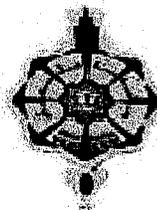


MS/530-08/01

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE



UNIVERSITE ABOU BEKR BELKAID-TLEMCEM
FACULTE DES SCIENCES
DEPARTEMENT DE PHYSIQUE



LABORATOIRE D'ETUDE ET PREDICTION DES MATERIAUX
UNITE DE RECHERCHE MATERIAUX ET ENERGIES RENOUVELABLES

MEMOIRE DE MASTER

EN

PHYSIQUE DE LA MATIERE CONDENSEE ET DES SEMI-CONDUCTEURS

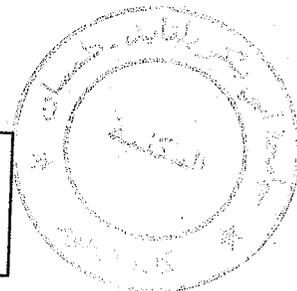
Inscrit Sous le N°:
Date le: 08 JUN 2012
Code: 7545

Présenté par :

M^{me} BENCHAIB LAMIA

THÈME

ETUDE DES ALLIAGES HEUSLER PAR
L'APPROCHE DU DATAMINING



Soutenu le : 28 / 06 / 2012 devant le jury :

Pr. A. Rahmoun	Président	Professeur	U. Sidi Bel Abbas
Mr. A. Sari	Examineur	Maitre Assistant (A)	U. Tlemcen
Pr. G. Merad	Encadreur	Professeur	U. Tlemcen
Mr. H. Zenasni	Co-encadreur	Doctorant	U. Tlemcen

Année universitaire 2011-2012



Bibliothèque sciences

BFS17645

Remerciements

Après les louanges à Dieu miséricordieux, guide de l'univers, que son salut et sa bénédiction soient portés sur son prophète Mohamed.

*Ce travail a été réalisé au sein du laboratoire de recherche d'Etudes et Prédiction des Matériaux (LEPM) affilié à l'Unité de Recherche des Matériaux et des Energies Renouvelables (URMER) de la faculté des sciences, sous la direction de **Monsieur Ghouti Merad**, Professeur à l'Université de Tlemcen. Je lui exprime particulièrement toutes mes reconnaissances pour m'avoir fait bénéficier de ces compétences scientifiques, ses qualités humaines et sa constante disponibilité.*

*Mes remerciements et ma gratitude s'adressent en particulier à **Monsieur Hocine Zenasni** pour les précieux conseils qui a bien voulu me fournir afin de réaliser ce travail, s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu me consacrer et sans qui ce mémoire n'aurait jamais vu le jour.*

*Mes remerciements les plus chaleureux s'adressent au **Mr A. Rahmoun** Professeur à l'université Sidi Bel Abbas d'avoir bien voulu évaluer ce mémoire en faisant l'honneur de présider mon jury.*

*Mes vifs remerciements s'adressent à **Monsieur A. Sari** Maître Assistant (A) à l'université de Tlemcen qui m'a fait l'honneur en acceptant spontanément d'examiner ce modeste travail.*

Leurs critiques et leurs remarques me permettront de clarifier plusieurs points importants de ce mémoire, je leur exprime mes vives reconnaissances.

Pour leur soutien, je remercie mes parents et ma famille, en particulier mon mari qui m'a épaulé durant toute la période d'exécuter ce travail.

J'exprime aussi mes sincères reconnaissances à tous mes enseignants pour leurs efforts fournis durant toute la période d'étude ainsi qu'à tous mes collègues et tous ceux qui ont collaboré d'une façon ou d'une autre à l'élaboration de ce travail.

Dédicaces

Avec grande joie et cœur ouvert je dédie ce modeste travail :

A la grâce de Dieu qui ma guidé.

*A mes très chers parents, qui ont su m'apporter amour et
tendresse et qui ont fait de moi ce que je suis aujourd'hui et je leurs serai
toujours redevable.*

A ma petite cerise sucrée, ma fille Chahrazed .

A mon très cher mari.

A mon frère Mouadh et mes sœurs.

A toute ma famille, petits et grands.

A mes amis qui me sont chers.

« L'imagination est plus importante que le savoir »

Albert Einstein

SOMMAIRE

Introduction Générale.....	1
Références	4

CHAPITRE I

GÉNÉRALITÉS SUR LES ALLIAGES HEUSLER

I.1- Introduction	6
I.2- La magnétorésistance géante et les têtes de lecture pour disques durs.....	6
I.3-La magnétorésistance tunnel et les mémoires magnétique à accès aléatoires.....	8
I.4- La semi-métallicité	9
I.5- Les semi-métaux ferromagnétiques.....	9
I.6-Les alliages Heusler.....	10
I.6.1- Définition	10
I.6.2- La température de Curie.....	13
I.6.3- Le paramètre de maille.....	14
I.6.4- Le moment magnétique.....	14
Références	16

CHAPITRE II

LES TECHNIQUES D'EXPLOITATION DES DONNÉES (DATA MINING)

II.1- Introduction	18
II.2- Principe.....	19
II.3- Algorithmes.....	19
II.3.1- Les méthodes non-supervisées	20
II.3.2- Les méthodes supervisées.....	20
II.3.3- Les méthodes de réduction de données	21
II.4- Analyse en composantes principales.....	21
II.4.1- introduction	21

II.4.2- Principes de l'analyse en composantes principales.....	22
II.4.3- Identification du modèle ACP	24
II.4.4- Détermination de la structure du modèle ACP.....	25
II.4.5- Conclusion	26
II.5- La méthode des moindres carrés partiels (PLS).....	27
II.5.1- Introduction – Problématique	27
II.5.1.1- Présentation.....	27
II.5.1.2- Historique.....	28
II.5.2- Intérêt de la Régression PLS	29
II.5.2-1- Défaillances du modèle linéaire	29
II.5.3- La méthode PLS linéaire	31
II.5.3-1 Modèle de base	31
II.5.3-2- La méthode	32
II.6 Conclusion	32
Références	34

CHAPITRE III

RÉSULTATS ET INTERPRÉTATIONS

III.1- Introduction.....	37
III.2- Base de données.....	37
III.2.1-Définition	37
III.2.2- Présentation des bases de données	38
III.3 Analyse des données	40
III.3.1- L'analyse en composante principale	40
III.3.2 Le modèle PLS (Partial Least Square)	54
III.3.2.a-Mesures de l'importance de variables	55
III.3.2.b-Paramètres et équations du modèle	55
III.3.2.c- Les résidus et prédictions.....	56
III.3.2.d- Prédiction de la température de curie pour quelques alliages d'Heusler.....	58
Références.....	61
Conclusion Générale	63

Introduction Générale

Pendant les trois dernières décennies, l'électronique a connu un développement drastique, mais peu d'attention a été dédiée à l'incorporation de matériaux magnétiques dans les dispositifs électroniques intégrés. Toutefois, un nouveau champ en électronique, la spintronique, a largement attiré l'attention récemment. Le fondement étant la vérité de base qu'un électron a un spin en plus de sa charge. Dans le contexte de l'électronique de spin, les spins des électrons et pas uniquement leur charge électrique, sont contrôlés dans l'opération de transfert d'information dans les circuits. Les matériaux magnétiques et les semi-conducteurs se sont développés séparément pendant longtemps avec les matériaux magnétiques principalement utilisés pour le stockage de données comme dans les disques durs et les dispositifs à base de semi-conducteurs pour le traitement de ces données comme dans les processeurs. C'est un défi de taille que d'intégrer les deux classes de matériaux pour le développement de dispositifs spintronique. Les dispositifs spintronique combinent les avantages des matériaux magnétiques et des semi-conducteurs pour être multilatéraux, rapides, et non volatiles.

Actuellement les progrès technologiques et industriels dans différents domaines dépendent fortement de l'avancement de la recherche dans le domaine des matériaux magnétiques (de nouvelles caractéristiques des matériaux doivent être prises en compte : la polarisation de spin, la bande de conduction, la symétrie des porteurs polarisés, le magnétisme des interfaces...). L'électronique de spin met donc en avant de nouvelles classes de matériaux comme les conducteurs ferromagnétiques et les semi-métaux (manganites, magnétites, alliages d'Heusler,...etc), certain nombre de matériaux magnétiques prioritaires, dont on peut citer particulièrement les semi-conducteurs magnétiques dilués (DMS) et les semi-métaux ferromagnétiques (HMFs).

L'une des applications de l'électronique de spin est l'utilisation des semi-métaux. Une des principales propriétés des ces métaux est la « semi-métallicité », c'est-à-dire, les électrons de conditions qui sont à 100% polarisés en spin en raison d'un écart au niveau de Fermi. Ce phénomène a simulé un grand intérêt dans le développement des matériaux qui possèdent cette propriété, une nouvelle classe a été prédite appelée les semi-métaux ferromagnétiques.

Les semi-métaux ferromagnétiques (Half-Metallic Ferromagnets –HMFs) ont attiré l'attention pour leurs applications potentielles en spintronique. Au niveau de Fermi, ils ont un gap énergétique dans une direction de spin et un caractère métallique pour l'autre direction. Ceci résulte de la polarisation de spin des électrons de conduction. Ils présentent de ce fait de bons candidats pour être utilisés comme des injecteurs de spin pour les MRAM (Magnetic Random Access Memories) et d'autres dispositifs spintroniques.

Une classe prometteuse parmi ces matériaux, sont les alliages Heusler, dont un certain nombre ont été prévus pour être des HMFs. Le premier matériau qui a été prédit pour être un semi-métal est le NiMnSb par De Groot et al [1]. Depuis ce temps, la famille des alliages d'Heusler fait objet de plusieurs études dans le but de concevoir de nouveaux HMFs. Depuis, un grand nombre d'alliages Heusler ont été déclarés posséder des propriétés intéressantes. Par exemple, Ishida et al. ont constaté que l'alliage Heusler total Co_2MnZ (Z : Si, Ge) est un semi-métal [2] ; par la suite, le Mn_2VAI , le Co_2MnX (X : Si, Ge, Sn) et le Ru_2MnZ (Z : Si, Ge, Sn, Sb) ont suscité plusieurs investigations [3]. Plus récemment, le Co_2CrAl , le Co_2CrGa et le Co_2FeSi ont montré un caractère HMFs suivant des calculs de structures électroniques [4]. Aussi, une forte magnétorésistance tunnel (TMR) a été observée dans la majorité des jonctions tunnels magnétiques (MTJs) basées sur le Co_2FeSi [5].

Dans notre travail nous allons utiliser un axe de recherche très étendu : le datamining, qui est novateur en science des matériaux.

Le Datamining est un processus d'extraction de connaissances valides et exploitables à partir de volumes de données même de tailles considérables. Jusqu'à présent, ses champs d'applications couvrent le domaine du marketing, information ... et tous autres domaines qui nécessitent l'utilisation des statistiques. D'un autre côté, en science des matériaux, une quantité phénoménale de données est produite chaque jour par différentes méthodes empiriques ou théoriques et forment une base de données énorme qui ne peut être exploitée de manière efficace. L'idée est d'utiliser des techniques de datamining pour extraire de nouvelles connaissances à partir de cette base de données.

Dans ce travail, nous nous intéressons aux deux familles d'alliages d'Heusler ; les alliages Heusler à base du Manganèse et les alliages Heusler à base du Cobalt, ces derniers ont un intérêt particulier à cause de leur température de Curie relativement élevée.

Ce manuscrit est organisé comme suit : le chapitre I est consacré a la présentation des alliages Heusler et leurs propriétés. Dans le chapitre II nous donnons une idée générale sur l'ensemble des techniques de datamining, notamment la PCA et la PLS. Dans le chapitre III nous exposons les bases de données, et les principaux résultats. Une conclusion générale est donnée à la fin.

Références

- [1] R. A. de Groot, F. M. Mueller, P. G. Van Engen, K.H.J. Buschow, Phys. Rev. Lett. 50 (1983) 2024.
- [2] Fujii S, Sugimura S, Ishida S, Asano S. J Phys : Condens Matter 1990 ; 2-8 583.
- [3] Hongzhi Luo et al. Journal of Magnetism and magnetic Materials 320 (2008) 421 – 428.
- [4] A. Kellow, N.E. Fenineche, T. Grosdidier, H. Aourag, C. Coddet, J. Appl. Phys. 94 (2003) 3292.
- [5] T. Marukame et al. Appl. Phys. Lett. 88 (2006) 262503.

CHAPITRE I
GENERALITES SUR LES ALLIAGES
HEUSLER

I.1- Introduction

La spintronique est une technologie émergente qui exploite non seulement la charge mais aussi le spin de l'électron pour coder l'information. Le spin qui était totalement négligé dans les applications de la microélectronique classique, donne lieu à des phénomènes physiques nouveaux qui offrent des perspectives intéressantes en matière d'intégrabilité, de vitesse de communication, de consommation et de non volatilité de l'information. C'est en ce sens que la spintronique a suscité un vif intérêt dans la communauté scientifique comme alternative à l'électronique classique qui est aujourd'hui confrontée à des obstacles physique majeurs dus à une réduction accrue de la taille des composants.

Pour réaliser un dispositif en spintronique, il est impératif d'avoir deux systèmes : un système qui peut générer un courant d'électrons spin polarisés, c'est-à-dire majoritairement composé d'électrons avec l'un des deux spins. Un tel système est dit "injecteur de spin" ; et un second système qui soit sensible à la variation du spin dans le courant, il sera le "détecteur de spin".

La spin-polarisation dans les matériaux non magnétiques peut être réalisée par application de larges champs magnétiques. Toutefois, la polarisation décroît rapidement sur une durée appelée "durée de vie du spin" (Spin lifetime). Cette durée est relativement courte dans les métaux, de l'ordre de la nanoseconde, alors qu'elle peut être nettement plus longue dans les semi-conducteurs, de l'ordre des microsecondes à basse température.

Ces durées de vie restent quand même trop courtes pour des applications utiles. C'est ainsi que deux volets se sont ouverts à l'investigation : les semi-conducteurs magnétiques dilués et les semi-métaux ferromagnétiques. Ce sont ces derniers qui nous intéressent dans ce chapitre.

I.2- La magnétorésistance géante et les têtes de lecture pour disques durs

La première manifestation d'un effet caractéristique de spintronique a été la magnétorésistance géante (Giant magneto-Resistance, ou GMR), découverte en 1988 à Orsay par l'équipe d'Albert Fert (prix Nobel 2007) [1]. Dans des structures artificielles alternant un métal magnétique et un métal non magnétique, un changement de résistance important est observé lorsque, sous l'effet d'un champ magnétique extérieur, les aimantations macroscopiques des couches magnétiques successives basculent d'un état

antiparallèle à un état parallèle. La figure I-1 montre la variation de la résistance dans une telle structure entre les deux cas extrêmes d'alignement. Des structures plus complexes ont ensuite été développées pour augmenter la sensibilité de l'élément magnétorésistif. Il s'agit de ce qu'on appelle des vannes de spin.

La magnétorésistance géante est l'outil adéquat pour détecter des champs magnétiques de très faible amplitude. L'application principale a été développée dans les années 90. Elle concerne les têtes de lecture pour disques durs où un petit élément magnétorésistif détecte la variation de champ magnétique généré par les transitions entre bits de données dans un support d'information et transcrit l'information sous la forme de pulses électriques. Aujourd'hui, la totalité de la production mondiale des têtes de lecture/écriture pour disques durs est constituée de têtes magnétorésistives de type vannes de spin.

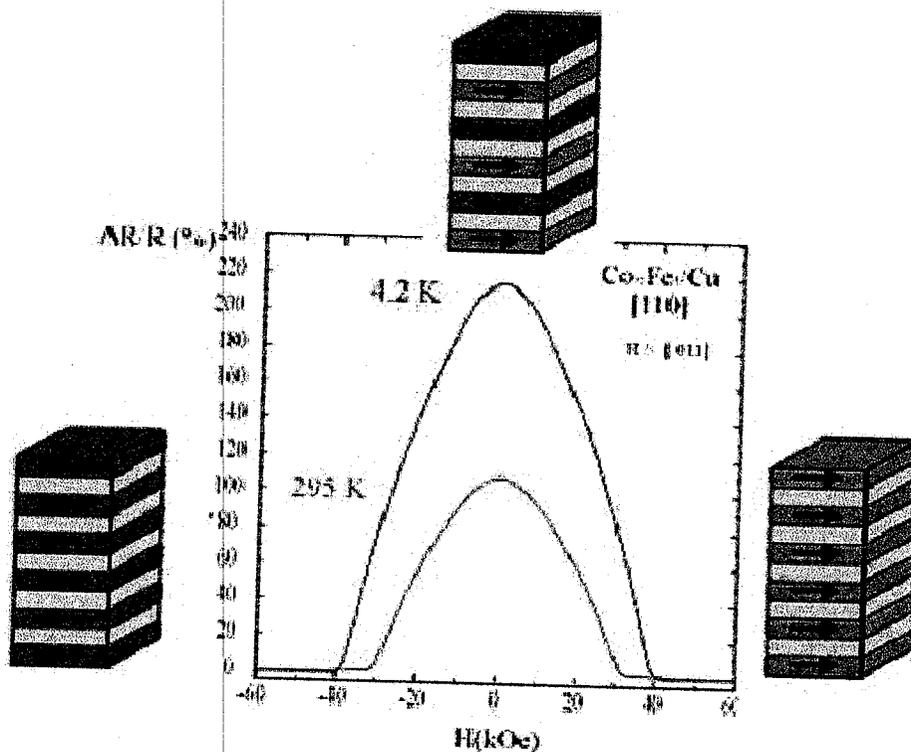


Figure I-1 Magnétorésistance géante dans une structure à base de multicouches d'alliage cobalt-fer avec du cuivre.

I.3-La magnétorésistance tunnel et les mémoires magnétique à accès aléatoires

Un effet de magnétorésistance similaire à la magnétorésistance géante, appelé magnétorésistance tunnel [2], a été observé dans des jonctions tunnel métal/isolant/métal, dans les quelles les deux électrodes métallique sont magnétique. Une variation importante de la résistance de la jonction lorsque les directions relatives des aimantations des couches ferromagnétiques varient. La forte variation de magnétorésistance à température ambiante, trois fois supérieure à la magnétorésistance géante, soit une variation relative proche de 50% associée à l'impédance élevée de (0.1 à 100 K Ω), proche de l'impédance caractéristique des composants semi-conducteurs usuels, permettent d'envisager une utilisation de ces structures comme cellules de stockage dans des mémoires non volatiles à accès aléatoire. Dans ces mémoires, appelées MRAM (Magnetic Random Access Memories), l'information n'est plus stockée sous la forme d'une charge dans une capacité, comme c'est le cas des mémoires semi-conductrices de type DRAM ou Flash, mais sous la forme d'une direction d'aimantation dans la jonction tunnel magnétique. Des prototypes fonctionnels de MRAM ont déjà démontré la validité de l'approche et la plupart des grands groupes industriels du domaine de la microélectronique projettent la mise rapide sur le marché de composants fonctionnels.

Paradoxalement, malgré l'avènement imminent de produits commerciaux, les mécanismes physiques en jeu sont encore mal compris. Bien que présentant le même comportement « macroscopique » que la magnétorésistance géante, le mécanisme physique à l'origine de la magnétorésistance tunnel est différent. Il apparait que la magnétorésistance tunnel dépend principalement de l'asymétrie de spin, mais également de la structure électronique de l'isolant et du caractère des liaisons électroniques à l'interface métal/isolant.

Un autre enjeu important émerge alors qui concerne la recherche de matériaux ferromagnétiques semi-métalliques, dans lesquels un seul type de spin est présent, et donc une asymétrie de spin de 100%.

Parmi ces matériaux, les manganites de type LaSrMnO₃ ont permis d'obtenir des amplitudes de magnétorésistance tunnel de 1800% correspondant à une polarisation en spin de 95%. Cependant la température maximale utile du LaSrMnO₃ est trop basse pour des applications dans l'électronique grand public. D'autres pistes sont actuellement à l'étude pour obtenir des matériaux semi-métalliques de température de Curie plus élevée

[3]. La semi-métallicité, accompagnée par un moment magnétique important à température ambiante, a été prédite dans les alliages Heusler par Groot et al [4].

I.4- La semi-métallicité

Le terme "métal" possède deux sens plutôt distincts. Tout d'abord il désigne les métaux purs et leurs alliages, ils ont une apparence lustrée et comportement mécanique assez rigide mais surtout très malléable et ductile. Les métaux sont de très bons conducteurs d'électricité et de chaleur et ne sont pas transparents à la lumière. Les éléments métalliques du tableau périodique ont un petit nombre d'électrons de valence (moins de 4 en général) qui sont faiblement liés à leur atome. La plus grande partie des propriétés des métaux est attribuée à leurs électrons. [5]

D'un autre côté, ce terme désigne un état de la matière, en l'occurrence l'état métallique, dans lequel elle conduit le courant électrique et la chaleur et possède des propriétés mécaniques particulières. De la matière qui contient des éléments métalliques n'est pas toujours dans un état métallique ; des exemples étant les oxydes, les carbures ou les nitrures. Ainsi, attester qu'un matériau est "métallique" revient à examiner sa structure de bandes énergétiques qui indiquerait l'absence d'un gap énergétique et la possibilité qu'auront les électrons d'être promus dans la bande de conduction lors de l'application d'un champ électrique.

De même, l'état semi-métallique est défini par rapport à une structure particulière des bandes énergétiques. Plus exactement, il s'agit d'un état de la matière où la conduction s'opère uniquement pour une partie des électrons de valence, les électrons de spin majoritaire ou spin up. Autrement dit, la structure de bandes énergétiques d'un matériau semi-métallique présente une asymétrie entre les états de spin up et de spin down avec un gap ou un pseudo gap énergétique au niveau de Fermi. Ceci donne lieu à des polarisations des électrons de conduction au niveau de Fermi qui peuvent atteindre les 100%.

I.5- Les semi-métaux ferromagnétiques

Les semi-métaux ferromagnétiques sont des solides qui présentent une asymétrie entre les états de spin up et de spin down avec un gap ou un pseudo gap énergétique au niveau de Fermi (figure I-2). Ceci résulte une spin-polarisation des électrons de conduction, l'écart peut se produire dans la majorité ou dans la minorité de l'état de spin. Cette propriété inhabituelle pourrait avoir des implications importantes pour les applications liées au

magnétisme et pour l'électronique de spin : conducteurs électriques, matériaux ferromagnétiques, où les électrons de conduction ont une grande mobilité et sont entièrement polarisés en spin, ces matériaux sont souhaitables pour de futures réalisations des couches minces, appareils spin électroniques ...etc.

En général, les semi-métaux sont ferromagnétiques, mais un matériau ferromagnétique n'est pas nécessairement un semi-métal. Les semi-métaux sont des cas extrêmes des ferromagnétiques ils n'ont pas seulement les électrons (3d) qui sont entièrement polarisés, mais aussi d'autres états (s et p notamment) qui ne se croisent pas au niveau de Fermi.

Les semi-métaux ferromagnétiques peuvent servir comme des injecteurs de spin idéal et des détecteurs, parce qu'ils ne peuvent transporter le courant que dans un seul sens. Ils peuvent être utilisés dans les mémoires magnétiques MRAM ainsi que d'autres dispositifs dont le fonctionnement dépend du spin comme les jonctions tunnel magnétiques. [6]

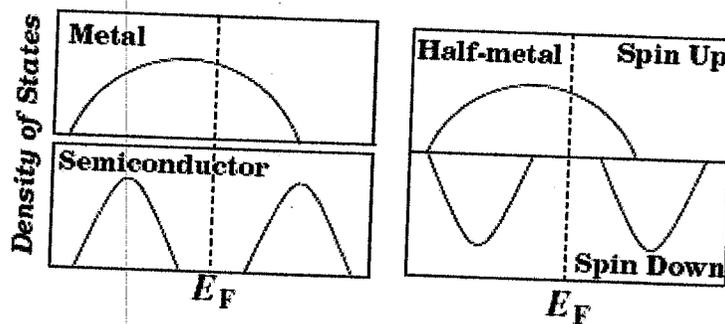


Figure I-2 représentations schématiques des densités d'états d'un métal, un semi-conducteur et un semi-métal au niveau de Fermi [7].

I.6- Les alliages Heusler

I.6.1- Définition

L'histoire d'une classe de matériaux la plus passionnante remonte à 1903 lorsque Fritz Heusler [8] a découvert un alliage avec une formule Cu_2MnAl se comportant comme un élément ferromagnétique, même si aucun de ses constituants ne soit magnétique. La famille de ce matériau remarquable, qui actuellement comprennent plus de 1000 composés est aujourd'hui connus par les **alliages Heusler**.

Les alliages Heusler sont des alliages ternaires composés de deux métaux de transition et un élément non magnétique. Il existe deux classes : les alliages dits Full Heusler de formule X_2YZ et ceux dits semi-Heusler de formule XYZ où, dans les deux cas, X et Y

sont les métaux de transitions et Z est l'élément non magnétique. La figure I-3 donne un aperçu sur des combinaisons possibles pour former ses alliages.

X₂YZ Heusler compounds

H 2.20																	He	
Li 0.98	Be 1.57											B 2.04	C 2.55	N 3.04	O 3.44	F 3.98	Ne	
Na 0.93	Mg 1.31											Al 1.61	Si 1.90	P 2.19	S 2.58	Cl 3.16	Ar	
K 0.82	Ca 1.00	Sc 1.36	Ti 1.54	V 1.63	Cr 1.66	Mn 1.55	Fe 1.83	Co 1.88	Ni 1.91	Cu 1.90	Zn 1.65	Ga 1.81	Ge 2.01	As 2.18	Se 2.55	Br 2.96	Kr 3.00	
Rb 0.82	Sr 0.95	Y 1.22	Zr 1.33	Nb 1.60	Mo 2.16	Tc 1.90	Ru 2.20	Rh 2.28	Pd 2.20	Ag 1.93	Cd 1.69	In 1.78	Sn 1.96	Sb 2.05	Te 2.10	I 2.66	Xe 2.60	
Cs 0.79	Ba 0.89	Hf 1.30	Ta 1.50	W 1.70	Re 1.90	Os 2.20	Ir 2.20	Pt 2.20	Au 2.40	Hg 1.90	Tl 1.80	Pb 1.80	Bi 1.90	Po 2.00	At 2.20	Rn		
Fr 0.70	Ra 0.90																	
		La 1.10	Ce 1.12	Pr 1.13	Nd 1.14	Pm 1.13	Sm 1.17	Eu 1.20	Gd 1.20	Tb 1.10	Dy 1.22	Ho 1.23	Er 1.24	Tm 1.25	Yb 1.10	Lu 1.27		
		Ac 1.10	Th 1.30	Pa 1.50	U 1.70	Np 1.30	Pu 1.28	Am 1.13	Cm 1.28	Bk 1.30	Cf 1.30	Es 1.30	Fm 1.30	Md 1.30	No 1.30	Lr 1.30		

Figure I-3 : Tableau périodique des éléments. Le grand nombre d'alliages d'Heusler peut être formé par la combinaison des différents éléments selon le schéma de couleur.

Les alliages Heusler cristallisent dans des structures cubiques du réseau de Bravais cubique à faces centrées. Les atomes de type Z occupent les sommets et les centres des faces, les atomes de type Y occupent les milieux des arêtes, alors que les atomes de type X occupent les positions tétraédriques : (i) les huit positions sont occupées dans le cas des alliages full Heusler, la structure cristalline est donc L2₁ ; (ii) la moitié de ces sites, quatre au total, sont occupées dans le cas des alliages semi Heusler, la structure cristalline est alors C1_b.

La figure I-4 présente les deux structures (a) C1_b et (b) L2₁.

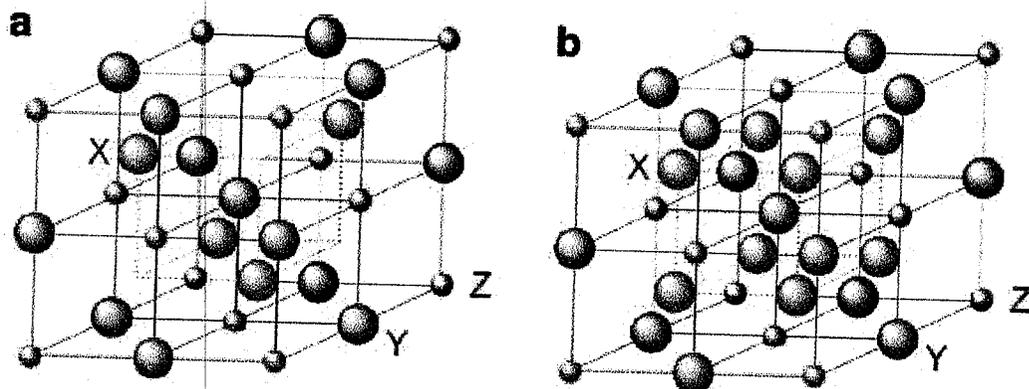


Figure I-4 Structures des alliages Heusler. (a) structure Semi Heusler C1b, (b) structure Full Heusler L21

Sous certaines conditions, notamment de température, un désordre atomique s'installe. En premier lieu, les atomes du type Y et ceux de type Z échangent leurs positions aléatoirement au point qu'un site atomique sera occupé par le type Y ou Z à probabilités égales. Les atomes de type X gardent leurs positions, et la structure cristalline devient dans ce cas de type B2. En second lieu, le désordre peut devenir suffisamment important pour que les atomes de type X soient aussi impliqués dans l'échange des sites atomiques. Le désordre devient total et la structure cristalline est de type A2. La figure I-5 montre les deux structures B2 et A2.

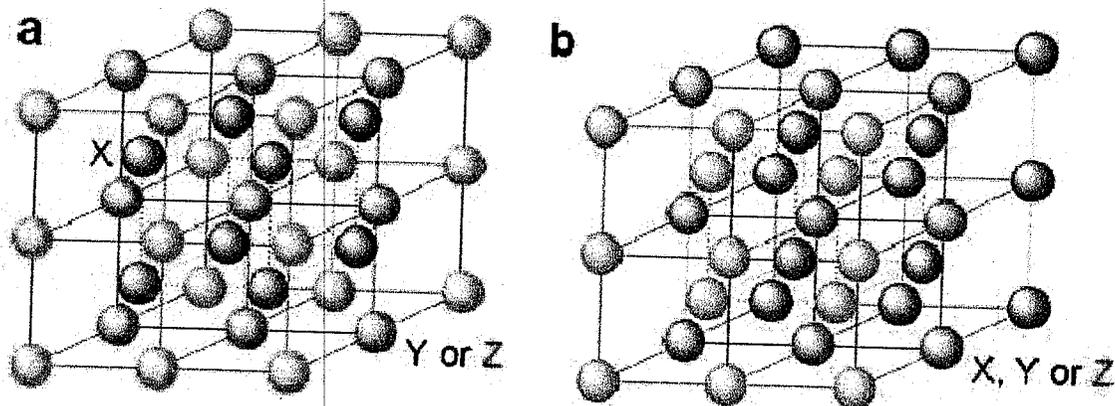


Figure I-5 Structures des alliages Heusler désordonnées : (a) la structure B2 et (b) la structure A

I.6.2- La température de Curie

Un point crucial dans le développement de matériaux semi-métalliques magnétiques est leur température de Curie T_c . Comparés aux semi-conducteurs magnétiques dilués (DMS), il est plus courant de trouver des alliages Heusler avec des températures de Curie supérieures à l'ambiante. Toutefois, il reste indispensable de prendre en compte ce facteur lors du développement de nouvelles compositions ; il est en effet clair que T_c dépend fortement de la composition. La figure I-6 présente les valeurs de la température de Curie pour une sélection de matériaux avec : (a) une structure Semi-Heusler et (b) une structure Full-Heusler. Nous pouvons y distinguer très nettement (figure 6-b) le Co_2FeAl avec une valeur de T_c proche de 1200°K , plus importante que celles de l'ensemble des autres matériaux représentés.

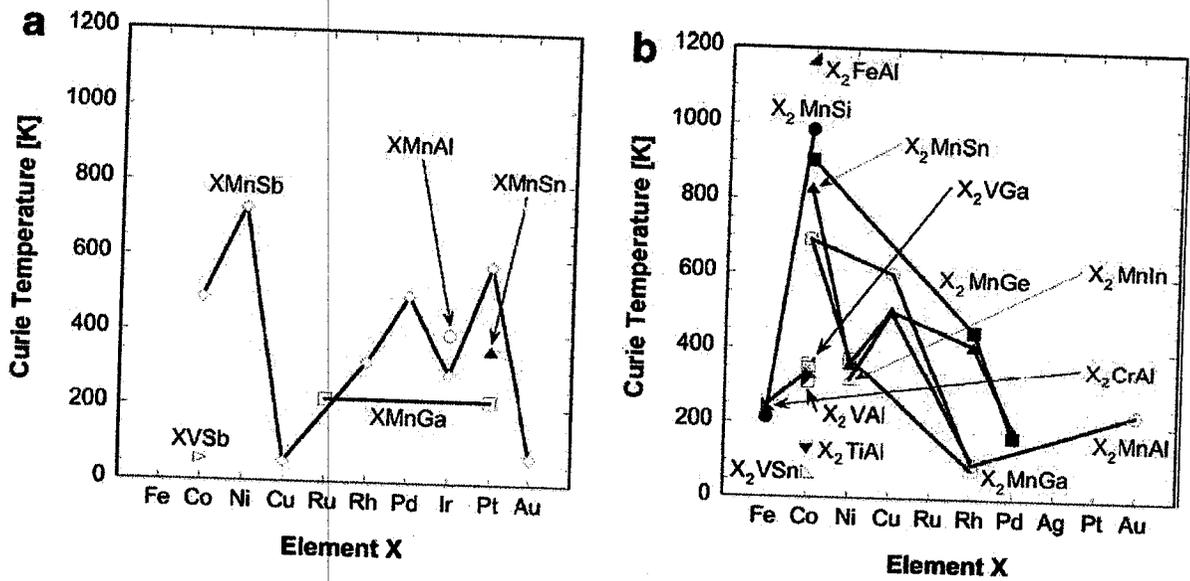


Figure I-6 Températures de Curie T_c pour un ensemble d'alliages (a) Semi Heusler et (b) Full Heusler [9].

I.6.3- Le paramètre de maille

Un autre point aussi important que le premier, voire plus, concerne la question de faire croître un alliage Heusler sur un substrat semi-conducteur sans altérer ses propriétés cristallines. Le meilleur moyen d'y arriver est très certainement d'avoir un accord de maille raisonnable entre les deux matériaux. Il a été démontré que plusieurs semi-conducteur III-V ont des paramètres de maille du même ordre que les alliages Heusler. Ce désaccord de maille très petit permet de réduire au minimum le désordre atomique à l'interface.

La figure I-7 présente les valeurs des paramètres de maille pour une sélection de matériaux avec: (a) une structure Semi-Heusler et (b) une structure Full-Heusler, comparés à ceux des principaux semi-conducteurs III-V.

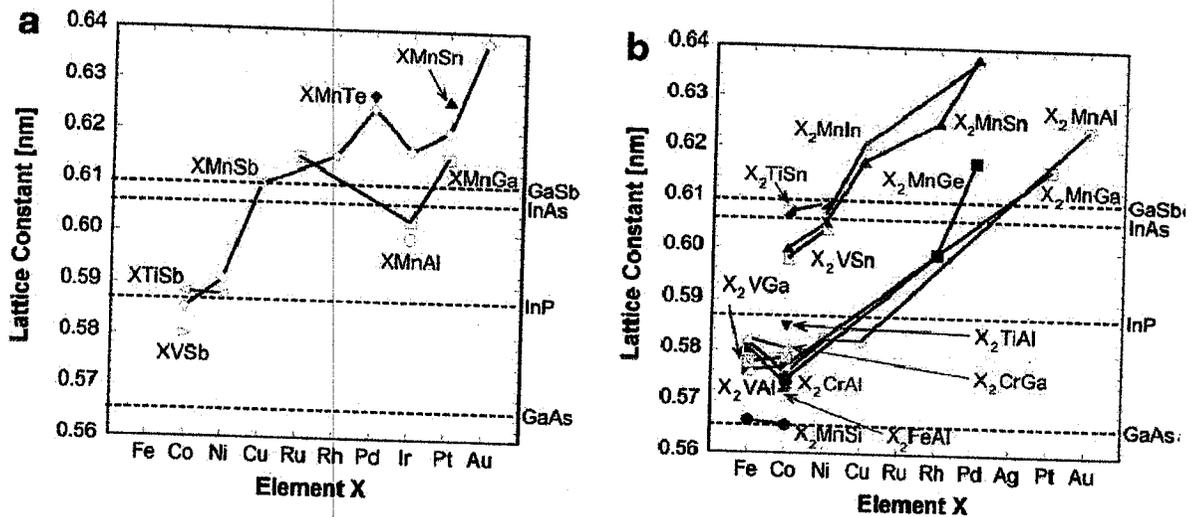


Figure I-7 Paramètres de maille pour un ensemble d'alliages (a) Semi Heusler et (b) Full Heusler comparés par ceux des semi-conducteurs III-V [9].

I.6.4- Le moment magnétique

La valeur du moment magnétique totale du matériau est également très importante, Généralement c'est cette propriété qui conditionne son utilisation, mais encore elle est très liée à la température de Curie. Souvent, autant le moment magnétique d'un matériau est important, autant sa température de Curie est plus élevée.

Une étude systématique [9] qui a porté sur 810 alliages a montré qu'une formule générale peut être employée pour estimer le moment magnétique total dans un alliage Heusler ; en l'occurrence :

$$M_t = Z_t - 18 \quad \text{pour les alliages semi Heusler}$$

$$M_t = Z_t - 24 \quad \text{pour les alliages full Heusler}$$

Z_t est le nombre d'électrons de valence dans le système.

La figure I-8 présente les valeurs des moments magnétiques pour une sélection de matériaux avec: (a) une structure Semi-Heusler et (b) une structure Full-Heusler. Les matériaux sont représentés suivant un nombre d'électrons de valence croissant. Il y est très clair que la formule est presque rigoureusement adéquate pour l'ensemble des alliages.

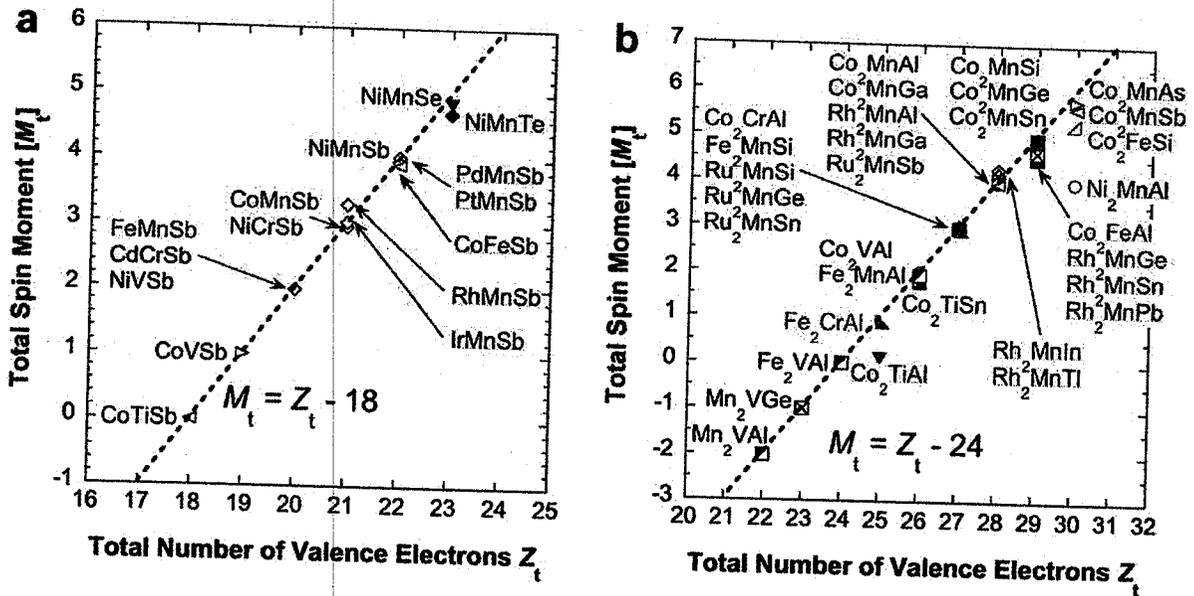


Figure I-8 Moments magnétiques pour un ensemble d'alliages (a) Semi Heusler et (b) Full Heusler [9].

Références

- [1] Le magazine d'information de l'université Paris-Sud, janvier- février 2008 numéro 69.
- [2] M. Jullière, Phys. Lett. A **54**, 225 (1975).
- [3] F. J. Albert, J. A. Katine, R. A. Burham, and D. C. Ralph, Appl. Phys. Lett. **77**, 3809 (2000).
- [4] Hongzhi Luo, Zhiyong Zhu, Guodong Liu, Shifeng Xu, Guangheng Wu, Heyan Liu, Jingping Qu, Yangxian Li, Journal of magnetism and magnetic Materials, Tianjin 300130, PR China, 3 July 2007.
- [5] Benabadji Mostefa Karim, Mémoire de Magister en Physique de la Matière Condensée et Des Semi-Conducteurs, Laboratoire d'Etude et de Prédiction des Matériaux, URMER, Université Abou Bakr Belkaid, 2009.
- [6] Seddik Larbi, Mémoire de Magister en Science et Génie des Matériaux, Université Djillali Liabès de Sidi-Bel-Abbes, 2010.
- [7] Iosif Galanakis and Peter H. Dederichs Institut für Festkörperforschung, Forschungszentrum Jülich, D-52425 Jülich, Germany.
- [8] F. Heusler, W. Starck, E. Haupt, Verh. Dtsh. Phys. Ges. **5**(1903) 219.
http://www.archive.org/stream/verhandlungende33unkngoog/verhandlungende33unkngoog_djvu.txt.
- [9] A. Hirohata, M. Kikuchi, N. Tezuka, K. Inomata. Current Opinion in Solid State and Material Science **10** (2006) 93-107.

CHAPITRE II

LES TECHNIQUES D'EXPLOITATION

DES DONNEES

(DATAMINING)

II.1- Introduction

L'exploration de données, aussi connue sous le nom fouille de données, datamining (forage de données) ou encore Extraction de Connaissances (ECD en français, KDD en Anglais), a pour objet l'extraction d'un *savoir 'utile'* ou d'une connaissance à partir de grandes quantités de données, par des méthodes automatiques ou semi-automatiques.

Le Datamining étant un processus d'extraction de connaissances valides et exploitables à partir de volumes de données même de tailles considérables, il a vocation à être utilisé dans un environnement professionnel et se distingue de l'analyse de données et des techniques de la statistique par les points suivants :

- Contrairement aux méthodes statistiques, le Datamining ne nécessite jamais que l'on établisse une hypothèse de départ qu'il s'agira de vérifier. C'est à partir des données elles-mêmes que se dégageront les corrélations intéressantes, l'algorithme n'étant là que pour les découvrir et non pour les établir. Le Datamining se situe à la croisée des statistiques, de l'intelligence artificielle et des bases de données.
- Les connaissances extraites par le Datamining ont vocation à être intégrées dans le schéma organisationnel de l'entreprise ou de l'entité considérée. Le Datamining impose donc d'être capable d'utiliser de manière opérationnelle les résultats des analyses effectuées, souvent dans des délais très courts. Le processus d'analyse doit permettre à l'organisation une réactivité (très) importante.
- Les données traitées sont issues des systèmes de stockage en place dans l'organisation et sont ainsi hétérogènes, multiples, plus ou moins structurées... Pour résumer ce sont des données dont la raison d'être n'est *a priori* pas l'analyse. Cela impose de disposer de systèmes performants de préparation ou de manipulation de données.

Le Datamining se propose donc de transformer en information, en connaissance, de grands volumes de données qui peuvent être stockées de manières diverses dans des bases de données relationnelles ou dans un (ou plusieurs) entrepôt de données (datawarehouse), mais qui peuvent aussi être récupérées de sources riches plus ou moins structurées comme Internet, ou encore en temps réel. Lorsque la source n'est pas directement un entrepôt de données, il s'agira très souvent de construire une base de données ou un datamart dédié à l'analyse et aux analystes. Cela suppose bien entendu d'avoir à sa disposition une palette d'outils de gestion de données (data management).

II.2- Principe

Le Datamining est un processus d'analyse dont l'approche est différente de celle utilisée en *statistique*. Cette dernière présuppose en général que l'on se *fixe une hypothèse* que les données vont nous permettre ou non de confirmer. Au contraire, le *Datamining* adopte une démarche beaucoup plus empirique et essaye ainsi de faire *émerger*, à partir des données brutes, des hypothèses que l'expérimentateur peut ne pas soupçonner, et dont il aura à valider la pertinence.

Le *Datamining* tente alors de réaliser un arbitrage entre validité scientifique, interprétabilité des résultats et facilité d'utilisation.

Plus qu'une théorie normalisée, le *Datamining* est un processus d'extraction de connaissances en suivant les étapes principales suivantes :

- Formaliser un problème que l'organisation cherche à résoudre en terme de données
- Accéder aux données idoines quelles qu'elles soient
- Préparer les données en vue des traitements et utilisations futurs
- Modéliser les données en leur appliquant des algorithmes d'analyse
- Evaluer et valider les connaissances ainsi extraites des analyses

Bien sûr ce processus n'est pas linéaire, il se veut cyclique, le *Datamining* participant dès lors à une rationalisation avancée du stockage de l'information et des données. Autrement dit, le *Datamining* ne consiste pas en une succession d'études ad hoc mais doit bien permettre de capitaliser des connaissances acquises sous forme de connaissances explicites, donc de structurer les contenus nécessaires à l'ingénierie des connaissances. Le *Datamining* se propose alors d'utiliser un ensemble d'algorithmes issus de disciplines scientifiques diverses (statistiques, intelligence artificielle, base de données) pour construire des modèles à partir des données, c'est-à-dire trouver des schémas « intéressants » (des 'patterns' ou motifs de conception en français) selon des critères fixés au départ, et extraire de ces données un maximum de connaissances utiles.

II.3- Algorithmes

Résoudre une problématique avec un processus de *Datamining* impose généralement l'utilisation d'un grand nombre de méthodes et algorithmes différents. On peut distinguer 3 grandes familles d'algorithmes :

- Les méthodes non-supervisées
- Les méthodes supervisées
- Les méthodes de réduction de données

II.3.1- Les méthodes non-supervisées

Elles permettent de travailler sur un ensemble de données dans lequel aucune des données ou des variables à disposition n'a d'importance particulière par rapport aux autres, c'est-à-dire un ensemble de données dans lequel aucune variable n'est considérée individuellement comme la cible, l'objectif de l'analyse. On les utilise par exemple pour dégager d'un ensemble d'individus des groupes homogènes (typologie), pour réaliser de la compression d'informations.

On peut citer quelques techniques disponibles pour ces méthodes:

- Techniques à base de Réseau de neurones : réseau de Kohonen, les réseaux Hebbienx...
- Techniques utilisées classiquement dans le monde des statistiques : classification ascendante hiérarchique, k-means et les nuées dynamiques (Recherche des plus proches voisins), les classification mixtes (Birch...), les classifications relationnelles...

II.3.2- Les méthodes supervisées

Leur raison d'être est d'expliquer et/ou de prévoir un ou plusieurs phénomènes observables et effectivement mesurés. Concrètement, elles vont s'intéresser à une ou plusieurs variables de la base de données définies comme étant les cibles de l'analyse.

Voici une liste non exhaustive des techniques disponibles :

- Techniques à base d'arbres de décision (Arbre de décision).
- Techniques statistiques de régressions linéaires et non linéaires au sens large : Régression linéaire, Régression linéaire multiple, Régression logistique binaire ou multinomiale, Analyse discriminante linéaire ou quadratique, modèle linéaire généralisé, régression PLS, régressions non paramétriques, équations structurelles ...
- Techniques à base de Réseaux de neurones : perceptron mono ou multicouches avec ou sans rétropropagation des erreurs, les réseaux à fonction radiale de base...
- Techniques à base d'algorithme génétique.
- Techniques à base d'Inférence bayésienne (Réseau bayésien).

II.3.3- Les méthodes de réduction de données

Elles permettent de réduire un ensemble de données volumineux à un ensemble de taille plus réduite, épuré de ce que l'on considérera comme de l'information non pertinente ou non signifiante, comme du bruit. Elles sont ainsi très souvent, mais pas systématiquement, utilisées en amont des techniques supervisées ou non supervisées. Elles sont notamment très complémentaires des techniques non supervisées classiquement utilisées dans le domaine des statistiques.

- Techniques d'Analyse factorielle : Analyse en composantes principales, analyse factorielle des correspondances, analyse des correspondances multiples, analyses factorielles (maximum de vraisemblance, moindres carrés non pondérés, avec ou sans rotation orthogonale ou oblique)...
- Techniques de positionnement : positionnement multidimensionnel...

II.4- Analyse en composantes principales

II.4.1- introduction

La représentation brute de l'information conduit à des vecteurs de caractéristiques de grandes dimensions, ce qui peut poser des problèmes de complexité de calcul et de capacité de stockage. Ainsi, il est parfois préférable de chercher à réduire la dimensionnalité d'un problème de classification pour améliorer ses performances. Le mathématicien Bellman, père de la programmation dynamique, a introduit l'expression «malédiction de la dimensionnalité» (curse of dimensionality), pour signifier que représenter les formes par des vecteurs de taille importante est source de problème (Bellman 1961) [1].

Il est évident que l'on se trouve d'emblée confronté à un problème de grande dimension. Plusieurs solutions sont préconisées dans la littérature pour réduire la taille de sa dimension Minami 1999[2]. Dans ce travail, nous avons utilisé l'analyse en composantes principales.

L'analyse en composantes principales (ACP) est une technique descriptive permettant d'étudier les relations qui existent entre les variables, sans tenir compte, a priori, d'une quelconque structure (Jolliffe 1986) [3], (Diamantaras 1996) [4]. Son utilisation pour l'exploitation des données remonte au début du siècle dernier. Elle est principalement issue des travaux de psychologues américains (Pearson 1901) [5], (Spearman 1904) [6], (Hotelling 1933) [7]. Le but de l'ACP est d'identifier la structure de dépendance entre des observations multivariées, afin d'obtenir une description ou une représentation compacte de ces dernières. L'idée de base de l'ACP est de réduire la dimension de la matrice des données, en retenant le

plus possible les variations présentes dans le jeu des données de départ. Cette réduction ne sera possible que si les variables initiales ne sont pas indépendantes et ont des coefficients de corrélation entre elles non nuls. Ces variables initiales sont transformées en de nouvelles variables, appelées composantes principales. Elles sont obtenues par combinaisons linéaires des précédentes et sont ordonnées et non corrélées entre elles. L'analyse en composantes principales cherche à identifier les vecteurs propres orthonormaux et leurs valeurs propres correspondantes de la matrice de dispersion des variables originelles.

Les vecteurs propres orthonormaux sont utilisés pour construire les composantes principales et les valeurs propres sont les variances des composantes principales correspondantes (Jolliffe 1986).

II.4.2- Principes de l'analyse en composantes principales

Dans le cadre de cette thèse, nous aborderons l'ACP comme une technique de réduction et de description des échantillons.

La théorie sous-jacente à l'analyse en composantes principales est vaste, nous ne passerons donc en revue que les points les plus importants. Du point de vue géométrique l'ACP consiste à effectuer une certaine rotation du repère des variables autour de leurs valeurs moyennes. Cette rotation transforme les n variables corrélées en l variables non corrélées. Notons que ce sont justement ces variables transformées que l'on a nommées les composantes principales. Supposons qu'on a un vecteur de données aléatoires $x_k = [x_1, \dots, x_m]^t \in \mathfrak{R}^m$ de moyenne nulle $E\{x(k)\} = 0$ et de matrice de covariance ou d'auto-corrélation :

$$\Sigma = E\{xx^t\} \in \mathfrak{R}^{m \times m}. \quad (\text{II.1})$$

Avec

$$E = \frac{1}{m-1}$$

En analyse en composantes principales, un vecteur caractéristique $t \in \mathfrak{R}^l$ est associé à chaque vecteur de données dont il optimise la représentation au sens de la minimisation de l'erreur d'estimation de x ou la maximisation de la variance de t . Les vecteurs t et x sont liés par une transformation linéaire $t = P^t x$,

Où la matrice de transformation $P \in \mathfrak{R}^{m \times l}$ vérifie la condition d'orthogonalité $P^t P = I_l$. Les colonnes de la matrice P forment les vecteurs de base orthonormés d'un sous-espace \mathfrak{R}^l de représentation réduite des données. La transformation linéaire s'apparente ainsi à une

projection de l'espace des données de dimension m vers un sous-espace orthogonal de dimension ℓ .

Les composantes t_j , avec $j = 1, \dots, \ell$, du vecteur caractéristique \mathbf{t} représentent les composantes projetées du vecteur de données \mathbf{x} dans ce sous-espace.

Au sens de l'ACP, la projection P est optimale si l'erreur quadratique d'estimation des vecteurs de données \mathbf{x} est minimale. Ce problème d'optimisation s'exprime par :

$$P_{opt} = \arg \min J_e(P) \quad (\text{II.2})$$

Où J_e représente le critère d'erreur d'estimation de l'ACP. Sous la contrainte d'orthogonalité de la matrice de projection $P^t P = I_\ell$, ce critère aura la forme :

$$\begin{aligned} J_e(P) &= \mathcal{E}\{\|\mathbf{x} - \hat{\mathbf{x}}\|^2\} & (\text{II.3}) \\ &= \mathcal{E}\{\|\mathbf{x} - P P^t \mathbf{x}\|^2\} \\ &= \mathcal{E}\{(\mathbf{x} - P \mathbf{t})^t (\mathbf{x} - P \mathbf{t})\} \\ &= \mathcal{E}\{\mathbf{x}^t \mathbf{x} - 2 \mathbf{t}^t \mathbf{x} + \mathbf{t}^t P P^t \mathbf{x}\} \\ &= \mathcal{E}\{\text{trace}(\mathbf{x} \mathbf{x}^t) - \mathbf{t}^t \mathbf{x}\} \\ &= \text{trace}(\Sigma) - \mathcal{E}\{\mathbf{t}^t \mathbf{x}\} \end{aligned}$$

Notons que la trace d'une matrice carrée est définie par la somme de ces éléments diagonaux. Du fait que la matrice de covariance Σ est indépendante de la matrice des paramètres P , minimiser J_e revient à maximiser le second terme J_v de son expression :

$$J_v(P) = \mathcal{E}\{\mathbf{t}^t \mathbf{x}\} = \sum_{j=1}^{\ell} \mathcal{E}\{t_j^2\} \quad (\text{II.4})$$

Ainsi, la minimisation de l'erreur quadratique d'estimation de \mathbf{x} est équivalente à la maximisation de la variance des projections t_j des données. En conséquence :

$$P_{opt} = \arg \min J_e(P) = \arg \max J_v(P) \quad (\text{II.5})$$

Le problème de l'ACP, considéré sous l'angle de la maximisation de la variance de projection des données, est celui de la détermination des vecteurs propres de la matrice de covariance Σ .

II.4.3- Identification du modèle ACP

Soit $\mathbf{x}_k = [x_1, \dots, x_n]^t$ un vecteur d'observations de n variables aléatoires.

On notera $\bar{\mathbf{x}}$ le vecteur correspondant à la moyenne des observations de \mathbf{x} et Σ la matrice de covariance de ces m observations. L'estimation des paramètres du modèle ACP se résume en une estimation des valeurs et vecteurs propres de la matrice de covariance Σ . Puisque la matrice de covariance Σ est carrée de dimension $n \times n$ et symétrique, on peut, selon l'algèbre linéaire, vérifier une relation de type :

$$P^t \Sigma P = L \quad (\text{II.6})$$

Où L est une matrice diagonale et P est une matrice orthonormale. Les éléments diagonaux de $L = \{l_1, \dots, l_n\}$ sont les valeurs propres de la matrice de covariance Σ alors que les vecteurs colonnes de P sont les vecteurs propres de Σ .

Les valeurs propres l_i peuvent être obtenues par la solution de l'équation caractéristique

$$\det[\Sigma - l_i I] = 0 \quad (\text{II.7})$$

Pour les n valeurs propres. Connaissant les valeurs propres l_i , on peut déduire les vecteurs propres de la matrice de covariance par la solution de

$$(\Sigma - l_i I) A_i = 0 \quad (\text{II.8})$$

et puis,

$$u_i = \frac{A_i}{\sqrt{A_i^t A_i}} \quad (\text{II.9})$$

pour les n vecteurs propres. Notons que l'équation (II.9) est aussi appelée équation de normalisation. Il est intéressant de constater qu'ici les éléments des vecteurs propres u_i sont les cosinus directeurs des axes générés par une rotation dans l'espace des observations. Nous pouvons donc calculer tout un ensemble de valeurs à partir des ces nouveaux axes. En d'autres termes, on peut transformer les variables $\mathbf{x} = [x_1, \dots, x_n]^t$ en un nouvel ensemble de variables $\mathbf{y} = [y_1, \dots, y_n]^t$ par l'application de

$$y_i = \mathbf{u}_i^t (\mathbf{x} - \bar{\mathbf{x}}) \quad (\text{II.10})$$

y_i est la $i^{\text{ème}}$ composante principale de l'ensemble des observations. Les nouvelles variables y_i ont évidemment une moyenne nulle et elles possèdent une variance donnée par l_i (valeur propre de Σ). Les vecteurs de colonnes u_i de P sont orthonormaux car ils sont orthogonaux et de modules unitaires. Donc, si l'on utilise ces vecteurs pour obtenir les composantes principales, tel que montré par l'équation (II.10), on obtiendra des composantes principales qui ne seront pas corrélées et de variances égales aux valeurs propres de la matrice de covariance Σ .

En appliquant une transformation sur les vecteurs colonnes u_i , on peut obtenir de nouvelles propriétés intéressantes. L'une de ces transformations consiste à effectuer une mise à l'échelle de u_i par les valeurs propres de la matrice de covariance Σ :

$$v_i = u_i \sqrt{l_i} \quad (\text{II.11})$$

Les composantes obtenues par l'utilisation de $V = [v_1, \dots, v_n]^t$ seront de même unité que les variables originales. Nous retrouvons également une transformation dans laquelle les vecteurs u_i sont divisés par les valeurs propres de la matrice de covariance S :

$$w_i = \frac{u_i}{\sqrt{l_i}} \quad (\text{II.12})$$

Dans ce cas, les composantes obtenues seront de variance unitaire. Enfin, nous soulignons que l'on peut avoir une mesure de variabilité expliquée par chacune des composantes principales en utilisant les vecteurs transformés de l'équation (II.11) selon :

$$var = v_i v_i^t \quad \text{avec } i=1, \dots, n. \quad (\text{II.13})$$

Les valeurs obtenues par l'équation (II.13) peuvent être utilisées pour indiquer le pourcentage de la variabilité des caractéristiques représentées par chacune des composantes principales.

II.4.4- Détermination de la structure du modèle ACP

L'analyse en composantes principales recherche une approximation de la matrice initiale des données X par une matrice de rang inférieur issue d'une décomposition en valeurs singulières. La question qui se pose alors, et qui a été largement débattue dans la littérature, concerne le choix du nombre de composantes principales qui doit être retenu. De nombreuses règles sont proposées pour déterminer le nombre de composantes à retenir [8-9-10-11]. La plupart de ces règles sont heuristiques et donnent un nombre de composants subjectif. Toutefois, dans le cadre de l'application de l'ACP à la réduction des données, le nombre de composantes a un

impact significatif sur la phase finale de classification. Si peu de composantes sont utilisées, on risque de perdre des informations contenues dans les données de départ en projetant certains échantillons dans le sous-espace des résidus et donc avoir des erreurs de modélisation, ce qui provoque une mauvaise caractérisation.

Si par contre beaucoup de composantes sont utilisées, il y a risque d'avoir des composantes retenues (les composantes correspondantes aux valeurs propres les plus faibles parmi celles retenues dans le modèle) qui sont porteuses de redondance, ce qui est indésirable.

Dans leurs travaux de recherche, Qin et al (Qin 1998) [12] ont proposé une technique basée sur la variance de l'erreur de reconstruction des mesures; ce critère permet de prendre en compte la notion de redondance entre les variables.

II.4.5- Conclusion

Ce que nous pouvons retenir de cette méthode est que l'ACP est une technique utile pour la compression et la classification des données. Le problème consiste à réduire la dimensionnalité d'un ensemble des données (échantillon) en trouvant un nouvel ensemble de variables plus petit que l'ensemble originel des variables, qui néanmoins contient la plupart de l'information sur l'échantillon. Par information nous voulons parler de la variation présente dans l'échantillon et donnée par les corrélations entre les variables originelles. Les nouvelles variables, appelées composantes principales (PC), sont non corrélées, et sont ordonnées par fraction de l'information totale que chacune contient [13-14].

En termes précis, l'usage principal de l'ACP est la réduction de la dimensionnalité tout en conservant le plus d'information possible.

1^{ière} PC: contient la plus grande quantité de variation.

2^{nde} PC: contient en deuxième lieu la plus grande quantité de variation.

Une fois que nous aurons appliqué cette méthode nous aurons besoin de faire usage d'une autre méthode qui n'est pas descriptive mais plutôt une méthode prédictive. En l'occurrence la méthode des moindres carrés partiels (PLS).

II.5- La méthode des moindres carrés partiels (PLS)

II.5.1- Introduction – Problématique

De nombreux problèmes industriels peuvent être décrits sous la forme d'un modèle de régression, où l'on possède des variables X sur lesquelles on peut plus ou moins agir et des variables Y que l'on ne peut qu'observer.

L'objectif est alors de décrire les relations entre Y et X , en l'absence de modèle théorique.

Le problème est que le nombre de variables X est souvent très important par rapport au nombre d'observations. La régression PLS (Partial Least Squares Regression) est une méthode d'analyse des données spécialement construite pour l'étude de ce type de problème [15].

II.5.1.1- Présentation

La régression PLS est une extension du modèle de régression linéaire multiple. Dans sa forme la plus simple, un modèle linéaire spécifie la relation (linéaire) entre une (ou des) variables dépendantes (réponses) Y et un ensemble de variables prédictives X telles que :

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p \quad (\text{II.16})$$

Où les b_i sont les coefficients de régression.

Un exemple ordinaire est l'estimation (prédiction) du poids d'une personne en fonction de sa taille ou de son sexe en se basant sur une expression donnant le poids (Y) en fonction de la taille et du sexe (X_1 et X_2). On peut utiliser une régression linéaire pour estimer les coefficients de régression respectifs à partir d'un échantillon de données mesurant la taille et le poids et en observant le sexe des individus.

Dans beaucoup de problèmes d'analyses de données, l'estimation de la relation linéaire entre deux variables est adéquate pour décrire les données observées et pour faire de bonnes prévisions pour de nouvelles observations.

Le modèle de régression multiple a été étendu de plusieurs façons afin de s'adapter aux problèmes d'analyse de données plus sophistiqués. Il sert donc de base pour de nombreuses méthodes multivariées comme l'analyse discriminante, la régression sur composantes principales (PCR) et la corrélation canonique.

La régression PLS est une technique récente qui généralise et combine les caractéristiques de l'analyse sur composantes principales et de la régression multiple.

Elle est particulièrement utile quand on a besoin de prédire un ensemble de variables dépendantes à partir d'un ensemble très grand de variables explicatives (prédicteurs) qui peuvent être très fortement corrélées entre elles.

Quand les prédicteurs sont peu nombreux, non significativement colinéaires et ont une relation connue avec les réponses, alors la régression linéaire multiple est la meilleure méthode pour utiliser les données. Cependant si l'une de ces trois conditions n'est pas vérifiée, la régression linéaire multiple peut être inefficace et inappropriée.

La méthode PLS est donc une méthode pour construire des modèles de prédiction quand les facteurs sont nombreux et très colinéaires. Notons que cette méthode met l'accent sur la prédiction de la réponse et pas nécessairement sur la mise en évidence d'une relation entre les variables.

Ceci signifie que la PLS n'est pas appropriée pour désigner les variables ayant un effet négligeable sur la réponse, mais quand le but est la prédiction et qu'il n'y a pas besoin de limiter le nombre de variables mesurées, la PLS est un outil très utile.

II.5.1.2- Historique

La régression PLS tire son origine des sciences sociales, plus précisément des sciences économiques, Herman Wold 1966 [16]. Elle devient très populaire en chimie grâce au fils d'Herman, Svante.

La régression PLS est née de l'association de l'algorithme NIPALS (Non linear Iterative Partial Least Squares) développé par H. Wold [16] pour l'analyse sur composantes principales et de l'approche PLS proposée par H. Wold [16] pour l'estimation des modèles d'équations structurelles sur les variables latentes. Il en résulte une représentation "classique" de la régression PLS sous la forme d'un algorithme (remarquons qu'il y a plusieurs versions possibles de l'algorithme NIPALS aboutissant aux mêmes résultats).

Une méthode d'estimation alternative pour les composantes de la régression est l'algorithme SIMPLS de Jong [17].

1966 : Herman Wold [16] publie un ouvrage introduisant le PLS, alors appelé Non Linear Iterative Partial Least Squares NIPALS.

1983 : Svante Wold (le fils d'Herman) et Harold Mertens [18] adaptent NIPALS au problème de régression avec trop de prédicteurs et appellent PLS cette adaptation de l'algorithme.

1990 : Stone et Brooks [19] introduisent la PLS dans le contexte du "Continuum Regression" (elle ajoute un paramètre continu λ , autorisant la méthode de modélisation à varier continûment entre MLR "Multiple Linear Regression" ($\lambda = 0$), PLS ($\lambda = 0.5$) et PCR "Principal Component Regression" ($\lambda = 1$)). Ce qui signifie une première projection sérieuse du PLS dans un contexte statistique.

Plus récemment la littérature concernant la méthode PLS s'est largement étoffé, surtout dans le journal "Chemometric". De plus un ouvrage entièrement en français dédié à cette méthode a été écrit par Tenenhaus en 1998 [20].

II.5.2- Intérêt de la Régression PLS

La régression PLS pour le modèle linéaire s'applique où l'on possède peu d'observations sur des variables très corrélées et en très grand nombre, c'est le cas pathologique de la régression linéaire, là où les méthodes habituelles ne fonctionnent pas (n étant très petit on ne peut pas tendre vers l'infini).

II.5.2-1- Défaillances du modèle linéaire

Supposons que l'on a :

- p variables explicatives
- c variables réponses
- un échantillon de n observations

Nous allons considérer dans en premier temps $c = 1$

On appelle :

- $T_{n \times p}$ la matrice des observations sur les variables explicatives centrées
- $Y_{n \times 1}$ les observations sur les variables réponses centrées.

On suppose ici que $\text{rang}(T) = p$ (ce qui est faux en général)

Soit $X_{n \times p}$ la matrice des variables centrées réduites, c'est à dire X^j la ji^{eme} colonne de X s'écrit :

$$X^j = \frac{T^j}{\sqrt{\sum_{i=1}^n (T_i^j)^2}} \quad (\text{II.17})$$

Soit V la matrice des covariances empiriques ($V = X^T X$)

Le modèle linéaire s'écrit :

$$Y = X\beta + \epsilon \quad (\text{II.18})$$

Où $\epsilon \approx N(0, \sigma^2 I)$.

Alors l'estimateur de β s'écrit :

$$\hat{\beta} = (XX^T)^{-1}Y^TY \quad (\text{II.19})$$

Et

$$E(\hat{\beta}) = \beta.$$

On sait alors que l'erreur quadratique moyenne vaut :

$$MSE = E(\|\hat{\beta} - \beta\|_2^2) = \sigma^2 \text{trace}(V^{-1}) \quad (\text{II.20})$$

Or on sait aussi que l'ensemble des valeurs propres de V^{-1} : $\lambda(V^{-1})$, peut s'écrire

$\lambda(V^{-1}) = \frac{1}{\lambda(V)}$, alors l'erreur quadratique moyenne se réécrit :

$$MSE = \sigma^2 \sum_{j=1}^p \frac{1}{\lambda_j(V)}. \quad (\text{II.21})$$

Or avec les problèmes de multi colinéarité, $\lambda_j(V) \sim 0$ et donc la MSE devient très grande et les coordonnées de $\hat{\beta}$ ne sont plus interprétables.

On peut facilement montrer que $\text{trace}(V^{-1}) = \sum_{j=1}^p \frac{1}{1-R^2_{-j}}$ où R^2_{-j} est le coefficient R^2 quand on effectue la régression de la variable explicative X^j sur les autres variables.

Or R^2_{-j} sera voisin de 1 si X^j est très corrélé avec les autres variables, donc $1 - R^2_{-j}$ sera proche de 0 et la MSE sera infinie.

On comprend maintenant mieux pourquoi le problème de multicollinéarité influe sur la variance de l'estimateur.

II.5.3- La méthode PLS linéaire

II.5.3-1 Modèle de base

- Principe

Comme dans la régression linéaire multiple, le but principal de la régression PLS est de construire un modèle linéaire

$$Y = XB + E \quad (\text{II.23})$$

où $B_{p \times c}$ sont les coefficients de régression et $E_{n \times c}$ le terme de bruit pour le modèle.

Usuellement les variables dans X et Y sont centrées en soustrayant leur moyenne, et réduites en divisant par leur écart type. La régression en composantes principales et la régression PLS produisent toutes les deux des facteurs de scores comme des combinaisons linéaires des

variables prédictives originelles, de telle manière qu'il n'y a pas de corrélation entre les facteurs scores utilisés par le modèle de régression prédictive.

Remarque : La régression sur composantes principales et la régression PLS diffèrent dans la méthode utilisée pour extraire les facteurs de scores.

En résumé, la régression sur composantes principales produit une matrice de poids W reflétant la structure de covariances entre les variables prédictives alors que la régression PLS produit une matrice de poids W reflétant les structures de covariance entre les prédicteurs et les réponses.

Pour établir le modèle, la régression PLS produit une matrice de poids $W_{p \times c}$ pour X telle que $T = XW$, c'est à dire les colonnes de W sont des vecteurs de poids pour les colonnes de X produisant la matrice de facteurs de score $T_{n \times c}$ correspondante.

Ces poids sont calculés de telle façon qu'ils maximisent la covariance entre la réponse et les facteurs de score correspondants.

La procédure OLS (Ordinary Least Squares) pour le régression de Y sur T est alors utilisée pour produire Q , tel que $Y = TQ + E$.

Une fois Q calculé, nous avons $Y = XB + E$ où $B = WQ$ et le modèle de prédiction est complet.

Une matrice supplémentaire nécessaire pour une description complète de la procédure de régression PLS est la matrice P des facteurs qui donne le modèle $X = TP + F$ où F est la partie non expliquée du score de X .

II.5.3-2- La méthode

La méthode PLS est une méthode de régression linéaire de c variables réponses sur p variables explicatives toutes mesurées sur les mêmes n individus. Les tableaux des observations, notés respectivement Y et X , de dimensions $n \times c$ et $n \times p$, sont supposés centrés et éventuellement réduits par rapport aux poids (p_1, \dots, p_n) . On note $D = \text{diag}(p_1, \dots, p_n)$ la matrice diagonale des poids.

L'intérêt de la méthode comparée à la régression sur composantes principales (RCP), réside dans le fait que les composantes PLS sur les X , notées t , sont calculées "dans le même temps" que des régressions partielles sont exécutées. Cette simultanité leur confère un meilleur pouvoir prédictif que celles de la RCP. La question est donc d'examiner comment cette simultanité est mise en œuvre.

Notons $E_0 = X$ et $F_0 = Y$ les tableaux centrés et réduits au sens de D qui en général est égal à $\frac{1}{n}I_n$. La méthode procède par étapes successives permettant le calcul des composantes

principales. On notera A le nombre total d'étapes, c'est à dire de composantes indicées par $k = 1, \dots, A$.

II.6 Conclusion

Depuis leur création, les méthodes d'exploration des données ont été étudiées par de nombreux chercheurs et ont été adaptées à de nombreuses situations. Cependant, il reste encore beaucoup de choses à développer.

On sait que la plupart de ces méthodes ont meilleure application dans le cas où il y a plus de prédicteurs que d'observations et qu'il y a une forte colinéarité entre ces prédicteurs, sachant qu'il est toujours plus intéressant de pouvoir étudier plus précisément la qualité d'un modèle lorsqu'il a été construit.

En effet, en développant des tests de validité et en examinant les résidus on peut certainement mieux juger de la qualité d'un modèle. Même si le contexte (peu d'observations, beaucoup de variables et qui sont très corrélées) permet de construire des modèles cohérents, rien ne permet d'affirmer *à priori* qu'ils seront de bonne qualité.

Un autre problème sur lequel il faudrait se pencher est celui des données aberrantes- parfois dues à des erreurs numériques ou empiriques, ou simplement des cas particuliers qui ne peuvent pas être traités conventionnellement (en anglais "outliers"). Comment sont-ils traités par la méthode, et quand peut-on dire qu'une observation est aberrante ? Dans le cas de la régression linéaires, un certain nombre de procédures ont été développées, mais sont-elles utilisables dans le cas de la régression PLS ou faut-il en développer d'autres ? De nombreuses autres questions concernant la régression PLS et quasiment toutes les technique de datamining restent en suspend et ces méthodes ont encore beaucoup de progrès à faire.

Cependant, il est clair à la lumière des résultats incontestablement utiles de l'utilisation de ces techniques dans divers domaines de la recherche, qu'elles offrent un moyen puissant pour l'analyse dont on ne doit pas manquer l'utilisation en science des matériaux pour profiter de leurs prérogatives tout en demeurant prudents quant aux interprétations des résultats.

Références

- [1] Bellman R., *"Adaptive Control Processes : A Guided Tour "*, Princeton University Press, 1961.
- [2] Minami K., Nakajima H., Mark, R., and Toyoshima, T., *"Realtime discrimination of ventricular tachyarrhythmia with fourier-transform neural network"*, IEEE Trans. Biomed. Eng., vol. 46, no 2 , Feb 1999.
- [3] Jolliffe I.T., *"Principal component analysis"*, Springer-Verlag, New York, 1986
- [4] Diamantaras K.I., Kung S.Y., *"Principal component neural networks. Theory and applications"*, John Wiley and Sons , 1996.
- [5] Pearson K., *"On lines and planes of closest fit to systems of points in space"*, The London, Edinburgh and Dublin Phylosophical Magasine and Journal of Science, Vol. 6, pp. 559-572, 1901.
- [6] Spearman C., *"General intelligence objectively determined and measured"*, American Journal of Psychology, Vol. 15, pp. 201-293, 1904.
- [7] Hotelling H., *"Analysis of a complex of statistical variables into principal components"*, Journal of Educational Psychology, Vol. 24, pp. 417- 441 1933.
- [8] Besse, P., and Ferre, L., *"Sur l'usage de la validation croisée en analyse en composantes principales"*, Revue de Statistique Appliquée, XLI (1), pp. 71-76, 1993.
- [9] Ferre L., *"Selection of components in principal component analysis :A comparaison of methods"*, Computational Statistics and Data Analysis,pp. 669-682 , 1995.
- [10] Himes D.M, and al., *"Determination of the number of principalcomponent for disturbance detection and isolation"*, Proceedings of ACC,Baltimore, 1994.
- [11] Valle S., and al., *"Selection of the number of principal components :The variance of the reconstruction error criterion with a comparison to othermethods"*, Industrial & Engineering Chemistry Research, vol. 38, pp. 4389-4401 ,1999.
- [12] Qin S. J. and Dunia R., *"Determining the number of principal componentsfor best reconstruction"*, Proc. of the 5th IFAC Symposium on Dynamicsand Control of Process Systems, PP. 359-364, Corfu, Greece, 1988.
- [13] Lindsay I Smith, *" A tutorial on Principal Components Analysis"* February 26, 2002.
- [14] HARKAT M-F., *"Détection et localisation de défauts par analyse en composantes principales"*. Doctorat de l'Institut Polytechnique de Lorraine, 30 juin 2003.

- [15] P. Bastien, V. Esposito Vinzi et M. Tenenhaus. "PLS generalised linear regression". *Computational Statistics & Data Analysis*, 48(1): 17-46, 2005.
- [16] Wold, H. (1966), "*Estimation of principal component and related models by iterative least squares*", *Multivariate Analysis*, ed. P.R. Krishnaiah, Newyork : Academic Press, 391-420.
- [17] de Jong, S. (1993), SIMPLS : "*An alternative approach to partial least squares regression*", *Chemometrics and Intelligent Laboratory Systems*, vol. 18, pp. 251-263.
- [18] Wold, S., Martens, H. & Wold, H. (1983), "*The multivariate calibration problem in chemistry solved by the PLS method*", in Proc. March 1982, *Lecture Notes in Mathematics*, Springer Verlag, Heidelberg, PP. 286-293.
- [19] Stone, M. & Brooks, R. (1990), "*Continuum regression : Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression*", *Journal of the Royal Statistical Society, Series B*, 52(2), 237-269.
- [20] Tenenhaus, M. (1998), "*La régression PLS, théorie et pratique*" Paris : Technip.

III.1- Introduction

Dans ce chapitre nous présentons les résultats liés à l'utilisation de l'analyse en composante principale (ACP) et la régression (PLS), appliquées aux alliages d'Heusler, en portant l'accent sur les alliages **Full-Heusler**. Afin d'avoir une cohérence dans notre base de données on a pris en considération seulement les alliages d'Heusler qui cristallisent dans la structure $L2_1$. Les calculs ont été faits avec cette technique en utilisant le logiciel XLSTAT (voir annexe). L'idée principale est que nous générons une base de données comprenant des différents propriétés structurales et magnétique telles le paramètre de maille (a), le nombre d'électrons de valence (NVE), le moment magnétique (M) et enfin la température de curie (T_c). Grâce à cette base de données nous allons identifier les tendances entre les différentes propriétés de ces alliages ainsi que leurs corrélations avec la température de curie.

III.2- Base de données

III.2.1-Définition

Une base de données, usuellement abrégée en BD ou BDD, est un ensemble structuré et organisé permettant le stockage de grandes quantités d'informations afin d'en faciliter l'exploitation.

Dans ce travail on a choisi de travailler sur deux familles d'alliages Heusler ferromagnétiques à base de Cobalt (Co) et l'autre antiferromagnétiques à base du Manganèse (Mn).

III.2.2- Présentation des bases de données

- Alliages à base du Cobalt

Composé	a (Å)	M (μB)	NVE
Co ₂ CrAs	5,797	4,93	29
Co ₂ CrBi	6,217	4,87	29
Co ₂ CrSb	6,021	4,9	29
Co ₂ MnAs	5,807	5,955	30
Co ₂ MnBi	6,1889	5,9775	30
Co ₂ MnSi	5,74	5,03	29
Co ₂ MnSn	5,996	5,085	29
Co ₂ MnSb	6,036	6,01	30
Co ₂ VSi	5,683	2,99	27
Co ₂ VSn	6,018	3,03	27
Co ₂ VGe	5,777	3,02	27
Co ₂ CrSi	5,649	3,9825	28
Co ₂ CrSn	5,992	3,995	28
Co ₂ MnGe	5,65	5,09	29

Tableau III-1 : alliages à base de cobalt (paramètre de maille, moment magnétique et nombre d'électrons de valence).

Composé	a (Å)	M (μB)	NVE	T (k)	[R]	r _Z (Å)
Co ₂ MnSi	5,74	4,96	29	970	[4]	1,1
Co ₂ MnSn	5,996	5,03	29	978	[5]	1,45
Co ₂ MnGe	5,65	5,09	29	985	[4]	1,25
Co ₂ MnAl	5,75	4,125	28	697	[4]	1,25
Co ₂ MnGa	5,77	4,17	28	700	[2]	1,3
Co ₂ VGa	5,78	2,05	26	454	[4]	1,3
Co ₂ TiAl	5,84	1,07	25	148	[3]	1,25
Co ₂ CrGa	5,71	3,045	27	495	[1]	1,3
Co ₂ TiGa	5,85	1,06	25	130	[3]	1,3
Co ₂ TiSn	6,088	2,09	26	370	[3]	1,45

Tableau III-2 : alliages à base de cobalt (paramètre de maille, moment magnétique, nombre d'électrons de valence, température de curie et rayon atomique de Z).

- Alliages à base du Manganèse

Composé	a (Å)	M (μB)	NVE	Mx (μB)	My (μB) [R]
Mn ₂ CrAs	5,732	1	25	-0,48	1,45 [6]
Mn ₂ CrGe	5,715	0,03	24	-0,5	0,53 [6]
Mn ₂ CrP	5,582	1	25	-0,32	1,28 [6]
Mn ₂ CrSb	5,943	1	25	-0,58	1,57 [6]
Mn ₂ CrSi	5,623	-0,03	24	-0,12	0,08 [6]
Mn ₂ CrSn	5,959	0,02	24	-1,6	1,6 [6]
Mn ₂ MnAs	5,731	2,01	26	-0,62	2,59 [6]
Mn ₂ MnGe	5,749	1,04	25	-1,52	2,53 [6]
Mn ₂ MnP	5,579	2,01	26	-0,48	2,44 [6]
Mn ₂ MnSi	5,621	1,15	25	-0,96	2,09 [6]
Mn ₂ MnSn	5,979	1,02	25	-2	3,01 [6]
Mn ₂ TiAs	5,822	-0,99	23	-1,292	0,281 [7]
Mn ₂ TiBi	6,202	-1,03	23	-1,494	0,43 [7]
Mn ₂ TiGe	5,876	-1,99	22	-2,564	0,626 [7]
Mn ₂ TiSb	6,075	-1,02	23	-1,418	0,365 [7]
Mn ₂ TiSi	5,783	-1,98	22	-2,486	0,435 [7]
Mn ₂ TiSn	6,136	-2,05	22	-2,828	0,711 [7]
Mn ₂ VSi	5,672	-1	23	-1,52	0,48
Mn ₂ VGe	5,767	-1	23	-1,6	0,551
Mn ₂ VSn	6,023	-1,02	23	-1,9	0,82

Tableau III-3 : alliages à base de manganèse (paramètre de maille, moment magnétique total, nombre d'électrons de valence, moment magnétique de X et moment magnétique de Y).

Composé	a (Å)	M (μB)	NVE	Tc (k)	[R]	Mx(μB)	My (μB)	rZ (Å)
Mn ₂ TiSi	5,783	-1,98	22	424	[8]	-2,468	0,435	1,1
Mn ₂ TiGe	5,876	-1,99	22	398	[8]	-2,564	0,526	1,25
Mn ₂ TiSn	6,136	-2,05	22	354	[8]	-2,828	0,711	1,45
Mn ₂ TiAs	5,822	-0,99	23	132	[8]	-1,292	0,281	1,15
Mn ₂ TiSb	6,075	-1,02	23	156	[8]	-1,418	0,365	1,45
Mn ₂ TiAl	5,95	-2,98	21	665	[8]	-3,66	0,57	1,25
Mn ₂ TiGa	5,96	-2,95	21	663	[8]	-3,68	0,6	1,3

Tableau III-4 : alliages à base de manganèse (paramètre de maille, moment magnétique total, nombre d'électrons de valence, température de curie, moment magnétique de X, moment magnétique de Y et le rayon atomique de Z).

III.3 Analyse des données

III.3.1- L'analyse en composante principale

❖ Choix des axes

Pour un plan factoriel donné, on regarde le pourcentage de l'information du nuage initial retranscrite par le plan factoriel. Ainsi, un axe expliquant moins de 10% sera rarement intéressant, les axes sont rangés dans l'ordre décroissant, de telle manière que le premier plan factoriel - constitué par les deux premiers axes factoriels - soit toujours celui qui est le plus riche en renseignement sur les propriétés du nuage étudié. Dans notre exemple, les axes F1 (PC1) et F2 (PC2) sont les plus riches en information (*fig. III.2*).

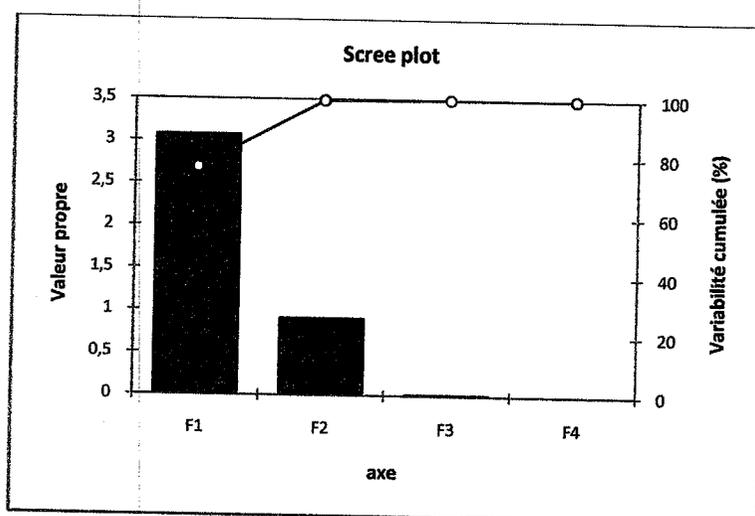


Figure III.1 : Représentation graphique des axes.

D'après les figures (III-2) et (III-3) on voit que PC1 capture 73.02% et 76.81% pour les bases sur les tableaux (III-1) et (III-2) respectivement. L'axe PC2 capture 26.94% et 22.78% pour les mêmes bases

❖ **Interprétation des données :**

Le premier résultat intéressant à identifier dans l'analyse en composantes principales est la matrice de corrélations.

Alliages à base du Cobalt

Les deux tableaux suivants correspondant aux matrices de corrélation des alliages d'Heusler ferromagnétiques à base du Co pour les nombre d'électrons de valence et la température de curie respectivement:

Variables	a (Å°)	M (μB)	NVE
a (Å°)	1	0,332	0,344
M (μB)	0,332	1	0,999
NVE	0,344	0,999	1

Tableau III-5 : Matrice de corrélation des alliages à base du Co.

Variabiles	a (Å)	M (μB)	NVE	T (k)	rz (Å)
a (Å)	1	-0,305	-0,304	-0,275	0,768
M (μB)	-0,305	1	0,999	0,985	-0,187
NVE	-0,304	0,999	1	0,988	-0,197
T (k)	-0,275	0,985	0,988	1	-0,181
rz (Å)	0,768	-0,187	-0,197	-0,181	1

Tableau III-6 : Matrice de corrélation des alliages à base du Co.

Pour le premier tableau (III-5), on remarque que le moment magnétique des alliages est fortement corrélé au nombre d'électrons de valence avec la valeur (0.999).

Pour le deuxième (III-6), d'une part, la température de curie est fortement corrélée au moment magnétique et au nombre d'électron de valence avec les valeurs (0.987) et (0.989) respectivement, et d'autre part, on constate que le paramètre de maille (a) est fortement corrélé avec le rayon atomique de l'élément z avec la valeur (0.768)

L'analyse des résultats de l'ACP pour les alliages d'Heusler ferromagnétiques est donnée par les graphes des observations « score plot » et des variables « loading plot ».

Les « score plot » donnent des informations sur les échantillons et traite la façon dont ils se comportent et les « loading plot » donnent des informations sur les relations qui existent entre les propriétés. Le résultat des « score plot » de cette analyse est montré sur les figures (III-2), (III-3), (III-4) et (III-5).

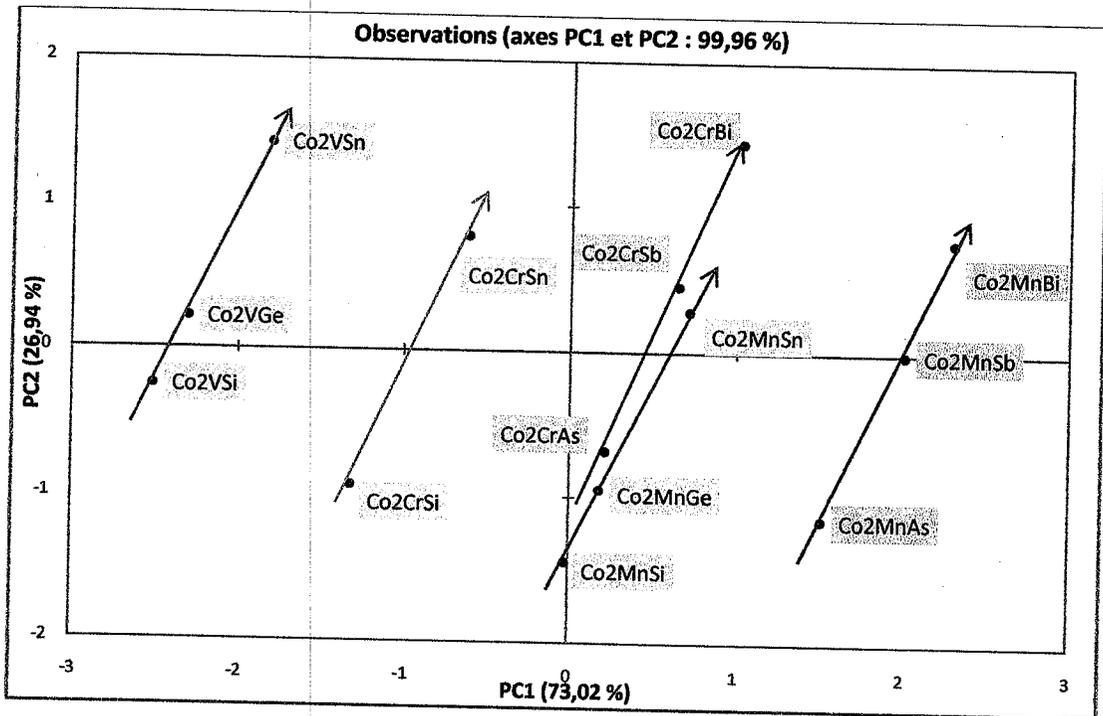


Fig. III-2 : Graphe des observations pour les alliages à base du Co.

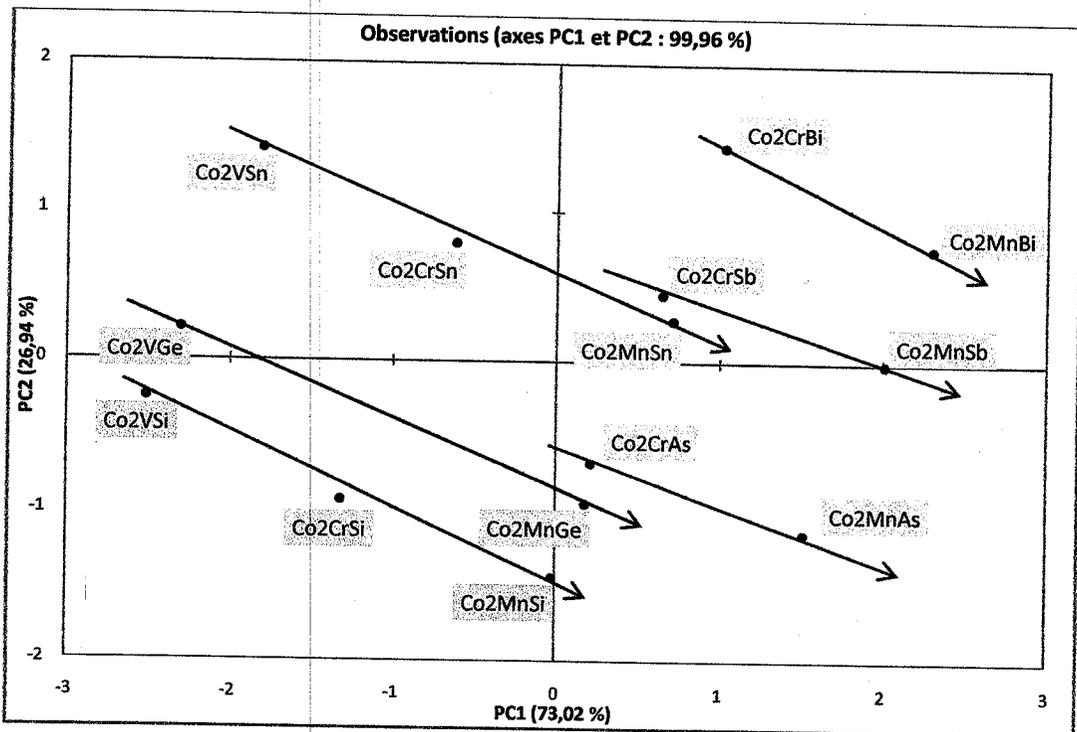


Fig. III-3 : Graphe des observations pour les alliages à base du Co.

Pour cette analyse le signe de chaque composante principale n'as aucune signification, en observant les deux graphes ci-dessus on remarque pour le premier que la variation selon l'axe PC2 est une variation du numéro atomique de l'élément Z, quand la valeur de l'axe PC2 augmente le numéro atomique de Z et le paramètre de maille de l'alliage augmentent.

Et pour le deuxième, on observe que la variation se fait plutôt selon l'axe PC1, quand la valeur de PC1 augmente le nombre d'électrons de valence et le moment magnétique de l'alliage augmentent.

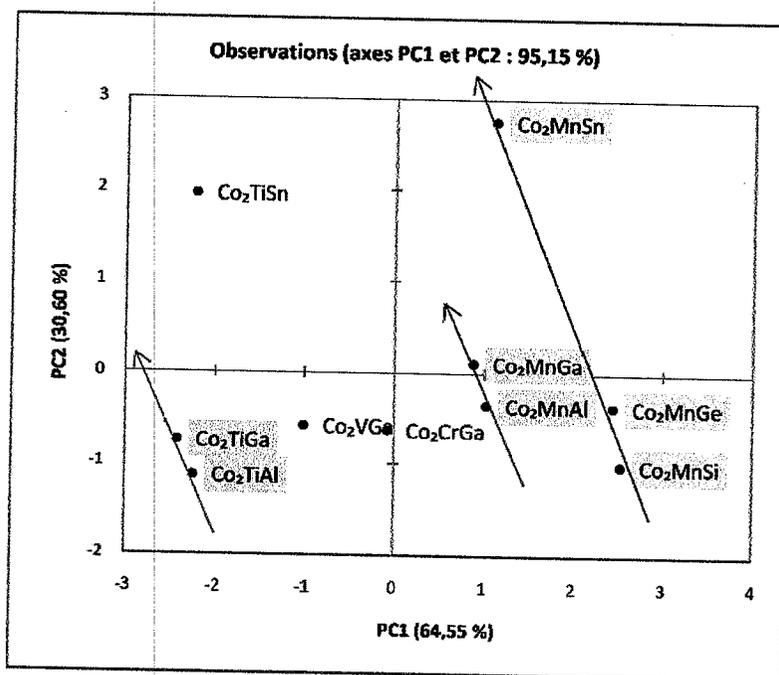


Fig. III-4 : Graphe des observations pour les alliages à base du Co.

Pour le graphe ci-dessus, l'augmentation de la valeur PC2 est toujours en fonction du rayon atomique de l'élément Z et le paramètre de maille.

Quand au graphe sur la figure (III-5), l'augmentation de la valeur PC1 est en fonction de l'augmentation du moment magnétique et la température de curie de l'alliage.

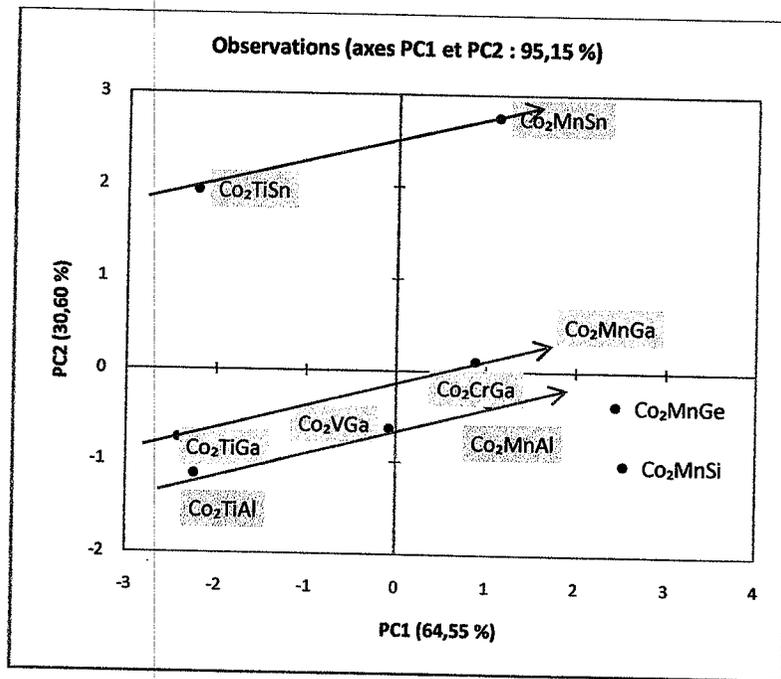


Fig. III-5 : Graphe des observations pour les alliages a base du Co.

Le graphe qui représente les variables (les propriétés) ou les descripteurs est le (loading plot) et celui qui représente les observations est (score plot), dans notre cas, une des observations représentées sont les alliages d'Heusler ferromagnétiques à base du Cobalt, les figures (III-6), (III-7) représentent la disposition des variables correspondant aux observations représentées sur les figures [(III-2), (III-3)] et [(III-4), (III-5)] respectivement. On peut classer les résultats en deux clusters pour les deux cas.

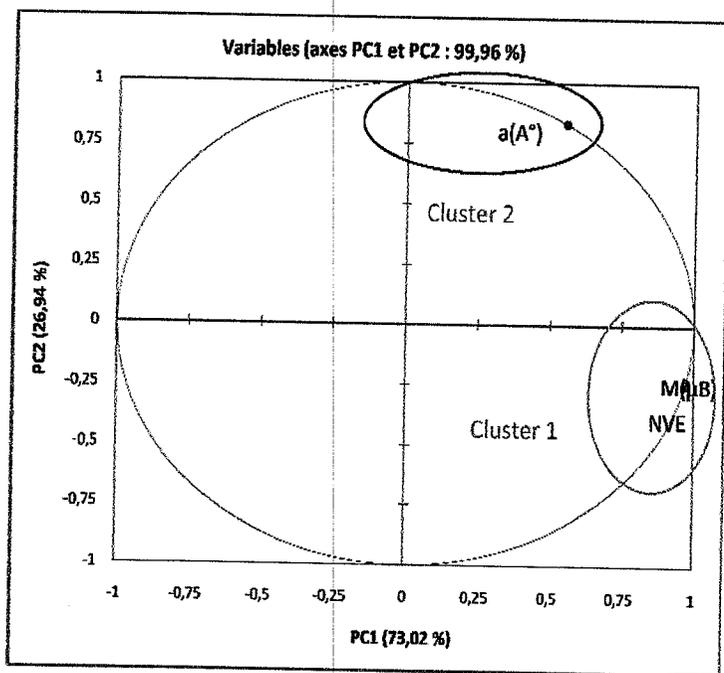


Fig. III-6 : Graphe des variables (loading plot) des alliages d'Heusler à base du Co

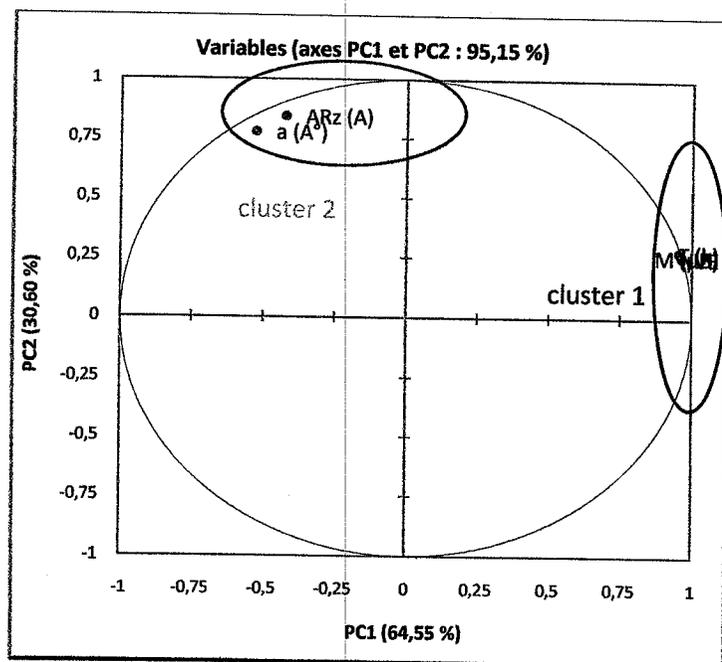


Fig. III-7 : Graphe des variables (loading plot) des alliages d'Heusler à base du Co

Les cercles de corrélation représentés par les figure (III-3-a) et (III-3-b) est aussi utile pour interpréter la signification des axes. Dans notre cas l'axe PC1 est clairement lié au cluster 1 et l'axe PC2 au cluster 2, par conséquent les propriétés dans le cluster 1 sont inversement corrélées avec les propriétés dans le cluster 2, alors que les propriétés dans le même cluster sont fortement corrélées. On peut déduire que la température de curie et le

moment magnétique sont fortement corrélés avec le nombre d'électrons de valence et inversement corrélés avec le paramètre de maille et le rayon atomique de z qui à leurs tour sont fortement corrélés l'un avec l'autre , en d'autres termes, plus le paramètre de maille de l'alliage et le rayon atomique de z sont petits plus sa température de curie et son moment magnétique sont grands, c'est le cas du Co_2MnGe avec un moment magnétique ($M=5.09\mu\text{B}$) et une température de curie ($T=985^\circ\text{K}$).

Les résultats précédents peuvent être confirmés par les graphes des biplot superposant les échantillons et leurs propriétés :

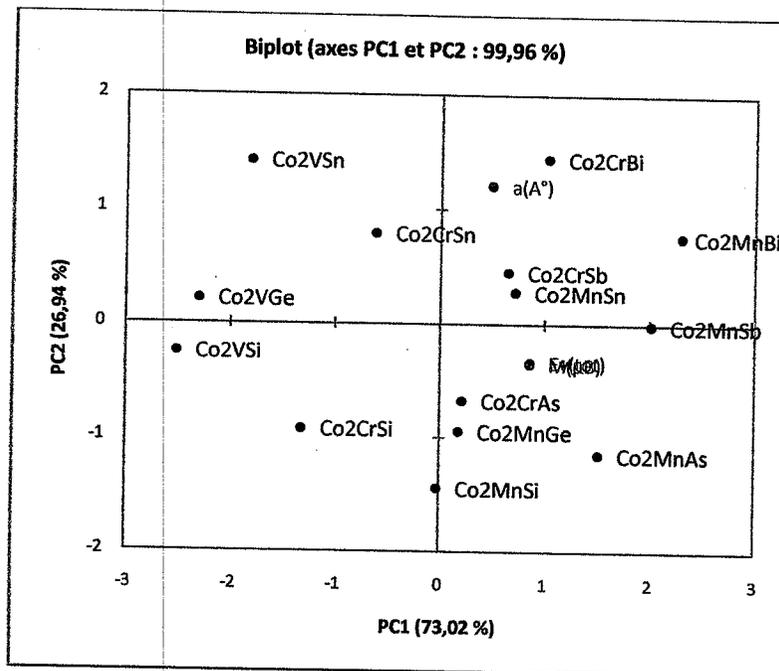


Fig. III-8 : graphe superposant les échantillons et les propriétés (biplot).

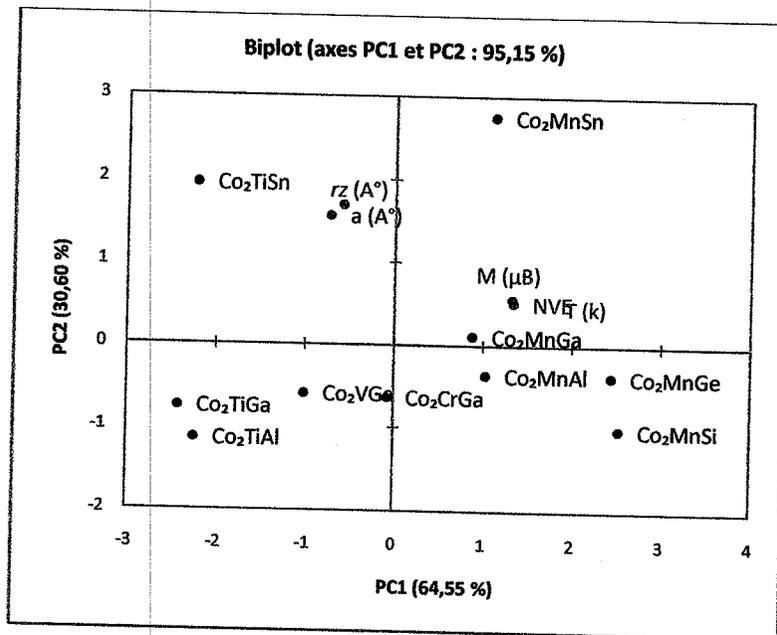


Fig. III-9: graphe superposant les échantillons et les propriétés (biplot).

Alliages à base du Manganèse

Les deux tableaux suivants correspondant aux matrices de corrélation des alliages d'Heusler antiferromagnétiques à base du Mn pour les nombre d'électrons de valence et la température de curie respectivement:

Variables	a (A°)	M _{tot} (μB)	NVE	My (μB)	Mx (μB)
a (A°)	1	-0,502	-0,493	-0,210	-0,594
M _{tot} (μB)	-0,502	1	1,000	0,812	0,731
NVE	-0,493	1,000	1	0,807	0,735
My(μB)	-0,210	0,812	0,807	1	0,194
Mx(μB)	-0,594	0,731	0,735	0,194	1

Tableau III-7 : matrice de corrélation des alliages à base du Mn (a, M, NVE, My et Mx)

Variables	a (Å)	M (μB)	NVE	Tc (k)	Mx(μB)	My (μB)	rz (Å)
a (Å)	1	-0,047	-0,021	-0,037	-0,130	0,528	0,976
M (μB)	-0,047	1	0,999	-0,991	0,995	-0,746	0,051
NVE	-0,021	0,999	1	-0,994	0,991	-0,725	0,076
Tc (k)	-0,037	-0,991	-0,994	1	-0,974	0,663	-0,124
Mx(μB)	-0,130	0,995	0,991	-0,974	1	-0,810	-0,029
My (μB)	0,528	-0,746	-0,725	0,663	-0,810	1	0,437
rz (Å)	0,976	0,051	0,076	-0,124	-0,029	0,437	1

Tableau III-8 : matrice de corrélation des alliages à base du Mn (a, M, NVE, Tc, Mx, My et rz).

On remarque dans le tableau (III-7) que le nombre d'électrons de valence est fortement corrélé avec les moments magnétiques Mx, My et le moment magnétique total M_{tot} avec les valeurs **(0.735)**, **(0.807)** et **(1.00)** respectivement et inversement corrélé avec le paramètre de maille (a) avec la valeur **(-0.502)**.

Dans le tableau (III-8) le nombre d'électrons de valence des alliages est fortement corrélé avec le moment magnétique Mx et le moment magnétique total M_{tot} avec les valeurs **(0.991)**, **(0.999)** respectivement. Le moment magnétique My est fortement corrélé avec La température de curie avec la valeur **(0.663)**, alors que cette dernière est inversement corrélée avec le moment magnétique total et le nombre d'électrons de valence avec les valeurs **(-0.991)** et **(-0.994)** respectivement. On remarque aussi que le paramètre de maille est fortement corrélé avec le rayon atomique de z avec la valeur **(0.976)**

L'analyse des résultats de l'ACP pour les alliages d'Heusler antiferromagnétiques à base du Mn est donnée par les graphes des observations (score plot) et des variables (loading plot) suivants :

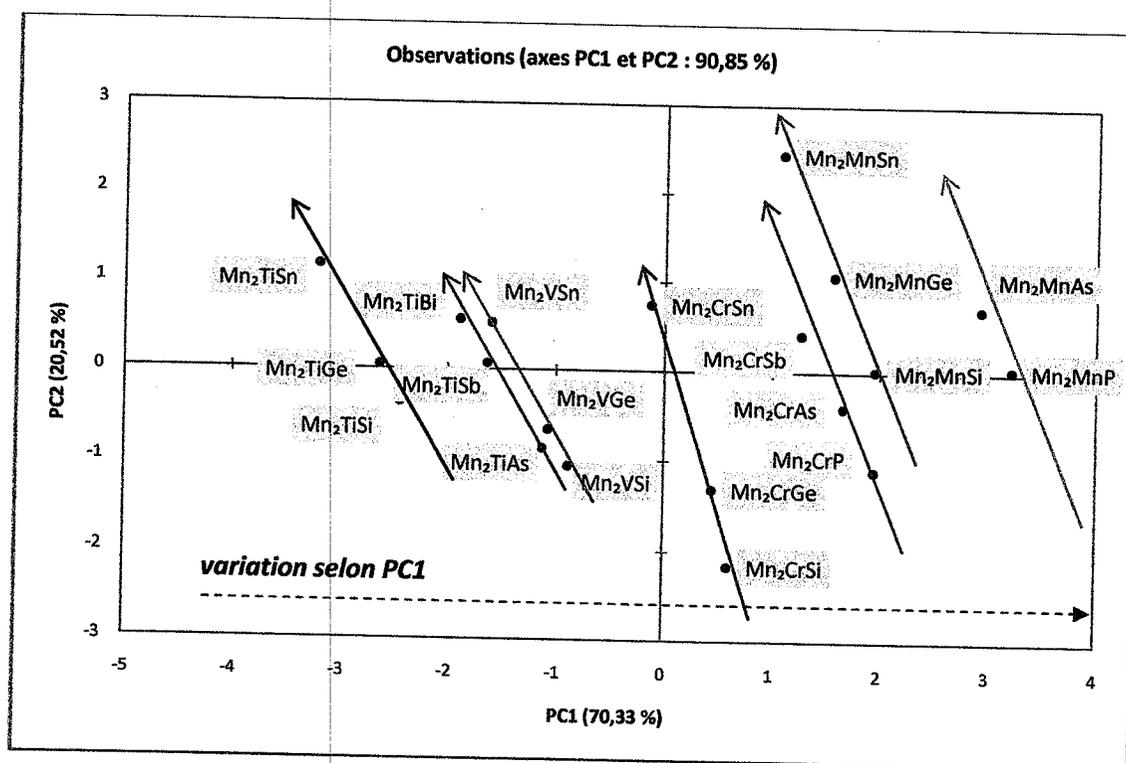


Fig. III-10 : Graphe des observations pour les alliages à base du Mn.

Sur le graphe ci-dessus, on constate que la variation est suivant l'axe PC2, quand la valeur du PC2 augmente le paramètre de maille de l'alliage augmente.

Par contre suivant l'axe PC1 l'augmentation de sa valeur suit l'augmentation du nombre d'électrons de valence et les moments magnétiques M_x , M_y et M_{tot} .

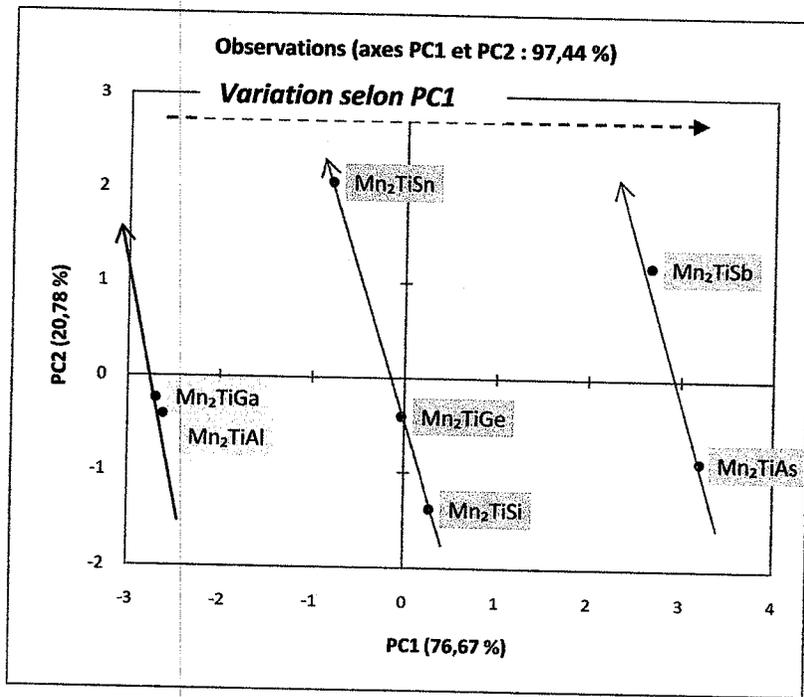


Fig. III-11 : Graphe des observations pour les alliages à base du Mn.

Le graphe (III-11) montre que l'augmentation de la valeur PC2 est en fonction de l'augmentation du paramètre de maille et le rayon atomique de l'élément Z. Et la valeur de l'axe PC1 augmente avec le nombre d'électrons de valence, le moment magnétique total et le moment magnétique de l'élément X. par contre la température de curie et le moment magnétique de l'élément Y diminue.

Comme précédemment, on peut classer les résultats en deux clusters pour le premier cas et trois clusters pour le deuxième cas :

On peut remarquer pour le graphe (III-12) que les moments magnétiques et le nombre d'électrons de valence appartiennent au même cluster1 clairement lié à l'axe PC1, alors ils sont fortement corrélés entre eux mais en même temps sont inversement corrélés avec le cluster 2 lié à l'axe PC2.

Pour le graphe (III-13), on remarque trois clusters distincts, cluster 1 et cluster 2 liés à l'axe PC1 et cluster 3 lié à l'axe PC2. Par conséquent, les propriétés dans le cluster 1 sont inversement corrélées avec les propriétés dans le cluster 2, alors que les propriétés dans le même cluster sont fortement corrélées entre eux.

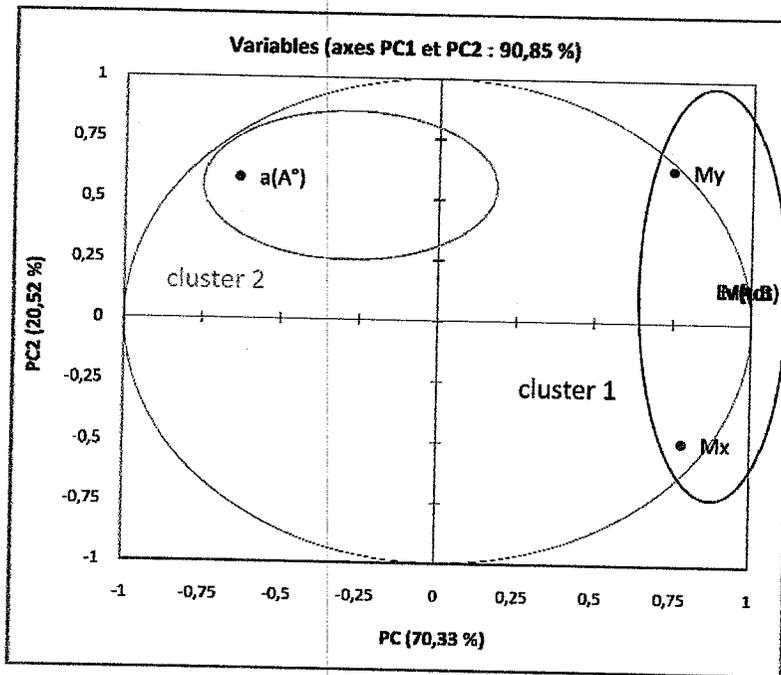


Fig. III-12 : Graphe des variables (loading plot) des alliages d'Heusler à base du Mn

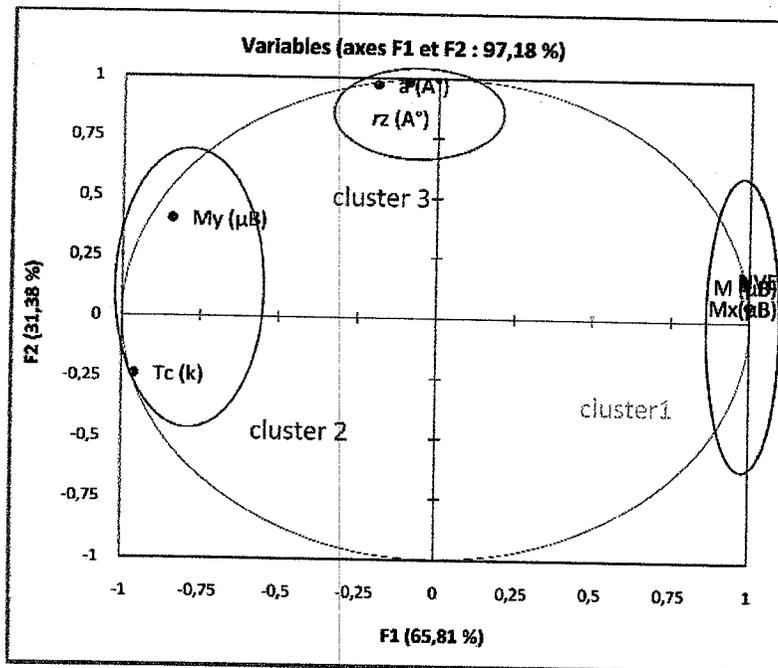


Fig. III-13 : Graphe des variables (loading plot) des alliages d'Heusler à base du Mn

Les résultats précédents peuvent être confirmés par les graphes des biplot superposant les alliages et leurs propriétés :

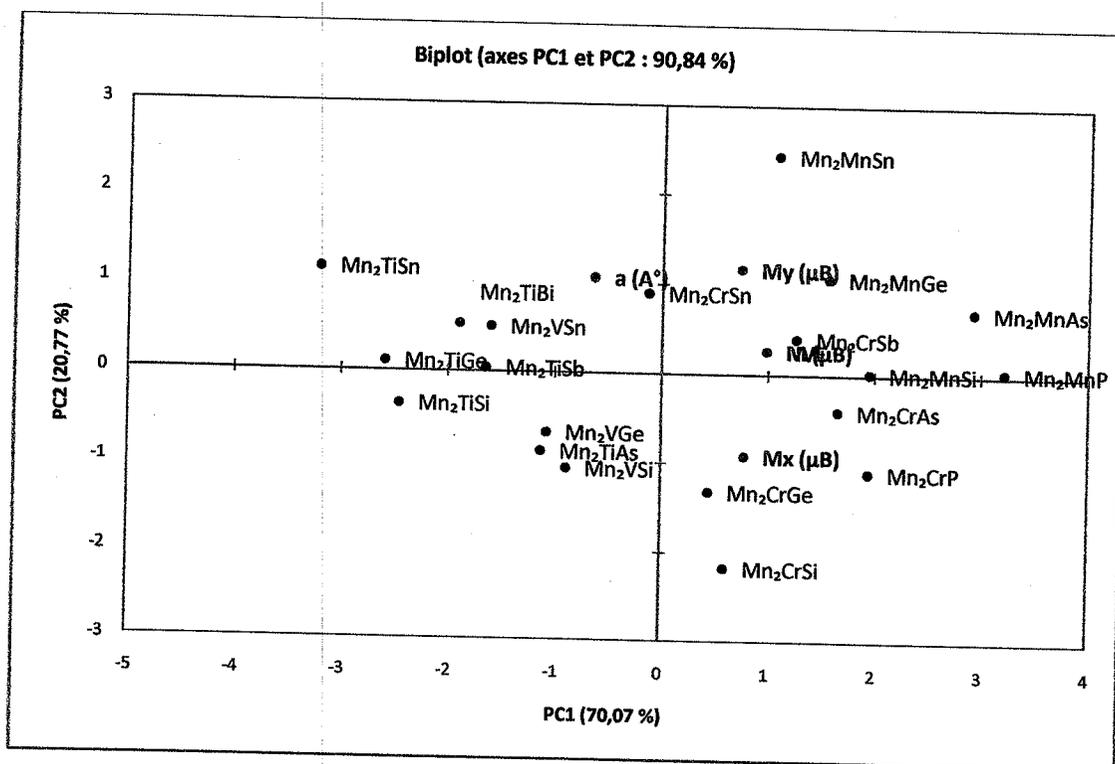


Fig.III-14 : graphe superposant les échantillons et les propriétés (biplot).

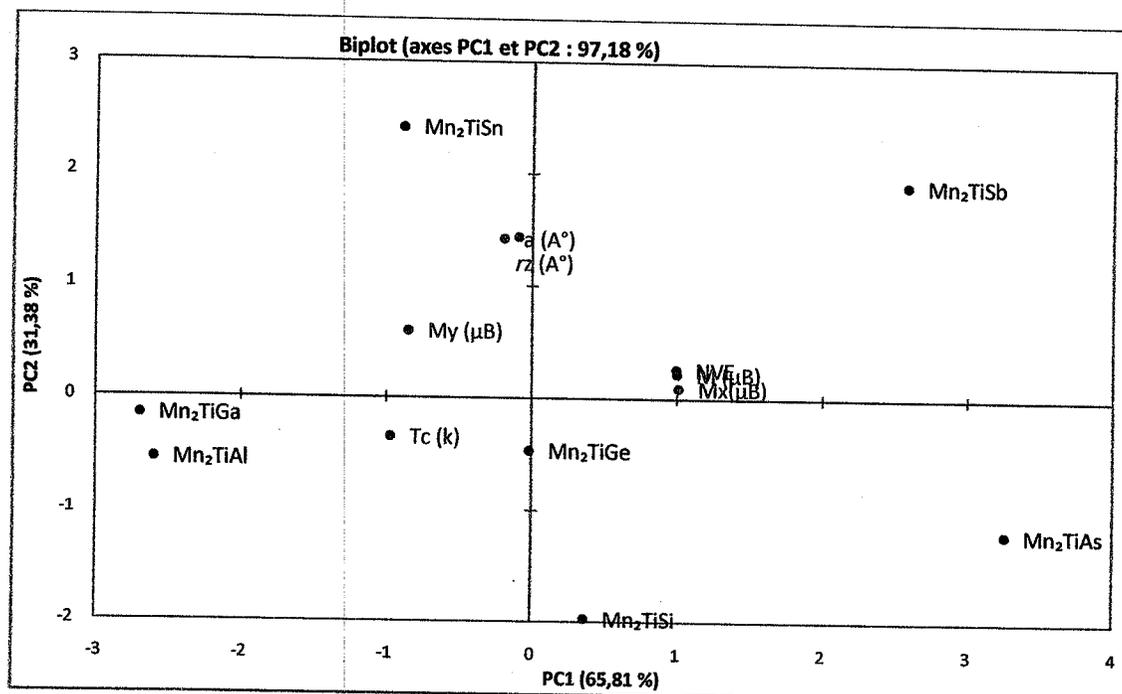


Fig.III-15 : graphe superposant les échantillons et les propriétés (biplot).

Comme dans l'expérience, les calculs prennent un temps considérables pour l'exécution. Employer la technique décrite ici peut mener à une réduction de nombre de calcul. La logique présentée ici peut être appliquée à n'importe quel système avec tout nombre d'échantillons et de descripteurs. La combinaison de l'informatique avec des données et des propriétés physiques calculées tiendra compte de la plus grande compréhension des relations entre structure et propriété. Avec cette connaissance, des matériaux peuvent alors être conçus pour optimiser les propriétés désirées.

L'utilisation de PCA ici démontre comment l'informatique peut être employée pour avoir une visualisation direct de l'information pour déterminer ce qu'est nécessaire et utile. En plus, les propriétés présentées ici peuvent être combinées avec les propriétés calculées.

III.3.2 Le modèle PLS (Partial Least Square)

Vu que la température est un paramètre crucial dans l'étude et la conception des alliages d'Heusler, nous allons essayer de prédire des températures de curie à partir des propriétés magnétique et structurales que nous avons utilisé dans notre base de données.

Tout d'abord ont a essayé de trouver une relation entre le moment magnétique et le nombre d'électrons de valence, pour cela on a fait appel à la méthode des moindres carrés partiels (PLS).

Après l'application de cette méthode nous avons remarqué qu'il existe une relation linéaire entre le moment magnétique et le nombre d'électrons de valence, dans notre cas donnée par l'équation suivante :

$$M_{\text{tot}} (\mu\text{B}) = -24,2673270440252 + 1,01154088050314 * \text{NVE}$$

Comme l'on peut distinguer, cette équation est similaire à celle de Slater-Pauling [9].

Certains alliages Heusler ont un nombre d'électrons de valence identique, mais des moments magnétique et températures de curie suffisamment différentes et pour ne pas créer des confuses, lors d'établissement de l'équation reliant la température de curie et le moment magnétique des alliages, on a opté pour rajouter un autre paramètre structurale qui a joué un rôle très important dans notre analyse en composantes principales (PCA), ce paramètre est le rayon atomique de l'élément Z, ce dernier qui a une contribution dans la

stabilité de la structure de l'alliage. Le rajout de ce paramètre va attribuer la différence dans nos prédictions.

III.3.2.a-Mesures de l'importance de variables

Dans le modèle PLS une variable X_k peut être importante pour la modélisation de Y , ces variables sont identifiées par des grands coefficients de régression b_k . Un résumé de l'importance d'une variable X pour X et Y à la fois est donné par le graphe des VIP (Variable Importance for the Projection) pour chacune des variables explicatives, sur chaque composante. Cela permet d'identifier rapidement quelles sont les variables explicatives les plus importantes sur l'ensemble des modèles (Figure III-16).

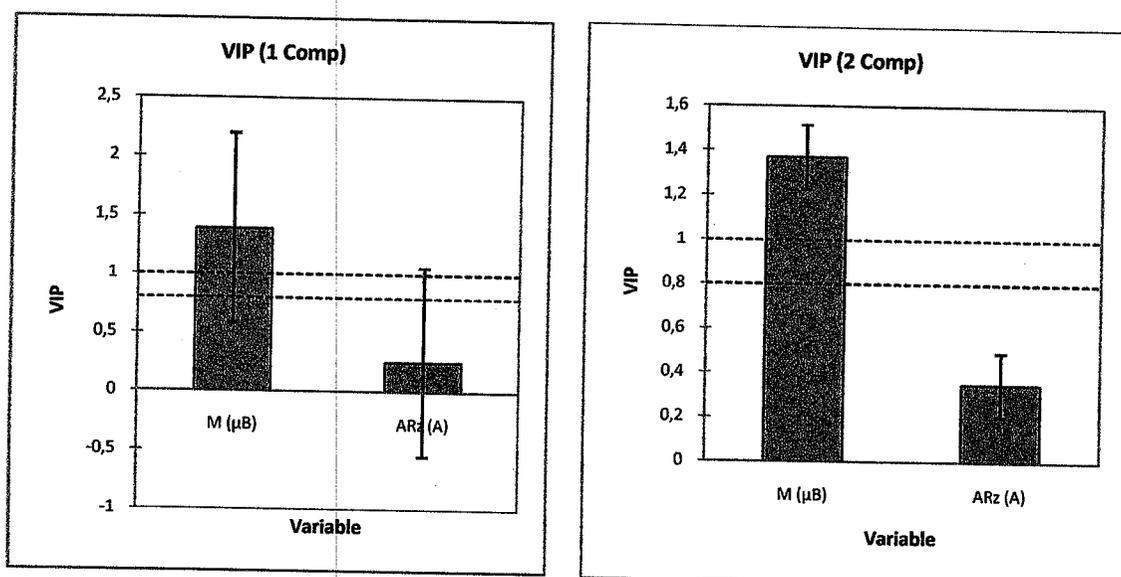


Fig.III-16 : exemple de VIP pour les deux premières composantes.

III.3.2.b-Paramètres et équations du modèle

Le tableau des paramètres des modèles correspondant à chacune des variables dépendantes est affiché sur le tableau III-9. Les équations sont alors données afin de les utiliser ultérieurement.

Variable	T (k)
Constante	-73,525
M_{tot} (μB)	199,193
r_z (A°)	11,633

Variable	Tc (k)
Constante	125,842
M_{tot} (μB)	-260,995
r_z (A°)	-193,561

Tableau III-9 : paramètre du modèle pour les alliages ferromagnétiques et antiferromagnétiques.

L'équation du modèle pour les alliages Heusler à base du cobalt :

$$T_c (k) = -73,5249158683449 + 199,192574872057 M_{tot} (\mu B) + 11,6327325186035 r_z (A^\circ)$$

L'équation du modèle pour les alliages Heusler à base du manganèse :

$$T_c (k) = 125,841573273148 - 260,994524073037 M_{tot} (\mu B) - 193,561404354373 r_z (A^\circ)$$

III.3.2.c- Les résidus et prédictions

Le modèle PLS permet aussi de donner des graphes correspondant aux coefficients normalisés ainsi que des graphes des résidus et prédictions permettant d'identifier la qualité du modèle. Nous obtenons alors des résultats plus riches en prenant les graphiques ci-dessous qui correspondent aux coefficients normalisés pour les deux modèles avec deux composantes et ceci pour la seule variable la température de curie.

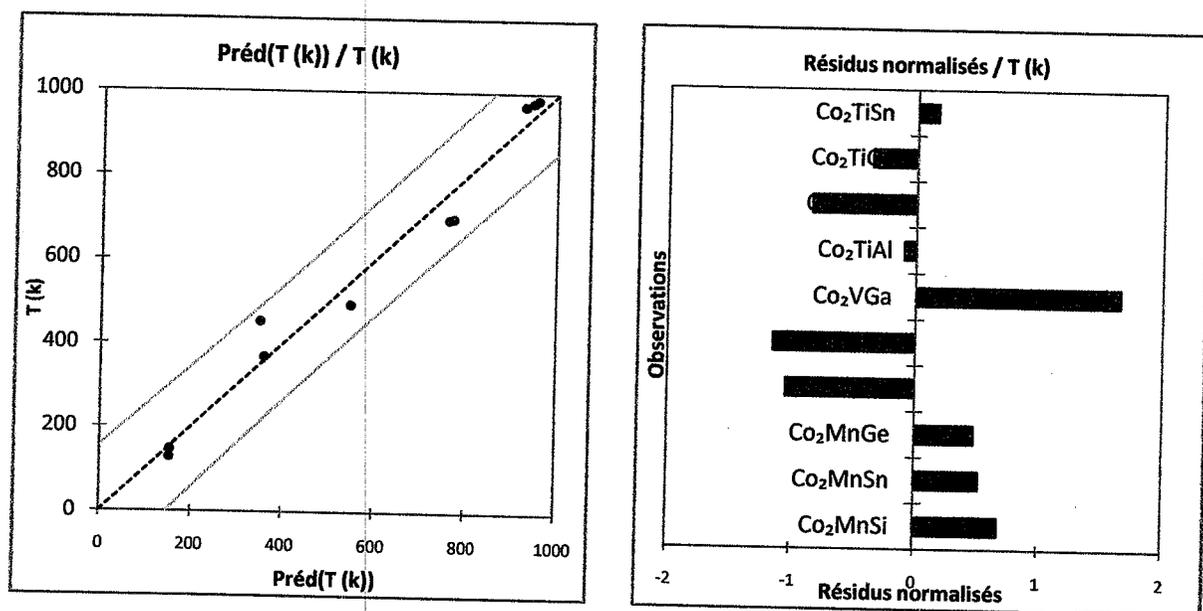


Fig. III-17 : Graphes des résidus et prédictions correspondant à la variable T_c pour les alliages à base du cobalt.

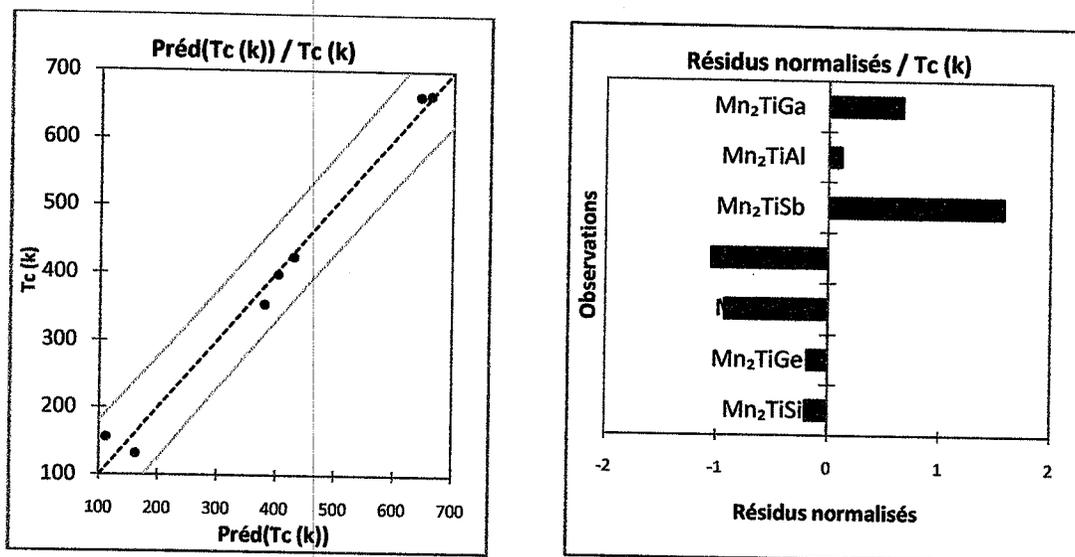


Fig. III-18: Graphes des résidus et prédictions correspondant à la variable T_c pour les alliages à base du manganèse.

L'analyse du modèle correspondant à la variable T_c nous permet de conclure que les ensembles des alliages à base du cobalt et à base du manganèse sont significatives et

influencent le modèle. Nous remarquons aussi dans le graphe des prédictions que la variation est quasiment linéaire entre les valeurs de T_c prédites en fonction de celle utilisée en entrée, ce qui indique bien la bonne qualité du modèle.

Observation	Poids	T_c (k)	T_c Préd (k)	Résidu
Co ₂ MnSi	1	970,000	927,266	42,734
Co ₂ MnSn	1	978,000	945,281	32,719
Co ₂ MnGe	1	985,000	954,906	30,094
Co ₂ MnAl	1	697,000	762,685	-65,685
Co ₂ MnGa	1	700,000	772,231	-72,231
Co ₂ VGa	1	454,000	349,942	104,058
Co ₂ TiAl	1	148,000	154,152	-6,152
Co ₂ CrGa	1	495,000	548,139	-53,139
Co ₂ TiGa	1	130,000	152,742	-22,742
Co ₂ TiSn	1	370,000	359,655	10,345

Tableau III-10 : Prédiction et résidus de la variable T_c pour les alliages à base du cobalt.

Observation	Poids	T_c (k)	T_c Préd (k)	Résidu
Mn ₂ TiSi	1	424,000	429,693	-5,693
Mn ₂ TiGe	1	398,000	403,269	-5,269
Mn ₂ TiSn	1	354,000	380,216	-26,216
Mn ₂ TiAs	1	132,000	161,631	-29,631
Mn ₂ TiSb	1	156,000	111,392	44,608
Mn ₂ TiAl	1	665,000	661,653	3,347
Mn ₂ TiGa	1	663,000	644,146	18,854

Tableau III-11 : Prédiction et résidus de la variable T_c pour les alliages à base du manganèse.

III.3.2.d- Prédiction de la température de curie pour quelques alliages d'Heusler

L'utilisation de PLS ici démontre comment l'informatique peut être employée pour utiliser l'information pour déterminer ce qui est nécessaire et utile, et employer alors cette

connaissance pour la conception de nouveaux matériaux. Dans notre cas, la méthode PLS a été utilisée pour prédire la température de curie pour les alliages d'Heusler, dont certains font l'objet de recherches et d'autres ne sont pas encore étudiés.

On peut effectuer une étude comparative avec des alliages d'Heusler qu'on a pu mesurer leurs températures de curie expérimentalement et cela sans qu'on ne sache peu d'information sur leurs propriétés magnétiques, en se basant uniquement sur notre approche.

Ces prédictions sont données dans le tableau 4 qui comporte les valeurs de la température de curie prédite et celles retrouvées par des différentes études.

Composé	NVE	r_z (A°)	T_c exp(k)	T_c prédite(k)
Co ₂ TiGa	25	1,3	130 [3]	140
Co ₂ TiSn	26	1,45	341 [3]	343
Co ₂ FeSi	30	1,1	1100 [4]	1134
Co ₂ FeGa	29	1,3	1093 [1]	937

Tableau III-12 : valeurs prédites des températures de curie pour quelques alliages à base du cobalt.

Le tableau III-12 révèle des températures expérimentales et prédites presque égales, ce qui montre une fiabilité de notre modèle pour des alliages d'Heusler ferromagnétique.

Composé	NVE	r_z (A°)	T_c exp(k [8])	T_c prédite(k)
Mn ₂ TiBi	23	1,6	-	77
Mn ₂ TiIn	21	1,55	630	608
Mn ₂ TiP	23	1	-	193

Tableau III-13 : valeurs prédites des températures de curie pour quelques alliages à base du manganèse.

Le tableau III-13 montre une température de curie des alliages d'Heusler antiferromagnétique, malheureusement on ne trouve pas assez d'études concernant cette classe des alliages Heusler et particulièrement leurs températures de curie, mais comme on le voit sur le tableau que la valeur de température du Mn₂TiIn trouvée et prédite sont en un

parfait accord, quand au Mn_2TiBi lorsqu'on a injecté sa valeur dans notre base de données ça à donner une bonne tendance sur notre ACP (voir le graphe ci-dessous).

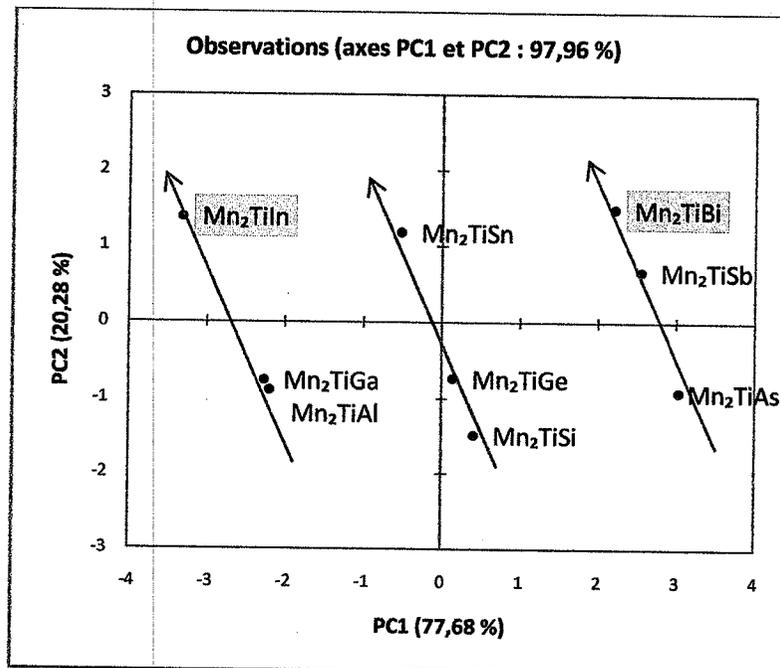


Fig. III-19 : l'aspect de notre ACP après injection des valeurs des températures prédites.

La méthode des moindres carrés partiels PLS (Partial Least Squares), qui est une méthode prédictive de régression linéaire à plusieurs variables. Nous a permis d'obtenir une loi linéaire reliant des variables réponses aux variables explicatives, dans notre cas il s'agit de trouver une loi de dépendance linéaire entre la température de curie et les propriétés structurales et magnétiques comme le moment magnétique totale de l'alliage (M_{tot}), le nombre d'électrons de valence (NVE) et le rayon atomique de l'élément Z (r_z).

Enfin, les résultats obtenus sont très intéressants et encourageants et peuvent ouvrir les portes dans le futur pour d'autres étudiants pour s'élargir et s'approfondir dans ce domaine de conception de nouveaux matériaux en utilisant les approches du datamining.

Références

- [1] International conference of magnetism. (ICM 2009).
- [2] B.S.D.ch.S Varaprasad et al. *J. Acta Materialia* 57 (2009) 2702-2709.
- [3] S.D. de Souza and R.N. SAXENA. Magnetic HF in Heusler alloys, *Hyperfine Interactions* 34 (1987) 431-434.
- [4] Jan Thoene, Stanislav Chadov, Gerhard Fecher, Claudia Felser and Jurgen Kubler. *J. Phys. D: Appl. Phys.* 42 (2009) 084013
- [5] Yasemin Kurtulus and Richard Dronskowski. (*Physical Review*) B 71, 014425 (2005).
- [6] S.Fujii, M. okada, S. Ishida and S. Asano, *J. Phys. Soc. JPN*, vol. 77, No,7(2008) 74702.
- [7] H. Zenasni; North African Conference On computational physics and chemistry 2010. Sidi Bel Abbas, Algeria, December, 12-14, 2010.
- [8] M Meinert, J M Schmalhorst, and G Department of Physics, Bielefeld University, 33501 Bielefeld, Germany (2010).
- [9] Iosif Galanakis and Peter H. Dederichs, *Lect. Notes Phys.* 676, 1–39 (2005).

**CONCLUSION GENERALE ET
PERSPECTIVES**

Conclusion Générale

Ce mémoire s'inscrit dans le cadre de l'**étude des alliages Heusler** en utilisant les techniques du Datamining (extraire des connaissances à partir des données).

Pour Conclure, nous présentons les principales motivations de notre travail. Actuellement, avec les progrès rapides dans la technologie et à tous les niveaux, la conception et la prédiction de nouveaux matériaux assistés par ordinateur sont devenus de plus en plus nécessaires.

L'analyse en composantes principales (PCA) et des moindres carrés partiels (PLS) qui sont des techniques informatiques d'analyse des données, ont été appliqués aux températures de curie pour différents alliages Heusler afin de découvrir des nouvelles règles ou lois.

L'utilité de cette approche d'informatique de matériaux ne se limite pas à l'interprétation des observations expérimentales qui existaient, mais elle pourrait être utilisée pour la conception de nouveaux alliages qui n'ont pas été étudiés expérimentalement.

Ce processus de conception de nouveaux matériaux se divise en trois étapes majeures :

- La première étape est la récolte de données afin de construire une base de données.
- La deuxième étape c'est l'analyse de ces données par des techniques du Datamining.
- La troisième et la dernière étape est la prédiction des matériaux avec des propriétés désirés pour concevoir de nouveaux alliages.

Dans un premier temps, nous avons construit une base de données qui tient compte des différentes propriétés structurales et magnétiques des alliages d'Heusler ferromagnétiques et antiferromagnétiques. Ces données tirées à partir des multiples calculs sont considérées comme données d'entrée pour notre calcul.

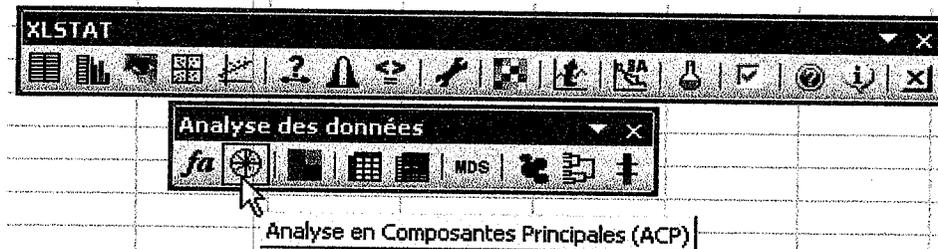
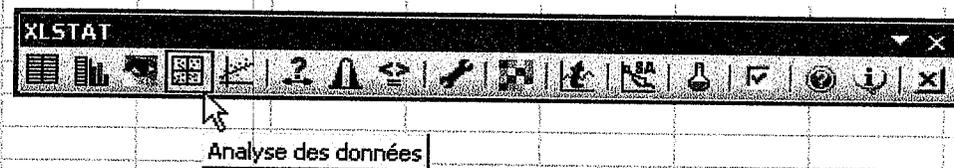
Dans un second temps, nous avons envisagé l'alternative de réduire ce nombre de données et les analyser en utilisant des modèles informatiques et des techniques d'exploitation des données, l'analyse en composante principale (ACP), et la régression aux moindres carrés partiels (PLS), ont été appliquées pour prédire des nouvelles propriétés.

Le calcul avec ces nouvelles approches produit des résultats comparables aux mesures expérimentales, et démontre comment l'informatique peut être employée pour utiliser l'information afin de déterminer ce qui est nécessaire et utile, et employer alors cette connaissance pour la conception de nouveaux matériaux. Les résultats obtenus sont très encourageants à approfondir l'étude dans ce domaine du Datamining.

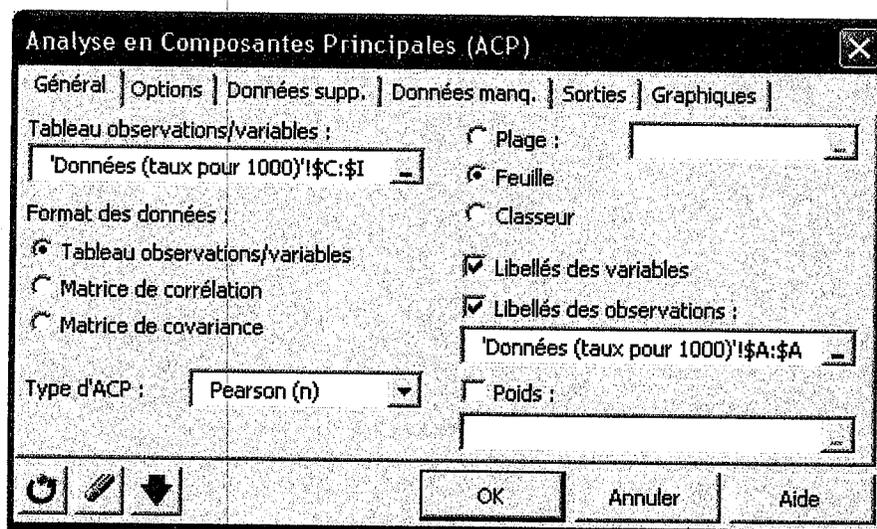
Ce travail en perspective devrait servir à l'intégration de l'informatique de matériaux dans la conception de nouveaux matériaux et augmenter aussi nos chances de découvrir de nouveaux matériaux aux propriétés exotiques, tel que les alliages quaternaires avec des températures de curie et des moments magnétique bien définis.

-Elargir notre étude pour d'autres classes d'alliages Heusler comme les alliages semi-Heusler.

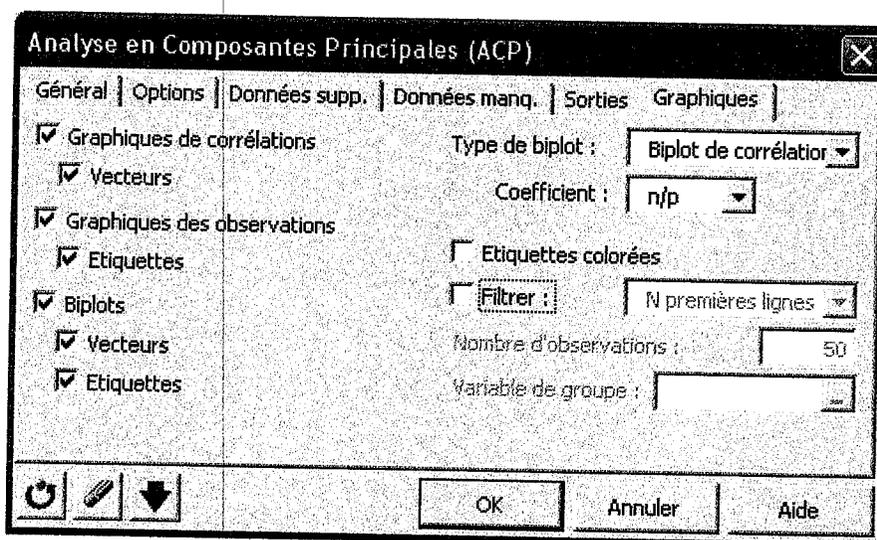
ANNEXE



- Une fois le bouton cliqué, la boîte de dialogue correspondant à l'Analyse en composantes principales apparaît. Nous pouvons alors sélectionner les données sur la feuille Excel.



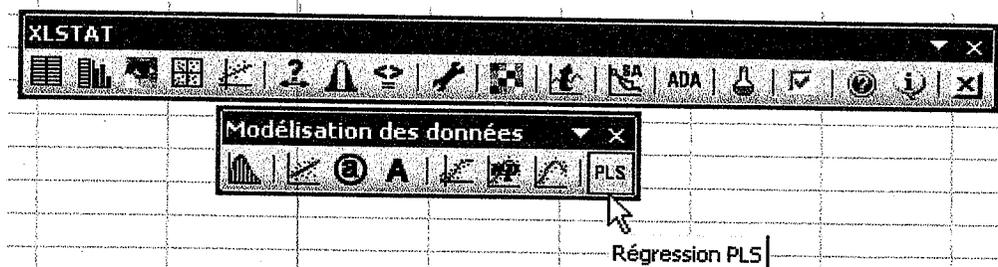
- Dans l'onglet "Graphiques", les options d' "Etiquettes" sont toutes activées afin que les libellés des variables et des observations soient bien affichés. L'option de filtrage des observations à afficher est aussi désactivée afin d'afficher toutes les observations. Lorsqu'il y a beaucoup d'observations, il est conseillé de ne pas afficher les étiquettes pour accélérer la génération du graphique, et de ne pas afficher toutes les observations afin de rendre le graphique plus lisible.



- Le premier résultat intéressant est la matrice des corrélations. Les tableaux des valeurs et vecteurs propres sont ensuite affichés. Tous les graphiques et tableaux sont mentionnés dans la partie résultat.

II.2. Faire une régression PLS avec XLSTAT

Nous lançons XLSTAT en cliquant sur le bouton  de la barre d'outils Excel, puis nous sélectionnons la commande Régression PLS de la barre d'outils "Modélisation des données".



- Une fois que nous avons cliqué sur le bouton, la boîte de dialogue apparaît. Sélectionnons au niveau des "Variables dépendantes" (les "Y" du modèle), les propriétés des matériaux. Ce sont en effet les données que l'on veut expliquer à travers des variables explicatives quantitatives (les "X" du modèle) que sont les nitrures et carbures.

Régression PLS

Général | Options | Validation | Prédiction | Données manq. | Sorties | Graphiques

Y / Variables dépendantes :

Quantitatives :

Qualitatives :

X / Variables explicatives :

Quantitatives :

Qualitatives :

Méthode : PLS PCR OLS

Plage :

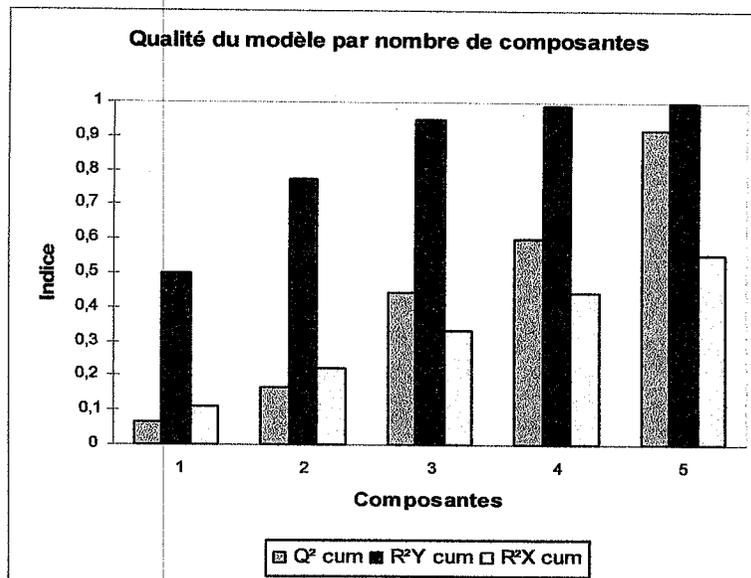
Feuille Classeur

Libellés des variables Libellés des observations :

Poids des observations :

OK Annuler Aide

- Le premier tableau et le graphique correspondant permettent de visualiser la qualité de la régression PLS en fonction du nombre de composantes retenues. Les variables explicatives sont en noir, et les variables dépendantes en bleu.



Indice	Comp1	Comp2	Comp3	Comp4	Comp5
Q ² cum	0,067	0,166	0,443	0,599	0,920
R ² Y cum	0,501	0,773	0,951	0,988	1,000
R ² X cum	0,111	0,222	0,333	0,444	0,556

- L'indice Q^2 cumulé est une mesure globale de la qualité de l'ajustement et de la qualité prédictive. XLSTAT-PLS a automatiquement retenu 5 composantes correspondant aux 5 propriétés du modèle. Les R^2Y cum et R^2X cum qui correspondent aux corrélations entre les composantes et les variables de départ sont proches de un dès la quatrième composante, ce qui indique que les composantes sont à la fois bien représentatives des X et des Y.

- Après quelques statistiques de base sur les différentes variables sélectionnées et la matrice de corrélations correspondante, les résultats propres à la régression PLS sont affichés. Tous les graphiques de corrélations et les tableaux correspondant sont mentionnés dans la partie résultat.