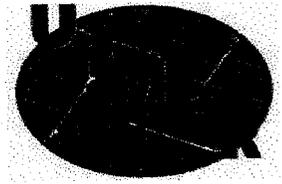
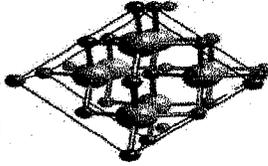


REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE MINISTERE
DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE

L
E
P
M



UNIVERSITE ABOU BEKR BELKAID-TLEMEN
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

Présenté par
Mme : MAMMAD Anissa

Thème

**Etude par datamining des propriétés structurales et élastiques
des intermétalliques de type AB_2**

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

Pr. A.Houari

Président

Université de Tlemcen

Dr. M. Benachour

Examineur

Université de Tlemcen

Pr.G. Merad

Encadreur

Université de Tlemcen

Mme.Sari.Aouatef

Co-Encadreur

Université de Tlemcen

Année Universitaire : 2011/2012

Remerciements

Je remercie ALLAH le Tout-puissant de m'avoir donné le courage, la volonté et la patience de mener à terme ce présent travail.

Ce travail a été réalisé dans le laboratoire de recherche d'Etudes et Prédiction des Matériaux (LEPM) affilié à l'Unité de Recherche des Matériaux et des Energies Renouvelables (URMER) de l'Université de Tlemcen.

*Je voudrais d'abord d'exprimer mes sincères remerciements à mon encadreur Monsieur **Ghouti Merad**, Professeur à l'université de Tlemcen Je lui présente toute ma gratitude pour l'attention et sa direction de ce travail .Je voudrais aussi exprimer mes profonde gratitude et mes sincères remerciements à Mme **Sari Aouatef** pour son co-encadrement, sa contribution à la réalisation et la finalisation de ce mémoire .J'y suis reconnaissante.*

*J'exprime mes sincères remerciements à monsieur **A. Houari** , Professeur à l'Université de Tlemcen, d'avoir accepté de présider le Jury de cette thèse.*

*Je remercie sincèrement monsieur **M. BENACHOUR** , Maitre de conférences à l'université de Tlemcen d'avoir accepté d'examiner ce travail et de faire partie de ce jury.*

Enfin, j'adresse mes vifs remerciements à toute la promotion matière condensée master 2, ainsi que tous mes amies qui m'ont encouragée et soutenue.

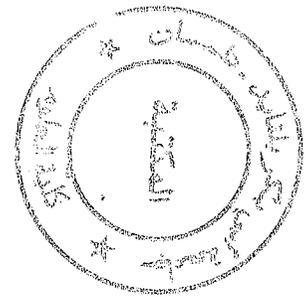
A mes parents,

A mon mari,

A mes frères,

A toute ma famille,

A mes amies.



SOMMAIRE

Introduction général	01
Références	03

Chapitre I : les composés intermétalliques.

I.1 Généralités sur les composés intermétalliques	04
I.1.1 Les composés de type AB	04
I.1.2 Les composés de types AB ₅	05
I.1.3 Les composés de type AB ₃	05
I.1.4 Composés de type ABC	06
I.1.5 Composés de type A ₂ B.....	06
I.1.6 Les composés de types AB ₂	07
I.1.6.1 Définition	07
I.1.6.2 Les structures cristallographiques des phases de laves de type AB ₂	08
A. Structure type MgCu ₂ ou C15	08
B. Structure type MgZn ₂ ou C14	09
C. Structure type MgNi ₂ ou C36	10
I.1.6.3 Propriétés et domaine d'application des phases de laves.....	11
Références	14

Chapitre II : Datamining et L'Analyse en composantes principales.

II.1 Datamining	15
II.1.1 Introduction	15
II.1.2 Principe de fonctionnement	15
II.1.3 Pour quelles applications ?	16
II.1.4 Les techniques mises en œuvre.....	16
II.1.5 Comment utiliser le datamining ?.....	16

1. Définition du problème	17
2. Collecte des données.....	17
3. Construire le modèle d'analyse.....	17
4. Etude des résultats.....	18
5. Formalisation et diffusion.....	18
II.1.6 Quelques méthodes statistiques	18
II.1.7 Difficultés du datamining	18
II.2 L'Analyse en composantes principales.....	20
II.2.1 Introduction.....	20
II.2.2 Définition	20
II.2.3 Principes de l'ACP.....	22
II.2.4 Formalisation de l'ACP.....	22
II.2.5 Conclusion.....	24
Références.....	25
<i>Chapitre III : Résultats et discussions.</i>	
III.1 Introduction.....	26
III.2 La base de données	26
III.3 Le schéma de l'ACP	28
III.4 Résultats	28
III.4 Conclusion	35
Conclusion générale.....	36
Annexe.	
I. Introduction.....	37
II. Utilisation du logiciel.....	37
III. Faire une Analyse en Composantes Principales (ACP) avec XLSTAT.....	38

Introduction

Introduction générale.

Une grande attention est apportée au développement de modèles prédictifs, ainsi des études et des recherches sont munies sur différents composés intermétalliques afin de mieux les exploiter dans divers domaines. Le thème de recherche développé dans le cadre de ce mémoire présente à la fois un caractère fondamental et un aspect statistique.

Les composés intermétalliques sont, au sens strict, des composés ordonnés formés par deux éléments métalliques ou plus avec une stœchiométrie précise. Il existe des milliers de composés intermétalliques. Parmi ces composés intermétalliques on trouve ceux de type AB_2 connus aussi sous l'appellation : *phase de Laves*. Grace à leurs applications dans différents domaines, ces matériaux ont attiré une très grande attention de la part des expérimentateurs et des théoriciens. Les phases de Laves sont classées en trois structures cristallines : structure hexagonale de type C14 ($MgZn_2$), la structure cubique de type C15 ($MgCu_2$), et structure C36 ($MgNi_2$) de type hexagonale.

Les phases de Laves forment le plus grand groupe de phases intermétalliques avec actuellement plus de 1400 représentants les plus connus [1]. Ils appartiennent tous à la famille des phases de Frank-Kasper [2,3]. Les phases de Laves ont été étudiées intensément afin de comprendre les aspects fondamentaux de la stabilité des phases.

D'autre part ces matériaux sont d'un intérêt particulier dans la recherche en métallurgie moderne en raison de leurs singulières propriétés physiques et chimiques, par conséquent, de nombreuses applications hypothétiques ou primitives ont été développées. Cependant, peu de connaissance pratique a été élucidée de l'étude de phase de Laves à ce jour. Un trait caractéristique est que la conductivité électrique presque parfaite. Par ailleurs, ils ne sont pas plastiquement déformable à température ambiante [4].

Le datamining est désormais une approche courante. Cette approche, qui permet de découvrir des schémas ou des connexions entre des données auparavant inconnus. Il se présente donc comme un outil incontournable dans un service marketing, et au sein des processus décisionnels d'une entreprise. Aujourd'hui, Il rassemble un faisceau de techniques statistiques qu'il convient d'utiliser au gré des problématiques – descriptives ou décisionnelles. Il s'assortit le plus souvent d'une méthode de travail pour ordonner au mieux hypothèses, modélisations et actions.

Dans ce mémoire, nous allons utiliser le datamining dans le domaine des sciences des matériaux, et combiner les atouts que peut offrir cette approche statistique avec les données expérimentales ou théoriques des propriétés mécaniques de quelques intermétalliques. Nous avons choisi l'une des techniques de datamining qui est l'analyse en composantes principale ACP.

Pour mener à bien notre étude le présent mémoire est structuré en trois chapitres :

- Le premier chapitre est dédié à une brève revue bibliographique dans laquelle sont exposés tour à tour les composés intermétalliques, leurs différents types, leurs structures, leurs domaines d'applications, et à la fin nous étudions en détail les phases de laves et leurs propriétés.
- Le deuxième chapitre présente une idée générale concernant le datamining et l'analyse en composantes principale ACP.
- Dans le troisième chapitre, nous exposons les résultats obtenus et nous développons les explications qui en dégagent.

Nous finissons ce mémoire par une conclusion générale.

Références

- [1] F. Stein, M. Palm and G. Sauthoff, *Intermetallics* 12 (2004) 713.
- [2] F.C. Frank and J.S. Kasper, *Acta Cryst.* 11 (1958) 184.
- [3] F.C. Frank and J.S. Kasper, *Acta Cryst.* 12 (1959) 483.
- [4] The Laves Phase Structures. nrl.navy.mil. Accessed on 2009-2-26.

Chapitre I

Les composés intermétalliques

1.1 Généralités sur les composés intermétalliques :

Les composés intermétalliques auxquels nous nous intéressons sont obtenus en combinant les deux types d'éléments A et B qui peuvent former des hydrures de stabilité intermédiaire à pression et à température ambiante. Les caractéristiques des composés intermétalliques AB_n vont dépendre de la nature des métaux mis en jeu ainsi que de la stœchiométrie du composé. L'indice n peut varier entre $n=1$ et $n=5$. Nous pouvons citer quelques classes des composés intermétalliques qui sont: les composés de type AB, AB_3 , AB_2 , AB_5 , A_2B , et les composés de type ABC.

Ces matériaux sont prometteurs et peuvent être utilisés dans différents domaines. Nous pouvons citer quelques exemples, une des classe de ces intermétalliques sont destinés aux applications à haute températures, grâce à leur propriétés mécaniques et à la résistance à la corrosion et à l'oxydation qui les caractérisent. Par ailleurs, des travaux entrepris dernièrement ont révélé des propriétés supraconductrices de ces composés. Nous ne pouvons pas passer sans citer ici un autre domaine d'application inédit, et d'actualité qu'est le stockage de l'hydrogène. D'autre part, des alliages à base de Magnésium ont été développés pour répondre à certains critères de légèreté dans le domaine de l'automobile, et aussi pour des applications de revêtements [1].

1.1.1 Les composés de type AB :

Les composés de formule AB cristallisent dans des structures de type CsCl cubique ou CrB orthorhombique dont le groupe d'espace est $Pm3m$ ou $Cmcm$ successivement. Ces intermétalliques sont appliqués dans le domaine du stockage de l'hydrogène car ils forment généralement des hydrures stables a température ambiante. Ils possèdent une capacité en réaction solide-gaz comprise entre 2 et 3 H/u.f. Le premier composé de type AB, étudié pour ses propriétés d'absorption de l'hydrogène a été ZrNi [2].

Par ailleurs, le composé TiNi est sollicité dans la composition des premiers accumulateurs électrochimiques utilisant des hydrures métalliques [3].

Le TiNi est moins sensible à la corrosion que TiFe et permet de faire diffuser l'hydrogène plus rapidement au sein de l'alliage. Ce composé a connu un regain d'intérêt au milieu des années 90 en raison de ces propriétés exceptionnelles d'alliage.

1.1.2 Les composés de types AB₅ :

Les composés de type AB₅ dits phases de Haucke cristallisent dans la structure hexagonale de type CaCu₅ dont le groupe d'espace est P6/mmm. Le composé LaNi₅ est l'exemple type des intermétalliques de type CaCu₅.

Les composés intermétalliques binaires et ternaires de type AB₅, cristallisent dans 34 différents types structuraux, mais la plupart des composés de stœchiométrie AB₅ optent principalement pour deux types de structures : une structure hexagonale de type CaCu₅ et une structure cubique de type AuBe₅.

Ces composés présentent un excellent comportement en milieu potassique (utilisation en électrochimie) et des capacités très proches de ce que laissent prévoir les caractéristiques mesurées en réaction solide-gaz. Le principal problème qu'a posé ce type de matériau a été sa médiocre cyclabilité (50 cycles à 80% de capacité pour LaNi₅).

1.1.3 Les composés de type AB₃ :

Ces composés cristallisent dans une structure rhomboédrique de type PuNi₃, structure qui contient un empilement de structure de type CaCu₅ et MgZn₂ dont les groupes d'espaces sont P6/mmm et P63/mmc successivement. Ces dernières années, de nombreuses études ont montré un réel intérêt pour le stockage de l'hydrogène dans les composés de type AB₃ [4].

Les composés AB₃ synthétisés jusqu'à présent ont montré des propriétés d'absorption/désorption d'hydrogène importantes grâce au fait qu'ils contiennent une partie des unités AB₂. On peut supposer que ce type de composés, peut devenir plus intéressant pour le stockage d'hydrogène que ceux de type AB₅. Les principales études actuelles consistent à substituer l'élément A par du Mg ou du Ca.

1.1.4 Composés de type ABC :

Les composés de type ABC cristallisent dans une structure orthorhombique de type TiNiSi dont le groupe d'espace est $Pnma$. L'hydruration de ces composés a une grande influence sur la structure cristalline et sur les propriétés magnétiques.

Les propriétés physiques des intermétalliques CeTX (T = élément de transition, X = élément np^1 ou np^2 avec np^1 : les éléments qui se trouvent dans le groupe IIIA et np^2 : les éléments qui se trouvent dans le groupe IVA) dépendent de la force d'hybridation entre les électrons 4f du cérium et ceux de la bande de conduction.

L'insertion d'hydrogène dans ces composés modifie à la fois le volume molaire et la densité d'états au niveau de Fermi, et donc la force d'hybridation [5].

1.1.5 Composés de type A_2B [6] :

Les composés avec la formule générale A_2B , cristallisant dans la structure de type PbFCl, ont été étudiés de manière exemplaire par des méthodes expérimentales et théoriques afin de démontrer la complexité de la liaison chimique dans ces phases intermétalliques.

La structure PbFCl est adoptée par environ 540 composés connus. Parmi les composés intermétalliques appartenant à la famille A_2B , le système magnésium nickel présente des propriétés d'absorption réversible de l'hydrogène intéressantes. Reily *et al.* ont étudié des propriétés d'hydrogénation de Mg_2Ni à la fin des années 60. Ce composé réagit avec l'hydrogène pour former Mg_2NiH_4 . Cependant, l'hydrure formé est très stable avec une pression de plateau d'environ 2 bar à 548 K et la cinétique de réaction est lente pour des températures inférieure à 440K.

La capacité électrochimique de Mg_2Ni est très faible à la température ambiante (8 mAh/g). Sous forme nanocristalline les propriétés électrochimiques de Mg_2Ni sont plus intéressantes et la capacité déchargée atteint 270mAh/g. Ces composés ont une courte durée de vie à cause de la corrosion du magnésium qui dans la potasse forme $Mg(OH)_2$. Le revêtement chimique ou électrochimique des particules Ni-P ou Ni permet d'améliorer la tenue en cyclage de ce composé.

Le travail que nous allons entreprendre, va porter sur une des classe des composés intermétalliques que nous avons cité au part avant, en l'occurrence les composés AB_2 communément appelés phases de Laves. Pour cette raison, nous allons étudier en détails la classe AB_2 .

1.1.6 Les composés de types AB_2 :

1.1.6.1 Définition :

L'appellation phases de Laves des composés intermétalliques AB_2 est issue du nom de Monsieur **Laves Fritz** qui fut le premier à les décrire. Ces phases sont caractérisées par un mélange en proportion spécifique de deux éléments métalliques, ce qui les distinguent des intermétalliques conventionnels. Elles représentent des structures compactes ordonnées avec comme formule chimique AB_2 pour les composés stœchiométriques et un rapport idéal des rayons $r_A/r_B = (3/2)^{1/2}$. Ces phases forment alors une structure périodique cristalline à différencier des structures des éléments constitutifs. Ces composés sont classés uniquement sur la base de la géométrie. Les phases de Laves cristallisent dans trois structures possibles définies comme suit dans la désignation Strukturbericht: cubique C15, hexagonale C14, et double- hexagonale C36.

Les deux dernières sont des formes uniques de l'arrangement hexagonal, mais partagent la même structure de base. En général, les atomes B sont ordonnés sous forme de tétraèdres autour des atomes A.

Pour la structure cubique, Les atomes B s'arrangent en tétraèdres, les tétraèdres étant joint entre eux par les sommets. Les atomes A occupent alors les espaces libres entre les tétraèdres formant entre eux un réseau de type diamant.

Les phases de Laves forment une très grande famille et de nombreuses substitutions par des éléments réduisant ou augmentant le volume de maille sont possibles.

1.1.6.2 Les structures cristallographiques des phases de laves de type AB_2 :

Comme nous l'avons mentionné, les composés AB_2 adoptent trois formes cristallographiques : la structure cubique C15 dite $MgCu_2$, la structure hexagonal C14 dite $MgZn_2$ et la structure double hexagonal C36 appelée $MgNi_2$.

A. Structure type $MgCu_2$ ou C15 :

Les atomes B s'arrangent en tétraèdres, les tétraèdres étant joint entre eux par les sommets. Les atomes A occupent alors les espaces libres entre les tétraèdres formant entre eux un réseau de type diamant. Le réseau est cubique à faces centrées et il y a 8 motifs AB_2 par maille. La représentation de cette structure est présentée en figure 1.

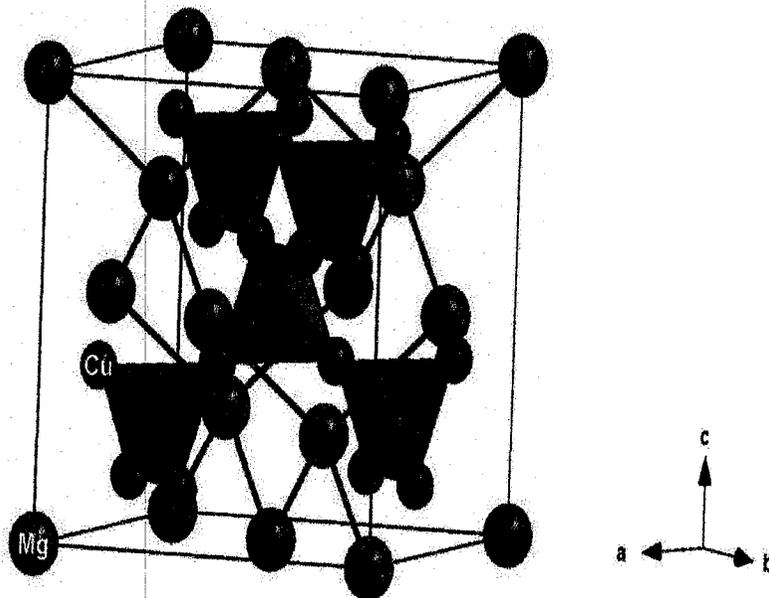


Fig.1 : Structure type $MgCu_2$ ou C15.

Les positions atomiques sont données dans le tableau 1 en description centrosymétrique (l'origine est placée au centre de symétrie c'est-à-dire décalée de $1/8 \ 1/8 \ 1/8$ par rapport à la figure 1).

Tableau 1 Données cristallographiques de la structure de type $MgCu_2$.

Structure type : $MgCu_2$ (C15)				
Groupe d'espace : Fd-3m				
Paramètres de maille : $a = 6,99 \text{ \AA}$				
Atome	Site	x	y	z
A	8a	1/8	1/8	1/8
B	16d	1/2	1/2	1/2

B. Structure type $MgZn_2$ ou C14:

Ce réseau est hexagonal. Les atomes B s'arrangent toujours en tétraèdres mais ces tétraèdres sont alternativement joints par les bases et par les sommets. Il y a cette fois 4 motifs par maille (figure2).

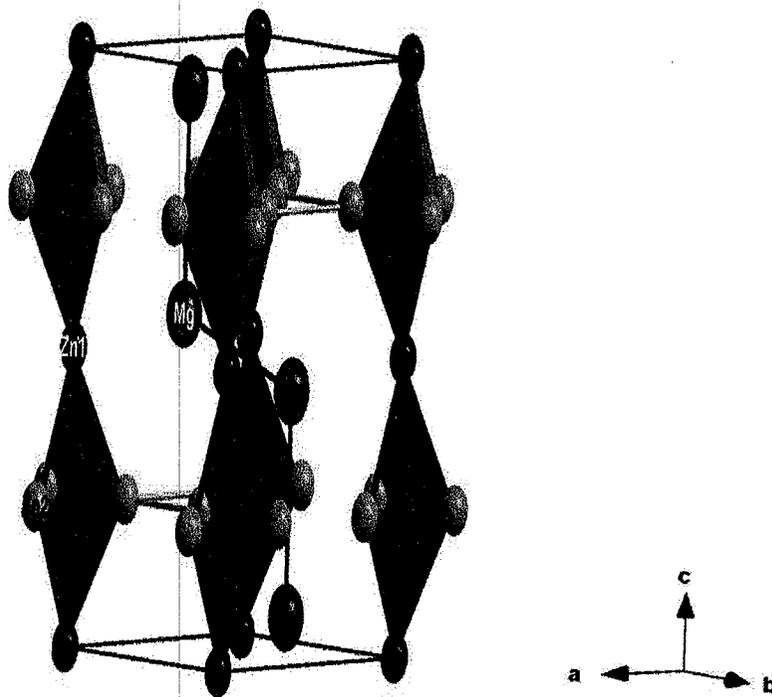


Fig.2 Structure type $MgZn_2$ ou C14.

Les positions atomiques idéales pour cette structure compacte sont données dans le tableau 2.

Tableau 2 Données cristallographiques de la structure de type $MgZn_2$.

Structure type : $MgZn_2$ (C14)				
Groupe d'espace : $P6_3/mmc$				
Paramètres de maille : $a = 5,2 \text{ \AA}$ $c = 8,6 \text{ \AA}$				
Atome	Site	x	y	z
A	4f	1/3	2/3	1/6
B1	2a	0	0	0
B2	6h	-1/6	-2/6	1/4

C. Structure type $MgNi_2$ ou C36 :

Cette structure est une combinaison de deux structures hexagonales. Dans le réseau la maille est doublée selon c par rapport à la structure C14.

On retrouve les chaînes des tétraèdres qui existent dans la structure C14 mais ces chaînes ne sont pas infinies et sont décalées les unes par rapport aux autres (figure3).

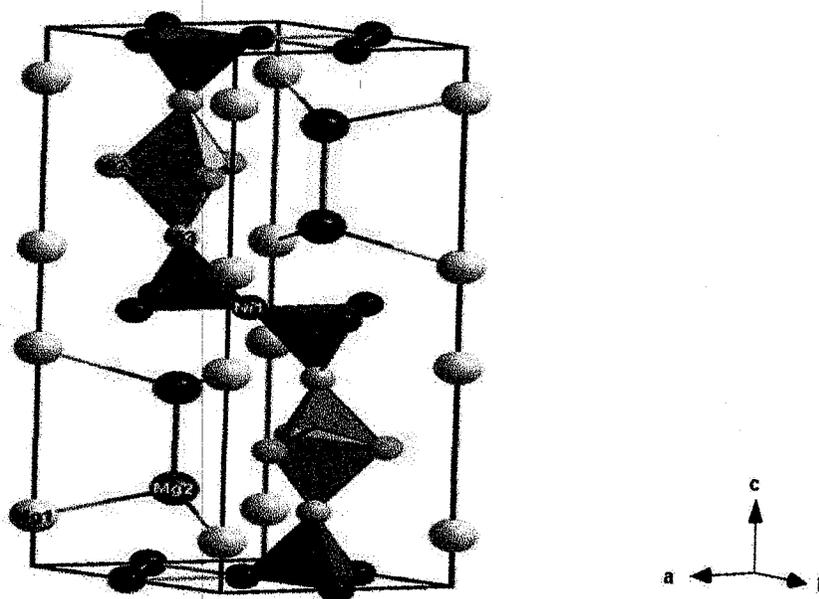


Fig.3 Structure type $MgNi_2$ ou C36

Tableau 3 : Données cristallographiques de la structure de type $MgNi_2$ ou C36.

Structure type : $MgNi_2$ (C36)				
Groupe d'espace : $P6_3/mmc$				
Ordre de grandeur des paramètres de maille : $a = 4,8 \text{ \AA}$ $c = 15,8 \text{ \AA}$				
Atome	Site	x	y	z
A1	4e	0	0	1/6
A2	4f	1/3	2/3	-1/6
B1	4f	1/3	2/3	1/8
B2	6g	1/2	0	0
B3	6h	1/6	2/6	1/4

1.1.6.3 Propriétés et domaine d'application des phases de laves :

Les phases de Laves sont un type important et attrayant de matériau à la fois scientifiquement et technologiquement. Ils sont sérieusement pris en considération pour de nombreuses applications pratiques.

Par exemple, $(Tb, Dy)Fe_2$ a été utilisé comme un transducteur magnéto en raison de sa grande magnéto-restriction [7]. Les composés $(Hf, Zr)V_2$ ont démontré propriétés supraconductrices avec une combinaison de haute température, haute densité critique de courant et une bonne résistance magnétique [8,9]. La supraconductivité est un phénomène caractérisé par l'absence de résistance électrique et l'expulsion du champ magnétique, notons l'existence de *l'effet Meissner* pour certains matériaux dits **supraconducteurs**. Puisque la supraconductivité permettrait notamment de transporter de l'électricité sans perte d'énergie, les applications potentielles sont donc stratégiques. Notons qu'un supraconducteur c'est un matériau conducteur de taille macroscopique ; caractérise par une température critique et un champ magnétique appliqué et un courant critique qui circule.

Par ailleurs, dans le domaine du stockage de l'hydrogène, on note que ces matériaux présentent des caractéristiques cinétiques d'absorption et de désorption relativement favorables mais également une capacité d'insertion très élevée grâce à l'existence de nombreux interstices au sein de la maille. De plus, l'existence d'une énergie de liaison hydrogène-métal faible permet d'observer l'absorption et la désorption de l'hydrogène à température voisine de l'ambiante et à pression atmosphérique.

La formation de cette liaison faible d'origine covalente ou ionique permet de rendre le phénomène de stockage réversible.

Les premières études de l'hydruration des composés AB_2 ont été effectuées par **Peblér et Gulbransen** [10]. ZrV_2 , $ZrCr_2$ et $ZrMn_2$ absorbent par voie solide-gaz près de 4 H/u.f. Cependant, la pression de dissociation de leurs hydrures est trop basse à température ambiante pour pouvoir avoir des applications pratiques. Les phases de Laves, tels que $Zr(Cr, Fe)_2$, ont été pris en considération pour les applications de stockage d'hydrogène grâce à la bonne cinétique de l'hydruration-déhydruration et absorbant l'hydrogène à haute capacités [11].

D'autre part, le $Mo(Co, Si)$ contribue à l'amélioration de la résistance à l'usure des matériaux « Tribaloy » [12]. Plus récemment, $HfCr_2$, $NbCr_2$ et $TiCr_2$ sont en cours d'élaboration pour les utilisations à haute température structurels en raison de leurs points de fusion élevés et leurs bonnes rétentions des propriétés mécaniques à des températures élevées [13,14].

En dépit de leurs propriétés utiles, la faible ductilité et les caractéristiques de ruptures fragiles à des températures ambiantes sont les principales limites pour des applications structurales de cette grande classe des alliages. La dureté élevée et la rupture fragile des phases de Laves sont dues à la configuration complexe atomique de chaque couche, et le manque de systèmes de glissement fonctionnant à des températures ambiantes.

Récemment, un mécanisme de cisaillement synchronisé (synchro-shear) [15] a été proposé pour faciliter la déformation plastique dans les phases de Laves.

La stabilité de phase et structures de défauts dans les phases de Laves, de $NbCr_2$, $NbFe_2$, $NbCo_2$ et $ZrCo_2$ ont été étudiés dans le but de la promotion

de la déformation de cisaillement synchronisé (synchro-shear) en augmentant la concentration des lacunes et la promotion du jumelage mécanique en contrôlant une stabilité de phase et réduisant l'énergie de défaut d'empilement. À l'heure actuelle, seuls des progrès limités ont été accomplis dans l'amélioration de la ténacité à la rupture de ces alliages en phase de Laves. Compte tenu de ce problème, des alliages à base d'un chiffon doux riche et renforcé en Cr sont des solutions. Les solides NbCr_2 , ZrCr_2 , ou des phases TaCr_2 ont été sélectionnés pour le développement des alliages [16,17- 18]. Parmi ces trois systèmes Cr- XCr_2 (X = Nb, Zr, et Ta), l'alliage Cr- TaCr_2 est plus résistant à l'oxydation de l'air et à la fissuration induite thermiquement.

Comme la plupart des phases de Laves sont fragiles à la température ambiante, les essais mécaniques sont fortement limités. Tous ces faits indiquent que plus de recherches doivent être entreprises sur les phases de Laves dans le but de mieux comprendre la stabilité de phase de base, la structure des défauts, mécanisme de déformation et le comportement à la rupture. En outre, les efforts de conception en alliage devrait être consacrée à surmonter la fragilité de ces phases.

L'objectif de cette étude est de déterminer la relation qui existe entre les différentes propriétés des matériaux de type AB_2 , afin d'être à même de choisir une composition optimale en fonction de l'application souhaitée. Dans ce but, nous allons utiliser une techniques statistiques, dite PCA que nous allons exposer dans le chapitre suivant.

Références

- [1] B.L. Mordike and T. Ebert. *Mater Sci Eng A*, **302** (2001), pp. 37–45. [SD-008]
- [2] .G.G.Libowitz ,H.F. Hayes ,T.R.P.Gibb. *Phys.Chem.*,62(1958)76.
- [3] M.A.Gutjahr,H.Buchner,K.D.Beccu, H. Saufferer, Proc. 8th Power Sources Symposium, Brighton, (1972) 79.
- [4] Cristina Iuliana (Pertrache) STAN, PHASE ET NOUVEAUX COMPOSES A BASE DE MAGNESIUM POUR LE STOCKAGE DE L'HYDROGENE ,24/10/2008.
- [5] Bernard CHEVALIER et Jean-Louis BOBET, Modification par hydruration des propriétés structurales et physiques des intermétalliques CeTX (T=Mn, Ni, Cu ; X=Al, Ga, In, Si, Ge, Sn) ,24/09/2004, N d'ordre : 2849
- [6] doxa.scd.univ-paris12.fr/theses/th0405792-Chapitre1.pdf. 1. L'accumulateur Nickel Métal- Hydrure (Ni-MH).
- [7] Moffett MB, Clark AE, Wun-Fogle M, Lindberg J, Teter JB, McLaughlin EA. *J Acoust Soc Am* 1991;89:1448.
- [8] Inoue K, Tachikawa K. *IEEE Trans Magnetics* 1979;15:635.
- [9] Olzi E, Maticotta FC, Setina P. *J Less-Common Met* 1988;139:123.
- [10] A. Pebler, E.A. Gulbransen, *Trans. Met. Soc. AIME*, 239 (1967) 1593.
- [11] Ivey D, Northwood D. *J Less-Common Met* 1986;115:23.
- [12] A Halstead , Rawlings RD. *J. Mater. Sci.* 1985;20:1248 & 1693
- [13] Anton DL, Shah DM. *Mater Res Soc Symp Proc* 1990;194:45.
- [14] Chen KC, Allen SM, Livingston JD. *J Mater Res* 1997;12:1472.
- [15] Hazzledine PM, Kumar KS, Miracle DB, Jackson AG. *MRS Symp Proc* 1992;288:591.
- [16] Takeyama M, Liu CT. *Mater Sci Engng* 1991;A132:61.
- [17] Liu CT, Tortorelli PF, Horton JA, Carmichael CA. *Mater Sci Engng* 1996;A214:23.
- [18] Brady MP, Zhu JH, Liu CT, Tortorelli PF, Walker LR, McKamey CG, Wright JL, Carmichael CA, Larson DJ, Miller MK, Porter WD. *Materials at High Temperatures* 1999;16(4):189±93.

Chapitre II

Datamining et L'Analyse en composantes principales

II.1 Datamining

II.1.1 Introduction :

Le Datamining que l'on peut traduire par "fouille de données" ou L'exploitation de données ou encore extraction de connaissances à partir des données (ECD) apparaît au milieu des années 1990 aux États-Unis comme une nouvelle discipline à l'interface de la statistique et des technologies de l'information : bases de données, intelligence artificielle, apprentissage automatique (« machine learning »).

Le Datamining est en fait un terme générique englobant toute une famille d'outils facilitant l'analyse des données contenues au sein d'une base décisionnelle de type Data Warehouse ou DataMart. Les techniques de datamining sont particulièrement efficaces pour extraire des informations significatives depuis de grandes quantités de données.

Le Datamining est un ensemble de techniques et de méthodes du domaine des statistiques, des mathématiques et de l'informatique permettant l'extraction du savoir, à partir d'un important volume de données brutes, de connaissances originales auparavant inconnues. Il s'agit de "fouilles" visant à découvrir "de l'information cachée" que les données renferment et que l'on découvre à la recherche d'associations, de tendances, de relations ou de régularités.

II.1.2 Principe de fonctionnement :

En peu de mots, le Datamining présente l'avantage de trouver des structures originales et des corrélations informelles entre les données. Il permet de mieux comprendre les liens entre des phénomènes en apparence distincts et d'anticiper des tendances encore peu discernables.

II.1.3 Pour quelles applications ?

A contrario des méthodes classiques d'analyses statistiques, le Datamining est particulièrement adapté au traitement de grands volumes de données. Avec l'augmentation de la capacité de stockage des supports informatiques, un maximum de renseignements sera capté, ordonné et rangé au sein du Data Warehouse « dépôt de données ». Avec le Datamining, ces bases de données sont exploitables.

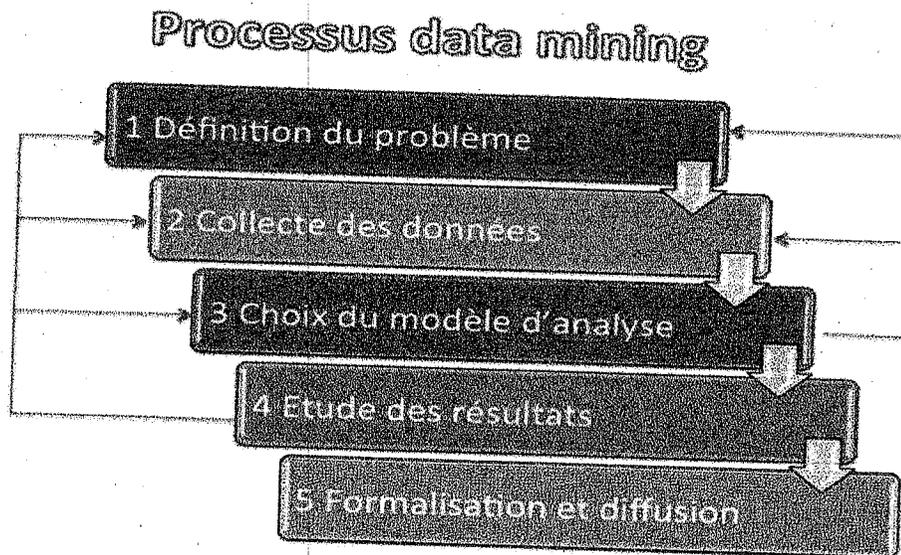
II.1.4 Les techniques mises en œuvre :

Les outils de datamining proposent différentes techniques à choisir en fonction de la nature des données et du type d'étude que l'on souhaite entreprendre, nous pourrions ici citer :

- Les méthodes utilisant les techniques de classification et de segmentation.
- Les méthodes utilisant des principes d'arbres de décision assez proches des techniques de classification.
- Les méthodes fondées sur des principes et des règles d'associations ou d'analogies.
- Les méthodes exploitant les capacités d'apprentissage des réseaux de neurones.
- les algorithmes génétiques pour les études d'évolution de populations.
- Algorithmes Naïve Bayes, séries chronologiques, régression linéaire...

II.1.5 Comment utiliser le datamining ?

Le principe : une démarche simplifiée et didactique qui se résume en 5 temps majeurs permet d'utiliser le datamining :



1. Définition du problème :

Quel est le but de l'analyse, que recherche-t-on ? Quels sont les objectifs ? Comment traduire le problème en une question pouvant servir de sujet d'enquête pour le datamining ? A ce sujet, se souvenir que l'on travaille à partir des données existantes, la question doit être ciblée selon les données disponibles.

2. Collecte des données :

C'est une phase absolument essentielle. On n'analyse que des données utilisables, c'est à dire "propres" et consolidées. On n'hésitera pas à extraire de l'analyse les données de qualité douteuse connues généralement comme étant *un bruit*. Bien souvent, les données méritent d'être retravaillées. S'assurer au final que la quantité de données soit suffisante pour éviter de fausser les résultats. Cette phase de collecte nécessite le plus grand soin.

3. Construire le modèle d'analyse :

Ne pas hésiter à valider vos choix d'analyse sur plusieurs jeux d'essais en variant les échantillons. Une première évaluation peut nous conduire à reprendre les points 1 ou 2.

4. Etude des résultats :

Il est temps d'exploiter les résultats. Pour affiner l'analyse on n'hésitera pas à reprendre les points 1, 2 ou 3 si les résultats s'avéraient insatisfaisants.

5. Formalisation et diffusion :

Les résultats sont formalisés pour être diffusés. Ils ne seront utiles qu'une fois devenus une connaissance partagée. C'est bien là l'aboutissement de la démarche. C'est aussi là que réside la difficulté d'interprétation et de généralisation...

II.1.6 Quelques méthodes statistiques :

Le champ théorique du datamining est riche, les techniques utilisées sont multiples, et les paramétrages possibles de ces techniques font apparaître une véritable explosion combinatoire. Contrairement à ce que certains éditeurs prétendent, il n'existe pas dans le cas général d'outil capable de guider le néophyte pour défricher la forêt des méthodologies, déjouer les pièges et conduire l'analyse en fonction du besoin et des données. Ultra-Fluide utilise des logiciels open source qui implémentent les techniques up-to-date telles que notamment :

- Etude des covariances.
- Régression multivariable.
- Clustering (maximum de vraisemblance, analyse discriminante, hiérarchique).
- Réseaux de neurones.
- Algorithme génétique.

II.1.7 Difficultés du datamining :

Les questions soumises au datamining sont souvent complexes, corrélativement d'ailleurs à la valeur ajoutée que l'on attend des réponses, elles nécessitent un œil expérimenté. Les outils nécessitent également une bonne dose de connaissances théoriques et un peu de courage pour s'approprier leurs documentations.

Le contraire ne pourrait que relever du miracle : comment imaginer détecter sans effort des informations cachées, généralement noyées dans du bruit, qui défient l'instinct, et dont l'exactitude ne se révèle que globalement ?

Le datamining offre cependant une apparence trompeuse. En effet, quelle que soit la méthode, quiconque manipule des données obtiendra toujours un résultat. C'est encourageant, et cela semble facile. Mais au delà d'une large variété de réponses raisonnables existent un nombre encore plus important de réponses fantaisistes (inadéquation entre les données, les objectifs, la méthode et les divers paramètres). Ces mirages sont monnaie courante, ils consistent simplement à faire apparaître des choses qui n'existent pas au lieu de choses qui sont cachées. Il est donc nécessaire d'élaguer ces mauvaises réponses pour choisir en bout de course, parmi les résultats raisonnables, ceux qui correspondent le mieux aux objectifs. Si l'on ne retenait qu'une seule recommandation, ce serait celle de ne jamais s'arrêter au premier résultat. Il faut calculer, recalculer, comparer, interpréter, affiner...

Il existe un autre conseil de bon sens pour éviter les mirages : se fonder sur des résultats robustes. Un résultat est valable, donc robuste, s'il est applicable à l'ensemble des ensembles de données autorisés par la classe de problèmes que l'on souhaite résoudre. Autrement dit si l'on travaille sur 100 données et que l'on obtient avec une méthode un résultat, alors on vérifie que l'application de la même méthode sur 100 autres données collectées selon le même processus produit un autre résultat raisonnablement proche du précédent.

En fait on procède souvent d'une manière légèrement différente : on dispose de 100 données du début à la fin, mais on torture longtemps 60 données seulement. Lorsque que l'on s'approche d'une solution plaisante on l'applique sans réfléchir aux 40 autres données, ou à l'ensemble des 100. On peut alors obtenir une première idée de la robustesse de la solution.

II.2 L'Analyse en composantes principales

II.2.1 Introduction :

L'ACP prend sa source dans un article de Karl Pearson publié en 1901. Le père du Test du χ^2 y prolonge ses travaux dans le domaine de la régression et des corrélations entre plusieurs variables. Pearson utilise ces corrélations non plus pour expliquer une variable à partir des autres (comme en régression), mais pour décrire et résumer l'information contenue dans ces variables.

Encore connue sous le nom de transformée de Karhunen-Loève ou de transformée de Hotelling, l'ACP a été de nouveau développée et formalisée dans les années 1930 par Harold Hotelling. La puissance mathématique de l'économiste et statisticien américain le conduira aussi à développer l'analyse canonique, et la généralisation des analyses factorielles dont fait partie l'ACP.

Les champs d'application sont aujourd'hui multiples, allant de la biologie à la recherche économique et sociale, et plus récemment le traitement d'images. L'ACP est majoritairement utilisée pour:

- décrire et visualiser des données ;
- les décorréler ; la nouvelle base est constituée d'axes qui ne sont pas corrélés entre eux ;
- les débruiter, en considérant que les axes que l'on décide d'*oublier* sont des axes *bruités*.

II.2.2 Définition :

L'Analyse en Composantes Principales (ACP) est une méthode de la famille de l'analyse des données et plus généralement de la statistique multivariée, qui consiste à transformer des variables liées entre elles (dites "corrélées" en statistique) en nouvelles variables décorrélées les unes des autres. Ces nouvelles variables sont nommées "composantes principales", ou axes. Elle permet au praticien de réduire

l'information en un nombre de composantes plus limité que le nombre initial de variables. Chaque composante principale est une combinaison linéaire des variables initiales. La mesure de la quantité d'informations représente sa variance. Les variances associées à chaque composante principale sont classées par ordre décroissant. La composante principale la plus informative est donc la première, et la moins informative est la dernière [1].

La méthode est basée sur l'hypothèse qu'il existe de fortes corrélations entre les données étudiées. On passe d'un certain nombre de variables potentiellement corrélées à un plus petit nombre de variables non corrélées, les "Composantes Principales". La 1ère "Composante Principale" absorbe le plus de variance possible, la 2ème "Composante Principale" absorbe le plus de variance possible parmi la variance restante, etc.. L'ACP permet d'analyser des données multivariées et de les visualiser sous forme de nuages de points dans des espaces géométriques [2].

L'analyse en composantes principales est utilisée pour réduire la dimension (le nombre de variables) d'un problème. Cette diminution du nombre de variables doit s'effectuer en perdant un minimum d'informations. Le but de l'Analyse en Composantes Principales (ACP) est de condenser les données originelles en de nouveaux groupements, appelées nouvelles composantes [1, 3].

Pour effectuer une Analyse en Composantes Principales, nous devons passer par les étapes suivantes:

- ✓ Normalisation des données pour être indépendants des paramètres P.
- ✓ Calcul d'une matrice de similarité C (bien souvent la corrélation).
- ✓ Recherche des éléments propres de C, qui donnent les axes principaux.
- ✓ Représentation des individus dans le nouvel espace (en ne considérant que les valeurs propres) [4].

L'A.C.P. permet d'explorer les liaisons entre variables et les ressemblances entre individus.

II.2.3 Principes de l'ACP [5] :

L'idée de l'ACP est de déterminer un nouveau repère de \mathbb{R}^p associé de manière naturelle à la structure du nuage considéré, de façon à pouvoir l'examiner plus commodément.

Pour s'affranchir des effets d'échelle dus à l'hétérogénéité éventuelle des variables, ces dernières sont en général normalisées, c'est à dire que chaque colonne est divisée par son écart-type; toutes sont dès lors exprimées dans la même échelle standard.

D'autre part, l'origine est placée au centre de gravité du nuage.

C'est le nuage ainsi transformé qui est en fait considéré; l'utilisateur n'a cependant pas à se préoccuper de ces transformations préalables, sauf demande contraire elles sont exécutées automatiquement par les logiciels d'ACP.

II.2.4 Formalisation de l'ACP [5] :

On note X la matrice à $n \times p$ données (portant les observations en ligne, éléments de \mathbb{R}^p , et les variables, quantitatives ou qualitatives en colonnes, éléments de \mathbb{R}^n), on suppose les colonnes de X préalablement centrées et réduites si nécessaire.

Soit u un vecteur (en colonne) unitaire de \mathbb{R}^p , le vecteur $X \cdot u$ de \mathbb{R}^n a pour composantes les produits scalaires des observations avec u , c'est-à-dire encore, les distances à l'origine des projections des observations selon la direction de u , tandis que l'inertie totale du nuage dans cette direction est donnée par le produit matriciel : $u' \times X' \times X \times u$.

La matrice symétrique $X' \times X$ est la matrice d'inertie du nuage, tandis que le produit $u' \times X' \times X \times u$, qui donne l'inertie dans cette direction, est l'application de la forme bilinéaire symétrique de matrice $X' \times X$ au vecteur unitaire u . On remarque que $X' \times X$ est simplement, au facteur $1/n$ près de la matrice des corrélations entre les variables-colonnes initiales (ou des covariances si on effectue une ACP non normée).

La recherche des directions principales u_k , c'est à dire des directions successives d'inertie maximale du nuage, se traduit donc par le problème de maximisation sous contrainte :

$$\text{Max } (u_k' \times X' \times X \times u_k) \text{ avec } u_k' \times u_k = 1$$

les vecteurs u_k successifs devant en outre être orthogonaux.

L'algèbre linéaire enseigne que les vecteurs propres normés successifs u_k , associés à la suite décroissante des valeurs propres (positives) de $X' \times X$: λ_k , apportent la solution du problème. Notons que la valeur propre λ_k mesure l'inertie dans la k -ième direction principale u_k :

$$u_k' \times X' \times X \times u_k = \lambda_k \times u_k' \times u_k = \lambda_k$$

Les vecteurs $c_k = X \times u_k$ de \mathbb{R}^n sont les composantes principales successives du nuage, centrées, de variances non corrélées respectives λ_k/n (de covariances : $c_k' \times c_k/n = \lambda_k \times u_k' \times u_k/n$, nulles). Ce sont les « nouvelles variables », dont les composantes donnent les coordonnées des points du nuage sur les axes factoriels.

Les diverses contributions, corrélations qui aident à l'interprétation, sont enfin facile à écrire, en fonction des λ_k , u_i et c_j . Ainsi, par exemple, la contribution de l'observation i à l'axe k est : $c_k(i)^2/\lambda_k$, où $c_k(i)$ désigne la i -ème composante de c_k ...

La présentation de l'ACP par les variables conduit par une autre voie au même problème mathématique : on cherche de nouvelles variables combinaisons linéaires des anciennes variables, non corrélées entre elles, et les plus corrélées possible avec l'ensemble de ces variables initiales, plus exactement telles que la somme des carrés des corrélations avec les anciennes variables soit maximale.

L'algèbre linéaire dit encore que les vecteurs propres normés successifs : y_k , associés à la suite décroissante des valeurs propres (positives) de $X \times X'$: μ_k , apportent la solution du problème, la valeur propre μ_k mesurant la corrélation totale maximisée pour la variable y_k .

De plus, les valeurs propres de $X' \times X$ et de $X \times X'$ sont les mêmes : $\lambda_k = \mu_k$, tandis que les vecteurs propres se déduisent par application de la matrice X (ou X'), ainsi les nouvelles variables y_k sont simplement les composantes principales $c_k = X.u_k$ normalisées.

Une autre présentation encore équivalente est par la recherche de nouvelles variables combinaisons linéaires normalisées des anciennes, non corrélées et de dispersion ou de variance maximale.

II.2.5 Conclusion

L'ACP est une technique de statistique descriptive dont le principe est simple mais qui met en oeuvre des calculs numériques importants, pour cette raison elle n'a pu se développer qu'avec l'apparition des ordinateurs.

L'ACP est à conseiller pour un premier examen, une mise en forme ou une présentation synthétique de données abondantes croisant des individus avec des variables.

On n'omettra cependant pas d'examiner préalablement les données par les méthodes statistiques usuelles (moyenne, écart-type, graphiques, corrélation, etc..).

Un reproche fréquemment adressé à l'ACP et aux techniques connexes est qu'elles ne révéleraient que des évidences. Le propos est injuste, mais il est rassurant que souvent les premiers axes retrouvent et confirment ce qui était déjà connu.

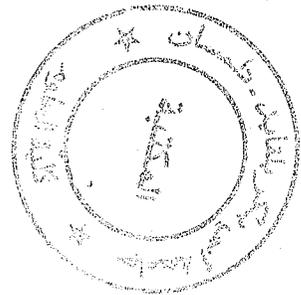
Comme avec les autres méthodes descriptives, il faut être très prudent pour inférer des modèles explicatifs ou causals à partir des configurations obtenues.

Références

- [1] Hotelling K (1933). Analysis of a Complex of Statistical Variables Into Principal Components, *Journal of Educational Psychology* pp. 417<441>.
- [2] J. Pages, B. Escofier : Analyses factorielles simples et multiples : Objectifs, méthodes et interprétation. DUNOD, 1990.
- [3] Pierre-Louis Gonzalez : Analyse en Composantes Principales.
- [4] Etude de données d'expression par Analyse en Composantes Principales Fonctionnelles BARRA VINCENT.
- [5] [http// www.cabannes.org/acp.pdf-France](http://www.cabannes.org/acp.pdf-France) . L'Analyse en composantes principales.

Chapitre III

Résultats et interprétations



III.1 Introduction :

Dans ce chapitre nous exposons les résultats liés à l'utilisation de l'Analyse en composantes principales (ACP), appliquées sur les propriétés structurales et mécaniques de quelques composés intermétalliques se présentant dans des phases de laves. L'ACP est un outil puissant pour analyser et décrire les relations entre les matériaux. Cette technique a été employée pour résoudre des problèmes en science des matériaux.

L'analyse en composantes principales permet de traiter un nombre très important de données et de dégager les aspects les plus intéressants de la structure de celles-ci.

Le but de l'ACP dans ce chapitre c'est de trouver des corrélations, des tendances, et de pouvoir tirer un savoir ou connaissances cachés initialement, et pourquoi pas prédire de nouveaux matériaux.

Les calculs ont été faits en utilisant le logiciel XLSTAT-Pro décrit dans l'annexe.

III.2 La base de données :

Le tableau [III-1] contient l'ensemble de données utilisées. Il est composé de 16 matériaux et 11 propriétés pour chaque composé. Les propriétés et leurs abréviations incluses dans l'analyse sont comme suit : les constantes élastiques (C_{11} , C_{12} , C_{44}), module de young (E), module de cisaillement (G), module compressibilité (B), le coefficient de Poisson ν , les paramètres de maille (a et c).

	C ₁₁ (G.Pascal)	C ₁₂ (G.Pascal)	C ₄₄ (G.Pascal)	E (G.Pascal)	G (G.Pascal)	V	B (G.Pascal)	a (Å)	C (Å)	c/a	C ₁₁ -C ₁₂ (G.Pascal)	Référence
TiCr ₂ C ₁₄	323	141	98	240	92	0.303	199	4.885	7.830	1.6127	182	[a]
ZrCr ₂ C ₁₄	272	132	20	91	32	0.415	176	5.094	8.103	1.5906	140	[a]
HfCr ₂ C ₁₄	302	139	51	161	59	0.362	191	5.046	8.095	1.6042	163	[a]
CaMg ₂ C ₁₄	62.043	17.074	17.951	55.738	18.990	0.201	31.063	6.232	10.093	1.6195	44.969	[b]
CaAl ₂ C ₁₄	112.747	25.705	36.515	106.507	38.153	0.208	55.808	5.700	9.280	1.6200	87.042	[b]
MgAl ₂ C ₁₄	92.105	47.403	17.173	51.405	18.342	0.295	53.294	5.458	8.758	1.6046	44.702	[b]
TiCr ₂ C ₁₅	292	158	80	200	75	0.336	208	6.857	6.857	1	134	[a]
ZrCr ₂ C ₁₅	252	146	54	146	54	0.366	182	7.145	7.145	1	106	[a]
HfCr ₂ C ₁₅	282	152	66	176	65	0.349	199	7.079	7.079	1	130	[a]
CaMg ₂ C ₁₅	59.521	15.454	22.034	52.388	22.034	0.217	30.815	7.079	7.079	1	44.067	[b]
CaAl ₂ C ₁₅	105.676	31.071	37.303	102.686	37.303	0.203	56.651	8.045	8.045	1	74.605	[b]
ZrMn ₂ C ₁₅	219	186	37	77	27	0.43	199	6.997	6.997	1	33	[c]
MgAl ₂ C ₁₅	95.772	47.709	24.261	66.734	24.806	0.321	62.090	7.667	7.667	1	48.063	[b]
CaMg ₂ C ₃₆	59.666	16.445	20.315	65.705	20.632	0.233	32.144	6.200	20.184	3.2554	43.221	[b]
CaAl ₂ C ₃₆	110.501	27.913	41.443	97.656	41.406	0.187	56.623	5.710	18.426	3.2269	82.588	[b]
MgAl ₂ C ₃₆	94.101	42.206	20.571	68.818	25.925	0.313	61.187	5.460	17.546	3.2135	51.895	[b]

Tableau (III-1) : Base de données utilisées dans l'analyse en composantes principales.

[a] Ab initio study of ground state properties of the Laves phase compounds TiC₂, ZrCr₂, and HfCr₂. Xing-Qiu Chen¹, W. Wolf², R. Podloucky¹, and P. Rogl¹. Institut für Physikalische Chemie, Universität Wien, Liechtensteinstrasse 22A, A-1090, Vienna, Austria. ²Materials Design s.a.r.l., 44, av. F.-A. Bartholdi, 72000 Le Mans, France. November 19, 2004.

[b] First-principles investigation of the binary AB₂ type Laves phase in Mg-Al-Ca alloy: Electronic structure and elastic properties. Wei-Yang Yu^a, Na Wang^a, Xiao-Bing Xiao^b, Bi-Yu Tang^{a, b, *}, Li-Ming Peng^c, Wen-Jiang Ding^{c, **}.

[c] Ab initio study of ground-state properties of the Laves-phase compound ZrMn₂. Xing-Qiu Chen, ¹W. Wolf, ²R. Podloucky, ¹P. Rogl, and M. Marsman. Received 15 April 2005; published 25 August 2005. PHYSICAL REVIEW B 72, 054440 (2005).

III.3 Le schéma de l'ACP:

L'Analyse en composantes principales (ACP) des 15 matériaux donne deux types de graphes : le premier est celui des variables appelé aussi (*loading plot*), dans notre cas c'est le diagramme des propriétés considérés. Le deuxième graphe est celui des observations, connu sous le nom (*scores plot*), c'est le diagramme des matériaux qui constituent la base de données.

Les axes principaux F1 et F2 nommés PC₁ et PC₂ dans l'approche de l'ACP correspondent aux vecteurs propres relatifs aux deux plus grandes valeurs propres de la matrice des covariances des données. Dans ce travail, PC₁ capture 58,48% du désaccord dans l'ensemble de données, et PC₂ capture 16,57% du désaccord. Ainsi, l'intégralité de l'information initialement comprise est réduite à 75.04%. Ce qui reste comme information est considéré moins utile et est communément appelé : *bruit*.

III.4 Les résultats:

Le graphe des propriétés (*loading plot*) est donné par la figure (III-1) et celui des observations (*scores plot*) de cette analyse est donné par la figure (III-2).

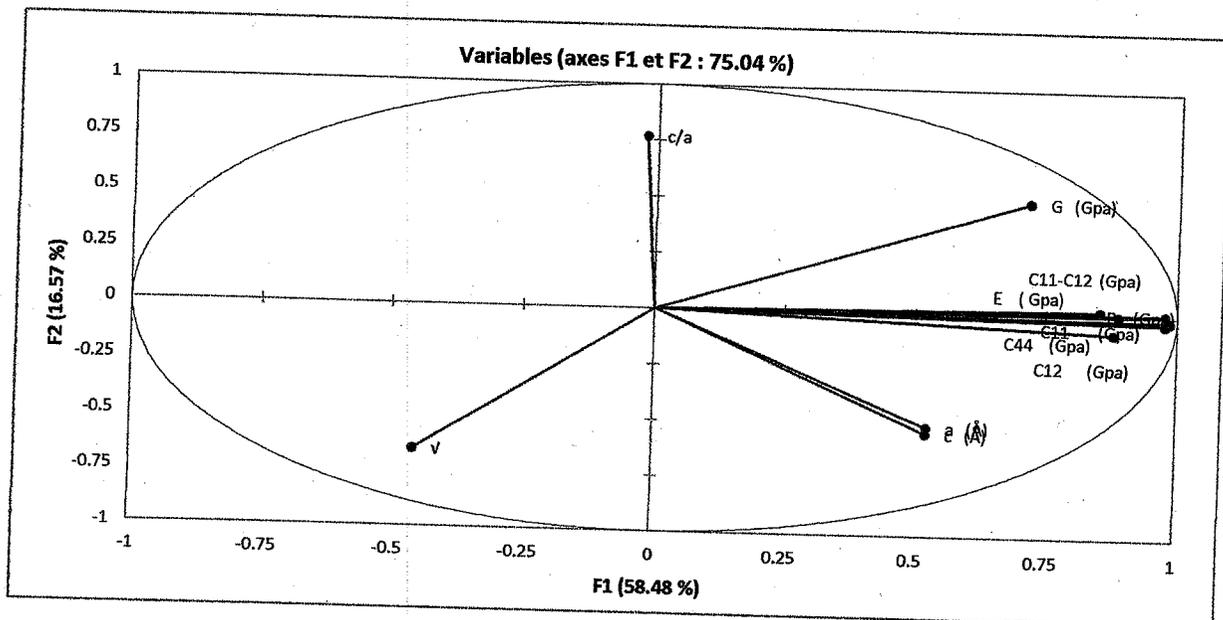


Figure (III-1) : graphe des propriétés (*loading plot*).

D'après la figure (III-1), on peut déduire que presque toutes les positions des propriétés mécaniques sont fortement corrélées (C_{11} , C_{12} , B , $C_{11} - C_{12}$, C_{44} , E) entre eux, tandis que les propriétés avec des valeurs négatives de $F1$ (c/a , v) sont inversement corrélées aux précédentes.

Par conséquent un groupement de matériaux dans le score plot, peut être dominé non pas par une mais par plusieurs propriétés déjà corrélées d'après le graphe des loading plot.

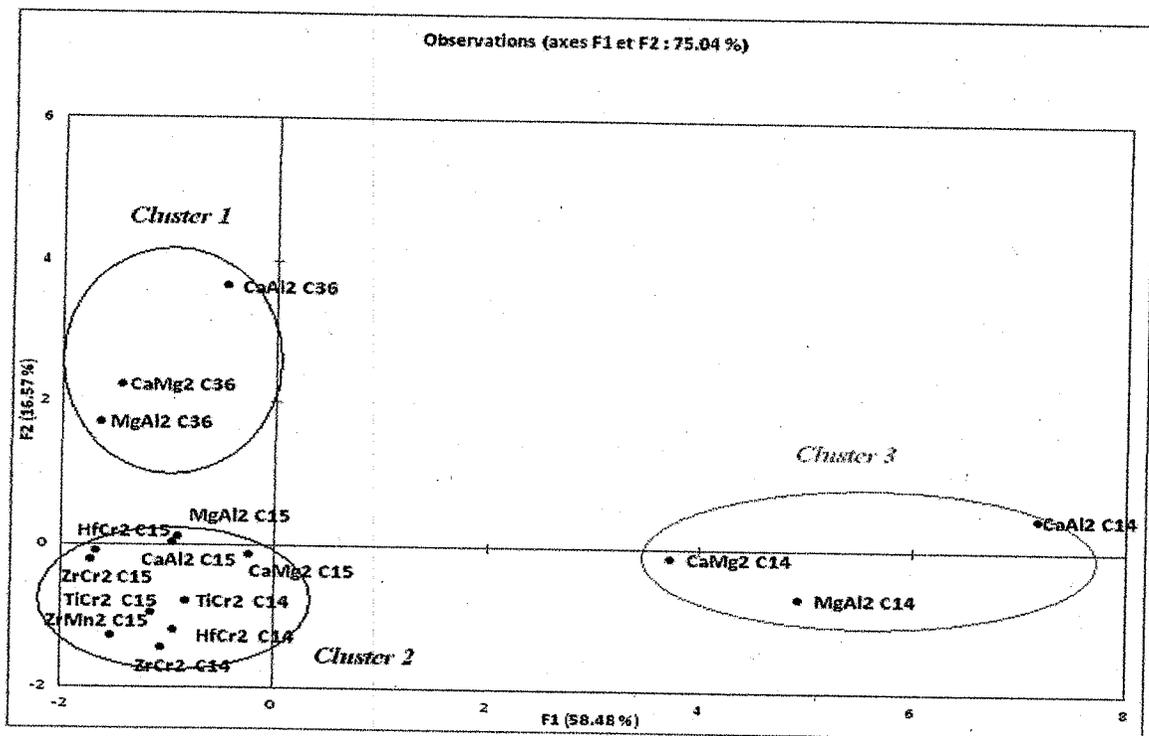


Figure (III-2) : graphe des observations (scores plot).

Sur le graphe bidimensionnel des score plot, les matériaux que nous avons étudié sont dispersés selon les axes F1 et F2. L'analyse est avant tout visuelle, et en observant la figure (III-2), nous pouvons distinguer 3 clusters des matériaux.

Cluster 1 : les matériaux cristallisant dans la structure C36.

Cluster 2 : les matériaux qui cristallisent dans la structure C15 et quelques uns cristallisant dans la C14.

Cluster 3 : les matériaux qui cristallisent dans la structure C14.

On remarque que le *cluster 1* se trouve dans le côté positif de l'axe vertical (PC2) et le *cluster 2* dans le côté négatif de (PC2), le *cluster 3* se situe à l'extrémité positive de l'axe horizontale (PC1),

Pour cette analyse, le signe de chaque composant principal (positif ou négatif) n'a aucune signification, c'est le positionnement (regroupement) qui compte.

Remarquons que le cluster 2 contient tous les composés adoptant la structure cubique (C15), et on trouve aussi les composés TiCr_2 , ZrCr_2 et HfCr_2 qui cristallisent dans C14. Notons que ces trois derniers matériaux contiennent du chrome, et peut être que c'est la raison pour laquelle ils sont dans le cluster des structures cubiques.

Pour mieux examiner les choses, nous allons considérer chaque cluster indépendamment des autres. Ainsi nous scrutons les composés du Cluster 3, et nous obtenons le graphe (III-3).

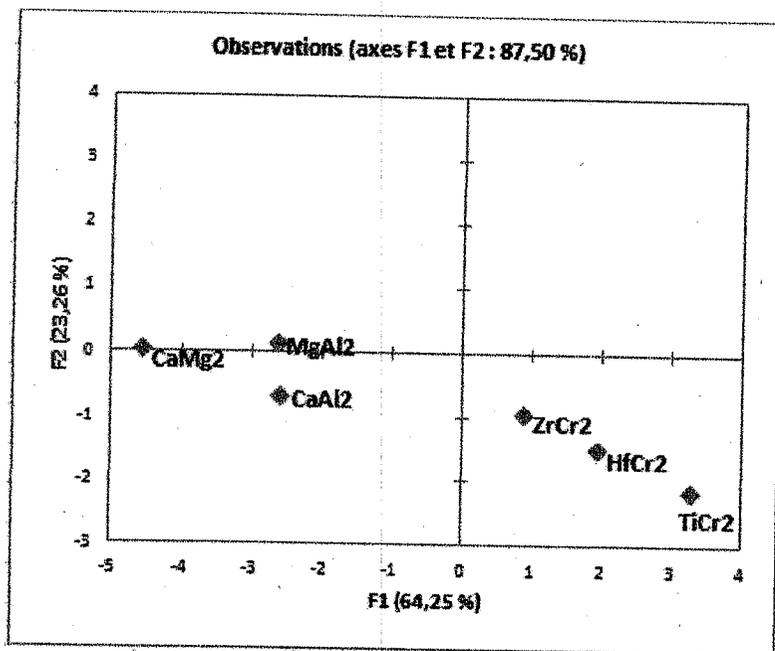


Figure (III-3) : graphe des observations des matériaux cristallisant dans la C14.

On remarque que les composés qui contiennent du chrome se trouvent du côté positif de F1, et le reste des composés se situent du côté négatif de F1. En comparant la position de chaque élément sur l'axe F1 avec la valeur du module de compression B

(voir tableau), on trouve que B augmente dans le sens positif de F1. Nous pouvons alors dire que l'axe F1 renferme l'information sur le module de compression.

La figure III-4 montre le score plot des matériaux qui cristallisent dans la structure cubique C15.

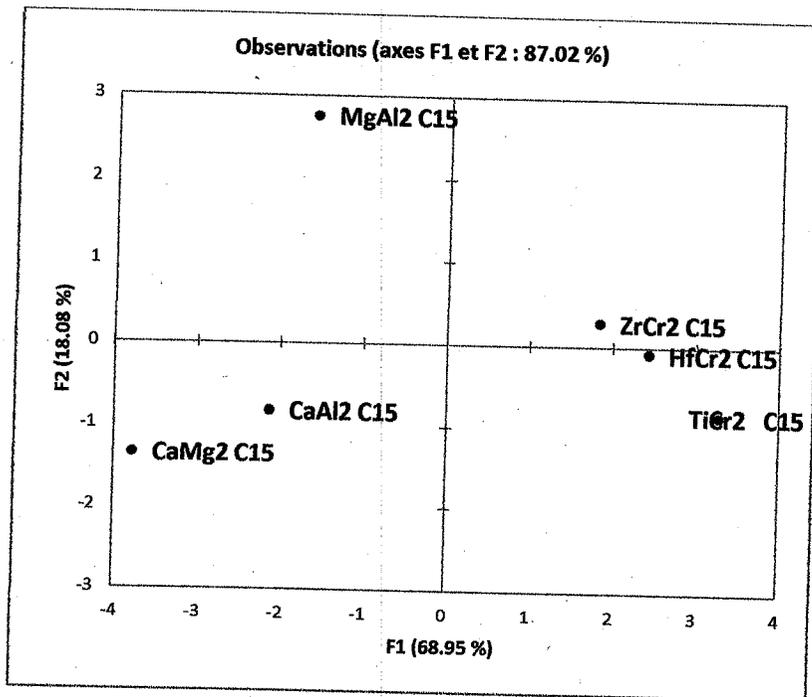


Figure (III-4) : graphe des observations des matériaux cristallisant dans la C15.

Le même raisonnement que précédemment est fait pour les matériaux qui cristallisent dans la structure cubique C15. Les composés qui contiennent du chrome se trouvent du côté positif de F1, et le reste des composés se situent du côté négatif de F1. Seulement, le raisonnement se fait ici par rapport à une autre propriété, en l'occurrence le module de Young E. En comparant la position de chaque élément sur l'axe F1 avec la valeur du module de young E (voir tableau), on trouve que E augmente dans le sens positif de F1. Nous pouvons alors dire que l'axe F1 renferme l'information sur le module de young.

Considérons maintenant, les matériaux purement hexagonaux (C14 et C36), le graphe correspondant est donné par la figure (III-5).

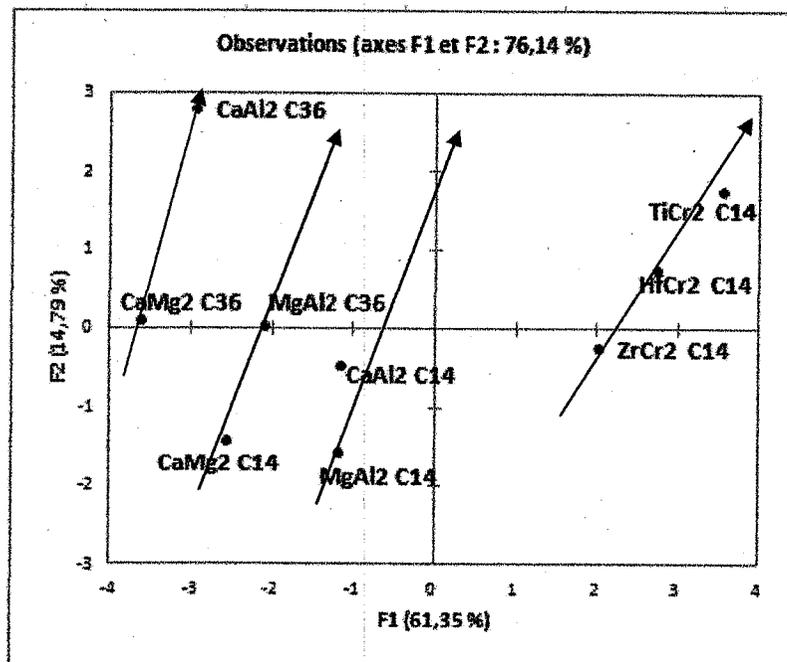


Figure (III-5) : graphe des observations des matériaux cristallisant dans la C14 et la C36.

Nous constatons d'abord que les composés à base de chrome sont toujours dégagés vers l'extrémité. D'autre part, les flèches indiquent le sens de C11-C12 ou bien du module de cisaillement G. Ainsi, CaMg_2 (C36) possède un module de cisaillement inférieur à celui de CaAl_2 (C36), et la valeur de C11-C12 de CaMg_2 (C36) est aussi inférieur à la valeur de C11-C12 de CaAl_2 (C36). Le même raisonnement est valable pour chaque flèche contenant deux ou trois matériaux.

Dans la figure qui suit (figureIII-6), nous avons prêté attention aux composés qui cristallisent dans la structure C15 et ceux qui cristallisent dans la structure C14 (figure III-6). Nous pouvons bien observer que les premiers se situent en dessous de F1, donc dans le côté négatif de F2, et ceux qui adoptent la C14 se situent dans le côté positif de F2. L'ACP ici permet de faire la distinction entre les deux structures. Ceci nous indique que l'axe F2 désigne la structure cristallographique.

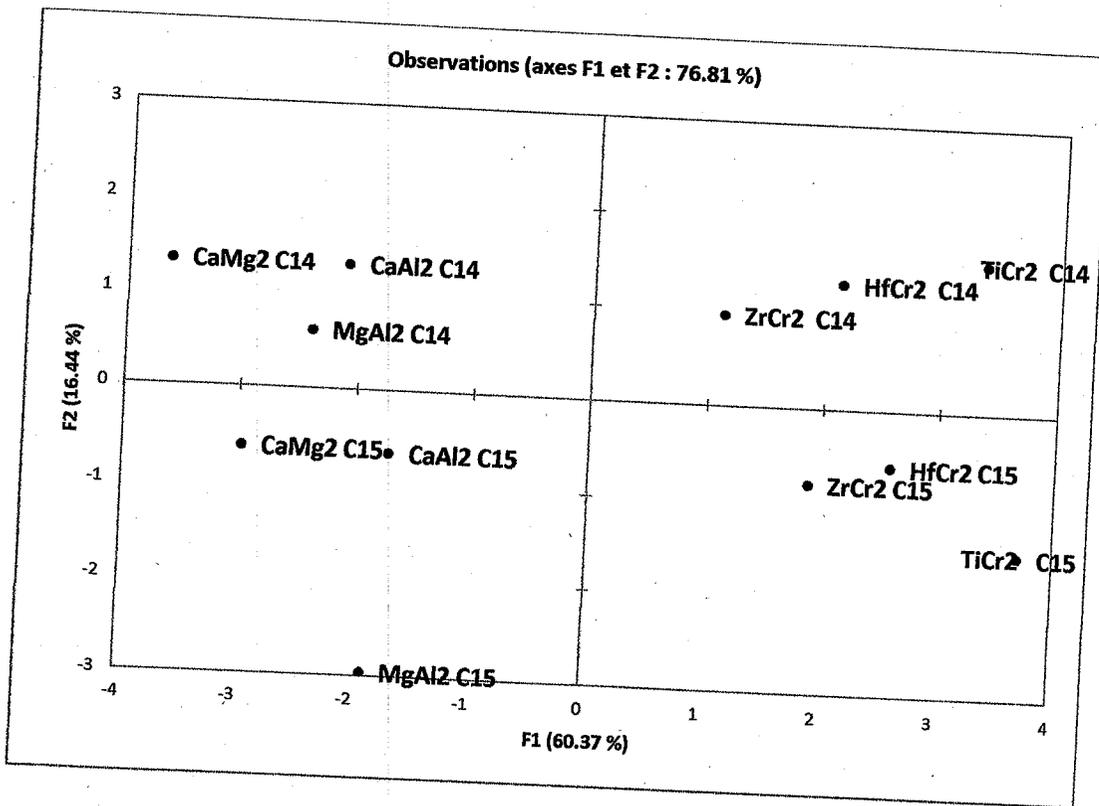


Figure (III-6) : graphe des observations des matériaux cristallisant dans la C14 et la C15.

La figure (III-7) montre la dispersion des composés cristallisant dans la structure C15 et C36. Nous pouvons distinguer 3 ilots. Le premier contient les matériaux adoptant la structure C36, le deuxième et le troisième incluent les matériaux cubiques.

Nous remarquons d'abord que les composés à base de chrome sont toujours dégagés vers l'extrémité. D'autre part, les tendances et les ilots sont compatibles avec les différentes valeurs de B et on a pour chaque ilot la valeur de B qui augmente dans le sens positive de F1 (voir tableau), pour CaMg₂ C15 par exemple, la valeur de B inférieur à celle de CaAl₂ C15 et aussi inférieur à celle de MgAl₂ C15 comme l'indique le flèche, et ainsi de suite pour les restes ilots.

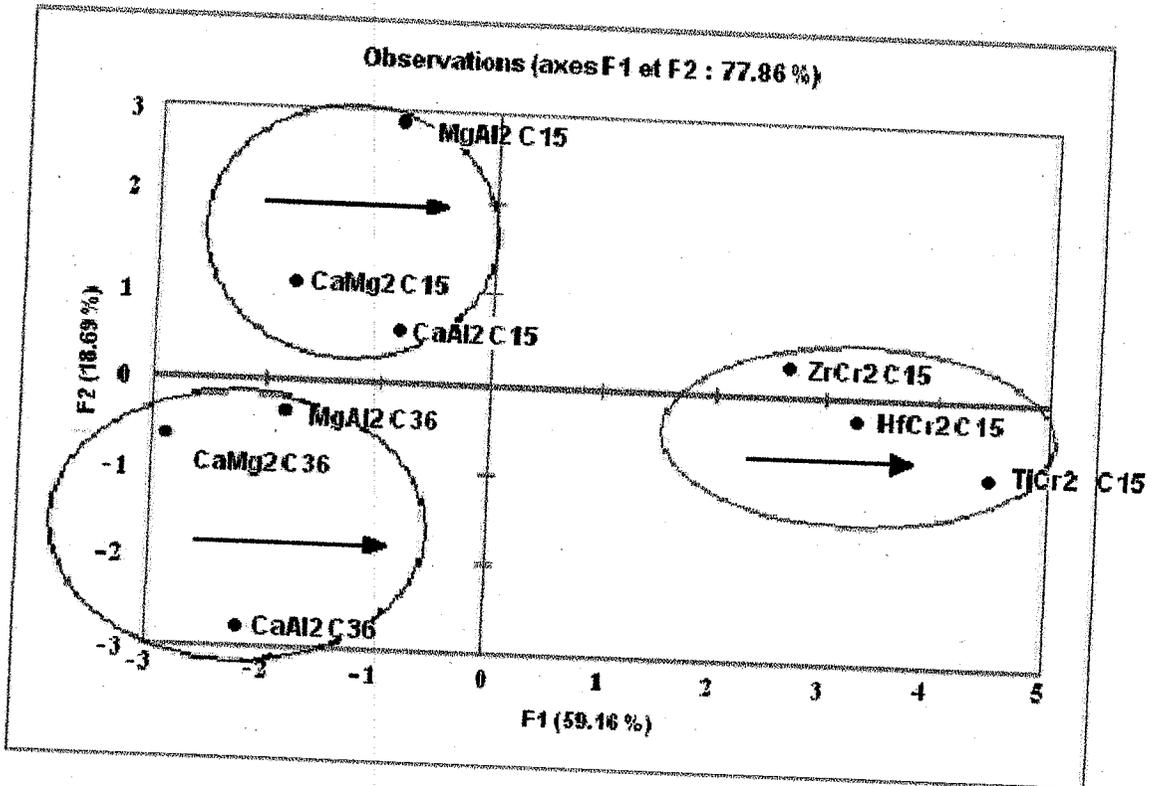


Figure (III-7) : graphe des observations des matériaux cristallisant dans la C15 et la C36.

III.5 Conclusion :

Dans ce chapitre nous avons rassemblé quelques composés des trois structures des phases de laves, en utilisant l'analyse en composante principale qui est une technique descriptive permettant d'étudier les relations qui existent entre les variables.

Nous avons montré comment se fait le clustering dans des groupes distincts, et leur classement suivant deux axes, appelés axes principaux (*PC1*, et *PC2*), en suite nous avons examiné chaque cluster séparément des autres pour mieux expliquer les choses.

Les résultats obtenus d'après l'application de la technique l'ACP sont les suivants :

- En considérant la grande base de données, c'est-à-dire pour l'ensemble des matériaux nous arrivons à les classer selon leurs structures et leurs propriétés.
- Pour les matériaux qui cristallisent dans la structure C14 nous constatons que les positions des matériaux sur l'axe F1 se fait selon le module de compression (B).
- Pour les matériaux qui cristallisent dans la structure C15 nous remarquons que les positions des matériaux sur l'axe F1 se fait maintenant selon le module de young (E).
- Pour l'ACP des matériaux qui cristallisent dans la C14 et la C36 les positions des matériaux se fait selon la valeur de C11-C12 et le module de cisaillement G.
- Pour l'ACP des matériaux qui cristallisent dans la C14 et C15 nous décelons que l'axe F2 indique la structure cristallographique.
- Pour l'ACP des matériaux qui cristallisent dans la C15 et la C36 nous remarquons que les tendances et les ilots se fait par rapport au module de compression (B) tout en différenciant la structure cristallographique.

De ce fait, l'ACP arrive à nous présenter et simplifier toutes les expositions de ces matériaux.

Conclusion Générale

Conclusion générale

Le but de ce mémoire consiste à étudier les phases de laves qui sont généralement des composés intermétalliques, et ils ont une grande importance en raison de leurs nombreuses applications pratiques dans divers domaines comme le stockage de l'hydrogène.

Les phases de laves cristallisent en trois structures : hexagonale de type C14 (MgZn_2), cubique de type C15 (MgCu_2), et C36 (MgNi_2) de type hexagonale.

Pour cela, nous avons montré dans le dernier chapitre les relations qui existent entre quelques composés de ces trois structures et qui forment la base de données. L'utilisation de l'une des techniques de datamining nous a permis de visualiser notre base de données sur des graphiques et facilité l'interprétation d'un grand nombre de données initiales. La méthode utilisée qui est l'ACP fait partie des méthodes d'analyses descriptives multivariées, elle permet de trouver des arrangements, des tendances et des corrélations entre les matériaux considérés et aussi entre les propriétés de ces matériaux.

Ce travail en perspective devrait aider à l'intégration de l'informatique de matériaux dans la conception de nouveaux matériaux et augmenter aussi nos chances de découvrir des nouvelles propriétés qui ne sont pas obtenues. Par ailleurs, ce travail peut être poursuivi en appliquant une autre technique de datamining qui soit prédictive.

I. Introduction .

XLSTAT est un logiciel d'analyse de données et de statistiques pour Microsoft Excel. Il offre de nombreuses fonctionnalités et des utilitaires pour faciliter la manipulation de vos données sous E **XLSTAT** permet notamment de réaliser des analyses (factorielles, ACP, AFD, AFC, etc.), diverses méthodes de régression et des tests (d'association, paramétriques). **XLSTAT** permet notamment de réaliser des analyses (factorielles, ACP, AFD, AFC, etc.), diverses méthodes de régression et des tests (d'association, paramétriques). **XLSTAT** offre de très nombreuses fonctionnalités qui font d'Excel un outil performant et facile d'accès pour répondre à la majorité de vos besoins en analyse de données et modélisation. **XLSTAT** fonctionne avec toutes les versions d'Excel, depuis la version 97 jusqu'à la version 2010 sous l'environnement Windows avec les versions X, 2004 et 2011 sous Mac OS.

Ce logiciel peut remplir plusieurs fonctions y compris la régression PLS et l'analyse en composantes principales (ACP) et ceci en faisant rentrer les variables explicatives (les X du modèle) qui sont indiquées par les composés des phases de laves dans notre cas, et les variables dépendantes (les Y du modèle) indiquées par les propriétés (C₁₁, C₁₂, C₄₄, a, c, E, B, G, V, C₁₁- C₁₂).

Dans cette annexe nous détaillons l'utilisation pratique du logiciel. Les tableaux des résultats et les graphes obtenus, sont décrits dans la partie précédente des résultats et discussions.

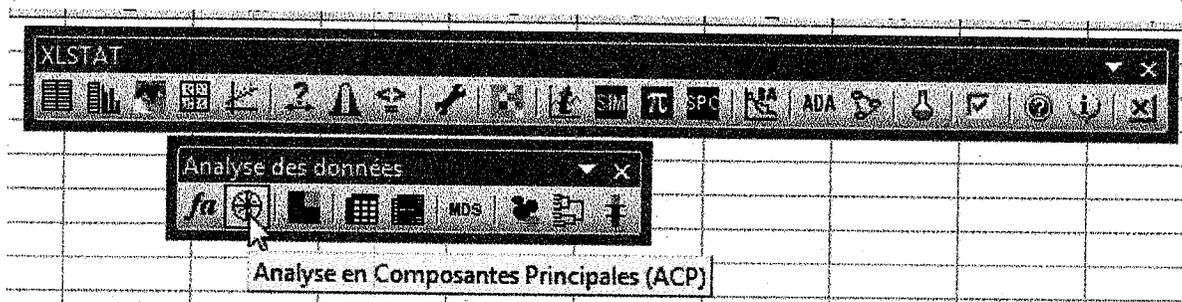
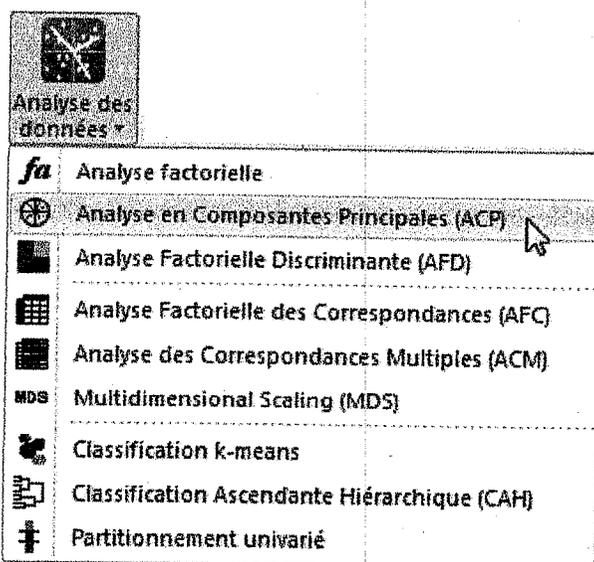
II. Utilisation du logiciel.

La figure suivante présente l'interface utilisateur du logiciel **XLSTAT**. La figure suivante présente l'interface utilisateur du logiciel **XLSTAT**. L'utilisation des différents boutons dépendra du type d'action désirée par l'utilisateur, parmi ces boutons, une correspond à l'analyse des données. En y accédant on pourra choisir l'ACP.

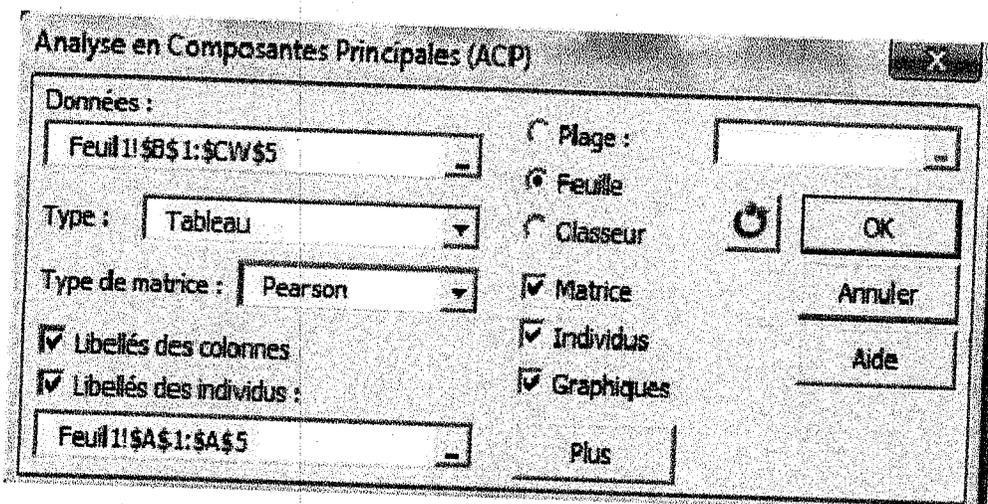


III. Faire une Analyse en Composantes Principales (ACP) avec XLSTAT.

Nous cliquons sur le menu XLSTAT/Analyse de données/Analyse en Composantes Principales, ou cliquons sur le bouton correspondant de la barre "Analyse de données" (voir ci-dessous).



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 « Général » réglez les paramètres suivants:
 1. Tableau observations/variables : Colonnes B à G
 2. Format des données : Tableau observations/variables
 3. Type d'ACP: Pearson (n)
 4. Libellés des variables : coché
 5. Libellés des observations : coché et sélectionnez la colonne A pour le nom des échantillons
 6. Feuille : pour obtenir les résultats dans une nouvelle feuille
- 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.