Model of Robust Regression with Least Absolute Deviation Regression and Ordinary Linear Programming Methods of gas chromatographic Retention indices of series compounds of pyrazines

Fatiha Mebarki Laboratory of Environmental security and Food; University Badji Mokhtar Annaba, Algeria
Khadija Amirat Laboratory of Environmental security and Food; University Badji Mokhtar Annaba, Algeria
Salima Ali Mokhnache Laboratory of Environmental security and Food; University Badji Mokhtar Annaba, Algeria
Djelloul Messadi Laboratory of Environmental security and Food; University Badji Mokhtar Annaba, Algeria

ABSTRACT

LAD regression diagnostics offers alternative dicapproaches whose main feature is the robustness. Here a non-parametric method for detecting influential observations is presented and compared with least squares regression method. Was applied to Model, separately, the retention indices of the same set of (89 pyrazines of test and 25 of validation) eluted in turn on OV-101 and Carbowax-20M, using theoretical molecular descriptors derived from DRAGON Software.

I-Introduction

Pyrazines are heterocyclic very present in our food. More than 80 derived from pyrazines were identified in a great number of cooked food, like the bread, the meat, the torrefied coffee, the cocoa or the hazel nusthey are very powerful aromatizing compounds

Mihara and Enomoto (1985), described a relation structure/re- tention for a unit of substituted pyrazines for which the increments of indices relating to various substituents on the cycle were given for a small series of substituents present. The method was then extended to integrate others substituents, by adding a term which takes account of the position on the cycle of a substituent compared to the others (Mihara & Masuda, 1987).In a similar approach, Masuda and Mihara (1986) describe the use of indices of connectivity modified to calculate in advance the indices of retention of a series of substituted pyrazines. The methods lead to good results, in so far as the increments of indices deter- mined in experiments available for the unknown compounds are implied, which constitutes their principal defect.

Stanton and Jurs (1989), used methodology QSRR to develop models connecting the structural characteristics of 107 variously substituted pyrazines, with their indices of retention obtained on two columns of very different polarities (OV-101 and Carbo- wax-20M). The equations were calculated using the multilinear regression, the choice of the explanatory variables (topological, electronic and physical properties) being realized by progres- sive elimination (Swall & Jurs, 1983), among the 85 individual molecular descriptors obtained for each whole molecule. The indices of retention (IR) obtained on each column were treated separately, while drawing from the same sets of descriptors. The models calculated with 6 explanatory variables provide high standards errors (S = 23 units of index - u.i. - on OV-101 and S = 36.33 u.i. on Carbowax -20 M) which do not predict good pre- dictive capacities for these models, and which let suppose non- linear relations between descriptors and property (IR) studied.

The objective of this work aims at using methodology QSSR, the approach Method LAD /Least square (LAD/OLS) , to model the indices of retention of (114) pyrazines reported from Davit T .Stanton and Peter C.Jurs (1989) and reported from Mihara and Enomoto (1985), the molecular descriptors being only calculated starting from the chemical structure of the compounds.

The linear statistical model for fixed purposes will be examined by two robust methods for the evaluation of the parameters of regression starting from estimates of the robust coefficients of regression most popular by the appendices. We based ourselves on the comparison between the two methods, the applicability (DA) will be discussed using the diagram of Williams who rep- resents the residues of prediction standardized according to the values of the levers (hi) (Eriksson et al., 2003;Tropsha et al., 2003).

II.Methodology

II.1Descriptor Generation. One used the molecular software of modeling Hyperchem 6.03, for to represent the molecules, then using semi-empirical method AM1 (Dewar et al., 1985;Holder 1998) to obtain the final geometries. It is established (Levine, 2000) that this Method gives good results when one treats small molecules (of less than one hundred atoms), like those consid- ered in this work.

The optimized geometries were transferred in the software dragon from data-processing software version 5.4[19], for the calculation of 1320 descriptors while operating on 89 pyrazines of test; subsets of descriptors were chosen by genetic algorithm, these descriptors can be separate in four categories: topological descriptors of The topological, geometrical, physical, and elec- tronic accounts of way and molecular indices of connectivity included. The geometrical descriptors included sectors of shade, the length with the reports/ratios of width, volumes of van der Waals, the surface, and principal moments of inertia. The calcu- lated descriptors of physical property included the molecular re- fringency of polarizability and molar. The electronic descriptors included most positive and most negative described by Kaliszan.

By employing the software Mobydigs (Todeschini et al., 2009) [21] and by maximizing the coefficient of prédictionQ2 and min- imal R2 of S (the error).

II.2Regression Analysis. The analysis of the multiple linear regressions was carried out with two methods by software Matlab(R2009a) for (LAD) and Minitab 16 for (OLS).

One considers the multiple model of regression given by [9]:

\[ y_i = \beta_0 + \sum_{j=1}^{p-1} \beta_j x_{ij} + e_i \]  

(1)

The detection of meaningless statements and ‘ with action lever- age according to the method of least squares is a problem which ‘ was largely studied. The diagnosis by the regression LAD offers alternative approaches whose principal characteristic is the robustness. In our study a non-parametric method to detect the meaningless statements and the point’s lever was applied and compared with the traditional method of diagnosis (least
The retention data for the 114 compounds chromatographed on the OV-101 and CRW-20M stationary phases were taken from (113 taken from Davit T. Stanton and Peter C. Jurs (1) and 1 compound (2-VinylPyrazine) taken from Mihara and Enomoto (29)] and are listed in table 1.

### IV. Results and discussion

An ideal model is one that has a high R value, allow standard error, and the fewest independent variables [1, 9]. The best models found has 3 descriptors for each stationary phase by using the software MobyDigs [21] are given below.

The criterion for identifying a compound as an outlier was that compound being flagged by three or more of six standard statistical tests used to detect outliers in regression analysis. These tests were (1) residual, (2) standardized residual, (3) Studentized residual, (4) leverage, (5) DFFITS, (6) Cook's distance. The residual is the difference between the actual value and the value predicted by the regression equation. The standardized residual is the residual divided by the standard deviation of the regression equation. The Studentized residual is the residual of a prediction divided by its own standard deviation.

Leverage allows for the determination of the influence of a point in determining the regression equation. DFFITS describes the difference in the fit of the equation caused by removal of a given observation, and Cook's distance describes the change in a model coefficient by the removal of a given point.

#### Table 1

<table>
<thead>
<tr>
<th>n°</th>
<th>Compounds</th>
<th>Retention Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Methylpyrazine</td>
<td>801</td>
</tr>
<tr>
<td>3</td>
<td>2,3-dimethylpyrazine</td>
<td>897</td>
</tr>
<tr>
<td>4</td>
<td>2,5-dimethylpyrazine</td>
<td>889</td>
</tr>
<tr>
<td>5</td>
<td>2,6-dimethylpyrazine</td>
<td>889</td>
</tr>
<tr>
<td>6</td>
<td>Trimethylpyrazine</td>
<td>981</td>
</tr>
<tr>
<td>7</td>
<td>Trimethylpyrazine</td>
<td>1067</td>
</tr>
<tr>
<td>8</td>
<td>Ethylpyrazine</td>
<td>894</td>
</tr>
<tr>
<td>9</td>
<td>2-ethyl-5-methylpyrazine</td>
<td>980</td>
</tr>
<tr>
<td>10</td>
<td>2-ethyl-6-methylpyrazine</td>
<td>972</td>
</tr>
<tr>
<td>11</td>
<td>2,5-dimethyl-3-ethylpyrazine</td>
<td>1059</td>
</tr>
<tr>
<td>12</td>
<td>2,6-dimethyl-6-ethylpyrazine</td>
<td>1064</td>
</tr>
<tr>
<td>13</td>
<td>2,3-dimethyl-5-ethylpyrazine</td>
<td>1066</td>
</tr>
<tr>
<td>14</td>
<td>2,3-diethylpyrazine</td>
<td>1065</td>
</tr>
<tr>
<td>15</td>
<td>2,3-diethyl-5-methylpyrazine</td>
<td>1137</td>
</tr>
<tr>
<td>16</td>
<td>Propylpyrazine</td>
<td>986</td>
</tr>
<tr>
<td>17</td>
<td>2-methyl-3-propylpyrazine</td>
<td>1072</td>
</tr>
<tr>
<td>18</td>
<td>2,3-dimethyl-5-propylpyrazine</td>
<td>1154</td>
</tr>
<tr>
<td>19</td>
<td>2,5-dimethyl-3-propylpyrazine</td>
<td>1132</td>
</tr>
<tr>
<td>20</td>
<td>2,6-methyl-3-propylpyrazine</td>
<td>1151</td>
</tr>
<tr>
<td>21</td>
<td>Isopropylpyrazine</td>
<td>949</td>
</tr>
<tr>
<td>22</td>
<td>2,3-dimethyl-5-isopropylpyrazine</td>
<td>1112</td>
</tr>
<tr>
<td>23</td>
<td>Butylpyrazine</td>
<td>1088</td>
</tr>
<tr>
<td>24</td>
<td>2-butyl-5-methylpyrazine</td>
<td>1121</td>
</tr>
<tr>
<td>25</td>
<td>Compounds</td>
<td>ov-101</td>
</tr>
<tr>
<td>26</td>
<td>3-butyl-3,5-dimethylpyrazine</td>
<td>1184</td>
</tr>
<tr>
<td>27</td>
<td>3-butyl-3,6-dimethylpyrazine</td>
<td>1196</td>
</tr>
<tr>
<td>28</td>
<td>5-butyl-2,3-dimethylpyrazine</td>
<td>1254</td>
</tr>
<tr>
<td>29</td>
<td>Isobutylpyrazine</td>
<td>1043</td>
</tr>
<tr>
<td>30</td>
<td>2,3-dimethyl-5-isobutylpyrazine</td>
<td>1200</td>
</tr>
<tr>
<td>31</td>
<td>2-isobutyl-3,5,6-trimethylpyrazine</td>
<td>1283</td>
</tr>
<tr>
<td>32</td>
<td>sec-butylpyrazine</td>
<td>1040</td>
</tr>
<tr>
<td>33</td>
<td>5-sec-butyl-2,3-dimethylpyrazine</td>
<td>1194</td>
</tr>
<tr>
<td>34</td>
<td>Pentylpyrazine</td>
<td>1192</td>
</tr>
<tr>
<td>35</td>
<td>2,3-dimethyl-5-pentylpyrazine</td>
<td>1352</td>
</tr>
<tr>
<td>36</td>
<td>Isopentylpyrazine</td>
<td>1157</td>
</tr>
<tr>
<td>37</td>
<td>2,3-dimethyl-5-isopentylpyrazine</td>
<td>1317</td>
</tr>
<tr>
<td>38</td>
<td>(2-methylbutyl)pyrazine</td>
<td>1151</td>
</tr>
<tr>
<td>39</td>
<td>2-(2-methylbutyl)-2,5,6-trimethylpyrazine</td>
<td>1363</td>
</tr>
</tbody>
</table>
40. (2-methyl-3-pentyl)pyrazine 1240 1240 1240 1240 1240
41. (2-methyl-3-pentyl)pyrazine 1121 1121 1121 1121 1121
42. 1-methylbutylpyrazine 1133 1133 1133 1133 1133
43. 2,5-dimethyl-5-(2-methylpentyl)pyrazine 1377 1377 1377 1377 1377
44. Hexylpyrazine 1293 1293 1293 1293 1293
45. Octylpyrazine 1495 1495 1495 1495 1495
46. 2-methyl-3-octylpyrazine 1546 1546 1546 1546 1546
47. 2-methyl-5-(2-methylbutyl)-3-octylpyrazine 1923 1923 1923 1923 1923
49. Methoxy pyrazine 877 877 877 877 877
50. 2-methoxy-3-methylpyrazine 954 954 954 954 954
51. 2-methoxy-5-methylpyrazine 969 969 969 969 969
52. 3-ethyl-2-methoxy pyrazine 1037 1037 1037 1037 1037
53. 3-isopropyl-2-methoxy pyrazine 1078 1078 1078 1078 1078
54. 5-isopropyl-3-methyl-2-methoxy pyrazine 1170 1170 1170 1170 1170
55. 5-sec-butyl-3-methyl-2-methoxy pyrazine 1250 1250 1250 1250 1250
56. 5-isobutyl-3-methyl-2-methoxy pyrazine 1257 1257 1257 1257 1257
57. 3-methyl-2-methoxy-5-(2-methylbutyl) pyrazine 1362 1362 1362 1362 1362
58. 3-methyl-2-methylthio-5-(2-methylpentyl) pyrazine 1444 1444 1444 1444 1444
60. Ethoxy pyrazine 959 959 959 959 959
61. 2-ethoxy-3-methylpyrazine 1029 1029 1029 1029 1029
62. 2-ethoxy-5-methylpyrazine 1047 1047 1047 1047 1047
63. 2-ethoxy-3-isopropylpyrazine 1101 1101 1101 1101 1101
64. 2-ethylthio-5-isopropyl-3-methylpyrazine 1143 1143 1143 1143 1143
65. 2-ethylthio-5-isobutyl-3-methylpyrazine 1134 1134 1134 1134 1134
66. 5-sec-butyl-2-ethoxy-3-methylpyrazine 1306 1306 1306 1306 1306
67. 2-ethoxy-3-methyl-5-(2-methylbutyl) pyrazine 1415 1415 1415 1415 1415
68. (methylthio)pyrazine 1076 1076 1076 1076 1076
69. 3-methyl-2-(methylthio)pyrazine 1151 1151 1151 1151 1151
70. 5-methyl-2-(methylthio)pyrazine 1163 1163 1163 1163 1163
71. 3-ethyl-2-(methylthio)pyrazine 1237 1237 1237 1237 1237
72. 3-isopropyl-2-(methylthio)pyrazine 1273 1273 1273 1273 1273
73. 3-isopropyl-3-(methylthio)pyrazine 1362 1362 1362 1362 1362
74. 5-sec-butyl-3-methyl-2-(methylthio) pyrazine 1441 1441 1441 1441 1441
75. 5-isobutyl-3-methyl-2-(methylthio)pyrazine 1446 1446 1446 1446 1446
76. 3-methyl-5-(2-methylbutyl)-2-(methylthio) pyrazine 1552 1552 1552 1552 1552
77. 3-methyl-5-(2-methylpentyl)-2-(methylthio) pyrazine 1638 1638 1638 1638 1638
78. (ethy lthio)pyrazine 1148 1148 1148 1148 1148
79. 2-ethylthio-3-methylpyrazine 1215 1215 1215 1215 1215
80. 2-ethylthio-3-isopropyl-3-methylpyrazine 1418 1418 1418 1418 1418
81. 5-sec-butyl-2-ethylthio-3-methylpyrazine 1494 1494 1494 1494 1494
82. 2-ethylthio-5-isobutyl-3-methylpyrazine 1496 1496 1496 1496 1496
83. 2-ethylthio-3-methyl-5-(2-methylbutyl) pyrazine 1602 1602 1602 1602 1602
84. 2-ethylthio-3-methyl-5-(2-methylpentyl) pyrazine 1686 1686 1686 1686 1686
85. Phenoxypyrazine 1415 1415 1415 1415 1415
86. 2-methyl-3-phenoxypyrazine 1465 1465 1465 1465 1465
87. 5-isopropyl-3-methyl-2-phenoxypyrazine 1620 1620 1620 1620 1620
88. 5-sec-butyl-3-methyl-2-phenoxypyrazine 1694 1694 1694 1694 1694
89. 5-isobutyl-3-methyl-2-phenoxypyrazine 1706 1706 1706 1706 1706
90. 3-methyl-5-(2-methylpentyl)-2-phenoxypyrazine 1807 1807 1807 1807 1807
92. (phenylthio)pyrazine 1606 1606 1606 1606 1606
93. 3-methyl-2-(phenylthio)pyrazine 1658 1658 1658 1658 1658
94. 3-isopropyl-3-methyl-2-(phenylthio) pyrazine 1806 1806 1806 1806 1806
95. 5-sec-butyl-3-methyl-2-(phenylthio) pyrazine 1874 1874 1874 1874 1874
96. 5-isobutyl-3-methyl-2-(phenylthio)pyrazine 1882 1882 1882 1882 1882
97. 3-methyl-5-(2-methylpentyl)-2-(phenylthio) pyrazine 2064 2064 2064 2064 2064
98. Acetylpyrazine 993 993 993 993 993
99. 2-acetyl-3-methylpyrazine 1061 1061 1061 1061 1061
100. 2-acetyl-5-methylpyrazine 1093 1093 1093 1093 1093

Volume : 5 | Issue : 1 | JANUARY 2016 • ISSN No 2277 - 8179
Research Paper
The definition of each descriptor is given table 2:

Table II
Definitions of Descriptors used in the Retention index Prediction Models [19].

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPC03</td>
<td>Molecular path count of order 03</td>
</tr>
<tr>
<td>GATS5e</td>
<td>Geary autocorrelation-lag 5/weighted by atomic Sanderson electronegativities</td>
</tr>
</tbody>
</table>

AFigp: Eigen value distance matrix sum from Poisson arizability weight (Barysz matrix)
Qpos: total positive charge
Se: sum of atomic Sanderson electronegativities
Mp: mean atomic polarizability (scaled on Carbon atom)
X1sol: salvation connectivity index chi-1

The coefficient of multiple determinations (R2) indicates the amount of variance in the data set accounted for by the model. The standard error of the regression coefficient is given in each case, and n indicates the number of molecules involved in the regression analysis procedure[1,9].

IV .1. The best models:
IR(OV-101):(MPC03,X1sol,GATS5e,AEigp,L3e,Qpos);S=20.892, R2=99,30,n=89 compounds.
IR(RWC) : Se, Mp, X1sol,DP01,Mor06v,Tm;S=22.64, R2=99.22,n=89 compounds.

The diagnostic statistics joined together in Table 3 make it possible to make comparisons and to draw several conclusions [21].

Table III
Diagnostiques Statistiques pour les Modèles Sélectionnés

<table>
<thead>
<tr>
<th>ID</th>
<th>Size</th>
<th>Models</th>
<th>R2</th>
<th>Q2</th>
<th>Qext</th>
<th>R2adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>OV-</td>
<td>101</td>
<td>MPC03 ,X1sol,GATS5e,AEigp,L3e,Qpos</td>
<td>99.30</td>
<td>99.12</td>
<td>98.99</td>
<td>96.94</td>
</tr>
</tbody>
</table>

Values of R2 and of R2 (adj) show, each time, quality of adjustment, whereas the very weak differences between R2 and Q2 inform about the robustness of the models which are, moreover, very highly significant (high values of the statistics F of Fisher).

Moreover, the similarity of SDEP and SDEC mean that the internal capacities of prediction models are not too dissimilar their capacities of adjustment.

The validation by bootstrap (QBOOT) confirms all at the same time the capacity of internal prediction and the stability of the models.

IV.2. Robust Regression:
Any robust method must be reasonably effective once compared to the estimators of least squares; if the fundamental distribution of the errors is normal and primarily more effective independent than the estimators of least squares, when there are peripheral observations. There are various robust methods for the evaluation the parameters of regression. The principal goal of this section is the method LAD (nap of the absolute values of the errors) whose coefficient of regression qualifies the robustness among the additional data [16].

IV.2.1. Comparison Robust Regression of OLS and LAD:
More particularly we will test 2 methods of estimate for the vector of the Parameters (β_0, β_1, ..., β_k) - Method of least squares ordinary, more known and the most used. - The method LAD (Sum of the absolute values of the errors.)
The large advantage of the method LAD is its robustness, i.e. that the estimators are not impact by the extreme values, (they are known as "robust"). It is thus particularly interesting to use the method LAD if one is in the presence of aberrant values in comparison with method OLS [8].

IV.2.1.1. Comparison of hyperplanes of regression: Column OV-101:
1/LAD:
\[ Y = -48.05 -10.14 \text{MPC03} + 337.87 \text{X1sol} -35.78 \text{GATS5e} -2.54 \text{AEigp} -38.51 \text{L3e} -156.88 \text{Qpos} \] (4)

2/OLS:
\[ Y = -31.2 - 7.77 \text{MPC03} + 300 \text{X1sol} - 24.9 \text{GATS5e} + 2.31 \text{AEigp} - 53.1 \text{L3e} - 62.6 \text{Qpos} \] (5)

Column CRW -20M:
1/LAD:
\[ Y = -242, 89 -42, 45 \text{Se} + 687, 45 \text{Mp} + 298,16 \text{X1sol} + 205, 42 \text{DP01} + 200,62 \text{Mor06v} + 8,04 \text{Tm} \] (6)

2/OLS:
\[ Y = -167 - 42, 8 \text{Se} + 755 \text{Mp} + 320 \text{X1sol} + 130 \text{DP01} + 163 \text{Mor06v} + 10,7 \text{Tm} \] (7)

Each equation on each column check the assumptions on the same linear statistical model for Fixes purposes for each method in comparison with the hyperplane calculated by LAD compared to the hyperplane calculated by the method of least squares.

It is noticed that \( \beta \) the calculated OLS are not very different from the regression with \( \beta \) the LAD on the two columns, except, \( \beta \)the calculated OLS is almost the same ones as for the regression with \( \beta \)the LAD on column CRW and \( \beta \)the calculated OLS is almost the same ones as for the regression with \( \beta \)the LAD on column OV-101.

It is thus relevant to remake a checking of the presences of aberrant values by using the following stage (figure 3):

The hyperplane of regression can radically change, with the change of the coefficients of the hyperplane.

IV.2.1.2. Graphical Comparisons of Alternative Regression Models
The field of application was discussed using the diagram of Williams.

Columns RW -20M

Method LAD and OLS (test, validation)
Fig.1 Diagram of Williams of the residues of prediction standardized according to the lever:

The analysis of the residues shows that the observations (82,25) residues raised but it (48)point influence in the two estimates and the observation( 12) point influence with the LAD estimate and lever by least square also observation 4 residue raised with OLS and not lever with LAD in the whole of validation on column OV -101 and on column CRW -20M the observations ( 45 ) not influence in the two estimates and observation 16 point influence in the two estimates in the whole of validation.

After elimination of the aberrant points collective between the two methods and after the secondary treatment one has the observation (12) point influence and the observations (1, 24) residues raised in the two estimates but it (25) observation 4 residue raised with OLS and not lever with LAD also the observation 4 residue raised in the whole of validation in the two estimates on column OV -101 and on column CRW -20M the observations ( 45 ) not influence in the two estimates and observation 16 point influence in the two estimates in the whole of validation and on column CRW -20M the observations ( 24 25 35 ) residues raised but it (84)point influence in the two estimates and observation 8 point influence in the two estimates in the whole of validation.

Thus finally the models in which the meaningless statements were removed become after elimination of the aberrant points collective [OV-101: test - (1, 12, 24), validation (4), CRW-20M: test - (24, 25, 35 84), validation (8)] between the two methods:

Column OV-101:
1/LAD:
\[ y = -48.05 -10.14 \text{MPC03} + 337.87 \text{X1sol} -35.78 \text{GATS5e} -2.54 \text{AEigp} -38.51 \text{L3e} -156.88 \text{Qps} \] (8)

2/OLS:
\[ y = -61.1 - 9.80 \text{MPC03} + 343 \text{X1sol} - 35.7 \text{GATS5e} - 2.80 \text{AEigp} - 40.7 \text{L3e} - 160 \text{Qpos} \] (9)

Column CW -20M:
1/LAD:
\[ Y = -242, 89 -42, 45 \text{Se} + 687, 45 \text{Mp} + 298, 16 \text{X1sol} + 205, 42 \text{DP01} + 200, 62 \text{Mor06v} + 8,04 \text{Tm} \] (10)

2/OLS:
\[ \text{IR (RCW)} = -192 - 42, 4 \text{Se} + 752 \text{Mp} + 305 \text{X1sol} + 155 \text{DP01} + 156 \text{Mor06v} + 13, 0 \text{Tm} \] (11)

It is noticed besides that \( \beta \) the OLS calculate more to approach which for the regression with \( \beta \) the LAD on the two columns into precise (\( \beta 1, \beta 3 \) and \( \beta 4 \)) the OLS calculate are almost the
same ones as for the regression with $\beta_1$, $\beta_3$ and $\beta_4$, the LAD and on the same order with $\beta_5$ and $\beta_6$) on OV 101 and $\beta_1$ the OLS calculate are almost the same ones as for the regression with $\beta_1$ the LAD on CRW-20M.

The analysis of the residuals shows that in this case, all the point of lad method between (-2, 2), but it the analysis of the residues of OLS method shows that the observations [OV-101: test (-6, 42), CRW-20M: test (-22, 24, 67, 78), validation (7, 13, 14)] the LAD estimate given good result On the other hand estimate OLS figure (4):

IV.2.1.3 Graphical Comparisons of Alternative Regression Models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data Set</th>
<th>Regression Method</th>
<th>Coefficient Remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>e stand lad vs HI1</td>
<td>SRES1 vs HI1</td>
<td>LAD</td>
<td>$\beta_2 (LAD) \approx \beta_2 (OLS)$</td>
</tr>
<tr>
<td>SRES vali vs HI vali</td>
<td>e stand lad vs HI1</td>
<td>OLS</td>
<td>$\beta_1 (LAD) \approx \beta_1 (OLS)$</td>
</tr>
<tr>
<td>X-Data</td>
<td>Y-Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-2$</td>
<td>$-1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0$</td>
<td>$1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$2$</td>
<td>$3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$4$</td>
<td>$5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$6$</td>
<td>$7$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$8$</td>
<td>$9$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10$</td>
<td>$11$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$12$</td>
<td>$13$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$14$</td>
<td>$15$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lastly, it is noted that LAD is a robust estimator but loses stability in the presence of points aberrant.

We note however the observation that the estimate the least square is near to the LAD estimate to which removed the aberrant values.

V. Conclusions:

The modeling of the indices of retention of 114 pyrazines (89 tests and 25 validations) eluted out of two columns various OV-101 and CRW-20M by two methods LAD and OLS are based on the following comparisons:

- The comparison of the equations of the hyperplanes:

L equations of OLS is closer to LAD after elimination of the aberrant points for the $\beta_2$ (LAD) $\approx$ ($\beta_2$ (OLS)) and the other coefficient remaining with the same order for column OV-101. Pour the column Crw-20m the $\beta_1$ (LAD) $\approx$ ($\beta_1$ (OLS)) and the other coefficient remaining with the same order after the secondary treatments for the checking of the presence of aberrant values (82, 48, 26, 25, 24, 12, 1) on column OV-101 and item (45, 82, 25, 24) for the column CRW-20M, and to be able to compare them By employing the following stage.

- Graph comparison: The applicability was discussed using the diagram of Williams in dependence.

Lastly, it is noted that LAD is a robust estimator but loses his stability in the presence of aberrant points.

It general this study is shown that results by the two estimates theoretical (equation) and graph give good results expressed by the models.

References:

[27] Related values.