

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**



**3^{ème} CONFÉRENCE INTERNATIONALE SUR L'ENERGIE, LES MATERIAUX, L'ÉNERGÉTIQUE APPLIQUÉE ET
LA POLLUTION (CIEMEAP'2016)**

CONSTANTINE, ALGERIE,

30-31 Octobre 2016

**3rd INTERNATIONAL CONFERENCE ON ENERGY, MATERIALS, APPLIED ENERGETICS AND POLLUTION
(ICEMAEP'2016)**

CONSTANTINE, ALGERIA,

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

Application de l'ACP à une base de données de liquides ioniques

***MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³;
ARBAOUI Khaldia⁴, AOUINE Rachida⁵***

Laboratoire Synthèse et catalyse, Univ Tiaret BP78 RP 14000 Tiaret Algeria,

¹nouraliman.m.h@gmail.com

²belarbi.hb@gmail.com

³m.debdab@hotmail.fr

⁴khalida.arbaoui@yahoo.fr

⁵lokatchikifafouch@gmail.com

RESUME

L'objectif de cette étude est de mettre en évidence le facteur le plus influant sur la conductivité électrique d'un liquide ionique en appliquant l'analyse en composantes principales (ACP). Pour atteindre cet objectif nous avons utilisé le logiciel XLSTAT que nous avons appliqué à une base de données de 500 liquides ioniques différents. Pour chaque liquide ionique, une dizaine de propriétés physico-chimiques sont répertoriées. Il ressort de cette étude que les liquides ioniques présentant une forte conductivité électrique sont caractérisés par une faible température de transition vitreuse, une forte température de cristallisation et un faible indice de réfraction. La conductivité d'un liquide ionique augmente avec la taille de l'anion et diminue avec l'augmentation de la taille du cation.

Mots clés : liquides ioniques ; Data mining ; ACP, base des données ; XLSTAT.

NOMENCLATURE

Symbols :

Td: température de décomposition, K

Tc: température de cristallisation, K

Tg: température de transition vitreuse, K

Tf: température de fusion, °C

Lettres grecques:

ρ : masse volumique, g cm⁻³

σ : conductivité électrique, S·m⁻¹

η : viscosité, cp

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

F: largeur de la fenetre électrochimique

Cathod : limite du potentiel cathodique, V

Anod: limite du potentiel anodique, V

M: masse molaire , g

n: indice de réfraction

1. *Introduction :*

Une attention particulière a été porte sur l'amélioration de la qualité d'énergie et aussi de leur stockage. Les besoins de stockage d'énergie dans les applications sont nombreux et leur nécessité se révèle de plus en plus forte. Donc représente le point dur dans l'ensemble de filière d'électricité^[3].

La recherche nécessite un model technico-économique de dispositif de stockage, les études est autour sur les accumulateurs électrochimiques ou batterie électrochimique seul dispositif commerciaux actuels capable de répondre au besoin du stockage pas totalement mais minimise considérablement la complexité du problème^[4]; concernant les batteries la question qui se pose comment maximiser la durée de vie et sa capacité énergétique et aussi des stratégies spécifique. D'après le bulletin de caractéristique de moyen de stockage ^[3], on remarque que le rendement des batteries électrochimique et batteries à circulation vaut 70% par rapport autre moyen. Donc la batterie dite à circulation d'électrolyte les composés chimiques responsables au stockage de l'énergie généralement sont des liquides ioniques^[5]. Notre objectif est de trouver un liquide ionique à partir de technique d'exploitation des données et tester comme un électrolyte.

2. *Méthodologie :*

Les résultats liés à l'utilisation de l'analyse en composantes principales (ACP), appliqué sur les résultats apportés par le modèle théorique de Dietl et al ^[6]. en portant l'accent sur les liquides ioniques. Afin d'avoir une cohérence dans notre base de données on décide de faire les calculs sur les liquides ioniques. Les calculs ont été faits avec une technique en utilisant le logiciel XLSTAT. nous avons construit une base de données pour les différents liquides ioniques, comportant des propriétés telles la température de décomposition « Td » en K°, la température de cristallisation « temp » en c°, la température de transition vitreuse « tg » en , ainsi que le point de fusion en c°, conductivité « con » en S.m-1 , la viscosité « n » en cp et la densité « p » en g.cm-1 , la masse molaire « M » en g/mol et fenetre électrochimique « f » ,limite cathodique « cathode » ,limite anodique

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

« anode » ; grâce a cette base de données , nous allons identifier les tendances entre les différentes propriétés de liquides ioniques, ainsi que leurs corrélation avec la conductivité.

2.1. Construction de la base :

Définition : Une base de données est un ensemble structuré et organisé permettant le stockage de grandes quantités d'informations afin d'en faciliter l'exploitation^[7].

Dans le cadre de cette étude, nous avons construit une base de données des propriétés physico-chimiques (12 propriétés) des liquides ioniques (500 liquides ioniques), qui ont été collectées à partir de plusieurs sources documentaires. Presque tous les types de propriétés physico-chimiques importantes sont recueillis dans cette base de données « voir annexetableau a ».

La base des données n'est pas complète donc il ya des données manquante. Pour remplir les manquant on utilise l'un des méthodes :

- a) **Remplacement par la moyenne^[7]**: Calculer la moyenne de la variable qui nous intéresse parmi les enregistrements qui ont une valeur. Cette valeur sera ensuite attribuée à tous les enregistrements où la valeur est manquante .les Avantages de cette méthode sont rapide et facile d'exécution et parmi les inconvénients : imprécis et perte de la variabilité dans les données.
- b) **Remplacement aléatoire^[7]**: Pour chacune des valeurs manquantes, on attribuera au hasard une des valeurs parmi l'ensemble des valeurs des enregistrements non manquants. les Avantages de cette méthode : permet de garder la variabilité dans les données et la moyenne de la population mais cette méthode est plus complexe à implanter et la valeur imputée pour chacun des clients n'est pas plus précise que le remplacement par la moyenne.
- c) **Utilisation de la régression^[7]** : On utilise la régression pour obtenir un estimé de la valeur possible en utilisant les données des enregistrements complets et de toutes l'information disponible. c'est une méthode la plus précise (meilleure) ; pas toujours possible (variables explicatives) et la plus complexe à implanter.

2.2. Méthode d'exploitation des données :

Traitement de base des données « tableau a. » peut être fait sous **XLSTAT** qui permet d'appliquer sur EXCEL les méthodes d'analyse des données.

a) Logiciel XLSTAT :

MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵

XLSTAT est un outil d'analyse de données et de statistiques pour Microsoft Excel le plus utilisé parce que c'est un logiciel sûr, rapide, facile à installer et à utiliser^[8]. Il offre de très nombreuses fonctionnalités qui font d'Excel un outil performant et facile d'accès pour répondre à la majorité des besoins en analyse de données et modélisation^[8].

Toutes les fonctions XLSTAT sont accessibles à partir d'une icône qui est ajoutée à la barre de menus de l'Excel. L'utilisation d'Excel comme interface rend ce logiciel simple à utiliser et efficace. La qualité des calculs est quant à elle identique à celle des logiciels scientifiques les plus renommés, et XLSTAT couvre l'essentiel des besoins du statisticien qu'il soit expert ou débutant^[12].

Des modules optionnels répondent à des besoins plus spécifiques (contrôle statistique des procédés, analyse des effets de doses en chimie et pharmacologie)^[9]. Il fonctionne avec toutes les versions d'Excel, depuis la version 97 jusqu'à la version 2007 sous les environnements Windows et Mac.

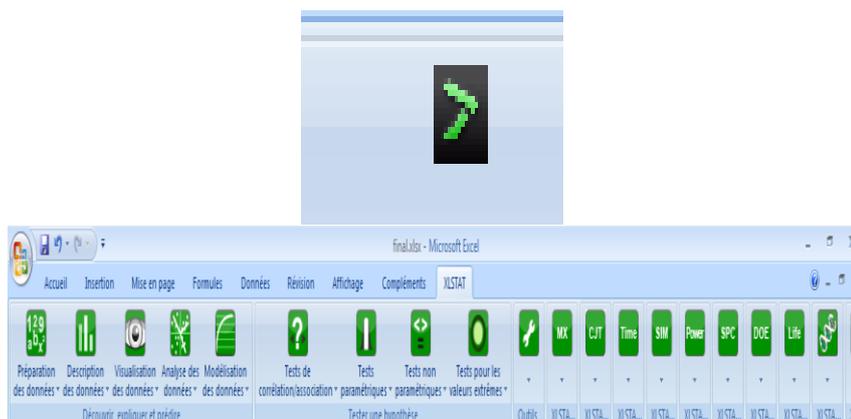


Figure 01 : logiciel XLSTAT

b) Application de l'analyse en composantes principales (ACP) :

L'Analyse en Composantes Principales (ACP) est une méthode très efficace pour l'analyse de données quantitatives (continues ou discrètes) se présentant sous la forme de tableaux à M observations / N variables^[8]. C'est une méthode descriptive qui permet d'effectuer une représentation simplifiée d'une série de variables intercorrélées. Cette technique utilise comme principe une transformation des variables initiales en de nouvelles variables non corrélées. Ces nouvelles variables sont appelées les Composantes Principales. Chaque composante principale est une combinaison linéaire des variables initiales^[9].

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

L'analyse en composantes principales est utilisée pour réduire la dimension (le nombre de variables) d'un problème. Cette diminution du nombre de variables doit s'effectuer en perdant un minimum d'informations. Le but de l'Analyse en Composantes Principales (ACP) est de condenser les données originelles en de nouveaux groupements, appelées nouvelles composantes ^[8,9]. L'ACP permet d'explorer les liaisons entre variables et les ressemblances entre individus.

Elle permet aussi^[9]:

- De visualiser et analyser rapidement les corrélations entre les N variables,
- De visualiser et analyser les M observations initialement décrites par N variables sur un graphique à deux ou trois dimensions, construit de manière à ce que la dispersion entre les données soit aussi bien préservée que possible,
- De construire un ensemble de P facteurs non corrélés

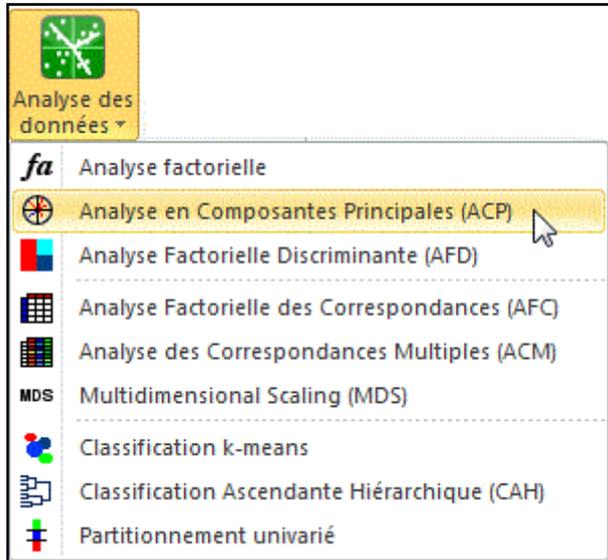
Les limites de l'Analyse en Composantes Principales viennent du fait que c'est une méthode de projection, et que la perte d'information induite par la projection peut entraîner des interprétations erronées. Des astuces permettent cependant d'éviter ces inconvénients^[12].

Il est également important de noter que l'ACP est un outil de **statistique exploratoire** et ne permet pas généralement de tester des hypothèses. Un avantage de cet aspect exploratoire est que l'ACP peut être exécutée à différentes reprises, en éliminant / rajoutant des observations ou des variables, à condition que ces manipulations soient justifiées dans les interprétations^[8].

Pour exécuter une analyse en composante principale avec le logiciel XLSTAT, il faut passer par les étapes suivantes :XLSTAT / Analyse de données / Analyse en Composantes Principales, ou cliquez sur le bouton correspondant de la barre **Analyse de données**^[9].



MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵



Une fois le bouton cliqué, la boîte de dialogue correspondant à l'**Analyse en composantes principales** apparaît. Vous pouvez alors sélectionner les données sur la feuille Excel. L'option Libellés des variables est activée, car la première ligne de données contient le nom des variables. Le Format des données choisi ici est Observations/Variables car c'est bien le format des données de départ. Le Type d'ACP choisi est Pearson (n), ce qui signifie que les calculs seront basés sur une matrice composée des coefficients de corrélation de Pearson, le coefficient de Pearson étant le coefficient de corrélation classiquement utilisé. Les matrices de covariance allouent plus de poids aux variables ayant des variances élevées. Les corrélations de Spearman peuvent être plus appropriées lorsque l'ACP est exécuté sur des variables aux distributions différentes^[9]. Dans l'onglet Sorties, on a choisi d'activer l'option Tester la significativité pour afficher en gras les corrélations significativement différentes de 0.

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
 AOUINE Rachida⁵**

Tm	Con	n	p	M	i	f
103,550	1,164	701,345	1,373	183,940	2,301	2,312
22,15						1,751
39,15						2,699
52,15						2,325
54,97						1,875
-9,85						2,703
60,84						5,200
53,37						2,335
-20,85						3,000
713,15						1,741
53,15						1,916
-13,85						2,183
0,28						1,668
39,15						2,350
-5,85	1,463	57,000	1,000	203,200	2,010	2,107
41,150	1,260	654,743	18,836	163,020	2,308	1,788

Analyse en Composantes Principales (ACP)

Général | Options | Données supp. | Prétraitement | Sorties | Graphiques

Statistiques descriptives Valeurs propres
 Corrélations Coordonnées des variables
 Tester la significativité Corrélations Variables/Facteurs
 Test de sphéricité de Bartlett Coordonnées des observations
Niveau de signification (%): 5 Contributions
 Kaiser-Meyer-Olkin Cosinus carrés

OK Annuler Aide

Dans l'onglet Graphiques, les options d'Etiquettes sont toutes activées afin que les libellés des variables et des observations soient bien affichés. L'option de filtrage des observations à afficher est aussi désactivée afin d'afficher toutes les observations.

Tm	Con	n	p	M	i	f
103,550	1,164	701,345	1,373	183,940	2,301	2,312
22,15						1,751
39,15						2,699
52,15						2,325
54,97						1,875
-9,85						2,703
60,84						5,200
53,37						2,335
-20,85						3,000
713,15						1,741
53,15						1,916
-13,85						2,183
0,28						1,668
39,15						2,350
-5,85	1,463	57,000	1,000	203,200	2,010	2,107
41,150	1,260	654,743	18,836	163,020	2,308	1,788

Analyse en Composantes Principales (ACP)

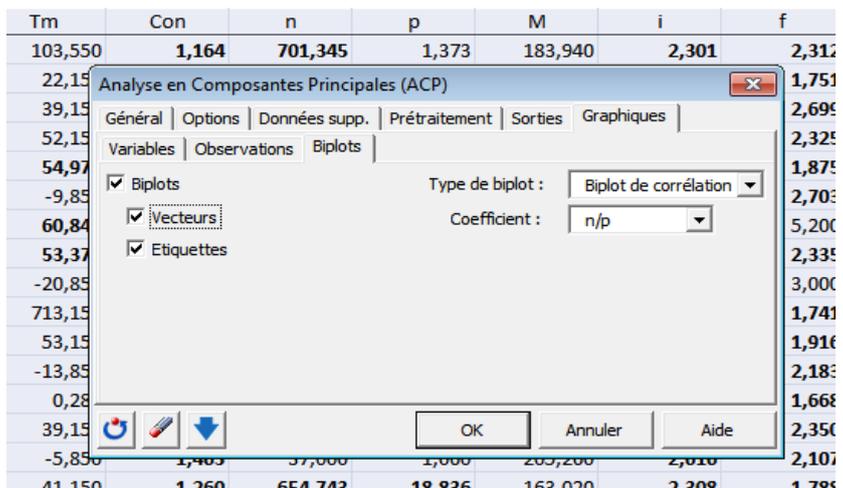
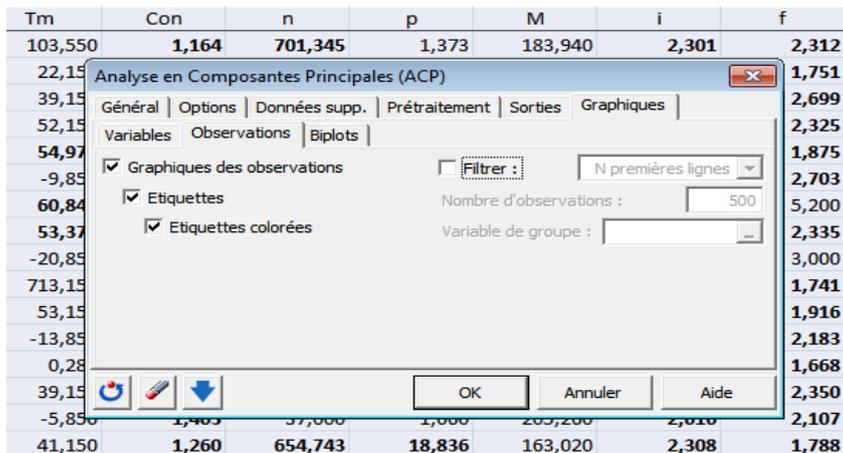
Général | Options | Données supp. | Prétraitement | Sorties | Graphiques

Variables | Observations | Biplots

Graphiques de corrélations Filtrer : Aléatoire
 Vecteurs Nombre de variables : 50
 Etiquettes colorées Variable de groupe : -

OK Annuler Aide

MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
 AOUINE Rachida⁵



Les calculs commencent lorsque vous cliquez sur le bouton **OK**. Si vous avez choisi dans le panneau des options XLSTAT l'option Demander la confirmation des sélections, XLSTAT vous demande de confirmer le nombre de lignes et de colonnes des sélections. Puis une nouvelle boîte vous permet de choisir les axes pour lesquels les graphiques doivent être affichés.

3. Résultats et discussion :

L'analyse en composante principale permet le développement d'un modèle mathématique basé sur les données expérimentales disponibles en fonction de la structure moléculaire, permettant ainsi d'estimer, d'étudier et d'identifier les propriétés physico-chimiques des liquides ioniques. Les résultats d'analyse des données expérimentales aux modèles proposés, obtenus par le logiciel XLSTAT, sont des résultats formés des tableaux et des figures^[10]. Les Données manquantes sont complétées par la méthode d'estimation Imputation multiple (MCMC) « voir l'annexe tableau b ».

MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵

L'analyse en composantes principales nous donne les axes principaux. Pour un plan factoriel donné, on regarde le pourcentage de l'information du nuage initial retranscrite par le plan factoriel. Ainsi, un axe expliquant moins de 10% sera rarement intéressant^[7], les axes sont rangés dans l'ordre décroissant, de telle manière que le premier plan factoriel constitué par les deux premiers axes factoriels soit toujours celui qui est le plus riche en renseignements sur les propriétés du nuage étudié. Dans notre exemple, les axes F1 (PC1) et F2 (PC2) sont les plus riches en information figure 02.

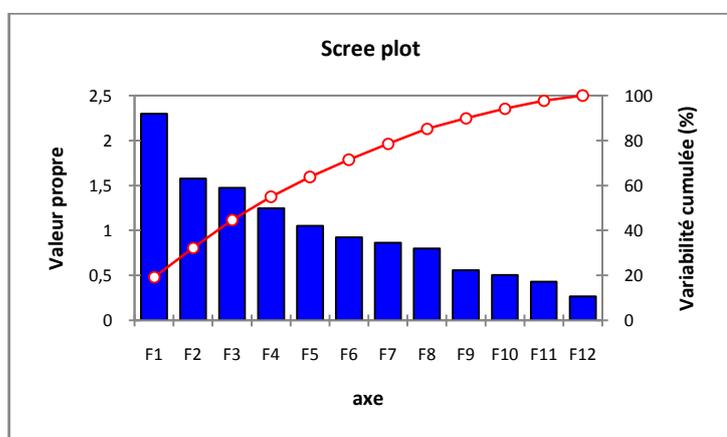


Figure 02 : Histogramme des valeurs propres. (Pourcentages de variance absorbée, pour les 12 axes principaux).

- Le premier résultat intéressant à identifier suit à une analyse en composantes principales est la matrice des corrélations qui permet d'analyser les relations bilatérales existant entre les différentes variables retenues. Cette matrice est représentée sur le tableau 01.

Variables	Temp	Tf	Tg	Tm	Con	n	p	M	I	f	cathode	Anode
Temp	1	0,281	-0,003	0,142	0,102	0,068	0,286	0,043	0,005	-0,014	-0,007	-0,076
Tf	0,281	1	0,038	0,150	-0,157	0,046	0,066	-0,034	0,017	-0,038	0,028	-0,010
Tg	-0,003	0,038	1	0,115	-0,036	0,162	-0,250	0,330	0,007	0,036	0,005	-0,004
Tm	0,142	0,150	0,115	1	0,023	0,134	0,069	-0,033	0,046	-0,086	0,042	-0,080
Con	0,102	-0,157	-0,036	0,023	1	0,002	0,046	-0,113	-0,103	0,065	-0,066	-0,047
N	0,068	0,046	0,162	0,134	0,002	1	0,025	-0,076	-0,022	-0,011	0,009	-0,193
P	0,286	0,066	-0,250	0,069	0,046	0,025	1	-0,033	0,052	-0,011	-0,031	0,035

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

M	0,043	-0,034	0,330	-0,033	-0,113	-0,076	-0,033	1	0,018	0,042	-0,019	0,030
I	0,005	0,017	0,007	0,046	-0,103	-0,022	0,052	0,018	1	-0,614	0,562	0,014
F	-0,014	-0,038	0,036	-0,086	0,065	-0,011	-0,011	0,042	-0,614	1	-0,703	0,196
cathode	-0,007	0,028	0,005	0,042	-0,066	0,009	-0,031	-0,019	0,562	-0,703	1	-0,052
anode	-0,076	-0,010	-0,004	-0,080	-0,047	-0,193	0,035	0,030	0,014	0,196	-0,052	1

Tableau 01 : Matrice de corrélation (Pearson (n))

La matrice de corrélation permet d'observer la relation positive entre point de fusion (Tm) et la viscosité (n) (0.13) et la température de congélation (Tf) (0.15), température de transition vitreuse (Tg) (0.12) et température de cristallisation (Temp) (0.14). Aussi une corrélation entre la conductivité (con) et la température de cristallisation (Temp) (0.10) ; et entre la viscosité (n) et la température de transition vitreuse (Tg) (0.16) ; entre la densité (p) et la température de cristallisation (Temp) (0.29) et aussi entre la masse molaire (M) et la température de transition vitreuse (Tg) (0.33) ; entre l'anode et fenêtre électrochimique (f) (0.2) et une forte corrélation entre l'indice de réfraction (i) et limite cathodique (cathode) (0.56). On constate aussi la relation négative entre la conductivité (con) et température de congélation (Tf) (-0.16) et la masse molaire (M) (-0.11) et l'indice de réfraction (i) (-0.1) et entre la fenêtre électrochimique (f) et l'indice de réfraction (i) (-0.61) et limite cathodique (-0.7) et entre limite anodique (anode) et la viscosité (n) (-0.2) et aussi entre la température de transition vitreuse (Tg) et la densité de liquides ioniques (p) (-0.25). Toutes ces corrélations entre variables vont conditionner la composition des axes factoriels dont le sens et la signification s'interpréteront en fonction de leur corrélation avec chaque variable.

La diagonalisation de la matrice de corrélation donne les valeurs propres présentées dans le tableau 02. Chaque facteur est caractérisé par une valeur propre, une part de la variance totale (% de variance), et une part de la variance cumulée (% cumulée).

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Valeur propre	2,300	1,576	1,473	1,248	1,053	0,926	0,864	0,801	0,557	0,506	0,432	0,266
Variabilité (%)	19,165	13,131	12,276	10,398	8,771	7,713	7,199	6,675	4,640	4,216	3,598	2,218
% cumulé	19,165	32,296	44,572	54,970	63,741	71,454	78,652	85,327	89,967	94,183	97,782	100,000

Tableau 02 : les valeurs propre des axes principaux.

MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵

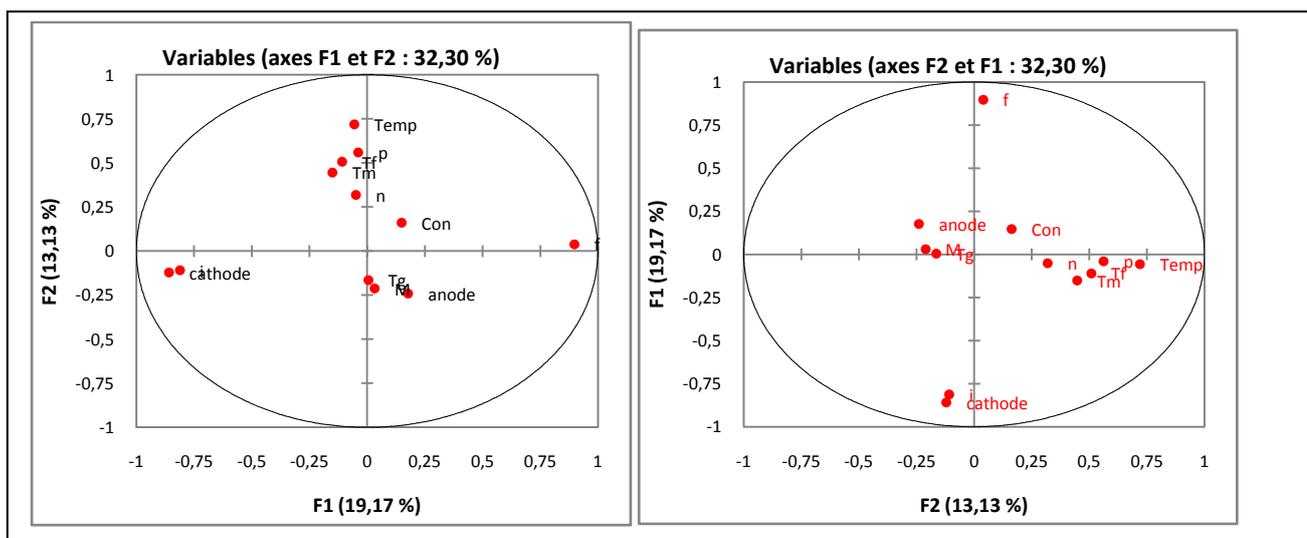
Ce tableau signifie que le premier axe (F1) permet d'expliquer (19.17%) de la variance totale du nuage de points, que le second axe (F2) permet d'expliquer (13.13%) de la variance totale. En projetant donc chaque individu sur un plan (F1, F2) on conserve 19.17+13.13 soit 32.30% de la variance totale.

Remarquons ici que la variance expliquée par l'axe F3 est faible (12.28 %) et F4 (10.40%), et remarquons aussi que l'essentiel de l'information est fourni par les 2 premiers axes, puisque la variance expliquée cumulée par est de 32.30%. Si les variables sont très corrélées entre elles, les premières valeurs propres sont importantes alors que si le lien linéaire entre les variables d'origines est faible, les pourcentages de variance expliquée par les axes seront faibles eux aussi.

La figure 02 nous permet d'évaluer les valeurs propres de manière schématique, La première composante principale (F1) correspond à la plus grande valeur propre, ainsi la deuxième composante (F2) correspond à la deuxième plus grande valeur propre. Généralement, l'étude se restreint à ces deux premiers axes, du fait qu'ils contiennent la majorité de la variance, le pourcentage de variance absorbée par les axes principaux est indiqué sur la figure (02).

A partir du tableau (02) et du graphique (02), on remarque que ces valeurs sont décroissantes de F1 (2.3) à F9(0.3) et qu'à partir F6 elles sont inférieures à 1.

L'analyse des résultats de l'ACP pour les liquides ioniques, est donnée par les graphes des observations "score plot" et des variables "loading plot". Le résultat des "Score plot" de cette analyse est montré sur les figures 03.



MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵

Figure 03-a : représentation des variables pour les liquide ioniques correspondant aux observations sur le plan factoriel (F1,F2).

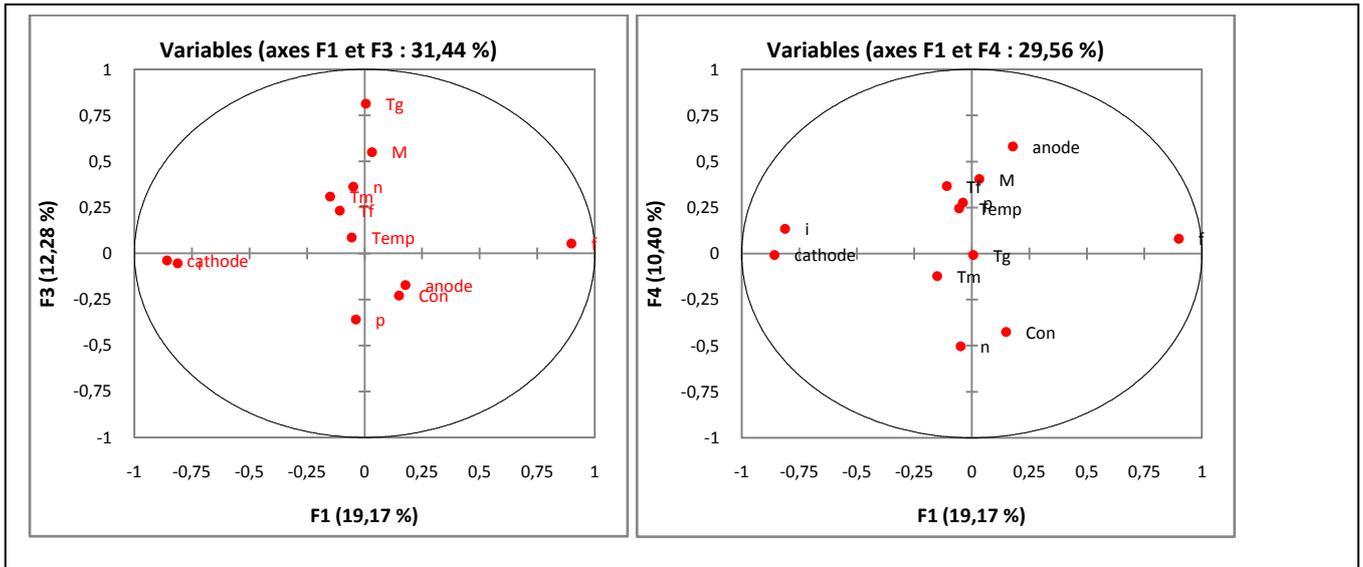
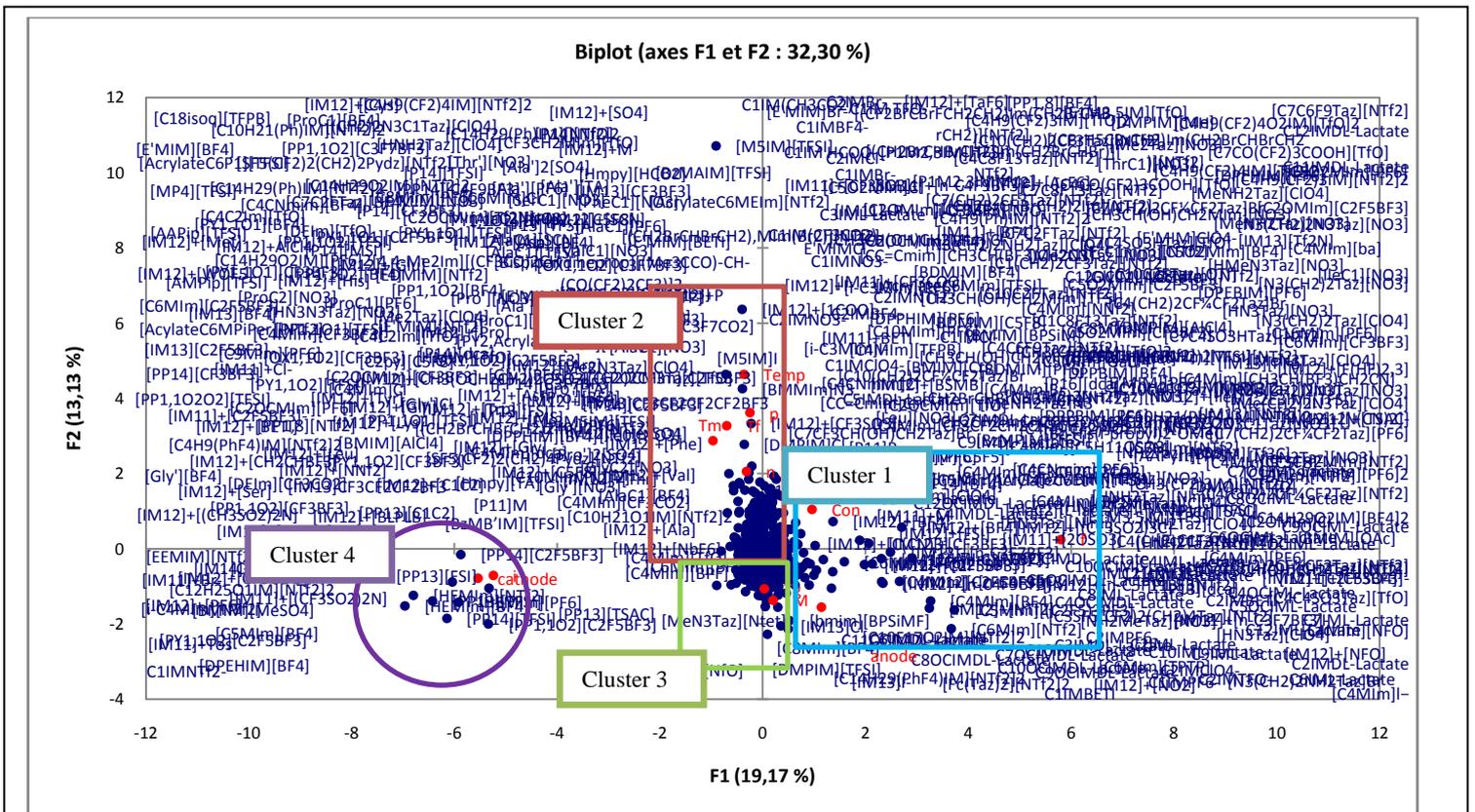


Figure 03-b : représentation des variables pour les liquide ioniques correspondant aux observations sur le plan factoriel (F1,F2)



MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵

● Actives □ Supplémentaires ● variables actives

Figure04 : graphe des observations de 500 liquides ioniques

Sur la figure 03-a On peut constater que l'axe F1 est fortement corrélé positivement à la fenêtre électrochimique, aussi il est fortement corrélé négativement à limite cathodique et l'indice de réfraction. Tandis que l'axe F2 est fortement corrélé négativement à la température de cristallisation et point de fusion et température de congélation et la densité de liquides ioniques.

Pour confirmer le fait qu'une variable est fortement liée à une composante principale, il suffit de consulter le tableau suivant de cosinus carrés.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Temp	0,003	0,518	0,008	0,060	0,077	0,011	0,043	0,038	0,239	0,002	0,000	0,000
Tf	0,012	0,258	0,054	0,135	0,201	0,007	0,136	0,046	0,148	0,002	0,001	0,000
Tg	0,000	0,027	0,662	0,000	0,038	0,024	0,000	0,031	0,000	0,210	0,007	0,000
Tm	0,023	0,199	0,096	0,015	0,002	0,312	0,019	0,323	0,002	0,009	0,000	0,000
Con	0,022	0,026	0,052	0,180	0,418	0,109	0,067	0,049	0,069	0,006	0,002	0,000
N	0,002	0,102	0,132	0,253	0,028	0,017	0,268	0,144	0,001	0,051	0,001	0,000
P	0,002	0,315	0,129	0,077	0,069	0,049	0,198	0,014	0,065	0,071	0,012	0,000
M	0,001	0,045	0,304	0,164	0,206	0,096	0,001	0,034	0,028	0,120	0,001	0,000
I	0,660	0,012	0,003	0,018	0,009	0,003	0,014	0,005	0,000	0,007	0,257	0,012
F	0,807	0,002	0,003	0,007	0,002	0,001	0,005	0,001	0,000	0,000	0,013	0,160
Cathode	0,737	0,015	0,001	0,000	0,002	0,003	0,001	0,012	0,000	0,009	0,135	0,085
Anode	0,031	0,058	0,029	0,339	0,001	0,293	0,112	0,104	0,004	0,017	0,004	0,009

Tableau 03 : Cosinus carrés des variables (propriétés physico-chimique)

Si le cosinus carrés est élevé, une variable est plus liée à l'axe. Dans notre cas l'axe F1 est lié à i, f, cathode les cosinus carrés i(0,66), f(0,81), cathode(0,74), l'axe F2 est lié à Temp, Tf, p avec des cosinus carrés Temp(0,52), Tf (0,26), p(0,32), l'axe F3 est lié à Tg, M avec des cosinus carré Tg(0,66), M(0,30) et l'axe F4

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

est lié à limite anodique avec cosinus carrés anode(0.34) et contrairement une variable est moins lié à l'axe si son cosinus carrés est proche de zéro. Dans ce cas point de fusion est moins lié à l'axe F2.

On peut remarquer que la conductivité et la masse molaire de liquides ioniques et l'indice de réfraction et température de congélation et la température de transition vitreuse sont inversement corrélés, ce qui nous permet de dire que les liquide ionique les plus visqueux et de masse molaire grande et d'indice de réfraction très important et de température de transition vitreuse élevé et point de congélation forte sont de faible conductivité électrique.

On outre il est claire que la conduction et point de fusion sont inversement corrélés ce qui confirme que la température de fusion forte les liquides ioniques est de faible conduction.

La température de cristallisation et la conductivité se trouvent dans le même coté cela signifie qu'ils sont fortement corrélés entre eux, dont les liquide ionique de grande température de décomposition sont conductibles.

A l'aide du tableau c de Coordonnées des observations (liquides ioniques) on constate que l'axe F1 est fortement lié à la variation de la taille de l'anion, quand la valeur de l'axe F1 augmente la taille des anions diminue. La même remarque pour l'axe F2 qui capture la variation de la taille de cation, quand la valeur de l'axe F2 augmente la taille des cations augmente.

En premier vue, il s'avère que l'ensemble des composés chimiques se séparent nettement en quatre groupes (clusters). Le cluster 1 est positif par rapport l'axe F1, il regroupe les liquides ioniques qui ont une forte conductivité électrique, la valeur de limite anodique élevé, de masse molaire faible et ont de température de transition vitreuse très faible.

Ceux avec un axe F2 négatif : le cluster 2, cluster 3 et cluster 4. Le cluster 1 correspond à liquides ioniques plus denses et visqueux, de température de fusion très importante et de faible conduction. Pour les liquides ioniques qui sont regroupé par cluster 3 ont une température de transition vitreuse élevé et plus massique ayant une faible conductivité électrique. Le cluster 4 regroupé les liquides ioniques qui ont une valeur de limite cathodique très importante et indice de réfraction élevé.

Parmi les 500 liquides ioniques en trouvent 8 liquides de bonne conductivité ionique ; la figure représente la structure et la formule chimique pour chaque liquide ionique de telle manière la valeur de la conductivité de liquides ioniques encadré en rouge sont des valeurs prédites.

MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
 AOUINE Rachida⁵

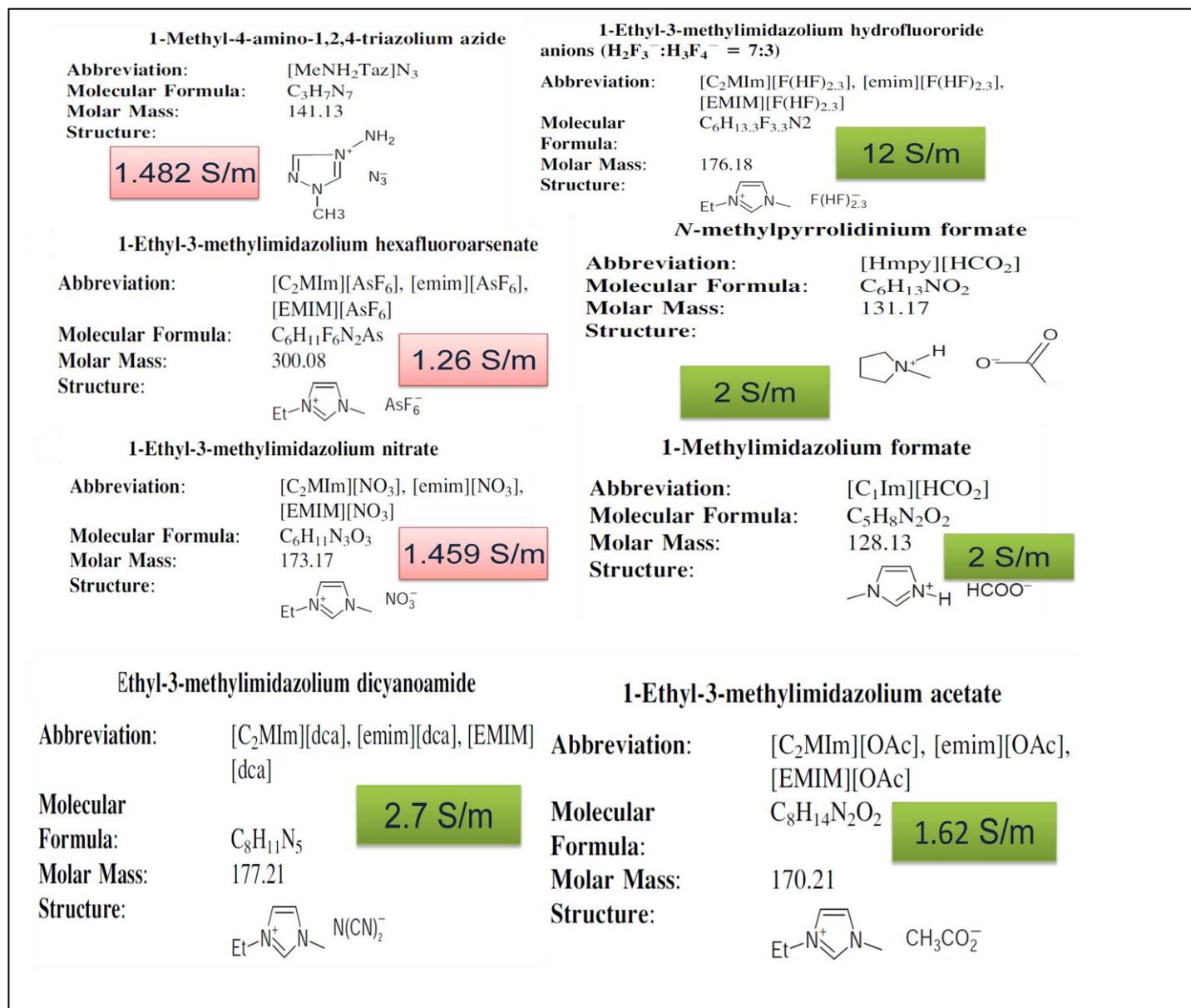


Figure05 : la structure et la formule chimique^[11] de chaque liquide ionique qui ont une bonne conductivité ionique.

4. Conclusion :

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

Dans ce travail d'étude on propose une nouvelle technique d'estimation de propriétés physico-chimiques des liquides ioniques basée sur la quelle les techniques du data mining seront appliquées en particulier, la méthode d'analyse en composantes principale (ACP). Pour avoir les résultats de cette estimation on a utilisé le logiciel XLSTAT, l'utilisation de ce logiciel est simple, facile ainsi que les étapes de régression ne prend pas beaucoup de temps. La technique analyse en composant principales (ACP) sert à décrire des tableaux "individus-variables quantitatives" de grande dimension (beaucoup de variables - c'est un traitement multi-varié et beaucoup d'individus) des outils statistiques pour résumer et sélection des composant principaux important. Quand le nombre de composés qui constitue la base des données est grand par rapport au nombre des interactions, la régression sera plus précise^[10].

D'après ce travail on peut conclure que la conductivité est liée à la variation de la taille des anions et les cations, elle est faiblement corrélée positivement à l'axe F1 et F2 ; lors que la conductivité augmente la taille de l'anion augmente par contre elle diminue par l'augmentation de la taille de cation.

On constate que les liquides ioniques $[C_1IM]^+[HCOO]^-$, $[IM_{12}]^+[1COO]^-$, $[IM_{12}]^+[F(HF)2.3]^-$, $[IM_{12}]^+[AsF_6]^-$, $[IM_{12}]^+[N(CN)2]^-$, $[IM_{12}]^+[NO_3]^-$, $[HmPY]^+[HCOO]^-$, $[MeNH_3Taz]^+[N]^-$ ont une forte conductivité parmi 500 liquides ioniques. Ces liquides ioniques ont une faible température de fusion sont moins visqueux et de température de transition vitreuse faible.

Référence :

[1]-Liquides ioniques : structure et dynamique Bachir AOUN soutenue le 14 Décembre 2010 l'université d'Orléans.

[2]-Janusz Kcprzyk, lakhmi C.jain ; DATA MINING ;2011,ISBN978-3-642-19720-8.

[3]- B. Multon, J. Ruer « stocker l'électricité, oui c'est indispensable et c'est possible, pourquoi, ou, comment ? » Publication ECRIN en contribution au débat national sur l'énergie, Avril 2003.

[4]-stockage de l'énergie dans les applications stationnaires, Bernard Multon, Goel Robin, Erika Eranbert, Hamed Ben Ahmed, 3 Mars 2012, Colloque Energie électrique : besoins, enjeux, technologies et applications, Belfort, 18 juin 2004.

**MOHAMED CHERIF Yamina¹; BELARBI El Habib²; DEBDAB Mansour³; ARBAOUI Khaldia⁴,
AOUINE Rachida⁵**

[5]-le stockage d'énergie, Enjeux, solution technique et opportunités de valorisation, Mars 2012, 89 rue Réaumur 75002 Paris.

[6]-Françoise fessank, Fabrice clérbot ; vers l'exploitation de grandes masses de données, DBLP, 28 february 2016.

[7] –Myriam Ikhlef, Geneviève lallich, Yannick Lintz ; conception et mise en place d'une base de données pour le service du récolement des dépôts au Musce du louvre, juin 2011

[8] - TALI Halima (2012/2013) thèse de master Développement d'une méthode prédictive de calcul des propriétés des liquides ioniques

[9] - le fichier original peut-être obtenu sur http://eire.census.gov/popest/states_dataset.csv).

[10] -XLSTAT 2007, Copyright © 2007, Addinsoft <http://www.addinsoft.com>

[11]-Suojiang Zhang, Xingmei Lu, Qing Zhou, Xiaohua Li, Xiamgping Zhang, Shucaï Li; Ionic liquids physicochemical properties; Radarweg29,PO BOX 211, 1000AE Amsterdam, the Netherlands Linacre House, Jordan Hill.

[12]-Ahemed El Atari, A.Elmoussaoui ; analyse de données sous logiciel XLSTAT, université Abedelemalek Essaadi (UAE), Géo-information, 21 mars 2015.