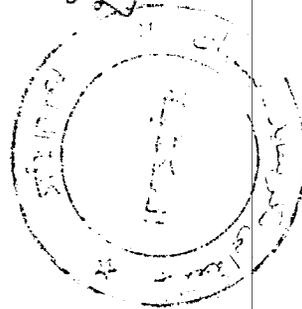
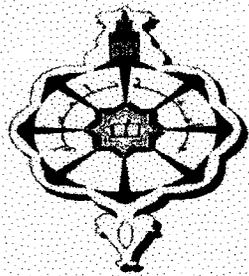


IN/530.4-09/
FACULTE DES SCIENCES

DEPARTEMENT DE PHYSIQUE



Mémoire :

Présenté par : Belahcene Brahim

Pour l'obtention du diplôme d'Ingénieur d'Etat en Physique

Spécialité : Matériaux et caractérisation

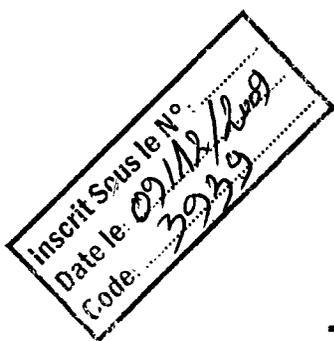
THEME :

***DATAMINING DES PROPRIETES
PHYSIQUES DES ELEMENTS DU
TABLEAU PERIODIQUE***

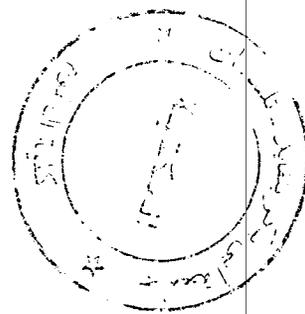
Soutenu en Mai 2009 devant le jury :

M ^r	G. Merad	Président	Professeur	Univ. Tlemcen
M ^{elle}	R. Mahiaoui	Examinateur	Maitre assistant (A)	Univ. Tlemcen
M ^{elle}	H. I. Faraoun	Encadreur	Maitre de Conférences (A)	Univ. Tlemcen





Remerciements



Je remercie ALLAH le tout puissant de nous avoir accordé la puissance et la volonté pour terminer ce travail.

Je tiens à exprimer ma reconnaissance à M^{elle} H.I. FARAOUN Maitre de Conférences (A) à l'université de Tlemcen, pour avoir dirigé ce mémoire.

J'exprime mes sincères remerciements à monsieur G. MERAD, Professeur à l'Université de Tlemcen, d'avoir accepter de présider le Jury de cette thèse.

Je remercie sincèrement M^{elle} R. Mahiaoui Maitre assistant (A) à l'université de Tlemcen, d'avoir accepter de faire partie du Jury.

A mes parents,

A mes frères,

A mes amis.

Table de matière

Introduction générale.....	2
----------------------------	---

Chapitre 1

Constitution de la base de données – Le tableau périodique des éléments.....	3
I.1. Historique.....	4
I.1.1. La théorie des triades.....	4
I.1.2. La détermination du « <i>poids atomique</i> ».....	4
I.1.3. Le tableau périodique de Mendeleïev.....	5
I.2. Construction de la base de données.....	6
I.2.1. La base de données.....	6
I.3. Définitions.....	10
I.3.1. Le nombre atomique.....	10
I.3.2. La masse atomique.....	10
I.3.3. Les états d'oxydation.....	11
I.3.4. L'électronégativité.....	11
I.3.4.1 Définitions.....	12
a. Définition de <i>Pauling</i>	12
b. Définition de <i>Mulliken</i>	12
c. Définition d' <i>Allred et Rochow</i>	12
d. Définition de <i>Parr</i>	12
I.3.5. L'énergie de première ionisation.....	13
I.3.6. Le paramètre du réseau.....	13
I.3.7. nombre d'atomes par cellule.....	13
I.3.8. Le volume molaire.....	13
I.3.9. Le module de Young E	14
I.3.10. Les constants élastiques et les modules d'élasticité.....	16
I.3.11. La conductivité thermique.....	17
I.3.12. La capacité molaire.....	17
I.3.13. L'entropie standard S_0	17
I.3.14. La différence d'enthalpie $H_{298}-H_0$	18

I.3.15. La température de fusion T_m	19
I.3.16. Température d'ébullition T_b	19
Conclusion I.....	20
Références I	20

Chapitre 2

Le datamining – la méthode d'Analyse en Composantes Principales ACP.....	22
II.1. Introduction	23
II.2. Le Datamining	23
II.3. Analyse En Composantes Principales « ACP »	23
II.3.1. Formulation de l'Analyse en Composantes Principales.....	24
a. Matrice des données	24
b. Centrage et réduction des données	25
c. Distances entre objets	25
d. Inertie totale I du nuage	25
II.3.2 Réalisation de l'analyse en composantes principales	25
a. Relations de transitions	26
b. Nuage des variables	26
II.4 Calcul d'une ACP avec MATLAB	27
Conclusion II.....	28
Références II.....	28

Chapitre 3

III.1. Introduction.....	30
III.2. Interprétation des résultats	30
III.3. Corrélation des énergies du système atomique et électronique	34
III.3.1. Définitions	34
III.3.2. La Base de données	34
III.3.3. Analyse	37
Conclusion III.....	39
Références III.....	39
Conclusion générale	40

Introduction générale

~ 1 ~

Introduction

La modernisation technologique exige de chercher de nouveaux matériaux répondant à des propriétés spécifiques. Pour cela, une approche empirique peut être mise en œuvre. Ainsi des bases de données de résultats expérimentaux peuvent être utilisées pour construire des modèles de prédictions des propriétés recherchées. Des modèles de prédictions efficaces permettent en effet de restreindre la recherche de matériaux aux cas les plus prometteurs.

La classification est une étape importante pour une analyse de données efficace. Elle consiste à regrouper les objets d'un ensemble de données en classes homogènes. Il existe deux types d'approches: la classification supervisée et la classification non supervisée. Elles se différencient par leurs méthodes et par leur but. La *classification supervisée* est basée sur un ensemble d'objet T (appelé ensemble d'apprentissage) dont la classe est connue, le but étant de découvrir la structure des classes à partir de l'ensemble T afin de pouvoir généraliser cette structure sur un ensemble de données plus large. La *classification non supervisée* consiste à diviser un ensemble de données D en sous ensembles, appelés classes, tel que les objets d'une classe sont similaires et que les objets de classes différentes sont différents, afin d'en comprendre la structure sous-jacente.

Le tableau périodique des éléments, également dit tableau de Mendeleïev correspond à des ordonnancements des éléments chimiques, définis de façon à respecter une périodicité bien spécifique. De tels ordonnancements peuvent par exemple être produits en parcourant la table périodique de haut en bas puis de gauche à droite. Des travaux récents semblent montrer que de tels ordonnancements peuvent être utilisés pour prédire efficacement la formation de composés chimiques. En réalité, le tableau de Mendeleïev est probablement le premier exemple de l'utilisation d'une technique d'analyse de données en science des matériaux.

Dans ce travail, nous nous sommes intéressés à la classification non supervisée des éléments du tableau périodique par la méthode d'analyse en composantes principales (ACP). Le but est d'analyser les corrélations entre les propriétés physiques des éléments et voir si la classification des éléments peut se faire autrement que suivant la classification de Mendeleïev.

Chapitre I

Constitution de la base de données – Le tableau périodique des éléments

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

Dans ce chapitre, nous allons présenter la base de données des propriétés physiques des éléments du tableau périodique utilisée dans notre travail.

I.1. Historique

Depuis l'antiquité, on connaît quelques corps simples comme le cuivre, l'or, le fer, l'argent ou le soufre. En 1700, seuls 12 corps simples (formés d'un seul élément) ont été isolés: l'antimoine, l'argent, l'arsenic, le carbone, le cuivre, l'étain, le fer, le mercure, l'or, le phosphore, le plomb et le soufre. Les techniques d'analyse évoluant, notamment grâce à l'apparition de la pile de Volta en 1800 permettent de faire des électrolyses, le nombre des éléments connus en 1850 sera multiplié par 5.

I.1.1. La théorie des triades

En étudiant les propriétés des éléments, les chimistes découvrent que certains d'entre eux possèdent des propriétés chimiques voisines. C'est ainsi que naît la théorie des triades. Une triade est un groupe de 3 éléments ayant des propriétés chimiques voisines.

- ⇒ En 1808 l'anglais Davy étudie la triade calcium, strontium et baryum
- ⇒ En 1818 le même Davy révèle les propriétés communes à la triade lithium, sodium et potassium
- ⇒ En 1817 le chimiste allemand Döbereiner suggère l'existence de la triade chlore, brome, iode.

Vers 1850, une vingtaine de triades sont identifiées et plusieurs tentatives de classification des éléments suivent mais sans succès.

I.1.2. La détermination du « poids atomique »

Au début du XIX^{ème} siècle, un savant de Manchester, John Dalton, introduit l'idée qu'à chaque élément correspond un atome à qui il assigne un « poids atomique ». Bien sûr il ne pouvait pas peser les atomes sur une balance. Mais, il déterminait ce « poids atomique » d'après la masse minimale de chaque élément qui entre en composition avec un autre.

Dalton choisit arbitrairement comme unité de référence le « poids atomique » de l'hydrogène égal à 1. Pour Dalton, les combinaisons chimiques (corps composés) résultaient de l'association de ces atomes en proportions fixes dont il déduisait les masses relatives expérimentalement. Pour caractériser un élément chimique, les chimistes utilisaient ce « poids atomique ».

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

- La distinction entre élément, atome et molécule est encore peu claire. Suite aux travaux de Gay-Lussac (1809) et Avogadro (1811), la distinction se fait entre atomes et molécules et dès 1814, Berzélius propose une notation à base de lettres pour représenter les éléments et détermine avec précision la masse d'un grand nombre d'éléments.
- Il faut attendre 1860, lors du premier congrès international de chimie à Karlsruhe pour que soient présentées et acceptées les notions d'atome et de molécule ainsi que la définition d'un système de masse atomique pour chaque élément. Le jeune chimiste russe Dimitri Mendeleïev assiste à ce congrès.

1.1.3. Le tableau périodique de Mendeleïev

Dimitri Ivanovitch Mendeleïev est né à Tobolsk, en Sibérie, le 8 février 1834 et est décédé à Saint-Petersbourg le 2 février 1907. Chimiste diplômé de l'Université de Saint-Petersbourg en 1856, il part pour Heidelberg et travaille avec des savants comme Robert Bunsen, Gustav Kirchhoff et le compositeur-chimiste Borodine. À Paris, ensuite, il se penche sur la densité des gaz et sur le spectroscope de Gustav Kirchhoff. En 1867, il rentre en Russie et est nommé professeur de chimie minérale à l'Université de Saint-Petersbourg.

L'histoire du tableau débute en 1817 avec Johann Dobereiner, un chimiste russe, qui établit déjà à l'époque un rapport entre la masse atomique des éléments et leurs propriétés. En 1865, John Alexander Newlands met en exergue une certaine répétition dans les propriétés des éléments. Il lance alors sa loi des octets ou des octaves : «Le huitième élément suivant un élément donné ressemble au premier comme la huitième note de l'octave ressemble à la première». Arrive ensuite la «*vis tellurique*» d'Alex Beguyer de Chancourtois, hélice ayant sur son axe vertical la suite des nombres entiers correspondant aux poids atomiques des éléments. Aucune de ces théories ne sera retenue. Mendeleïev planche sur un travail semblable, étudiant les 63 éléments chimiques connus à l'époque. Se faisant, il note qu'en rangeant ces éléments selon leur masse atomique croissante apparaît alors une nette périodicité concernant leurs propriétés. Afin de mettre en évidence cette périodicité, il range les 63 éléments verticalement selon la masse atomique croissante, changeant de colonne lorsque les propriétés particulières réapparaissent.

Présenté en 1869 à la société russe de chimie sous l'intitulé: «La dépendance entre les propriétés des masses atomiques des éléments», le tableau de l'époque est conçu au départ

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

Symbole	Nombre atomique Z	masse atomique	Les états de l'oxydation	ELECTRONEGATIVITY %a		Paramètre réseau	Nombre d'atomes par cellule	Le volume molaire	Module de Young E	S ₁₁ 'constante de élastique'	S ₄₄	S ₁₂	C ₁₁ module élasticité
H	1	1.008	1	2.2	13.	377.1	2	13.26	0.2	2930	9090	1240	0.042
He	2	4.003	1	3.2	24.	357.7	2	32.07	0.1	2327	4738	1102	0.031
Li	3	6.939	1	0.97	5.3	350.9	2	13	11.5	315	104	-144	13.4
Be	4	9.012	2	1.47	9.3	228.5	2	4.88	286	3.45	6.16	-0.28	292
O	8	15.99	2	3.5	13.	540.3	4	8	0	1280	3640	2.6	0.275
Na	11	22.98	1	1.01	5.1	420.9	2	0.565	6.8	549	233	-250	7.59
Mg	12	24.31	2	1.23	7.6	320.9	2	1.23	44.4	22	60.9	-7.8	59.3
Al	13	26.98	3	1.47	5.9	361.4	4	10	70.2	16	35.3	-5.8	108
Si	14	28.09	4	1.74	8.1	543.1	8	12.06	112	12.7	-2.15	165.6	79.6
S	16	32.06	6	2.44	10.	540.3	128	15.49	17.8	74.6	121	-13.1	14.22
K	19	39.10	1	0.91	4.3	532.1	2	0.64	3.52	1339	526	-620	3.69
Ca	20	40.06	2	1.04	6.1	558.8	4	25.86	19.6	104	71	-42	22.8
Sc	21	44.96	3	1.2	6.5	330.8	2	15.04	75.2	12.5	36.1	-4.3	99.3
Ti	22	47.9	4	1.32	6.82	295.03	2	10.55	1.83E-	9.69	21.5	-4.71	160
V	23	50.94	5	0.525	6.7	302.3	2	8.34	127	6.75	23.2	-2.31	230
Cr	24	52	6	1.56	6.7	288.4	2	7.23	145	3.05	9.98	-0.49	348
Mn	25	54.94	7	1.6	7.4	892.1	58	7.38	196	6.25	7.95	-2.23	336.5
Fe	26	55.85	3	1.64	7.9	286.6	2	7.09	211	7.67	8.57	-2.83	230
Co	27	58.93	3	1.75	7.8	250.7	2	6.62	204	5.11	14.1	-2.37	295
Ni	28	58.71	3	1.75	7.6	352.4	4	6.59	220	7.67	8.23	-2.93	247
Cu	29	63.54	2	1.75	7.7	361.4	4	7.09	1.3	4760	7.02	128	46.8
Zn	30	65.37	2	1.66	9.3	266.4	2	9.17	92.7	8.22	25.3	0.6	85.6
Ga	31	69.72	3	1.82	5.9	451.9	8	11.81	9.81	12.2	28.6	-4.4	100
Ge	32	72.59	4	2.02	7.8	565.9	8	13.64	79.9	14.9	-2.64	129	67.1
As	33	74.59	5	0.81	9.7	413.2	2	12.95	22	46.71	44.91	36.94	123.6
Kr	36	83.8	7	3.1	13.	564.5	4	45.28	1.8	618	744	-226	2.85
Rb	37	85.47	1	0.89	4.1	570.3	2	55.79	2.35	1330	625	-600	2.96
Sr	38	87.62	2	0.99	5.6	608.4	4	34.5	15.7	218	135	-90	10.94
Y	39	88.91	3	1.11	6.2	364.8	2	19.89	66.3	15.4	41.1	-5.1	77.9
Zr	40	91.22	4	1.22	6.6	323.1	2	14.02	68	10.1	30.1	-4	144
Nb	41	92.91	5	0.893	6.7	330.0	2	10.84	104	6.56	35.2	-2.29	245
Mo	42	95.94	6	1.3	7.0	314.7	2	9.39	330	2.63	9.2	-0.68	465
Tc	43	99	7	0.55	7.2	273.8	2	8.6	407	5.7	-1.1	433	177
Ru	44	101.1	8	1.42	7.3	270.5	2	8.14	432	2.09	5.53	-0.58	563
Rh	45	102.9	4	1.35	7.4	280.3	4	8.29	379	3.46	5.43	-1.1	413
Pd	46	106.4	4	1.35	8.3	389.0	4	5.23	121	13.7	14.1	-6	221
Ag	47	107.8	1	1.42	7.5	408.6	4	3.62	80	23	22	-9.8	122
Cd	48	112.4	2	1.46	8.9	297.8	2	13	62.3	12.4	53.1	-1.2	114.1
In	49	114.8	3	1.49	5.7	459.9	2	15.71	10.6	148.8	153.7	-46	45.1

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

Sn	50	118.6	4	1.72	7.3	581.9	4	16.24	1.83E-	0.33	41.6	45.6	31.2
Sb	51	121.7	5	1.046	8.6	450.6	2	18.2	54.4	16.1	38.9	-6.1	101
Xe	54	131.3	0	2.4	12	613.2	4	33.03	45.32	690	708	-271	2.93
Cs	55	132.9	1	0.86	3.8	614.1	2	70.96	1.69	1190	690	-450	1.6
Ba	56	137.3	2	0.97	5.2	502.3	2	38.21	12.8	157	105	-61	12.6
La	57	138.9	3	1.08	5.5	377.4	4	22.6	39.2	51.7	55.7	-19.2	34.5
Ce	58	140.1	4	-1.08	5.5	516.1	4	17	44.1	62.8	57.8	-22.3	26
Pr	59	140.9	4	-1.07	5.4	367.2	4	20.8	35.2	26.6	73.6	-11.3	49.4
Nd	60	144.2	3	-1.07	5.5	365.8	4	20.59	37.9	23.7	66.5	-9.5	54.8
Gd	64	157.2	3	1.11	6.1	363.3	2	19.9	56.3	18	16.1	48.1	67.8
Tb	65	158.9	4	-1.1	5.8	360.5	2	19.31	57.5	17.4	46	-5.2	69.2
Dy	66	162.5	3	-1.1	5.9	359.1	2	19	63.1	16	41.2	-4.6	74
Ho	67	164.9	3	-1.1	6.0	357.7	2	18.75	67.1	15.3	38.6	-4.3	76.5
Er	68	167.2	3	-1.11	6.1	355.9	2	18.44	73.3	14.1	36.4	-4.2	84.1
Yb	70	173.0	3	-1.06	6.2	548.4	4	24.84	18.4	89.2	56.4	-31.9	18.6
Lu	71	174.9	3	-1.14	5.4	350.5	2	17.78	68.4	14.3	37.3	-4.2	86.2
Hf	72	178.4	4	-1.23	6.8	319.4	2	13.41	138	7.16	18	-2.48	181
Ta	73	180.9	5	0.322	7.5	330.3	2	10.87	185	6.89	12.1	-2.58	264
W	74	183.8	6	-1.4	7.8	316.5	2	9.53	407	2.45	6.24	-0.69	523
Re	75	186.2	7	0.15	7.8	276.0	2	8.86	520	6.21	-0.8	683	273
Ir	77	192.2	6	1.44	8.9	383.9	4	8.57	528	2.28	3.9	-0.67	580
Pt	78	195.0	4	1.44	8.9	392.3	4	9.1	170	7.35	13.1	-3.08	347
Au	79	196.9	3	-1.42	9.2	407.8	4	10.19	78	23.4	23.8	-10.8	191
Hg	80	200.5	2	-1.44	10	300.5	1	14.81	25	154	151	-119	36
Tl	81	204.3	3	1.44	6.1	345.6	2	17.24	7.89	104	139	-83	41.9
Pb	82	207.1	4	1.55	7.4	405.9	2	18.26	0.0000	0.44	93.7	68	-43
Bi	83	208.9	5	0.946	7.2	308.6	2	21.44	34	26	114	-7.9	63.4

C_{44} module élasticité	La conductivité thermique λ	La capacité molaire C_p	L'entropie standard S_0	La différence d'enthalpie H298- H0	La température de fusion T_m	Le changement d'enthalpie à la fusion ΔH_m	Température d'ébullition T_b	Le changement d'enthalpie à l'ébullition ΔH_b
0.11	0.171	14.418	130.68	8.4684	13.81	0.12	20.3	0.46
0.0217	0.143	20.786	126.152	6.1965	0.95	0.021	4.215	0.082
9.6	84.7	24.77	29.12	4.632	453.69	3	1620	147.7
163	200	16.44	9.5	1.95	1560	7.895	2741	291.58
2.06	0.0245	14.69	205.147	8.68	54.361	0.444	90.18	6.2
4.3	141	28.24	51.3	6.46	370.87	2.597	1156	99.2
16.4	171	24.895	32.671	4.9957	923	8.4768	1366	127.4
28.3	237	24.392	28.3	4.54	933.47	10.711	2790	294
63.9	83.7	20	18.81	3.217	1687	50.208	3505	383.3

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

8.27	0.269	22.7	32.07	4.412	388.36	1.721	882	9.62
1.9	102.4	29.58	64.68	7.088	336.86	2.3208	1040	79.1
14	200	25.94	41.588	5.736	1115	8.5395	1773	153.6
27.7	15.8	25.52	34.644	5.2174	1814	14.0959	3104	314.2
46.5	21	25.02	30.72	4.824	1941	14.146	3631	410
43.1	30.7	24.9	30.89	4.507	2183	21.5	3690	451.8
100	93.7	23.44	23.543	4.05	2180	21.004	2952	344.3
115	29.7	26.28	32.22	4.9957	1519	12.9089	2335	226.7
117	80.2	25.1	27.28	4.489	1811	13.806	3139	349.6
71	100	24.811	30.04	4.7655	1768	16.2	3184	376.6
122	83	26.07	29.796	4.787	1728.3	17.4798	3157	369.24
0.34	1401	24.443	33.15	5.004	1357.77	13.263	2843	300.7
39.6	121	25.44	41.631	5.657	692.68	7.322	1180	115.3
35	33.5	26.15	40.727	5.572	302.91	5.5898	2478	258.7
48.3	58.6	33.347	31.09	4.636	1211	36.9447	3107	331
22.6	50	24.65	35.689	5.117	1090	24.4429	61.3	34.8
1.35	9.49E-03	20.786	164.085	6.1965	115.765	1.64	119.8	9.05
1.6	58.2	31.062	76.776	7.489	312.47	2.1924	970	75.7
7.41	35.3	26.4	55.694	6.568	1050	7.431	1685	137.19
24.3	17.2	26.53	44.788	5.9835	1795.15	11.3942	3611	363.3
33.4	22.7	25.36	39.181	5.5663	2127.85	20.9978	4203	561.3
28.4	59	24.69	36.27	5.22	2750	30	5017	683.2
109	142	23.932	28.56	4.589	2893	37.4798	4952	582.2
199	185	23.56	32.985	4.356	2430.01	33.2912	4538	592.9
181	117	24.06	28.614	4.6024	2607	38.589	4423	595.5
184	150	24.98	31.556	5.36	2237	26.5935	3970	493.3
70.8	71.8	25.98	37.823	5.4685	1828	16.736	3237	357.6
45.5	429	25.36	42.551	5.7446	1234.93	11.297	2436	250.6
19	96.8	25.98	51.8	6.247	594.22	6.1923	1040	97.4
6.51	81.6	26.732	57.65	6.61	429.75	3.283	2346	231.45
72.3	66.6	27.17	51.18	6.323	505.08	7.0194	2876	295.8
39.5	25.9	25.23	45.522	5.8702	903.78	19.874	1860	165.8
1.41	0.0051	20.744	169.575	6.197	161.391	3.1	165.03	12.65
1.44	35.9	32.18	85.23	7.711	301.59	2.096	947	67.7
9.5	18.4	28.09	62.5	6.91	1000	7.119	2118	141.5
18	13.5	27.11	56.902	6.6651	1193	6.1965	3730	413.7
17.3	11.4	26.4	69.454	7.2002	1072	5.4601	3699	414.2
13.6	12.5	27.2	73.931	7.4182	1204	6.8869	3785	296.8
15	16.5	27.45	71.006	7.1337	1289	7.1421	3341	273
20.8	10.6	37.02	68.009	9.0076	1506.15	9.668	3569	359.4
21.8	11.1	28.91	73.8	9.4266	1632	10.1504	3496	330.9
24.3	10.7	28.16	74.956	8.8659	1685.15	11.3505	2835	230
25.9	16.2	27.15	75.019	7.9956	1745	11.757	2968	242.5
27.4	14	28.12	73.178	7.3923	1802	19.9033	3136	261.4
17.7	38.5	26.74	59.831	6.7111	1097	7.6567	1467	128.83
26.8	16.4	26.86	50.961	6.389	1936	18.6481	3668	355.9
55.7	23	25.3	43.56	5.845	2506	27.196	4963	575.5
82.6	60.7	25.3	41.472	5.6819	3290	36.5682	5778	743.1
160	164	24.27	32.618	4.97	3693	52.3137	5828	806.8

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

1.16	71.2	25.31	36.482	5.333	3458	34.075	5869	714.8
256	147	24.98	35.505	5.2677	2719	41.124	4701	604.1
76.5	71.6	25.85	41.631	5.7237	2041.5	22.175	4100	509.8
42.2	317	25.42	47.488	6.0166	1337.33	12.552	3130	334.4
12.9	8.34	27.983	75.9	9.342	234.32	2.2953	629	59.2
7.2	46.1	26.32	64.3	6.832	577	4.1422	1746	164.1
14.8	35.2	26.51	64.8	6.87	600.61	4.7739	2019	177.58
11.5	7.87	25.52	56.735	6.4266	544.55	11.2968	1837	174.1

I.3. Définitions :

I.3.1. Le nombre atomique

Le numéro atomique (Z) est le terme employé en chimie et en physique pour représenter le nombre de protons du noyau d'un atome. Un atome peut être schématisé en première approche par une agglomération compacte (noyau atomique) de protons (p^+) et de neutrons (n), autour de laquelle circulent des électrons (e^-). Dans un atome de charge électrique neutre, le nombre d'électron est lui aussi égal au numéro atomique (Z protons de charge $+e$, plus Z électrons de charge $-e \Rightarrow$ charge totale = 0 Coulomb).

À l'origine, ce numéro déterminait la position d'un élément chimique dans le tableau périodique. Quand Dimitri Mendeleïev ordonna les corps chimiques connus en fonction de leurs propriétés, il s'aperçut que le classement dans l'ordre strict de masse atomique pouvait conduire à des situations illogiques : par exemple, l'iode et le tellure doivent être inversés par rapport à leur masse atomique. Ce classement selon les propriétés chimiques détermine le numéro atomique et transmissible tout de suite, approximativement proportionnel à la masse de l'atome, mais comme le prouvent les quelques inversions relevées, représentatif d'autres propriétés que la masse.

I.3.2. La masse atomique

La masse atomique est un nombre sans dimension qui est assez proche du nombre de nucléons dans le noyau de l'atome.

La masse d'un atome s'exprime le plus souvent en UMA, unité de masse des atomes qui équivaut à la masse de l'isotope léger de l'Hydrogène.

La masse molaire d'un atome est en fait la masse d'une mole d'un élément chimique donné et vaut autant de grammes que la valeur de la masse atomique : c'est la masse d'une mole, soit de $6,022 \times 10^{23}$ atomes du même élément.

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

La masse atomique est également définie pour un élément chimique, c'est alors la moyenne pondérée des masses atomiques des isotopes au prorata de la fréquence de ces isotopes tels qu'on les trouve dans la nature.

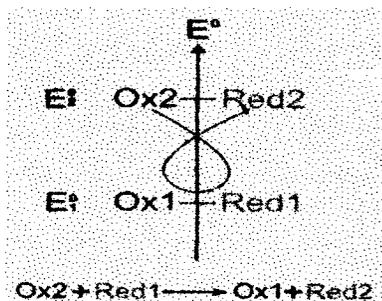
I.3.3. Les états d'oxydation

Une réaction d'oxydoréduction est une réaction chimique au cours de laquelle se produit un transfert d'électrons. On appelle *oxydant* l'atome qui capte les électrons et *réducteur* l'atome qui les cède au cours de la réaction.

Dans une oxydoréduction,

- l'élément qui perd un ou des électron(s) est appelé réducteur,
- l'élément qui capte un ou des électron(s) est appelé oxydant.

Le réducteur s'oxyde (réaction d'oxydation), l'oxydant se réduit (réaction de réduction).



L'oxydoréduction se compose donc de deux demi-réactions : une oxydation et une réduction.

- *Oxydation*
réducteur(1) = oxydant(1) + ne
- *Réduction*
oxydant(2) + ne^- = réducteur(2)

Oxydoréduction (« somme » de l'oxydation et de la réduction)
oxydant(2) + réducteur(1) \rightarrow oxydant(1) + réducteur(2)

I.3.4. L'électronégativité

L'électronégativité d'un élément est une grandeur qui caractérise sa capacité à attirer les électrons lors de la formation d'une liaison chimique avec un autre élément (À titre d'exemple, le fluor est l'élément le plus électronégatif tandis que le francium est le moins électronégatif). La différence d'électronégativité entre ces deux éléments détermine la nature de la liaison chimique : covalente lorsque la différence est faible (inférieure à 1,7), ionique lorsque cette différence est plus importante.

- Les gaz inertes n'ont pas d'électronégativité
- Plus un atome est petit, plus son électronégativité est grande

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

L'électronégativité est notée χ_A où A est le symbole de l'élément considéré. Plus χ_A est grand et plus l'élément est susceptible d'attirer des électrons en lui.

1.3.4.1 Définitions

Il existe plusieurs définitions de l'électronégativité (*Pauling, Mulliken, Parr, Allred et Rochow*) ce qui a conduit à construire plusieurs échelles.

a. Définition de *Pauling*

La différence d'électronégativité entre les éléments A et B a pour expression

$$\Delta\chi_{AB} = 0,102 * ((E_{AB} - 1/2(E_{AA} + E_{BB}))^{1/2} \quad (1.1)$$

où E_{AB}, E_{AA} et E_{BB} sont les énergies de liaison des molécules diatomiques A-B, A-A et B-B.

Cette définition ne donne que la différence entre deux électronégativités. On a donc besoin d'une origine qui a été fixée arbitrairement en donnant la valeur de 4 à l'électronégativité du fluor (élément le plus électronégatif de la classification).

b. Définition de *Mulliken*

L'électronégativité d'un élément est la moyenne de son affinité électronique A_e et de son énergie d'ionisation E_I

$$\chi = \frac{A_e + E_I}{2} \quad (1.2)$$

c. Définition d'*Allred et Rochow*

L'électronégativité d'un élément a pour expression

$$\chi = \frac{Z_{eff} \cdot e^2}{r_{cov}^2} \quad (1.3)$$

où Z_{eff} est la charge effective du noyau, e la charge élémentaire et r_{cov} le rayon covalent de l'élément.

d. Définition de *Parr*

L'électronégativité est l'opposée de la dérivée de l'énergie de l'atome par rapport au nombre d'électrons.

$$\chi = -\left(\frac{\partial E}{\partial n}\right) \quad (1.4)$$

Les échelles d'électronégativité les plus utilisées sont l'échelle de *Mulliken*, l'échelle d'*Allred-Rochow* et l'échelle de *Pauling*.

Les électronégativités des atomes impliqués dans une liaison tendent à être égales (principe d'égalisation des électronégativités de *Sanderson*, 1951). L'égalisation des électronégativités est réalisée par le transfert de densité électronique vers l'atome le plus électronégatif.

L'électronégativité permet d'estimer le caractère ionique d'une liaison à l'aide de la relation de *Pauling*

$$I_{AB} = 100(1 - e^{-\Delta\chi^2/4}) \quad (1.5)$$

Ou de celle de Haney et Smith

$$I_{AB} = 16\Delta\chi + 3.5\Delta\chi^2 \quad (1.6)$$

L'électronégativité est également la notion à l'origine de la polarité de certaines molécules. En effet, dans une molécule, lorsque les atomes de part et d'autre de la liaison covalente ont des électronégativités différentes, l'atome le plus électronégatif attire davantage les électrons. Les éléments dont l'électronégativité est faible sont fréquemment dits électropositifs.

I.3.5. L'énergie de première ionisation

C'est l'énergie nécessaire pour extraire le premier électron (c'est-à-dire l'électron dont l'énergie de liaison est la plus faible) de la structure atomique.

Formule générique décrivant l'ionisation à un électron : $X \rightarrow X^+ + e^-$

& (I.5.1, I.5.2, I.5.3, I.5.4, I.5.5) Réf : [12-13-14]

I.3.6. Le paramètre du réseau

Le paramètre du réseau est une grandeur caractéristique de la structure cristalline de l'élément. Il représente la distance entre deux nœuds consécutifs constituant la maille et qui sont l'image l'un de l'autre par translation du réseau.

I.3.7. nombre d'atomes par cellule

Le nombre d'atomes par cellule dépend de la structure cristalline. Autant ce nombre est important, autant le matériau est plus dense et plus compact.

& (I.5.6, I.5.7) Réf : [5-10]

I.3.8. Le volume molaire

En chimie, le volume molaire d'une substance est le volume occupé par une mole de cette substance. Son unité est le litre par mole ($L \cdot mol^{-1}$).

Dans les unités du système international, le volume molaire s'exprime en mètres cubes par mole ($m^3 \cdot mol^{-1}$). Il est souvent exprimé dans les conditions normales de température et de pression (CNTP).

Le volume molaire est une entité physico-chimique qui augmente avec l'élévation de la température. ; Il Peut se calculer en utilisant la formule :

$$V_m = \frac{V}{n} \quad (1.7)$$

Cette formule s'applique aux gaz uniquement avec

V le volume en litres et ;

n la quantité de matière en moles.

En utilisant la formule de calcul de la masse molaire M :

$$n = \frac{m}{M} \quad (1.8)$$

On peut déterminer une seconde formule :

$$m = \frac{V \times M}{V_m} \quad (1.9)$$

On peut aussi utiliser la relation des gaz parfaits :

$$PV = nRT \quad (1.10)$$

Afin de déterminer :

$$V_m = \frac{RT}{P} \quad (1.11)$$

I.3.9. Le module de Young E Le module de Young ou module d'élasticité longitudinale est la constante qui relie la contrainte de traction et la déformation longitudinale pour un matériau isotrope.

Le physicien britannique Thomas Young (1773-1829) avait remarqué que le rapport entre la contrainte de traction appliquée à un matériau et la déformation qui en résulte (un allongement relatif) est constant, tant que cette déformation reste petite et que la limite d'élasticité du matériau n'est pas atteinte. La loi d'élasticité est la loi de Hooke :

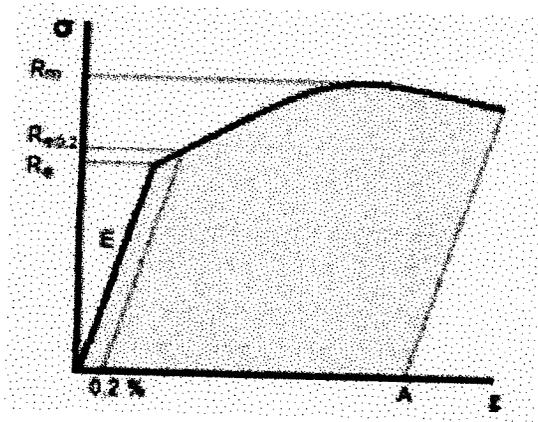


Diagramme contrainte déformation

$$\sigma = E\varepsilon \quad (1.12)$$

Où :

- σ est la contrainte (en unité de pression),
- E est le module de Young (en unité de pression),
- ε est l'allongement relatif (sans dimensions).

Le module de Young est la contrainte mécanique qui engendrerait un allongement de 100% de la longueur initiale d'un matériau (il doublerait donc de longueur), si l'on pouvait l'appliquer réellement : dans les faits, le matériau se déforme de façon permanente, ou se rompt, bien avant que cette valeur soit atteinte.

D'après l'équation aux dimensions, le module de Young est homogène à une pression, ou plus précisément une contrainte. L'unité internationale est donc le pascal (Pa). Cependant, en raison des valeurs élevées que prend ce module, il est en général donné en mégapascal (MPa) ou Newton par millimètre carré (N/mm²).

Dans le cas d'un matériau cristallin et certains matériaux amorphes, le module de Young exprime la « force de rappel » électrostatique qui tend à maintenir les atomes à distance constante. Il peut s'exprimer en fonction de la dérivée seconde du potentiel interatomique.

I.3.10. Les constants élastiques et les modules d'élasticité

La loi de Hooke établit que la déformation est directement proportionnelle à la contrainte pour des déformations suffisamment faibles. Par conséquent, les composantes de la déformation sont des fonctions linéaires des composantes de la contrainte.

$$\begin{aligned}
 e_{xx} &= s_{11}X_x + s_{12}Y_y + s_{13}Z_z + s_{14}Y_z + s_{15}Z_x + s_{16}X_y \\
 e_{yy} &= s_{21}X_x + s_{22}Y_y + s_{23}Z_z + s_{24}Y_z + s_{25}Z_x + s_{26}X_y \\
 e_{zz} &= s_{31}X_x + s_{32}Y_y + s_{33}Z_z + s_{34}Y_z + s_{35}Z_x + s_{36}X_y \\
 e_{yz} &= s_{41}X_x + s_{42}Y_y + s_{43}Z_z + s_{44}Y_z + s_{45}Z_x + s_{46}X_y \\
 e_{zx} &= s_{51}X_x + s_{52}Y_y + s_{53}Z_z + s_{54}Y_z + s_{55}Z_x + s_{56}X_y \\
 e_{xy} &= s_{61}X_x + s_{62}Y_y + s_{63}Z_z + s_{64}Y_z + s_{65}Z_x + s_{66}X_y
 \end{aligned}
 \tag{1.13}$$

Inversement, les composantes de la contrainte sont des fonctions linéaires des composantes de déformation :

$$\begin{aligned}
 X_x &= C_{11}e_{xx} + C_{12}e_{yy} + C_{13}e_{zz} + C_{14}e_{yz} + C_{15}e_{zx} + C_{16}e_{xy} \\
 Y_y &= C_{21}e_{xx} + C_{22}e_{yy} + C_{23}e_{zz} + C_{24}e_{yz} + C_{25}e_{zx} + C_{26}e_{xy} \\
 Z_z &= C_{31}e_{xx} + C_{32}e_{yy} + C_{33}e_{zz} + C_{34}e_{yz} + C_{35}e_{zx} + C_{36}e_{xy} \\
 Y_z &= C_{41}e_{xx} + C_{42}e_{yy} + C_{43}e_{zz} + C_{44}e_{yz} + C_{45}e_{zx} + C_{46}e_{xy} \\
 Z_x &= C_{51}e_{xx} + C_{52}e_{yy} + C_{53}e_{zz} + C_{54}e_{yz} + C_{55}e_{zx} + C_{56}e_{xy} \\
 X_y &= C_{61}e_{xx} + C_{62}e_{yy} + C_{63}e_{zz} + C_{64}e_{yz} + C_{65}e_{zx} + C_{66}e_{xy}
 \end{aligned}
 \tag{1.14}$$

Les quantités S_{ij} sont appelées constantes d'élasticité ; les quantités C_{ij} sont appelées constantes de rigidités élastique ou modules d'élasticités. Les S_{ij} ont la dimension : [surface] / [force] ou [volume] / [énergie]. Les C_{ij} ont la dimension [force] / [surface] ou [énergie] / [volume].

& (I.5.8, I.5.9, I.5.10) Réf : [5-6-7]

1.3.11. La conductivité thermique

La conductivité thermique est une grandeur physique caractérisant le comportement des matériaux lors du transfert thermique par conduction. Elle représente la quantité de chaleur transférée par unité de surface et par une unité de temps sous un gradient de température.

Mathématiquement, la conductivité thermique λ peut donc s'écrire comme la somme de deux contributions :

$$\lambda = \lambda_e + \lambda_p \quad (1.15)$$

Où

- λ_e est la contribution des électrons
- λ_p est la contribution des vibrations des atomes (phonons)

1.3.12. La capacité molaire

La capacité thermique (ou capacité calorifique) d'un corps est une grandeur permettant de quantifier la possibilité qu'a un corps d'absorber ou restituer de l'énergie par échange thermique au cours d'un changement de température. La capacité calorifique est l'énergie qu'il faut apporter à un corps pour augmenter sa température de un Kelvin. Elle s'exprime en Joule/Kelvin. C'est une grandeur extensive : plus la quantité de matière est importante plus la capacité calorifique est grande. Toute chose étant égale par ailleurs, plus la capacité thermique d'un corps est grande, plus grande sera la quantité d'énergie échangée au cours d'une transformation s'accompagnant d'une variation de la température de ce corps.

On appelle capacité calorifique molaire à volume constant $C_{Vm}(V, T)$, le rapport de la quantité d'énergie transmise par chaleur Q_v nécessaire pour faire monter la température d'une mole de corps pur d'un degré.

1.3.13. L'entropie standard S_0

Le troisième principe de la thermodynamique stipule que l'entropie d'un corps pur est nulle à la température de 0°K. Donc on peut calculer l'entropie molaire d'un corps pur de façon absolue. Ce calcul permet de dresser des tables thermodynamiques où sont reportées les

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

entropies molaires standards de la plupart des composés, définies à la température de référence de 298°K, notée S_{289}^0 ou plus simplement S_0 .

I.3.14. La différence d'enthalpie $H_{298}-H_0$

L'enthalpie (du préfixe *en-* et du grec *thalpein* : « chauffer ») est une fonction d'état extensive de la thermodynamique, dont la variation permet d'exprimer la quantité de chaleur mise en jeu pendant la transformation à pression constante (isobare) d'un système thermodynamique au cours de laquelle celui-ci reçoit ou fournit un travail mécanique.

Considérons une transformation isobare au cours de laquelle le système passe d'un état A à un état B d'équilibres en échangeant de la chaleur Q_p et du travail uniquement par l'intermédiaire des forces de pression $W_{f,p}$.

Le premier principe permet d'écrire :

$$\Delta U = U_B - U_A = Q_p + W_{f,p} \quad (1.16)$$

U étant la fonction d'état énergie interne

À pression constante le travail des forces de pression est égal à :

$$W_{f,p} = -p\Delta V = -p(V_B - V_A) \quad (1.17)$$

$$U_B - U_A = Q_p - p(V_B - V_A) \quad (1.18)$$

D'où :

$$Q_p = (U_B + pV_B) - (U_A + pV_A) \quad (1.19)$$

On définit ainsi une nouvelle fonction d'état, la fonction enthalpie $H(U,p,V)$

$$H = U + pV = ST + pV \quad (1.20)$$

Il s'ensuit que :

$$Q_p = H_B - H_A = \Delta H \quad (1.21)$$

Par conséquent, à pression constante, la chaleur mise en jeu, qui n'est pas une fonction d'état, devient égale à la variation de la fonction d'état enthalpie H. La variation de cette

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

fonction ne dépend que de l'état final et de l'état initial du système et est indépendante du chemin suivi par la transformation.

C'est l'intérêt de l'application de la fonction enthalpie dans les cas très courants de transformations effectuées à l'air libre, à pression atmosphérique constante.

Cette propriété est à la base de la calorimétrie à pression constante. Par abus de langage on confond souvent les termes chaleur et enthalpie.

L'enthalpie a la dimension d'une énergie, et s'exprime en joules dans le Système International.

La propriété mathématique induite pour toute fonction d'état implique que sa différentielle est totale exacte c'est-à-dire qu'elle est égale à la somme des différentielles partielles par rapport à chaque variable.

I.3.15. La température de fusion T_m

Le point de fusion d'un corps représente la température à laquelle coexistent son état solide et son état liquide. Il s'agit d'une constante pour une substance pure donnée.

T_m est habituellement calculée sous pression atmosphérique normale (1 atmosphère).

Il est donc souvent possible, en connaissant son point de fusion, de déterminer la nature d'une substance (identification), bien que cette information ne soit pas suffisante pour permettre une identification complète d'une nouvelle molécule.

I.3.16. Température d'ébullition T_b

La température la plus élevée que peut atteindre un corps avant de s'évaporer, sous forme gazeuse, librement, se nomme le *point d'ébullition*. Cette température se calcule à la pression atmosphérique de 1 atmosphère (101.3 kPa).

Il est possible de définir de façon équivalente le point d'ébullition comme étant la température à laquelle la pression de vapeur du liquide est de 1 atmosphère.

& (I.5.11, I.5.12, I.5.13, I.5.14, I.5.15, I.5.16) Réf : [8-9-10]

Conclusion I

La détermination des propriétés physique des éléments métallique et non métallique est une étape incontournable du développement de nouveaux matériaux non seulement pour l'identification, et la caractérisation des matériaux mais aussi et surtout pour la compréhension, donc l'optimisation, des phénomènes physiques mis en jeu.

Dans ce chapitre nous avons présenté quelques propriétés physiques des matériaux, qui nous allons présenter leurs valeurs expérimentales dans le chapitre II, comme une base de donnés.

Références I

- [1]. W. Martienssen (Ed.): *Numerical Data and Functional Relationships in Science and Technology, Landolt–örnstein, New Series III and IV* (Springer, Berlin, Heidelberg 1970–2003).
- [2]. R. Blachnik (Ed.): *Elemente, anorganische Verbindungen und Materialien, Minerale, D'Ans-Lax, Taschenbuch für Chemiker und Physiker, Vol. 3, 4th edn.* (Springer, Berlin, Heidelberg 1998).
- [3]. D. R. Lide (Ed.): *CRC Handbook of Chemistry and Physics, 80th edn.* (CRC Press, Boca Raton 1999).
- [4]. Lehrstuhl für Werkstoffchemie, T.H. Aachen: *Thermodynamic Properties of Inorganic Materials, Landolt– Börnstein, New Series IV/19* (Springer, Berlin, Heidelberg 1999).
- [5]. Charles Kittel (trad. Nathalie Bardou, Évelyne Kolb), *Physique de l'état solide (Solid state physics)*, (1998).
- [6]. H.B Huntington, « Elastic constants of crystal » *Solid state physics* 7, 213 ,(1958).
- [7]. A.E.HLOVE, *A treatise on the mathematical theory of elasticity*, Dover Paperbak,(1944).
- [8]. *Thermodynamique, fondements et applications*, J.Ph. Pérez, A.M. Romulus, p155 édition Masson .
- [9]. Harris Benson. *Physique 1 : mécanique. 3^e édition.* Éditions du Renouveau Pédagogique, Saint-Laurent, Québec , p. 519 , (2004) .
- [10]. David R .Lideh ; *Handbook of chemistry and Physics 88th Edition* ,Editor-in-Chief;(2007-2008).

Chapitre I : Constitution de la base de données – Le tableau périodique des éléments

- [11]. Henry Guerlac (1961) *Lavoisier — the crucial year. The background and origin of his first experiments on combustion in (1772)* .
- [12]. Schüring, J., Schulz, H. D., Fischer, W. R., Böttcher, J., Duijnisveld, W. H. (editors)(1999). *Redox: Fundamentals, Processes and Applications*, Springer-Verlag, Heidelberg, 246 pp. ISBN 978-3540665281.
- [13]. *Thermoelectric Handbook*, Ed. Rowe DM - Chemical Rubber Company, Boca Raton (Floride) (1995).
- [14]. GS Nolas (et al.), *Thermoelectric, basic principles and new materials developments*, Springer (2001).
- [15]. GD Mahan (et al.), *Thermoelectric materials: new approaches to an old problem*, *Physics Today*, Vol. 50, p42, (1997).

Chapitre II

Le datamining – la méthode d'Analyse en Composantes Principales ACP

Chapitre II : Le datamining – la méthode d'Analyse en Composantes Principales ACP

II.1. Introduction

La classification est une étape importante pour l'analyse des données, elle consiste à regrouper les objets d'un ensemble de données en classes homogènes [1-2].

Pour estimer le potentiel de prédiction de la classification des éléments du tableau périodique, on s'intéresse à l'une des méthodes du Datamining, c'est la méthode d'Analyse en Composante Principale « ACP ».

II.2. Le Datamining

Le datamining est apparu dans les années 90, pour extraire des connaissances nouvelles à partir d'une base de données. Le datamining tire, à partir d'une vaste base de données, des conclusions nouvelles quant aux données entrées [6-7-8].

Il existe de nombreux logiciels de datamining tels que Statistica, XLSTAT, SPSS, Clémentine, SAS ou Oracle... Tous ont la méthodologie en commun : il faut trouver un domaine, construire une base (data warehouse), définir des questions ouvertes et le problème général. Une fois ces étapes accomplis, en rentrant (input) dans le logiciel les requêtes de la base, il en sortira, en principe, les réponses aux questions posées (après une brève analyse des résultats - output).

Dans ce travail, nous avons programmé la méthode de classification ACP par le langage Matlab.

II.3. Analyse En Composantes Principales « ACP »

L'Analyse en Composante Principale (ACP) fait partie des analyses descriptives multivariées.

Le but de cette analyse est de résumer le maximum d'informations possibles en perdant le moins possible pour :

- Faciliter l'interprétation d'un grand nombre de données initiales ;
- Donner plus de sens aux données réduites ;

L'ACP permet donc de réduire des tableaux de grandes tailles en un petit nombre de variables tout en conservant un maximum d'information. Les variables de départ sont dites 'métriques'.

Plusieurs méthodes sont préconisées dans la littérature pour réduire la taille de dimension de la grande base de données [3].

Chapitre II : Le datamining – la méthode d'Analyse en Composantes Principales ACP

Le nombre de données recueillies pour d'écrire un phénomène peut être trop important pour permettre un traitement direct de toute l'information disponible. Bien souvent l'information totale, constituée de p variables (mesures, observations, ...), est redondante et le but de l'algorithme de l'ACP est d'extraire $q < p$ "facteurs" (qui sont des combinaisons linéaires des p variables) contenant tout ou une grande partie de l'information contenue dans les p variables de départ.

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

Les vecteurs propres orthonormaux sont utilisés pour construire les composantes principales et les valeurs propres sont les variances des composantes principales correspondantes [4].

II.3.1. Formulation de l'Analyse en Composantes Principales [9]

a. Matrice des données

La matrice suivante représente la forme générale d'une classe (cluster) des données.

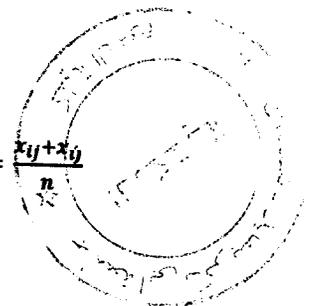
$$\begin{pmatrix} x_{11} & \dots & x_{1p} \\ \vdots & x_{ij} & x_{ij} & \vdots \\ x_{m1} & x_{ij} & x_{mp} \end{pmatrix} \quad (2.1)$$

$$\begin{pmatrix} \bar{x}_1 & \bar{x}_i & \bar{x}_p \\ s_1 & s_i & s_p \end{pmatrix}$$

m objets (ou individus), p descripteurs (ou variables)

s Coefficient de simple concordance (Sokal et Michener, 1958)

$$s(i, i) = \frac{x_{ij} + x_{ij}}{n}$$



b. Centrage et réduction des données

Pour centrer les données d'une matrice on utilise la formule suivante :

$$X_{jj} = \frac{(x_{jj} - \bar{x}_i)}{s_i} \quad (2.2)$$

c. Distances entre objets

L'écart ou la distance entre les individus ou les objets est donnée par l'expression suivante :

$$d(i, \ell)^2 = \sum_{j=1}^p p_j (x_{ij} - x_{\ell j})^2 \quad (2.3)$$

ou p_j est une pondération qui est en général égale à 1.

d. Inertie totale I du nuage

On note I le moment d'inertie du nuage des individus

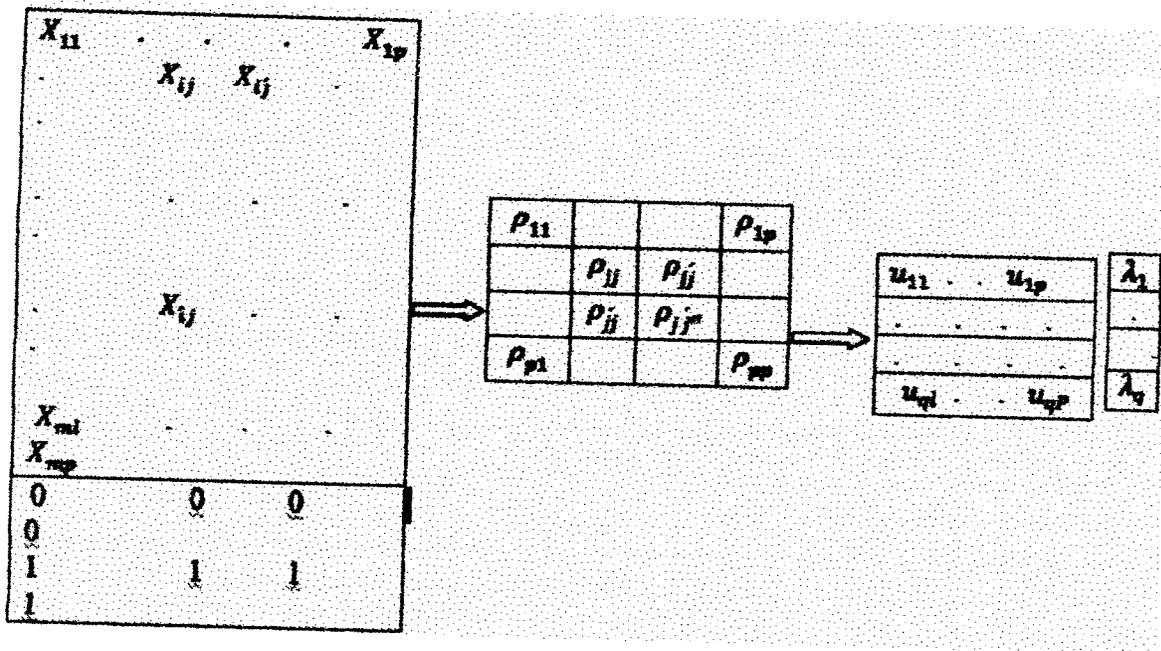
$$I = \frac{1}{m} \sum_{j=1}^p \sum_{i=1}^m \left(\frac{x_{ij} - \bar{x}_j}{s_j} \right)^2 = p \quad (2.4)$$

Ce moment d'inertie totale est intéressante car c'est une mesure de la dispersion du nuage des individus par rapport à son origine. Si ce moment d'inertie est grand, cela signifie que le nuage est très dispersé, tandis que s'il est petit, cela signifie que alors le nuage est très concentré sur son centre de gravité.

II.3.2 Réalisation de l'analyse en composantes principales

Le but : trouver des axes orthogonaux (indépendants) qui ont la propriété d'extraire le maximum de la variance projetée (ou inertie projetée) des individus ou objets.

- calcul de la matrice des données centrées-réduites ;
- calcul de la matrice des corrélations variables/variables ;
- calcul de la matrice des vecteurs propres et du vecteur des valeurs propres ;



Avec $q \leq p$

a. Relations de transitions

$$F_q(i) = \frac{1}{\sqrt{\lambda_q}} \sum_{j=1}^p \left(\frac{x_{ij} - \bar{x}_j}{s_j} \right) G_q(j) \quad (2.5)$$

ou F est la projection de l'individu i sur l'axe factoriel q , G les projections des p variables sur q , et λ la valeur propre associée à la composante q .

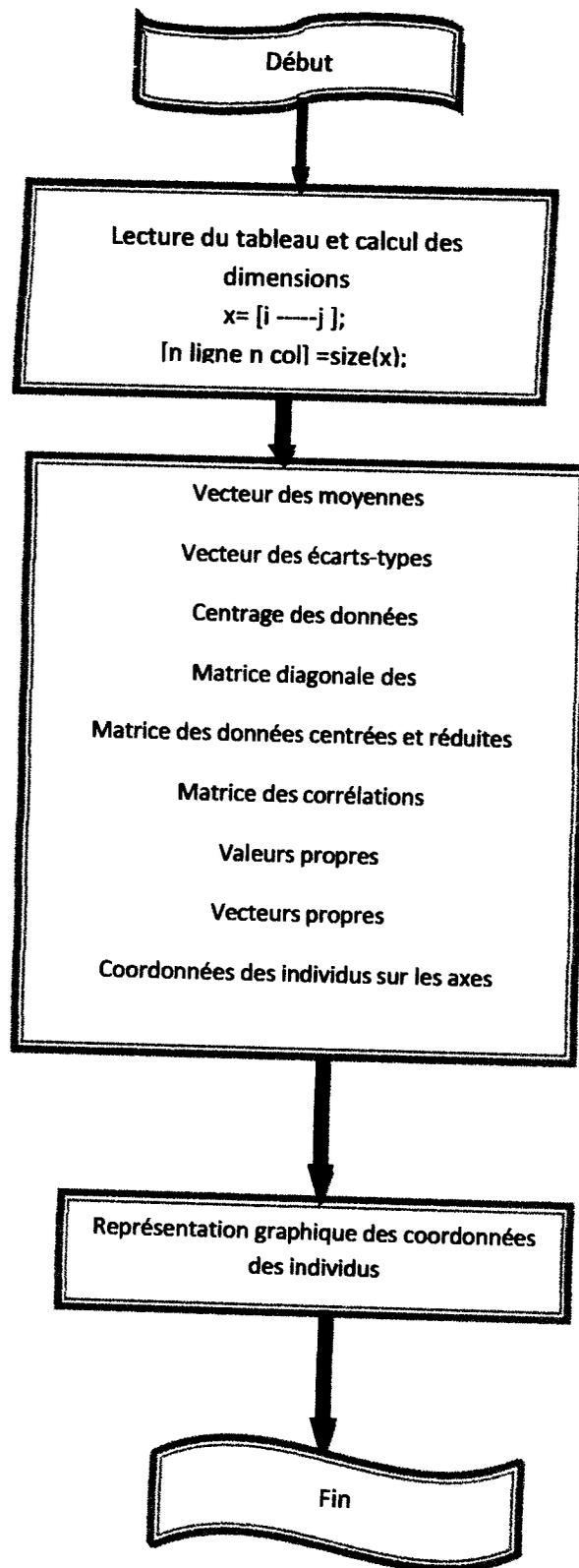
$$G_q(j) = \frac{1}{\sqrt{\lambda_q}} \sum_{i=1}^m \left(\frac{x_{ij} - \bar{x}_j}{s_j} \right) F_q(i) \quad (2.6)$$

b. Nuage des variables

On peut alors définir l'angle $(\widehat{j, j})$ entre deux vecteurs par son cosinus :

$$\cos(j, j) = \sum_{i=1}^m \left(\frac{x_{ij} - \bar{x}_j}{s_j} \right) \left(\frac{x_{ij} - \bar{x}_j}{s_j} \right) = \text{cor}(j, j) \quad (2.7)$$

II.4 Calcul d'une ACP avec MATLAB



Algorithme de base de la méthode ACP

Conclusion II

Dans ce chapitre nous avons présenté la technique utilisée dans la classification des éléments du tableau périodique, on a montré que la méthode de l'ACP permet de trouver un axe u , issu d'une combinaison linéaire des X_n , tel que la variance du nuage autour de cet axe soit maximale. Les nouvelles variables, appelées composantes principales (PC), sont non corrélées, et sont ordonnées par fraction de l'information totale que chacune contient [1-5].

Références II

- [1] H. Hotelling, "Analysis of a complex of statistical variables into principal components", *Journal of Educational Psychology*, vol.24, pp.417-441, 1933.
- [2] L. Ferre, "Selection of components in principal component analysis: A comparison of methods", *Computational Statistics and Data Analysis*, pp.669-682, 1995.
- [3] Minami K., Nakajima H., Mark, R., and Toyoshima, T., "Realtime discrimination of ventricular tachyarrhythmia with fourier-transform neural network", *IEEE Trans. Biomed. Eng.*, vol. 46, no 2, 1999.
- [4] Jolliffe I.T., "Principal component analysis", Springer-Verlag, New York, 1986
- [5] 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.
- [6] Pieter Adriaans and Dolf Zantinge _ *Data Mining* _ Addison _ Wesley, 1996.
- [7] Jean-Michel Franco and EDS-Institut Prometheus. *Le Data Warehouse. Le Data Mining.* yrolles, 1997.
- [8] J- Han and M-Kamber. *Data Mining: Concepts and Techniques.* Morgan Kaufmann, 2000.
- [9] Rapport, 'Statistiques et Probabilités', Tome 2, A.I.D.E.P, Dunod.

Chapitre III

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

III.1. Introduction

Dans ce chapitre nous allons présenter les résultats graphiques des corrélations entre les individus, obtenues à partir des calculs numériques sous Matlab effectués sur la base de la théorie présentée dans le chapitre précédent.

La première étape consiste à déterminer la signification physique des axes de projection (PC1, PC2). La seconde étape consiste en l'interprétation des corrélations entre les éléments, qui peut être tirée des deux diagrammes : loadings, ou diagramme des variables, et scores, ou diagramme des propriétés.

III.2. Interprétation des résultats

La figure 1 présente le diagramme des propriétés (loadings plots). Il permet de déceler les corrélations entre les 22 propriétés physiques des éléments, que nous avons noté Var1... Var22 pour une meilleure lisibilité du schéma. La liste des variables est la suivante :

Var1 : Nombre atomique Z	Var12 : La constante d'élasticité S_{12}
Var2 : Masse atomique	Var13 : Le module d'élasticité C_{11}
Var3 : Les états de l'oxydation	Var14 : Le module d'élasticité C_{44}
Var4 : L'électronégativité	Var15 : La conductivité thermique
Var5 : L'énergie de 1 ^{ère} l'ionisation	Var16 : La capacité molaire C_p
Var6 : Le paramètre réseau	Var17 : L'entropie standard S_0
Var7 : Le nombre d'atomes par cellule	Var18 : La différence Enthalpie $H_{298}-H_0$
Var8 : Le volume molaire	Var19 : La température de fusion
Var9 : Le module de Young	Var20 : Le changement d'enthalpie à la fusion
Var10 : La constante d'élasticité S_{11}	Var21 : La température d'ébullition
Var11 : La constante d'élasticité S_{44}	Var22 : Le changement d'enthalpie à l'ébullition

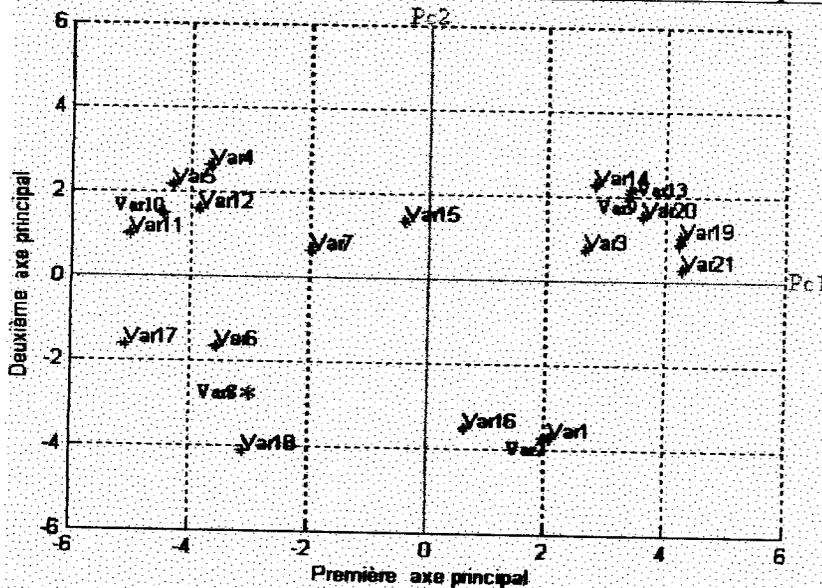


Figure 1 Corrélation entre les propriétés physiques – Scores plot

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

Nous pouvons remarquer un certain nombre de corrélations et d'anti corrélations. Certaines de ces corrélations sont triviales et assez simples à expliquer. D'autres sont le reflet de principes et relations physiques intrinsèques.

Dans la liste des corrélations, nous trouvons :

- Var1 corrélé à Var2 : le nombre atomique est corrélé à la masse atomique. Ceci est assez normal du moment que la masse atomique comprend la somme des masses des neutrons et des protons. Le nombre de ces derniers étant justement le nombre atomique Z .
- Var 1 et Var2 corrélés à Var 16 : c'est à dire le nombre et la masse atomiques sont corrélés à la capacité calorifique. Ceci peut s'expliquer comme suit : La capacité calorifique exprime l'aptitude du matériau à assimiler la chaleur. Le mode principale d'assimilation étant les phonons, ou vibrations du réseau. Autant la masse atomique est importante, autant il faudra de l'énergie pour exciter les phonons ; autrement dit autant la chaleur absorbée pour une augmentation de température donnée est plus importante.
- Var21 est corrélé à Var 19 : la température de fusion est corrélée à la température d'ébullition. Ceci est encore trivial du moment que l'ébullition se produit après la fusion.
- Var19 et Var 21 corrélés à Var20 : les températures de fusion et d'ébullition sont corrélées au changement d'enthalpie à la fusion. Cette dernière grandeur exprime l'énergie qu'il faut fournir pour que le matériau passe complètement de l'état solide à l'état liquide. Elle est de plus en plus importante que les liaisons interatomiques sont plus fortes. Ceci explique sa corrélation avec la température de fusion qui est toujours plus importante lorsque les liaisons interatomiques sont plus rigides.
- Var13 et Var14 corrélés à Var 9 : les constantes élastiques C_{11} et C_{12} sont corrélées au module de Young. Les deux constantes expriment la réponse du matériau à une déformation à l'échelle atomique. Elles sont autant plus importantes que les liaisons interatomiques sont plus fortes. Le module de Young est la réponse macroscopique du matériau à une déformation de traction. Indiscutablement le comportement mécanique macroscopique est une expression du déplacement des atomes à plus petite échelle lors de l'application d'une contrainte externe.
- Var4 corrélé à Var5 : l'électronégativité est corrélé à l'énergie de 1^{ère} l'ionisation. Un matériau électronégatif a tendance à attirer les électrons des autres atomes dans une liaison. A plus forte raison, les électrons de sa dernière couche sont fortement liés et ne sont que très difficilement cédés. Ceci s'exprime par une énergie d'ionisation élevée.
- Var6 corrélé à Var8 : le paramètre du réseau est corrélé au volume molaire. Ceci est tout à fait normal vue que le volume est fonction de l'espacement interatomique.

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

- Var17 corrélé à Var6 et Var8 : l'entropie est corrélée au paramètre du réseau et par conséquent au volume molaire. Ceci implique que l'augmentation du volume molaire augmente le désordre dans le système.
- Var10 corrélé à Var11 et Var12 : les constantes d'élasticité S_{11} , S_{12} et S_{44} sont toutes corrélées. Ce fait est cohérent du moment que toutes ces grandeurs sont liées directement à la manière dont varie la structure lorsqu'une contrainte externe est appliquée. Autrement dit, elles découlent toutes de la liaison interatomique.

D'un autre côté, nous trouvons les anti-corrélations :

- Var14 inversement corrélés à Var6 et Var8 : le module de Young est inversement corrélé au paramètre du réseau et au volume molaire. On sait d'expérience que le module de Young est proportionnel au module de compression qui lui est mesure de la compressibilité du réseau. Donc autant il est important autant c'est difficile de comprimer le réseau. Lorsque le paramètre du réseau et donc le volume molaire diminuent, le volume par atome diminue également, et donc la possibilité de rapprocher les atomes diminue. Ceci explique qu'autant le paramètre de réseau diminue autant le module de compression, et donc le module de Young augmente.
- Var1 et Var2 inversement corrélés à Var4 et Var5 : le nombre et la masse atomiques sont inversement proportionnels à l'électronégativité et l'énergie de première ionisation. Autant le nombre d'électrons augmente dans l'atome, autant la distance entre les électrons de valence et le noyau augmente, et donc autant ces électrons sont moins liés et l'énergie nécessaire pour les arracher (énergie de première ionisation) diminue. Aussi, autant le nombre d'électrons augmente, autant l'atome est moins électronégative car son noyau est de plus en plus écranté par les électrons et n'attire plus fortement les électrons des atomes voisins.

La figure 2 représente le diagramme des variables, c'est-à-dire les éléments dans ce cas (Loading plot). Il permet de voir comment sont distribués les éléments sur les axes PC1 et PC2 lorsque toutes leurs propriétés sont prises en compte.

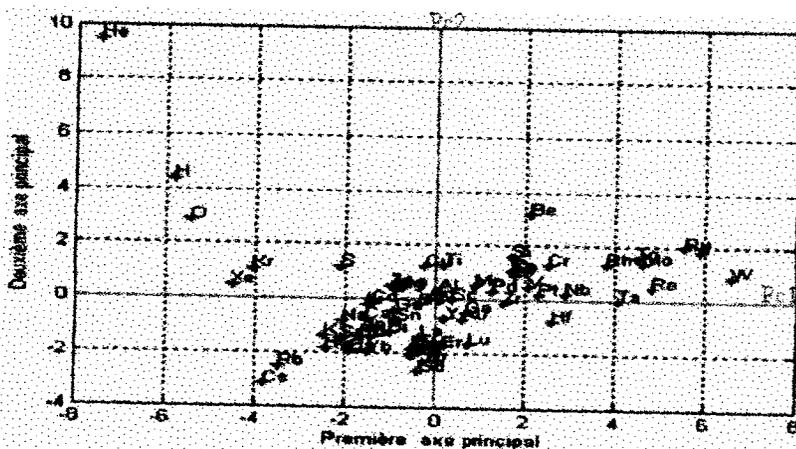


Figure 2 Corrélation entre les éléments – Loading plot

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

Vue sous cette forme, la classification des éléments n'est pas très constructive. En effet, autant le diagramme des propriétés reflète un grand nombre de corrélations physiques, autant le diagramme des variables semble condensé et peu lisible.

Nous pouvons voir que l'hélium, l'hydrogène, l'oxygène et quelques gaz rares (Xe, Kr) sont bien visibles sur le côté du diagramme. Aussi, le tungstène occupe une position particulière et se distingue avec quelques éléments métalliques (Ru, Rh, Ta...). Le reste des éléments, c'est-à-dire la majeure partie sont concentrés au centre du diagramme au point que la lisibilité est quasi nulle.

Pour y voir plus clair, nous avons tracé ce diagramme d'une manière astucieusement différente. La figure 3 est la représentation tridimensionnelle du diagramme des variables sur laquelle chaque élément est représenté par un cube plus ou moins éloigné dans la direction perpendiculaire pour permettre d'apercevoir tous les éléments.

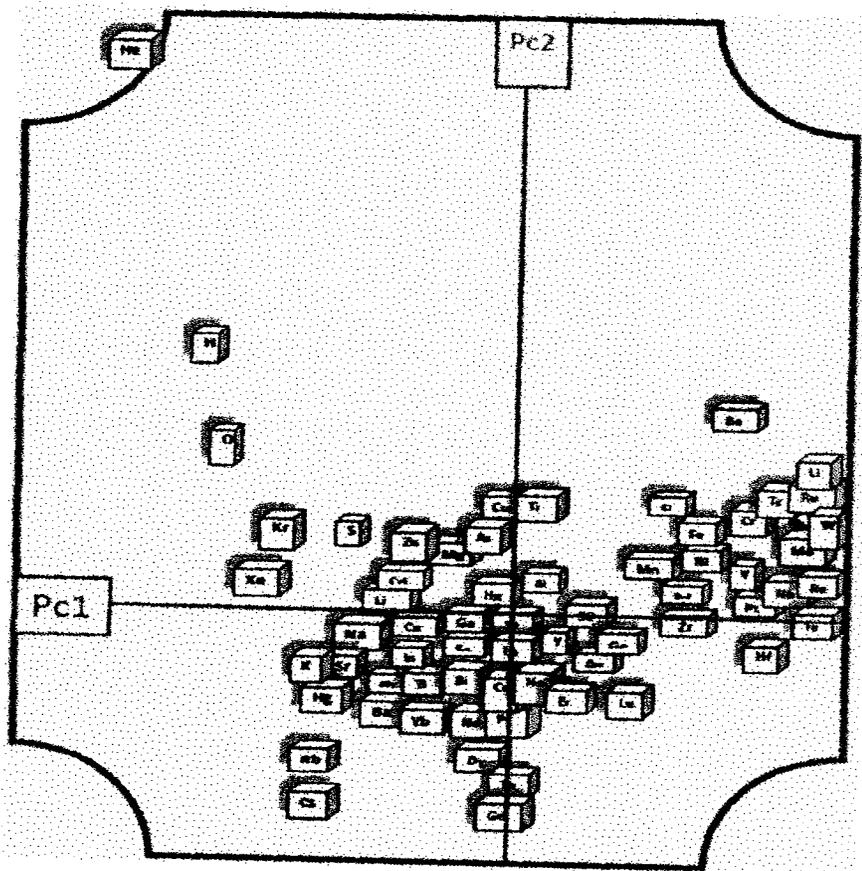


Figure 3 Représentation tridimensionnelle du diagramme des variables

Bien que la figure soit suffisamment claire, il reste à noter que la distribution des éléments dans le graphe ne suit pas une logique particulière. Il est très probable que le grand nombre de propriétés utilisées dans notre analyse est responsable de la variabilité des résultats et du fait que cette classification des éléments, aussi originale soit elle n'est pas très avantageuse.

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

Suite à cette introspection, nous avons décidé de refaire une analyse sur l'ensemble des éléments de la classification périodique en incluant cette fois dans notre base de données un ensemble plus restreint de propriétés et qui sont des propriétés intrinsèques de base. Plus particulièrement, nous avons inclus cette fois : l'énergie totale de l'atome, l'énergie cinétique des électrons, l'énergie de l'interaction coulombiennes et l'énergie d'échange et corrélation du système électronique.

III.3. Corrélation des énergies du système atomique et électronique

III.3.1. Définitions

Un système solide est un ensemble d'atomes composés de noyaux et d'électrons. Leurs interactions sont décrites par l'équation de Schrödinger qui doit être résolue suivant les lois de la mécanique quantique.

Cette résolution étant difficile, on a trouvé des moyens de la résoudre pour un système solide cristallin avec un ensemble d'approximations. Ces approximations s'inscrivent toutes dans le cadre de la théorie de la fonctionnelle de densité DFT.

L'énergie totale du système est calculée en évaluant l'énergie d'interaction coulombienne, l'énergie cinétique des électrons et un terme d'énergie dite d'échange et corrélation qui est ajouté pour restituer l'énergie mise en œuvre par les électrons pour respecter le principe de Pauli (énergie d'échange) et l'énergie des interactions complexes entre les électrons et qui ne sont pas incluses dans le terme coulombien (énergie de corrélation).

III.3.2. La Base de données :

Les valeurs de ces énergies pour l'ensemble des éléments du tableau périodique ont été déterminées par l'Institut National des Standards et Technologies (NIST) aux états unies et sont disponibles pour consultations sur le web.

Nous avons collecté dans notre base de données les énergies totale, coulombienne, cinétique et d'échange et corrélation pour 93 éléments allant de l'hydrogène à l'uranium. Le tableau 1 donne les valeurs que nous avons utilisées et qui sont issue de la référence [2].

Tableau 1 Base de données des énergies électroniques [2]

Élément	Energies totales	Energies cinétiques	Energies coulombiennes	Energies d'échange et corrélation
H	-0.445671	0.425027	0.282827	-0.920999
He	-2.834836	2.767922	1.99612	-6.625564
Li	-7.335195	7.237484	3.990755	-16.911041
Be	-14.447209	14.309424	7.115257	-33.357034
B	-24.344198	24.161047	11.503002	-56.484587
C	-37.425749	37.190391	17.627997	-87.515412
N	-54.025016	53.731727	25.799845	-127.416229
O	-74.473077	74.116881	36.331102	-177.152578

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

F	-99.099648	98.676131	49.535385	-237.688846
Ne	-128.233481	127.738667	65.726488	-309.988206
Na	-161.44006	160.896571	79.732497	-388.546372
Mg	-199.139406	198.541505	95.67329	-477.899149
Al	-241.315573	240.663489	112.670733	-577.205757
Si	-288.198397	287.48774	131.767814	-687.900675
P	-339.946219	339.173459	153.065007	-810.4001
S	-396.716081	395.878143	176.68795	-945.141155
Cl	-458.664179	457.758293	202.772337	-1092.568993
Ar	-525.946195	524.969812	231.458124	-1253.131982
K	-598.20059	597.17213	257.381524	-1421.104115
Ca	-675.742283	674.657334	285.20613	-1601.463209
Sc	-758.679275	757.5252	318.555762	-1797.853811
Ti	-847.277216	846.050533	355.232355	-2008.740849
V	-941.678904	940.377057	395.162964	-2234.347114
Cr	-1042.030238	1040.642716	441.867987	-2478.305541
Mn	-1148.449372	1146.990867	485.085158	-2731.133992
Fe	-1261.093056	1259.553429	535.295832	-3003.082484
Co	-1380.091264	1378.468819	589.169216	-3291.260194
Ni	-1505.580197	1503.873332	646.833302	-3596.066045
Cu	-1637.785861	1635.9828	713.928477	-3923.329381
Zn	-1776.57385	1774.693645	774.056522	-4257.166094
Ga	-1921.846456	1919.894783	833.754978	-4603.576824
Ge	-2073.807332	2071.782093	895.945638	-4965.789424
As	-2232.534978	2230.434545	960.716547	-5344.04275
Se	-2398.11144	2395.934307	1028.192719	-5738.619928
Br	-2570.6207	2568.365427	1098.492305	-6149.802515
Kr	-2750.14794	2747.813142	1171.723688	-6577.865761
Rb	-2936.337293	2933.940454	1239.168864	-7013.68955
Sr	-3129.453161	3126.990706	1308.954372	-7465.641817
Y	-3329.520604	3326.984254	1384.603788	-7937.187681
Zr	-3536.737751	3534.12421	1464.236106	-8426.900586
Nb	-3751.196175	3748.49657	1551.501122	-8938.494135
Mo	-3973.013235	3970.231601	1638.977454	-9465.024542
Tc	-4202.188857	4199.331793	1724.900741	-10004.79062
Ru	-4438.981228	4436.031047	1824.80092	-10573.32364
Rh	-4683.301031	4680.264572	1923.261557	-11155.54306
Pd	-4935.368406	4932.234516	2031.645838	-11762.87581
Ag	-5195.031215	5191.8186	2131.618463	-12377.29448
Cd	-5462.390982	5459.099922	2233.419308	-13008.90242
In	-5737.309064	5733.94412	2334.432447	-13654.87065
Sn	-6019.953353	6016.512431	2438.225987	-14319.01448

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

Sb	-6310.376268	6306.85786	2544.726589	-15001.36474
Te	-6608.631413	6605.034216	2653.969369	-15702.06148
I	-6914.773092	6911.09591	2766.012734	-16421.26955
Xe	-7228.856107	7225.097817	2880.919348	-17159.15943
Cs	-7550.55771	7546.734488	2988.242134	-17904.90606
Ba	-7880.111578	7876.220134	3098.114585	-18668.85311
La	-8217.57523	8213.607378	3215.018109	-19455.49476
Ce	-8563.360285	8559.301423	3351.833535	-20278.17841
Pr	-8917.664369	8913.501128	3508.213147	-21136.96521
Nd	-9280.311037	9276.05444	3654.948015	-22003.0647
Pm	-9651.484134	9647.133137	3806.516511	-22890.9423
Sm	-10031.25909	10026.81273	3962.921783	-23800.75332
Eu	-10419.71078	10415.16816	4124.189856	-24732.67422
Gd	-10816.65388	10812.02673	4273.725204	-25670.16534
Tb	-11222.94198	11218.20435	4461.481521	-26663.60914
Dy	-11637.86966	11633.03335	4637.607532	-27663.02195
Ho	-12061.77055	12056.8348	4818.79662	-28685.33803
Er	-12494.7183	12489.6824	5005.110005	-29730.76571
Tm	-12936.78649	12931.64974	5196.611124	-30799.51525
Yb	-13388.04859	13382.8103	5393.365097	-31891.7984
Lu	-13848.23038	13842.90353	5575.225315	-32987.4068
Hf	-14317.49372	14312.0769	5760.587419	-34104.62931
Ta	-14795.8842	14790.37645	5949.357668	-35243.46006
W	-15283.44882	15277.84931	6141.634533	-36404.0875
Re	-15780.23602	15774.54396	6337.535878	-37586.72292
Os	-16286.29541	16280.51003	6537.173416	-38791.57439
Ir	-16801.67747	16795.79805	6740.648988	-40018.84238
Pt	-17326.57638	17320.59285	6958.348706	-41279.08998
Au	-17860.79094	17854.71137	7170.546095	-42552.58722
Hg	-18404.27422	18398.10839	7374.994996	-43837.04718
Tl	-18956.95763	18950.71039	7576.942724	-45137.45422
Pb	-19518.99315	19512.66268	7781.994459	-46459.62774
Bi	-20090.41445	20083.99952	7989.999458	-47803.48411
Po	-20671.25654	20664.75611	8200.970142	-49169.10374
At	-21261.55522	21254.96835	8414.956548	-50556.60634
Rn	-21861.34687	21854.6727	8632.016039	-51966.12038
Fr	-22470.31965	22463.57381	8837.303934	-53382.42581
Ra	-23088.68808	23081.86753	9045.587062	-54820.47465
Ac	-23716.49695	23709.59435	9262.611883	-56286.01651
Th	-24353.83223	24346.84566	9483.654939	-57774.56758
Pa	-25001.29138	24994.2006	9743.171707	-59321.16943
U	-25658.41789	25651.23118	9991.594178	-60876.21062

III.3.3. Analyse

La figure 4 représente le diagramme des propriétés (score plot) obtenu avec la base de données présentée plus haut.

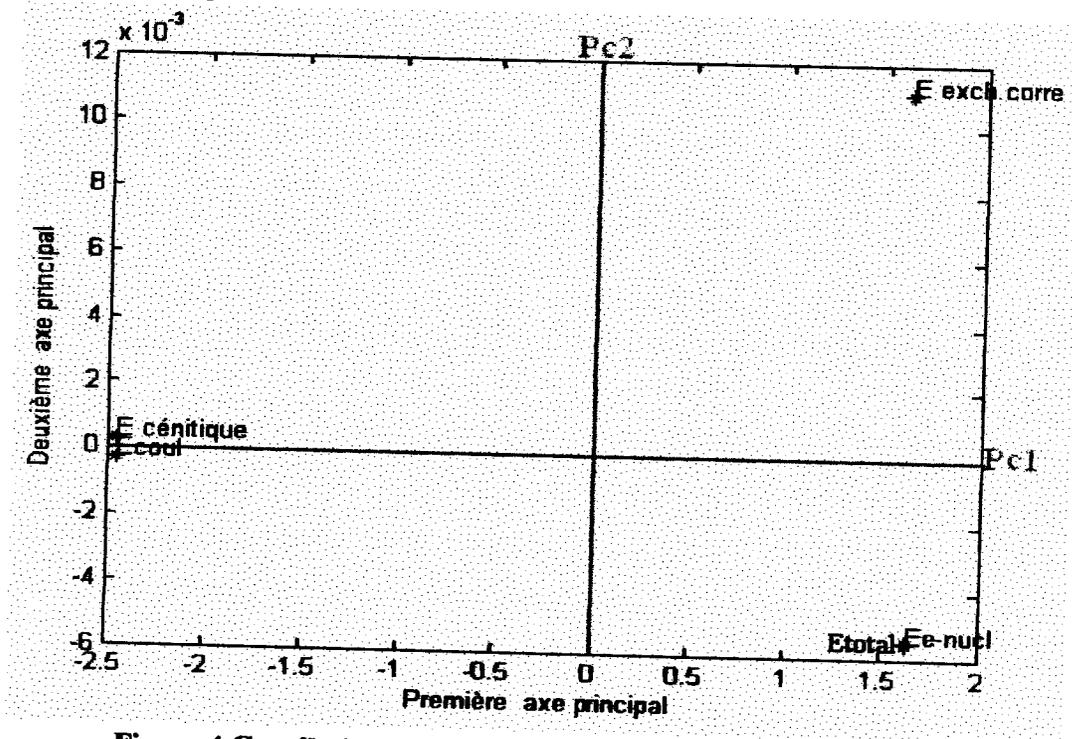


Figure 4 Corrélation entre les propriétés physiques – Scores plot

Nous pouvons voir que les énergies sont réparties pratiquement uniquement suivant l'axe PC1. L'énergie d'interaction électrostatique électron noyau et l'énergie totale sont du côté positif de PC1 alors que l'énergie cinétique et l'énergie d'interaction électrostatique électron-électron sont du côté négatif. L'énergie d'échange et corrélation est également du côté droit, mais avec une petite composante suivant PC2.

La figure 5 représente le diagramme des variables (loadings plot) qui comporte l'ensemble des éléments du tableau périodique. A cause du nombre élevé des éléments, nous sommes amenés à zoomer sur quelques zones pour mieux observer la disposition des éléments.

Nous pouvons voir sur le graphe que les éléments du tableau périodique sont alignés suivant leur numéro atomique, à partir de l'hydrogène et jusqu'à l'uranium dans le sens inverse de l'axe PC1. Aussi, suivant PC2, ils forment une sorte de cloche vers la direction positive de PC2. Nous pouvons expliquer ceci comme suit : la direction négative de l'axe PC1 est corrélée avec l'énergie d'interaction électrons-électron et avec l'énergie cinétique des électrons, ainsi, les éléments avec plus d'électrons évoluent dans cette direction.

Pour la courbure de la distribution des éléments, nous pouvons dire que pour les éléments les plus lourds, l'énergie d'interaction coulombienne compte plus que l'énergie d'échange et corrélation car il y a trop d'électrons, et c'est également le cas des éléments les plus légers où il n'y a pas assez d'électrons pour que l'échange et corrélation soient importants. Pour les

Chapitre III : Datamining des propriétés physiques des éléments du tableau périodique

éléments du milieu, l'énergie d'échange et corrélation compte plus que l'énergie de coulomb, ce qui fait que les positions des éléments sont décalées vers le haut.

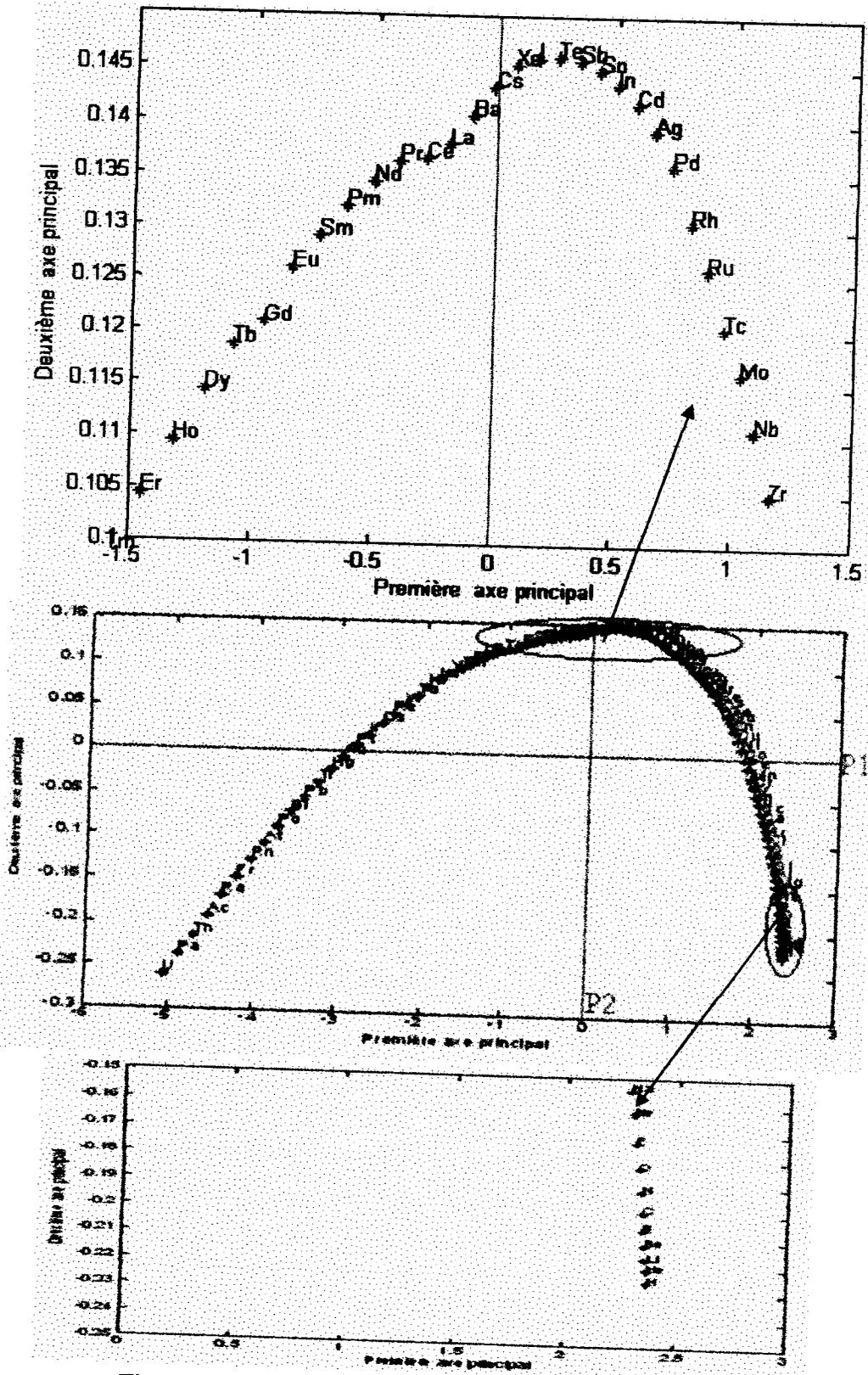


Figure 5 Corrélation entre les éléments – Loading plot

Conclusion III

- ✓ Nous avons présenté l'utilisation d'une méthode d'exploitation de données connus sous le nom de l'Analyse en Composantes Principales ACP.
- ✓ Nous avons constitué une base de données d'un grand nombre de propriétés physiques des éléments du tableau périodique.
- ✓ L'ACP nous a permis de visualiser cette base de données sur des graphiques et d'examiner les relations qui existent entre ces propriétés.
- ✓ Nous avons appliqué l'ACP sur une base de données des énergies totale, cinétique, électrostatique et d'échange et corrélation de tous les éléments du tableau périodique
- ✓ Les éléments lourds sont caractérisés par l'énergie cinétique et l'énergie d'interaction électron-électron.
- ✓ Les éléments légers sont caractérisés par L'énergie d'interaction électron noyau et l'énergie totale.
- ✓ l'énergie d'échange et corrélation a un impact particulier sur la distribution des éléments sur le loading plot.

Références III

- [1]. *W. Martienssen (Ed.): Numerical Data and Functional Relationships in Science and Technology, Landolt-örnstein, New Series III and IV (Springer, Berlin, Heidelberg 1970-2003).*
- [2]. *Svetlana Kotochigova, Zachary H. Levine, Eric L. Shirley*, Mark D. Stiles, and Charles W. Clark, " Atomic Reference Data for Electronic Structure Calculations", NIST Standard Reference Database 141 Inquiries or comments: charles.clark@nist.gov.*

Conclusion générale

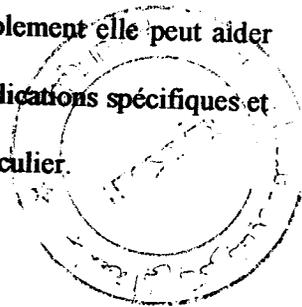
Le mémoire que nous présentons comprend l'utilisation de l'une des techniques de Datamining pour analyser des tableaux de données sur un ensemble de matériaux, en l'occurrence l'ensemble des éléments du tableau périodique.

L'idée de base était de repérer et représenter les corrélations qui existent entre les propriétés physiques des éléments, et éventuellement voir la classification des éléments sous un lumière différente que celle de la classification de Mendeleïev.

Nous avons rassembler une base de données comprenant certaines des propriétés cristallographiques, mécaniques, thermodynamiques, moléculaires, ioniques, atomiques, électromagnétiques, optiques et électroniques des éléments du tableau périodique. Cette base nous a servit pour appliquer la technique d'extraction de données dite ACP ; Analyse en Composantes Principales.

La technique d'analyse en composantes principales nous a permis de visualiser notre base de données sur des graphiques et d'examiner les relations qui existent entre les variables qui forment cette base de données, elle nous a également permis d'identifier des corrélations et des anti-corrélations entre les propriétés physiques.

Notre conclusion principale est que l'analyse en utilisant les outils de datamining peut être très utile pour percevoir les inter-relations entre les propriétés, et probablement elle peut aider significativement à analyser les informations dont on dispose en vue d'applications spécifiques et ce en dévoilant les bonnes questions à poser et les cas dignes d'intérêt particulier.



ANNEXE

Propriété .physique	Unité
Nombre atomique Z	
masse atomique	" g.mol ⁽⁻¹⁾ "
Les états de l'oxydation	
Electronégativité χ_a	
1ère énergie de l'ionisation	eV
Paramètre réseau	pm
Numéro d'atomes par cellule	
Le volume molaire	V.mol
Module élasticité E	GPa
S11 de l'acquiescement	1/TPa
S44 de l'acquiescement élastique	1/TPa1/TPa
S12 de l'acquiescement élastique	GPa
C11 de la raideur élastique	GPa
C44 de la raideur élastique	W/(mK)
La conductivité thermique λ	J/(molK)
La capacité molaire c p	J/(molK)
L'entropie standard S0	kJ/mol
La différence Enthalpie H298- H0	K
La température de fusion Tm	kJ/mol
Enthalpie changent ΔH_m	K
Température d'ébullition Tb	kJ/mol
Enthalpie changent ΔH_b	microHartree
Energies totales	microHartree
Energies cinétiques	microHartree
Energies coulombiennes	microHartree
Energies d'échange et corrélation	microHartree

1 microHartree = 0.03 meV = 0.0006 kcal/mole