Rev. Energ. Ren. Vol.5 (2002) 23-32

Hierarchical Classification of the Photovoltaic Materials by Data Analysis Method

M. Benosman, A. Zerga and B. Benyoucef

Laboratoire de Matériaux et des Energies Renouvelables Département de Physique - Faculté des Sciences
Université de Tlemcen - B.P : 119 Tlemcen, Algérie

Abstract – *In Photovoltaic, many semiconductors materials are used. These materials possess different optical, electrical and mechanical properties. However, the present technologies used in the solar cell manufacture, depend essentially for the material and impose an elevated price. Therefore, the choice of the material is very important and should be developed very finely.*
In this paper, we first present the theory of some important photovoltaic materials, than secondly, we developed a statistical analysis method. This method allows to class hierarchically the different photovoltaic materials (Si, GaAs, CdTe, CdSe, InP...).
In this method, the influence of the physical and technological parameters such as the bandgap, the carrier lifetime, the carrier mobility, the coefficient of reflection…are taken into account in order to analyze the related importance of each parameter on the hierarchical classification.
This method is based on the ordinal statistical analysis of the data, called Analysis in Principal Component (A.P.C). The photovoltaic data are taken from the recent literature and are represented on tables. A graphic visualization of this classification is presented while using the factorial analysis.

Résumé – *En photovoltaïque, plusieurs matériaux semi-conducteurs sont utilisés. Ces matériaux possèdent des propriétés optiques, électriques et mécaniques différentes. Cependant, les différentes techniques d'élaboration des cellules solaires dépendent essentiellement du matériau utilisé et le pris de revient. A ce fait, le choix du matériau est très important et doit être développé très finement. Dans ce travail, la première étape consiste dans l'étude théorique de quelques matériaux photovoltaïques, ainsi que leurs paramètres. Dans une deuxième démarche, on présente une méthode d'analyse statistique, nous permettant de faire une classification hiérarchique des différents matériaux photovoltaïques (Si, GaAs, CdTe, CdSe, InP,…) selon leurs caractéristiques et performances optimales.*
Dans cette optique, on prend en considération l'influence de quelques paramètres physiques et technologiques tels que la durée de vie des porteurs, l'énergie de gap, la mobilité, le coefficient de réflexion, afin d'examiner le degré d'importance de chaque paramètre sur le matériau utilisé.
La méthode utilisée est une méthode d'analyse ordinale de données appelée Analyse en Composantes Principales (A.C.P) souvent utilisée en analyse statistique. Les données photovoltaïques étudiées sont présentées sur des tableaux de nombres. L'analyse factorielle permet de mieux visualiser ces tableaux et de les présenter sous forme de graphique. Nous obtenons ainsi une hiérarchisation des matériaux selon leurs caractéristiques et performances optimales.

Keys words: Hierarchisation- Solar cell- Photovoltaic materials- Analysis in principal component method- Statistical analysis.

1. INTRODUCTION

The research on the production of the solar electricity is motivated both by improvement of the performances for different photovoltaic used materials and by a reduction of the technological realization cost. The present stake of the development for the photovoltaic path is to find a compromise between the best photovoltaic materials with a low price and high performances. Indeed, every material is characterized by typical properties, which can influence considerably the price and the performances of the solar cells. Consequently, the

classical method for the choice of optimal photovoltaic material is difficult. This can justify the interest to use the mathematical statistics method for a hierarchically classification of all photovoltaic materials.

The classical methods are based on linear models and can not provide an explicit diagram of all situations and their evolution, nevertheless another method should looked for. In the past, the statistical methods were intended to validate or estimate the numeric parameter values fixing the random variable distribution. These variables are governed by the natural phenomena with very complex mechanisms. Actually, we can investigate in recent methods where the part for the observation and the collection of data is very important.

In this study, we regroup all essential parameters of the photovoltaic materials, which influence the solar cell performance. The method "Analyses in Principal Components" is one method of the statistical analysis, which allows us to visualize the table of data and to present them in diagram. This method is based on the notion of the distance. Once these distances are calculated, we dispose the individuals (materials) on a diagram, in the same manner as the characters (photovoltaic data). This results in a classification of the different photovoltaic materials according to their characteristics and performances.

2. THEORY

In this work, we have developed numerical program, which allows us to represent the photovoltaic data in tables of numbers and in diagram. The algorithm of this program is based on the A.P.C. Method. However, we present many parameters influencing the solar cells performances.

2.1. Carrier mobility

It is defined by the following equation [1]:

$$\mu = V/E \tag{1}$$

V is the velocity of the carrier of charges and E is the electric field.

Table 1: Electron Mobility ($cm^2/V.sec$) at 300K as a function of impurity concentration.

Impurity concentration (cm^{-3})	Photovoltaic materials				
	Ge	Si	GaAs	CdTe	CdSe
10^{14}	4600	1900	7000	1000	620
10^{16}	3700	1200	6000	900	580
10^{18}	2000	320	3000	800	500

The temperature dependence of mobility as produced by impurity scattering is:

$$\mu \approx (m^*)^{-1/2}\, T^{3/2} / N_I \tag{2}$$

m is the mass of the electron or the hole, T is the temperature and N_I is the ionized impurity density. We calculate the carrier mobility according to the concentration of impurities for different photovoltaic materials. We get the following results, which are presented in Table 1 and in Table 2.

2.2. Carrier lifetime

It is defined by the following equation [2]:

$$\tau_p = \frac{1}{\sigma_p V_{th} N_r}\left[\left(1+\frac{N_c}{n_{n_0}}\right)\exp\left[-(E_c-E_r)/KT\right]+\left(\frac{\sigma_p}{\sigma_n}\right)\left(\frac{N_v}{n_{n_0}}\right)\exp\left[-(E_r-E_v)/KT\right]\right]^{-1} \quad (3)$$

Table 2: holes Mobility (cm²/V.sec) at 300K as a function of impurity concentration.

Impurity concentration (cm^{-3})	Semiconductors			
	Ge	Si	GaAs	CdTe
10^{14}	1900	600	400	100
10^{16}	1300	460	320	80
10^{18}	360	200	160	60

where n_{n0} is the free electron concentration in the n type of the semiconductor; V_{th} is the thermal velocity; E_c and E_v are respectively, the energy of the conduction and valence bands, σ_n and σ_p are the capture cross-section for electron and hole. Similar expressions can be obtained for the electron lifetime in the p type material semiconductor.
Then, the minority carrier lifetime becomes:

$$\begin{cases} \tau_p = \dfrac{1}{\sigma_p V_{th} N_r} \\ \tau_n = \dfrac{1}{\sigma_n V_{th} N_r} \end{cases} \quad (4)$$

Table 3 resumes the lifetime of electrons and holes for the studied semiconductors.

Table 3: Minority carrier lifetime (sec) in solar cells at low carrier concentration levels.

Semiconductors	Electrons Life time	Holes Life time
Ge	$0,5\ 10^{-4}$	$0,5\ 10^{-4}$
Si	$0,8\ 10^{-4}$	$0,8\ 10^{-4}$
Inp	$4\ 10^{-8}$	$4\ 10^{-8}$
GaAs	$4\ 10^{-8}$	$4\ 10^{-8}$
CdTe	$1,6\ 10^{-7}$	$1,6\ 10^{-7}$
AlSb	$9\ 10^{-8}$	$9\ 10^{-8}$
CdSe	$1,6\ 10^{-8}$	$1,6\ 10^{-8}$
Gap	$4\ 10^{-8}$	$4\ 10^{-8}$

2.3. Reflection

The fraction of light reflected on the surface of the material upon which the light is normally incident, is given by:

$$R=\frac{(n_2-n_1)^2+k_2^2}{(n_2+n_1)^2+k_2^2} \tag{5}$$

where (n_1, n_2) are the real index of refraction and (k_1, k_2) are respectively the extinction coefficient for air and photovoltaic material. The values of the reflection coefficients for the photovoltaic materials used in this work are given in Table 4.

Table 4: Reflection coefficient of some photovoltaic materials [3].

Materials	Ge	Si	InP	GaAs	CdTe	AlSb	CdSe	Gap
R	0,36	0,30	0,262	0,286	0,207	0,272	0,184	0,294

2.4. Performances

In order to establish the hierarchical classification, it is necessary to calculate the optimal performances for different photovoltaic materials. The short circuit current, the open circuit voltage, the fill factor and the efficiency, conventionally define these performances. The results we obtain by theoretical calculus are presented on the Table 5.

Table 5: Representation of studied material performances.

Semiconductors	Ge	Si [4]	Inp [5]	GaAs	CdTe[6]	CdSe	Gap
I_{CC} (mA/cm²)	16	40,7	24,8	29,7	25	84	29
V_{C0} (V)	0,18	0,69	0,86	0,88	0,84	0,35	1,39
FF(%)	57	95	83	82	74	73	86
η(%)	10	22	17,9	24	15,8	16	17,4

However for the purpose of this study using Analysis in Principal Components Method, it is required to regroup all parameters influencing solar cells. These are resumed in Table 6.

3. DESCRIPTION OF THE APC METHOD

The APC method consists in establishing a hierarchical classification for different materials relatively to their characteristics and their optimal performances (Table 6). This method is based on a measure of a distance d between two elements i and i′ of our data resumed in table T, where d is expressed by:

$$d^2(i,i')=\sum_{1\le j\le m}\left(X_{ij}-X_{i'j}\right)^2 \tag{6}$$

In the data given in Table T, the lines are indexed by i and the columns by V. The size of this table is (n,m) for respectively lines and columns. The magnitudes Xij represent the measure of the j^{th} variable of V on the i^{th} element of I.

In the A.P.C. method, the contribution of a variable to a distance d depends on the dispersion for its values. This dispersion is bound to the nature of the variable. In our numerical code, we calculate these distances for all couples of the population. Then, we represent the individuals (materials) and the characters (parameters) on a diagram, in order to respect these distances. However, we can obtain any scheme, which suitably represent the calculated distances. This provides any curve defined only by individuals of the population and their distances. In summary, we calculate the eigen values and the eigen vectors of any matrix.

Table 6: Representation of all influencing parameters at 300K

Photovoltaic materials	Ge	Si	InP	GaAs	CdTe	CdSe	GaP
V_1:Gap energy (e.V)	0.67	1.106	1.29	1.39	1.45	1.74	2.24
Density of mass V_2: Band of conduction V_3: Band of valence	 0.55 0.31	 1.08 0.56	 0.07 0.4	 0.072 0.5	 0.11 0.35	 0.13 0.4	 0.13 0.8
Density of state x 10^{19} cm^{-3} V_4 :Band of conduction V_5:Band of valence	 1.03 0.43	 2.82 1.04	 0.465 0.636	 0.465 0.888	 0.0917 0.520	 0.118 0.636	 0.118 1.80
V_6 :Intrinsic concentration	2.4 10^{13}	1.45 10^{10}	2.63 10^{7}	4.5 10^{6}	1.52 10^{6}	7.07 10^{3}	0.764
V_7: Electron mobility (cm^2/V.s)-$N_D=10^{14}/cm^3$	3950	1500	4000	8500	1000	600	120
V_8 : Hole mobility (cm^2/V.s)-$N_A=10^{14}/cm^3$	1900	600	650	400	100	-	100
V_9 : Electron diffusion coefficient (cm^2/s)	102.2	38.81	103.5	219.94	25.87	15.25	3.10
V_{10} : Hole diffusion coefficient (cm^2/s)	49.16	15.52	16.81	10.35	2.58	-	2.58
V_{11}: Reflection coefficient	0.36	0.30	0.262	0.286	0.207	0.184	0.294
V_{12}: Electron life time (s)	5 10^{-5}	8 10^{-5}	4 10^{-8}	4 10^{-8}	16 10^{-8}	16 10^{-9}	4 10^{-8}
V_{13}: Hole life time (s)	5 10^{-5}	8 10^{-5}	4 10^{-8}	4 10^{-8}	16 10^{-8}	16 10^{-9}	4 10^{-9}
V_{14} : Electron length diffusion (cm)	32 10^{-4}	34 10^{-4}	10.1 10^{-4}	14.8 10^{-4}	50.8 10^{-5}	39.4 10^{-5}	15.8 10^{-6}
V_{15} : Hole length diffusion (cm)	22 10^{-4}	21 10^{-4}	41 10^{-5}	32.1 10^{-5}	16 10^{-5}	-	16 10^{-6}
V_{16} : Current of short-circuit (mA/cm^2)	16	40.7	24.8	29.7	25	84	29
V_{17} : Tension of open circuit (V)	0.18	0.69	0.86	0.88	0.84	0.35	1.39
V_{18} : Fill factor	0.57	0.95	0.83	0.82	0.74	0.73	0.86
V_{19} : Efficiency	0.10	0.22	0.179	0.24	0.158	0.16	0.17

This matrix is associated to the calculated distances between the individuals of the population. The plane is determined by the two eigen vectors corresponding to the two greater eigen values and provides the scheme which check the best calculated distances [7,8].

In general, it is impossible to plot a diagram, which appropriately checks the calculated distances, for example, the problem in representing a pyramid on a plane surface. However, we can found a diagram, which follows the best-calculated distance, and distort less possible, the figure defined by the individuals of the population and their distances [9].

In addition, this figure is not realizable only in the space with more than three dimensions, where it is impossible to represent it physically. Therefore, it is necessary to have a simplified scheme. This method allows us to visualize the table of data and to present them on diagrams. Indeed, the two individuals with the same " profile " across the variables of description found at the same point on the diagram. The set of all these points representing these individuals or characters form " clusters ". This cluster is particularly shaped. The calculations define the axes of " larger elongation " on which will be carried out the projection. This calculation will provide the cluster with the best plane representation. It is this projection which we have analyzed.

However, it is necessary to note that, the two axes represented on the diagram do not respectively correspond to the characters and the individuals. Those axes are the result of complex calculation and must be globally interpreted. It necessary to note that this method requires the calculation of the correlation coefficient ρ [10] between two vectors. The coefficient is a quantified evaluation of the importance that each parameter has on the used material. The correlation coefficient is given by the following equation:

$$\rho = \frac{\mathrm{cov}(x, y)}{\sigma_x \, \sigma_y} \tag{7}$$

where σ_x, σ_y are type distance of the vectors x and y respectively and cov(x, y) is the covariance between vectors x and y respectively.

4. RESULTS AND DISCUSSION

First, we present some graphical representations on individuals and characters. These graphical representations allow us to visualize the data given in the previous table T. We get the following results:

For the variation of the electron mobility according to the concentration of the impurities (Table 1), the representation of characters is presented in Figure 1.

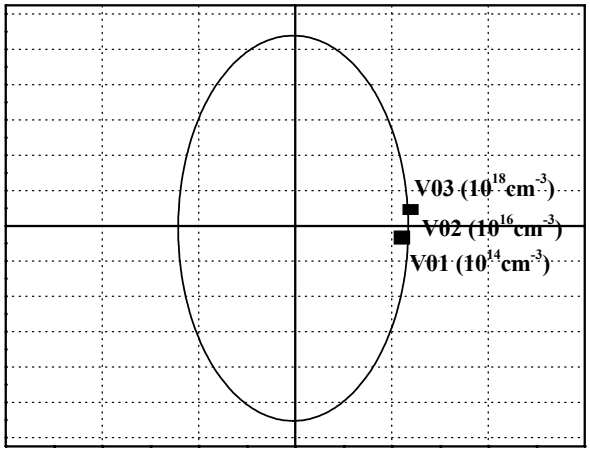

Fig. 1: Variation of the electron mobility according to the impurity concentration. Presentation of the characters in the principal plan.

In this figure, we note that the two vectors V_{01} and V_{02} respectively corresponding to the impurity concentrations 10^{16} cm^{-3} and 10^{18} cm^{-3}, have a correlation coefficient of ρ=0,995848. This shows that they are presented on the same point where the importance of the variation of the mobility in this range of impurity concentration.
For the same case, the representation of individuals (materials) is presented on Figure 2.

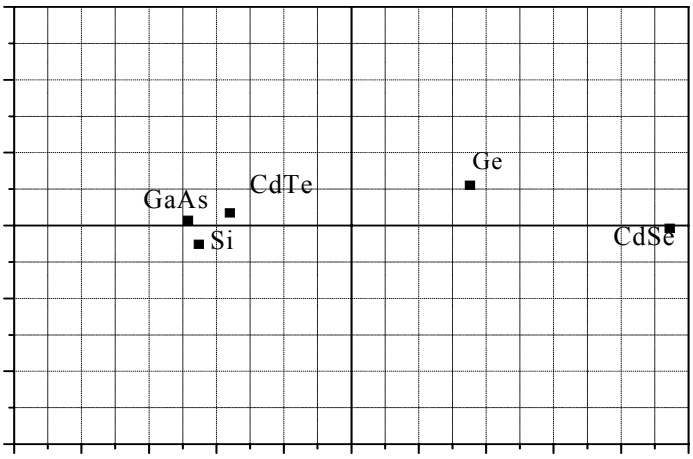

Fig. 2: Variation of the electron mobility Vs the impurity concentration
Presentation of the individuals(materials) in the principal plan.

Figure 2 shows that the three materials Si, GaAs and CdTe are independent of the other points and are regrouped on the diagram. This explains that these three materials have a strong presence on the diagram. Therefore, for this variation of electron mobility according to the impurity concentration, the three cited materials are better than the two others Ge and CdSe.
In the case of hole mobility variation according to the impurity concentration (Table 2), the presentation of characters is shown on figure 03 and the individuals in Figure 4.

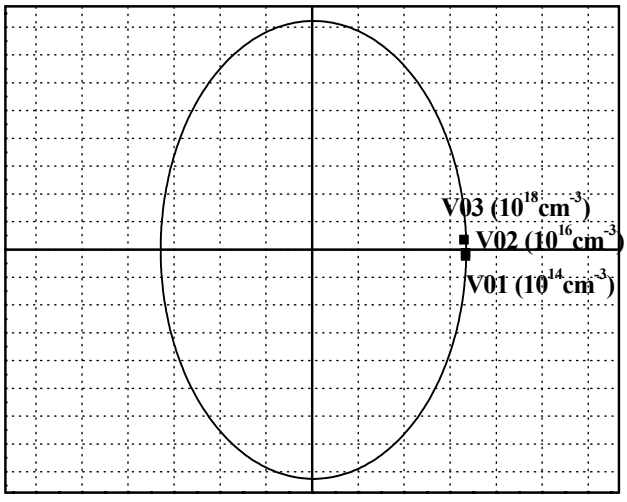

Fig. 3: Variation of holes mobility Vs the impurity concentration.
Presentation of the characters in the principal plan.

Similarly as the above study, the variation of the hole mobility according to the impurity concentration (Figure 3), the coefficient of correlation between the two vectors V_{01} and V_{02} respectively corresponding to concentrations of 10^{14} cm^{-3} and $10^{16}cm^{-3}$, is equal to 0,999211.

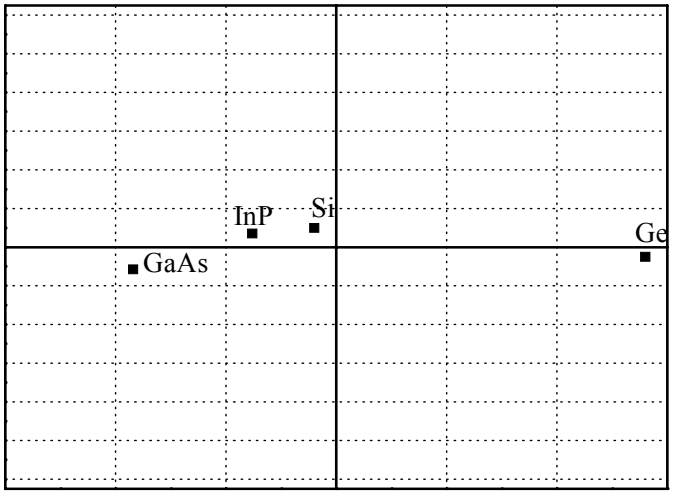

Fig. 4:Variation of holes mobility Vs the impurity concentration.
Presentation of the individuals in the principal plan.

This result leads to the same conclusion as given in Figure 1. In the representation of the individuals (materials), we observe the presence of the Si near the critical point, which represent the origin. In a first conclusion, we can consider the Si as the most important photovoltaic material. Therefore we can classify the different materials according to their hole mobility as follows: Si, GaAs, CdTe and Ge. To establish the hierarchical classification, it is necessary to present the parameters and the performances on the same diagram (Table 6).
Figures 5 and 6 show the classification of materials according to their performances.

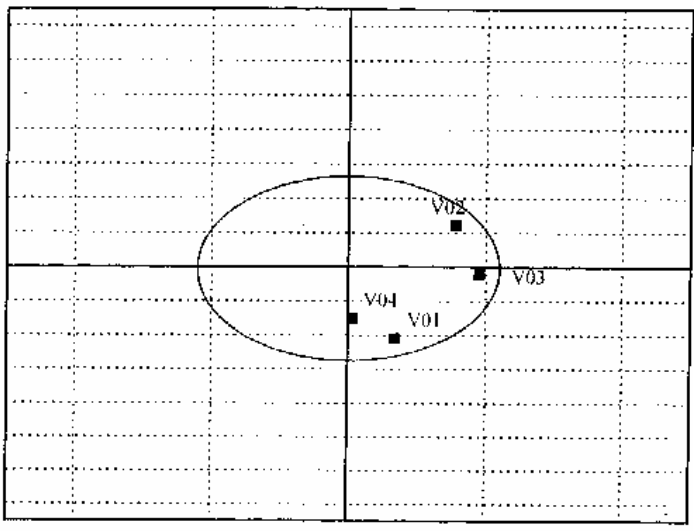

Fig. 5: Representation of the performances on the principal plan

Figure 5 shows the importance of the four performances of the studied photovoltaic materials according to Table 5. We note that the important parameter is the efficiency. This parameter is represented by the vector V_{04} and plotted on the vertical axis. In the statistical analysis, this case permit us to distinguish between individuals or characters, and gives more information on the individual (studied material).

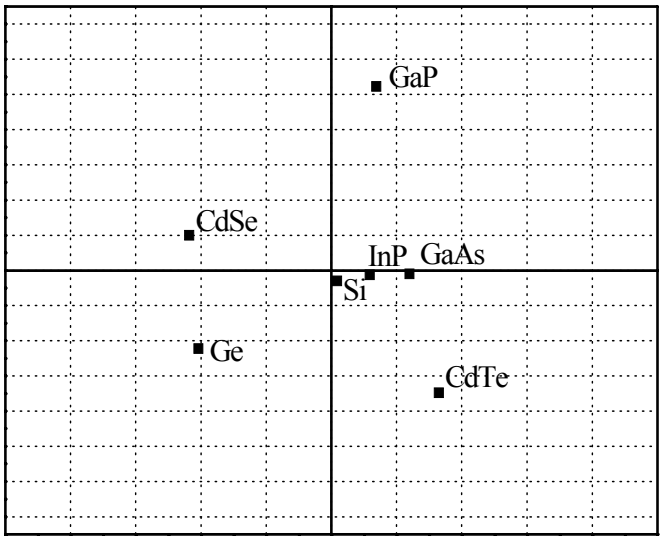

Fig. 6: Representation of individual on the principal plan according to the performance.

Figure 6 show that the Si, InP and GaAs materials, are plotted on the horizontal axis, near to the critical point. They are also better regrouped than the other materials. Their distribution is more uniform. We can therefore deduce that the most interesting material is the silicon. Although, the GaAs has a greater efficiency however it cost is higher.
In the previous calculations, we have established a hierarchical classification of the most advanced photovoltaic materials according to their characteristic and performances (Table 6).

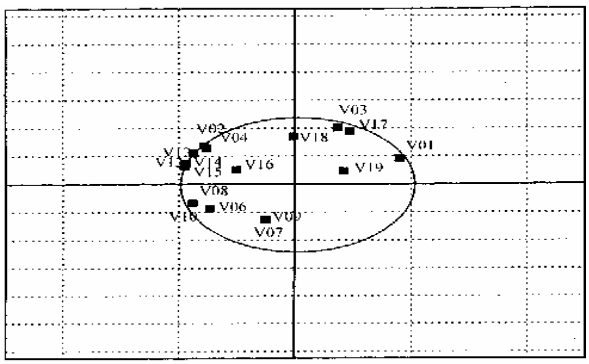

Fig. 7: Representation of the all characters on the principal plan

Figure 7 shows that the vectors V_{12}, V_{13}, V_{14}, V_{15} are respectively corresponding to the electron lifetime, hole lifetime, the electron diffusion length and the hole diffusion length, are very closely related. We note a predominant performances of the photovoltaic cell represented by vectors V_{16}, V_{17}, V_{18} and V_{19} respectively corresponding to the short circuit, the open circuit voltage, the fill factor and the efficiency that are near of the axis. The correlation coefficient confirm these results since $\rho(V_{12},V_{13}) = 0{,}582$ and $\rho\ (V_{14},V_{15}) = 0{,}973$.

In the purpose to establish a classification of the individuals, these are plotted in figure 08. In this diagram, the GaAs, InP, CdTe materials have a strong presence compared to Si. These materials possess some very interesting properties such as having a profile or a very important information. Therefore, the statistical analysis allows us to make a obvious classification of the different photovoltaic materials according to their characteristics and properties.
The results we obtained shows the following classification: Si- GaAs- InP- CdTe- GaP- CdSe- Ge. It to noted that in spite that the efficiency of silicon is lower than that of the GaAs,

it remains the most used material because of their very interesting properties in the solar cell development.

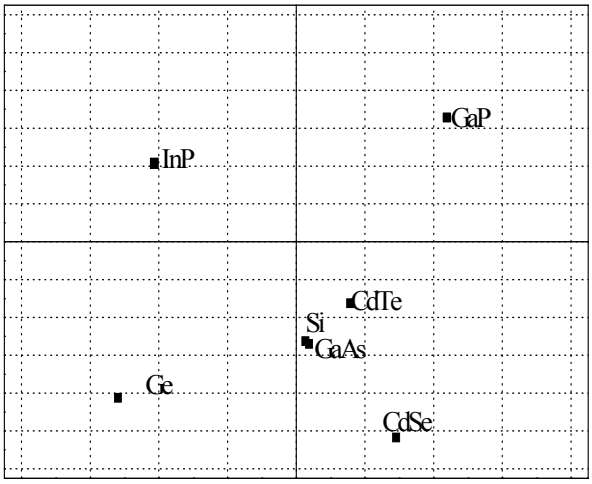

Fig. 8: Representation of the individuals (materials) in the principal plane.

All these results can be found in tables of numbers. This similarity in results clearly shows such diagrams of factorial analysis are the visualizations of tables of numbers.

5. CONCLUSION

In this paper, we established a hierarchical study of the different photovoltaic materials. The results of calculations and the analysis of data show that the A.P.C method is a simple method, very effective in the classification. It allows us to have a global visualization of the table of numbers and represent them in diagram.

The important result of this study is the possibility to hierarchically class the photovoltaic materials, Si, GaAs, InP, CdTe, GaP, CdSe and Ge, with respect to actual performances. Finally, we have confirmed that the silicon presents a great importance to be used in photovoltaic. This is not only for its technology simplification, best physical performances but also for its processing price.

REFERENCES

[1] Jacob Millman- Arvin Grabel. "Dispositifs à Semi-conducteurs" 1989.
[2] Harold J. Hovel. "Semiconductors and Semimetals"- Volume 11. Academic Press New York. Solar Cells. 1975.
[3] R. C. Neville, " Solar Energy Conversion : The Solar Cell " 1978.
[4] H. Ohtsuka, Y. Nagata, M. Sakamoto. "Optimization of Thermal Processing and Devices for High Efficiency C- Si Solar Cells". Solar Energy Materials and Solar Cells. 48 (1997) 137-143.
[5] T. Veda, M. Mohri."Improvement of Conversion Efficiency by $In_{0.52}Al_{0.48}Ad$ Windows Layers for $P^{+}InP$ Splar Cells". Solar Energy Materials and Solar Cells 50. 1998.
[6] C. Frekides, J. Britt." CdTe Solar Cells With Efficiency Over 15%". Solar Energy Materials and Solar Cells, 35. 1994.
[7] I. C. Lerman. "Les présentations Factorielles de la Classification". Gauthier Villars, Paris. 1970.
[8] Mme Nora. "Analyse Factorielle d'un nuage de points", Publication Interne, Université ParisVI.
[9] I. C. Lerman. "Classification et Analyse Ordinale des Données" Edition Dunod 1981.
[10] "Statistiques et probabilités", Tome II, A.I.D.E.P, Edition Dunod.