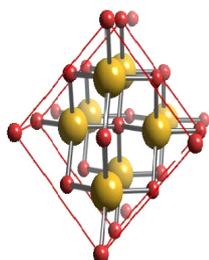


REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE ABOU-BAKR BELKAID DE TLEMCCEN



FACULTE DES SCIENCES
Département de Physique



Laboratoire d'Etude et Prédiction de Matériaux
Unité de recherches Matériaux et Energies Renouvelables

MEMOIRE DE MASTER

EN

PHYSIQUE de la MATIERE CONDENSEE

Présentée par :

Mme Benyacoub Amina

Thème

**Prédiction des nouveaux matériaux thermoélectriques
dans le cadre de génome de matériaux**

Soutenue le : 30/06/2012 devant la commission d'examen :

Pr. G.Merad

Président

Université de Tlemcen

Pr. H.I.Faraoun

Examineur

Université de Tlemcen

Pr. H. Aourag

Encadreur

Université de Tlemcen

Remerciements

*Ce travail a été effectué à l'Université Abou Bekr Belkaid de Tlemcen, sous la houlette de Monsieur **Aourag Hafid**, Professeur à l'Université de Tlemcen que je lui présente mes sincères remerciements pour ses conseils précieux et sa disponibilité aussi généreuse.*

*Je tiens à exprimer mon profonde gratitude à Monsieur **G. MERAD**, Professeur à l'Université de Tlemcen, de m'avoir accordé l'honneur de présider le jury de notre mémoire.*

*Mes remerciements les plus sincères à M^{elle} **Feraoun Houda**, Professeur à l'Université de Tlemcen d'avoir bien voulu examiner ce travail.*

J'adresse ainsi mes sentiments de reconnaissance à tous ceux qui ont aidé pour affiner ce mémoire.

Dédicace

Très reconnaissantes et redevables à l'égard de toute personne ayant contribué à mon éducation, instruction et formation. Je tiens vivement à remercier le Bon Dieu de m'avoir alloué le courage, la volonté et les capacités pour réaliser ce travail aussi modeste soit-il.

Aux deux bougies (mes Chers Parents) qui m'ont toujours accompagné durant tout mon parcours.

A mon marie Mourad.

A mes très chers frères et sœurs.

A toutes mes amies.

A tous ceux qui me sont chers.

Amina et Meriem

Introduction générale.....	01
Chapitre I : Généralités sur les intermétalliques	
I. Développement des études sur les composés intermétalliques.....	03
II. C'est quoi un composé intermétallique.....	03
II.1 Définition.....	03
II.2 Les différentes phases des intermétalliques.....	04
II.3 Les phases ordonnées et désordonnées.....	05
II.4 Symétrie cristalline et notation.....	05
II.5 Réseaux cristallographiques propres.....	06
II.5.1 Liaison covalente.....	08
II.5.2 Liaison ionique.....	09
II.5.3 Liaison métallique.....	10
III.1 Les propriétés des composés intermétalliques.....	11
III.1.1 Propriétés physiques.....	11
III.1.2 Propriétés magnétiques.....	12
III.1.3 Propriétés électriques.....	12
III.1.4 Propriétés mécaniques.....	12
III.1.5 Propriétés thermiques.....	13
III.1.5.1 la chaleur de formation.....	13
III.1.5.2 les enthalpies de formations.....	13
III.2 Effet sur les propriétés.....	14
III.2.1 Propriétés physiques.....	14
III.2.2 Propriétés de conduction.....	14
III.2.3 Effet sur les propriétés mécaniques.....	14
III.2.4 Effet de combinaisons intermétalliques.....	15
IV. Les diagrammes de phases de quelques intermétalliques.....	15
IV. 1 Les aluminures de titane.....	16
V. Les applications des intermétalliques.....	17
Référence.....	20

**Chapitre II :Les techniques d'exploitation de données
(DATA MINING)**

II.1 Introduction.....	21
II.2 Définition.....	21
II.3 Principe.....	21
II.4 Comment on l'utilise ?.....	22
II.5 Les techniques mises en œuvre.....	22
II.6 Méthodes globales.....	22
II.6.1 Les méthodes supervisées.....	22
II.6.2 Les méthodes non-supervisées.....	23
II.6.3 Les méthodes de réductions de données.....	23

II.7 Méthodes statistiques.....	23
II.7.1 Analyse en composante principale (ACP).....	24
II.7.1.1 projeter la réalité sur un plan.....	24
II.7.1.2 Définition de l'ACP.....	24
II.7.1.3 principe de l'ACP.....	24
II.7.1.4 Résumer les données.....	25
II.7.1.5 Les composantes principales.....	26
II.7.1.6 Caractères des composantes principales.....	26
II.7.1.7 propriétés de l'analyse en composante principale.....	26
II.7.1.8 Transformer les données_ centrage et réduction.....	27
II.7.1.9 Détermination de la structure du modèle de l'ACP.....	27
II.7.1.10 Interprétation des résultats de l'ACP.....	28
II.7.1.11 Application de l'ACP.....	29
II.7.1.12 l'ACP en résumé.....	30
II.7.1.13 conclusion.....	30
II.7.2-La régression (PLS).....	31
II.7.2.1Présentation.....	31
II.7.2.2 Interprétation du modèle PLS.....	32
II.7.2.3La régression sur composantes principales.....	33
II.7.2.4 Intérêt de la Régression PLS.....	33
II.7.2.5 Conclusion.....	33
Référence.....	34
Chapitre III :Résultats et discussions	
III.1 Introduction.....	35
III.2 Description de la méthode d'analyse de composantes principales.....	35
III.3 Résultats.....	35
III.3.1 Définition d'une base de données.....	35
III.3.2Base de données 1.....	35
III.3.3Techniques de calculs.....	37
III.3.3.1L'analyse en composantes principales (ACP).....	37
III.3.3.2 Le modèle PLS (Partial Least Square).....	43
III.3.4 Base de données 2.....	52
III.3.4.1 L'analyse en composantes principales (ACP).....	54
III.3.4.2 Le modèle PLS (Partial Least Square).....	57
Références.....	60
Conclusion générale.....	61
Annexe.....	62

Introduction générale



Les matériaux ont de tout temps défini le niveau de développement de notre civilisation. Aujourd'hui, ils jouent un rôle déterminant dans toutes les mutations technologiques.

Une grande partie du travail a été fait dans les 10 à 15 dernières années pour développer et caractériser les intermétalliques et de développer des technologies de traitement. En réponse à la nécessité d'une faible densité, les alliages à hautes performances pour une utilisation dans les composants de cellules et les moteurs à turbine (NRC, 1993, 1996), par exemple, des efforts concertés ont été réalisés ces dernières années pour améliorer les propriétés des alliages intermétalliques, en particulier alliages à base d'aluminiums (par exemple, TiAl, Ti₃Al, NiAl, Ni₃Al, FeAl).

Le terme « intermétalliques » désigne la classe de matériaux obtenus par combinaison de plusieurs éléments métalliques et dont la structure cristalline est différente de celle des métaux qui le constituent. Pour désigner ces composés, de nombreux auteurs utilisent soit le terme « d'alliages intermétalliques », soit celui de « composés intermétalliques ». Ces composés se forment du fait des forces de liaison plus importantes entre atomes de natures différentes comparées à celles des liaisons entre atomes identiques.

La nature non directionnelle des liaisons métalliques est perdue partiellement donnant lieu à des structures chimiquement ordonnées qui procurent à ces composés des propriétés particulières. Initialement, ces composés étaient employés comme phases durcissantes, puis ils ont été isolés afin de les développer comme principal composant dans différentes applications.

L'objet de notre Mémoire est de trouver des corrélations entre les propriétés structurales et thermodynamiques des composés intermétalliques ternaires en utilisant deux techniques bien connues en datamining tel que l'analyse en composantes principales (PCA) et la régression des moindres carrés partiels (PLS), les données peuvent être examinées de sorte que les tendances et les corrélations deviennent évidentes. En plus, le nombre de propriétés exigées pour expliquer un système peut être réduit à un nombre minimum de sorte que le problème de créer des quantités suffisantes de données et d'analyser ces données soit réduit.

Le travail développé dans ce mémoire a été réparti sur trois chapitres :

Dans Le premier chapitre nous rappelons quelques généralités sur les composés intermétalliques, un aperçu sur leurs propriétés particulièrement structurales et thermodynamiques, aussi leurs applications, leurs réseaux cristallographiques et on site un diagramme de phase de quelques alliages.

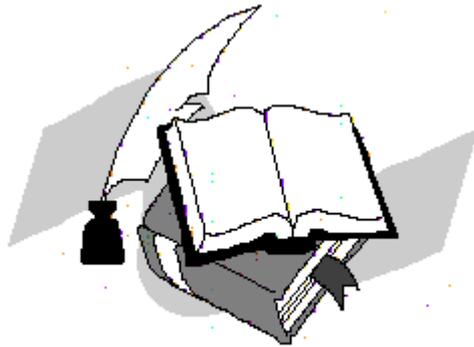
Dans Le deuxième chapitre nous donnons une idée générale sur l'ensemble des techniques de datamining, leur historique et leurs champs d'application nous prenons en considération l'identification du modèle ACP et le principe de la régression PLS (Partial Least Square).

Dans le troisième chapitre nous traitons les résultats liés à l'application des approches du datamining, notamment l'analyse en composantes principales (ACP) et la régression (PLS) Une discussion détaillée de la pertinence de notre procédure est également présentée en se basant sur la comparaison de nos résultats avec les données de la littérature.

Enfin, une conclusion générale récapitule les principaux résultats obtenus et propose des perspectives de travaux futurs.

CHAPITRE I

Généralités sur les intermétalliques



I. Dévoilement des études sur les composés intermétalliques :

Avant les années 70, très peu de progrès ont été faits dans les études et les développements sur les composés intermétalliques en raison de leur fragilité intrinsèque à température ambiante, cela limitait leurs utilisations industrielles. A la fin des années 70, quelques progrès remarquables sur les études de certains composés intermétalliques ont déclenché à nouveau l'intérêt du monde entier pour les alliages intermétalliques [1]. En 1976, les chercheurs américains ont trouvé que, par addition d'élément et par la technique de métallurgie des poudres, la ductilité et la résistance d'alliage à base de TiAl et de Ti_3Al pouvaient être améliorées [2]. Ensuite, ils ont trouvé qu'à la température ambiante, après avoir substitué partiellement au Co du Ni ou du Fe, le composé Co_3Al pouvait être transformé d'une structure hexagonale à une structure $L1_2$, ce qui présente une certaine ductilité [3]. En 1979, les chercheurs américains et japonais ont indiqué presque en même temps que, par l'addition du bore B, la ductilité du composé Ni_3Al pouvait être significativement améliorée [4].

Depuis ces 20 dernières années, les composés intermétalliques sont largement étudiés, les plus grands progrès concernent les composés intermétalliques de types A_3B et AB dans les trois systèmes tels que le Ni-Al, Fe-Al et Ti-Al. Les composés intermétalliques tels que le Ni_3Al , NiAl, TiAl, et Ti_3Al présentent de bonne tenue en température, ils pourraient donc être utilisés comme matériaux de structure mécanique à haute température. Actuellement, certains éléments de turboréacteurs fabriqués en alliages de Ni_3Al , de TiAl et de Ti_3Al sont en train d'être testés en dimensions réelles. En particulier, les composants en alliages de Ti_3Al sont déjà utilisés en pratique. Les composés de Fe-Al seront également prometteurs dans le domaine industriel grâce à leurs bas coûts d'accès, à leurs ressources mondiales et à leurs hautes résistances à l'oxydation. Ils attirent donc beaucoup de chercheurs dans le monde entier [5].

Ces dernières années, des nouveaux systèmes de composés intermétalliques ayant des structures beaucoup plus complexes et des points de fusion beaucoup plus élevés tels que Nb -Al, Be-Nb, Mo-Si deviennent une nouvelle orientation d'étude et une nouvelle tendance de développement comme des nouveaux systèmes de matériaux de structure en température [6].

II. C'est quoi un composé intermétallique :

II.1 Définition

Un alliage forme un intermétallique à certaines compositions et conditions de pression. Il s'agit d'une alternance d'atomes qui forment une structure périodique. Les intermétalliques se forment lorsque les atomes de l'alliage ont une forte différence d'électronégativité ; la

densité électronique devient alors hétérogène, la liaison n'est plus strictement métallique mais devient partiellement covalente ou ionique.

Initialement, les alliages intermétalliques formaient des précipités dans un alliage plus complexe (par exemple dans les super-alliages à base nickel comme les Inconels). Ces précipités jouent un rôle important dans le renforcement mécanique des alliages (durcissement structural). Par la suite, ces phases ont été isolées, caractérisées, et on a commencé dans les années 1960 à les utiliser comme constituant massif de pièces et non plus comme phase mineure.

Du fait de l'ordre chimique, ces alliages se déforment assez mal plastiquement : l'ordre chimique devant être conservé, le déplacement au niveau de la dislocation doit être de deux atomes au lieu d'un seul habituellement. Les dislocations sont souvent dissociées et apparaissent comme deux dislocations liées, on les appelle de ce fait des « super-dislocations ». Leurs propriétés mécaniques sont donc souvent appréciables.

Les composés intermétalliques définissent aussi comme :

Une phase qui cristallise avec une structure différente de celles des éléments constituants.

C'est une indication de l'existence de changement de nature des liaisons, en général vers des liaisons moins métalliques et plus covalentes (c'est-à-dire vers des transferts d'électrons plus importants), ainsi que d'autres facteurs de stabilité.

II.2 Les différentes phases des intermétalliques :

Lorsque deux éléments A et B sont miscibles en toutes proportions et forment une solution solide continue pour certains domaines de concentration, il peut se former des surstructures. Les composés A_3B , AB , AB_3 sont des exemples de ces nouvelles phases. Le nom de phases intermédiaires est plutôt réservé à celles qui possèdent une structure différente de celle des métaux de base ou des solutions solides terminales. On parle de composé intermétallique lorsque la phase intermédiaire n'existe qu'à l'état ordonné. Ces composés ont un caractère métallique plus ou moins marqué. Leur stabilité dépend de différents facteurs : facteur de valence (ou de concentration électronique), facteur de taille, facteur électrochimique (différence d'électronégativité). La prédominance de ces facteurs permet de classer les phases intermédiaires : phases de Hume Rothery (la structure de certains alliages dépend du nombre d'électrons par atome des composants), phases de Laves, phases complexes de type o, et composés semi-métalliques (composés d'un métal avec un élément des sous-

groupes B, ou présence des atomes interstitiels H, B, C, N, O en solution d'insertion dans un métal).

Pour préciser les termes, nous parlerons de « composé intermétallique » quand la composition stoechiométrique est atteinte, et de « phase ou alliage intermétallique » s'il agit d'une solution solide sur la base du composé.

➤ **Phases de Hume Rothery :**

- la phase bêta β : cubique centré, caractérisée par un rapport **3/2**
- la phase gamma γ : cubique complexe, caractérisée par un rapport **21/13**
- la phase epsilon ϵ : hexagonale, caractérisée par un rapport **7/4**.

II.3 Les phases ordonnées et désordonnées

L'alliage formera une solution solide ordonnée, les atomes tendant à s'entourer d'atomes de l'autre espèce chimique. Les composés intermétalliques sont généralement ordonnés à basse température (cas notamment des alliages Ti-Al-Nb, développés pour les moteurs d'avion pour remplacer les superalliages beaucoup plus lourds à propriétés équivalentes). Ils sont également durs et fragiles car les dislocations, pour se déplacer, doivent respecter les voisinages en termes d'espèces chimiques. A haute température, le terme d'entropie l'emporte et la plupart des solutions solides sont désordonnées.

Quelle différence entre phases ordonnées et composés intermétalliques ?

- Il y a une forte différence de structures par rapport aux éléments constituants.
- Il existe un petit, voire très petit, domaine d'existence en composition.
- Dans le cas des phases ordonnées il existe une transition ordre-désordre.

II.4 Symétrie cristalline et notation

Lorsque l'on regarde une maille sans prêter attention à la nature chimique, on voit ressortir une structure cristalline ; cependant, du fait de l'alternance chimique, la structure réelle peut être différente.

Par exemple, dans le cas du FeAl en phase B2, on a en apparence une structure cubique à corps centré, mais l'alliage a en réalité une structure cubique simple.

Pour prendre en compte cette « sur structure chimique », on utilise souvent la notation Strukturbericht. [7]

Exemples d'alliages intermétalliques :

- Aluminiures de fer :
 - FeAl B2
 - Fe₃Al
- alimuniures de nickel :
 - NiAl β
 - Ni₃Al
- alumiures de titane :
 - TiAl β,γ
 - Ti₃Al α₂
- alliages de terres rares pour applications magnétiques :
 - samarium-fer Sm₂Fe₁₇
 - samarium-cobalt
 - samarium-fer-cobalt
 - gadolinium-fer Gd₂Fe₁₇
 - samarium-fer-gadolinium SmGdFe₁₇

II.5 Réseaux cristallographiques propres

Les composés intermétalliques possèdent des réseaux cristallographiques propres, différents de ceux des éléments constituants, comme le Fe₃C et Cu₂Mg voir figure (1.1). C'est une indication de l'existence de changement de nature des liaisons, en général vers des liaisons moins métalliques et plus covalentes ou ioniques (ils conservent leur caractère métallique, c'est-à-dire que tous les électrons de valence ne sont pas engagés il reste des électrons libres).

Une raison importante pour avoir une compréhension approfondie des liaisons interatomiques dans les solides est que dans certains cas, le type de liaison permet d'expliquer quelques propriétés des matériaux. Si l'on considère par exemple le carbone, qui peut exister en forme de graphite ou de diamant. Alors que le graphite est relativement friable, le diamant est le matériau le plus rigide connu. Cette disparité drastique est directement attribuée aux types des liaisons interatomiques qui se trouvent dans les deux variantes du carbone.

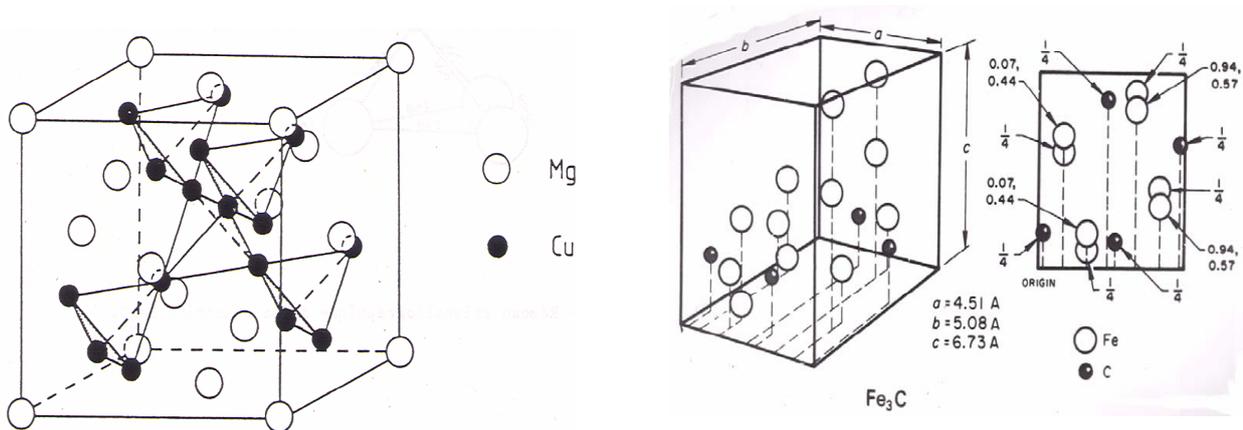


Figure (1.1) : Exemples de matériaux intermétalliques Fe_3C et Cu_2Mg . [8]

Deux atomes séparés par une distance très grande par rapport à leurs diamètres constituent deux systèmes indépendants. En les rapprochant suffisamment l'un de l'autre, il se produit une superposition de leurs orbitales respectives qui modifie la forme du potentiel dans lequel évoluent les électrons. Il en résulte un changement des orbitales elles-mêmes et, par conséquent, une variation de l'énergie des électrons. A ce stade, les deux atomes forment un système.

Ces atomes exercent l'un sur l'autre une force répulsive si l'énergie du système augmente quand ils se rapprochent. Dans le cas contraire la force est attractive. L'existence d'une force attractive est fréquente, comme en témoigne le nombre extrêmement élevé de molécules différentes et de substances se présentant sous forme condensée, c'est-à-dire solide ou liquide.

Selon leurs positions respectives dans la table de Mendéléev, les atomes établissent entre eux des liens de valence de natures différentes. L'étude de ces liens à partir d'équations fondamentales de la physique quantique n'a jusqu'ici été entreprise avec succès que dans les milieux cristallins et les molécules simples.

On distingue trois types de liens de valence forts ainsi nommés en raison de l'énergie de cohésion élevée qui leur est associée : ionique, covalente (homopolaire) et métallique.

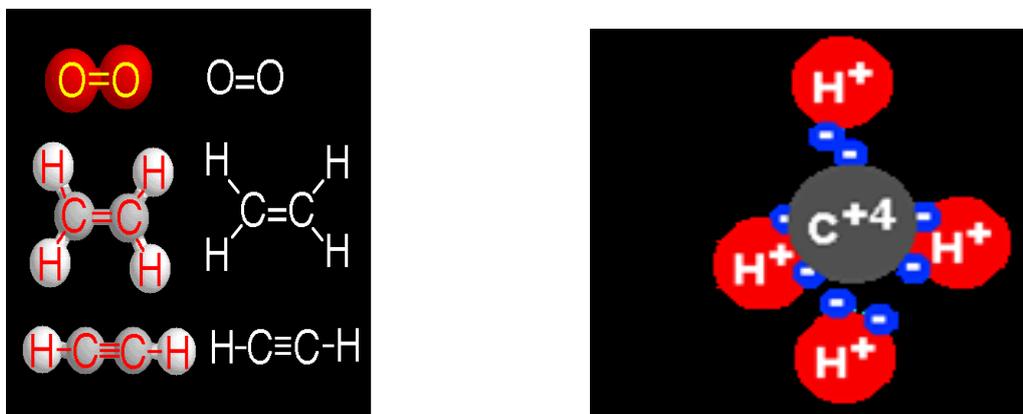
Pour chaque type, la liaison implique les électrons de valence et la nature des liaisons dépend des structures électroniques des atomes. En général, chacun de ces trois types de

liaisons découle de la tendance des atomes à avoir une structure électronique stable pareille à celles des gaz rares en remplissant entièrement leurs couches externes.

Il est rare qu'une liaison donnée appartienne strictement à l'un des types mentionnés. La plupart du temps, elle présente les caractères de deux, voire trois types simultanément, avec prédominance éventuelle de l'un deux ; Par exemple, dans le cristal d'arséniure de gallium GaAs, la caractéristique covalente est environ deux fois plus apparente que le caractère ionique alors que dans NaCl le caractère ionique domine à 94% le caractère covalent. Le silicium est purement covalent.

II.5.1 Liaison covalente

- Partage des électrons accumulation de charge entre deux cœurs atomiques relativement positifs.
- liaison forte (200-700kJ/mole)
- partage de 1 ou plusieurs paires d'électrons pour compléter leur couche électronique extérieure à 8 électrons
- Ex. molécules d'éléments non métallique : O₂, N₂, CH₄, diamant
- Molécules avec atomes dissimilaires: CH₄, H₂O, HNO₃, SiO₂



orientée ou directionnelle

Figure (1.2) : exemple de liaison covalente CH₄ [9]

➤ **Caractéristiques des cristaux covalents :**

- Leur point de fusion est élevé car leur énergie de cohésion est grande : 714 kJ/mol [10] pour le diamant qui se sublime dès 3652°C et 1186kJ/mole [10] pour le carbure de silicium qui fond à 2600°C [10].
- Ils sont très peu déformables en raison du caractère directionnel de la liaison, duquel résulte encore un arrangement spatial relativement peu dense des atomes (faible facteur de remplissage atomique).
- Ce sont des isolants ou des semi-conducteurs, car les électrons de valence sont fortement liés aux atomes et la conductivité par ions est négligeable.

II.5.2 Liaison ionique :

Transfert de charge de l'élément le moins électronégatif (le plus métallique) vers l'éléments le plus électronégatif (le moins métallique); formation duo (+-) et attraction coulombienne.

.les sels (NaCl, CaCl₂)

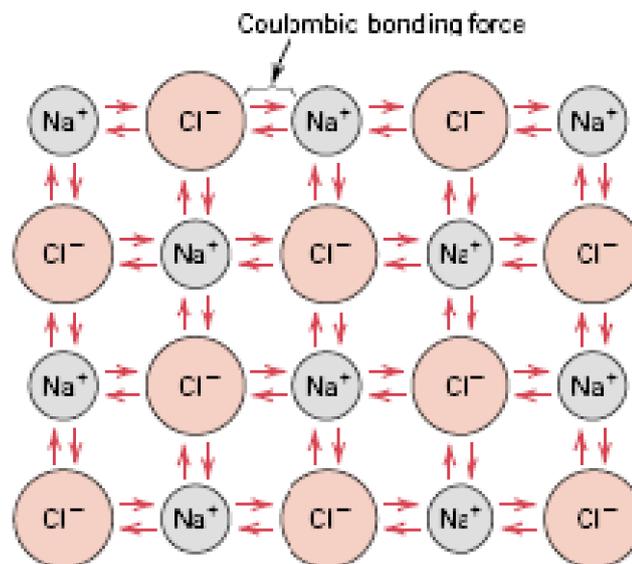


Figure (1.3) : Exemple de matériau possédant une liaison ionique (NaCl).

➤ Caractéristiques des cristaux ioniques :

- Les cristaux ioniques sont durs, car leur énergie de cohésion est élevée, de l'ordre de 750kJ/mole pour les composés halogène-métal alcalin. Par comparaison, l'énergie de cohésion typique de la valence de Van der Waals est de l'ordre de 10kJ/mole [10]. Pour la même raison, leur point de fusion est élevé.

- Leur conductivité électronique est faible car les électrons sont solidement attachés aux ions. A haute température une conduction ionique apparaît.
- Leur permittivité est élevée et varie faiblement avec la température et la fréquence.

II.5.3 Liaison métallique :

Ions positives dans une mer d'électrons (de valence) négatifs.

- énergie de liaison : 70 à 850 kJ/mole.
- Les électrons périphériques ne sont plus attachés à un atome.
- Ils sont délocalisés dans le solide sous forme de **nuage** qui assure la cohésion.

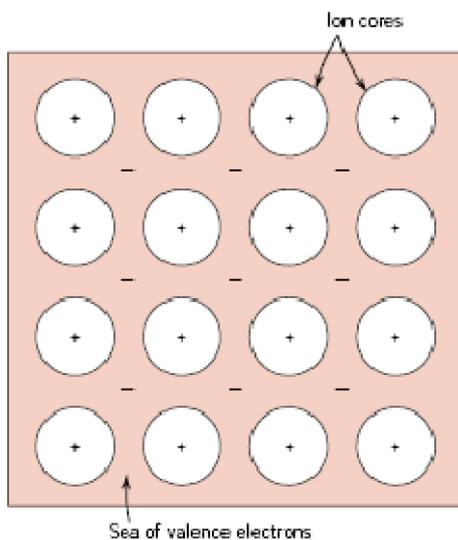


Figure (1.4) : configuration de la distribution des ions métalliques dans le gaz d'électrons libres.

➤ Caractéristiques des cristaux métalliques :

- La liaison métallique est caractérisée par une énergie de cohésion en moyenne plus faible que la liaison covalente ou ionique. A titre d'exemple, cette énergie vaut 311kJ/mole dans l'aluminium et 386kJ/mole dans le fer [10].
- Les cristaux métalliques sont ductiles. Cette propriété découle du caractère non directionnel de la liaison métallique.
- Moyennant certaines analogies physico-chimiques, des métaux différents peuvent être mélangés, parfois en toutes proportions, et former des alliages. Ceci provient du fait que le nombre exact d'électrons disponibles pour établir la liaison n'influence pas celle-ci de manière importante.

- La haute conductivité électrique des cristaux métalliques a déjà été mentionnée. Elle s'accompagne d'une haute conductibilité thermique, liée elle aussi à la mobilité des électrons.
- Propriétés optiques : reflet métallique.
- Structure cristalline compacte.

III. Les propriétés des composés intermétalliques

Nous observerons les matériaux à trois niveaux d'échelle :

- ✓ Échelle atomique
- ✓ Échelle microscopique
- ✓ Échelle macroscopique

Il existe des relations étroites entre les propriétés macroscopiques des matériaux et la structure rencontrée aux différents niveaux d'échelle, et qu'il nous faut connaître.

Les propriétés physiques, chimiques, mécaniques des matériaux permettent de distinguer les grandes familles. Ces propriétés dépendent de la structure du matériau aux différentes échelles. Il faut donc maîtriser les structures pour maîtriser les propriétés.

Il faut connaître les structures aux différentes échelles pour pouvoir agir dessus ce qui implique la compréhension des mécanismes qui font évoluer les structures.

Il est à noter que la plupart des progrès technologiques importants sont presque tous liés à l'amélioration des propriétés des matériaux. [11].

III.1 propriétés

Les composés intermétalliques présentent des propriétés intéressantes pour de nombreuses applications.

III.1.1 Propriétés physiques

Lors de la formation des composés intermétalliques les principales propriétés physiques à savoir la dureté, la conductivité thermique et électrique, la susceptibilité magnétique..., seront modifiées de manières significatives de celles de leurs métaux constituants.

Aussi que la dureté des intermétalliques est plus grande que celles des métaux indiquant leurs grandes fragilités. De même le module d'élasticité est plus grand, provoquant une ductilité

inférieure. Les coefficients de dilatation thermique et électrique sont généralement basse pour ces composés.

Les composés à grande maille ont souvent une très grande dureté, accompagnée d'une fragilité très marquée ; ils sont souvent utilisés sous forme de pièces frittées (carbures pour outil de coupe).

III.1.2 Propriétés magnétiques

Les composés intermétalliques sous forme de particules dispersées dans une matrice métallique (phénomène de précipitation) modifient de façon notable les propriétés de cette matrice. De nombreux alliages sont durcis par précipitation : précipitation des carbures (aciers spéciaux), des composés intermétalliques (alliages légers, aciers à hautes performances). certains composés intermétalliques sont ferromagnétiques comme SmCo_5 ...

III.1.3 Propriétés électriques

Il ont une densité électronique souvent anisotrope (caractère partiellement covalent de la liaison). Des composés covalents comme InSb et GaAs sont utilisés comme semi-conducteurs, d'autres (de type A_3B) ont des températures critiques de supraconduction élevées : 18,05 K pour Nb_3Sn et 17,9 K pour V_3Si .

III.1.4 Propriétés mécaniques

Parmi les propriétés attribuées aux intermétalliques en ingénierie sont leurs résistances spécifiques élevées (module d'élasticité/densité).

Le module d'élasticité pour plusieurs intermétalliques est grand à température élevée, il compris entre 150 GPa pour les matériaux qui ont de basse densités entre 2 à 5g/cm^3 .

Le module d'élasticité des matériaux ordinaires diminue plus rapidement contrairement aux intermétalliques autant la température s'élève. Cela résulte de leurs liaisons partiellement non métalliques et leurs températures de fusion élevées.

Les constantes élastiques (module de compressibilité et module de cisaillement) sont fortement liées aux propriétés mécaniques et sont importantes dans la conception et la sélection des matériaux. Le module de compressibilité des intermétalliques est moins sensible aux traitements thermophysiques ou aux propriétés microstructurales. Car il mesure la force interatomique intrinsèque, il est souvent utilisé pour estimer les potentiels interatomiques utilisés dans les simulations informatiques. Il est proportionnel à l'énergie de cohésion. C'est

des critères utilisés pour caractériser la dureté. Ainsi, il est important de prévoir ou d'estimer les constantes élastiques pour des études théoriques ou des applications industrielles.

Les eutectiques contrôlés possèdent de remarquables propriétés mécaniques : ce sont des composés intermétalliques sous forme de longues baguettes ou de grandes lamelles, régulièrement disposées parallèlement à une même direction, dans une matrice plus ductile (principe du renforcement par fibres).

En outre les propriétés mécaniques des matériaux sont fortement corrélés au type de liaison entre atomes car il détermine l'énergie de cohésion E_{coh} de la matière, la force nécessaire pour écarter les uns des autres jusqu'à rupture.

Les composés à grande maille ont souvent une très grande dureté, accompagnée d'une fragilité très marquée ; ils sont souvent utilisés sous forme de pièces frittées (carbures pour outil de coupe).

III.1.5 Propriétés thermiques

Les propriétés thermiques concernent la capacité de stockage et de transfert de l'énergie d'agitation thermique.

III.1.5.1 La chaleur de formation

La chaleur de formation des composés intermétalliques binaires terre-rare-élément de transition peut être calculée à partir du modèle semi-empirique de Miedema. Dans la structure des composés RM_5 (R = terre-rare, M = élément de transition) la contraction du rayon atomique des atomes de terre-rare augmente avec la différence d'électronégativité * entre les constituants et est reliée à l'enthalpie de formation des composés. L'état de valence de Cédans Ce-M peut être prévu à partir du calcul de l'enthalpie de formation; la valeur de la différence d'enthalpie entre les deux états de valence (Ce^{3+} ou Ce^{4+}) est environ 80 kJ (mol Ce)⁻¹.

III.1.5.2 Les enthalpies de formation

L'enthalpie de formation H et aussi une autre grandeur qui peut nous servir pour identifier le comportement mécanique d'un matériau, c'est une mesure de l'énergie mise en

jeu, lors de la formation d'un composé chimique à partir des corps simples correspondants, purs, pris dans l'état stable à la température considérée T.

Les enthalpies de formation des alliages liquides et d'un certain nombre de composés intermétalliques des systèmes binaires (Ca,Sn) et (Ca,Pb) ont été mesurées par calorimétrie "Calvet". (Ca,Sn), l'enthalpie de formation de CaSn est plus exoénergétique que celle de CaPb, ce qui donne une indication sur les stabilités respectives de ces deux phases dans les alliages ternaires à basse température.

III.2 Effet sur les propriétés

III.2.1 Propriétés physiques

- ✓ Les intermétalliques ayant leur réseau propre ont une densité propre.

III.2.2 Propriétés de conduction

- ✓ Les intermétalliques sont de mauvais conducteurs. Leur présence dans les applications de conduction n'est donc pas souhaitée.
- ✓ Il faut toutefois signaler qu'il vaut mieux qu'une impureté présente dans un conducteur soit combinée dans un intermétallique, plutôt qu'en solution.
 - Dans le premier cas une partie du conducteur conduit mal le courant : c'est celle où l'intermétallique est présent.
 - Dans le deuxième cas, tout le conducteur conduit mal le courant [8]

III.2.3 Effet sur les propriétés mécaniques

Les liaisons des composés intermétalliques sont principalement de nature ionique ou covalente. Il n'y a dès lors pas de possibilité de déformation plastique.

- ✓ Ces composés sont donc durs, mais cassants.
- ✓ En général, ils sont peu souhaités, car ils fragilisent les métaux.
- ✓ Ils sont, par contre, très bénéfiques lorsqu'ils sont disséminés dans une phase métal pur ou solution.

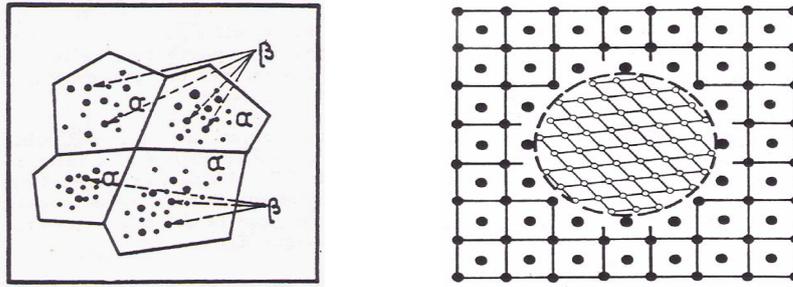


Figure (1.5) : germination de phase intermétallique dans une matrice métallique.[8]

III.2.4 Effet de combinaisons intermétalliques disséminées dans une matrice métallique

Lorsque les composés intermétalliques sont disséminés dans une phase métallique pure ou solution, ils freinent le mouvement des dislocations, mais ne l'entravent pas complètement.

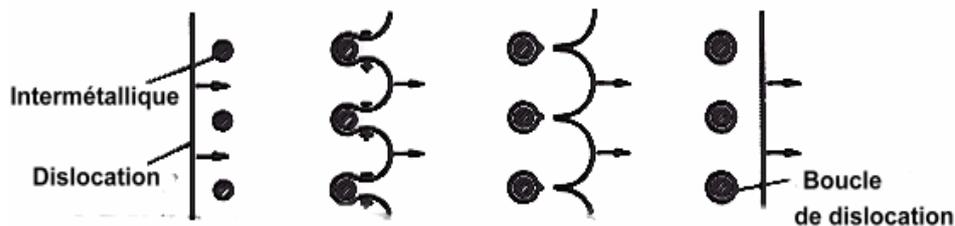


Figure (1.6) : mécanisme de freinage des dislocations par les phases intermétalliques.[8]

IV. Les diagrammes de phases de quelques intermétalliques

Un diagramme de phases (ou diagramme d'équilibre) permet de résumer les constitutions d'équilibre d'un système d'alliage. Un diagramme de phases permet de prédire, pour un mélange donné, la constitution des phases en présence, en équilibre les unes avec les autres.

Deux facteurs physiques ont une influence sur la nature et la composition des phases présentes : (1) la température qui joue un rôle particulièrement important lors de la coulée et dans les modifications des propriétés mécaniques des alliages dentaires, et (2) la pression qui est habituellement négligée car elle n'a d'influence qu'à des niveaux extrêmement élevés.

Deux types de transformations peuvent être retrouvés dans les diagrammes de phases.

La stabilité des combinaisons des intermétalliques est satisfaite que dans un domaine de concentration, car il existe des zones du diagramme de phase où il y a mélange de cristaux d'intermétalliques et de cristaux de solution. Il peut y avoir plusieurs combinaisons intermétalliques dans un domaine de concentration qui ne comprend pas nécessairement le minimum d'énergie libre du composé A_xB_y .

Les transformations à l'état solide se réalisent par le passage d'une lacune de solubilité, la recristallisation et les transformations allotropiques (changement de phases). Au cours du refroidissement, on peut observer l'apparition ou la disparition des composés intermétalliques. Ce phénomène s'opère lorsque les lacunes de solubilité augmentent avec l'abaissement des températures.

IV.1 Les aluminures de titane

- Nouvelle famille (intermétalliques) qui suscite un grand intérêt.
- Ils sont basés sur les composés Ti_3Al , $TiAl$.
- Structures ordonnées bonne stabilité à haute température, mais fragile à température ambiante.

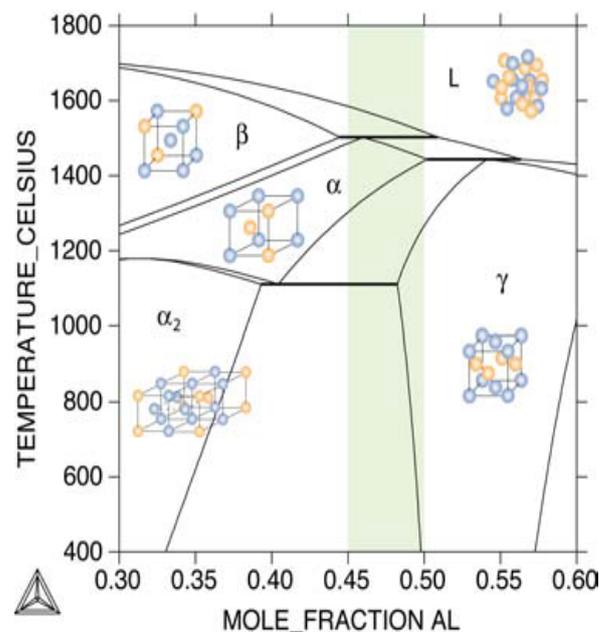


Figure (1.7) : diagramme de phase Ti-Al[12]

V. Les applications des intermétalliques

❖ Une grande partie du travail a été fait dans les 10 à 15 dernières années pour développer et caractériser les intermétalliques et de développer des technologies de traitement. En réponse à la nécessité d'une faible densité, les alliages à hautes performances pour une utilisation dans les composants de cellules et les moteurs à turbine (NRC, 1993, 1996), par exemple, des efforts concertés ont été réalisés ces dernières années pour améliorer les propriétés des alliages intermétalliques, en particulier alliages à base de aluminiums (par exemple, TiAl, Ti₃Al, NiAl, Ni₃Al, FeAl,

❖ Les intermétalliques ont donné lieu à divers développements de nouveaux matériaux. voici quelques exemples :

L'inclusion des alnico et des matériaux de stockage d'hydrogène dans les piles au nickel-métal hydrure.

❖ Ni₃Al, qui est la phase de durcissement dans les familles des superalliages à base de nickel, et des aluminiums de titane ont également l'intérêt dans les applications de pales de turbine., Propriétés et exemples

- Les matériaux magnétiques e.g. alnico; sendust; Permendur, FeCo
- Les supraconducteurs e.g. A15 phases; niobium-étain
- stockage de l'hydrogène ex. AB₅ composés (batteries nickel-métal hydrure)
- Les alliages à mémoire de forme e.g. Cu-Al-Ni (alliages de Cu₃Al et nickel); Nitinol (NiTi)
- Les matériaux de revêtement e.g. NiAl
- à haute température : Les matériaux structurelles e.g. aluminium de nickel, Ni₃Al
- Les amalgames dentaires qui sont des alliages intermétalliques de Ag₃Sn et Cu₃Sn
- contact de grille / couche barrière pour la microélectronique, par exemple TiSi₂

❖ L'applications pratiques des intermétalliques, terrestres et aérospatiales;

Les alliages intermétalliques polyphasiques à usage dentaire;

L'applications de matériaux magnétiques durs basés sur des intermétalliques, en particulier ceux comportant des métaux des terres rares.

❖ *Applications biomédicales des composés intermétalliques dérivés du TiNi*

Le composé intermétallique TiNi est très résistant à la corrosion et son comportement mécanique en déformation ressemble à celui des tissus vivants (os et tendons). Ces propriétés mécaniques si particulières se dérivent d'une transformation cristallographique qui s'opère à des températures proches de l'ambiante et qui lui confère des propriétés d'alliage à mémoire de forme. De ces faits, le composé TiNi est très utilisé dans les applications biomédicales : orthopédie, orthodontie, guide de cathéter, microchirurgie.... Or, lors d'un séjour prolongé dans le milieu humain des soucis liés à la stabilité chimique et mécanique du TiNi apparaissent. En particulière, la réactivité du TiNi avec l'hydrogène provenant du milieu aqueux joue un rôle néfaste. Lors du temps de séjour des implants de TiNi dans le corps humain, l'hydrogène peut se dissoudre dans le composé, affaiblir sa tenue mécanique et perturber son comportement d'alliage à mémoire de forme. La compréhension des propriétés d'hydrogénation des composés base TiNi, qui est assez méconnue, est donc primordiale. Dans nos recherches actuelles, nous nous intéressons à la modification des propriétés d'hydrogénation du TiNi par des substitutions partielles des métaux de transition dans les sous-réseaux de Ti et/ou de Ni. Dans le cadre des applications biomédicales, nous poursuivons des substitutions qui réduisent la capacité et/ou la cinétique d'absorption en hydrogène du composé TiNi. Figure (1.8)

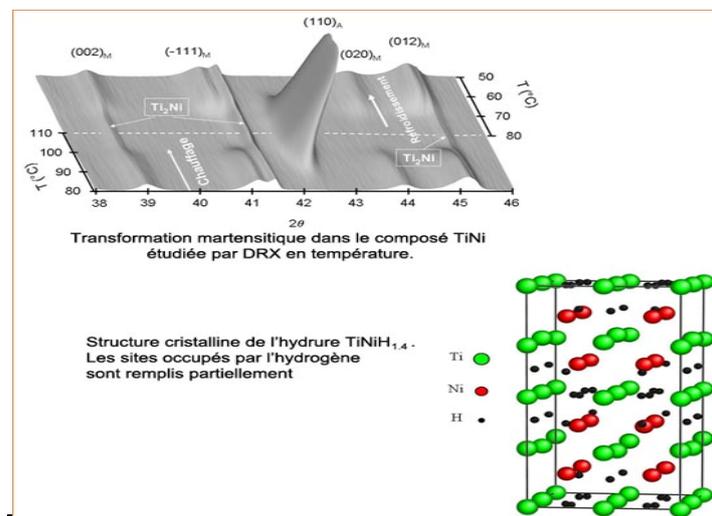


Figure (1.8) : l'hydruure $TiNiH_{1,4}$

❖ Matériaux pour l'aéronautique

Notre objectif est de développer de nouveaux matériaux et de fabriquer des pièces près des côtes pour diverses applications dans le secteur aéronautique. Nous travaillons sur des alliages intermétalliques (Siliciures au Niobium ou Aluminures de Titane) pour des aubes de turbine, des composites à matrice métallique pour des paliers lisses, et des matériaux magnétiques pour des aimants de rotors. Ces études sont développées en partenariat avec Turbomeca, SnecmaMotors, Mecachrome, Liebherr Aerospace et L'onera.[13]

Références :

- [1] F. H. Froes, C. Suryanayana, D. Eliezer, J. Mater. Sci., 27 (1992), pp. 5113.
- [2] H. A. Lipsitt, Aviation Week, 81 (1976), pp. 81.
- [3] C. T., Liu, J. Nucl. Metal., 85-86 (1979), pp. 907.
- [4] Aoki K., O. Izumi, Nippon Kinzoku Gakkaishi, 43 (1979), pp. 1190.
- [5] Chen Y. S., Material Science and Engineering, 8 (1990), pp. 30.
- [6] R. C. Westwood, Metal. Trans. B, 19 (1988), pp. 155.
- [7] *Intermetallic Compounds, Vol. 3: Progress*, edited by J. H. Westbrook and R. L. Fleischer, John Wiley & Sons, Chichester 2002, 1086 pages.
- [8] JPC_phys Mat_chapitre4b.
- [9] Performances des matériaux et leur structure « Amor Guidoum »
- [10] Philippe Robert, 'Matériaux de L'électrotechnique', Vol II, Traité D'Electricité de l'école polytechnique fédérale de Lausanne publié sous la direction de Jacques Neiryck, Editions Georgi.
- [11] Précis de métallurgie : élaboration, structures-propriétés, normalisation

Jean Barralis et Gérard Maeder.

- [12] Les alliages de titane : Elisabeth Aeby-Gautier LSG2M CNRS UMR 7584 - Ecole des Mines de Nancy.
- [13] COLLONGUES (R.). – *La non-stoechiométrie*. 230 p., bibl., Masson (1971).

CHAPITRE II

Les techniques d'exploitation des données (Datamining)



II.1 Introduction :

Le développement des capacités de stockage et les vitesses de transmission des réseaux ont conduit les utilisateurs à accumuler de plus en plus de données. Certains experts estiment que le volume des données double tous les ans.

Par ailleurs, la question de l'analyse de grands volumes de données s'est imposée. Que doit-on faire avec ces données coûteuses à collecter et à conserver ? Le datamining, est le domaine de recherche tentant de répondre à cette question en mettant au point des outils de fouille de données capables d'analyser ces grands volumes de données.

Le « datamining », est considéré aujourd'hui comme un champ à la fois scientifique et industriel et c'est devenu une discipline très en vogue. Cette émergence n'est pas le fruit du hasard mais le résultat de la combinaison de nombreux facteurs à la fois technologiques et économiques.

II.2 Définition :

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

Connue aussi sous l'expression de **fouille de données, forage de données, prospection de données, datamining**, ou encore extraction de connaissances à partir de données, « ECD » en français, « KDD » "Knowledge Discovery in Databases" en anglais.

C'est-à-dire «Les tentatives de découvrir les règles et les modèles à partir de données».

II.3 Principe :

En peu de mots, l'outil de prospection Datamining est à même 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.4 Comment on l'utilise ?

Au contraire des méthodes classiques d'analyses statistiques, Cet instrument d'analyse 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 seront captés, ordonnés et rangés au sein du Data Warehouse. Comportement des acheteurs, caractéristiques des produits, historisation de la production, désormais plus rien n'échappe à la collecte.

Avec le Datamining, ces "tera-nesque" bases de données sont exploitables.

II.5 Les techniques mises en œuvre

Différentes techniques sont proposées. Elles sont à choisir en fonction de la nature des données et du type d'étude que l'on souhaite entreprendre

- 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
- Et pour les études d'évolution de populations, les algorithmes génétiques
- Algorithmes Naïve Bayes, séries chronologiques, régression linéaire...

II.6 Méthodes globales :

Résoudre une problématique aussi complexe, que celle incluse dans le processus de datamining, contenant une énorme quantité de données impose l'utilisation de méthodes et d'algorithmes puissants, on distingue trois grandes familles d'algorithmes :

II.6.1 Les méthodes supervisées :

Ceux sont des méthodes qui expliquent et prévoient un ou plusieurs phénomènes observables. Concrètement elles vont s'intéresser à une ou plusieurs variables de la même base de données définies comme étant les cibles de l'analyse. Citons quelques techniques disponibles pour cette classe de méthodes :

- Techniques à base statistiques de régressions linéaires et non linéaires.
- Techniques à base de réseaux de neurones : perception mono ou multicouches.

- Techniques à base d'algorithme génétique.
- Techniques à base d'Inférence bayésienne (Réseau bayésien).

II.6.2 Les méthodes non-supervisées :

Dans ces méthodes, on n'accorde aucune importance particulière à aucune des données ou variables, toutes les données sont traitées de la même manière et ont le même poids, c'est-à-dire qu'aucune des variables n'est considérée comme la cible. Cette méthode servira pour réaliser la compression de l'information ou alors pour dégager un ensemble d'individus des groupes homogènes (typologie). Nous allons citer quelques algorithmes utilisés :

- Techniques classiques utilisées dans la statistique : classification ascendante hiérarchique, classifications mixtes (Birch), classifications relationnelles...
- Techniques à base de Réseau de neurones : réseau de Kohonen...

II.6.3 Les méthodes de réductions de données :

Ces méthodes permettent de réduire un ensemble de données volumineux en un ensemble épuré de ce que l'on considèrera comme de l'information non pertinente (bruit), elles sont très souvent, mais pas systématiquement utilisées en amont des techniques supervisées ou non supervisées. Les Techniques les plus utilisées sont :

- Techniques de positionnement : positionnement multidimensionnel...
- Techniques d'analyse factorielle: Analyse en composante principale, analyse des correspondances multiples...

II.7 Méthodes statistiques :

Le datamining se propose d'utiliser un ensemble de techniques issues de discipline scientifiques diverses (statistiques, intelligence artificielle, base de données) pour construire des modèles à partir de données, et extraire de ces données un maximum de connaissances utiles. Dans notre travail on a utilisé deux de ces méthodes, l'ACP (analyse en composante principale) et la PLS (Partial Least Squares régression).

II.7.1 Analyse en composante principale (ACP) :

II.7.1.1 Projeter la réalité sur un plan

Nous avons l'habitude de dessiner ou photographier la réalité. et naturellement passons d'un espace à 3 dimensions à un espace à 2 dimensions. Mais selon le point de vue, l'information retenue ne sera pas la même.

L'ACP nous propose un point de vue permettant de voir au mieux les individus d'un tableau.

II.7.1.2 Définition de l'ACP

L'analyse en composantes principales est la technique de visualisation en axes principaux la plus ancienne et la plus répandue. Conçue pour la première fois dans un cadre limité par le psychologue Karl Pearson en 1901[1], elle fut intégrée par la suite à la statistique mathématique par Harold Hotelling en 1933[2].

- L'ACP est une méthode descriptive.
- Son objectif est de représenter sous forme graphique l'essentiel de l'information contenue dans un tableau de données quantitatif.
- Dans un tableau de données à j variables, les individus se trouvent dans un espace à j dimensions.
- L'ACP nous propose un point de vue permettant de voir au mieux les individus d'un tableau.

II.7.1.3 Principes de l'ACP:

Matrices des données : n objet(ou individus), p descripteurs(ou variables) :

On applique usuellement une ACP sur un ensemble de p variables aléatoires X_1, \dots, X_p connue à partir d'un échantillon de n réalisations conjointes de ces variables. Cet échantillon de ces p variables aléatoires peut être structuré dans une matrice M à n lignes et p colonnes.

$$X = \text{individus} \dots \begin{matrix} & \begin{matrix} 1 & \dots & p \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ n \end{matrix} & \begin{bmatrix} x_1^1 & \dots & x_1^p \\ \vdots & & \vdots \\ x_n^1 & \dots & x_n^p \end{bmatrix} \end{matrix}$$

Pour avoir une image de l'ensemble des unités, on se place dans un espace affine en choisissant comme origine un vecteur particulier de R^p , par exemple le vecteur dont toutes les coordonnées sont nulles. Alors, chaque unité sera représentée par un point dans cet espace.

L'ensemble des points qui représentent les unités est appelé traditionnellement « nuage des individus ».

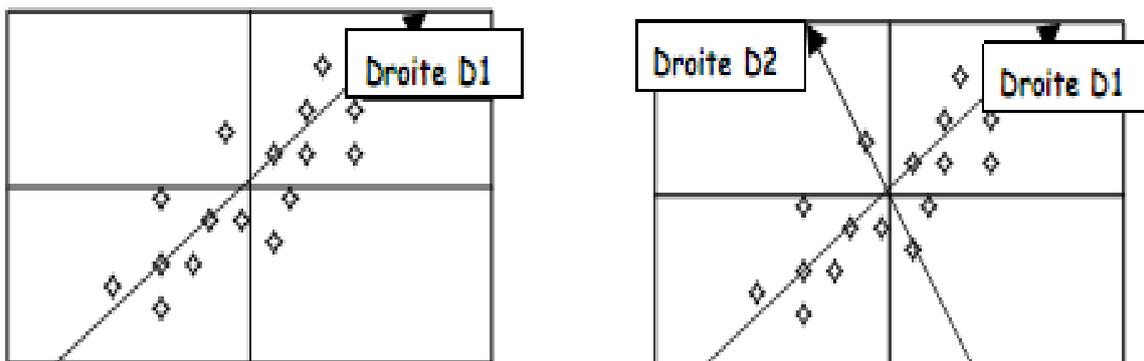
En faisant de même dans R^n , chaque variable pourra être représentée par un point de l'espace affine correspond. L'ensemble des points qui représentent les variables est appelé « nuage des variables ».

On constate, que ces espaces étant de dimension supérieure en général à 2 et même 3, on ne peut visualiser ces présentations. L'idée générale des méthodes factorielles est de trouver un système d'axes et de plans tel que les projections de ces nuages de points sur ces axes et ces plans permettent de reconstituer les positions des points les uns par rapport aux autres, c'est-à-dire avoir des images les moins déformées possible.

II.7.1.4 Résumer les données

Lorsqu'on projette les données sur un plan, on obtient un graphique déformé de la réalité.

- ❖ Le rôle de l'ACP est de trouver des espaces de dimensions plus petites minimisant ces déformations. [3]
- ❖ On utilise un espace à 2 dimensions (un plan). Ce plan est appelé le plan principal. Il est constitué de deux droites perpendiculaires.
- ❖ La méthode consiste à calculer la première droite D1 de façon à maximiser les carrés des distances de projection des points sur la droite.



II.7.1.5 Les composantes principales

- Les droites D1 et D2 sont des caractères synthétiques obtenus par des combinaisons linéaires avec les variables d'origines. [3]
- Ces droites sont appelées *composantes principales*, ou *axes principaux*.
- La première composante principale doit "capturer" le maximum d'inertie du tableau des données. La variance des individus doit être maximale.
- Il reste un résidu non expliqué par cette première composante. C'est sur ce résidu qu'est calculée la deuxième composante principale.

II.7.1.6 Caractères des composantes principales :

- La première composante principale "capture" le maximum d'inertie du tableau des données. [3]
- La deuxième composante principale est un complément, une correction de la première.
- La deuxième composante principale doit avoir une corrélation linéaire nulle avec la première (orthogonalité).
- Il n'y a pas de redondance d'information entre deux composantes principales.
- On calcule les autres composantes de la même manière.

II.7.1.7 Propriétés des composantes principales:

▪ Nombre :

Bien que l'objectif soit en général de n'utiliser qu'un petit nombre de Composantes Principales, l'ACP en construit initialement p , autant que de variables originales. Ce n'est que par la suite que l'analyste décidera du nombre de Composantes à retenir.

"Retenir k Composantes Principales" veut dire "Remplacer les observations originales par leur projections orthogonales dans le sous-espace à k dimensions défini par les k premières Composantes Principales.

▪ Orthogonalité :

Les Composantes Principales définissent des directions de l'espace des observations qui sont deux à deux **orthogonales**. Autrement dit, l'ACP procède à un changement de repère orthogonal, les directions originales étant remplacées par les Composantes Principales.

▪ Décorrélation :

Les Composantes Principales sont des variables qui s'avèrent être deux à deux décorrélées.

II.7.1.8 Transformer les données - Centrage et réduction :

- L'importance que prendront les variables dans le calcul des composantes principales est fonction de leur ordre de grandeur.
- Une variable ayant un écart-type important aura plus de poids qu'une variable de faible écart-type.
- Des variables de fort écart-type "construiront" les premières composantes.
- Les calculs ne sont pas faux, mais la lecture des résultats d'une ACP peut devenir compliquée.
- C'est pour remédier à ça qu'il convient de centrer et réduire les variables. [3]

II.7.1.9 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 [4-5-6-7]. 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) [8] 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.7.1.10 Interprétation des résultats de l'ACP

En ACP normée, les variables projetées sur chaque plan factoriel se trouvent à l'intérieur d'un cercle de rayon unité. Plus une variable est projetée vers le bord du cercle, mieux elle est représentée. Par ailleurs, deux variables bien représentées et proches l'une de l'autre sont corrélées positivement tandis que deux variables qui s'opposent sont corrélées négativement. Une orthogonalité entre deux variables traduit l'absence de corrélation linéaire. Notons que l'ACP ne mesure que des liens linéaires entre variables. Avant de conclure sur l'existence ou l'absence de relations entre variables, il est donc utile d'examiner l'allure de leurs nuages de corrélation.[11]

Le diagramme en bas à droite de la figure 1 indique le pourcentage de l'inertie totale du nuage des variables porté par chaque axe factoriel. Ce pourcentage est respectivement de 66%, 10% et 5% pour les trois premiers axes qui sont retenus pour la représentation graphique. La première composante principale sépare les variables de concentration (concentration moyenne, minimale et maximale et concentrations des quatorze semaines) de la densité de population. Compte tenu des différences de concentration entre ville et campagne, cette opposition était prévisible. On observe de plus une corrélation positive entre la concentration d'ozone et l'altitude. Aucune corrélation n'apparaît avec les coordonnées de l'espace.

Les autres composantes sont essentiellement liées à la position géographique des tubes passifs, représentée par les coordonnées X et Y, et à la densité de population. Cela se vérifie aisément en examinant le plan F2-F3 dans lequel toutes les variables de concentration sont projetées vers le centre du cercle.

Notons que le nuage des variables de concentration associées aux quatorze semaines de mesure est dispersé autour de la concentration moyenne. Quelques périodes se démarquent, telle la semaine 8 qui correspond à une chute des concentrations liée à une diminution brutale de température.

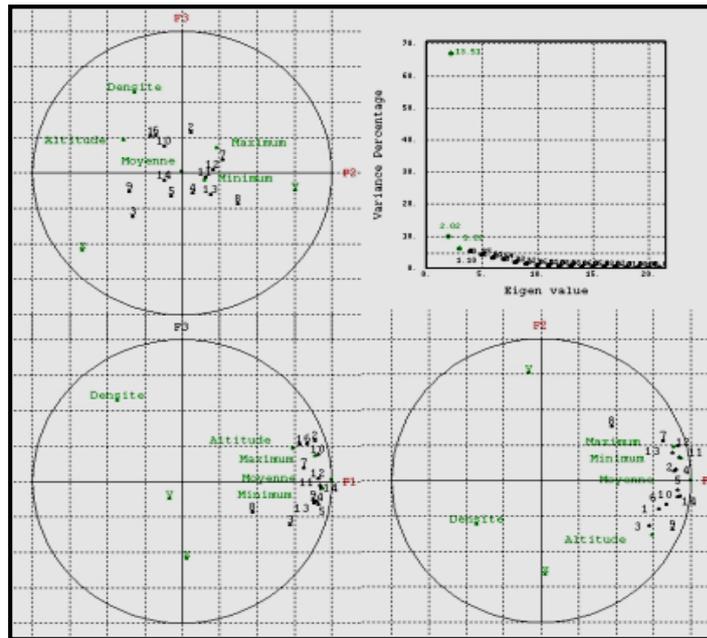
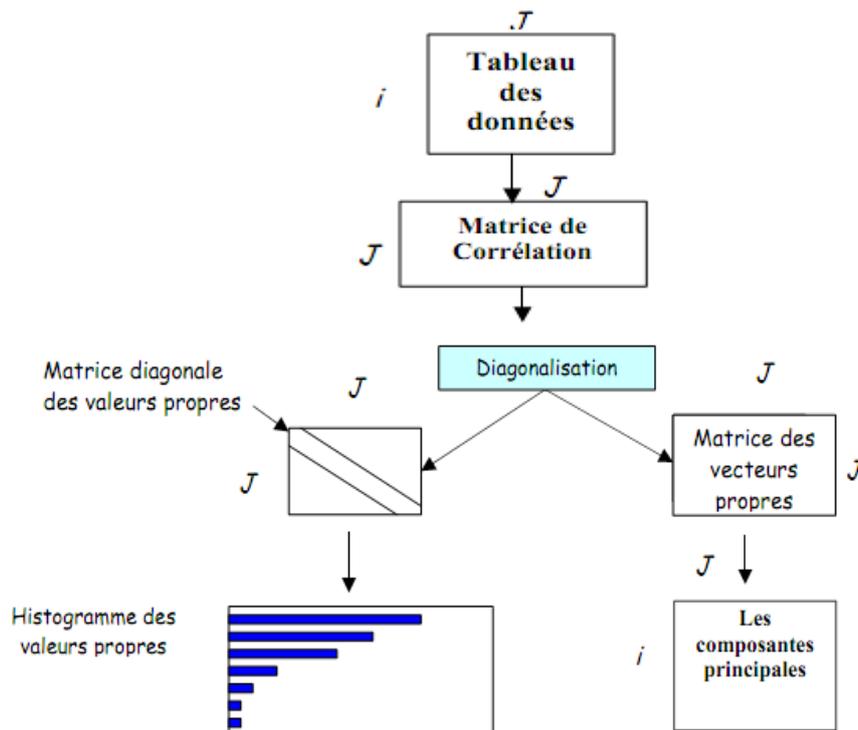


Figure 1 : Exemple d'ACP. Données d'ATMO Auvergne : 14 semaines consécutives de mesure de l'ozone (été 2001) et variables auxiliaires. Projection du nuage des variables sur les deux premiers plans factoriels [11]

II.7.1.11 Applications de l'Analyse en Composantes Principales :

L'utilisation la plus commune de l'ACP est de fournir des données décrites par un grand nombre de variables quantitatives des représentations **planes** (et donc interprétables visuellement) aussi fidèles que possible. Pour cela, on projette ces données sur des **plans factoriels**, chaque plan étant défini par une paire de Composantes Principales prises parmi les premières CP.

II.7.1.12 L'ACP en résumé [3]



II.7.1.13 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 [4].

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.7.2 La régression (PLS) :

La régression PLS (Partial Least Squares régression) est une technique d'analyse et de prédiction relativement récente. Elle a été conçue pour faire face aux problèmes résultants de l'insuffisance de l'utilisation de la régression linéaire classique, qui trouve ses limites dès lors que l'on cherche à modéliser des relations entre des variables pour lesquelles il y a peu d'individus, ou beaucoup de variables explicatives en comparaison au nombre d'individus (le nombre de variables explicatives pouvant excéder très largement le nombre d'individus), ou encore lorsque les variables explicatives sont fortement corrélées entre elles.

Cette méthode fut proposée par Wold S., Albano C., Dunn III W.J., Esbensen K., Hellberg S., Johansson E. et Sjöström M. en 1983 [9]. Elle fut principalement développée par Svante Wold, qui a dirigé de nombreuses thèses portant sur le sujet. De plus, ce dernier, associé à Nouna Kettaneh-Wold et à plusieurs collaborateurs, ont développé le logiciel d'analyse des données SIMCA-P, logiciel centré sur la régression PLS.

II.7.2.1 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 :

$$Y = b_0 + b_1X_1 + b_2X_2 \dots \dots + b_pX_p$$

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.7.2.2 Interprétation du modèle PLS :

Le modèle PLS, nous permet de former de nouvelles variables x, t_k , qui sont des combinaisons linéaires des x et qui sont utilisées comme prédicteurs de Y . Ainsi PLS est basé sur un modèle linéaire.

Tous les paramètres, t , u , w , et q sont déterminés par un algorithme PLS. Les scores t et u contiennent des informations sur les objets et les corrélations/ non corrélations entre ces différents objets [10].

Les poids w et q donnent des informations sur la manière dont la quelle les différentes variables se combinent pour former des relations quantitatives entre X et Y , et par conséquent fournir une interprétation des scores t et u . Ces poids w expriment à la fois les corrélations positives entre X et Y et les corrélations nécessaires pour prédire Y à partir de X .

La régression PLS permet ainsi de conserver toutes les variables explicatives tout en obtenant une équation de régression cohérente.

II.7.2.3 La régression sur composantes principales :

La PLS est une technique quantitative de décomposition spectrale étroitement liée à la régression sur composantes principale (PCR). Cependant, dans la PLS, la décomposition est faite d'une manière légèrement différente.

Au lieu de décomposer d'abord la matrice spectrale X en un ensemble de vecteurs propres et de scores, et de les régresser contre les Y dans une étape séparée, la PLS utilise l'information de Y en même temps que le processus de décomposition.

Ceci implique que les variables expliquant le mieux les Y seront plus fortement pondérées. De plus, les vecteurs propres et les scores calculés en utilisant la PLS seront différents de ceux de PCR. L'idée principale de la PLS est de donner le plus d'information possible sur Y dans les premiers vecteurs construits.

II.7.2.4 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).

- Simplicité de son algorithme.
- Ni inversion, ni diagonalisation de matrices, mais seulement une succession de régressions simples, autrement dit des calculs de produits scalaires.
- On peut donc traiter de très grands ensemble de donnés.
- La régression PLS donne en pratique d'excellente prévision, même dans le cas d'un petit d'observations et d'un grand nombre de variables.

II.7.2.5 Conclusion

La régression PLS permet de rapprocher le statisticien du chercheur dans les problèmes de modélisation. En générale ce dernier souhaite en effet conserver dans son modèle toutes les variables importantes tout en obtenant des équations de régression cohérentes. Lorsque, en régression multiple il y a multi colinéarité et / ou un nombre important de variables explicatives par rapport au nombre d'observation, la solution la plus courante consiste à exclure des variables explicatives par des méthodes pas à pas. La régression PLS permet dans ces situations de conserver toutes les variables explicatives tout en obtenant une équation de régression cohérente.

Référence:

[1] Pearson K. (1901):*Phil. Mag.* 2, n°11, 559-572.

“On lines and planes of closest fit to systems of points in space”

[2] Hotelling H. (1933):”*J. Educ. Psy.* 24, 417-441, 498-520.

“Analysis of a complex of statistical variables into principal components”.

[3] Jean Thioulouse, Anne-Beatrice Dufour and Daniel Chessel (2004). ade4:

Analysis of Environmental Data: Exploratory and Euclidean methods in

Environnementale sciences. R package version 1.3-3.

[4] 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.

[5] Ferre L., "Selection of components in principal component analysis: A comparison of methods", *Computational Statistics and Data Analysis*, pp. 669-682, 1995.

[6] Himes D.M, and al., "Determination of the number of principal component for disturbance detection and isolation", *Proceedings of ACC*, Baltimore, 1994.

[7] Valle S., and al., "Selection of the number of principal components: The variance of the reconstruction error criterion with a comparison to other methods", *Industrial & Engineering Chemistry Research*, vol. 38, pp. 4389-4401, 1999.

[8] Qin S. J. and Dunia R., “Determining the number of principal components for best.

[9] S. Wold, M. Sjostrom, L. Eriksson, "**PLS-regression: a basic tool of chemometrics** ", *Chemometrics and Intelligent Laboratory Systems*, 58 (2001) 109-130.

[10] A. Hoskuldsson, "**Variable and subset selection in PLS regression**", *Chemometrics and Intelligent Laboratory Systems* 55 (2001) 23–38.

[11] **INERIS** DRC-03-45597-LRI-LMa –LCSQA-n°140_v1.

CHAPITRE III
Résultats et discussions



Chapitre III

Résultats et discussions

III.1- Introduction

Dans ce chapitre nous présentons, les résultats liés à l'utilisation de l'Analyse en composantes principales (ACP) et la régression (PLS), appliquée sur un ensemble des composés intermétalliques. Le but étant de détecter les tendances et les corrélations qui existent entre leurs propriétés structurales et thermodynamiques, en se basant sur une analyse d'une base de données expérimentales et théoriques disponible sur ces matériaux.

Les calculs ont été faits avec ces deux techniques en utilisant le logiciel XLSTAT décrit en annexe.

III.2- Description de la méthode d'analyse de composantes principales :

L'ACP se réalise sur une population donnée. Dans notre étude « les composés intermétalliques ». La ACP est un outil puissant pour l'analyse des matériaux et a été employé pour résoudre des problèmes en science des matériaux. La ACP est efficace pour découverte des informations de base qui peuvent être perdues quand on a une grande quantité de données à exploiter.

III.3- Résultats :

III.3.1- Définition d'une base de données :

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

III.3.2- Base de données 1 :

Le tableau [III-1] représente la base de **donnée1** il contient l'ensemble de données utilisées. Dans notre cas les individus statiques constituant cette population sont les 29 composés intermétalliques, et nous avons retenus comme variables : le volume molaire (V), la densité moyenne d'électrons (ρ_{ws}), le module de compressibilité (B1). (Tableau III-1).

nombres	Composés	V	η_{ws}	B ₁
1	Al ₄₈ Co ₅ Ru ₄₇	8,917	4,304	165,2
2	Al ₅₀ Co _{1.5} Ru _{48.5}	9,026	4,256	163,5
3	Al ₅₀ Co ₅ Ru ₄₅	8,972	4,226	160,2
4	Al ₅₂ Co _{1.5} Ru _{46.5}	9,08	4,18	158,6
5	Al ₅₂ Co ₅ Ru ₄₃	9,028	4,148	155,4
6	Co _{78.97} Al _{14.48} Ni _{6.55}	6,962	4,968	171,8
7	Co _{80.75} Al _{12.68} Ni _{6.57}	6,926	5,017	174,3
8	Co _{82.84} Al _{12.5} Ni _{4.66}	6,92	5,025	174,7
9	Co _{82.85} Al _{10.58} Ni _{6.57}	6,885	5,074	177,3
10	Co _{83.78} Al _{13.3} Ni _{2.92}	6,932	5,007	173,8
11	Co _{84.89} Al _{10.39} Ni _{4.72}	6,88	5,081	177,6
12	Co _{87.12} Al _{8.29} Ni _{4.59}	6,841	5,138	180,6
13	Co ₆ Ru ₄₁ Ta ₅₃	9,627	4,942	235,1
14	Co ₁₅ Ru ₃₅ Ta ₅₀	9,368	4,951	229,7
15	Co ₂₀ Ru ₃₀ Ta ₅₀	9,28	4,921	224,7
16	Co ₂₆ Ru ₂₄ Ta ₅₀	9,223	4,845	216,5
17	Cu _{2.7} AlMn _{0.3}	7,777	3,839	114,6
18	Cu _{2.8} AlMn _{0.2}	7,766	3,861	115,8
19	Cu _{1.29} Al _{0.52} Ni _{0.07}	7,818	3,873	117,3
20	Cu _{1.30} Al _{0.54} Ni _{0.05}	7,841	3,853	116,4
21	Cu ₂₅ Ni ₂₅ Zn ₅₀	7,955	3,681	107,8
22	Cu ₃₀ Ni ₂₀ Zn ₅₀	8,03	3,622	105,3
23	Cu ₃₅ Ni ₁₅ Zn ₅₀	8,091	3,57	103,1
24	Cu ₄₀ Ni ₁₀ Zn ₅₀	8,135	3,525	101,1
25	Cu ₄₀ Ni ₅ Zn ₅₀	8,162	3,489	99,3
26	Cu _{66.5} Zn _{20.8} Al _{12.7}	7,796	3,785	111,7
27	Cu _{67.7} Zn _{19.4} Al _{12.9}	7,776	3,805	112,6
28	Cu _{68.7} Al ₁₇ Zn _{14.3}	7,801	3,809	113,2
29	Cu ₂ MnAl	7,852	3,685	106,6

Tableau (III-1) : Base de données 1 utilisée dans l'analyse en composantes principales [1], [2], [3], [4], [5], [6], [7]

III.3.3- Techniques de calculs

III.3.3.1- L'analyse en composantes principales (ACP):

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 renseignements 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).

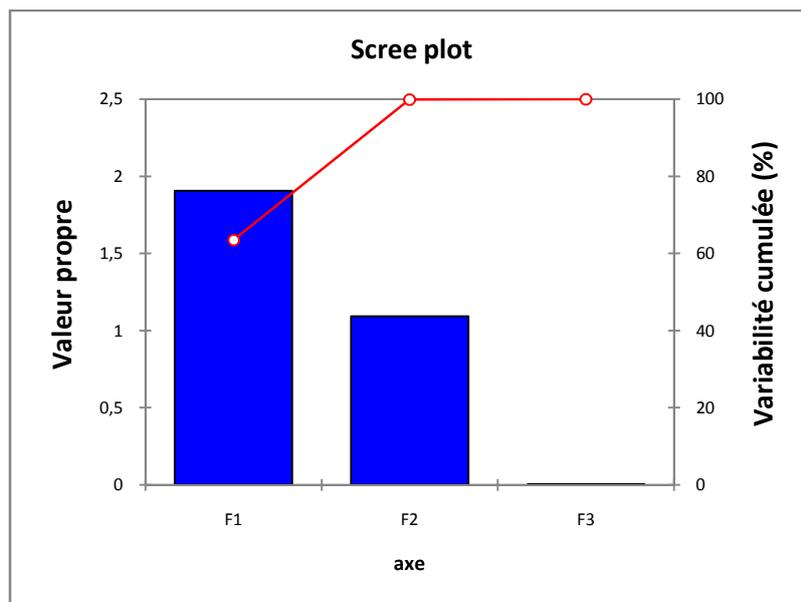


Figure III.1 : Représentation graphique des axes

L'analyse en composantes principales nous déduit les axes principaux, dans notre cas la première composante principale (PC1) est le vecteur propre qui correspond à la plus grande valeur propre, et (PC2) est le vecteur propre correspondant à la deuxième plus grande valeur propre. Par définition Les vecteurs propres sont les coefficients à affecter aux variables initiales pour obtenir les composantes principales.

Matrice de corrélation

La matrice de corrélation est un résultat intéressant en composantes principales, elle permet d'analyser les relations bilatérales existant entre les différentes variables retenues.

Le tableau suivant correspond à la matrice de corrélation des composés intermétalliques :

Variables	V	ws	B1
V	1	-0,153	0,305
ws	-0,153	1	0,891
B1	0,305	0,891	1

Tableau [III.2] : Matrice de corrélation des composés intermétalliques ternaires

A partir du tableau [III.2] on remarque que le module de compressibilité B1 et la densité moyenne d'électrons η_{ws} , sont fortement corrélés avec la valeur (**0.891**).

L'analyse des résultats de l'ACP pour les composés intermétalliques, 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).

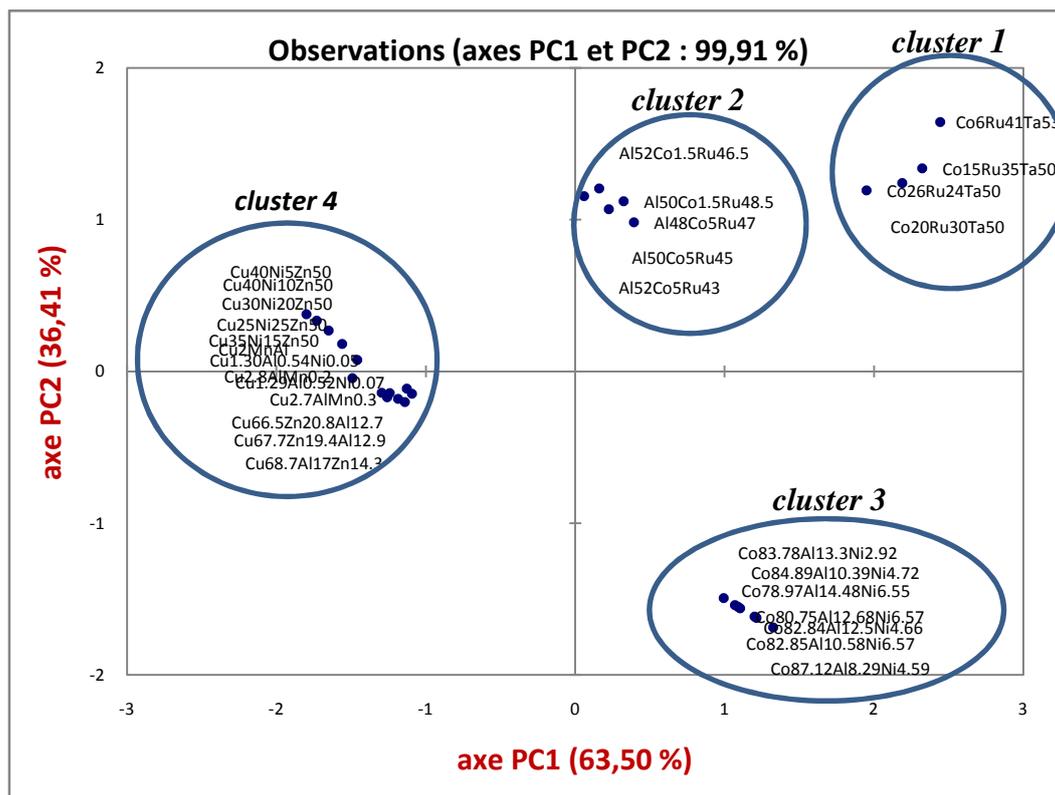


Figure (III.2): Graphe des observations pour 29 composés intermétalliques (Scores plot)

On remarque que l'axe PC1 capture 63,50% de la variance pour l'ensemble des données, et PC2 capture 36,41% de la variance comme nous le voyons dans la figure [III.2]. Aucune autre composante principale PCs n'est incluse dans ces discussions parce qu'elles ne fournissent pas des informations significatives.

En premier vue, il s'avère que l'ensemble des composés se séparent nettement en quatre groupes (clusters).

Ceux avec un PC1 positif : cluster1, cluster2, cluster3. ces trois clusters correspondant à des intermétalliques à base de CoRuTa et AlCoRu et CoAlNi respectivement. Et ceux avec un PC1 négatif : cluster4 il contient des intermétalliques a base de Cu : [CuNiZn, CuAlZn, CuZnAl, CuAlMn, CuAlNi, CuMnAl].

Nous remarquons que ces 4 clusters à travers PC1 se regroupe par classe ayant a peut près un même module de compressibilité (bulk modulus) et une même densité électrique à cette effet plus le PC1 augmente plus le bulk augmente et la densité augmente. Nous remarquons le PC2 par contre donne la variation de la concentration du deuxième élément dans l'intermétallique car le PC1 donne la concentration du premier élément de

l'intermétallique. Par conséquent les propriétés intéressantes dans les trois clusters, un volume molaire moyenne et un grand module de compressibilité B1 et une grande densité d'électrons η_{ws} . Et pour le quatrième cluster, les trois propriétés B1, η_{ws} et V sont très petites.

La lecture pour chacun des facteurs retenus des corrélations avec les trois variables permet ensuite de déterminer leur signification concrète.

Le graphe qui représente les variables ou les descripteurs (loadings plot) correspond à l'axe qui représente les observations (scores plot), dans notre cas les observations sont les composés intermétalliques. La figure (III-3) représente la disposition des variables correspondants aux observations représentées sur la figure (III-2).

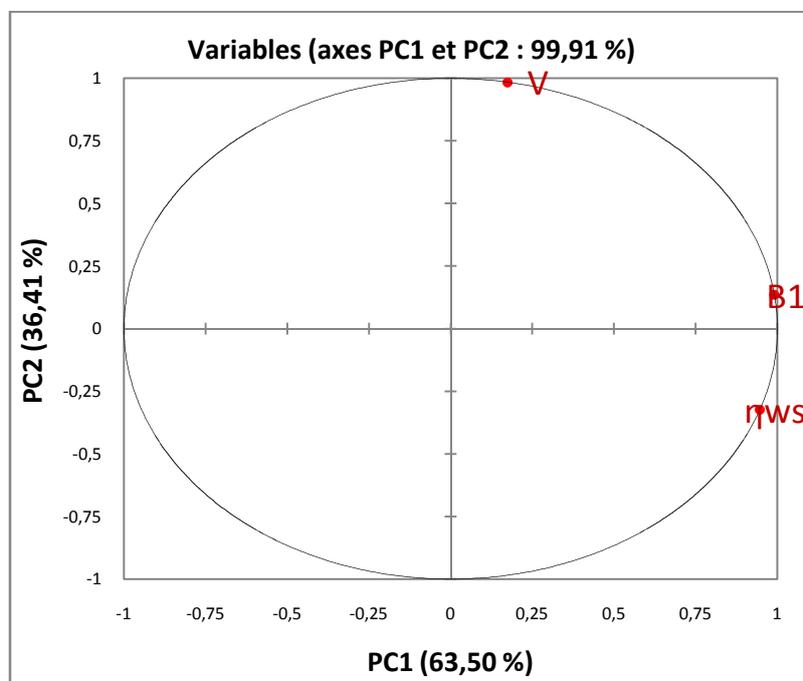


Figure (III-3) : Graphe des variable (loadings plot)

Pour confirmer le fait qu'une variable est fortement liée à une composante principale, il suffit de consulter la table suivante des cosinus carrés :

	PC1	PC2	PC3
V	0,030	0,969	0,000
η_{ws}	0,894	0,104	0,001
B1	0,980	0,018	0,001

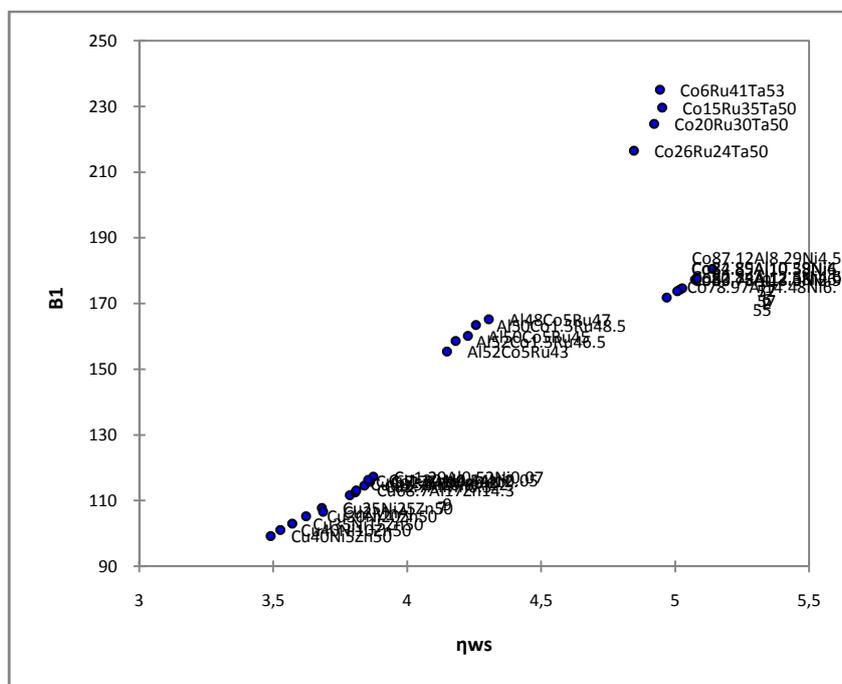
Tableau [III.3] : Cosinus carrés des variables.

Plus le cosinus est élevé (en valeur absolue), plus la variable est liée à l'axe. Plus le cosinus est proche de zéro, moins la variable est liée à l'axe. Dans notre cas, l'axe PC1 est clairement lié à B1 et η_{ws} , avec des cosinus **0,980** et **0,894** respectivement, et l'axe PC2 est fortement lié à la variable V avec des cosinus **0,969**.

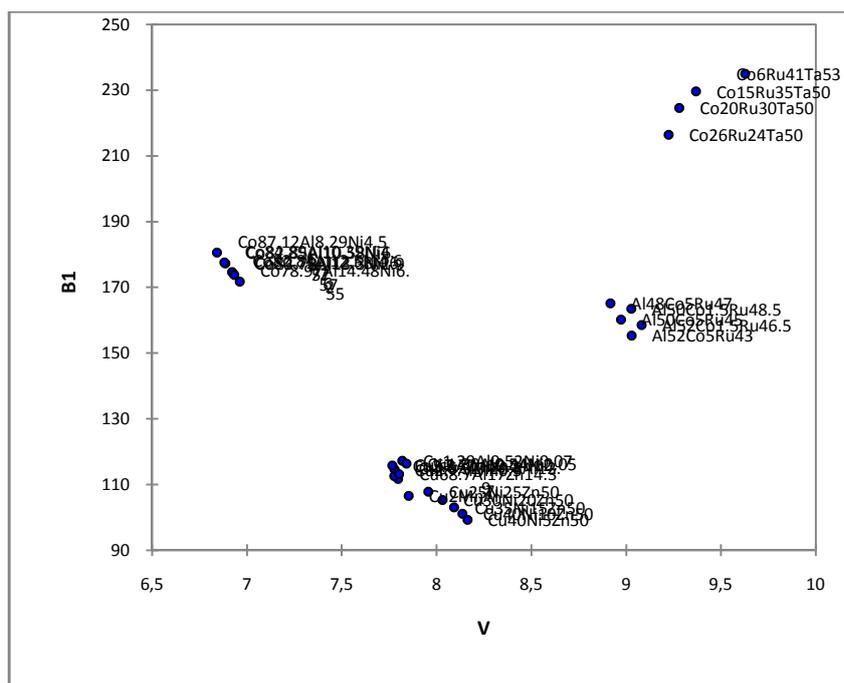
Les vecteurs propres sont affichés dans un cercle de rayon 1. On va regarder pour chaque axe principale quelles sont les variables qui se suivent et s'opposent. Par ailleurs, deux variables bien représentées et proches l'une de l'autre sont corrélées positivement tandis que deux variables qui s'opposent sont corrélées négativement. Une variable d'importance au plus elle sera représentée proche de cercle. Dans le cas contraire, une variable n'expliquant pas un axe se trouve vers le centre.

On remarque que l'axe PC1 est clairement lié positivement aux deux variables B1 et η_{ws} , et l'axe PC2 est fortement liée positivement à la variable V c'est-à-dire l'axe PC2 suit la variation de V, les propriétés dans la même direction de variation sont fortement corrélées entre eux, on peut dire que le module de compressibilité est bien corrélé avec la densité moyenne d'électrons. Cela confirme que les matériaux avec une grande densité sont compressibles. On peut constater aussi que les composés avec un grand volume molaire sont moins compressibles.

Cette tendance est montrée dans les deux figures (III.4), (III.5) qui représentent les variations de B1 en fonction de η_{ws} et en fonction de V.

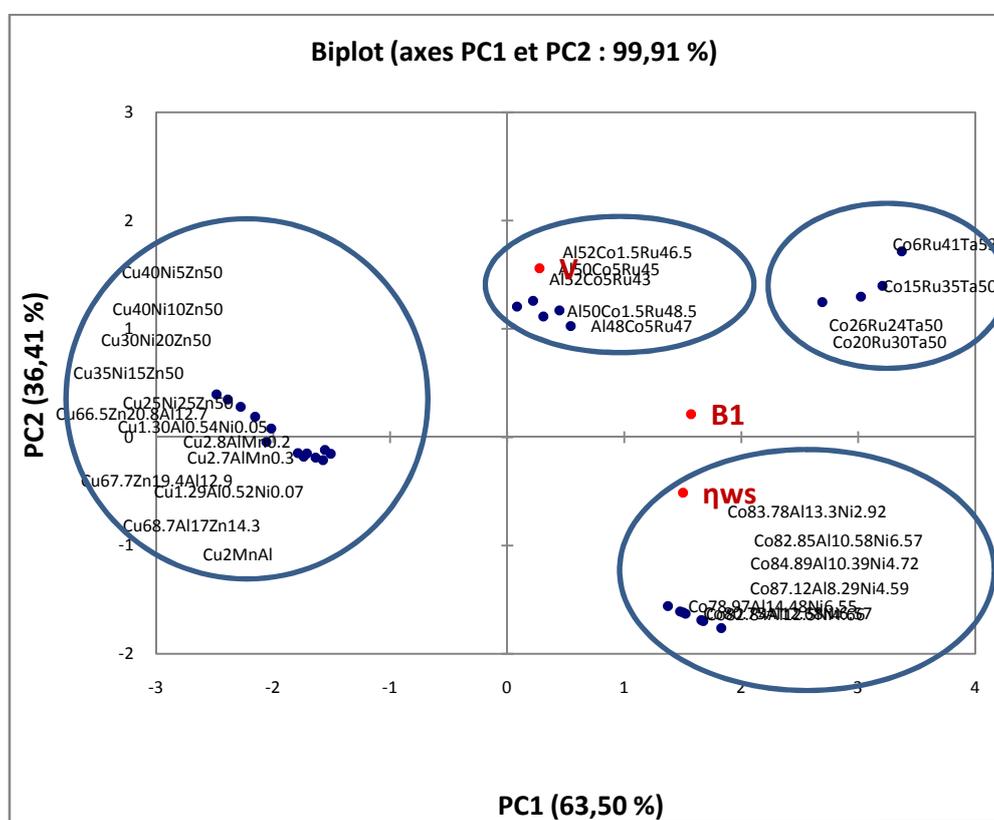


Figure(III.4) les variations de B1 en fonction de η_{ws}



Figure(III.4) les variations de B1 en fonction de V

Les résultats précédents peuvent être confirmés en regardant le graphe suivant des "Biplot" superposant les échantillons et les propriétés :



Figure(III.6) Graphe superposant les observations et les variables (bibplot)

En se basant sur le graphe (III.6) on peut déduire que le module de compressibilité est lié à la variation de la densité moyenne des électrons, ainsi on allant vers les PC1 positifs, les deux propriétés augmentent de façon proportionnelle et avec un volume molaire moyenne, on peut voir cette tendance pour cluster1, cluster2 et cluster3. Par contre quand allant vers les PC1 négatifs les trois propriétés diminuent, par exemple le cluster4.

III.3.3.2- Le modèle PLS (Partial Least Square)

1- Le graphique des corrélations :

L'un des premiers résultats du PLS c'est le graphe des corrélations (Figure [III.7]) ou plus précisément le cercle de corrélations, qui traduit bien les corrélations à chaque groupe de variables X et Y. Ce graphe des observations X (X- Scores) nous montre bien les 29 intermétalliques regroupés selon leurs propriétés similaires formant quatre clusters.

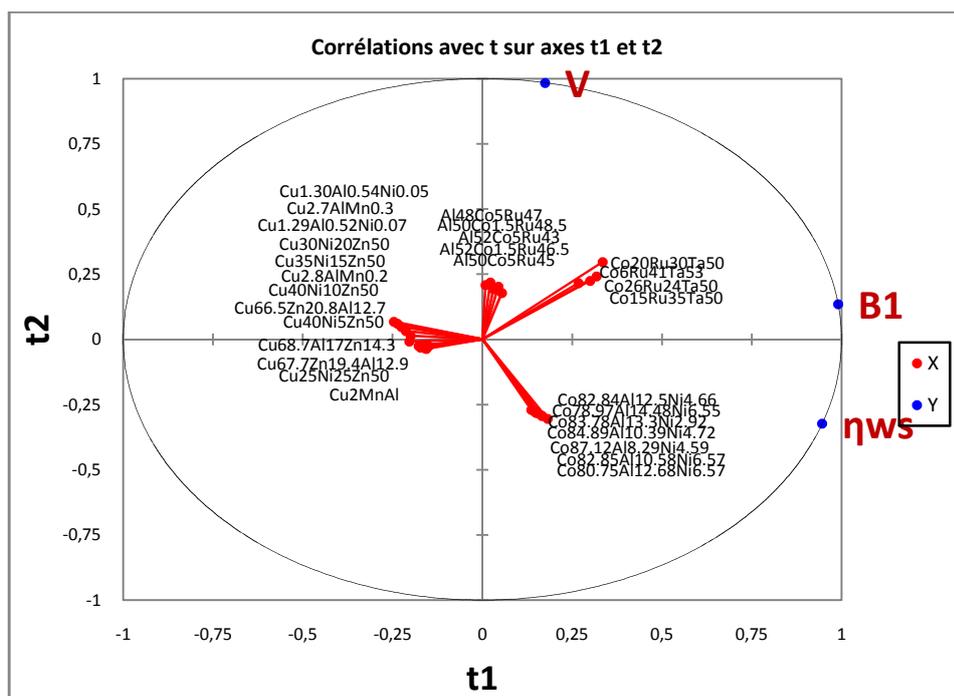


Figure (III.7) : Graphe des observations (t1, t2) des Composés intermétalliques.

On note que les variables dépendantes sont bien représentatives dans le graphique, elles sont globalement bien corrélées avec les variables explicatives (les intermétalliques).

2- Les résidus et prédictions

✓ Variable B1 :

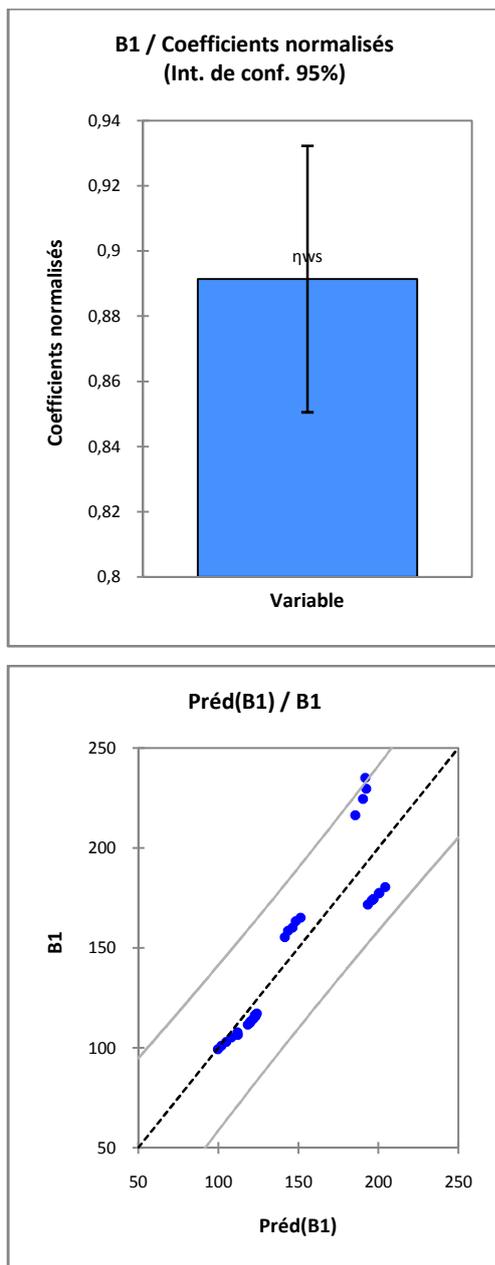


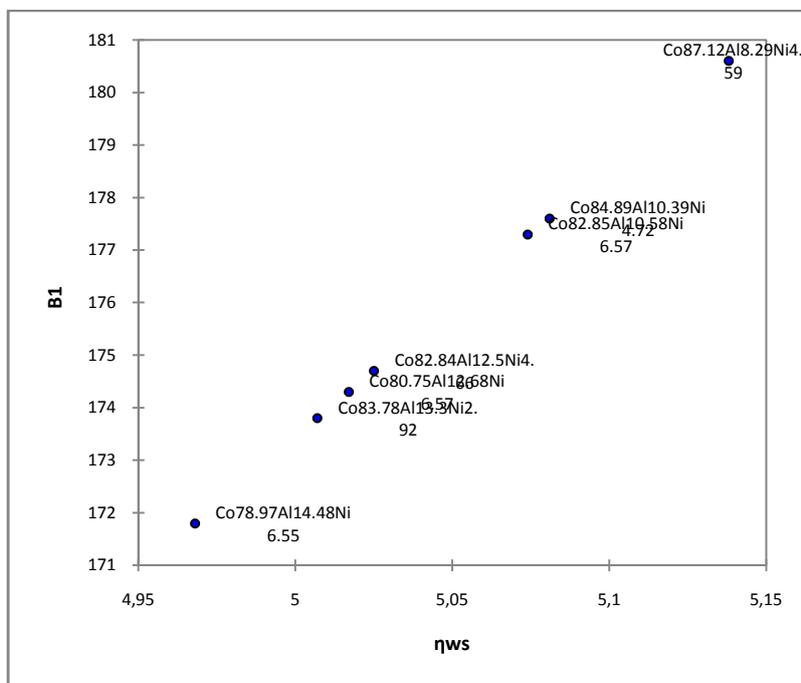
Figure [III.8] : graphe de résidus et Prédiction correspond à la variable B1

Observation	Préd(B1)	Résidu
Al48Co5Ru47	151,210	13,990
Al50Co1.5Ru48.5	148,169	15,331
Al50Co5Ru45	146,268	13,932
Al52Co1.5Ru46.5	143,354	15,246
Al52Co5Ru43	141,327	14,073
Co78.97Al14.48Ni6.55	193,276	-21,476
Co80.75Al12.68Ni6.57	196,380	-22,080
Co82.84Al12.5Ni4.66	196,887	-22,187
Co82.85Al10.58Ni6.57	199,991	-22,691
Co83.78Al13.3Ni2.92	195,747	-21,947
Co84.89Al10.39Ni4.72	200,435	-22,835
Co87.12Al8.29Ni4.59	204,046	-23,446
Co6Ru41Ta53	191,629	43,471
Co15Ru35Ta50	192,199	37,501
Co20Ru30Ta50	190,298	34,402
Co26Ru24Ta50	185,483	31,017
Cu2.7AlMn0.3	121,751	-7,151
Cu2.8AlMn0.2	123,145	-7,345
Cu1.29Al0.52Ni0.07	123,905	-6,605
Cu1.30Al0.54Ni0.05	122,638	-6,238
Cu25Ni25Zn50	111,741	-3,941
Cu30Ni20Zn50	108,004	-2,704
Cu35Ni15Zn50	104,709	-1,609
Cu40Ni10Zn50	101,858	-0,758
Cu40Ni5Zn50	99,578	-0,278
Cu66.5Zn20.8Al12.7	118,330	-6,630
Cu67.7Zn19.4Al12.9	119,597	-6,997
Cu68.7Al17Zn14.3	119,851	-6,651
Cu2MnAl	111,995	-5,395

tableau (III.1) : Coefficients normalisés (variables B1)

- D’après le tableau des coefficients de la variable B1, nous remarquons que les résidus ont des grandes valeurs pour tous les éléments, donc on peut dire que le modèle n’est pas bon pour tous l’ensemble des intermétalliques. Ce résultat peut être confirmé par des graphes correspondant aux coefficients normalisés ainsi que des graphes des résidus et prédictions (figure [III.8]). Donc on ne peut pas prédire une relation entre les deux propriétés B1 et η_{ws} , pour tous l’ensemble des intermétalliques. Mais on peut trouver des simples relations pour chaque cluster.

- **Le modèle PLS Pour(le cluste) à base de CoAlNi**



Figure(III.9) la variation de B1 on fonction de η_{ws}

On peut remarquer a partir la figure(III.9) que la variation de B1 on fonction de η_{ws} est positivement linéaire pour les intermétalliques à base de CoAlNi. B1 augmente avec l’augmentation de η_{ws} .

➤ Donc on peut prédire une simple relation entre B1 et η_{ws} .

❖ **Paramètres et équations du modèle :**

Le tableau suivant représente les paramètres des modèles correspondant aux deux variables B1 et η_{ws} .

Variable	B1
Constante	-85,712
η_{ws}	51,829

Tableau (III.2) : paramètre du modèle

❖ Equation du modèle :

$$B1 = -85,7116121025879 + 51,8289800260016 * \eta_{ws}$$

❖ Les résidus et prédictions :

Observation	Préd(B1)	Résidu
Co78.97Al14.48Ni6.55	171,775	0,025
Co80.75Al12.68Ni6.57	174,314	-0,014
Co82.84Al12.5Ni4.66	174,729	-0,029
Co82.85Al10.58Ni6.57	177,269	0,031
Co83.78Al13.3Ni2.92	173,796	0,004
Co84.89Al10.39Ni4.72	177,631	-0,031
Co87.12Al8.29Ni4.59	180,586	0,014

Tableau (III.3) : Coefficients normalisés (variables B1)

D’après le tableau des coefficients de la variable B1, nous remarquons que les résidus sont presque nuls pour tous les éléments, ce qui nous fait dire que le modèle est bon.

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 le modèle avec 2 composantes et ceci pour la variable B1.

✓ Variable B1 :

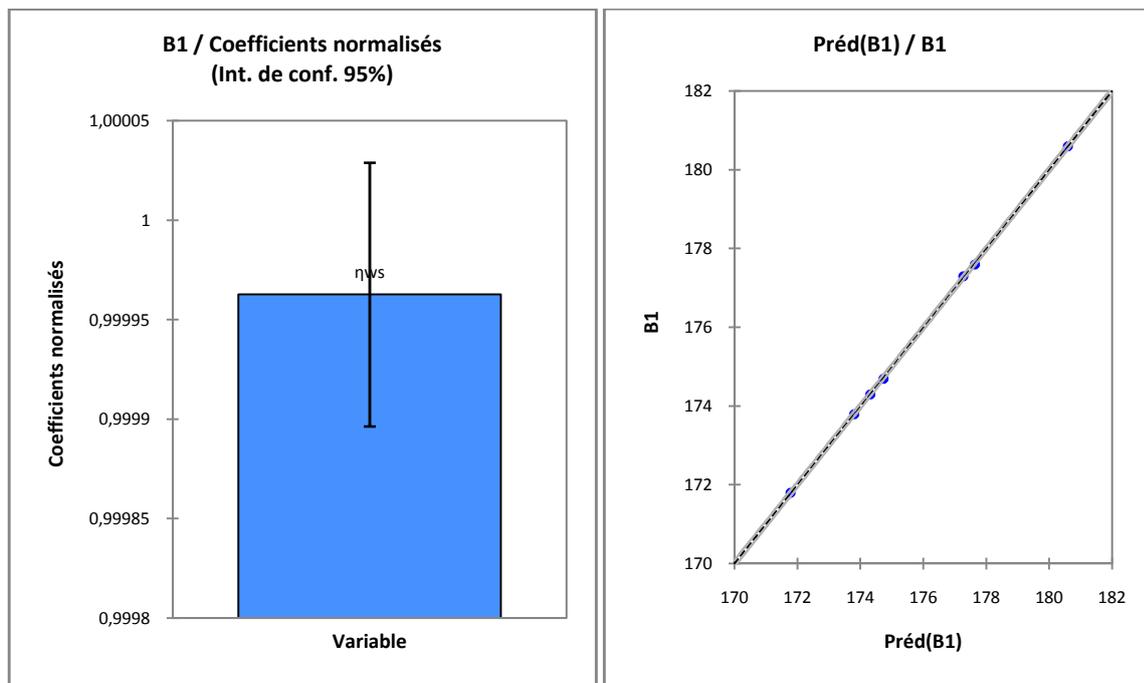
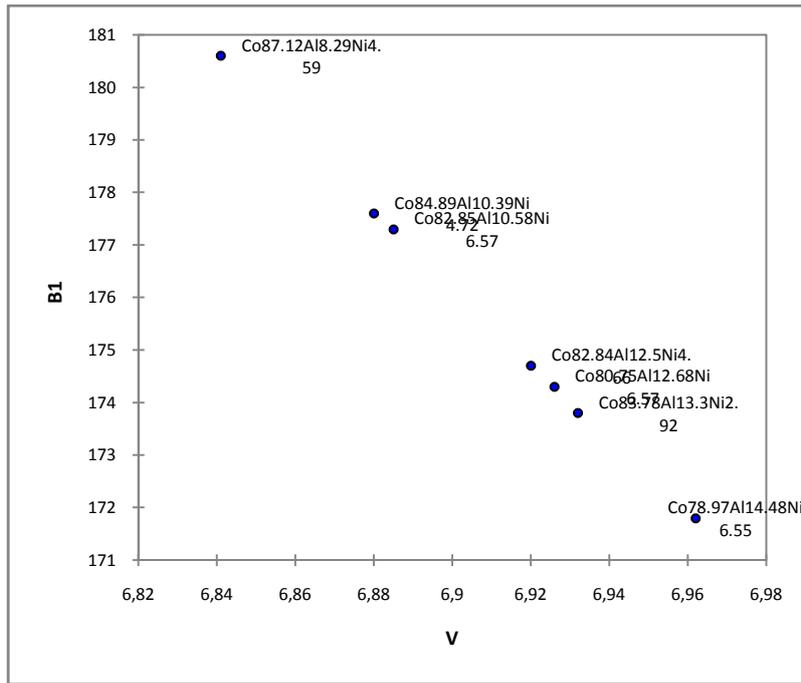


Figure [III.10] : graphe de résidus et Prédiction correspond à la variable B1

L'analyse du modèle correspondant à la variable B1 nous permet de conclure que L'ensembles des intermétalliques a base de CoAlNi, sont significatives et influencent le modèle. Nous remarquons aussi dans le graphe des prédictions que la variation est linéaire entre les valeurs de B1 prédites en fonction de celle utilisée en entrée, ce qui indique bien la bonne qualité du modèle.

- On peut aussi prédire une relation simple entre les deux variables B1 et V pour les intermétalliques à base de CoAlNi.



Figure(III.11) la variation de B1 on fonction de V

On se basant sur le graphe dans la figure (III.11). On remarque que la variation de B1 en fonction de V est linéaire, on voit aussi que les deux propriétés B1 et V sont inversement proportionnelle. Donc on peut prédire une relation entre le bulk modulus et le volume molaire.

❖ Paramètres et équations du modèle :

Le tableau suivant représente les paramètres des modèles correspondant aux deux variables B1 et V.

Variable	B1
Constante	679,789
V	-72,983

Tableau (III.4) : paramètre du modèle

❖ Equation du modèle :

$$B1 = 679,788644957059 - 72,9826772576721 * V$$

❖ Les résidus et prédictions :

Observation	Préd(B1)	Résidu
Co78.97Al14.48Ni6.55	171,683	0,117
Co80.75Al12.68Ni6.57	174,311	-0,011
Co82.84Al12.5Ni4.66	174,749	-0,049
Co82.85Al10.58Ni6.57	177,303	-0,003
Co83.78Al13.3Ni2.92	173,873	-0,073
Co84.89Al10.39Ni4.72	177,668	-0,068
Co87.12Al8.29Ni4.59	180,514	0,086

Tableau (III.5) : Coefficients normalisés (variables B1)

D’après le tableau des coefficients de la variable B1, nous remarquons que les résidus sont presque nuls pour tous les éléments, ce qui nous fait dire que le modèle est bon.

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 le modèle avec 2 composantes et ceci pour la variable B1.

✓ Variable B1 :

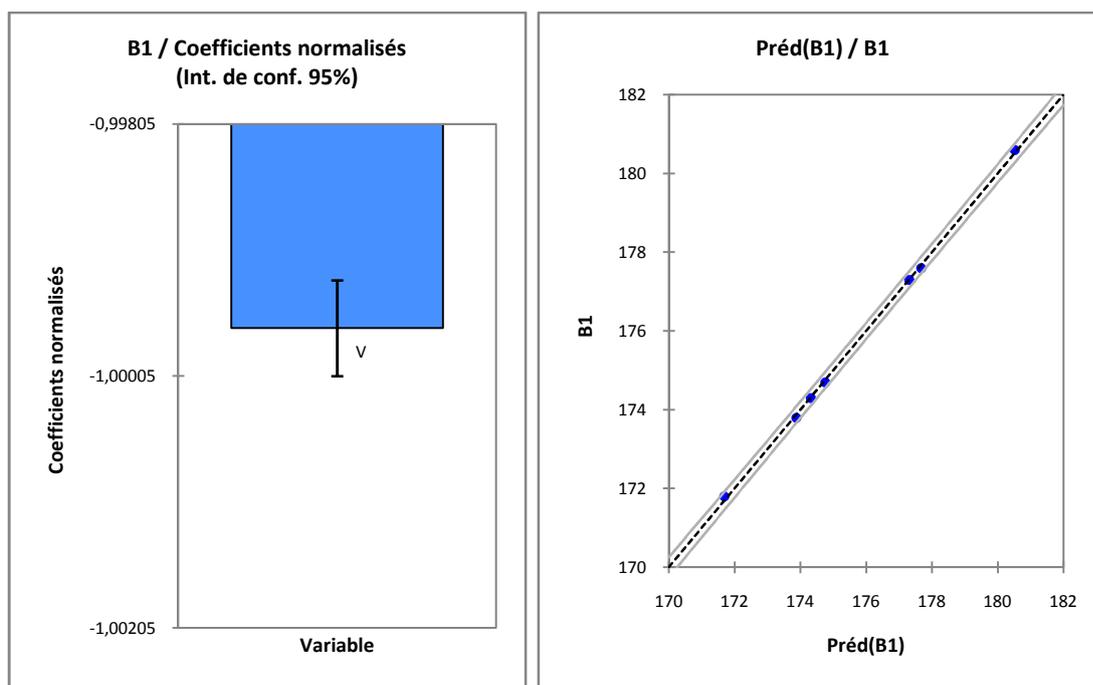
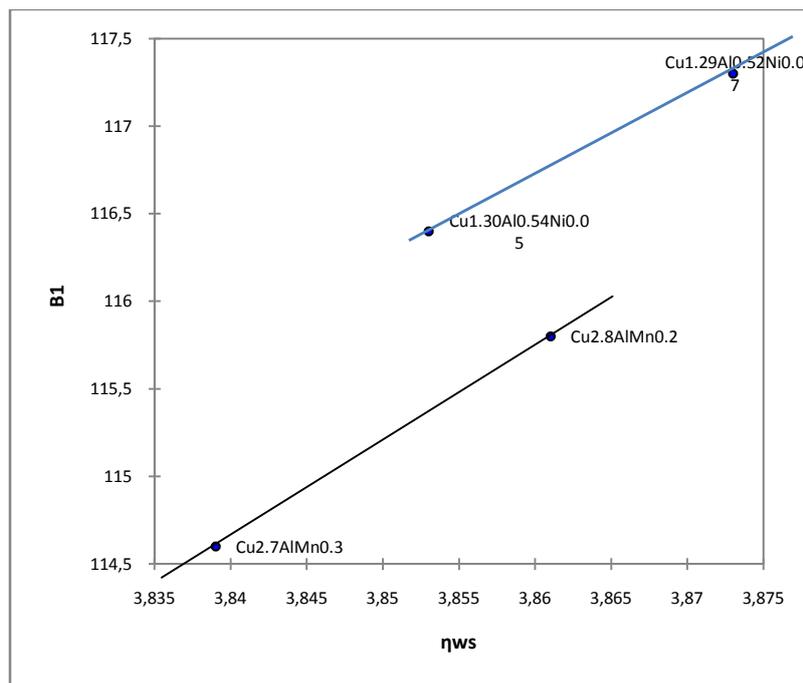


Figure [III.12] : graphe de résidus et Prédiction correspond à la variable B1

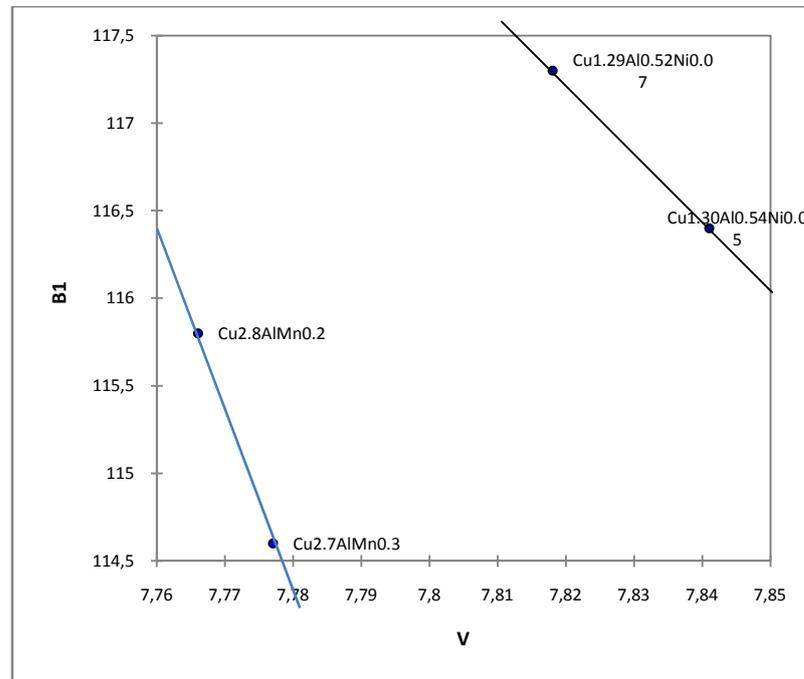
L'analyse du modèle correspondant à la variable B1 nous permet de conclure que L'ensembles des intermétalliques a base de CoAlNi, sont significatives et influencent le modèle. Nous remarquons aussi dans le graphe des prédictions que la variation est linéaire entre les valeurs de B1 prédites en fonction de celle utilisée en entrée, ce qui indique bien la bonne qualité du modèle.

- Pour le cluster 4 a base de [CuAlMn, CuAlNi, CuNiZn]



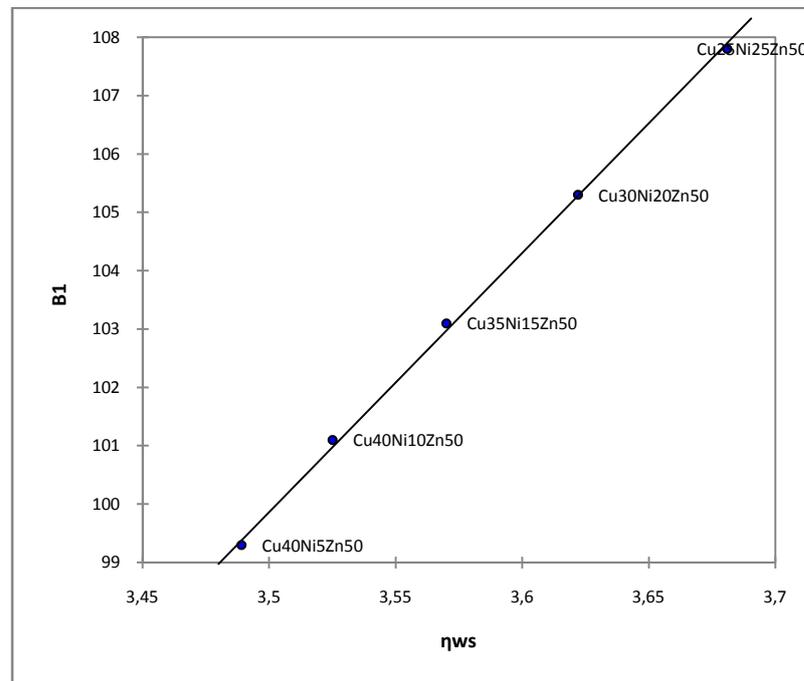
Figure(III.13) la variation de B1 on fonction de η_{ws} pour les intermétallique à base de CuAlMn et CuAlNi

- On remarque clairement que le module de compressibilité augmente avec l'augmentation de la densité moyenne d'électrons.



Figure(III.14) la variation de B1 on fonction de V pour les intermétallique à base de CuAlMn et CuAlNi

- On peut remarquer que la variation du module de compressibilité on fonction de volume molaire est inversement linéaire.



Figure(III.15) la variation de B_1 on fonction de η_{ws} pour les intermétallique à base de CuNiZn

On remarque clairement a partir la figure (III.15) que la variation de la concentration de deuxième élément de l'intermétallique a une influence remarquable sur le module de compressibilité B_1 .

III.3.4 Base de données 2 :

Pour cette étape nous avons construit une base de donnée comprenant : les mêmes matériaux que la base1 mais en ajoutant quelques propriétés.

Le tableau (III.1) contient l'ensemble de données utilisées. Tout l'ensemble de données est composé de 5 propriétés. Les propriétés et leurs abréviations incluses dans l'analyse sont comme suit : le volume molaire (V), la densité d'électrons (η_{ws}), l'énergie de cohésion (E_{ABC}), l'enthalpie de formation (H), le module de compressibilité (B_2).

nombres	composés	V	ws	E _{ABC}	H	B ₂
1	Al ₄₈ Co ₅ Ru ₄₇	8,917	4,304	483,66	-20,41	181,1
2	Al ₅₀ Co _{1.5} Ru _{48.5}	9,026	4,256	485,11	-18,22	178,7
3	Al ₅₀ Co ₅ Ru ₄₅	8,972	4,226	477,2	-21,03	175,8
4	Al ₅₂ Co _{1.5} Ru _{46.5}	9,08	4,18	478,65	-18,81	173,5
5	Al ₅₂ Co ₅ Ru ₄₃	9,028	4,148	470,74	-21,58	170,5
6	Co _{78.97} Al _{14.48} Ni _{6.55}	6,962	4,968	410,22	-16,21	181,2
7	Co _{80.75} Al _{12.68} Ni _{6.57}	6,926	5,017	411,96	-14,87	183
8	Co _{82.84} Al _{12.5} Ni _{4.66}	6,92	5,025	412,06	-14,27	182,6
9	Co _{82.85} Al _{10.58} Ni _{6.57}	6,885	5,074	414	-13,08	185
10	Co _{83.78} Al _{13.3} Ni _{2.92}	6,932	5,007	411,22	-14,15	181,20
11	Co _{84.89} Al _{10.39} Ni _{4.72}	6,88	5,081	414,11	-12,58	184,7
12	Co _{87.12} Al _{8.29} Ni _{4.59}	6,841	5,138	416,14	-10,59	186,6
13	Co ₆ Ru ₄₁ Ta ₅₃	9,627	4,942	706,4	-25,07	253,2
14	Co ₁₅ Ru ₃₅ Ta ₅₀	9,368	4,951	682,1	-27,2	248,2
15	Co ₂₀ Ru ₃₀ Ta ₅₀	9,28	4,921	670,8	-27,76	242,8
16	Co ₂₆ Ru ₂₄ Ta ₅₀	9,223	4,845	657,24	-27,92	234,1
17	Cu _{2.7} AlMn _{0.3}	7,777	3,839	329,7	-5,73	117,8
18	Cu _{2.8} AlMn _{0.2}	7,766	3,861	331,05	-4,61	118
19	Cu _{1.29} Al _{0.52} Ni _{0.07}	7,818	3,873	336,94	-5,42	119,7
20	Cu _{1.30} Al _{0.54} Ni _{0.05}	7,841	3,853	336,05	-4,49	118,4
21	Cu ₂₅ Ni ₂₅ Zn ₅₀	7,955	3,681	256	-14,45	120
22	Cu ₃₀ Ni ₂₀ Zn ₅₀	8,03	3,622	251,4	-13,73	115,7
23	Cu ₃₅ Ni ₁₅ Zn ₅₀	8,091	3,57	246,8	-12,24	111,3
24	Cu ₄₀ Ni ₁₀ Zn ₅₀	8,135	3,525	242,2	-9,75	106,7
25	Cu ₄₀ Ni ₅ Zn ₅₀	8,162	3,489	237,6	-5,89	102,3
26	Cu _{66.5} Zn _{20.8} Al _{12.7}	7,796	3,785	292,01	-0,61	112
27	Cu _{67.7} Zn _{19.4} Al _{12.9}	7,776	3,805	294,88	-0,61	112,9
28	Cu _{68.7} Al ₁₇ Zn _{14.3}	7,801	3,809	305,01	-0,68	113,6
29	Cu ₂ MnAl	7,852	3,685	320,25	-1,7	107,6

Tableau [III.1] : Base de données 2 [1], [2], [3], [4], [5], [6], [7]

III.3.4.1 L'analyse en composantes principales (ACP)

Le résultat des "Scores plot" de cette analyse est montré sur la figure suivante :

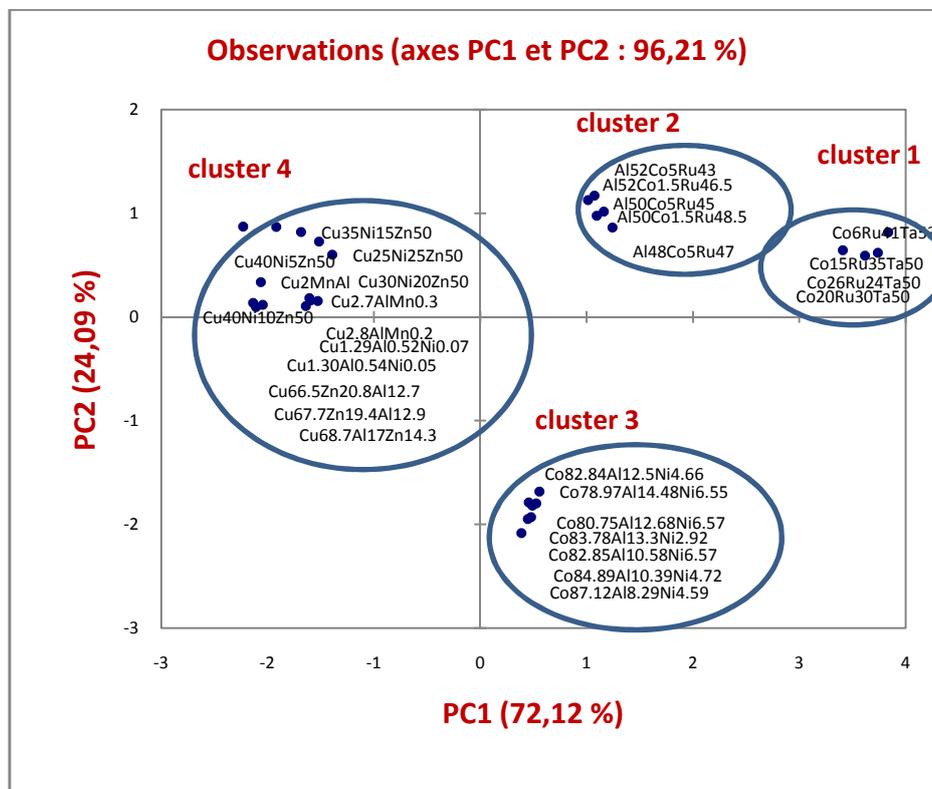


Figure [III.1] : Graphe des observations pour les matériaux intermétalliques (scores plot)

- Dans notre cas, PC1 capture 72.12% de la variance pour l'ensemble des données, et PC2 capture 24.09 % de la variance comme nous le voyons dans les figures [III.1]. Aucune autre composante principale PCs n'est incluse dans ces discussions parce qu'elles ne fournissent pas des informations significatives.
- En analysant la figure des scores (III.1), on remarque le même résultat avec le graphe de scores pour la base 1, juste des petites variations dans les positions des matériaux. Cela veut dire que quatre groupes ou clusters émergent, ceux avec un PC1 positif : 3 clusters contenus des matériaux intermétalliques à base de CoRuTa et AlCoRu et CoAlNi respectivement. Et avec (V) moyenne et un grand (B2) et une grande densité et une grande énergie de cohésion. et ceux avec un PC1 négatif les matériaux à base de

Cu : [CuNiZn, CuAlZn, CuZnAl, CuAlMn, CuAlNi, CuMnAl] que nous notons cluster 4. Et avec une grande valeur d'enthalpie de formation.

- Nous remarquons que ces 4 clusters à travers PC1 se regroupe par classe ayant à peu près un même module de compressibilité (bulk modulus) et une même densité électrique et une même énergie de cohésion à cette effet plus le PC1 augmente plus le bulk augmente et la densité augmente et l'énergie de cohésion augmente. Nous remarquons le PC2 par contre donne la variation de la concentration du deuxième élément dans l'intermétallique car le PC1 donne la concentration du premier élément de l'intermétallique.
- La figure [III.2] montre les graphes des variables « loadings » correspondant aux échantillons représentés sur la figure [III.1]. Les axes des deux graphes de loadings et scores sont identiques ainsi l'information dans les deux graphiques peut être comparée directement.

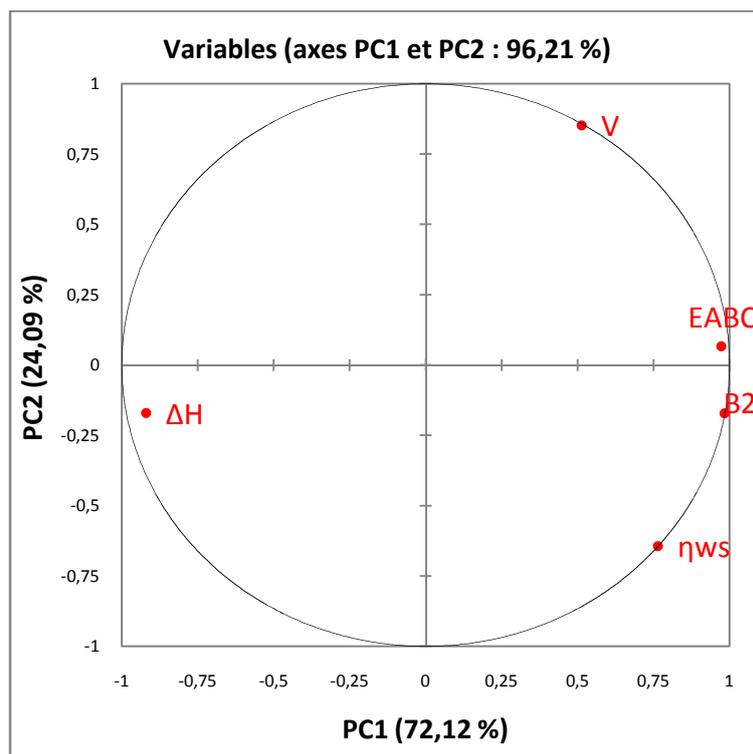


Figure (III-2) : Graphe des variables (loadings plot)

- Le graphe représente les variables ou les descripteurs (loadings plot).

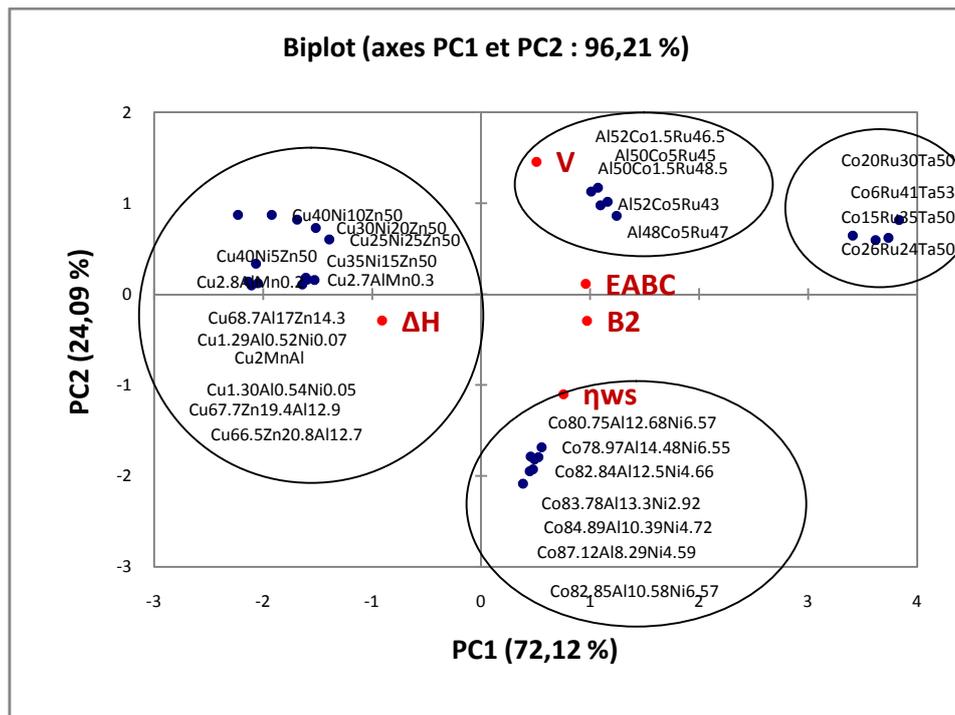


Figure (III.4) : Graphe superposant les échantillons et les variables (bibplot) des composés intermétalliques

En se basant sur les graphes (III.4) on peut constater que :

- Cluster1 et cluster3 sont fortement corrélés avec l'énergie E_{ABC} et le module de compressibilité B_2 et la densité η_{ws} , cela veut dire que ces matériaux ont un grand module de compressibilité et une grande densité et énergie de cohésion. Par contre ces deux clusters sont inversement corrélés avec l'enthalpie de formation.
- Cluster2 bien lié au volume molaire V , c'est-à-dire que ces matériaux possèdent un grand volume molaire. Aussi qu'une corrélation inverse avec ΔH et η_{ws} .
- Cluster4 est fortement lié avec l'enthalpie de formation et une corrélation inverse avec les quatre propriétés, tel que, E_{ABC} , B_2 , η_{ws} et V , cela confirme que ces matériaux sont incompressibles. Par conséquent, le module de compressibilité est un des critères utilisés pour caractériser la dureté.

III.3.4.2 Le modèle PLS (Partial Least Square)

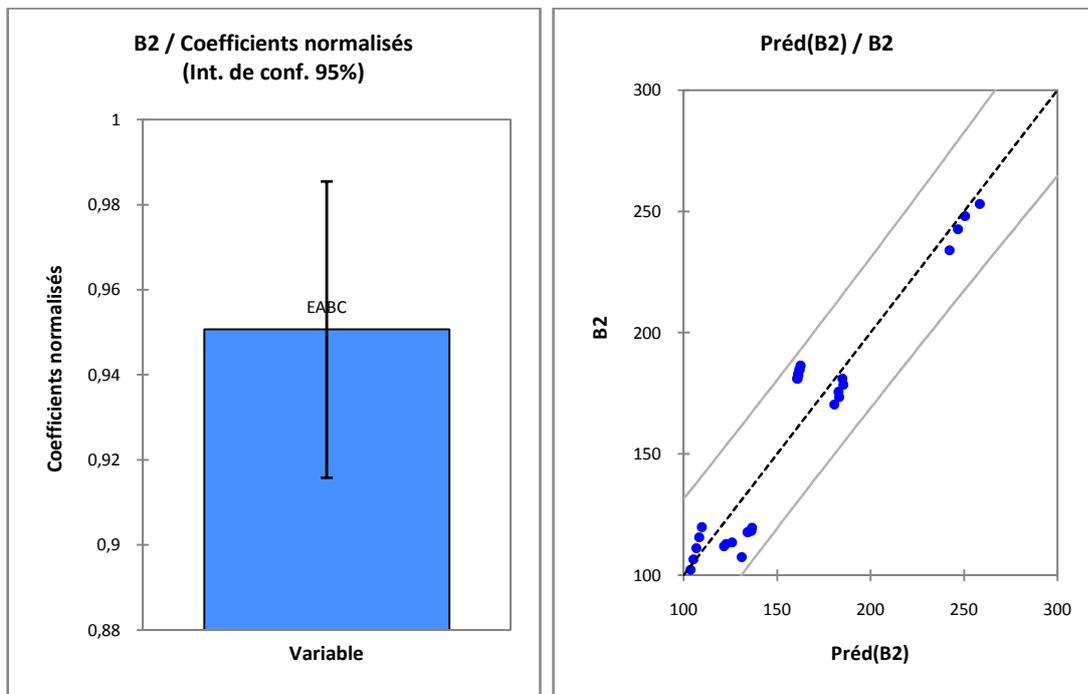
La régression PLS, nous donne des tableaux de prédictions et résidus qui permettent de vérifier que les valeurs d'une propriété sont très bien reproduites par ces variables.

Observation	Préd(B2)	Résidu
Al48Co5Ru47	184,800	-3,700
Al50Co1.5Ru48.5	185,278	-6,578
Al50Co5Ru45	182,667	-6,867
Al52Co1.5Ru46.5	183,146	-9,646
Al52Co5Ru43	180,535	-10,035
Co78.97Al14.48Ni6.55	160,556	20,644
Co80.75Al12.68Ni6.57	161,130	21,870
Co82.84Al12.5Ni4.66	161,163	21,437
Co82.85Al10.58Ni6.57	161,804	23,196
Co83.78Al13.3Ni2.92	160,886	20,314
Co84.89Al10.39Ni4.72	161,840	22,860
Co87.12Al8.29Ni4.59	162,510	24,090
Co6Ru41Ta53	258,329	-5,129
Co15Ru35Ta50	250,308	-2,108
Co20Ru30Ta50	246,577	-3,777
Co26Ru24Ta50	242,101	-8,001
Cu2.7AlMn0.3	133,975	-16,175
Cu2.8AlMn0.2	134,421	-16,421
Cu1.29Al0.52Ni0.07	136,365	-16,665
Cu1.30Al0.54Ni0.05	136,071	-17,671
Cu25Ni25Zn50	109,646	10,354
Cu30Ni20Zn50	108,127	7,573
Cu35Ni15Zn50	106,609	4,691
Cu40Ni10Zn50	105,090	1,610
Cu40Ni5Zn50	103,572	-1,272
Cu66.5Zn20.8Al12.7	121,533	-9,533
Cu67.7Zn19.4Al12.9	122,481	-9,581
Cu68.7Al17Zn14.3	125,825	-12,225
Cu2MnAl	130,856	-23,256

Tableau (III.2) : Coefficients normalisés (variables B2)

D'après le tableau des coefficients de la variable B2, nous remarquons que les résidus sont non nulle pour tous les éléments, donc on peut dire que le modèle n'est pas bon pour tous l'ensemble des intermétalliques. Ce résultat peut être confirmé par des graphes correspondant aux coefficients normalisés ainsi que des graphes des résidus et prédictions (figure [III.5]). Donc on ne peut pas prédire une relation entre B2 et E_{ABC} pour tous l'ensemble des intermétalliques.

La variable «B2»



Figure(III.5) : Graphes des résidus et prédictions correspondant à la variable «B2 »

Référence:

- [1] Every AG, McCurdy AK. Landolt-Bomstein New Series 1989; 3/ 29a:11–222.
- [2] Balazyuk VN, Geshko ET, Mikhalchenko VP, Sharlai BM. Fiz Met Metalloved 1976; 42:854.
- [3] Murakami Y. J Phys Soc Japan 1972;33:1350.
- [4] Nakamura M. In: Westbrook JH, Fleischer RL, editors. Intermetallic compounds vol. 1, Principles. John Wiley & Son Co; 1994. p. 873.
- [5] Fleischer RL, Field RD, Briant CL. Metall Trans 1991; 22A:403.
- [6] Fleischer RL. ISIJ International 1991; 31:1186.
- [7] Fleischer RL. J Mater Res 1993; 8:49.

Conclusion



Dans le cadre de ce mémoire nous nous sommes intéressés à l'étude des intermétalliques. Ces matériaux ont des propriétés très intéressantes.

Pour effectuer notre étude, nous nous sommes basés sur une nouvelle manière de représenter et d'analyser l'information. Il s'agit d'un nouvel axe en devenir baptisé « Génome des matériaux », destinée à la conception de nouveaux matériaux en utilisant les techniques du datamining. Pour l'extraire des connaissances à partir des données.

L'objectif principal est de montrer comment on peut utiliser l'informatique dans la physique des matériaux et ainsi interpréter des observations expérimentales massives.

Et on peut aussi découvrir de nouveaux schémas, et de nouvelles corrélations entre les différentes propriétés des matériaux.

Cette étude nous a permis de mieux comprendre les corrélations entre les propriétés structurales et thermodynamiques des composés intermétalliques. Par l'utilisation des deux techniques de datamining : l'analyse en composante principale (ACP) et la technique de moindres carrés partielles (PLS).

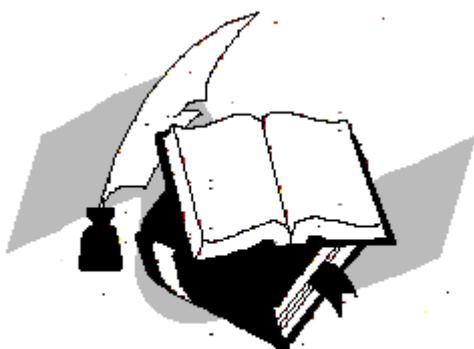
Le traitement de ces données avec l'ACP et la PLS grâce au logiciel XLSTAT nous a permis dans un premier temps :

De déduire, grâce à l'approche de l'ACP (analyse en composante principale), pour les intermétalliques la forte corrélation qui existe entre le module de compressibilité et la densité d'électrons moyenne et l'énergie de cohésion, par contre une corrélation inverse entre ces propriétés et l'enthalpie de formation.

De constater que le module de compressibilités est lié à la variation de la concentration des éléments de l'intermétallique.

La deuxième étape de notre travail a été consacrée à l'approche PLS (la régression des moindres carrés) qui nous a permis de prédire des relations simples entre ces propriétés.

Annexes



I. Introduction

Le logiciel XLSTAT est un outil statistique d'analyse de données pour Microsoft Excel le plus complet et le plus utilisé, et il est destiné aux tests multidimensionnels sur les moyennes et les variances, XLSTAT offre de très nombreuses fonctionnalités qui font d'Excel un outil performant et facile d'accès pour répondre à la majorité de nos besoins en analyse de données et modélisation. XLSTAT fonctionne avec toutes les versions d'Excel, depuis la version 97 jusqu'à la version 2007 sous l'environnement Windows.

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 des composés intermétalliques dans notre cas, et les variables dépendantes (les Y du modèle) indiquées par les propriétés (B1, B2, w_s , V, H, E_{ABC}).

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. L'utilisation des différents boutons dépendra du type d'action désirée par l'utilisateur, l'une correspond à la régression PLS et l'autre à l'ACP.

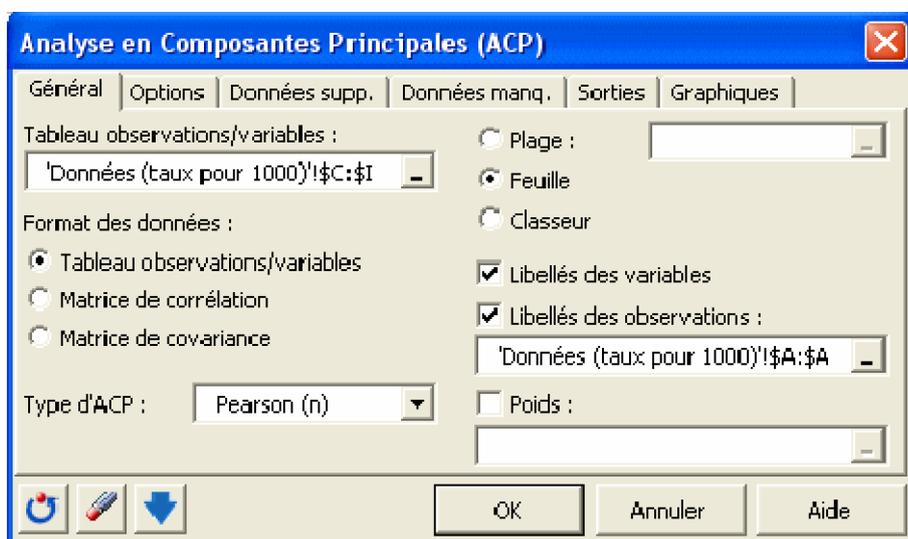


II.1 Faire une Analyse en Composantes Principales (ACP) avec XLSTAT

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 "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

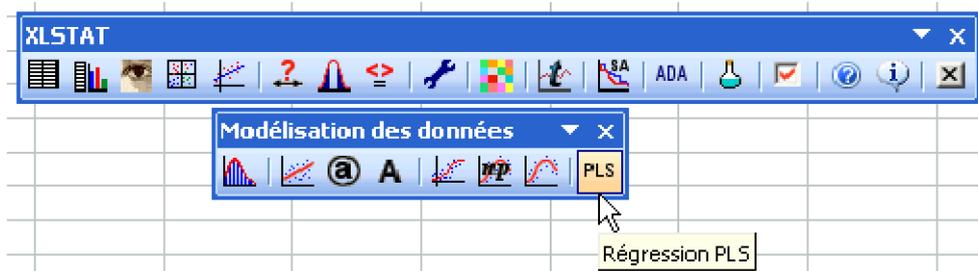
génération du graphique, et de ne pas afficher toutes les observations afin de rendre le graphique plus lisible.



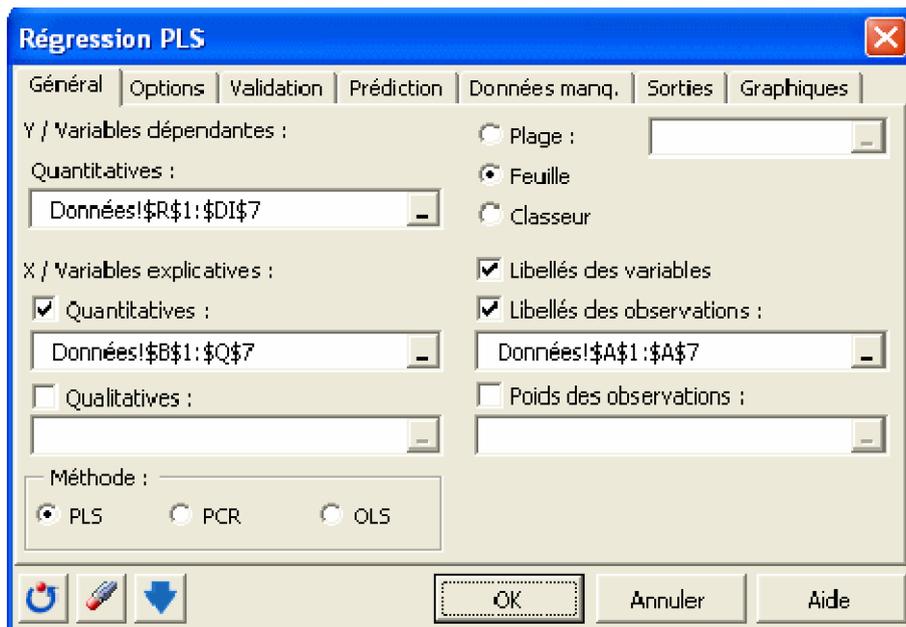
- Le premier résultat intéressant est la matrice des corrélations. Les tableaux des valeurs et vecteurs propres sont ensuite affichés. Tous les graphiques et tableaux sont mentionnés dans la partie résultat.

II.2 Faire une régression PLS avec XLSTAT

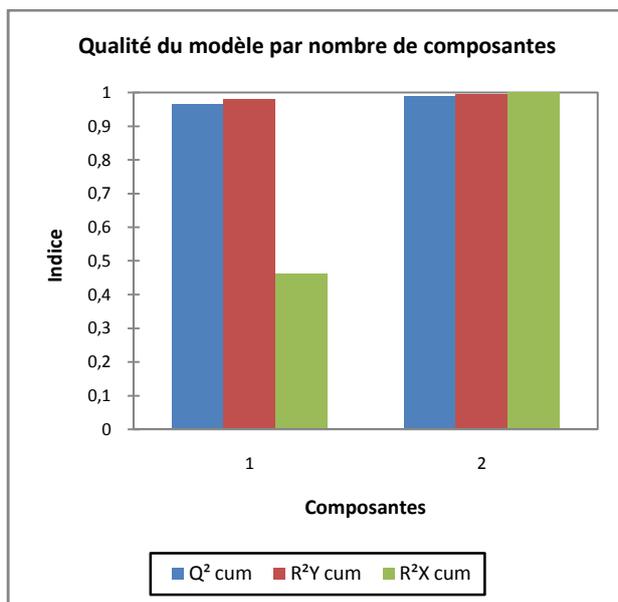
Nous lançons XLSTAT en cliquant sur le bouton  de la barre d'outils Excel, puis nous sélectionnons la commande Régression PLS de la barre d'outils "Modélisation des données".



- 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 intermétalliques.



- 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
Q ² cum	0,966	0,989
R ² Y cum	0,979	0,994
R ² X cum	0,461	1,000

