which is represented by proactive and strategic management and life cycle planning to support a system. The second type is short-term forecasting, which is represented by a reactive management, such as reducing inventories to avoid the dormant parts, and forecasting the economic order quantity (EOQ) to cover a specific period.

5.3.1 Obsolescence management

The effective management of obsolescence requires three mitigation strategies that may be defined as: reactive, proactive, and strategic.

Reactive management involves immediately executing solutions when parts become obsolete. Possible solutions in this level include four categories currently adopted according to (F. R. Rojo et al., 2010), which are : 1-Existing stock—this represents the least expensive approach, where the only cost is the functional test of the components. It is a stock held by the supply chain that contains the original spare components. The existing stock should be sufficient to ensure supply throughout the remaining life cycle of the system (Bartels, Ermel, Pecht, & Sandborn, 2012b; Pinçe & Dekker, 2011; F. J. Romero Rojo et al., 2012; F. R. Rojo et al., 2010); 2- Last Time Buy—when a supplier informs the end of life of a product, the manufacturer decides to purchase a quantity (last purchase) sufficient to support production until redesign. It is a short-term strategy frequently used by companies; 3-Alternate Part—consists of replacing the part with another one with the same performance, dimensions and mechanical characteristics; 4-Redesign—consists of redesign of an obsolete component.

Proactive management approach is based on the proactive monitoring of life cycle information on parts in order to prevent the risk of obsolescence leading to production shutdowns and costly design. Obsolescence forecasting, the focus of this paper, is essential to the proactive management level.

Finally, strategic management represents long-term strategic planning for part obsolescence in order to satisfy the system's requirements. In other words, it is a combination of the proactive and reactive approaches, in order to minimize the cost of the life cycle. Part obsolescence forecasting can be the input for the strategic management level (P. Sandborn, 2008).

5.3.2 Obsolescence forecasting

Forecasting part obsolescence plays a key role in proactive and strategic management levels. Obsolescence forecasting can be broken down into two groups. The first is based on obsolescence risk estimation risk and the second method is based on the life cycle forecasting. Obsolescence risk forecasting is used to predict the probability that a component will become obsolete, by creating a scale to indicate the levels of the likelihood of a part becoming obsolete (C. Josias et al., 2004; F. R. Rojo et al., 2012; van Jaarsveld & Dekker, 2011a). In this context, (F. R. Rojo et al., 2012) conducted a Delphi study to analyze the risk of obsolescence. They developed a risk index using some indicators, which are: years to end of life, the number of sources available, and the consumption rate versus the availability of stock. Another approach developed by (C. Josias et al., 2004) aims to create a risk index by measuring the manufacturers' market share, number of sources available for each part, and the company's risk level. Furthermore, the Markov model was used to estimate and manage the obsolescence risk (van Jaarsveld & Dekker, 2011a), wherein the authors used product demand data history to estimate the risk of obsolescence using the Markov chain and Poisson process. However, (Hu & Bidanda, 2009) focus on the decision making through the life cycle of a product using the Markov decision process (MDP) to estimate the obsolescence risk. (Yosra Grichi, Yvan Beauregard, et al., 2018; C. Jennings, Wu, & Terpenny, 2016b) created models to estimate the obsolescence risk using machine learning algorithms as well as a meta-heuristic approach to

improve the classification by random forest algorithm. Finally, (Zolghadri et al., 2018) propose a method to estimate the obsolescence risk using a Bayesian model.

Alternatively, Life cycle forecasting refers to a process that predicts the length of time during which the product will be in production. (Solomon et al., 2000) were the first to introduce the life cycle forecasting method. They conducted a study to predict the obsolescence date from the life cycle curve. During the lifecycle, most products go through five stages that refer to the demand changes, which are named as introduction, growth, maturity, decline and phase out (obsolescence). This period represents the product availability in the market when it can be purchased by customers. Everything before the introduction stage represents the product development period. Another method based on data mining was developed by (P. Sandborn, 2017; P. A. Sandborn et al., 2007), where the obsolescence date was obtained by applying Gaussian method to predict future sales over time. Moreover, other researchers have introduced regression analysis and time series method to predict the date of obsolescence (Gao et al., 2011; Jungmok & Namhun, 2017). Croston method and neural network approach were used to forecast intermittent demand in order to minimize the impact of obsolete and dormant stock (Babai, Dallery, Boubaker, & Kalai, 2019; Kourentzes, 2013). Lastly, (Y Grichi et al., 2017; C. Jennings et al., 2016a) have used a data-driven method by creating machine learning algorithms to forecast the obsolescence date.

This paper mainly discusses the life cycle curve approach to forecasting. Two main papers are critically discussed in this work, which are (Song & Zipkin, 1996; van Jaarsveld & Dekker, 2011a). Jaarsveld used a continuous Markov model for spare parts demand with two states, where state one represents the part which is moving (having demand), while the second state represents the part which is dropped (dead). The two states of the Markov process are driven by a compound Poisson process. On the other hand, Song used Markov chain with the Poisson process to model the demand behavior. Both reference papers used the same basic model: compound Poisson process, to model arrival order. They assumed that the transition from high demand states to lower demand states culminates in an obsolescence state. However, these methods are not sufficient to forecast the life cycle curve, and therefore it is necessary to

introduce a new obsolescence forecasting approach because of the lack of scalability and accuracy in these current methods. The novel approach will take into account all the stages of the life cycle. Nevertheless, the extensions of calculation for more than two states will complicate the estimation of the model parameters. In this aspect, hidden Markov theory is used to estimate significantly the parameters of the model (van Jaarsveld & Dekker, 2011a).

5.3.3 HMM: Baum-Welch

The hidden Markov model is useful for modeling sequential data with a few parameters using discrete hidden states, and the estimation procedures are usually based on the EM algorithms (Viterbi algorithm).

One main feature of this algorithm is the joint usage of forward probabilities ($\alpha_t(i) = P(O_0, \ldots, O_t, q_t = S_i|M)$) and backward probabilities,$\beta_t(i) = P(O_{t+1}, \ldots, O_T, q_t = S_i|M)$ used for evaluation problems, to compute the probability of transition from state *i* to state *j,* $\xi_t(i,j) = P\big(q_t = S_i; q_{t+1} = S_j \big| O, M\big)$.

For HMM training, we use time series data of sample $\hat{\lambda}_a(t)$ as observed Markov emissions (*O*). This data came from repeatedly observed (or simulated) flow demand sequences.

The main functioning of this algorithm is presented as: given a set *O* of observed emissions (data of sample$\hat{\lambda}_a(t)$) and an initial HMM, M_0, we can compute a re-estimated model, M_1, with the main property:

$$P(O|M_1) \geq P(O|M_0) \tag{5.1}$$

Baum-Welch algorithm implies repeating the estimation procedure until some precision condition is reached.

$$\left| \frac{P(O|M_{i+1}) - P(O|M_i)}{P(O|M_i)} \right| < c \tag{5.2}$$

The output of Baum-Welch algorithm is the estimated model $\widehat{M}$.

5.4 Proposed framework

Throughout this research, the literature has shown an evaluation of the most promising approaches to forecast obsolescence. In fact, Markov process comprises an extraordinarily rich and flexible class of models. We suggest how this technique can be adopted to model obsolescence using the formalism of state transition diagrams. In this section, the obsolescence forecasting approach based on Markov model is presented, and a flowchart of the proposed method is illustrated in Figure 5.2. The main objective is to develop a Markov Poisson process model to forecast the obsolescence risk time performed with HMM. The hidden Markov will be applied in this case to estimate the model parameters. In order to estimate the transition probability, five Markov states are used. The transient state is forward and the duration in one particular state follows an exponential distribution. However, the transition from state to state may occur when there is a change in demand order.

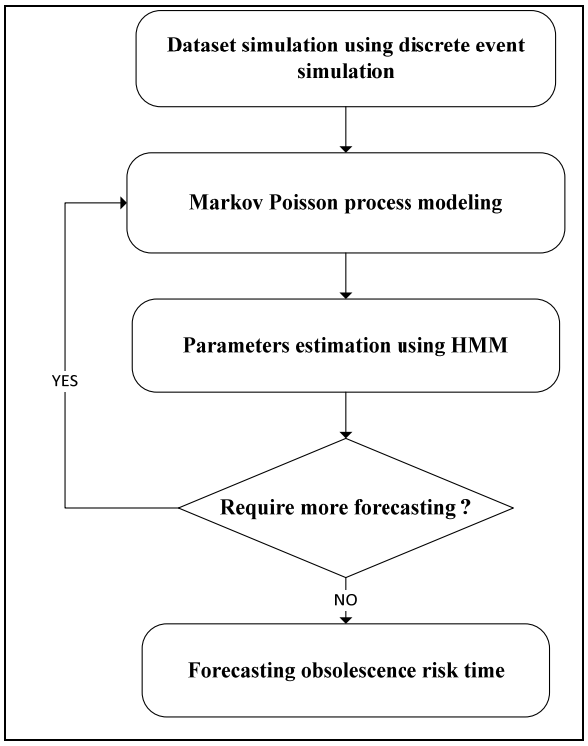

Figure 5.2 Flowchart of the proposed method

In each state, the demand follows a homogeneous Poisson process with constant intensity λ_i., i.e., constant rate of arrival of orders and the sequence of observed states s_i can be characterized by the corresponding intensity λ_{s_i}. Thus, the transition probability is estimated with Poisson distribution with changing order rate. Moreover, we represent the world as a finite-state.

We consider five states leading to obsolescence, as illustrated in Figure 5.3.

State 5: new product comes into the market (demand is slow but increasing)

State 4: increase in demand order.

State 3: demand is high and price low.

State 2: demand decrease: competitors announce a new product.

State 1: demand dead or obsolescence, which presents the last state or absorbing state.

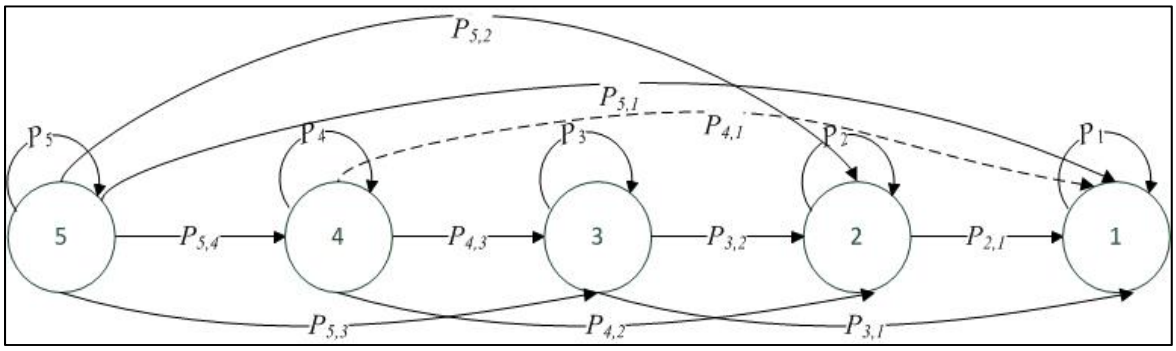

Figure 5.3 Probability transition between states

The transition between states may jump over several stages. For example, the transition from state 5 to 1 may be caused by a competitor's market integration. It can be also be caused by several independent future events, such as the introduction of several competing products that can cause a decline in the demand.

The proposed methodology is as follows: **Step 1** presents the data simulation from Markov hidden Poisson process as a Markov modulated Poisson process. To do this, we set up a transition rate matrix (Q), initial distribution (delta), and the set of Poisson intensities (lambda). The transition rate matrix is cyclic, so we can restart a new obsolescence cycle once we are in obsolescence state in order to simulate many obsolescence cases. Moreover, we assume that there is a greater chance of transitioning to the nearest state than to others, i.e., if we are in state 5 there is more chance of transition to state 4 rather than to states 3, 2, or 1. Once the data

is simulated, **step 2** presents the development of Markov hidden Poisson process; we estimate sample intensity function of the observed order arrivals. The next step is using the sample data. In fact, we assume that we do not know the HMM that generates the data (in real world cases, we only see the order arrival data), and we perform parameter estimation of a Poisson HMM. The estimation is done by EM (Baum-Welch) algorithm of HHM. Once we have the estimated model, obsolescence risk time is estimated from repeated simulations using estimated Poisson HMM. With simulated data from fitted Poisson HMM, we can identify the sojourn time of each state and the time until we reach obsolescence of each cycle. Therefore, we can develop the distribution histogram of obsolescence risk time. Finally, the obsolescence risk time is estimated as the mean of all obsolescence cycles in the simulated data.

5.4.1 Model

the obsolescence forecasting can be performed by the following steps given as a Markov chain model. We assume that the order rate depends on the state of a continuous time Markov process. For the reasons given below, our demand Markov state starts in a fixed state, namely state 5. The probability of transition from i=5 to j=1 in time t is given by equation (5.3):

$$(X_t = 1 | X_0 = 5) = (P^t)_{5,1} \tag{5.3}$$

In this case, we cannot directly observe the sequence of Markov state of the demand flow. However, we can observe the following two main features:

(1) Time of order arrivals: they conform to a discrete time ordered set of points on [0,S) and as a whole they are a point process realization. We can count the number of orders until time "t" by a counting variable N_t, which presents the sum of order arrivals until time t. This variable is discrete, non-decreasing and has jumps of size 1 (due to the fact that two different orders cannot arrive at the same time).

(2) Each arrival order has a random positive and discrete size, which is related to the amount of parts that are required on one specific order.

The combination of points 1 and 2 conforms a parts order arrival process with arrival time epochs given by the point process in (1) and jumps of positive and discrete random size of (2).

The modeling problem implies a double stochastic structure:

First level: At any time "*t*" we observe a compound Poisson process realization Z_t, where:

$$Z_t = \sum_{i=1}^{N_t} S_i \tag{5.4}$$

Where:

Z_t : Is the cumulation of parts requirement till time *t*.

N_t : Is the cumulation of orders requirement till time *t*.

S_i : Is the size (amount of parts) of i^{th} order (size order), and represents the amount of spare parts that are included in an individual order.

Based on Equation (5.4), N_t is a homogeneous Poisson process with intensity parameterλ. N_t is a random variable with probability mass function given by:

$$P\{N_t = k\} = \frac{e^{-\lambda t}(\lambda t)^k}{k!} \tag{5.5}$$

With the probability of a cumulation of k orders in the time interval [0,t]. The HPP drives the sum of Z_t process, i.e., N_t fixes the number of cumulated orders at any time "*t*". The sequence of order sizes S_i is positive independent and identically distributed random variables.

The intensity parameter "λ" discussed in level 1 is not unique. We have a set of five $\lambda's$, each one corresponding to a demand state:

$$\lambda \in \{\lambda_1, \dots, \lambda_5\} \text{ and } \lambda_1 > \cdots > \lambda_5 \tag{5.6}$$

At any time the demand process can be in a particular state $i \in (1, . ., 5)$ following a Markov chain structure where $X_t=i$, transition matrix = P_{5x5} and with trivial initial distribution $\Pi = (X0 = 5) = 1$ (we suppose that all demand cycles began at state 5 and finished on state 1 according to the equation (5.3)).

X_i : is an HMM. This means that we cannot observe the Markov states; we can only observe the order arrivals. The transition matrix is characterized by the following features:

1) $P_{1,1}$=1 and $P_{1,j}$=0 for j=2,..,5 , i.e., once it is in state 1 the sequence will continue in this state forever, because it is an absorbing state. No backward transitions can be done.

2) $$P_{i,j} = \begin{cases} P_{i,j} > 0 & if \ \ i \geq j \\ P_{i,j} < 0 & if \ \ i < j \end{cases} \quad (5.7)$$

By combining Markov model and compound, we can obtain the equation (5.4) as:

$$Z_t = \sum_{i=1}^{N_{X_t}} S_i \quad (5.8)$$

Where, N_{X_t} is an HPP with intensityλ_{X_t}. The intensity is driven by the Markov chain.

5.4.2 Parameters estimation

As shown in the previous discussion, it is not possible to observe the sequence of Markov process; only the sequence of orders can be observed. To discover the Markov sequence, it is necessary to make some data transformations: First, the time of arrival orders should be taken into account, and the sample-counting variable$\widehat{N}_t$.Second developed, estimating the sample intensity function $\hat{\lambda}_a(t)$ for a time window a, given by the question below:

$$\hat{\lambda}_a(t) = \frac{\widehat{N}_{t+a} - \widehat{N}_t}{a} \quad (5.9)$$

For a regular partition of time windows [0,S] of demand cycle, i.e., a sequence of disjointed intervals :

[0,a],(a,2a],….,((n-1)a,S].

$\hat{\lambda}_a(t)$ can be viewed as a discrete sequence of λ at epochs a, 2a, 3a… S. The $\hat{\lambda}_a(t)$ conforms to a discrete time series of sample intensities.

Each discrete time can be associated to Markov, and these times considered as transitions points. Some transitions will be to the same state (P_{ii}) and the sequence of sample $\hat{\lambda}_a(t)$ will

have similar values. The duration in a particular state *i*, ΔT_{Si} , is a random variable with the following distribution:

$$P(\Delta T_{Si} = n) = (p_{ii})^n \tag{5.10}$$

The probability of *n* identical sequences of Markov states, i.e., *n* transitions to itself.
The probability of a transition to another state at duration *n* and *n*-1 duration in state *i* is:

$$\sum_{j \neq i} P\left(\Delta T_{Si} = n - 1; S_j\right) = (P_{ii})^{n-1}(1 - P_{ii}) \tag{5.11}$$

The former approach is the basis for a hidden Markov model. The time series of sample $\hat{\lambda}_a(t)$ are emissions from a Markov state. Due the random nature of $\widehat{N}_t$ we will not observe identical values of $\hat{\lambda}_a(t)$ during the time duration of a particular state. These fluctuations conform to a probability density of emission.

$$b_i(O_k) = P\{O_k(t) | X_t = i\} \tag{5.12}$$

Where $b_i(O_k)$ is the probability of emission O_k given the Markov Chain with state *i*. The emission is $\hat{\lambda}_a(t)$.
The estimation procedure consists of applying the Baum-Welch algorithm under the previous model restrictions to the sequence $\hat{\lambda}_a(t)$. This algorithm is the HMM version of an EM algorithm.

5.5 Numerical example

This section presents a numerical example to forecast life cycle curves and estimate the obsolescence risk time. R programming is adopted to develop the model.
In the first phase, the data is generated by discrete event simulation with repeated simulations of sequences of Markov chain. The transition matrix is first given by Figure 5.4. This transition matrix is fixed for simulation purposes. The main assumptions are: 1) the process will be in a

specific state for a relatively long time, then the *P_ii* values are close to 1. 2) The transitions to other states occur only in forward direction starting from state 5.

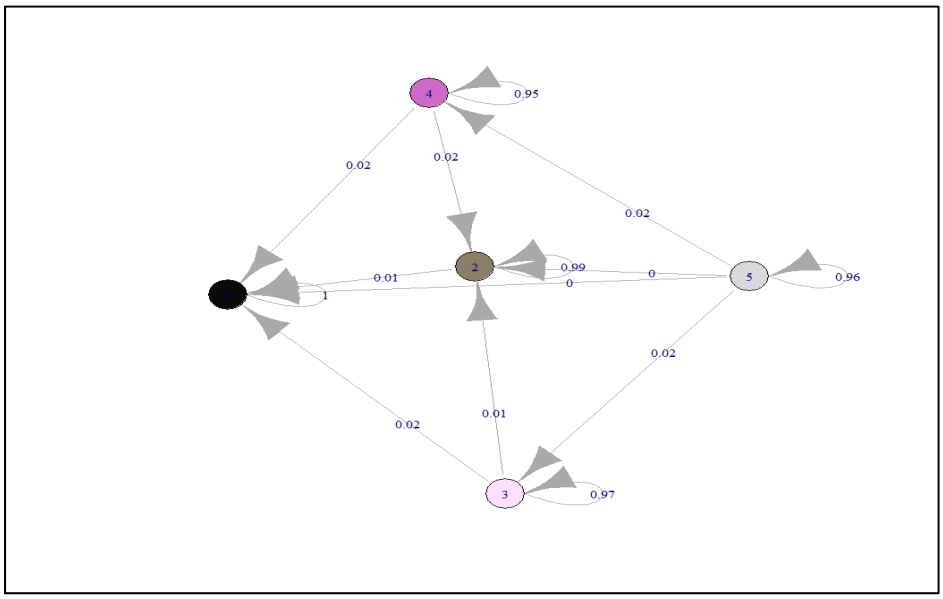

Figure 5.4 Transition matrix of markov chain

The second phase is focused on the visualization of the sample process. For each Markov epoch in the simulation stage, the HPP is generated with lambda related to Markov states. This HPP has a duration given by the time window's length "a".

Figure 5.5 presents a plot of the first cumulated order arrival sequence generated from data simulation. Then, we present the empirical intensity in a reduced time window.

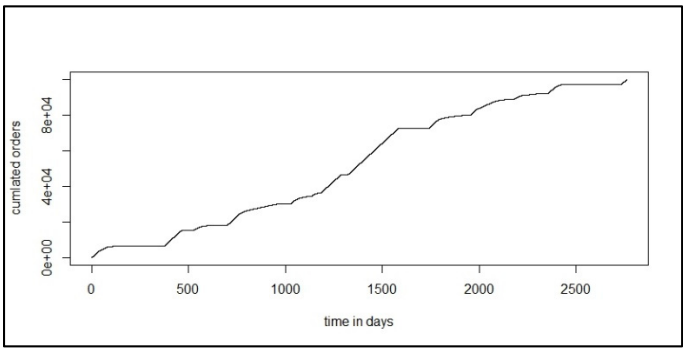

Figure 5.5 Cumulation of arrival orders

Figure 5.6 shows the obsolescence risk time, where the first window is set to 200-time units (in days) in which we can see only one obsolescence cycle. On the other hand, the second plot is set to 2000-time units' length, where we can see a sequence of nine obsolescence cycles. This figure shows the forecast results fitted by the Markov Poisson process model with the uncertainties of forecasting.

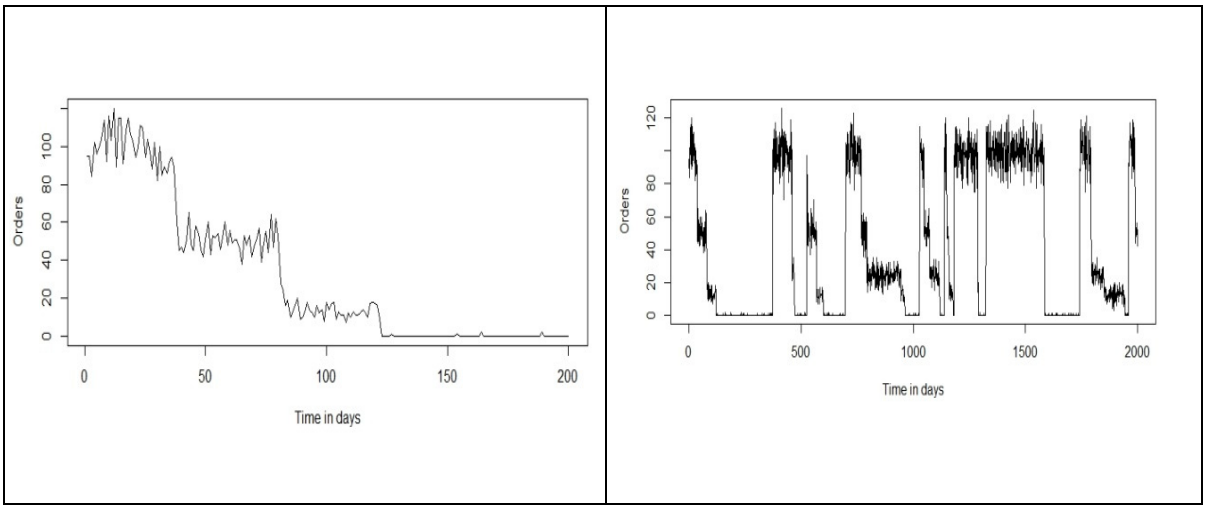

Figure 5.6 Order Vs. time in days

The third step is the estimation of model parameters using hidden Markov Poisson process from simulated data. The estimation is done by EM algorithm for HMM. At this point, we assume that we do not know the model, and we only know the order arrival data.

The obsolescence forecasting is estimated from order arrivals data, which is presented by the estimation of expected time for obsolescence transition, from state 5 at t=0, given by Markov model. As already discussed, the last stages of the life cycle curves are the obsolescence stages. Obsolescence risk time is given by Figure 5.7. The expected obsolescence time is estimated in 130 days in this case.

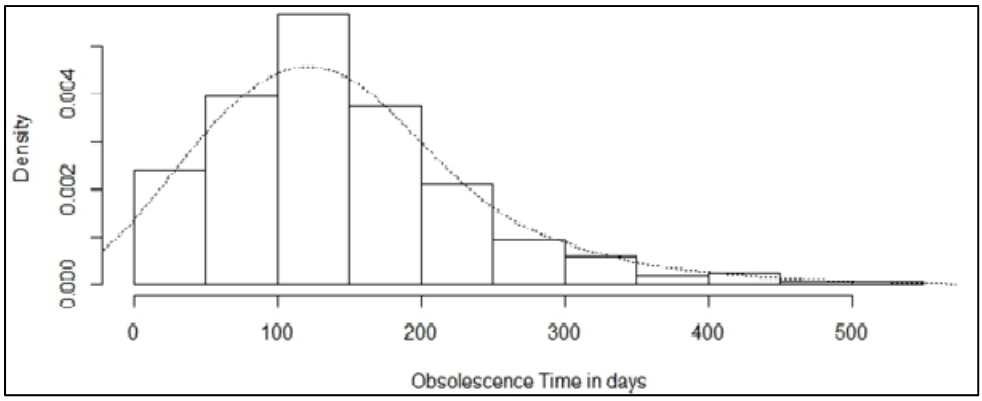

Figure 5.7 Distribution of obsolescence risk time

The accuracy of the model is calculated based on the coefficient of variation:

$$Coefficient\ of\ Variation\ (CV) = \frac{std}{Mean\ value} * 100 = 6\% \tag{5.13}$$

5.6 Discussion

In the previous section, the obsolescence forecasting based on Markov chain and homogeneous compound Poisson process method was presented and applied using data simulation. The Paper introduced as well the hidden Markov theory to estimate the model parameters.

The first part of the model was the generation of data; therefore, a transition matrix and the Poisson intensities were set up. Once the data was simulated, the intensity function of the observed order arrival was estimated. Next, the estimation of the parameters of the Poisson HMM was performed. From the simulated data fitted from Poisson HMM, the sojourn time was identified for each state. After obtaining the full life curve from the proposed predictive method, the expected obsolescence time was estimated. On the other hand, the estimation of the life cycle curve proved complicated because the transitions between phases may jump over stages that can be caused by several independent future events as discussed by (Song & Zipkin, 1996; van Jaarsveld & Dekker, 2011a).

The forecast result in Figure 5.7 presents the predicted life cycle curve, which indicates an increase in the demand represented by the availabilities of sales data, while the right-hand side

of the curve represents a decrease in the demand. The expected obsolescence time is estimated at 130 days in this case. For the accuracy of the life cycle curve forecasting, the method was evaluated using standard deviation to measure the amount of variability of the obsolescence zone, which was presented by 6%.

As discussed in previous sections, the proposed method can overcome the big limitation of Jaarsveld's method which presented by an approach to forecast obsolescence risk performed by Markov chain with only two states. This approach was limited and not adequate to forecast the whole life cycle since it considers only one transition between two states. The gap and limitation in the current approach were examined and improved in the proposed method. This approach is proven more accurate to forecast the obsolescence risk time.

5.7 Conclusion and future work

Research on obsolescence is growing due to the serious challenges that complex systems face every day. However, firms which are facing to manage these problems must react. Forecasting obsolescence is the key enabler for all proactive and strategic management approaches to addressing obsolescence. Through an accurate obsolescence forecasting, firms can reduce significantly the cost and save millions of dollars.

This paper introduced an approach to estimate obsolescence time based on simulation of demand data using Markov homogeneous Poisson process performed with hidden Markov to estimate the model parameters. The results of forecasting are presented in figures 5.5 and 5.6. The proposed method consists of three phases: data simulation, Markov homogeneous Poisson process modeling to forecast the life cycle curves, and hidden Markov to estimate the parameters of the model. As discussed regarding Table 5.1, the proposed method can forecast obsolescence risk time taking into account the whole life cycle curve performed by the demand rate. A numerical example was provided to demonstrate the capability of this method.

The main theoretical contribution of this paper was introducing Markov model and hidden Markov theory for obsolescence life cycle forecasting after an in-depth analysis of gaps in the actual approaches (Jungmok & Namhun, 2017; van Jaarsveld & Dekker, 2011a). The method

was examined for its ability to forecast the obsolescence risk time using sales data as a requirement with high accuracy.

This paper provides an approach enabling firms that are facing this kind of problems to enhance their forecasting and to ensure support to the long-life systems. The proposed method can be applied by any company using an enough data. The life cycle forecasting could also be useful to help make lifetime buy or last buy orders more accurate.

Along with these strengths, the proposed method's limitations are also identified. Like all obsolescence forecasting frameworks, this approach has some limitations that may compromise the validity of the estimations. First, the paper considered data simulation; however, if the data used to build the model does not represent the current real world, the model will not be very effective. In obsolescence, there is an extremely high chance of this occurring due to rapid innovation. Other datasets and real-world problems for obsolescence forecasting can be tested using this method. Second, the proposed method is more complicated than Jaarsveld's method and more sensitive to data because no fixed distribution is assumed.

Moreover, there are some other market factors for which manufacturers should consider carrying out an additional risk assessment, which can cause a lot of variations in the life cycle. Another limitation is based on the Markov property, i.e., the next state depends only on the current state; this will not influence the past state sequence.

It would be interesting to extend the method by considering other factors to improve the accuracy of the forecasting. Real data application might be useful to test the feasibility of the model. Another direction for future research is to perform some sensitivity analysis to test output variations with changes in the model parameters.

CHAPITRE 6

DISCUSSION

Ce chapitre présente une synthèse des développements scientifiques des méthodes et outils réalisés tout au long de la thèse pour la prévision de l'obsolescence, en réponse aux objectifs de doctorat mentionnés à l'introduction. À travers ce chapitre, les avantages et limitations de la recherche seront exposés.

6.1 Justification du problème de recherche

La problématique de la recherche étudiée dans cette thèse se concentre sur le manque de modèles et méthodes efficaces pour la prévision de l'obsolescence. En effet, le problème d'obsolescence a été évoqué initialement à travers un besoin industriel d'une entreprise avionique partenaire. La littérature a confirmé par la suite que ce problème existe dans plusieurs secteurs tels que l'automobile, l'électronique, etc., et qu'il n'est pas seulement limité au secteur aéronautique. Néanmoins, la littérature a surtout mis l'accent sur l'obsolescence des composants aéronautiques, car ce problème émerge fréquemment dans des systèmes qui ont un cycle de vie beaucoup plus long que leurs composants. Cet écart rend ces industries très fragiles relativement au problème d'obsolescence et entraîne d'énormes retards ainsi que des coûts supplémentaires.

Ce qui a élargi l'écart du cycle de vie entre les systèmes et leurs composants est l'introduction de nouvelles technologies qui ont révolutionné l'industrie électronique. Cette évolution a créé des milliers de nouveaux composants avec un cycle de vie très court. De ce fait, l'obsolescence peut par conséquent se produire lorsque la technologie n'est plus offerte sur le marché.

L'obsolescence est un problème majeur pour plusieurs entreprises, motivé ainsi par la littérature, et qui nécessite d'être résolu. Actuellement, plusieurs entreprises ne disposent pas de méthodes efficaces pour prévoir l'obsolescence; par conséquent, la prévision de l'obsolescence est nécessaire dans la gestion proactive, ce qui justifie cette recherche.

6.2 Synthèse des développements scientifiques

Cette recherche s'est intéressée aux développements scientifiques de trois méthodes de prévision de l'obsolescence présentées comme des articles scientifiques dans les chapitres 3 à 5. Pour répondre aux objectifs de doctorat, la recherche s'est divisée en deux volets. Le premier a été consacré à la prévision du risque d'obsolescence et le deuxième s'est intéressé à la prévision du cycle de vie d'obsolescence.

6.2.1 Prévision du risque d'obsolescence

Le premier développement scientifique a répondu aux deux premiers objectifs de doctorat, soit la prévision du risque de l'obsolescence, présentée aux chapitres 3 et 4. La conception scientifique a ainsi permis de proposer deux modèles prédictifs basés sur l'apprentissage machine.

La première partie de la conception a évoqué l'intégration de l'apprentissage machine supervisé pour un problème de classification. Comme nous l'avons présenté au chapitre 3, l'étude a permis le développement d'un modèle de prévision de risque de l'obsolescence. Initialement, l'étude a fait appel à plusieurs algorithmes pour comparer leurs performances. D'après les résultats obtenus, il s'est avéré que le modèle forêt aléatoire a donné les meilleurs résultats. À partir de ces résultats, l'étude s'est concentrée sur l'amélioration de la performance du modèle prédictif. Pour cela, une approche d'optimisation métaheuristique a été combinée avec le modèle forêt aléatoire, dans le but d'optimiser les paramètres du modèle et de sélectionner les variables appropriées. Les résultats présentés au chapitre 3 ont démontré que l'intégration GA-RF améliore significativement la précision et la performance du modèle. L'approche proposée se montre aussi efficace avec l'intégration d'une autre approche d'optimisation PSO, présentée à l'annexe III.

La deuxième partie du développement scientifique, présentée au chapitre 4, a permis de développer une méthode stratégique qui propose une intégration combinée entre l'apprentissage machine supervisé et l'apprentissage machine non supervisé pour évaluer le

risque d'obsolescence. La méthode a été établie sur deux niveaux. Le premier niveau présente un problème de classification déjà discuté précédemment, tel que présenté au chapitre 3. Au deuxième niveau, une technique d'apprentissage automatique non supervisé a été proposée pour partitionner les données en différentes classes qui décrivent le niveau du risque en fonction du modèle prédictif obtenu au premier niveau. En effet, le problème de classification de la première phase a été utilisé dans le clustering en se basant sur k-means pour obtenir des classes définies comme un risque « faible », « moyen » ou « élevé » à partir des probabilités.

Par ailleurs, le choix des algorithmes s'est basé principalement sur le critère de complexité en temps, défini par le temps d'exécution des modèles développés. En effet, le temps d'exécution dépend de la cardinalité et de la qualité des données qui sont introduits dans les algorithmes, ainsi que le choix des paramètres choisis. Plus la cardinalité des données et grande plus le temps d'exécution est plus long. Basé sur l'ensemble des données collecté dans cette étude et les paramètres choisis pour les algorithmes, le temps d'exécution des modèles développés n'a pas dépassé les 60 minutes, et ce pour une cardinalité de près d'un millier de données.

L'approche proposée offre non seulement l'avantage de la classification de l'obsolescence, mais elle fait également appel à l'apprentissage machine non supervisé pour classifier les données qui sont prédites comme non obsolètes sur une échelle de risque. Cette intégration combinée s'est avérée efficace pour améliorer les performances des modèles prédictifs dans la gestion du risque d'obsolescence. Elle peut aider à choisir la meilleure approche de mitigation stratégique. Les avantages et l'efficacité de ces méthodes prédictives ont été démontrés par une étude de cas sur deux ensembles de données.

Limites et recommandations

Une des principales limites dans cette étude sont les données utilisées pour entrainer les modèles. Premièrement, les données sont collectées à partir de différents sites Internet, ce qui a provoqué initialement des données manquantes et qui ne sont pas à jour, dans certains cas.

Lors du traitement d'informations, il s'est avéré que la répartition des données n'est pas proportionnelle entre celles qui sont étiquetées comme obsolètes et celles qui sont non obsolètes. En effet, dans l'ensemble de données, il se trouve que la majorité des données sont étiquetées comme non obsolètes, ce qui a perturbé un peu la validité des résultats dans l'un des ensembles de données utilisés. D'autre part, l'échantillon n'était pas assez étendu pour aider les algorithmes à s'entraîner correctement. La raison pour laquelle la collecte des données était passable (mais tolérable) se résume au manque d'informations dans plusieurs données. La majorité des données n'avaient pas d'étiquettes claires (obsolètes ou non obsolètes), mais comportaient plutôt les mots « échantillon » ou « contacter le vendeur », ce qui a conduit à supprimer toutes les données ayant ce genre d'étiquettes lors du traitement des données.

Il existe également une limite liée aux données. En effet, les résultats de l'étude ont été appliqués uniquement sur des ensembles de données dont les caractéristiques n'ont pas d'indice paramétrique évolutif clair. Cela signifie que leur évolution dans le temps n'est pas claire ou connue et demeure inexpliquée. Par exemple, dans le cas des cellulaires, plusieurs caractéristiques, telles que la taille de la mémoire, sont censées évoluer. Ainsi, dans certains nouveaux cellulaires, on peut trouver des mémoires de petite taille, mais qui sont très bien vendus dans le marché jusqu'à aujourd'hui, étant donné qu'ils coûtent moins cher avec une mémoire de petite taille.

Bien que l'analyse de conception prédictive combinée proposée au chapitre 4 montre un certain potentiel et des opportunités, il existe cependant certaines limites et des pistes à explorer. Dans cette recherche, ML non supervisé est appliqué pour classer les données sur une échelle de risque basée sur la probabilité de toutes données prédites comme non obsolètes.
Des tests semblables à ceux de C.L Josias (2009) peuvent être investis à cet égard, car ils visent à analyser initialement la corrélation entre les différentes variables pour développer par la suite un indice de risque basé sur les différents facteurs qui influencent l'obsolescence dans certains composants technologiques. L'équation de l'indice de risque pourra notamment aider à améliorer la classification des données dans une échelle de risque d'obsolescence.

Plusieurs aspects peuvent être pris en compte pour des recherches futures. Le modèle peut être reproduit avec des données plus larges des composants électroniques avec plus de variables dont l'évolution s'avère claire. De plus, d'autres algorithmes d'optimisation peuvent également être utilisés, comme les colonies de fourmis.

Par ailleurs, la qualité du regroupement peut être évaluée avec d'autres méthodes pour valider les résultats. Aussi, il serait intéressant d'intégrer le clustering (regroupement) dans la première phase de tri de données. En effet, l'apprentissage non supervisé peut être appliqué sur deux étapes différentes : l'étape de regroupement des ensembles de données puis l'étape de regroupement de risque d'obsolescence. Ceci peut renforcer notre confiance dans le fait que cette stratégie peut produire une méthode robuste de prévision du risque d'obsolescence.

6.2.2 Prévision du cycle de vie de l'obsolescence

Le deuxième développement scientifique a répondu au troisième objectif de doctorat, soit la prévision du cycle de vie de l'obsolescence, présentée au chapitre 5. La conception scientifique a permis de proposer un modèle de prévision du cycle de vie basé sur la chaîne de Markov et le processus de poisson composé. Comme nous l'avons présenté au chapitre 5, l'étude a permis le développement d'un modèle pour estimer la date d'obsolescence basée sur la simulation des données du taux de la demande en utilisant le processus de poisson homogène. La méthode a s'est basée sur trois étapes : la simulation des données, la modélisation du processus de poisson homogène de Markov pour prévoir la courbe de cycle de vie et finalement, l'intégration de Markov cachée pour estimer les paramètres du modèle.

L'avantage de cette étude est que, contrairement à la méthode de van Jaarsveld & Dekker, (2011), elle utilise toutes les phases de cycle de vie par l'introduction de la théorie de Markov cachée. La méthode proposée a été examinée pour sa capacité à prévoir la date d'obsolescence en utilisant les données de ventes, et ce, avec une grande précision. En outre, cette méthode offre l'avantage de prévoir la possibilité des changements brusques du taux de la demande dans chaque état de cycle de vie.

Limites et recommandations

Avec ces points forts, il existe des limites dans cette étude et des pistes de recherche à explorer. Comme toute approche, cette recherche présente certaines limites qui peuvent compromettre la validité des estimations. Premièrement, la partie expérimentale a été examinée avec une simulation de données. Cependant, en utilisant des données réelles, le modèle peut produire des résultats plus précis, efficaces et plus proches de la réalité. Des études expérimentales avec d'autres études de cas sur des problèmes réels seront fortement suggérées. De plus, il existe certains autres facteurs du marché pour lesquels les fabricants devraient envisager d'effectuer une évaluation de risque supplémentaire, ce qui peut entraîner de nombreuses variations dans le cycle de vie.

Une autre limite vient du fait que la méthode proposée est plus compliquée que celle de Jaarsveld, donc elle est plus sensible aux données, car aucune distribution fixe n'est supposée. De plus, une limite qui dépend de la propriété de Markov, c'est-à-dire que l'état suivant dépend uniquement de l'état actuel, n'influencera pas la séquence d'état passée.

Il serait intéressant d'étendre la méthode en considérant d'autres facteurs pour améliorer la précision des prévisions. Des tests avec des données réelles seraient requis afin de tester la faisabilité du modèle. Une autre direction pour la recherche future consiste à effectuer une analyse de sensibilité pour tester les variations de sortie avec des changements dans les paramètres du modèle.

6.3 Généralisation et validation des méthodes

Les méthodes proposées dans cette thèse ont pour objectif de prévoir le risque et le cycle de vie de l'obsolescence des composants électroniques. Ces méthodes ont suivi les étapes de la science de conception discutées dans le chapitre 2. Le Tableau 6.1 présente un résumé des développements scientifiques réalisés dans ces travaux de recherche, ainsi que les principaux travaux développés dans la littérature au cours des 10 dernières années. Dans ce tableau,

chaque méthode est caractérisée en fonction du type de prévision d'obsolescence, et associée aussi à des facteurs d'évolutivité les plus courants dans la prévision de l'obsolescence : l'apprentissage machine, les données de ventes, la manipulation humaine et la capacité de prévoir sur une grande échelle. En effet, l'interprétation et l'avis des experts dans le processus de prévision de l'obsolescence peuvent entrainer des erreurs d'estimation, et par conséquent, l'inexactitude des prises de décision. La collecte des données de ventes, quant à elle, nécessite un suivi périodique de données de ventes. L'historique des données doit être disponible et archivé dans la base de données de l'entreprise. Idéalement, les méthodes qui ne nécessitent pas de données de vente ou de la manipulation humaine et qui devraient être capables de prévoir l'obsolescence d'un grand échantillon de composants, sont les meilleures approches pour la prévision de l'obsolescence.

Tableau 6.1 Résumé des développements scientifiques et les facteurs évolutifs associés à chaque méthode

Méthodes	ML supervisé	ML non supervisé	Prévision du risque de l'obsolescence	Prévision du cycle de vie de l'obsolescence	Données de ventes	Manipulation humaine	Grande échelle de données	Méthode améliorée
GA-RF PSO-RF	✓	-	✓	-	-	-	✓	**
PSO-RF avec k-protos	✓	✓	✓	-	-	-	✓	**
Markov /processus de poisson composé	-	-	-	✓	✓	-	✓	**
Rojo et al. (2012)	-	-	✓	-	-	✓	-	-
Zheng (2011)	-	-	-	✓	✓	✓	-	-
Jaarsveld et al. (2011)	-	-	✓	-	✓	-	✓	*
Jenning (2016)	✓	-	✓	✓	-	-	✓	*
Zolghadri et al. (2018)	-	-	✓	-	-	-	-	-
Josias et al (2004) et (2009)	-	-	✓	-	-	-	-	-

** Méthodes améliorées par rapport aux méthodes * disponibles

Pour valider la qualité des résultats des deux méthodes GA-RF et PSO-RF, l'ensemble de données utilisé par Jennings et al. (2016) pour entrainer l'algorithme de la forêt aléatoire, a été appliqué également dans ces travaux de recherche. Les résultats affirment que les méthodes développées améliorent considérablement la précision de la classification de l'algorithme. Les résultats de comparaison sont présentés dans le Tableau 6.2.

Tableau 6.2 Comparaison des mesures de précision

	RF (développé par Jennings et al (2016)	**GA-RF**	**PSO-RF**
Mesure de précision $\frac{TP+TN}{TP+TN+FP+FN}$	92.9%	93.3%	96%

Le deuxième développement scientifique s'est intéressé à l'intégration de PSO-RF avec un algorithme de l'apprentissage non supervisé. Cette approche innovante vise non seulement à faire des classifications avec PSO-RF, mais aussi à catégoriser un ensemble de données dans une échelle de risque. L'introduction de l'apprentissage machine non supervisé pour la prévision du risque de l'obsolescence n'a jamais été appliqué auparavant.

En termes de validation, l'approche proposée a été comparée qualitativement aux méthodes développées par Rojo et al. (2012), et Josias et al. (2009) pour la prévision du risque de l'obsolescence en utilisant une échelle de risque (voir Tableau 6.3). En effet, ces méthodes exigent des facteurs d'obsolescence comme le nombre de fournisseurs disponibles, le niveau du stock et le taux de consommation ainsi que le nombre d'années restant pour le composant. Ces facteurs doivent être disponibles pour pouvoir classifier un composant comme étant à risque faible, moyen ou élevé. Toutefois, l'approche proposée ne l'exige pas.

Tableau 6.3 Comparaison entre la méthode développée et les diverses approches disponibles

	Échelle de risque	**Facteurs clés**	**Regroupement d'un grand nombre de composants**
Méthode développée	✓	-	✓
Rojo et al. (2012)	✓	✓	-
Josias et al. (2004), (2009)	✓	✓	-

Le troisième développement scientifique s'est intéressé au développement d'un modèle de prévision du cycle de vie. Alors que la méthode de Jaarsveld et al (2011) considère deux états du cycle de vie, soit en production ou obsolète pour estimer le risque d'obsolescence, la méthode développée prévoit la période exacte avant que le produit ne soit obsolète. De plus, contrairement à la méthode de Zheng et al. (2011), la méthode proposée a une capacité de prédire la courbe de cycle de vie sans l'exigence d'avoir des facteurs paramétriques évolutionnaires comme la méthode de Zheng l'exige. L'autre avantage d'appliquer la chaîne de Markov par rapport aux autres méthodes est lié à la propriété puissante de Markov qui implique que les propriétés des variables aléatoires liées à l'avenir ne dépendent que de l'information de l'état présent, et non de l'information provenant des états passés.

Validation des méthodes par rapport au besoin industriel

L'entreprise partenaire gère l'obsolescence de façon réactive. À la vue de ce manque, l'entreprise a cherché à implanter un outil efficace de prévision qui va l'aider à réduire les coûts supplémentaires engendrés par l'obsolescence et à minimiser son impact potentiel sur la chaine de production. Pour évaluer le risque et pour bien cibler leurs besoins et les livrables requis, un questionnaire a été élaboré à cet égard en collaboration avec Héloïse Conrad (Conrad, 2016). Le questionnaire a été réalisé et approuvé par un expert en obsolescence au sein de l'industrie et par des professeurs chercheurs (voir annexe IV). Les résultats du questionnaire ont servi à l'identification des facteurs de l'obsolescence en termes de prévision et à mieux comprendre leur impact sur la chaine de production. De plus, le questionnaire a permis de développer des modèles pour prévoir l'obsolescence en réponse à leurs attentes.

Le premier livrable qui a été communiqué avec l'entreprise est l'identification des principaux facteurs d'obsolescence qui nuisent à leur gestion de l'obsolescence. Après une validation de ces facteurs, des lignes directrices ont été élaborées et validées afin d'améliorer leur gestion de l'obsolescence (Figure 6.1). Ces communications et directives ont permis par la suite d'établir les caractéristiques qui forment les modèles afin de répondre à leurs besoins.

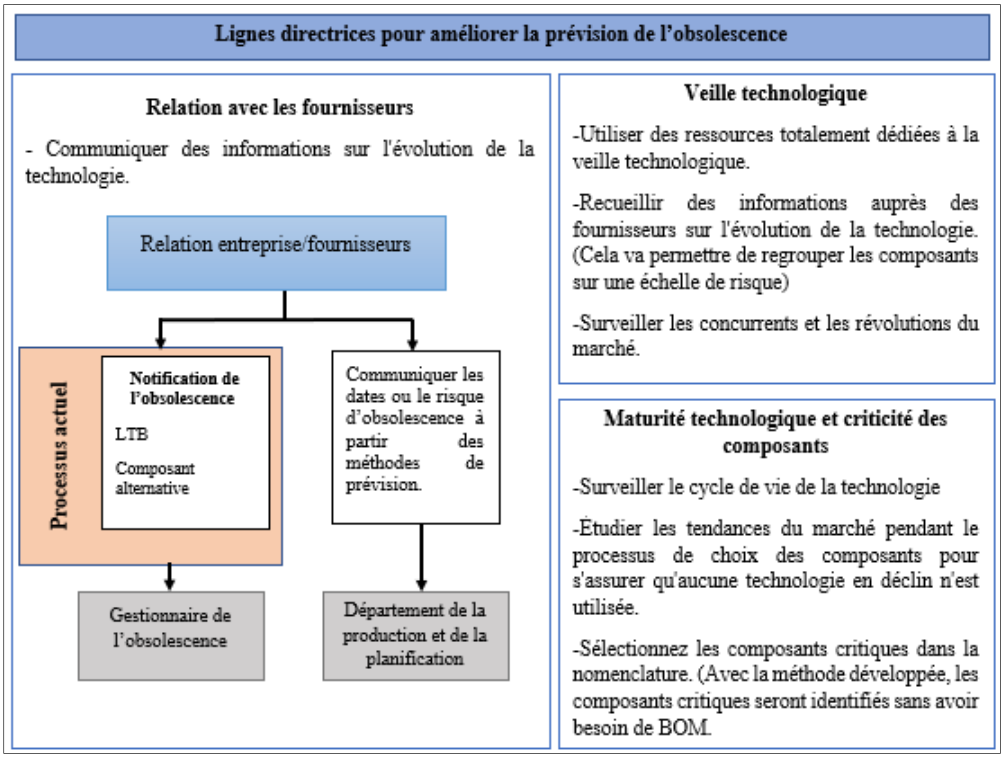

Figure 6.1 Lignes directrices pour améliorer la prévision de l'obsolescence

Les méthodes développées dans cette thèse ont démontré des résultats suffisants pour qu'une mise en application dans l'entreprise est fortement envisagée. Le travail effectué dans cette recherche offre à l'entreprise plus qu'une méthode de prévision, soit la prévision du risque par la classification ou la prévision de la date d'obsolescence à partir de l'historique des données de vente. La méthode sera choisie selon les données disponibles. Les solutions proposées semblent convenables par rapport au besoin industriel et pourront donc être validées dans l'entreprise.

Par ailleurs, avant de procéder à l'implémentation de ces modèles, notamment les algorithmes de l'apprentissage machine, des outils doivent être mises en place dans l'entreprise. D'abord, une plateforme de base de données peut être implanter dont laquelle tous les composants électroniques avec leurs spécifications techniques détaillées seront collectés. Cette plateforme doit aussi être surveillé, entretenu et modernisé périodiquement. Pour cela, un gestionnaire de données spécialisé en datamining doit être consacré pour gérer ces données. De plus, des

communications fréquentes avec les fournisseurs doivent être établis pour échanger l'information sur le cycle de vie des composants.

Il existe notamment des bases de données disponibles tel que SiliconExpert, une grande plateforme qui contient des milliers de composants électroniques avec leurs BOM. Elle offre également des approches pour la gestion de l'obsolescence et des alertes sur l'évolution technologique. L'implantation d'une plateforme similaire peut être envisagée à cet égard, et les algorithmes développés peuvent ainsi être intégrés dans cette plateforme. La mise à jour des statuts des composants ainsi que leurs cycles de vie doivent être suivis et actualisés fréquemment. Aussi, les algorithmes doivent s'exécuter à chaque fois qu'une mise à jour est faite. L'amélioration et la maintenance des modèles de prévision est envisagé si d'autres facteurs seront pris en considération dans la modélisation.

CONCLUSION

La recherche sur l'obsolescence se développe progressivement en raison de graves défis et de la pression auxquels sont confrontées les entreprises face à ce problème. La prévision de l'obsolescence est le principal catalyseur dans la gestion proactive et stratégique pour lutter contre l'obsolescence. Grâce à la prévision, les entreprises peuvent émettre un plan d'action proactif afin de réduire considérablement les coûts engendrés par les problèmes d'obsolescence.

Cette dissertation discute des méthodes, outils et connaissances développés pour la prévision de l'obsolescence des composants électroniques. Les méthodes proposées sont basées principalement sur l'analyse de données à grande échelle. L'étude présentée dans cette thèse a permis de répondre à plusieurs questions et de proposer des méthodes innovantes relativement à la problématique de recherche et afin de satisfaire les objectifs du doctorat.

La thèse a introduit trois approches pour la prévision de risque et de cycle de vie de l'obsolescence. Pour la prévision du risque, l'étude s'est basée principalement sur l'apprentissage machine supervisé et non supervisé (chapitres 3 et 4) pour un problème de classification du risque d'obsolescence. En ce qui concerne la prévision du cycle de vie, cette thèse a permis le développement d'une approche pour estimer la date d'obsolescence basée sur la variation de la demande en utilisant le processus de poisson composé et le modèle de Markov. D'un point de vue pratique, cette méthode améliore considérablement la gestion de l'obsolescence, par la prise de décision de la stratégie de mitigation adéquate.

Retombées scientifiques

Cette thèse a permis de valoriser la culture scientifique sous forme de publications d'articles dans des revues scientifiques. Ces travaux de recherche ont permis d'enrichir amplement les connaissances à propos de l'obsolescence, ses facteurs et ses impacts à partir d'une revue exhaustive de la littérature qui a mis l'accent sur les lacunes dans les travaux scientifiques

existants. Les méthodes développées dans cette thèse permettront un renforcement de la recherche scientifique autour de l'obsolescence des composants électroniques par l'appui de trois principales contributions. La première contribution concerne l'introduction de l'apprentissage machine supervisé pour la prévision du risque de l'obsolescence et d'optimiser la classification du modèle par l'intégration d'une approche métaheuristique en proposant une approche innovante de combinaison entre la Forêt aléatoire et l'algorithme génétique. La deuxième contribution scientifique, quant à elle, est justifiée par l'introduction d'une méthode efficace pour la prévision du risque de l'obsolescence basée sur deux apprentissages machine, soit l'apprentissage supervisé et non supervisé. La différence entre ce travail et d'autres travaux antérieurs est que ces travaux de recherche impliquent l'apprentissage supervisé et l'apprentissage non supervisé simultanément. Finalement, la troisième contribution repose sur l'introduction d'une méthode amélioré pour prévoir le cycle de vie de l'obsolescence.

Retombées industrielles

Les travaux de recherche réalisés dans le cadre de cette thèse, et les résultats encourageants permettront une meilleure évaluation de l'obsolescence et pourront offrir un peu de répit aux industries qui sont soumises à de grandes pressions afin de leur assurer un soutien aux systèmes à longue durée de vie. Les méthodes développées entraîneront un changement de perception chez les entreprises pour faire face à l'obsolescence et les aideront à évaluer le risque dans le marché des composants électroniques afin de suivre la cadence de l'évolution technologique. Le projet contribuera à mettre en évidence des perspectives pour l'avenir des industries dans plusieurs secteurs, notamment l'électronique et l'aéronautique, et à mieux piloter ce genre de projet d'une façon stratégique.

Les méthodes développées offrent à l'entreprise une démarche améliorée pour sa gestion de l'obsolescence. Par ailleurs, l'implémentation de ces méthodes selon les données disponible à l'entreprise est envisagée dans un futur proche.

ANNEXE I

OBSOLESCENCE FORECASTING STRATEGY OF TECHNOLOGICAL COMPONENT-A COMPARATIVE STUDY OF ALGORITHMS

Yosra Grichi, Yvan Beauregard, Thien-My Do

Département de génie mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

Article présenté à la conférence « 7th IESM conference », tenue à Saarbrücken, Germany, october 2017

Abstract

Obsolescence is a major issue that is driven by many factors, mainly through technological advancement, in which the life cycle of the components is often shorter than that of their systems. Basically, obsolescence problems are often sudden and not planned, causing delays and extra costs. Moreover, forecasting appears to be one of the most efficient solutions. By using a case study, this paper presents a comparison of the different algorithms which are: Back Propagation Neural Network (BP-NN), AdaBoost, Support Vector Machine (SVM), Random Forest (RF) and Decision tree that is capable of forecasting obsolescence risk of a large sample of electronic components with a high degree of accuracy. Using supervised learning, the data set is divided into 2 groups: 2/3 for training and 1/3 for testing and validation. In addition, a database is developed to collect all the data necessary for training. Finally, algorithms are compared between them to estimate the best predictor based on confusion matrix.

Related work

As defined by Moore's law, the rapid evolution of electronic components continues to grow, which stipulates that semiconductor density doubles approximately every 18 months (Homchalee & Sessomboon, 2014; P. Sandborn, 2008; Tomczykowski, 2003). In fact, when we talk about technological obsolescence, we immediately think of electronics, this evolution creates new electronic components every year with short lifetimes. In USA, the industry has grown at a rapid rate since the 1990s. New technologies are introduced in the market with increased speed.

Today, the short life cycle of technology and the lack of forecasting represent a challenge for several companies which need to take into account the risk of obsolescence; on the other hand, the modeling of obsolescence is considered as a complex problem that requires a good

knowledge about the components affected by obsolescence. In the same way, it is necessary to know the lifetime of the components and their technological evolutions and therefore know all the factors that influence the obsolescence of components. Moreover, forecasting appears to be one of the most efficient solutions.
In the literature, there are two types of forecasting, which are forecasting of the obsolescence risk and forecasting of the obsolescence date. These prediction methods are based on a human knowledge that is used to estimate obsolescence or market survey; however these methods do not produce satisfactory results in most cases (C. P. Jennings, 2015). Forecasting obsolescence risk was a reactive nature based on the resolution of the problem once noticed. It's used to predict the probability that a component is still in production or not (F. R. Rojo et al., 2012; Van Jaarsveld & Dekker, 2011b). Rojo (2012) has developed a methodology to evaluate the risk level of obsolescence based on availability of the stock, consumption rate and the number of sources available (F. R. Rojo et al., 2012) .
In the literature, the majority of the works are based on the life cycle forecasting for a single type of which become obsolete every day. To resolve this problem, machine learning appears as an optimal method for forecasting. In fact, machine learning is a data analysis method that automates the recognition of data models without human manipulation (Zurada, 1992) capable of forecasting of a large sample of electronic components with a high degree of accuracy. During the last decades, machine learning has attracted the attention of many researchers in several disciplines and has been applied in a number of areas. According to Wu and Kumar (X. Wu et al., 2008), ten top algorithms have been identified by IEEE « International Conference On Data Mining », include, AdaBoost, SVM, K-Means, decision tree, etc. These algorithms are proven as a good predictor.

The current challenges to solve this problem are described below:

- What are the main algorithms that are best adapted to obsolescence forecasting?
- How can these approaches forecast a large number of data?
- How is the obsolescence data applicable in the model?

Proposed methodology

Throughout this research, the literature has showed how to evaluate the most promising approaches to obsolescence prediction based on selection criteria such as the capacity and stability of the algorithms. Based on these criteria, all the algorithms have been the subject of this research.
In fact, regardless the field of application, the learning stages are always the same:
Step 1 is the extraction of the data either by collecting data from the surveys or by extracting data from the available databases. Step 2 is to extrapolate the data by choosing the attributes and specifications of the data that the model will require. Step 3 is the distribution of the sample used to estimate the error. The sample will be split into 2 groups: 2/3 for training and 1/3 for the test. Note that it is possible to subdivide the data into 3 parts: learning, validation and testing. Step 4 often used as an adjustment sample, and it is used to adjust the parameters of the data mining methods. Note that the absence of a variable to explain, learning is unsupervised, otherwise learning is supervised.

Once the data is collected, the next step (step 5) the learning set will be introduced into the algorithms to create the predictive model. Then, the parameters of the model will be estimated, as the learning duration, the number of variables, etc. After that, supervised learning algorithm is developed to train the data and create predictive models. This algorithm must identify the status (active or obsolete) for each component in the test set. The precision of the algorithms is calculated by comparing the current state with the state predicted by the model. Finally, step 6 is the optimization of these parameters and the comparison of the optimal models will be concluded from the estimation of the forecast error for each model.

Case study

Data collection

The data is collected for the case study from micron site. Micron technology is a global leader in the semiconductor industry that provides information on the technical specifications, technology, Brand and the status of (production/ end-of-life) of different electronic components. The data contains more than 700 components with a known label, including many technical specifications show in figure 2 such as voltage (v), cycle time (ns), clock rate (HZ), depth (Mb), etc. Once the database is built, then the data is divided into two random groups divided into training (2/3) and test (1/3) to calculate the accuracy of the model. Figure-A I-1 shows an overview of the adopted methodology to build the forecasting models.

Once the data is constructed, after that the next step is the training of the data set using the five algorithms chosen and applied in the case study: BP-NN, SVM, AdaBoost, RF and Decision Trees. These algorithms have been proven as the most relevant in prediction.
K-means will not be studied in this case because we are in the case of supervised learning (output label is known), while K-means is based on unsupervised learning.

Data collection

First, we distinguish the predictive variables (INPUT = 13 variables) and the target variable (OUTPUT = end-of-life / production). 734 observations, of which 70% (513) are for the learning sample data and 30% (221) are for the test.
The first algorithm used was BP-NN. The classification of the neural networks was used with R3.3.3 using the “nnet” package. The neural networks used in this study were constructed with 3 hidden layers. The probability that each part is available or obsolete is considered as output. The following algorithm applied was SVM. This algorithm used the SVM classification function from the “e107” package with R3.3.3.

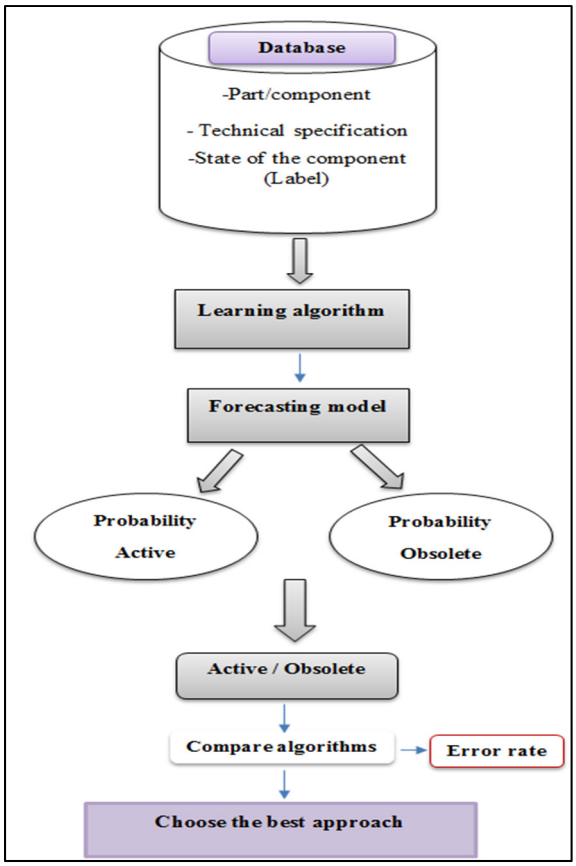

Figure-A I-1 Methodology adopted

The third algorithm was AdaBoost. The classification of this algorithm was done with R3.3.3 from the “ada” package that implements the boost algorithm.
After, RF was built with R3.3.3 from the 'random Forest' package. Type of random forest is classification, number of trees = 500, Number of variables tested at each split = 3
Finally, Decision trees algorithm was developed from the 'rpart' package.
The accuracy of the model was initially present by a confusion matrix (Table-A I-1), for training size of 70% and 30% for testing. The error for the model was retained to 21% (BP-NN), 22% (SVM), 25% (AdaBoost), 19% (RF) and 24% (Decision tree).
The confusion matrix shows how many components were correctly classified. The numbers in the cells (production, production) and (end-of-life, end-of-life) are correctly classified and all other cells present a classification error.

Table-A I-1 Confusion matrix

	Actual	End-of-life	Production	Overall Error
BP-NN	**End-of-life**	12	9	21%
	Production	16	60	
SVM	**End-of-life**	9	12	22%
	Production	9	67	
AdaBoost	**End-of-life**	53	26	25%
	Production	32	110	
RF	**End-of-life**	10	11	19%
	Production	7	69	
Decision tree	**End-of-life**	53	26	24%
	Production	27	115	

Results analysis

Table-A I-2 represents an average accuracy to evaluate model performance contain the prediction accuracy of the classification algorithms for different training data set sizes, from 50% to 90%, to see how much of the data portion influence the accuracy of the model.
The results show that RF shows better accuracy compared to the other algorithms at 70% of training size. Therefore, the prediction model created using SVM has such high precision for a 90% of training size. Note that all the algorithms show an increase of accuracy when the training size increases. Fig.5 shows a graphical display of confusion matrix of algorithms which shows that RF represents the best classifier with a slight difference with the other algorithms.

Table-A I-2 Average accuracy

	BB-NN		SVM		BOOST		RF		**Decision tree**	
Training size	Training (%)	Testing (%)	Training (%)	Testing (%)	Training (%)	Testing (%)	Training (%)	Testing (%)	Training (%)	Testing (%)
50%	94%	69%	87%	77%	90%	72%	94%	77%	85%	73%
60%	93%	70%	84%	74%	89%	73%	93%	77%	83%	77%
70%	91%	79%	85%	78%	89%	75%	91%	81%	82%	76%
80%	91%	76%	84%	78%	88%	78%	91%	84%	89%	80%
90%	90%	74%	92%	91%	89%	84%	90%	88%	90%	89%

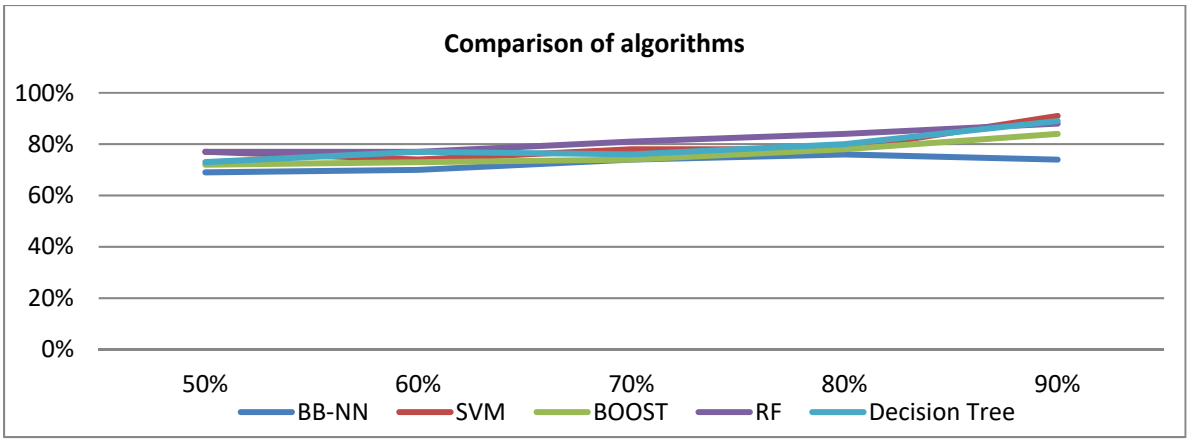

Figure-A I-2 Accuracy graph of algorithms

Limitations and future research

Research on obsolescence is growing, because of the serious challenge to complex systems. Machine learning methodology presented in this research can predict the risk of obsolescence with a high degree of accuracy without the need for human manipulation or market survey. Random Forest appears as the best predictor to forecasting obsolescence risk. However, the exploitation of the model can be carried out on new data.

Several aspects can be considered for future research. First, the sample is small, and several information are missing. The larger the data the more the model gives more accurate results. Information can be gathered from a larger sample to minimize the error rate.
The standard BP-NN has a major disadvantage: "Overfitting problem". This means that the network can lose its capability. It happens when the neural network learns the specific details of the inputs and not their general characteristic found in the present and future data.
To resolve this problem, the Bayesian-Multilayer Perceptron (B-MLP) will be developed. The Bayesian regularization method is used to reduce the error rate. This integration of the error function will require to Neural Network to obtain smaller weights and biases to reach a greater capacity of generalization. B-MLP will be compared to RF model to estimate the best predictor.

ANNEXE II

A RANDOM FOREST METHOD FOR OBSOLESCENCE FORECASTING

Yosra Grichi, Yvan Beauregard, Thien-My Do

Département de génie mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

Article présenté à la conférence « Industrial Engineering and Engineering Management (IEEM) », tenue à Singapore 2017

Abstract

Driven by the frequent technological changes and innovation, obsolescence has become a major challenge that cannot be ignored in which the life cycle of the components is often shorter than that of their systems. Basically, obsolescence problems are often sudden and not planned which causes delays and extra costs. On the other side forecasting appears to be one of the most efficient solutions to solve this problem.
This paper aims to provide new light and help industries to generate different solutions to the problems of obsolescence. Specifically, it presents a framework for forecasting the obsolescence based on random forest (RF) algorithm which has proven as the best predictor for forecasting obsolescence risk based on a previous comparative study with a high degree of accuracy.

Introduction

The rapidly evolving technologies have aggravated the obsolescence problem, which is a major issue that is noticed when an alternative component is no longer in stock or the supplier no longer produces it. Or because demand has dropped, OEMs are obliged to abandon manufacturing.
As defined by Moore's law, the rapid evolution of electronic components continues to grow, which stipulates that semiconductor density doubles approximately every 18 months [3-5]. In fact, when we talk about technological obsolescence, we immediately think of electronics. This evolution creates new electronic components every year with short lifetimes. In USA, the industry has grown at a rapid rate since the 1990s. New technologies are introduced in the market with increased speed. Today, the short life cycle of technology and the lack of forecasting represent a challenge for several companies which need to consider the risk of obsolescence. Moreover, the modeling of obsolescence is considered as a complex problem that requires a good knowledge about the components affected by obsolescence. In the same way, it is necessary to know the lifetime of the components and their technological evolutions

and therefore know all the factors that influence the obsolescence of components. On the other side, forecasting appears to be one of the most efficient solutions.
Obsolescence forecasting is one of the best solutions in the management obsolescence approach that assists manufacturers in identifying part obsolescence before it occurs and helping companies to enhance forecasting in order to ensure support for part/system in service. In the literature, most of the works are based on the life cycle forecasting for a single type part. Moreover, it is based on human knowledge to estimate obsolescence. To resolve this problem, machine learning appears as an optimal method for forecasting. In fact, machine learning is a data analysis method that automates the recognition of data models without human manipulation [6, 7]. It is forecast a large sample of electronic components with a high degree of accuracy. In previous work [8], we have presented a comparison of the different algorithms which are: artificial neural network (ANN), AdaBoost, support vector machine (SVM), random forest (RF) and decision tree that are capable of forecasting obsolescence risk of a large sample of electronic components to estimate the best predictor based on confusion matrix. However, Random Forest has appeared as the best predictor to forecast the obsolescence risk.
Therefore, this paper presents a model to predict obsolescence based on Random Forest algorithm, which can help manufacturers to make the best decision on their components and thus, increase the profit.

Literature review

Obsolescence can impact industries due to its negative effects conducted by functional, technological or logistical obsolescence [9]. The management of the obsolescence is traditionally, which is based on the resolution of the unavailability of a part facing obsolescence. Few studies are devoted to forecasting of this issue in proactive manner [10].
The obsolescence forecast can be divided into two types; the first one is the long-term forecasting (1 year or longer) that allows a proactive management and life cycle planning to support a system. While the second type of forecast is the short term. This type forecasts can be observed from the supply chain, which may include reducing the number of sources, reducing inventories, price increases that may in some cases be accompanied by a reduction in component availability [11].
To overcome the problems caused by obsolescence, many researches were generated in the past trying to create models that can forecast obsolescence risk effectively.

Obsolescence forecasting

Forecasting obsolescence was reactive in nature, which based on the resolution of the problem once noticed. The most classical reactive approaches are LTB (Last Time Buy) and exiting stock [12]. In the literature, there are two types of forecasting methods, which are forecasting of the obsolescence risk and forecasting of the obsolescence date (life cycle forecasting).

Forecasting of the obsolescence risk

It's used to predict the probability that a component is still in production or not [13-15]. According to the literature, a few studies have been devoted to the prediction of the risk of obsolescence. In this context, Rojo et al. [14] conducted a Delphi study to analyze the risk of obsolescence through a team of experts and developed a methodology to forecast the risk level of obsolescence based on obsolescence's indicators which are; years to end of life, number of sources available and consumption rate versus availability of the stock (more the stock is lower, more the risk of obsolescence is higher).
Another approach developed by Josias et al. [15] aims to create a risk index based on some variables like number of manufacturers and life cycle stage.

Life cycle forecasting

Solomon et al. [16] are the first to introduce the life cycle forecasting method. In their paper, the authors conducted a study to predict the obsolescence date from the life cycle curve, which includes six stages: introduction, growth, maturity, saturation, decline, and phase out. Other current method developed by Sandborn using data mining from a combination of sales data of parts and prediction using the Gaussian method [17]. Moreover, other authors have introduced regression analysis to predict the date of obsolescence [18].

Random forest

Random forests introduced by Breiman [19] are an integration of tree predictors so that every tree depend on the values of a random vector separately and through similar distribution for the whole trees in the forest. The tree classifier of a forest has a generalization error which relies on the strength correlation of each single tree in the forest [19].
Classification accuracy is improved significantly due to the increases of a group of trees and letting them choose the most popular class. A primary example is bagging [19] where to raise every tree an arbitrary selection (without replacement) is came from the training set examples. Further example is random split selection where arbitrarily the split is selected from among the K best splits at each single node.
Random Forest's algorithm consists of rotating a large number of decision trees randomly constructed and then generating them. Bootstrap sampling (OOB: Out-Of-Bag sampling) is used in RF to have a better estimate of the distribution of the original data set [19]. Indeed, bootstrapping means, rather than using all the data to build the trees, we can randomly select for each tree a subset of the data. In statistical terms, if the trees are uncorrelated, this reduces the forecast variance. The main advantage of random forests is their resistance to variances and biases.

Forecasting with random forest

The random forest algorithm is used in the regression case to predict a continuous dependence and classification variable in order to predict a categorical dependent variable.

For the regression type, random forest consists of a set of simple prediction trees; each can produce a numerical response when presented with a subset of explanatory variables or predictors.
For the classification type, A categorical variable with N modalities will be broken down into a disjunctive array (with N-1 variables) according to a 0-1 coding scheme. Thus, a categorical variable with N modalities can be considered as a set of N-1 variables, of which only one will assume the value 1 for a given observation. In fact, the ability to make predictions on a random subset of predictive variables is one of the strengths of the Random Forest module, which makes it particularly well suited to processing data sets with extremely high predictive variables. This random feature selection encourages systems diversity, and by the end it enhances classification performance. The random forest can be constructed by sampling arbitrarily a feature subset for every system, and/or by sampling arbitrarily a training data subset with regard to every system. The majority vote combines the final prediction. Finally, the random forest attains a favorable and vigorous performance with various applications [20].

Proposed framework

Throughout this research, the literature and previous study have showed how to evaluate the most promising approaches to obsolescence prediction based on selection criteria such as the accuracy of the algorithms. According to these criteria, Random Forest has been selected to be the subject of this research. Figure-A II-1 illustrates the research model showing the input and the output of this model. This research attempts to prove the accuracy of RF algorithm to forecast obsolescence.

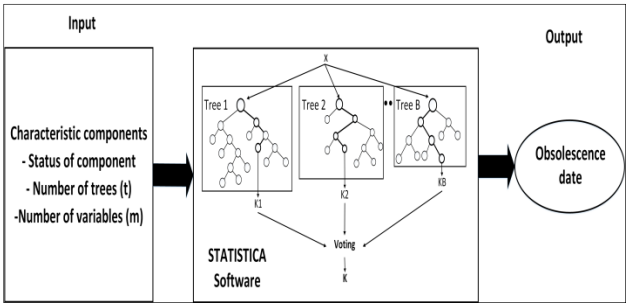

Figure-A II-1Research model

However, with the random forest method, we will model the obsolescence.
The training of RF follows several steps. First step is the extraction of the data either by collecting data from the available database. Next step is to extrapolate the data by choosing the attributes and specifications of the data that the model will require. In step 3, we randomly subdivide this set of observations into three subsets. The first subset consists of 70% of the data and we use it to form the models, this is called the learning set. While the second subset consists of 15% of the data and we use it to evaluate and validate our models. The third subset consists of 15% of the data and we use it to test our forecasting model.

Once the data is collected and the features are extracted, step 4 presents a sample data using Bootstrap which grow from the training data. After that, the learning set will be introduced into the RF to create the predictive model. Then, the parameters of the model will be estimated. For each terminal node of the tree, these steps will be repeated until the number of trees will be specified and the minimum node size is obtained.
Next step is the estimation of OOB error for the model and finally the output will be represented as an ensemble of trees (Tb).

Experimental studies

For RF model in obsolescence forecasting, three parameters are required. In fact, at each node, the draw of the variables is done without replacement and uniformly among all the p explanatory variables (each variable has a probability 1/p to be chosen). The number m ($m \leq p$) is fixed at the beginning of the forest construction and is identical for all trees. By default, the number ntree is set to 500 and mtry is equal to p/3, as suggested by Breiman.

Proposed model validation

A total of thousand experiment data of cell phones characteristics are used in this study, obtained from Connor et al. [7]. The data provides information such as brand, date of introduction and the status (production/ end-of-life) of different cellphones. Also, the data contains many technical specifications such as screen size, GPS, keyboard, etc. Some data used in this paper are categorical variables (yes or no), the other are continuous. These predictive variables present the input of RF. The target variable (dependent variable) is the obsolescence date.
The simulation of RF is divided in several steps. First, the data are pre-processed for cleaning and formatting. After that, the data was trained, validated and tested from the RF algorithm. The accuracy of the model was compared to another approach. In this case, STATISTICA software is used to simulate the model.
In the RF prediction process, the estimation is defined as the average of the whole of trees.
The parameter number of predictors determines the number of independent variables to be considered for each node at each tree. In STATISTICA, the optimal value of this parameter is given by log2 (M + 1) where M represents the number of predictors variables at input.
The simulation procedure of model parameters is used as follows:
B = 100 (number of trees)
m = 3 (Number of variables in each split)
nmin = 5 (Minimum node size)

Experimental results

Multiple Linear Regression (MLR) model is chosen to be compared to Random Forest in order to validate the accuracy of our model. The analysis below shows the results of the regression

model with the introduction date as the dependent variable (2). This analysis shows the coefficient of the regression model and the correlations between predictors.
The p-values of all variables are less than 0.01, with R = 75%. The criterion of selecting the predictors to enter the model was the strong relationship between them. In fact, additional analysis not shown herein indicates that there is a correlation between predictors.
The SPSS software was used for statistical analysis.

The regression equation is:

$$Y = 2010.07 + 0.9X_1 - 0.31X_2 + 0.34X_3 + 0.27X_4 + 0.13X_5 - 0.03X_6 - 0.55X_7 + 3.2E^{-5}X_8 + 0.73X_9 - 0.009X_{10}$$

With:
Y: Introduction date
X1 Headphone, X2 = MP3, X3 = Wav, X4 = Vibration, X5 = GPS, X6 = Java, X7 = Email, X8 = Resolution, X9 = Screen size, X10 = Wight.

Table-A II-1Analysis of variances

Model	DF	SS	MS	F	P
Regression	12	743.39	61.94	58.81	0.000
Residual Error	565	595.11	1.053		
Total	577	1338.5			

Table-A II-2 Standard error RF Vs. MLR

Model	RF		MLR
Standard Error	Train	Test	1.06
	0.05	0.15	

Conclusion

This paper presents a new random forest approach for predicting obsolescence. Compare to a benchmark model, RF was more accurate model with a high degree of accuracy presented by an estimate error of 0.15 (Table II). However, this model can help manufacturers especially for electronic industries that evolution has created new electronic components every year with short. Moreover, the exploitation of this model can be carried out on new data.
To improve RF algorithm, features that do not have a correlation can be removed from the training and the same optimal parameters can be repeated in order to obtain more accuracy rate of forecast. Naïve Bayes is widely used for probabilistic forecasting, characterized by its robustness and efficiency.

ANNEXE III

A NEW APPROACH FOR OPTIMAL OBSOLESCENCE FORECASTING BASED ON RANDOM FOREST TECHNIQUE AND METAHEURISTIC PARTICAL SWARM OPTIMIZATION

Yosra Grichi, Thien-My Do, Yvan Beauregard

Département de génie mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

Article présenté à la conférence « International Conference on Industrial Engineering and Operations Management », tenue à Paris, France, Juillet 2018

Abstract

Obsolescence is highly complex problem due to the influence of many factors such as competitive market pressure, technological advancement and short life cycle of technological components. Basically, obsolescence problems are often sudden and not planned, causing delays and extra cost. To overcome this problem, forecasting appears to be one of the most efficient solutions. Indeed, many studies have been conducted to create models that can effectively forecast obsolescence. In addition, applying machine learning techniques have attracted many attentions and have been widely used to predict obsolescence risk and life cycle. Popular algorithms such as random forest, has been reporting satisfactory performance. To improve the accuracy of machine learning algorithms for obsolescence forecasting, this paper proposes a new optimization approach for obsolescence forecasting based on random forest (RF) and Particle Swarm Optimization (PSO). In fact, parameters optimization and features selection of RF have an important effect on it is predictive accuracy and PSO presents one kind of effective method for RF parameters and features choosing. To examine the effectiveness of this approach, this paper presents a comparison between PSO-RF and GA-RF (random forest based on genetic algorithm). Experimental results show that PSO-RF outperformed GA-RF with 96% of accuracy.

Keywords
Obsolescence, forecasting, optimization, machine learning, random forest, particle swarm optimization

Introduction

Obsolescence issue occurs in systems that have a longer lifecycle than their components, such as in automotive, avionics, military, etc. The negative effects of obsolescence on the production performances have been studied in the literature and represent a major challenge in long term (F. R. Rojo et al., 2012; F. R. Rojo et al., 2010; P. Sandborn, 2013; P. Sandborn et al., 2011). Rapid technological is one of the factors that increase the rate of obsolescence. The electronic industry has emerged as the fastest growing sector and has spread widely around the world. As defined by Moore's law, the rapid evolution of electronic components continues to grow, which stipulates that semiconductor density doubles approximately every 18 months (Voller & Porté-Agel, 2002). This evolution creates new electronic components every year with short lifetimes. In the USA, the industry has grown at a rapid rate since the 1990s. New technologies are introduced in the market at increasing rates. Today, the short lifecycle and the lack of forecasting represent a challenge for several companies that need to take into account the risk of obsolescence. Moreover, obsolescence modeling is a complex problem that requires a good knowledge of parts. In the same way, it is necessary to know the lifetime of the components, their technological evolutions and, therefore, know all the factors that influence the obsolescence of components to understand how obsolescence may occur. These factors can include component criticality, technological watching, technological maturity, and the number of sources among others.

Obsolescence forecasting appears to be one of the best solutions in the obsolescence management as it assists manufacturers to identify part obsolescence. Through obsolescence forecasting, companies can ensure support for parts in service. There are two methodologies of obsolescence forecasting: long-term forecasting (1 year or longer), which allows a proactive management and life cycle planning to support a system, and short-term forecasting, which can be observed from the supply chain. Short term forecasting may involve reducing the number of sources, reducing inventories, and increasing the price (Bartels, Ermel, Sandborn, et al., 2012).

To overcome the problems caused by obsolescence, many studies have been conducted to create models that can effectively forecast obsolescence. Statistical methods such as regression, Partial least square regression (PLS), logistical regression and Gaussian method have been previously employed in many works (Gao et al., 2011; Jungmok & Namhun, 2017; Solomon et al., 2000). Indeed, machine learning has attracted the attention of many researchers in various disciplines, and has been applied in obsolescence recently (Foruzan, Scott, & Lin, 2015; C. Jennings et al., 2016a; Dazhong Wu, Jennings, Terpenny, Gao, & Kumara, 2017; X. Wu et al., 2008; Yun et al., 2010). Random forest algorithm is a kind of machine learning method that has been used in many areas and has shown a high degree of satisfactory classifications accuracies (Y Grichi et al., 2017). However, the performance of the classification may be reducing due to the irrelevant and redundant features in the dataset. In fact, not all of the features are useful for classification. Indeed, the optimization of the model's parameters and choosing the right features can maximize the classification accuracy.

In order to improve the accuracy of obsolescence forecasting, this study attempts to improve the classification accuracy rate of RF to better forecasting the obsolescence risk by developing an approach based on particle swarm optimization (PSO) and random forest (RF). PSO algorithm has applied to some machine learning algorithms such as, SVM, Neural network,

etc. However, PSO appears as an optimal approach compared with other optimization techniques like Genetic Algorithms (GAs) and Ant colony, which require a few parameters to adjust. However, the developed PSO-RF approach can optimize the parameter of RF as well as identify the right features extraction to improve the classification accuracy rate of RF. The accuracy of PSO-RF will be compared to GA-RF.

Literature review

Potential obsolescence forecasting strategies: Background

Forecasting obsolescence is reactive in nature and was based on the resolution of the problem once noticed. The most classical approaches include lifetime or last-time buy (F. R. Rojo et al., 2010). There are two types of forecasting methods, namely forecasting of the obsolescence risk and forecasting of the obsolescence date (life cycle forecasting). Obsolescence risk forecasting is used to predict the probability that a component still in production (C. Josias et al., 2004; F. R. Rojo et al., 2012). A few researchers focus on the prediction of the risk of obsolescence. In this context, Rojo et al. (2012) conducted a Delphi study to analyze the risk of obsolescence. They developed a risk using some indicators, which are years to end of life, the number of sources available, and the consumption rate versus availability of the stock. Another approach developed by Josias et al. (2004) aims to create a risk index by measuring the number of sources, life cycle stage (introduction, growth, maturity, decline, end of life), and market risk. (van Jaarsveld & Dekker, 2011a) developed a method based on historical demand data to estimate the risk of obsolescence. The risk of obsolescence was estimated based on Markov Chain. Last, (Y Grichi et al., 2017; C. Jennings et al., 2016a) have used data-driven method by create machine learning algorithms to forecast the obsolescence risk of a large number of parts. Alternatively, for life cycle forecasting, Solomon et al. (2000) were the first to introduce the life cycle forecasting method. In their paper, the researchers conducted a study to predict the life stage of apart from the life cycle curve, which included six stages: introduction, growth, maturity, saturation, decline, and obsolescence. Another method was developed by (P. Sandborn, 2007) using data mining with Gaussian method to predict the zone of obsolescence. This zone is given between $+2.5\sigma$ and $+3.5\sigma$ and gives time intervals for the period for the part will become obsolete. Moreover, other researchers have introduced regression analysis to predict the date of obsolescence (Gao et al., 2011).

Random forest

Introduced by Breiman (2001), Random forests are an integration of tree predictors where every tree depends on the values of a random vector separately (Breiman, 2001). A similar distribution applies for all the trees in the forest. The tree classifier of a forest has a generalization error which relies on the strong correlation between all trees in the forest. Classification accuracy increases significantly when the group of trees is enlarged. A primary example is bagging, where to raise every tree, an arbitrary selection (without replacement) is done from the set examples. Another example is random split selection where arbitrarily, the split is selected from among the K best splits at every single node. A random forest algorithm consists of rotating many decision trees that are randomly constructed and then generating

them. Bootstrap sampling (OOB: Out-Of-Bag sampling) is used in RF to have a better estimate of the distribution of the original data set. Indeed, bootstrapping means randomly selecting a subset of the data for each tree rather than using all the data to build the trees. In statistical terms, if the trees are uncorrelated, this reduces the forecast variance. The main advantage of random forests is their resistance to variances and biases. The random forest algorithm is used in the regression case to predict a continuous dependence and classification variable in order to predict a categorical dependent variable. For the regression type, a random forest consists of a set of simple prediction trees; each is capable of producing a numerical response when presented with a subset of explanatory variables or predictors. The error in this forecast is called Out Of Bag (OOB). For the classification type, a categorical variable with N modalities is broken down into a disjunctive array (with N-1 variables) according to a 0-1 coding scheme. Thus, a categorical variable with N modalities can be considered as a set of N-1 variables, of which only one will assume the value 1 for a given observation. In fact, the ability to make predictions on a random subset of predictive variables is one of the strengths of the Random Forest module, which makes it particularly well suited to processing data sets with extremely high predictive variables. This random feature selection encourages systems diversity, and by the end, it enhances classification performance. The random forest is constructed by sampling arbitrarily the features subset as well as the training subset with regard to every system. The majority vote combines the final prediction. Finally, the random forest attains a favorable and vigorous performance with various applications (Cheng et al., 2012; J. Friedman et al., 2001).

Particle swarm optimization

PSO called swarm intelligence or collective intelligence is developed by Eberhart and Kennedy in 1995 (Kennedy & Eberhart, 1995; Y. Shi, 2001). The overall behavior of the PSO is not programmed in advance but emerges from the sequence of elementary interactions between individuals. In this context, many researchers have applied the PSO in several machine learning for optimization (Lin et al., 2008; Xiaodan, 2017). The Optimization method of PSO can be iterative; each particle consists of changing the velocity toward its fitness value and global version of PSO. For the movement, the particle must decide on its next movement (its new speed) by linearly combining three pieces of information: its current speed V_{ij}^{n} (velocity), its best performance already found P_{ij}^{n} and which is known as the personal best position (*pbest*), and the best performance of its neighbors or informants P_{gj}^{n} known as the global best position (*gbest*) (See equation below). The iteration of the velocity and position of the particles are getting by following equations:

$$V_{ij}^{k+1} = w * \ V_{ij}^{k} + c_1 * \ r_1 \left(P_{ij}^{k} - x_{ij}^{k}\right) + \ c_2 * r_2 (P_{gj}^{k} - \ x_{ij}^{k})$$

$$x_{ij}^{k+1} = \ x_{ij}^{k} + V_{ij}^{k+1}$$

Where $P_{ij} = (p_{i1}, p_{i2}, \dots p_{im})$ and $P_{gj} = (p_{g1}, p_{g2}, \dots p_{gm})$. k is the number of iterations, x_{ij}^{k} present the particle position. Positive coefficient r_1 and r_2 are random number, generated uniformly in the range [0 1]. w presents the inertia coefficient of PSO algorithm. This weight is updated according to the following equation (H.-L. Chen et al., 2011):

$$w = w_{min} + (w_{max} - w_{min}) \frac{(t_{max} - t)}{t_{max}}$$

Where t_{max} presents the maximum number of iterations. Usually, the inertia coefficient is generated in the range [0.4 0.9] (H.-L. Chen et al., 2011). The coefficients c_1 and c_2 are calculated as follows:

$$c_1 = (c_{1f} - c_{1i}) \frac{t}{t_{max}} + c_{1i}$$

$$c_2 = (c_{2f} - c_{2i}) \frac{t}{t_{max}} + c_{2i}$$

Withc_{1f},c_{1i}, c_{2f} , and c_{2i} are a positive constant.
The position of the particles as well as their initial velocity must be initialized randomly according to a uniform law.
The original process for implementing the local version of PSO is as follows:
Step 1: Initialize randomly a population. Step 2: Measure the fitness of each particle in the population (calculate fitness score for each particle using selected features). Step 3: Update the velocity and position of each particle by looking for the best performance for each particle (local optimum). If the current fitness is better than the previous fitness, the previous *pbest* is replaced with the current *pbest*. Finally, in step 4: continues until the process converges. Stop the algorithm if the termination criterion is satisfied; otherwise, return to step 2.

Proposed PSO-RF optimization approach

In this paper, the PSO-RF model is developed and applies for forecasting of obsolescence risk. This main objective of this paper is to apply particle swarm optimization to enhance the classification performance of random forest algorithm by searching for the optimal parameters for RF and discovering the best subset of features as well.
The proposed PSO-RF optimization approach consists of 4 steps as follows:
Step 1: The initialization of the data processing. The dataset is initialized to construct the RF model based on supervised learning (known data), also the data is split into two groups (training and testing) randomly.
Step 2: The initialization of PSO and RF parameters. For the PSO, the number of generations, population size (number of particles) and so one, are initialized. The position and velocity set to x^0_i and v^0_i respectively, are determined as well. Each particle has (1) dimension (d) which is the length of features and the number of parameters to be optimized. (2) Position (x): position of the i^{th} particle. The initialization the parameters of RF (number of trees and the number of variables to split on at each node (mtry)) are included into the algorithm as well.
Step 3: PSO is adopted to construct PSO-RF model. In fact, PSO is looking for the optimal solution of particles by evaluation of fitness based on the update particle velocity and its position. However, if the current fitness is better than the previous fitness, the previous *pbest*

is replaced with the current *pbest* until to find the optimal solution (if current *gbest* is better than previous *gbest*, then replace *gbest* score and *gbest* particle. If the particle is already created and evaluated, then generate a new one. As introduced by Kennedy, the PSO is searching in a discrete space (0 or 1). Each feature in the PSO represents a binary bit (0 or 1), where 1 represents a selected feature while 0 represents a feature that is not selected. The features selected are based on the particle's position. **Step 4**: presents the training subset of features selection and parameters optimization is introduced in the RF model; therefore, the PSO-RF is obtained. For each terminal node of the tree, these steps are repeated until the specified number of trees is reached and the minimum node size is obtained. Next, the Out of Bag (OOB) error for the model is estimated. For classification, OOB error is estimated as the proportion of times that the categorical variable is not equal to the true class prediction. Finally, the output is represented as an ensemble of trees$\{T_b\}$. To make prediction at a new point x, let $\widehat{c_b}(x)$ be the class predicting of the b^{th} random forest tree. The equation is given as follows:

$$\widehat{C_{rf}^{b}} = majority\ vote\{\widehat{c_b}(x)\}.$$

Finally, the precision of the model is calculated by comparing the current state with the state predicted by the model using a confusion matrix. The fitness function is calculated as follows:

$$f = accuracy\ (\text{AUC})$$

The architecture of the proposed method is illustrated in Figure-A III-1.

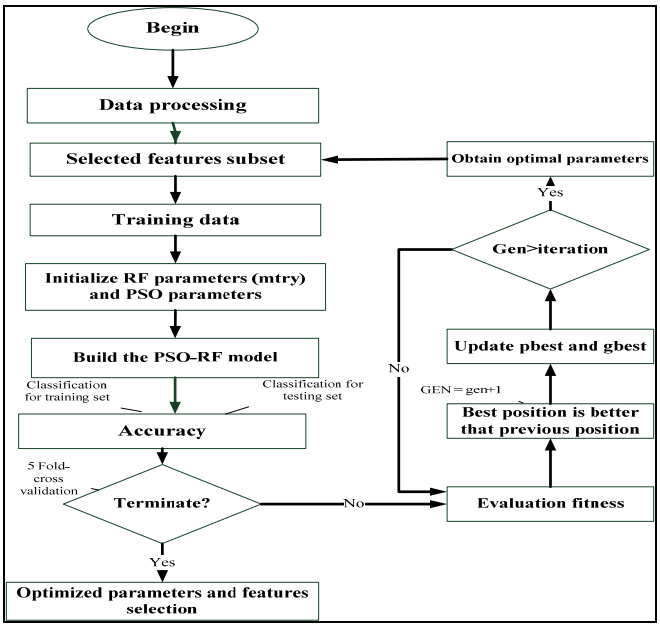

Figure-A III-1 Architecture of PSO-RF model

Numerical case study

The R programming is adopted to develop the PSO-RF. To measure the performance of PSO-RF approach, a dataset is used, taken from (C. Jennings et al., 2016a). A total of 999 instances provided information about cellphones and thirteen attributes were used as predictive variables. The output present two class (in production/end-of-life). About 70% of the data are randomly selected as the training set for constructing the model while 30% of the data are used as the test set to validate the model accuracy. To obtain a better estimation of classification accuracy, the K-Folds cross-validation method presented by (Salzberg, 1997) was applied and set to 5.

For features selection, we have introduced an assumption for the model: selected features should be greater or equal to 2. The parameters setting for PSO is obtained as follows: number of iteration and number of populations (particles) are set to 10 and 50 respectively. In fact, the iteration tries to generate new particle, however when there is so many particles which were already used, it will try to find other ones. If the solution is already near to the optimal, the result will not be change. Based on the experimental results, PSO is faster to find the solution with small iteration. As suggested by (Ratnaweera, Halgamuge, & Watson, 2004): $c_{1i} \leftarrow 2.5$, $c_{1f} \leftarrow 0.5$, $c_{2i} \leftarrow 0.5$, $c_{2f} \leftarrow 2.5$, $w_{max} \leftarrow 0.9$, $w_{min} \leftarrow 0.4$.

To examine the effectiveness of this approach, the PSO-RF is benchmarked with GA-RF (random forest based on genetic algorithm). The characteristic of GA is as follows: maximum generations, population, crossover and mutation are set to 20, 50, 0.8 and 0.1 respectively.

Experimental results and discussion

The results in this paper are described in terms of accuracy (AC), error rate, sensitivity (SE), specificity (SP), and Cohen's KAPPA, which are calculated by the following equations (Woods & Bowyer, 1997; Woods, Kegelmeyer, & Bowyer, 1997):

$$AC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$error\ rate = 1 - AC$$

$$SE = \frac{TP}{TP + FN}$$

$$SP = \frac{TN}{TN + FP}$$

$$K = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$

TP, TN, FP and *FN* are defined as true positive, true negative, false positive and false negative. For KAPPA equation, Pr(a) presents the probability of success of classification (accuracy) and Pr(e) presents the probability of success due to chance. In order to validate the accuracy of the proposed PSO-SVM algorithm, the results obtained by PSO-RF is compared with the GA-RF

(RF with genetic algorithm) developed by (Y Grichi, Yvan Beauregard, & Thien-My Dao, 2018). The classification accuracy rates of GA-RF are cited from their original papers that was achieved a good predictive performance. The accuracy of PSO-RF, GA-RF and RF was initially presented by a confusion matrix (See Table 1) for testing set. The comparisons of the algorithms are shown in Table-A III-1 and Table-A III-2.

Table-A III-1 Confusion matrix

Algorithms	Predict		
	Actual	Available	Discontinued
GA-RF	Available	**166**	12
	Discontinued	8	**113**
PSO-RF	Available	**170**	9
	Discontinued	4	**116**
RF	Available	**163**	15
	Discontinued	11	**110**

Table-A III-2 Accuracy measures (testing sample)

Accuracy measure	RF (%)	GA-RF (%)	PSO-RF (%)
Accuracy	91.3	93.3	**96**
No Information Rate	58.2	58.2	58.2
Kappa	82.1	86.2	**91**
Sensitivity	88	90.4	**97.7**
Specificity	93.7	**95.4**	92.8
Error rate	8.7	6.7	**4**
Balanced Accuracy ((Sensitivity+ Specificity)/2)	90.8	92.9	**95.25**

Table-A III-3 Accuracy measures (training sample)

Accuracy measure	RF	GA-RF	PSO-RF
Accuracy	94.7	**98.6**	**98.6**
No Information Rate	58.3	58.3	0.583
Kappa	89.1	97.1	97
Sensitivity	91.4	98	**98.8**
Specificity	97.1	**99**	98.3
Error rate	5.3	**1.4**	**1.4**
Balanced Accuracy	94.3	98.5	**98.55**

The result obtained from the PSO-RF model proves that the prediction through parameters optimization and choosing the best small features can improve significantly the classification accuracy of the RF. The experiments results were compared with RF-PSO and RF. PSO-RF approach yielded a higher classification accuracy rate compared to the other approaches. Thus, PSO-RF yielded more appropriate subset with few iterations. Figure-A III-2 shows the Receiver Operator Characteristic curve, that presents the false positive rate vs. true positive rate. These curves present as follows: green, blue, orange for PSO-RF, RF and GA-RF respectively.

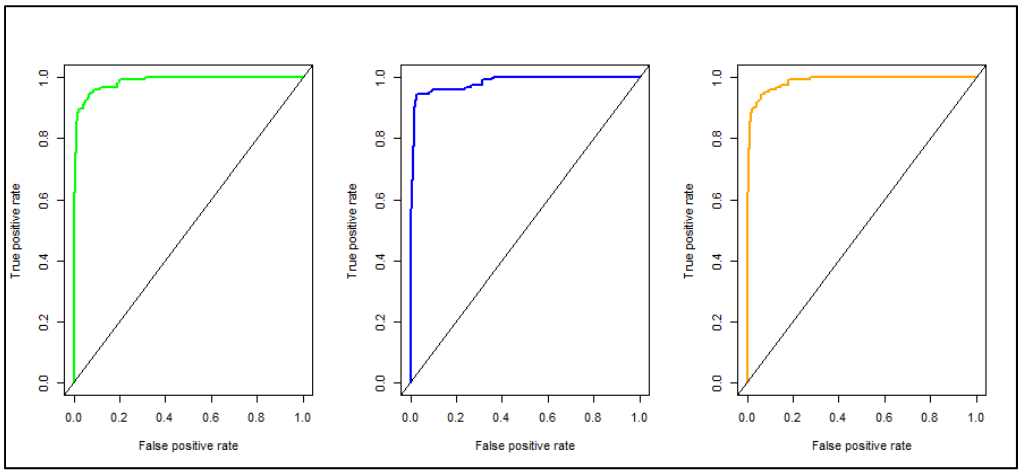

Figure-A III-2 ROC curve

The use of feature selection and parameter optimization were found to improve the classification accuracy rate for random forest to improve the forecasting of obsolescence risk. Experimental results show that PSO-RF has better performance than that of GA-RF.

Conclusion

This paper presents an improved approach for obsolescence forecasting risk with a high degree of accuracy based on machine learning and meta-heuristic PSO. PSO search for the optimal parameter value for RF to obtain a subset of beneficial features. The optimal set features were adopted for the training and testing of RF model to improve the classification accuracy of the model. In order to validate this approach, PSO-RF was compared to RF and GA-RF. Experimental results show that PSO-RF outperformed GA-RF with 96% of accuracy.
For future work, larger data with more features can give more accurate results. Other datasets and real-world problems for obsolescence forecasting can be tested using this approach.
Other optimization algorithms can also be used, such as ant colony, which is widely used for optimization, and compare it with the existing approach.

ANNEXE IV

QUESTIONNAIRE

This survey was written by Yosra Grichi and Héloïse Conrad, thesis and master students at the École de Technologie Supérieure (ETS) in Montreal.
This questionnaire was approved by the Research Ethics Committee (REC) of the ETS.

Confidentiality:

In accordance with the recommendations of the ethics committee, we inform you that no name of respondent or company will be disclosed. The data and results will be treated anonymously. The results will be communicated by email, providing to your company an overview of habits and practices of the aviation industry concerning forecasting and managing of obsolescence.
Please, complete the survey and return it by e-mail. You can **directly** answer in the survey (Word), so you do not need to print.

Company name :		
Name of the respondent :		
Function of the respondent :		
Seniority of the respondent :		
E-mail :		
Telephone :		

Context:

This survey aims to know the habits of company in forecasting technological obsolescence, in terms of organization, procedures, time management and relationship with suppliers.

Definitions:

Obsolescence here stands for: an obsolete component is a component that is no longer available from the usual supplier (production stop due to changes in technology, have suspended business activities or other)
Suppliers are here suppliers of electronic components or avionics subsystems.
The term **component** refers here, depending on your company:

- Basic electric or electronic components,
- Electronic sub-systems already assembled.

Questionnaire:

Part 1 – In your company

1. Has your company had one or more obsolescence problems?
   ☐ **Yes** ☐ **No**

2. What is the frequency of occurrence of obsolescence issues?
   - ☐ Less than 1 per month – Precise the number of cases per year:
   - ☐ 1 to 2 cases per month
   - ☐ 3 to 5 cases per month
   - ☐ 5 to 10 cases per month
   - ☐ More than 10 cases per month
3. What is the major factor do you usually take into consideration to predict obsolescence?
   - ☐ Relationship with suppliers
   - ☐ Number of sources
   - ☐ Technological maturity
   - ☐ Stock and consumption rate ratio
   - ☐ Technological Watch
   - ☐ Proactive measures
   - ☐ Criticality of component
   - ☐ I don't know
   - ☐ Other:

Obsolescence management

4. What is your company's preferred approached to deal with obsolescence issues?
   - ☐ Reactive approach (Method for managing the problem of obsolescence once it has appeared):
     - ☐ « Last Time Buy » (=final order purchase)
     - ☐ Find an equivalent (same "form, fit and function")
     - ☐ Find an alternative (different performance)
     - ☐ Re-design
     - ☐ Emulation (make the component you need on your own)
   - ☐ Proactive approach (Methods to anticipate the occurrence of obsolescence issues):
     - ☐ « Risk mitigation Buy »
     - ☐ Contracting with suppliers (to ensure a continuous supply)
     - ☐ Design for obsolescence
   - ☐ Other:
5. Does your company employ obsolescence managers?
   ☐ Yes ☐ No ☐ I don't know
6. Does your company use an obsolescence monitoring software?
   ☐ Yes ☐ No ☐ I don't know

7. Does your company use technology roadmapping, indicating what technologies to use in short and long-term to reach business goals?
   ☐ Yes ☐ No ☐ I don't know
8. Do you apply a component obsolescence plan strategy on the long term?
   ☐ Yes ☐ No ☐ I don't know
9. If YES, what obsolescence forecasting method does your company use?
   ☐ Qualitative:
      ☐ Delphi
      ☐ Market survey
      ☐ Other:
   ☐ Quantitative:
      ☐ Mathematic models – Explain:
      ☐ Other:
   ☐ Other:
   ☐ I don't know
10. If YES, how do you evaluate this method?

Poor	Passable	Correct	Efficient	Excellent
☐	☐	☐	☐	☐

11. Does your company consider obsolescence at the design phase (Design for Obsolescence)?
    ☐ Yes ☐ No ☐ I don't know
12. The standardization of avionic subsystems and modular architecture make it easier to deal with obsolescence by facilitating the replacement of all components and subsystems.
    Does your company work on the standardization of avionic components or subsystems?
    ☐ Yes ☐ No ☐ I don't know
13. If YES, does your company involve its suppliers in the standardization project?
    ☐ Yes ☐ No ☐ I don't know

<u>Stock and consumption rate</u>

14. Does your company have stock for most of the components you use?
    ☐ Yes ☐ No
15. If YES, this stock is calculated to support maintenance and production for:
    ☐ 1-3 months
    ☐ 4-6 months
    ☐ 7-12 months
    ☐ 1-2 years

☐ ≥2 years

☐ I don't know

16. Are there extra stocks for critical components compared to standard components stocks?
    ☐ Yes ☐ No ☐ I don't know
17. Indicate the level of component consumption compared to the available stock:
    ☐ Low: High stock and low consumption rate

    ☐ Medium: Low stock and low consumption rate -OR- High stock and high consumption rate

    ☐ High: Low stock and high consumption rate

Technology maturity

18. Do you have information on the age of technologies used to build critical components?
    ☐ Yes ☐ No ☐ I don't know
19. If yes, does your company use this information to better predict component obsolescence, know which ones are at risk?
    ☐ Yes ☐ No ☐ I don't know

Criticality of component

A **component criticality** refers to the impact a malfunction of the component has on the system it is implemented in.

20. Concerning obsolescence forecasting, does your company proceed differently for critical components compared to standard components?
    ☐ Yes ☐ No ☐ I don't know
21. If YES, according to your company, what are the criteria for a component to have a particular treatment?
    ☐ Very complex component

    ☐ Custom-made component

    ☐ Expensive component

    ☐ Critical component: requires a lot of certification and qualification tests

    ☐ Other:

22. If YES, what does the specific treatment for critical components consist in?
    ☐ Informal watch with suppliers to gather information on the state and the future of the production

    ☐ Increased stock monitoring

    ☐ Other:

Part 2 – Supply chain and suppliers

Relationship with suppliers

23. How would you qualify the relationship between your company and its suppliers in terms of quality of communication?

Poor	Passable	Correct	Efficient	Excellent
☐	☐	☐	☐	☐

24. For the majority of the components, how many suppliers do you have?
    - ☐ One supplier
    - ☐ Two suppliers
    - ☐ Three or more suppliers

25. Is there a formal process with suppliers to forecast obsolescence?
    - ☐ Yes, for all suppliers
    - ☐ Yes for some of suppliers
    - ☐ No
        a. If YES, since how long has it been set up ?
            - ☐ 1 year or less
            - ☐ 2-5 years
            - ☐ 5-9 years
            - ☐ 10+ years
        b. If YES, the process: ☐ Implies the supplier.
            ☐ Is internal to your company.
        c. If YES, does the process indicate (check the right answers):
            - ☐ The mean of communication
            - ☐ The communication frequency
            - ☐ The person to contact
                - ☐ In your company
                - ☐ From your supplier
            - ☐ A minimum period between the obsolescence notice given by the supplier and the actual production stop

26. How does your supplier usually notice you for a component obsolescence?
    - ☐ By e-mail
    - ☐ By post mail
    - ☐ By telephone

☐ By fax

☐ Other:

27. Concerning obsolescence notices, are there regular contacts:
    In your company: ☐ Yes ☐ No ☐ I don't know

    Function:

    Department:

    At your supplier's: ☐ Yes ☐ No ☐ I don't know

    Function:

    Department:

28. What is the time margin between the time you receive the obsolescence notice and the production stop of this component?
    Usually: months

    Exceptional cases (the shortest margin ever met): months

29. When your supplier sends you an obsolescence notice for a component, does he inform you about:
    a. A maximum date to order a « Last Time Buy » :
       ☐ Yes ☐ No ☐ I don't know
    b. A suggested alternative replacement components (that may perform the same functions as the old component but with some different characteristics) :
       ☐ Yes ☐ No ☐ I don't know

30. Does your major supplier inform you about:

    a. The state of your components (that is to say whether the components you order and use are in the expanding, maturity or declining phases) :

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

    b. Changes in technology (that is to say whether the technologies used in components or subsystems are still valid or will be replaced with a more innovative technology) :

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

    c. The evolution of the production plan based on the emerging technologies in the short-run (from now on to 5 years):

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

d. The supplier's long-term strategy (10-20 years) :

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

31. Does your firm inform your major supplier about:

a. New products to be designed and produced:

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

b. Changes in purchase order:

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

c. Delivery delay:

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

d. Inventory level:

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

e. Production planning:

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

f. Future demand forecasting:

Never	Annually	Quarterly	Monthly	Weekly	Daily
☐	☐	☐	☐	☐	☐

32. This question aims at determining to what extent does your company use technology (hardware and software) to support information exchange with your suppliers.
Does your company use:

- ☐ Electronic Data Interchange (EDI) system
- ☐ Enterprise Resource Planing (ERP) systems
- ☐ Forecast or demand management system
- ☐ Other- Explain:

Technological Watch

Technological watch is about dedicating full time or partial time resources to monitor the evolution of technologies to update your company's knowledge and inform the design department.

33. **Does your company carries out technology watch with suppliers and competitors?**
    - ☐ Yes
    - ☐ No
    - ☐ Other, explain:
    - ☐ I don't know
34. How does your company select its suppliers?
    Explain the main criteria:

BIBLIOGRAPHIE

Adetunji, O., Bischoff, J., & Willy, C. J. (2018). Managing system obsolescence via multicriteria decision making. *Systems Engineering, 21*(4), 307-321. doi:10.1002/sys.21436

Ahmad, A., & Khan, S. S. (2019). Survey of state-of-the-art mixed data clustering algorithms. *IEEE Access, 7*, 31883-31902.

Ansari, Z., Azeem, M., Ahmed, W., & Babu, A. V. (2015). Quantitative evaluation of performance and validity indices for clustering the web navigational sessions. *arXiv preprint arXiv:1507.03340*.

Asgharbeygi, N., & Maleki, A. (2008). *Geodesic k-means clustering.* Paper presented at the 2008 19th International Conference on Pattern Recognition.

Babai, M. Z., Dallery, Y., Boubaker, S., & Kalai, R. (2019). A new method to forecast intermittent demand in the presence of inventory obsolescence. *International Journal of Production Economics, 209*, 30-41. doi:10.1016/j.ijpe.2018.01.026

Bartels, B., Ermel, U., Pecht, M., & Sandborn, P. (2012a). Case Study Hardware Forecasts and Trends. In *Strategies to the Prediction, Mitigation and Management of Product Obsolescence* (pp. 77-142): John Wiley & Sons, Inc.

Bartels, B., Ermel, U., Pecht, M., & Sandborn, P. (2012b). Reactive Obsolescence Management. In *Strategies to the Prediction, Mitigation and Management of Product Obsolescence* (pp. 157-191): John Wiley & Sons, Inc.

Bartels, B., Ermel, U., Sandborn, P., & Pecht, M. G. (2012). *Strategies to the prediction, mitigation and management of product obsolescence* (Vol. 87): John Wiley & Sons.

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

Budiaji, W., & Leisch, F. (2019). Simple K-Medoids Partitioning Algorithm for Mixed Variable Data. *Algorithms, 12*(9), 177.

Chen, H.-L., Yang, B., Wang, G., Liu, J., Xu, X., Wang, S.-J., & Liu, D.-Y. (2011). A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. *Knowledge-Based Systems, 24*(8), 1348-1359.

Chen, K.-Y., & Wang, C.-H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management, 28*(1), 215-226.

Cheng, Y.-Y., Chan, P. P., & Qiu, Z.-W. (2012). *Random forest based ensemble system for short term load forecasting.* Paper presented at the Machine Learning and Cybernetics (ICMLC), 2012 International Conference on.

Conrad, H. (2016). *Élaboration d'un plan de transition et de mise en oeuvre pour améliorer la gestion de l'obsolescence dans une entreprise du secteur aéronautique.* École de technologie supérieure,

Dianati, M., Song, I., & Treiber, M. (2002). *An introduction to genetic algorithms and evolution strategies*. Retrieved from

Du, J., Liu, Y., Yu, Y., & Yan, W. (2017). A prediction of precipitation data based on support vector machine and particle swarm optimization (PSO-SVM) algorithms. *Algorithms, 10*(2), 57.

Fortin, F., & Gagnon, J. (2010). *Fondements et étapes du processus de recherche : méthodes quantitatives et qualitatives*. Montréal: Chenelière éducation.

Foruzan, E., Scott, S. D., & Lin, J. (2015). *A comparative study of different machine learning methods for electricity prices forecasting of an electricity market.* Paper presented at the 2015 North American Power Symposium (NAPS), 4-6 Oct. 2015, Piscataway, NJ, USA.

Friedman, H. P., & Rubin, J. (1967). On some invariant criteria for grouping data. *Journal of the American Statistical Association, 62*(320), 1159-1178.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1): Springer series in statistics Springer, Berlin.

Gao, C., Liu, X., & Wang, X. (2011). *A model for predicting the obsolescence trend of FPGA.* Paper presented at the 2011 9th International Conference on Reliability, Maintainability and Safety: Safety First, Reliability Primary, ICRMS'2011, June 12, 2011 - June 15, 2011, Guiyang, China.

Glover, F. (1986). Future paths for integer programming and links to ar tifi cial intelli g en ce. *Computers operations research, 13*(5), 533-549.

Gonzalez, R. T., Padilha, C. A., & Barone, D. A. C. (2015). *Ensemble system based on genetic algorithm for stock market forecasting.* Paper presented at the 2015 IEEE Congress on Evolutionary Computation (CEC).

Gravier, M. J., & Swartz, S. M. (2009). The dark side of innovation: Exploring obsolescence and supply chain evolution for sustainment-dominated systems. *Journal of High Technology Management Research, 20*(2), 87-102. doi:10.1016/j.hitech.2009.09.001

Grichi, Y., Beauregard, Y., & Dao, T.-M. (2018). Optimization of Obsolescence Forecasting Using New Hybrid Approach Based on the RF Method and the Meta-heuristic Genetic Algorithm. *American Journal of Management, 18*(2), 27-38.

Grichi, Y., Beauregard, Y., & Dao, T.-M. (2019). An approach to obsolescence forecasting based on Hidden Markov Model and Compound Poisson Process. *International Journal of Industrial Engineering, 1*(2), 111-124.

Grichi, Y., Beauregard, Y., & Dao, T. (2017). *A random forest method for obsolescence forecasting.* Paper presented at the Industrial Engineering and Engineering Management (IEEM), 2017 IEEE International Conference on.

Grichi, Y., Dao, T.-M., & Beauregard, Y. (2018). A new approach for optimal obsolescence forecasting based on the random forest (RF) technique and meta-heuristic particle swarm optimization (PSO).

Guoqiang, Z., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting, 14*(1), 35-62. doi:10.1016/S0169-2070(97)00044-7

Handl, J., Knowles, J., & Kell, D. B. (2005). Computational cluster validation in post-genomic data analysis. *Bioinformatics, 21*(15), 3201-3212.

Hassan, S., Khosravi, A., & Jaafar, J. (2013). *Bayesian Model Averaging of Load Demand Forecasts from Neural Network Models.* Paper presented at the 2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2013), 13-16 Oct. 2013, Los Alamitos, CA, USA.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 75-105.

Hippert, H. S., & Taylor, J. W. (2010). An evaluation of Bayesian techniques for controlling model complexity and selecting inputs in a neural network for short-term load forecasting. *Neural Networks, 23*(3), 386-395. doi:http://dx.doi.org/10.1016/j.neunet.2009.11.016

Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*: University of Michigan Press.

Homchalee, R., & Sessomboon, W. (2013). *A comparison of forecasting models using multiple regression and artificial neural networks for the supply and demand of Thai ethanol.*

Paper presented at the 2013 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 10-13 Dec. 2013, Piscataway, NJ, USA.

Hu, G., & Bidanda, B. (2009). Modeling sustainable product lifecycle decision support systems. *International Journal of Production Economics, 122*(1), 366-375.

Huang, C.-M., Romero, J. A., Osterman, M., Das, D., & Pecht, M. (2019). Life cycle trends of electronic materials, processes and components. *Microelectronics Reliability, 99*, 262-276. doi:10.1016/j.microrel.2019.05.023

Huang, J. Z., Ng, M. K., Rong, H., & Li, Z. (2005). Automated variable weighting in k-means type clustering. *IEEE transactions on pattern analysis and machine intelligence, 27*(5), 657-668.

Huang, Z. (1998). Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery, 2*(3), 283-304.

Jennings, C., & Terpenny, J. P. (2015). Taxonomy of Factors for Lifetime Buy. *IIE Annual Conference. Proceedings*, 430-436. Retrieved from http://search.proquest.com/docview/1791989172?accountid=27231

Jennings, C., Wu, D., & Terpenny, J. (2016a). Forecasting obsolescence risk and product life cycle with machine learning. *IEEE Transactions on Components, Packaging and Manufacturing Technology, 6*(9), 1428-1439. doi:10.1109/TCPMT.2016.2589206

Jennings, C., Wu, D., & Terpenny, J. (2016b). *Forecasting obsolescence risk using machine learning.* Paper presented at the ASME 2016 11th International Manufacturing Science and Engineering Conference, MSEC 2016, June 27, 2016 - July 1, 2016, Blacksburg, VA, United states.

Jennings, C., Wu, D., & Terpenny, J. (2017). *Receiver operating characteristic analysis for forecasting obsolescence risk.* Paper presented at the 67th Annual Conference and Expo of the Institute of Industrial Engineers 2017, May 20, 2017 - May 23, 2017, Pittsburgh, PA, United states.

Jennings, C. P. (2015). *Forecasting obsolescence risk and product lifecycle with machine learning.* Iowa State University,

Jilani, T. A., Amjad, U., Jaafar, J., & Hassan, S. (2012). *An improved heuristic-based fuzzy time series forecasting model using genetic algorithm.* Paper presented at the 2012 International Conference on Computer & Information Science (ICCIS), 12-14 June 2012, Piscataway, NJ, USA.

Jinglin, D., Yayun, L., Yanan, Y., & Weilan, Y. (2017). A Prediction of Precipitation Data Based on Support Vector Machine and Particle Swarm Optimization (PSO-SVM) Algorithms. *Algorithms, 10*(2), 57 (15 pp.). doi:10.3390/a10020057

Josias, C., Terpenny, J. P., & McLean, K. J. (2004). *Component obsolescence risk assessment.* Paper presented at the IIE Annual Conference and Exhibition 2004, May 15, 2004 - May 19, 2004, Houston, TX, United states.

Josias, C. L. (2009). Hedging future uncertainty: A framework for obsolescence prediction, proactive mitigation and management.

Jungmok, M., & Namhun, K. (2017). Electronic part obsolescence forecasting based on time series modeling. *International Journal of Precision Engineering and Manufacturing, 18*(5), 771-777. doi:10.1007/s12541-017-0092-6

Kennedy, J. (2010). Particle Swarm Optimization. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 760-766). Boston, MA: Springer US.

Kennedy, J., & Eberhart, R. (1995). *PSO optimization.* Paper presented at the Proc. IEEE Int. Conf. Neural Networks.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *science, 220*(4598), 671-680.

Kourentzes, N. (2013). Intermittent demand forecasts with neural networks. *International Journal of Production Economics, 143*(1), 198-206. doi:10.1016/j.ijpe.2013.01.009

Lin, S.-W., Ying, K.-C., Chen, S.-C., & Lee, Z.-J. (2008). Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert Systems with Applications, 35*(4), 1817-1824.

Livingston, H. (2000). *GEB1: Diminishing manufacturing sources and material shortages (DMSMS) management practices.* Paper presented at the Proceedings of the DMSMS Conference.

Lobo, F. G., Goldberg, D. E., & Pelikan, M. (2000). *Time complexity of genetic algorithms on exponentially scaled problems.* Paper presented at the Proceedings of the 2nd Annual Conference on Genetic and Evolutionary Computation.

Macas, M., Moretti, F., Lauro, F., Pizzuti, S., Annunziato, M., Fonti, A., . . . Giantomassi, A. (2014). *Importance of Feature Selection for Recurrent Neural Network Based Forecasting of Building Thermal Comfort.* Paper presented at the Adaptive and Intelligent Systems Third International Conference (ICAIS 2014), 8-10 Sept. 2014, Cham, Switzerland.

Mellal, M. A. (2020). Obsolescence A review of the literature. *Technology in Society, 63*. doi:10.1016/j.techsoc.2020.101347

Memmedli, M., & Ozdemir, O. (2011). *Forecasting neural network-based fuzzy time series with different neural network models.* Paper presented at the 11th WSEAS International Conference on Signal Processing, Computational Geometry and Artificial Vision, ISCGAV'11, 11th WSEAS International Conference on Systems Theory and Scientific Computation, ISTASC'11, August 23, 2011 - August 25, 2011, Florence, Italy.

Mladenović, N., & Hansen, P. (1997). Variable neighborhood search. *Computers & operations research, 24*(11), 1097-1100.

O'Dowd, R. J. (2010). *A Survey of Electronics Obsolescence and Reliability.*

Pan, Y., Xue, W., Zhang, Q., & Zhao, L. (2011). *A forecasting model of RBF neural network based on genetic algorithms optimization.* Paper presented at the 2011 Seventh International Conference on Natural Computation.

Park, H.-S., & Jun, C.-H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications, 36*(2), 3336-3341.

Parvin, A. J., & Beruvides, M. G. (2017). *Forecasting technology obsolescence: Assessing the existing literature, a systematic review.* Paper presented at the 2017 International Annual Conference of the American Society for Engineering Management, ASEM 2017, October 18, 2017 - October 21, 2017, Huntsville, AL, United states.

Pince, C., & Dekker, R. (2011). An inventory model for slow moving items subject to obsolescence. *European Journal of Operational Research, 213*(1), 83-95. doi:10.1016/j.ejor.2011.02.013

Pinçe, Ç., & Dekker, R. (2011). An inventory model for slow moving items subject to obsolescence. *European Journal of Operational Research, 213*(1), 83-95. doi:10.1016/j.ejor.2011.02.013

Ratnaweera, A., Halgamuge, S. K., & Watson, H. C. (2004). Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. *IEEE Transactions on evolutionary computation, 8*(3), 240-255.

Ren, M., Liu, P., Wang, Z., & Pan, X. (2016). *An improved mixed-type data based kernel clustering algorithm.* Paper presented at the 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD).

Rojo, F. J. R., Baguley, P., Shaikh, N., Roy, R., & Kelly, S. (2012). TOMCAT: An Obsolescence Management Capability Assessment Framework. *Journal of Physics:*

Conference Series, 364(1), 012098. Retrieved from http://stacks.iop.org/1742-6596/364/i=1/a=012098

Rojo, F. J. R., Roy, R., Shehab, E., & Cheruvu, K. (2012). A study on obsolescence resolution profiles. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 226*(1), 167-177. doi:10.1177/0954405411407565

Rojo, F. R., Roy, R., & Kelly, S. (2012). *Obsolescence risk assessment process best practice.* Paper presented at the Journal of Physics: Conference Series.

Rojo, F. R., Roy, R., & Shehab, E. (2010). Obsolescence management for long-life contracts: state of the art and future trends. *International Journal of Advanced Manufacturing Technology, 49*(9-12), 1235-1250. doi:10.1007/s00170-009-2471-3

Rokach, L., & Maimon, O. (2007). Introduction of Decision trees. *Data Mining with Decision Trees: theory and application*, 5.

Rousseeuw, P. J., & Kaufman, L. (1990). Finding groups in data. *Hoboken: Wiley Online Library*.

Salzberg, S. L. (1997). On Comparing Classifiers: Pitfalls to Avoid and a Recommended Approach. *Data Mining and Knowledge Discovery, 1*(3), 317-328. doi:10.1023/a:1009752403260

Sandborn, P. (2007). Software obsolescence-Complicating the part and technology obsolescence management problem. *IEEE Transactions on Components and Packaging Technologies, 30*(4), 886-888.

Sandborn, P. (2008). Strategic management of DMSMS in systems. *DSP Journal*, 24-30.

Sandborn, P. (2013). Design for Obsolescence Risk Management. *Procedia CIRP, 11*(0), 15-22. doi:http://dx.doi.org/10.1016/j.procir.2013.07.073

Sandborn, P. (2017). Forecasting technology and part obsolescence. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 231*(13), 2251-2260. doi:10.1177/0954405415598923

Sandborn, P., Prabhakar, V., & Ahmad, O. (2011). Forecasting electronic part procurement lifetimes to enable the management of DMSMS obsolescence. *Microelectronics Reliability, 51*(2), 392-399. doi:10.1016/j.microrel.2010.08.005

Sandborn, P. A., Mauro, F., & Knox, R. (2007). A data mining based approach to electronic part obsolescence forecasting. *IEEE Transactions on Components and Packaging Technologies, 30*(3), 397-401. doi:10.1109/TCAPT.2007.900058

Sandborn, P. S. a. P. (2002). Determining Optimum Redesign Plans for Avionics Based on Electronic Part Obsolescence Forecasts.

Shen, Y., & Willems, S. P. (2014). Modeling sourcing strategies to mitigate part obsolescence. *European Journal of Operational Research, 236*(2), 522-533. doi:http://dx.doi.org/10.1016/j.ejor.2014.01.025

Shi, Y. (2001). *Particle swarm optimization: developments, applications and resources.* Paper presented at the evolutionary computation, 2001. Proceedings of the 2001 Congress on.

Shi, Z., & Liu, S. (2020). Optimal inventory control and design refresh selection in managing part obsolescence. *European Journal of Operational Research, 287*(1), 133-144. doi:10.1016/j.ejor.2020.04.038

Singh, A., Yadav, A., & Rana, A. (2013). K-means with Three different Distance Metrics. *International Journal of Computer Applications, 67*(10).

Solomon, R., Sandborn, P. A., & Pecht, M. G. (2000). Electronic part life cycle concepts and obsolescence forecasting. *IEEE Transactions on Components and Packaging Technologies, 23*(4), 707-717. doi:10.1109/6144.888857

Song, J.-S., & Zipkin, P. H. (1996). Managing inventory with the prospect of obsolescence. *Operations Research, 44*(1), 215-222.

Tomczykowski, W. J. (2003). *A study on component obsolescence mitigation strategies and their impact on RM.* Paper presented at the The International Symposium on Product Quality and Integrity; Transforming Technologies for Reliability and Maintainbility Engineering, January 27, 2003 - January 30, 2003, Tampa, FL, United states.

Van der Laan, M., Pollard, K., & Bryan, J. (2003). A new partitioning around medoids algorithm. *Journal of Statistical Computation and Simulation, 73*(8), 575-584.

van Jaarsveld, W., & Dekker, R. (2011a). Estimating obsolescence risk from demand data to enhance inventory control—A case study. *International Journal of Production Economics, 133*(1), 423-431. doi:https://doi.org/10.1016/j.ijpe.2010.06.014

Voller, V. R., & Porté-Agel, F. (2002). Moore's Law and Numerical Modeling. *Journal of Computational Physics, 179*(2), 698-703. doi:http://dx.doi.org/10.1006/jcph.2002.7083

Ward, C. D., & Sohns, C. W. (2011). Electronic component obsolescence. *IEEE Instrumentation and Measurement Magazine, 14*(6), 8-12. doi:10.1109/MIM.2011.6086892

Woods, K., & Bowyer, K. W. (1997). Generating ROC curves for artificial neural networks. *IEEE Transactions on medical imaging, 16*(3), 329-337.

Woods, K., Kegelmeyer, W. P., & Bowyer, K. (1997). Combination of multiple classifiers using local accuracy estimates. *IEEE transactions on pattern analysis and machine intelligence, 19*(4), 405-410.

Wu, D., Jennings, C., Terpenny, J., Gao, R., & Kumara, S. (2017). *Data-Driven Prognostics Using Random Forests: Prediction of Tool Wear.* Paper presented at the ASME 2017 12th International Manufacturing Science and Engineering Conference collocated with the JSME/ASME 2017 6th International Conference on Materials and Processing.

Wu, D., Jennings, C., Terpenny, J., & Kumara, S. (2016, 5-8 Dec. 2016). *Cloud-based machine learning for predictive analytics: Tool wear prediction in milling.* Paper presented at the 2016 IEEE International Conference on Big Data (Big Data).

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., . . . Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems, 14*(1), 1-37. doi:10.1007/s10115-007-0114-2

Xiaodan, W. (2017). *Forecasting short-term wind speed using support vector machine with particle swarm optimization.* Paper presented at the 2017 International Conference on Sensing, Diagnostics, Prognostics and Control (SDPC), 16-18 Aug. 2017, Los Alamitos, CA, USA.

Xiaozhou, M., Thornberg, B., & Olsson, L. (2014). Strategic Proactive Obsolescence Management Model. *IEEE Transactions on Components, Packaging and Manufacturing Technology, 4*(6), 1099-1108. doi:10.1109/TCPMT.2014.2316212

Yang, J., & Honavar, V. (1998a). Feature subset selection using a genetic algorithm. In *Feature extraction, construction and selection* (pp. 117-136): Springer.

Yun, C., Ping, L., & Li, Y. (2010). *Aftermarket demands forecasting with a Regression-Bayesian-BPNN model.* Paper presented at the 2010 IEEE International Conference on Intelligent Systems and Knowledge Engineering (ISKE 2010), 15-16 Nov. 2010, Piscataway, NJ, USA.

Zaabar, I., Beauregard, Y., & Paquet, M. (2018). *Using Multicriteria Decision Making Methods to Manage Systems Obsolescence.* Paper presented at the 2018 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2018, December 16, 2018 - December 19, 2018, Bangkok, Thailand.

Zhang, X., Zou, D., & Shen, X. (2018). A novel simple particle swarm optimization algorithm for global optimization. *Mathematics, 6*(12), 287.

Zheng, L., Nelson, I. I. I. R., Terpenny, J., & Sandborn, P. (2012). Ontology-Based Knowledge Representation for Obsolescence Forecasting. *Journal of Computing and Information Science in Engineering, 13*(1), 014501-014501. doi:10.1115/1.4023003

Zolghadri, M., Addouche, S.-A., Boissie, K., & Richard, D. (2018). Obsolescence prediction: a Bayesian model. *Procedia CIRP, 70*, 392-397.

Zurada, J. M. (1992). *Introduction to artificial neural systems* (Vol. 8): West St. Paul.