



RELATION STRUCTURE – RETENTION CHROMATOGRAPHIQUE POUR LES PHENOLS

H. Nacer, D. Messadi.

*Laboratoire de sécurité environnementale et alimentaire
Institut de chimie.*

Université Badji Mokhtar Annaba. BP 12 El-Hadjar- Annaba. Algérie

E.mail: nacerhouria@yahoo.fr

Résumé :

En tenant compte de la polarité des phases stationnaires, on a pu prédire les valeurs des indices de rétention sur deux colonnes polaires (NGA, OV-225) à partir des valeurs obtenues sur une colonne non polaire (SE-30)

Nous avons adopté le modèle qui conduit à la plus faible somme des carrés des écarts.

Nous avons cherché à corrélérer les indices de rétention obtenus sur chacune des phases stationnaires avec divers descripteurs moléculaires (différents indices topologiques, réfraction moléculaire, moment dipolaire...)

Abstract :

According to the polarity of stationary phases, we could predict the retention indices on two polar columns (NGA, OV-225) from the values obtained on non polar (SE-30). column

The model which leads to the lowest squared intervals sum was adopted.

We searched to correlate the retention indices obtained on each column with diverse molecular descriptors (different topologic indices, refractivity, dipolar moment ...)

Mots clés:

Indices de rétention, chimiométrie, chromatographie en phase gazeuse, descripteurs moléculaires, corrélation

1. Introduction :

La prévision des caractéristiques de composés à partir de leur structure est importante et trouve des applications dans des domaines aussi variés que la sécurité environnementale et alimentaire, la formulation, le génie des procédés etc....

Les phénols et leurs dérivés [1] sont des polluants substantiels de l'environnement [2], c'est la raison pour laquelle nous nous sommes intéressés à la résolution par chromatographie gazeuse de 43 phénols diversement substitués en utilisant la technique des phases multiples.

Trois colonnes de polarités différentes, mettant en jeu différents processus de rétention, ont été exploitées:

-Une phase non polaire constituée de diméthylpolysiloxane (SE-30)

-Deux phases polaires, l'une constituée de polynéopentylglycoladipate (NGA) et l'autre de 3-cyanopropylméthylpolysiloxane (OV-225). 57 descripteurs moléculaires ont été testés avec 38 composés phénoliques sélectionnés à partir

des 43 composés initiaux, après nettoyage des données.

Un descripteur moléculaire est une entité numérique caractérisant une propriété donnée d'une molécule et obtenue à partir de sa structure moléculaire.

2. Calculs des modèles statistiques :

Les modèles ont été calculés en utilisant une méthode de sélection pas-à-pas à l'aide du logiciel de traitement des données Minitab [3], pour chacune des 3 colonnes. SE-30, NGA et OV-225.

Nous avons repris les données publiées par Grzybowski et al. [4], en nous assignant plusieurs buts:

Tout d'abord, sur la colonne SE-30, essayer de corrélérer un plus grand nombre de phénols avec les descripteurs moléculaires soit déjà disponible dans la littérature scientifique soit facile à obtenir.

Relation structure- rétention chromatographique pour les phénols.

Nous avons pu ainsi corrélérer, sur la colonne SE-30, les indices de rétention [5] de 30 phénols. Valeur propre de la matrice de distance d'ordre 1 (VP (Ax1)), énergie de la plus haute orbitale moléculaire occupée (HOMO) [6] calculée à l'aide du logiciel de modélisation moléculaire Hyperchem Pro-6.02 et le nombre d'atomes d'oxygène [nbreO].

En introduisant le moment dipolaire calculé (μ_{cal}), nous avons pu corrélérer 26 phénols avec une amélioration sensible des paramètres statistiques [7]:

L'erreur standard est pratiquement divisée par deux ;

$R^2 = 98.9\%$, $R^2_{adj} = 98.6\%$, $R^2_{pred} = 99.99\%$, $F = 450.86$; la statistique de Durbin-Watson: $d = 1.96$ [8-9].

Nous avons noté cependant une détérioration du facteur d'inflation moyen de la variance <VIF> qui passe de 1.5 à 6.5 mais qui reste toutefois acceptable puisque inférieur à 10 [10];

2.1. Graphes et discussion :

Sur la fig.2.1 on remarque une droite de régression de coefficient de corrélation $R = 95.66\%$ supérieur à la valeur critique $R^2_{cr} = 92.84\%$ donnée par les tables statistiques pour un niveau de signification $\alpha = 0.05$ et $n = 26$, individus ce qui met en évidence la normalité des résidus.

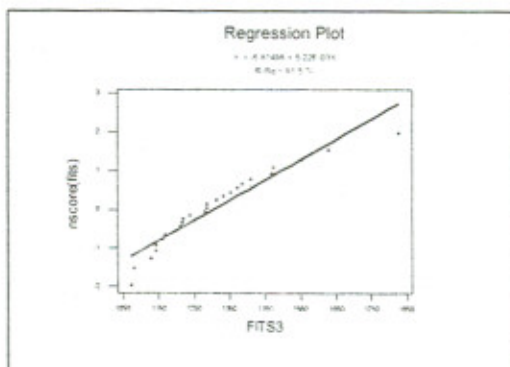


Fig.2. Dispersion des scores normaux en fonction des indices de rétention prédits pour la colonne SE-30 : \hat{I}_{SE-30}

La Fig.2.2 représente le graphique des valeurs estimées en fonction des valeurs observées,

avec 3 descripteurs moléculaires:

elle montre une droite de régression de coefficient de corrélation $R^2 = 99.4\%$, cela signifie que les indices de rétention estimés sur la colonne SE-30 sont expliqués à 99.40% par les indices de rétention expérimentaux I_{SE-30} et montre aussi que les points s'alignent bien sur la droite de régression.

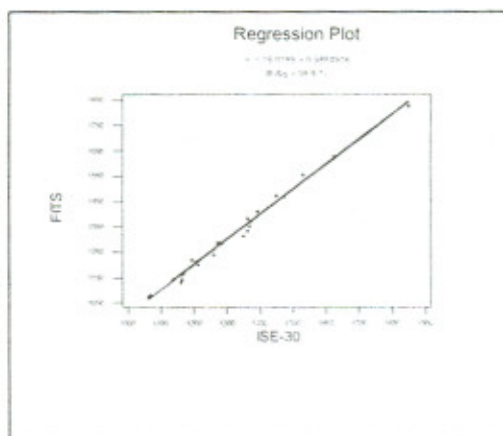


Fig. 2.2 : Dispersion des indices de rétention estimés en fonction des valeurs des indices de rétention observées I_{SE-30}

Ainsi le modèle à 4 régresseurs permettrait d'expliquer environ 99.99 % de la variabilité des observations estimées, en comparaison des données originales expliquées à 98.8% par un ajustement par les moindres carrés.

La capacité prédictive du modèle est satisfaisante.

L'équation du modèle est :

$$I_{SE-30} = -1054(\pm 158.10) + 161(\pm 16.10)\mu_{cal} + 45.2(\pm 2.62)VP(Ax1) + 138(\pm 18.34)HOMO + 74.1(\pm 16.90)nbreO$$

2.2. Calculs de modèles statistiques sur les colonnes polaire NGA et OV-225 :

Nous avons procédé de la même manière que précédemment au calcul de modèles statistiques linéaires [8-9] sur les colonnes NGA et OV-225.

Sur la colonne polaire OV-225 nous avons pu calculer 7 modèles possibles, alors que sur la colonne NGA, 5 modèles seulement ont pu être obtenus.

Relation structure- rétention chromatographique pour les phénols.

3. Calcul des indices de rétention sur les colonnes polaires à partir des grandeurs correspondantes déterminées sur la colonne non polaire SE-30 fournies par le meilleur modèle :

$$I_{NGA}^1 = f_1 (I_{SE-30}^{exp} \dots) \quad 3.1$$

$$I_{NGA}^2 = f_2 (I_{SE-30}^{esti} \dots) \quad 3.2$$

Pour la prédiction des indices de rétention sur les colonnes polaires NGA et OV-225, nous avons procédé par injection dans chacune des équations de régression des modèles obtenus sur les colonnes NGA et OV-225, des valeurs des indices de rétention expérimentaux et estimés suivant les équations:

Le tableau 3.1 regroupe les valeurs des indices de rétention observées I_{NGA} et celles calculées (I_{NGA}^1, I_{NGA}^2) sur la colonne polaire NGA à partir des équations 3.1 et 3.2, les résidus $ei1 = I_{NGA} - I_{NGA}^1$ et $ei2 = I_{NGA} - I_{NGA}^2$, ainsi que les pourcentages des erreurs relatives.

Tableau 3.1 : Valeurs des $I_{SE-30(cal)}$ et $I_{SE-30(exp)}$ du meilleur modèle injectées dans l'équation de régression du modèle obtenu pour la colonne NGA :

$$I_{NGA} = 172(\pm 69.32) + 23.4 (\pm 3.30) VP (Ax*3) + 83(\pm 12.09).^1\chi - 36.7 (\pm 7.92) NBREC + 2.43 (\pm 0.06) I_{SE30}.$$

| Nom des composés: ...phénol | I_{NGA} | I_{NGA}^1 | I_{NGA}^2 | ei1 | ei2 | (ei1/ I_{NGA})% | (ei2/ I_{NGA})% |
|--|-----------|-------------|-------------|---------|---------|-----------------------|-----------------------|
| 4-Me | 1813 | 1787.79 | 1818.16 | 25.215 | -5.160 | 1.39079 | -0.28461 |
| 3-Me | 1782 | 1800.95 | 1834.31 | -18.949 | -52.313 | -1.06337 | -2.93565 |
| 2,4-(Me) ₂ | 1825 | 1842.75 | 1856.99 | -17.750 | -31.989 | -0.97258 | -1.75284 |
| 3-Et | 1898 | 1880.17 | 1802.78 | 17.825 | 95.221 | 0.93916 | 5.01690 |
| 4-Et | 1890 | 1880.41 | 1830.42 | 9.594 | 59.579 | 0.50761 | 3.15233 |
| 3,5-(Me) ₂ | 1877 | 1750.47 | 1741.46 | 126.525 | 135.541 | 6.74083 | 7.22113 |
| 2,3-(Me) ₂ | 1857 | 1832.92 | 1830.49 | 24.077 | 26.507 | 1.29653 | 1.42739 |
| 4-Cl | 2058 | 2071.76 | 2138.61 | -13.756 | -80.606 | -0.66843 | -3.91670 |
| 3-Cl | 2061 | 2086.19 | 2141.42 | -25.187 | -80.421 | -1.22207 | -3.90203 |
| 2, 4,6-(Me) ₃ | 1778 | 1814.85 | 1829.14 | -36.847 | -51.136 | -2.07240 | -2.87602 |
| 3-(O-Me) | 2083 | 2008.67 | 1979.07 | 74.334 | 103.932 | 3.56861 | 4.98952 |
| 2, 3,5-(Me) ₃ | 1920 | 1919.14 | 1863.30 | 0.864 | 56.705 | 0.04497 | 2.95338 |
| 3-Br | 2214 | 2215.11 | 2249.71 | -1.107 | -35.710 | -0.04998 | -1.61291 |
| 4-Br | 2191 | 2226.35 | 2254.48 | -35.345 | -63.485 | -1.61320 | -2.89752 |
| 3-Me-4-Cl | 2135 | 2143.16 | 2137.62 | -8.165 | -2.624 | -0.38242 | -0.12291 |
| 2, 4, 6-(Cl) ₃ | 2067 | 2070.87 | 1981.47 | -3.872 | 85.527 | -0.18733 | 4.13776 |
| 2, 4, 5-(Cl) ₃ | 2158 | 2106.74 | 2158.11 | 51.258 | -0.112 | 2.37525 | -0.00520 |
| 3-(OH) | 2576 | 2543.58 | 2468.76 | 32.422 | 107.242 | 1.25861 | 4.16311 |
| 2-(O-Me)-4-(CH ₂ CH=CH ₂) | 1923 | 1880.47 | 1891.97 | 42.529 | 31.035 | 2.21158 | 1.61387 |
| 3,5-(Cl) ₂ | 2343 | 2340.10 | 2299.79 | 2.898 | 43.212 | 0.12371 | 1.84431 |
| 2-(O-Me)-4-Pr | 1884 | 1937.72 | 1979.30 | -53.719 | -95.296 | -2.85131 | -5.05818 |
| 2-(O-Me)-4-CHO | 2235 | 2213.65 | 2277.65 | 21.353 | -42.654 | 0.95537 | -1.90844 |
| 2,6-(O-Me) ₂ -4-Me | 2106 | 2146.55 | 2131.05 | -40.552 | -25.049 | -1.92555 | -1.18940 |
| 2-(O-Me)-4-(CO-Me) | 2326 | 2327.65 | 2378.63 | -1.646 | -52.628 | -0.07078 | -2.26259 |
| 2,6-(O-Me) ₂ -4-Pr I | 2256 | 2247.14 | 2254.55 | 8.857 | 1.446 | 0.39260 | 0.06408 |
| 2,6-(O-Me) ₂ -4-(CO-Me) | 2683 | 2759.43 | 2705.27 | -76.430 | -22.265 | -2.84867 | -0.82986 |

Nous avons établi la dépendance I_{NGA} :

- 1^{ère} en fonction de I_{NGA}^1
- 2^{ème} en fonction de I_{NGA}^2

On constate (fig.3.1) que les indices de rétention I_{NGA}^1 sont expliqués à 98.6% par les I_{NGA} alors que les I_{NGA}^2 sont expliqués à 96.38% par les I_{NGA} (fig3.2).

Relation structure- rétention chromatographique pour les phénols.

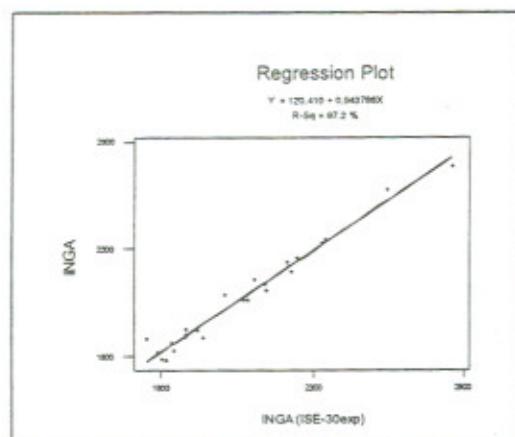


Fig. 3.1. Dispersion de I_{NGA} en fonction de I_{NGA}^1

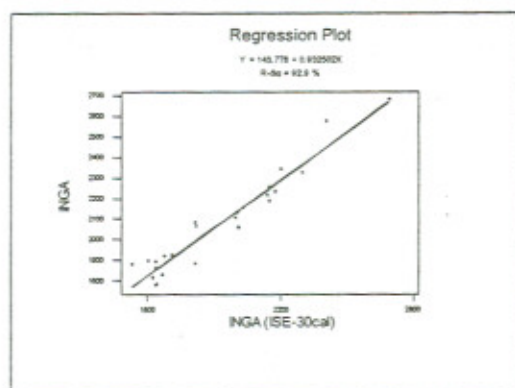


Fig.3.2. Dispersion de I_{NGA} en fonction de I_{NGA}^2

De même que précédemment, Le tableau 3.2 regroupe les valeurs des indices de rétention observées I_{OV-225} et celles calculées ($I_{OV-225}^1, I_{OV-225}^2$) sur la colonne polaire OV-225 à partir des équations (3.3) et (3.4), les résidus $ei_1 = I_{ov-225} - I_{OV-225}^1$ et

$$ei_2 = I_{ov-225} - I_{OV-225}^2.$$

Ainsi que les pourcentages des erreurs relatives.

$$I_{OV-225}^1 = f_3(I_{SE-30}^{exp} \dots) \quad (3.3)$$

$$I_{OV-225}^2 = f_4(I_{SE-30}^{esti} \dots) \quad (3.4)$$

Tableau 3.2 : Valeurs des $I_{SE-30(cal)}$ et $I_{SE-30(exp)}$ du meilleur modèle injectés dans l'équation de régression du modèle pour OV-225 :

$$I_{OV-225} = -166(\pm 58.2) + 2.36(\pm 0.05) I_{SE-30} - 20.5(\pm 2.40) VP(Ax^3) - 9.38(\pm 2.17) \text{ref.mol} - 43.1(\pm 8.3) \chi_1 + 14.7(\pm 5.21) \chi_2$$

| Nom des composés: ...phénol | I_{OV-225} | I_{OV-225}^1 | I_{OV-225}^2 | ei1 | ei2 | (ei1/ I_{OV-225})% | (ei2/ I_{OV-225})% |
|--------------------------------|--------------|----------------|----------------|----------|----------|-----------------------|-----------------------|
| 4-Me. | 1654 | 1628.57 | 1658.07 | 25.4296 | -4.070 | 1.53746 | -0.24609 |
| 3-Me. | 1648 | 1641.69 | 1674.09 | 6.3085 | -26.094 | 0.38279 | -1.58339 |
| 2, 4-(Me) ₂ . | 1660 | 1692.25 | 1706.08 | -32.2481 | -46.078 | -1.94266 | -2.77576 |
| 3-Et. | 1742 | 1741.95 | 1666.78 | 0.0527 | 75.219 | 0.00302 | 4.31795 |
| 4-Et. | 1746 | 1742.47 | 1693.92 | 3.5346 | 52.080 | 0.20244 | 2.98280 |
| 3, 5-(Me) ₂ . | 1706 | 1686.62 | 1677.86 | 19.3842 | 28.140 | 1.13624 | 1.64946 |
| 2, 3-(Me) ₂ . | 1693 | 1697.18 | 1694.82 | -4.1771 | -1.817 | -0.24673 | -0.10733 |
| 4-Cl. | 1922 | 1905.58 | 1970.50 | 16.4217 | -48.502 | 0.85441 | -2.52351 |
| 3-Cl. | 1911 | 1918.74 | 1972.39 | -7.7430 | -61.386 | -0.40518 | -3.21224 |
| 2, 4, 6-(Me) ₃ . | 1621 | 1710.33 | 1724.21 | -89.3349 | -103.212 | -5.51110 | -6.36716 |
| 3-(O-Me). | 1940 | 1865.08 | 1836.33 | 74.9207 | 103.666 | 3.86189 | 5.34358 |
| 2, 3, 5-(Me) ₃ . | 1823 | 1814.71 | 1760.48 | 8.2888 | 62.522 | 0.45468 | 3.42960 |
| 3-Br. | 2069 | 2051.27 | 2084.88 | 17.7256 | -15.881 | 0.85673 | -0.76756 |
| 4-Br. | 2054 | 2062.01 | 2089.34 | -8.0125 | -35.341 | -0.39009 | -1.72061 |
| 3-Me-4-Cl. | 2025 | 2001.62 | 1996.24 | 23.3812 | 28.762 | 1.15463 | 1.42035 |
| 2, 4, 6-(Cl) ₃ . | 1928 | 1959.90 | 1873.08 | -31.8996 | 54.925 | -1.65454 | 2.84880 |
| 2, 4, 5-(Cl) ₃ . | 2039 | 2006.03 | 2055.92 | 32.9722 | -16.918 | 1.61708 | -0.82973 |
| 3-(OH). | 2371 | 2382.25 | 2309.59 | -11.2535 | 61.411 | -0.47463 | 2.59008 |

Relation structure- rétention chromatographique pour les phénols.

| | | | | | | | |
|---|------|---------|---------|----------|----------|----------|----------|
| 2-(O-Me)-4-(CH ₃ CH=CH ₂). | 1848 | 1846.45 | 1857.61 | 1.5493 | -9.613 | 0.08384 | -0.52021 |
| 3, 5-(Cl) ₂ . | 2217 | 2215.21 | 2176.06 | 1.7924 | 40.945 | 0.08085 | 1.84685 |
| 2-(O-Me)-4-iso-Pr. | 1810 | 1907.81 | 1948.19 | -97.8095 | -138.189 | -5.40384 | -7.63476 |
| 2-(O-Me)-4-CHO. | 2199 | 2167.44 | 2229.60 | 31.5608 | -30.602 | 1.43523 | -1.39161 |
| 2,6-(O-Me) ₂ -4-Me. | 2076 | 2095.76 | 2080.70 | -19.7585 | -4.702 | -0.95176 | -0.22648 |
| 2-(O-Me)-4-(CO-Me). | 2283 | 2352.67 | 2402.19 | -69.6733 | -119.186 | -3.05183 | -5.22059 |
| 2,6-(O-Me) ₂ -4-Pr. | 2254 | 2238.45 | 2245.64 | 15.5536 | 8.356 | 0.69004 | 0.37070 |
| 2, 6-(O-Me) ₂ -4-(CO-Me). | 2685 | 2769.46 | 2716.85 | -84.4553 | -31.851 | -3.14545 | -1.18625 |

Nous avons établi la dépendance de I_{OV-225}

- 1^{ère} en fonction de I_{OV-225}^1
- 2^{ème} en fonction de I_{OV-225}^2

On constate sur la fig.3.3 que les indices de rétention calculés I_{OV-225}^1 sont expliqués à 98.6% par I_{OV-225} alors que sur la fig.3.4, I_{OV-225}^2 sont expliqués à 96.38% par I_{OV-225} .

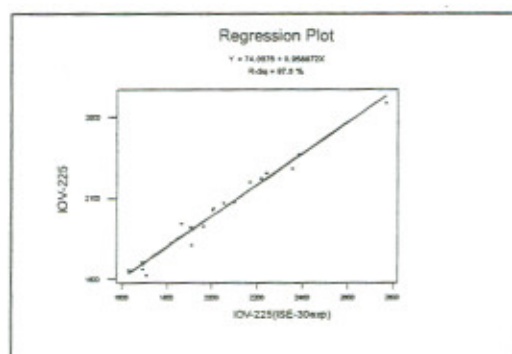


Fig.3.4. Dispersion de I_{OV-225} en fonction de I_{OV-225}^1 .

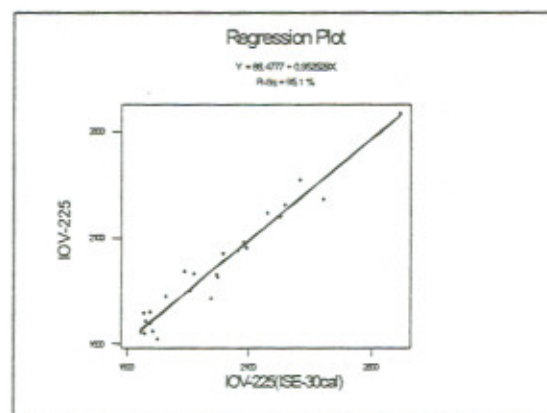


Fig.3.4. Dispersion de I_{OV-225} en fonction de I_{OV-225}^2 .

Conclusion :

Nous avons corrélé sur une colonne non polaire SE-30, les indices de rétention de 38 phénols avec des descripteurs moléculaires convenables.

Nous avons pu ainsi trouver un modèle satisfaisant construit sur 26 individus.

Nous avons calculé sur la colonne polaire OV-225, 7 modèles possibles, alors que sur la colonne polaire NGA, 5 modèles seulement ont pu être obtenus.

Grâce au modèle sélectionné sur la colonne non polaire SE-30, nous avons pu prévoir à l'avance, les indices de rétention sur les deux colonnes polaires NGA et OV-225.

Relation structure- rétention chromatographique pour les phénols.

Tableau des symboles et abréviations.

| Symbole | Désignation |
|---------------------|--|
| d | Statistique de Durbin-Watson |
| I | Indice de rétention |
| I _{OV-225} | Indice de rétention sur la colonne OV-225 |
| I _{NGA} | Indice de rétention sur la colonne NGA |
| I _{SE-30} | Indice de rétention sur la colonne SE-30 |
| Î _{SE-30} | Indices de rétention prédits pour la colonne SE-30. |
| NGA | Colonne polaire constituée de polynéopentyl glycol adipate |
| OV-225 | Colonne polaire constituée de 3-cyanopropyl polysiloxane |
| SE-30 | Colonne non polaire constituée de diméthylpolysiloxane |
| μ cal | Moment dipolaire calculé |
| VP (Ax1) | Valeur propre de la matrice de distance d'ordre 1 |
| HOMO | Energie de la haute orbitale moléculaire occupée |
| ¹ χ | connectivité moléculaire, |
| NbreO | Nombre d'atome d'oxygène |
| NbreC | Nombre d'atome de carbone |

Références bibliographiques :

1. Source : Bremond R. et Vuichard R. (1973) *Paramètres de la qualité des eaux*, Ministère de la protection de la nature et de l'environnement, SPEPE, Paris, 179 p.
2. K. Verschuere, "Handbook of environmental Data on Organic chemicals", 2nd Edition, Van Nostrand Reinhold Company, New York (1983).
3. Minitab release 14.1, logiciel d'analyse et traitement des données 2003.
4. J. Grzybowski., H lamparczyk. A Nasal., A Radecki., *J.chromatogr.* , **196, 217** (1980).
5. R. Kaliszan, HD Höltinge , *J.chromatogr.* , **234, 303** (1982).
6. L. B. Kier, L. H. Hall, W. J. Murray, M. Randić', *J. Pharm. Sci.*, **64** (1975) 1971.
7. N.R. Draper, H. Smith, "Applied regression analysis", J. Wiley, New York (1981).
8. J. Durbin, G.S. Watson, *Biometrika*, **37, 409** (1950).
9. J. Durbin, G.S. Watson, *Biometrika*, **38, 159** (1951).
10. M. Rehaïlia, "Modèles linéaires statistiques : Modèles à effets fixes", Office des Presses Universitaires, Alger (1995).