CHAPTER 5

RESULTS AND DISCUSSIONS

The retention indices of the examined compounds and their parameters are shown in Table 5.1. Most of the compounds for all stationary phases elute in increasing order according to their boiling points. The few exceptions that exist have nearly the same boiling points or are isomers.

5.1 Principal Component Analysis (PCA)

By the mathematical technique of "eigenanalysis" in PCA, the data matrix will be decomposed into two matrices. The "score" matrix gives the information about the nature of the compounds, and the "loading" matrix gives the information about the nature of parameters. These results can be visualized by the score and loading plots. Figure 5.1 and 5.2 are the results for family 1 (compounds with one carbon atom and mixed number of hydrogen atoms) as the example. From the score plot, it can be seen that all the compounds in this family are classified by the number of hydrogens and fluorines. Compounds with the same number of fluorines will arrange themselves from bottom to top and from right to left according to the increasing number of hydrogens (i.e. compounds 10 to 40), while compounds with the same number of hydrogens will arrange themselves from top to bottom and from right to left according to the increasing number of fluorines (i.e. compounds 10 to 14). The loading plot shows that the retention indices of the three columns (assigned by no.1,2 and 3) are clustered together with BP (5), $^0\chi$ (6) and $^1\chi$ (7). This means that I,BP,$^0\chi$ and $^1\chi$ can explain the same behavior of the system or behaves in the same way. In this way one can have the idea that the retention index can be basically predicted by the parameters BP,$^0\chi$ and/or $^1\chi$. The parameters nF (12), nH (14) and D (15) are separated from the other parameters. The position of nF and nH confirms the result of the score plot that the compounds can be classified by these parameters in the way as described above. The position of parameter D (dipole moment) can give the idea that it may explain or behaves differently compared to the other parameters. D is known as a quantum chemical parameter which describes the electronic properties while the other parameters describe the structural or bulk properties. The score and loading plots of all products with known D values, on the three columns (Figure 5.3 and 5.4), also show the same results as for family 1. The classification of all products is based on nF (15) and nH (17). The retention indices (1,2 and 3) can be basically predicted by BP(5),$^{0}\chi$(6) and $^{1}\chi$ (7). The parameter D (19) also differs from the other parameters.

Looking at the score plot, one can see that the relationship of the compounds with nF and nH appears linear. This leads to the idea of a relationship between the retention index and nF. Figure 5.5 to 5.18 show the linear relationship within a family between the retention index of the compounds and nF on every column, except for CF_4 (14) (Figure 5.5 to 5.7). The reason for CF_4 being an "outlier" can be explained by its very low boiling point (-128 °C) compared to all other compounds. With the limitation of the operating conditions for the columns and instrument, the temperature of the columns could not decrease below -60 °C. So in this case CF_4 will not be retained in the column at the proper condition. Therefore CF_4 can be defined as an outlier which can increase the uncertainty of the model. That deleted from the raw data.

When using the PCA model to predict the retention index (Table 5.2), it is found that the model from PCA does not give a good prediction. This can be seen from the results of the tested products of all families on the apolar column as an example (Table 5.3 and Figure 5.19). The difference between the measured and predicted indices varies from about 5 to 400. This bad prediction is a result of the choice of data matrix on which the PCA analysis is performed. As mentioned in chapter 3, each row in the matrix concerns a product and each column concerns an index or a solute parameter. In its searching for the principal components that best describe the total variance in the data matrix, PCA searches the vectors (factors) that explain the variation not only in the indices but also in the solute parameters. PCA cannot make a distinction between dependent and independent variables. So it cannot give a good prediction.

Table 5.1 The retention indices and the parameters of the examined products

CODE	INDEX			PARAMETERS						
	I-NP	I-SP	I-P	MW	BP	CHI(0)	CHI(1)	CHI(2)	CHI(3)-p	CHI(3)-c
CFC WITH C 1 ATOM										
10	650	684.4	693.1	153.8	76.8	5.281	2.39	4.286	0	3.41
# 11	479	500	520	137.4	23.8	3.862	1.681	1.742	0	0.375
12	311.7	312.5	345.5	120.9	-29.8	2.443	0.972	0.205	0	-0.26
13	174.1	139.3	200	104.5	-81.4	1.024	0.262	-0.326	0	0.084
## 14	143.2	200	200	88	-128	-0.394	-0.447	0.15	0	-0.022
20	602.3	694.3	686.3	119.4	61.2	4.163	2.07	2.474	0	0.854
21	453.6	521.5	529.3	102.9	8.9	2.744	1.251	0.516	0	0.16
22	307.3	342.9	381.8	86.5	-40.8	1.325	0.432	0.28	0	0.03
23	174.1	200	233.3	70	-82.1	-0.093	-0.387	0.087	0	-0.006
# 30	520.7	602.1	628.2	84.9	40.1	3.098	1.69	1.01	0	0
32	203.8	241.9	327.3	52	-51.7	0.26	-0.316	0.035	0	0
40	333.9	389.5	435.5	50.5	-24	2.195	1.195	0	0	0
41	174.1	213.3	300	34	-78.5	0.776	-0.224	0	0	0

CODE	PARAMETERS								
	CHI(4)-p	CHI(4)-c	R	Q	nF	nCl	nH	nC	D
10	0	0	3.39	2.91	0	4	0	1	0
# 11	0	0	3.0172	2.624	1	3	0	1	0.5
12	0	0	2.5557	2.328	2	2	0	1	0.5
13	0	0	2.172	1.82	3	0	0	1	0.5
## 14	0	0	1.8016	1.84	4	0	0	1	0
20	0	0	2.87	2.41	0	3	1	1	1.1
21	0	0	2.4377	2.124	1	2	1	1	1.3
22	0	0	1.9921	1.832	2	1	1	1	1.4
23	0	0	1.5781	1.548	3	0	1	1	1.6
# 30	0	0	2.2564	1.988	0	2	2	1	1.8
32	0	0	1.4285	1.42	2	0	2	1	2
40	0	0	1.6671	1.568	0	1	3	1	1.9
41	0	0	1.2782	1.288	1	0	3	1	1.8

tested products
outlier product

Table 5.1 (continued)

CODE	INDEX I-NP	I-SP	I-P	PARAMETERS MW	BP	CHI(0)	CHI(1)	CHI(2)	CHI(3)-p	CHI(3)-c
CFC-NO H ATOM										
110	1061.9	1123.1	1132	236.8	185	8.171	3.836	6.079	3.214	3.85
112	698.4	725	768.1	203.8	92.8	2.791	3.014	1.977	1.174	-0.244
# 112a	695.7	721.4	760.8	203.8	91.5	2.791	3.014	2.984	0.671	1.834
113	532.3	536.2	600	187.4	47.6	3.915	1.707	0.933	0.405	-0.027
113a	530.7	534.8	584.2	187.4	45.7	3.915	1.707	2.318	-0.601	1.957
114	366.8	355.9	419.4	170.9	3.6	2.496	0.998	-0.111	0.14	-0.182
# 115	214	200	266.7	154.5	-38.7	1.077	0.289	-0.148	-0.125	-0.059
116	100	100	100	138	-78.1	-0.342	-0.421	-0.185	0.113	0.064

CODE	PARAMETERS CHI(4)-p	CHI(4)-c	R	Q	nF	nCl	nH	nC	D
110	7.684	0.854	5.2802	4.368	0	6	0	2	0
112	0.969	-0.16	4.3573	3.776	2	4	0	2	#
# 112a	0.367	0.442	4.4003	3.788	2	4	0	2	#
113	-0.095	-0.065	3.9389	3.492	3	3	0	2	0.4
113a	-0.584	0.424	4.0461	3.564	3	3	0	2	#
114	-0.181	0.03	3.553	3.28	4	2	0	2	0.5
# 115	0.109	0.012	3.1825	3.02	5	1	0	2	0.8
116	-0.05	-0.006	2.812	2.76	6	0	0	2	#

tested products

CODE	INDEX I-NP	I-SP	I-P	PARAMETERS MW	BP	CHI(0)	CHI(1)	CHI(2)	CHI(3)-p	CHI(3)-c
CFC WITH H 1 ATOM										
120	953.3	1061.8	1070.1	202.3	162	7.053	3.462	4.693	2.143	2.282
121	783.3	884	904.2	185.8	116.6	5.635	2.752	2.587	1.474	0.421
# 122	626.5	721.2	745.2	169.4	71.9	4.216	2.043	1.489	0.447	0.283
123	483.6	561.2	590.9	153	27.1	2.797	1.333	1.286	-0.401	0.389
123a	486.4	552.8	600	153	28.2	2.797	1.224	0.1	0.182	-0.158
124	344.8	407.6	441.9	136.5	-12	1.378	0.514	0.008	-0.163	-0.035
125	209.8	239.8	309.1	120	-48.5	-0.041	-0.305	-0.219	0.075	0.044

Table 5.1 (continued)

CODE	CHI(4)-p	CHI(4)-c	R	Q	nF	nCl	nH	nC	D
120	3.842	0.427	4.7007	3.686	0	5	1	2	1
121	1.308	-0.08	4.2393	3.572	1	4	1	2	!
# 122	-0.022	0.015	3.8371	3.324	2	3	1	2	!
123	-0.15	-0.003	3.4666	3.064	3	2	1	2	!
123a	-0.168	0.015	3.3916	3.032	3	2	1	2	!
124	0.081	-0.003	3.0211	2.772	4	1	1	2	!
125	-0.025	-0.003	2.6071	2.488	5	0	1	2	!

tested product

CFC WITH H 2 ATOMS

CODE	I-NP (INDEX)	I-SP	I-P	MW	BP	CHI(0)	CHI(1)	CHI(2)	CHI(3)-p	CHI(3)-c
# 130	878.1	1037.6	1047.5	167.9	146.3	5.936	3.094	3.243	1.429	0.714
130a	826	922	962	167.9	130.5	5.988	2.992	3.833	1.071	1.925
131	715.6	856.4	885.6	151.4	102.5	4.517	2.274	1.791	0.581	0.29
131a	676.7	765.7	820.7	151.4	88	4.569	2.282	1.636	0.647	0.064
132	567.1	681.9	729.5	135	59	3.322	1.455	0.546	0.236	-0.134
132b	535.5	606.8	674.8	135	46.8	3.15	1.573	0.445	0.224	-0.091
# 133a	401.1	455.2	525	118.5	6.1	1.732	0.863	0.26	-0.2	0.032
134	276.5	335.1	412.9	102	-19.7	0.26	-0.183	-0.24	0.05	0.025
134a	245.7	305.2	354.5	102	-26.5	0.313	-0.14	-0.241	0.038	0.032

CODE	CHI(4)-p	CHI(4)-c	R	Q	nF	nCl	nH	nC	D
# 130	1.707	0	4.1212	3.368	0	4	2	2	1.5
130a	1.281	0.427	4.1055	3.448	0	4	2	2	!
131	0.187	0	3.6757	3.076	1	3	2	2	!
131a	0.267	-0.08	3.6441	3.152	1	3	2	2	!
132	-0.13	0	3.2302	2.784	2	2	2	2	!
132b	-0.145	0.015	3.2419	2.904	2	2	2	2	!
# 133a	0.045	-0.003	2.8714	2.644	3	1	2	2	!
134	-0.011	0	2.4022	2.216	4	0	2	2	!
134a	-0.008	-0.003	2.4575	2.36	4	0	2	2	!

tested products

Table 5.1 (continued)

CODE	I-NP	INDEX I-SP	I-P	MW	BP	PARAMETERS CHI(0)	CHI(1)	CHI(2)	CHI(3)-p	CHI(3)-c
CFC WITH H 3 ATOMS										
140	742.3	866.3	907.8	133.4	113.5	4.87	2.634	2.289	0.714	0.357
# 140a	630.6	683.4	737.1	133.4	74.1	5.086	2.293	3.936	0	1.925
141	599.5	711.1	781.4	116.9	75.7	3.451	1.814	0.73	0.29	-0.067
141b	488.6	531.6	604.5	116.9	32	3.667	1.583	1.53	0	-0.064
# 142	459	550.9	637.9	100.5	35.1	2.032	0.995	0.334	-0.134	0.013
142b	342.9	382.1	459.4	10.5	-9.2	2.248	0.874	0.132	0	-0.091
143	316.7	395.5	500	84	5	0.614	-0.008	-0.245	0.025	0.013
143a	208.6	229.1	327.3	84	-47.6	0.829	0.165	-0.26	0	0.032

CODE	CHI(4)-p	CHI(4)-c	R	Q	PARAMETERS nF	nCl	nH	nC	D
140	0.427	0	3.526	2.984	0	3	3	2	1.4
# 140a	0	0.427	3.5412	3.032	0	3	3	2	1.7
141	-0.08	0	3.0805	2.656	1	2	3	2	#
141b	0	-0.08	3.0798	2.736	1	2	3	2	#
# 142	0.015	0	2.6665	2.372	2	1	3	2	2.1
142b	0	0.015	2.6776	2.488	2	1	3	2	2.2
143	-0.003	0	2.2526	2.088	3	0	3	2	#
143a	0	-0.003	2.3071	2.228	3	0	3	2	2.3

tested products

CODE	I-NP	INDEX I-SP	I-P	MW	BP	PARAMETERS CHI(0)	CHI(1)	CHI(2)	CHI(3)-p	CHI(3)-c
CFC WITH H 4&5 ATOMS										
150	623.9	727.2	791.6	98.9	83.5	3.805	2.19	1.195	0.357	0
150a	561.2	638.1	693.2	98.9	57.3	3.968	1.957	2.205	0	0.357
151a	431.5	484.3	561.1	82.5	16.1	2.549	1.138	0.407	0	-0.067
# 152a	270.5	318.5	419.4	66	-24.7	1.13	0.319	-0.229	0	0.013
160	422.1	481.6	542.1	64.5	13.1	2.902	1.552	0.845	0	0
161	278.1	315.6	406.5	48	-37.1	1.483	0.549	-0.158	0	0

CODE	CHI(4)-p	CHI(4)-c	R	Q	PARAMETERS nF	nCl	nH	nC	D
150	0	0	2.9308	2.828	0	2	4	2	1.8
150a	0	0	2.9617	2.532	0	2	4	2	2
151a	0	0	2.5162	2.24	1	1	4	2	#
152a	0	0	2.1022	1.956	2	0	4	2	2.3
160	0	0	2.3665	2.112	0	1	5	2	2
161	0	0	1.9526	1.828	1	0	5	2	2

tested product

Table 5.1 (continued)

CODE	INDEX I-NP	I-SP	I-P	PARAMETERS MW	BP	CHI(0)	CHI(1)	CHI(2)	CHI(3)-p	CHI(3)-c
CFC WITH C-DOUBLE BONDS										
1110	795.1	837	865.3	165.8	120.8	5.781	2.89	3.819	2.857	2.026
# 1112-c	459.7	474.6	530.7	132.9	21.1	2.943	1.472	0.704	0.472	-0.024
1112-t	461.6	‡	‡	132.9	22	2.943	1.472	0.704	0.472	-0.024
1112a	449.7	460.5	520	132.9	19	2.943	1.472	1.711	-0.535	0.982
1113	303.7	305.7	364.6	116.5	-27.9	1.524	0.762	0.154	-0.217	-0.043
1120	683.6	741.4	766.5	131.4	88	4.663	2.463	2.784	1.429	1.162
1121	530.4	574.1	615.9	115	35.1	3.244	1.753	1.117	0.581	0.137
1122	355	383.6	429	98.5	-17.7	1.825	1.044	0.457	-0.267	0.119
# 1123	213.9	231.4	302.7	82	-61	0.407	0.225	-0.362	0.05	-0.059
1130-c	591.1	671.7	710.8	96.9	60.1	3.545	2.047	1.594	0.714	0.229
1130-t	555	605.9	629.5	96.9	48.4	3.545	2.047	1.594	0.714	0.229
1130a	508.4	544	574.7	96.9	37	3.598	1.902	2.405	0	1.013
# 1131a	337	358.3	406.5	80.5	-25	2.179	1.193	0.553	0	-0.012
1132	294.3	336.2	412.9	64	-28	0.707	0.408	-0.075	0.025	-0.06
1132a	176.5	200	237	64	-82	0.76	0.483	-0.291	0	-0.031
# 1140	372.2	411.8	441.9	62.5	-13.9	2.48	1.507	0.976	0	0.149
1141	231.4	209.5	300	46	-72.2	1.061	0.687	-0.183	0	-0.028

CODE	CHI(4)-p	CHI(4)-c	R	PARAMETERS Q	nF	nCl	nH	nC	D
1110	3.415	0.357	3.8245	3.381	0	4	0	2	‡
# 1112-c	-0.26	-0.067	2.9967	2.813	2	2	0	2	‡
1112-t	-0.26	-0.067	2.9967	2.813	2	2	0	2	‡
1112a	-0.26	0.185	2.9967	2.813	2	2	0	2	‡
1113	0.12	-0.027	2.5828	2.529	3	1	0	2	‡
1120	0.714	0.179	3.2616	2.848	0	3	1	2	0.9
1121	-0.16	-0.033	2.8477	2.564	1	2	1	2	‡
1122	0.03	0.006	2.4338	2.28	2	1	1	2	‡
# 1123	-0.006	0.006	2.0199	1.996	3	0	1	2	‡
1130-c	0	0	2.6987	2.315	0	2	2	2	1.8
1130-t	0	0	2.6987	2.315	0	2	2	2	0
1130a	0	0.179	2.6993	2.436	0	2	2	2	‡
# 1131a	0	-0.033	2.2854	2.152	1	1	2	2	‡
1132	0	0	1.8709	1.747	2	0	2	2	2.4
1132a	0	0.006	1.8715	1.868	2	0	2	2	1.4
# 1140	0	0	2.1364	1.9	0	1	3	2	1.5
1141	0	0	1.7225	1.616	1	0	3	2	1.4

tested products

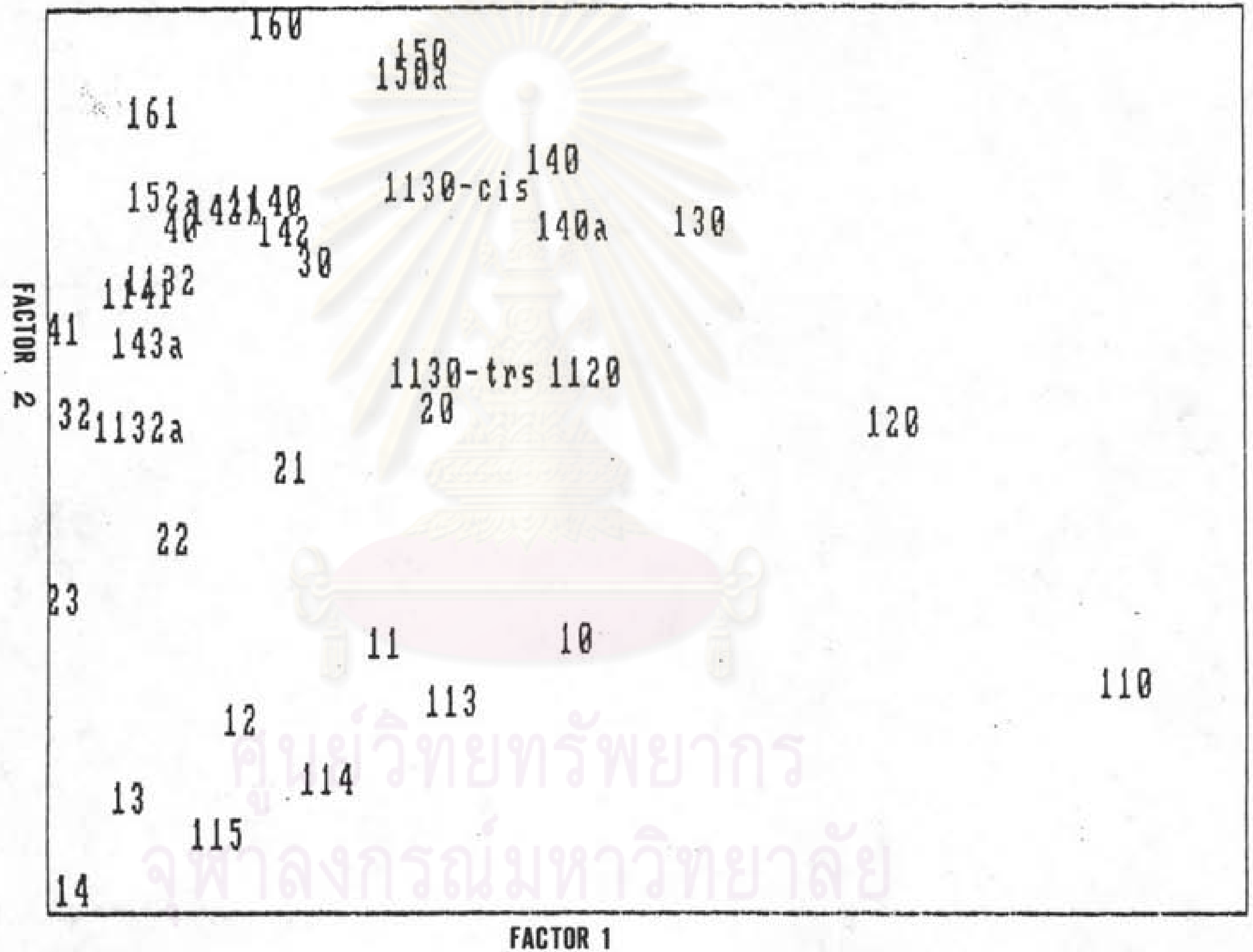
Plotted "PARVUS : ESS" Solvay-NOH

Fig. 5.1 The score plot of family 1

1 : I-NP
2 : I-SP
3 : I-P
4 : MW
5 : BP
6 : CHI(0), $^{0}\chi$
7 : CHI(1), $^{1}\chi$
8 : CHI(2), $^{2}\chi$
9 : CHI(3)-C, $^{3}\chi_{c}$
10 : R
11 : Q
12 : nF
13 : nCl
14 : nH
15 : D

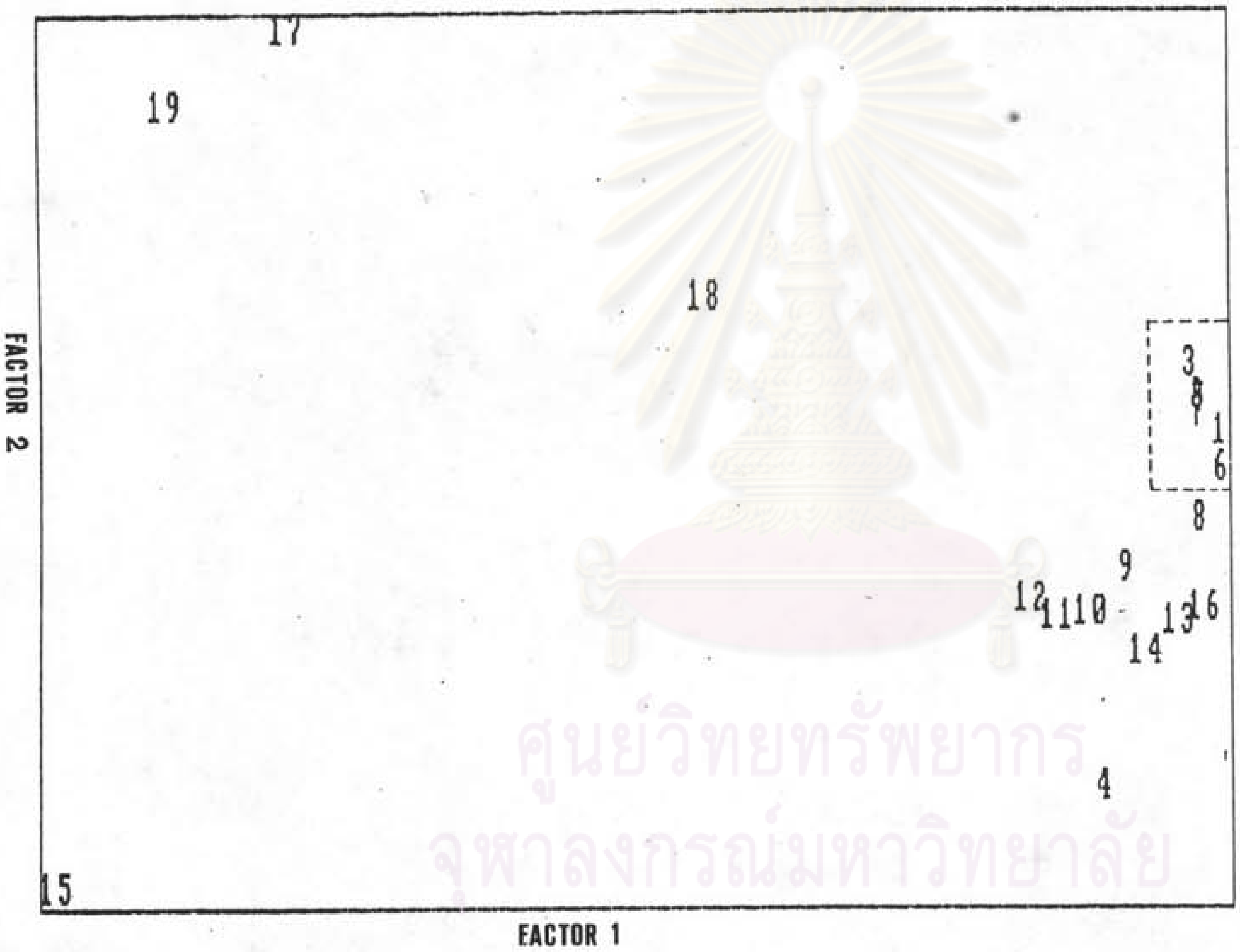
Plotted "PARVUS : ESS" Solvay-NOH

Fig. 5.2 The loading plot of family 1



Plotted "PARVUS : ESS" Solvay-NOH

Fig. 5.3 The score plot of the products which have the D values



1 : I-NP
2 : I-SP
3 : I-P
4 : MW
5 : BP
6 : CHI(0), $^{0}\chi$
7 : CHI(1), $^{1}\chi$
8 : CHI(2), $^{2}\chi$
9 : CHI(3)-P, $^{3}\chi_{P}$
10 : CHI(3)-C, $^{3}\chi_{C}$
11 : CHI(4)-PC, $^{4}\chi_{PC}$
12 : CHI(4)-C, $^{4}\chi_{C}$
13 : R
14 : Q
15 : nF
16 : nCl
17 : nH
18 : nC
19 : D

Plotted "PARVUS : ESS" Solvay-NOH

Fig. 5.4 The loading plot of retention indices and all parameters

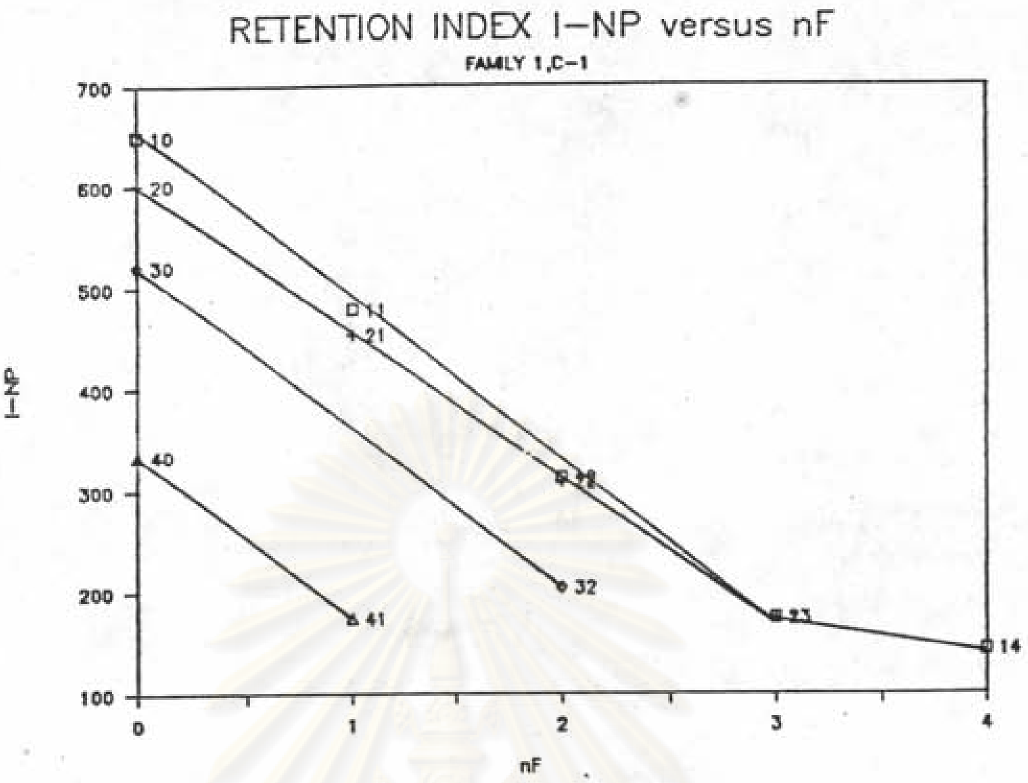


Fig. 5.5 Relationship between I and nF of family 1 on apolar column

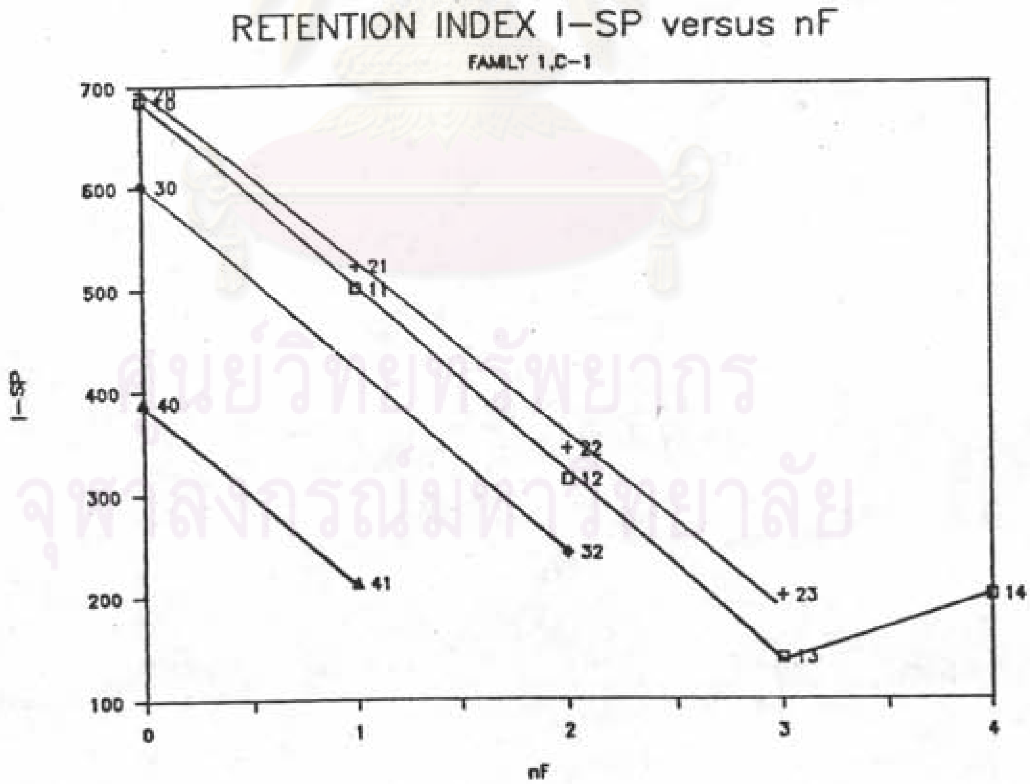


Fig. 5.6 Relationship between I and nF of family 1 on slightly polar column

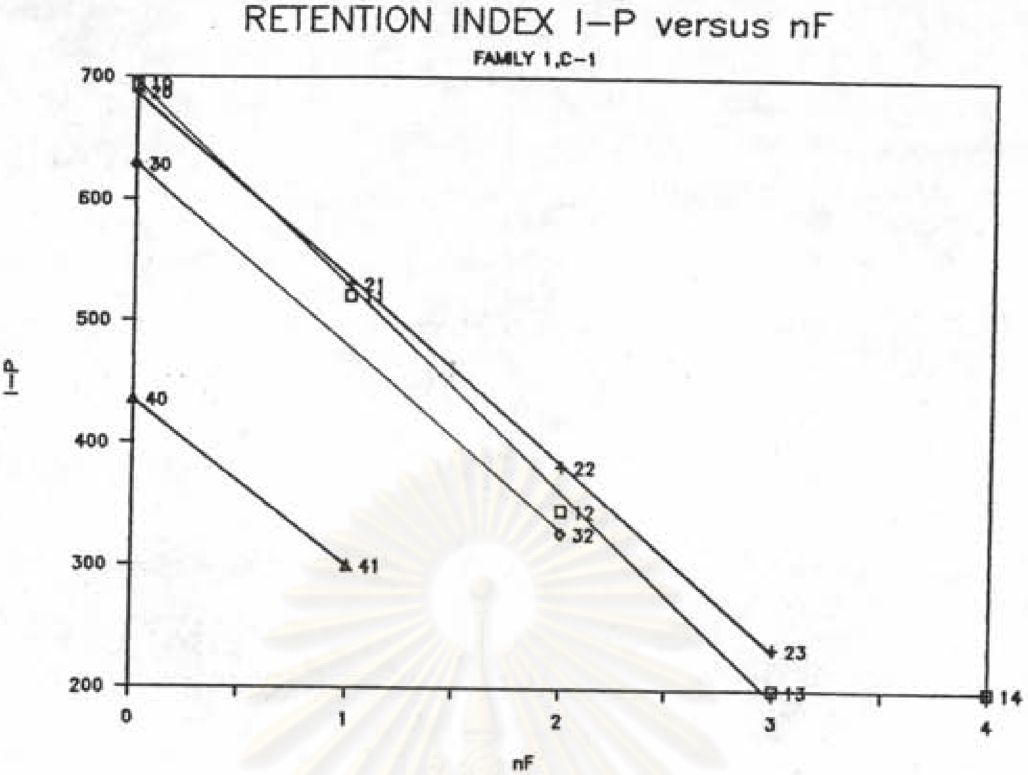


Fig. 5.7 Relationship between I and nF of family 1 on polar column

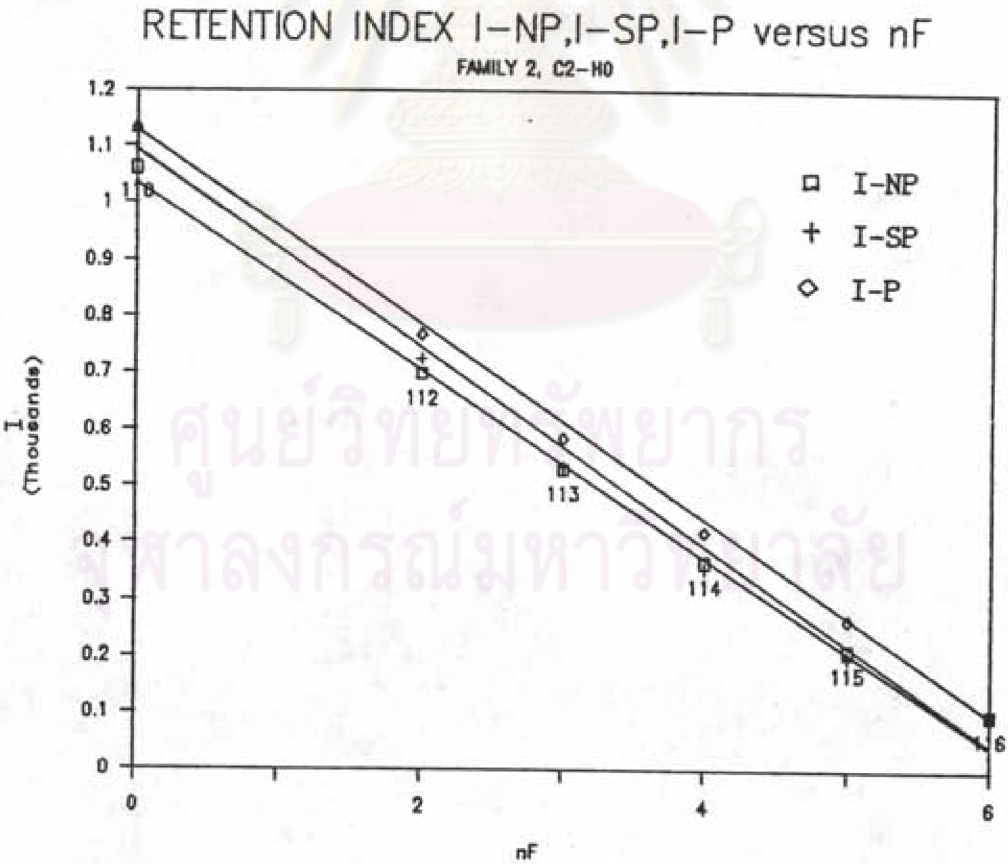


Fig. 5.8 Relationship between I and nF of family 2 on three columns

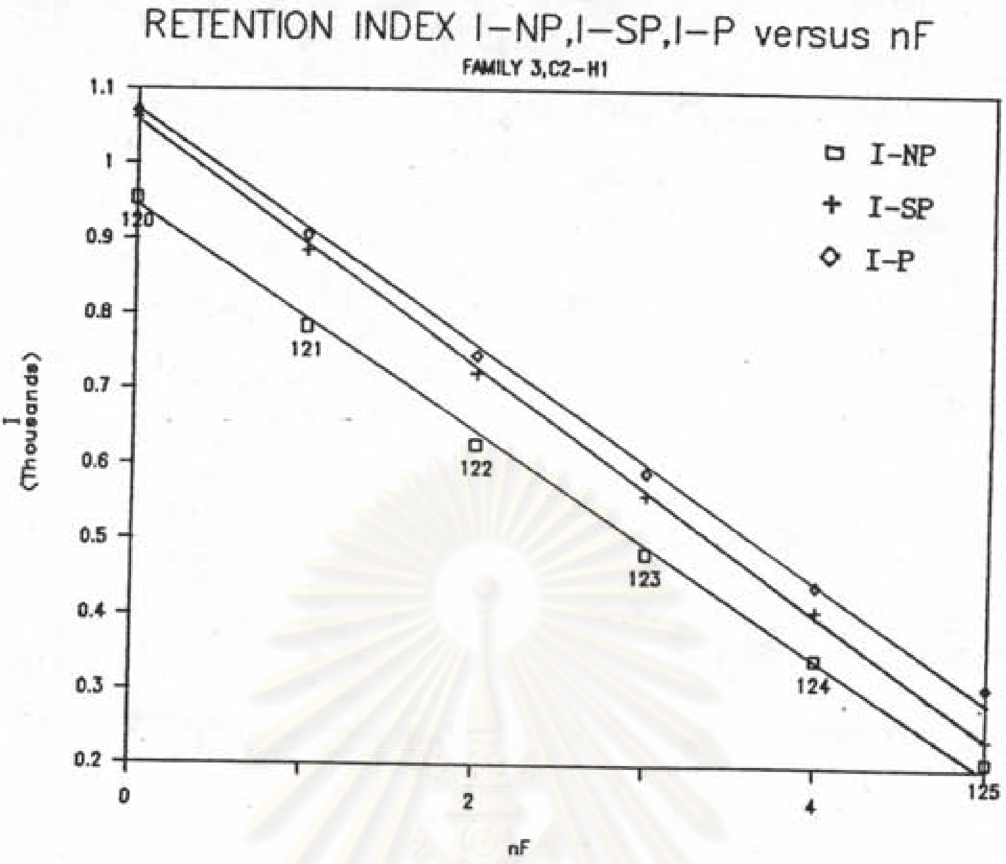


Fig. 5.9 Relationship between I and nF of family 3 on three columns

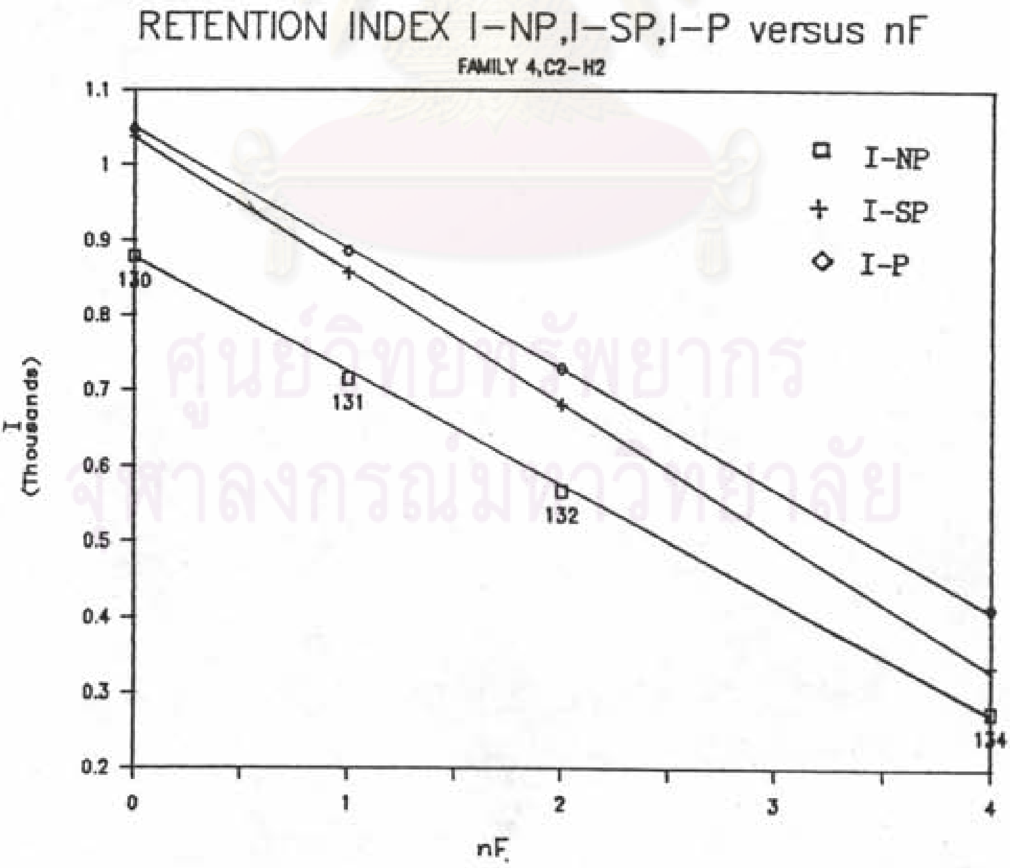


Fig. 5.10 Relationship between I and nF of family 4 on three columns

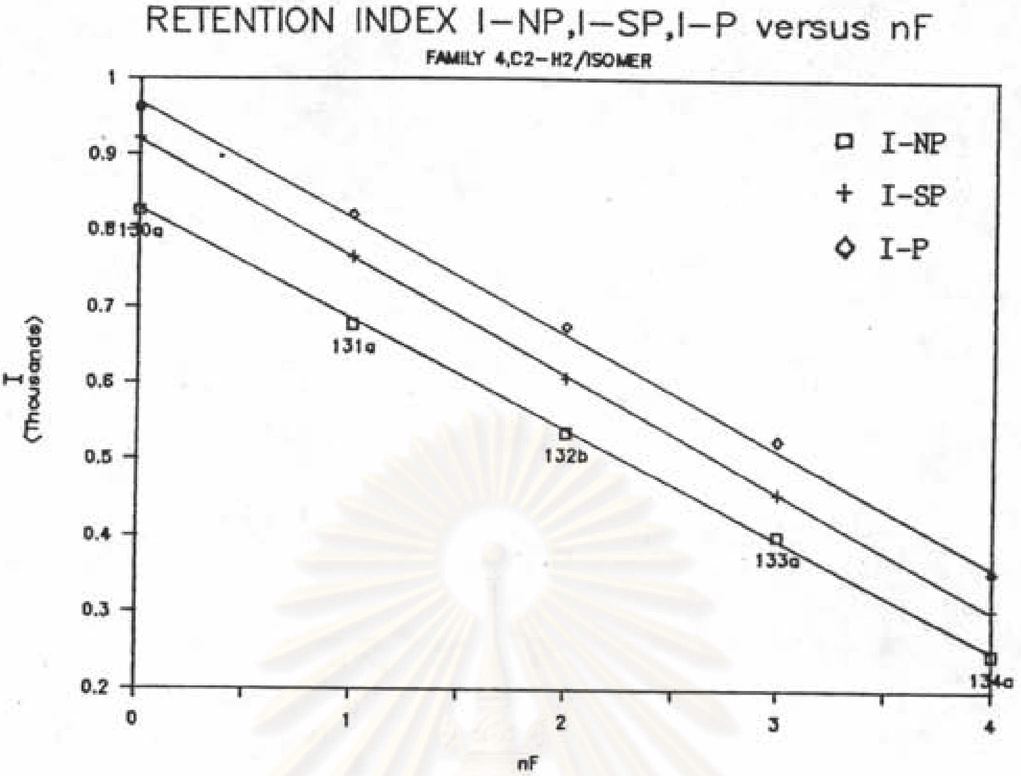


Fig. 5.11 Relationship between I and nF of family 4 (isomer) on three columns

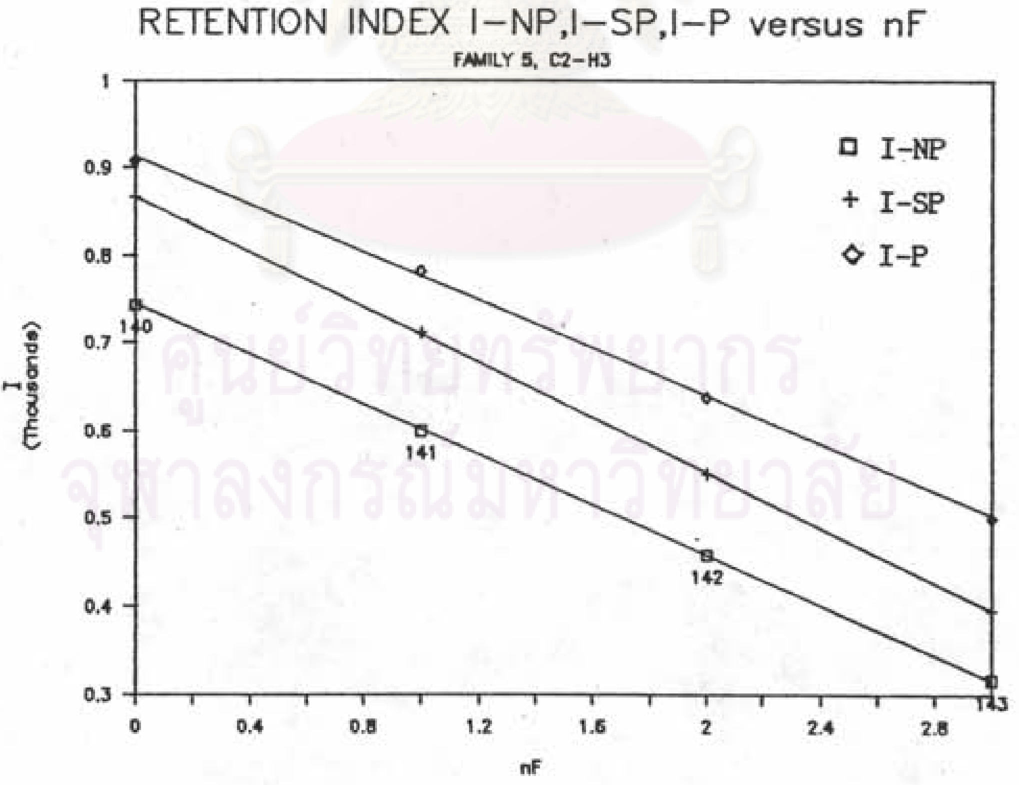


Fig. 5.12 Relationship between I and nF of family 5 on three columns

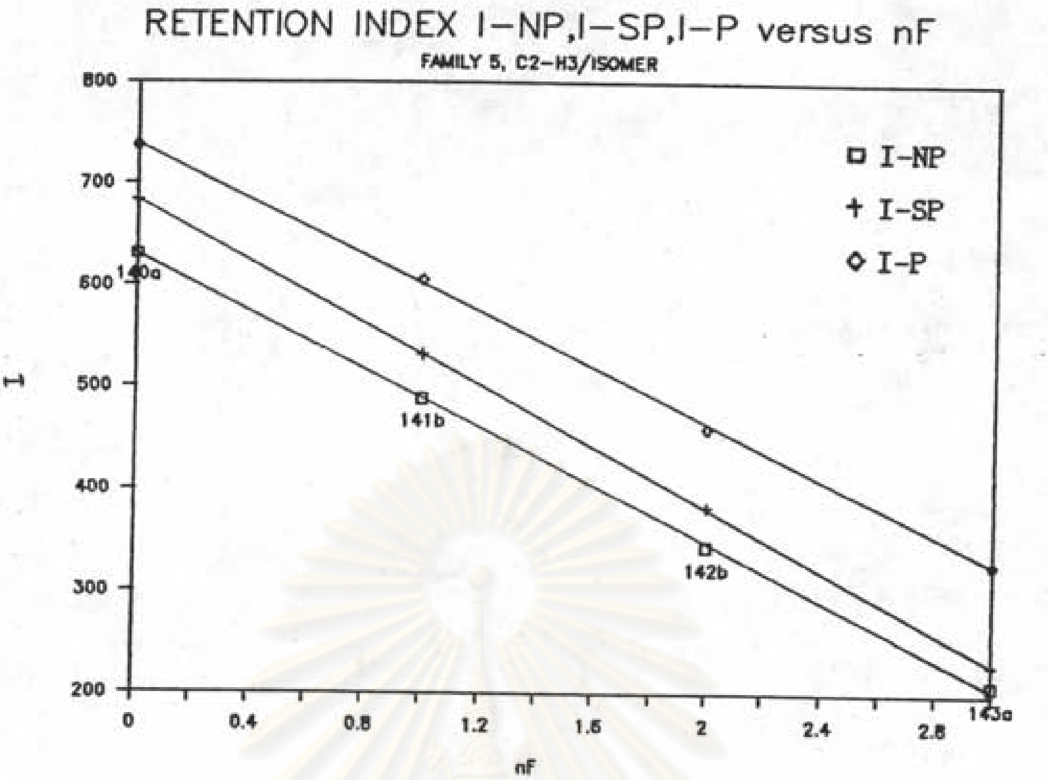


Fig. 5.13 Relationship between I and nF of family 5 (isomer) on three columns

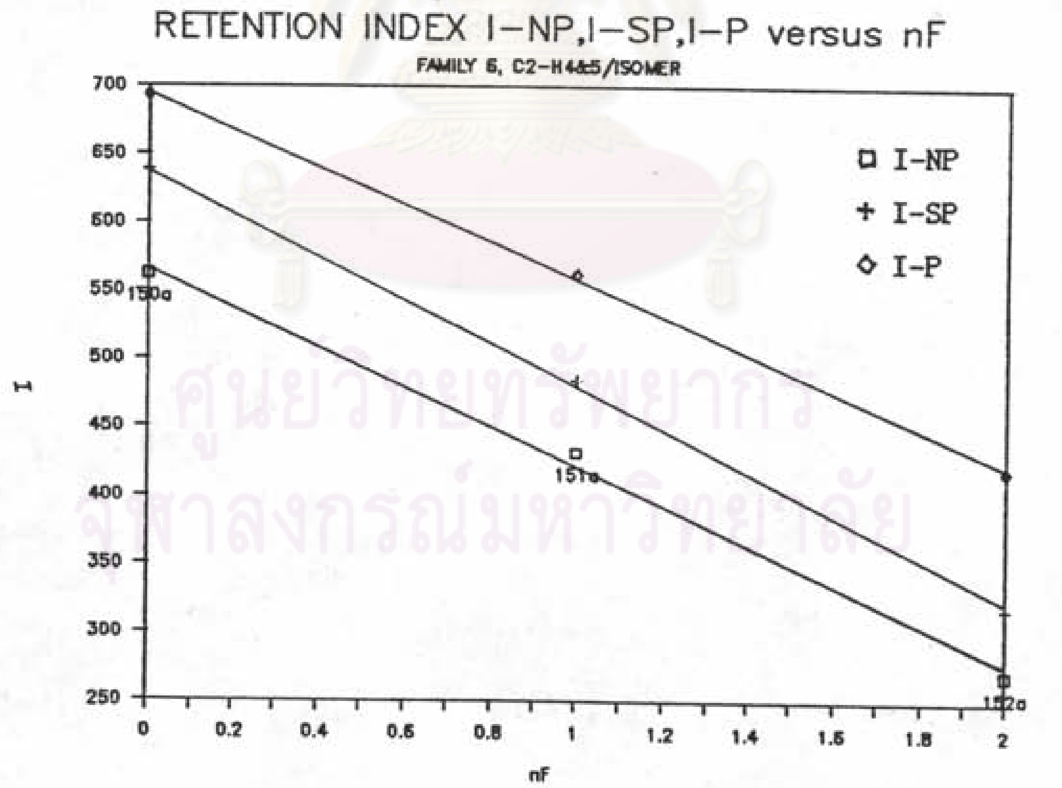


Fig. 5.14 Relationship between I and nF of family 6 (isomer) on three columns

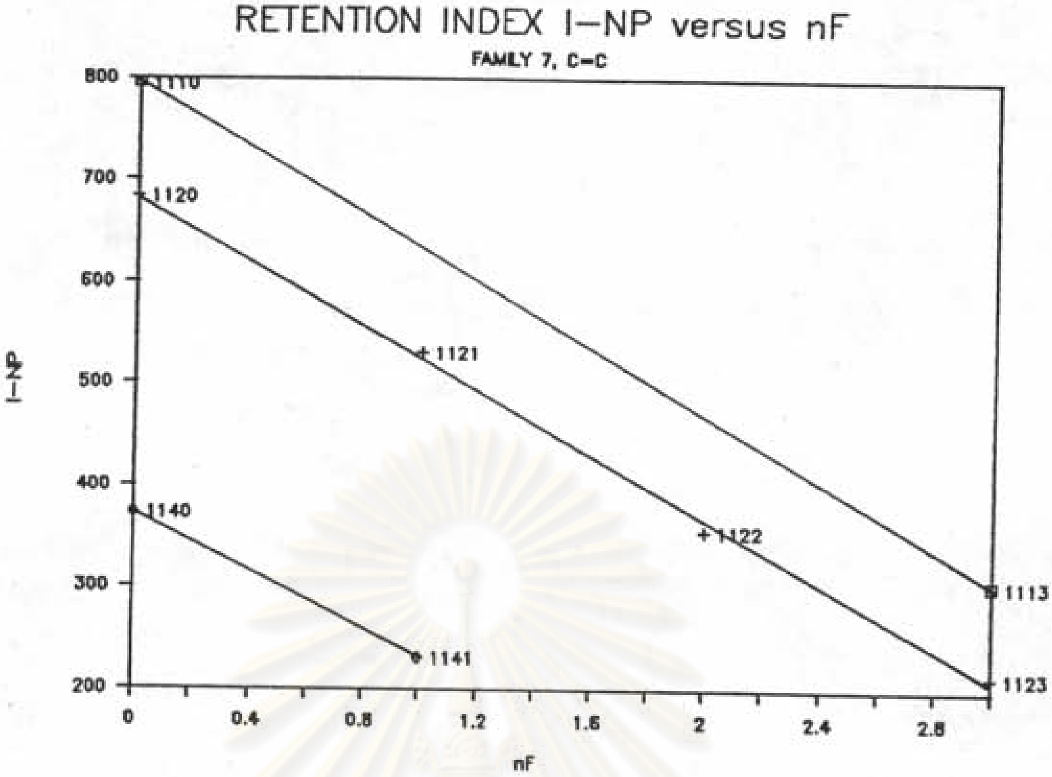


Fig. 5.15 relationship between I and nF of family 7 on apolar column

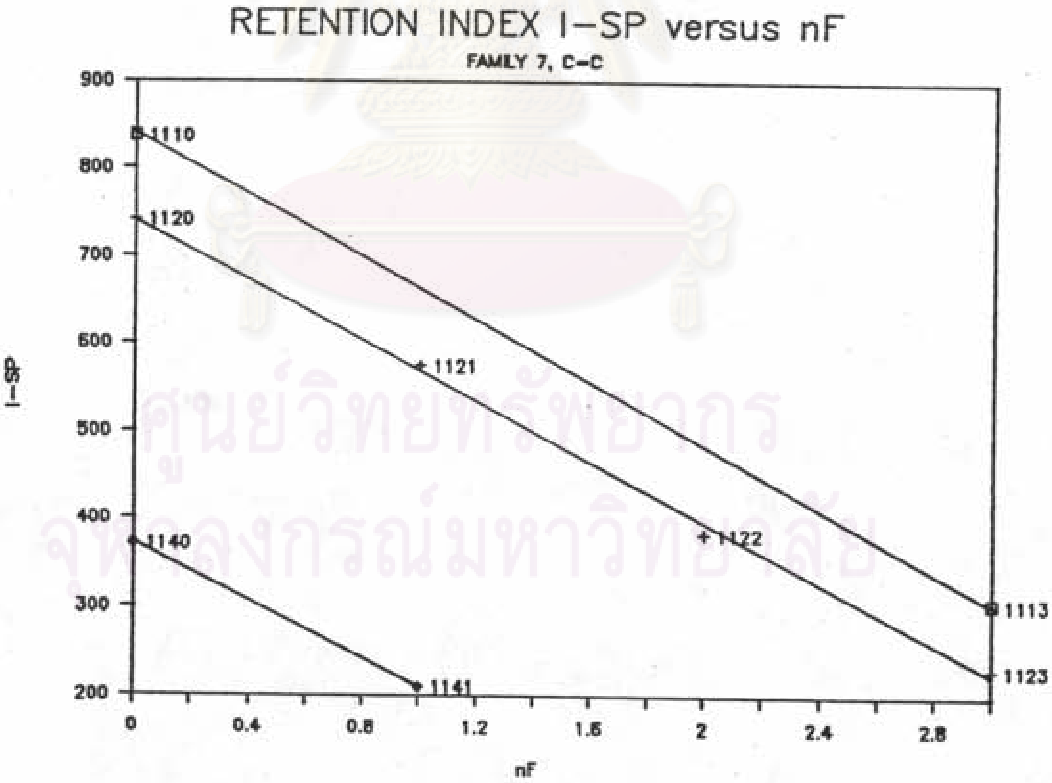


Fig. 5.16 Relationship between I and nF of family 7 on slightly polar column

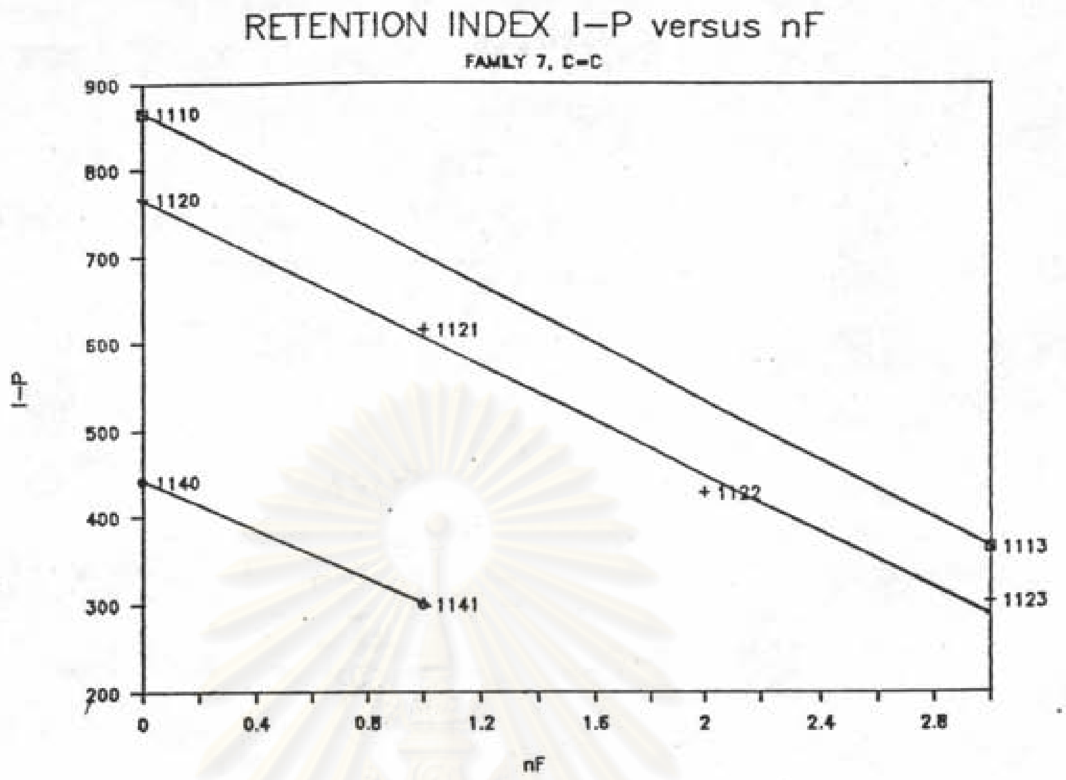


FIg. 5.17 Relationship between I and nF of family 7 on polar column

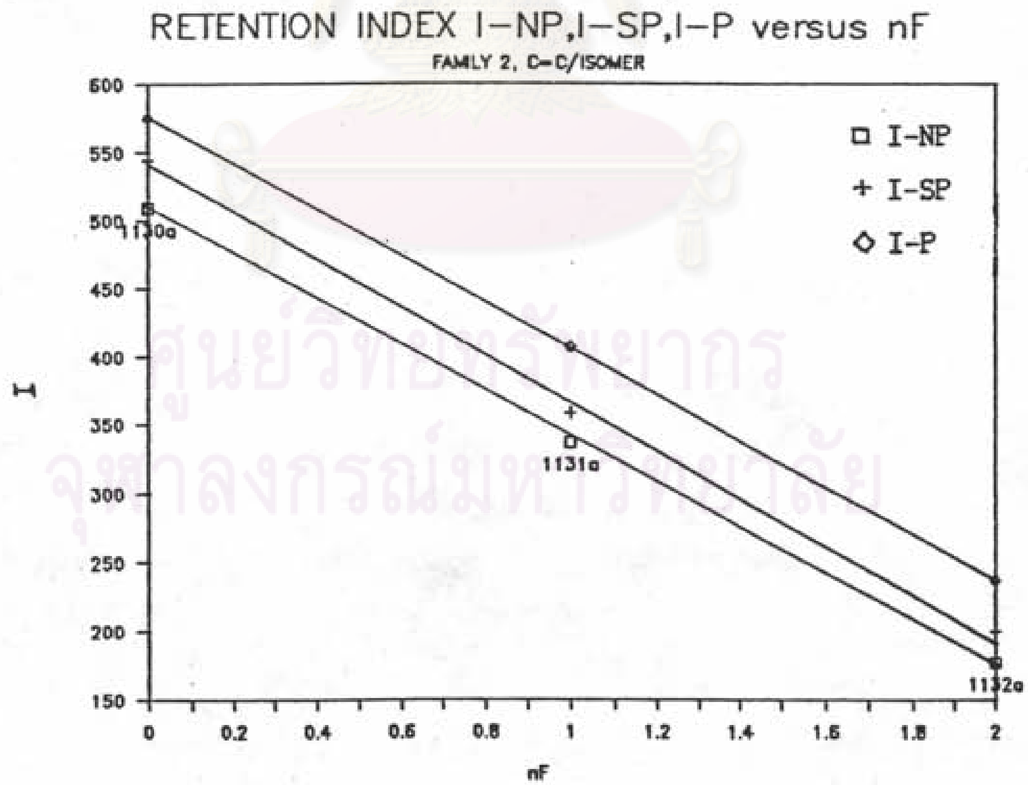


Fig. 5.18 Relationship between I and nF of family 7 (isomer) on three columns

Table 5.2 Summary of models from PCA analysis for apolar column

FAMILY	No.OF FACTORS	FACTORS	MODELS	RMS ERROR
1	4	BP,MW,nH,$^3\chi_c$	$I = 0.87\ BP + 2.89\ MW + 47.85\ nH + 40.92\ ^3\chi_c$	1.96
2	4	$^1\chi, ^4\chi_{pc}, ^4\chi_c$,nF	$I = 217.59\ ^1\chi + 10.98\ ^4\chi_{pc} + 164.80\ ^4\chi_c + 33.1\ nF$	7.91
3	2	$^3\chi_c$,nF	$I = 418.75\ ^3\chi_c + 97.87\ nF$	24.26
4	4	$^1\chi, ^4\chi_{pc}, ^4\chi_c$,nF	$I = 274.08\ ^1\chi + 29.22\ ^4\chi_{pc} - 100.44\ ^4\chi_c + 67.76\ nF$	27.04
5	3	$^3\chi_p, ^4\chi_c$,nF	$I = 1126.9\ ^3\chi_p + 1654.53\ ^4\chi_c + 232.71\ nF$	286.4
6	1	nH	$I = 93.14\ nH$	209.7
7	5	BP,$^4\chi_{pc}, ^4\chi_c$,nF,nH	$I = 5.35\ BP + 26.27\ ^4\chi_{pc} + 172.05\ ^4\chi_c + 159.05\ nF + 150.54\ nH$	23.89
2-6	5	BP,$^4\chi_{pc}, ^4\chi_c$,nF,nH	$I = 5.82\ BP - 4.54\ ^4\chi_{pc} + 24.54\ ^4\chi_c + 84.16\ nF + 73.46\ nH$	43.25
1-7	5	BP,$^4\chi_{pc}, ^4\chi_c$,nF,nH	$I = 3.69\ BP + 63.74\ ^4\chi_{pc} - 128.42\ ^4\chi_c + 152.57\ nF + 64.64\ nH$	44.10

Table 5.3 The predicted results of tested products for apolar column using PCA model

$$I = 3.69\,BP + 63.74\,{}^{4}\chi_{PC} - 128.42\,{}^{4}\chi_{C} + 152.57\,nF + 64.64\,nH$$

Products	Imeas	Ipred	ΔI (Imeas-Ipred)
11	479.0	240.4	238.6
30	520.7	341.9	178.8
112a	695.7	609.4	86.3
115	214.0	625.4	-411.5
122	626.5	631.8	-5.3
130	878.1	778.0	100.1
133a	401.1	613.8	-211.7
140a	630.6	412.5	218.1
142	459.0	629.6	-170.6
152a	270.5	472.6	-202.1
1112-cis	459.7	375.0	84.7
1123	213.9	296.1	-82.2
1131a	337.0	193.9	143.1
1140	372.2	142.7	229.5

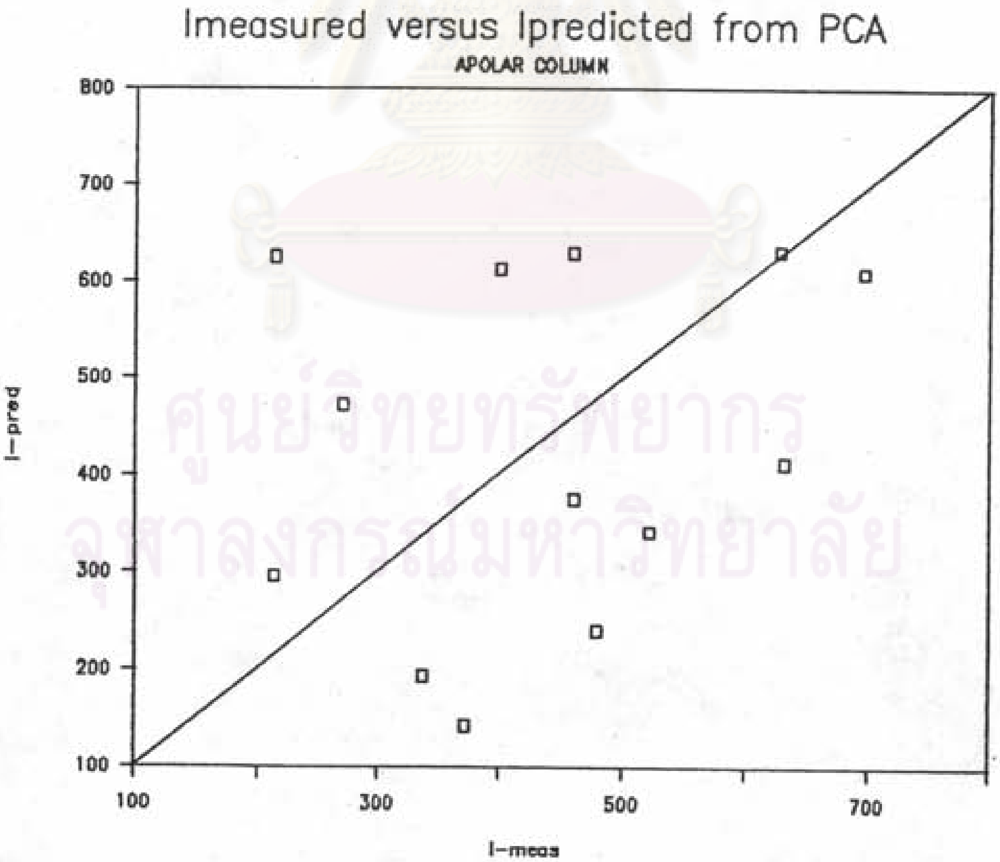
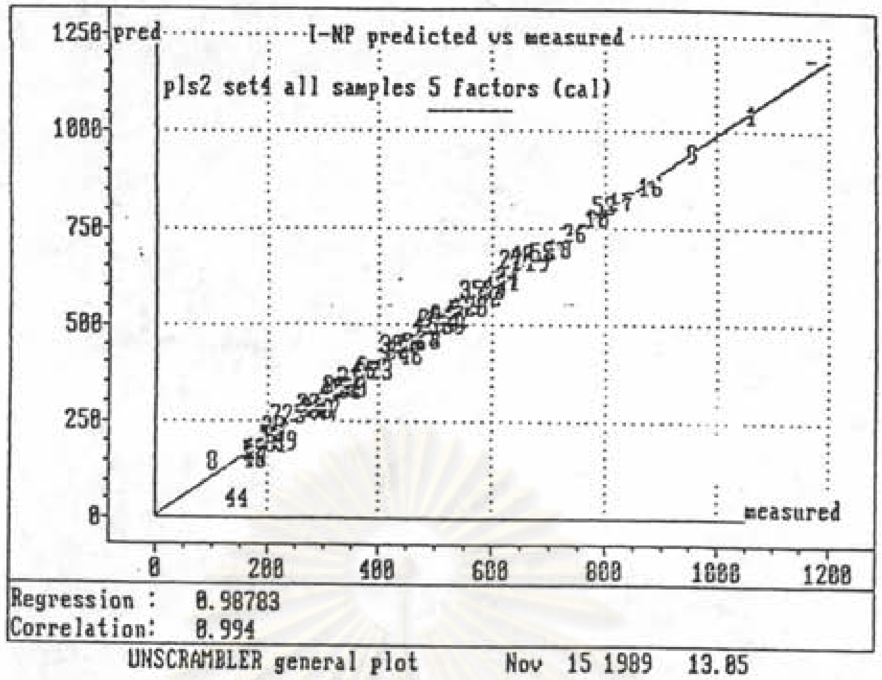


Fig. 5.19 The PCA results of tested products on apolar column

5.2 Partial Least-Squares Analysis (PLS)

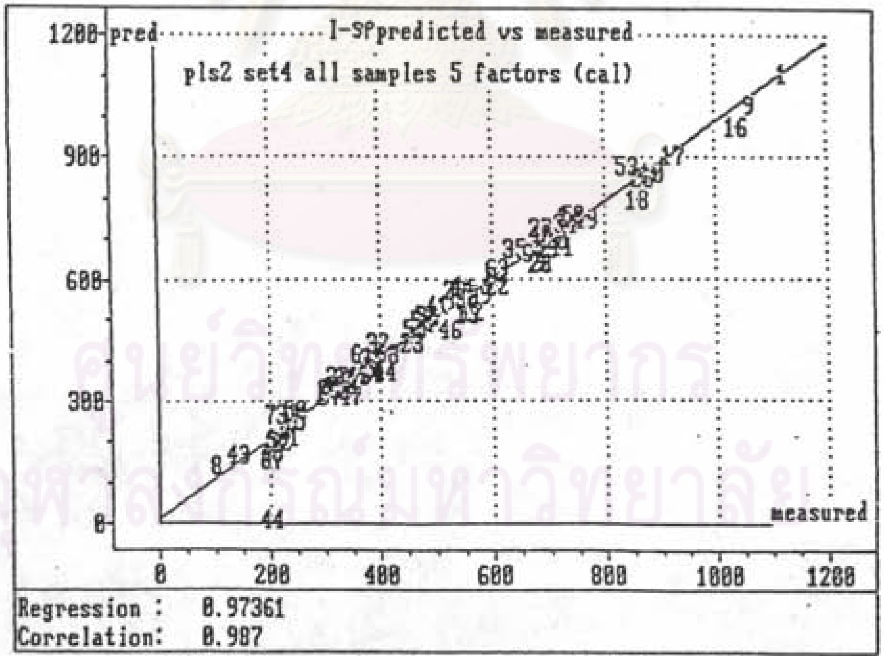
Better results are obtained from partial leas-tsquares analysis (PLS). The predictions are rather good. This can be seen from Figure 5.20 and 5.21 for the prediction of all compounds (both training and testing sets) on the apolar and slightly (medium) polar columns. They show that all the compounds fit with the 45° line except CF_4 (no.44). It means that the predicted and measured values match well. These good prediction are a result of the PLS principal which divides the data matrix into two matrices. One is the independent matrix, X and the other one is the dependent matrix, Y. Based on the same principal as PCA, PLS will search for the principal components for both the independent and dependent matrices to make the matrix X fit the matrix Y. That is the reason why PLS can give a good prediction. Furthermore, PLS can give the additional information of the nature of products and parameters by score and loading plots. Figure 5.22 and 5.23 show the score and loading plots from PLS of family 1 on the apolar column which are the same results as PCA. The loading plot with the exploded view of the parameters with the three retention indices (Figure 5.24 and 5.25) are also shown here.

Because PLS can give a rather good prediction of the retention index, it will be tried to use to predict the missing D values for the examined products. To do this, the D values will be put in the Y-matrix (as dependent variable) and all other variables (also the I values) are in the X-matrix (as independent variables). Figure 5.26 shows the loading plot for this condition. The result (Table 5.4 and Figure 5.27) shows that the predictions are rather good but not so perfect and it can be seen that 110 (point 13) and 1130-trans (point 60) are outliers. It means that the values of D cannot be explained well by these parameters.



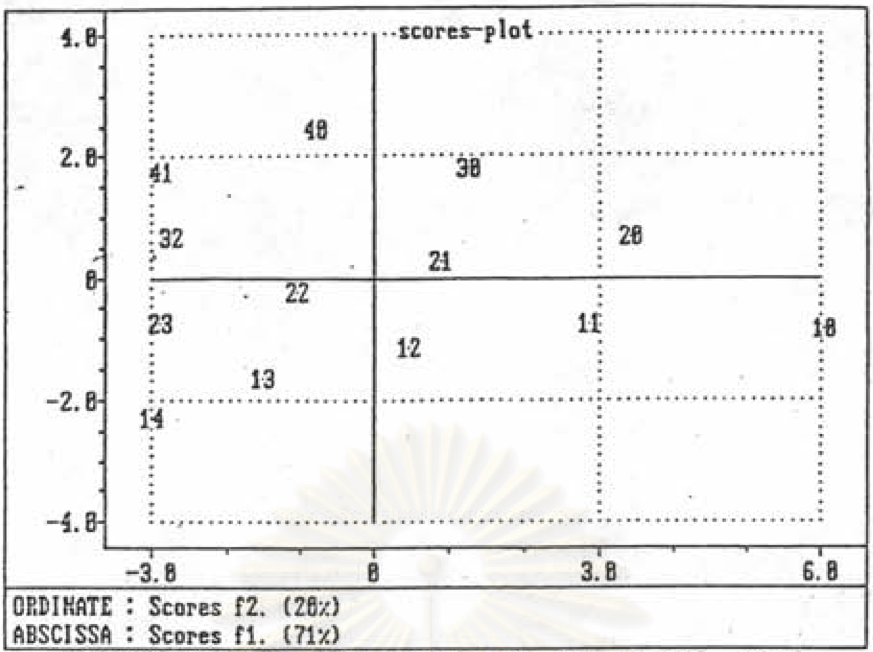
Licensee : DUPHAR B.V., ANALYTICAL DEP.

Fig. 5.20 The prediction results of all products on apolar column using PLS analysis



Licensee : DUPHAR B.V., ANALYTICAL DEP.

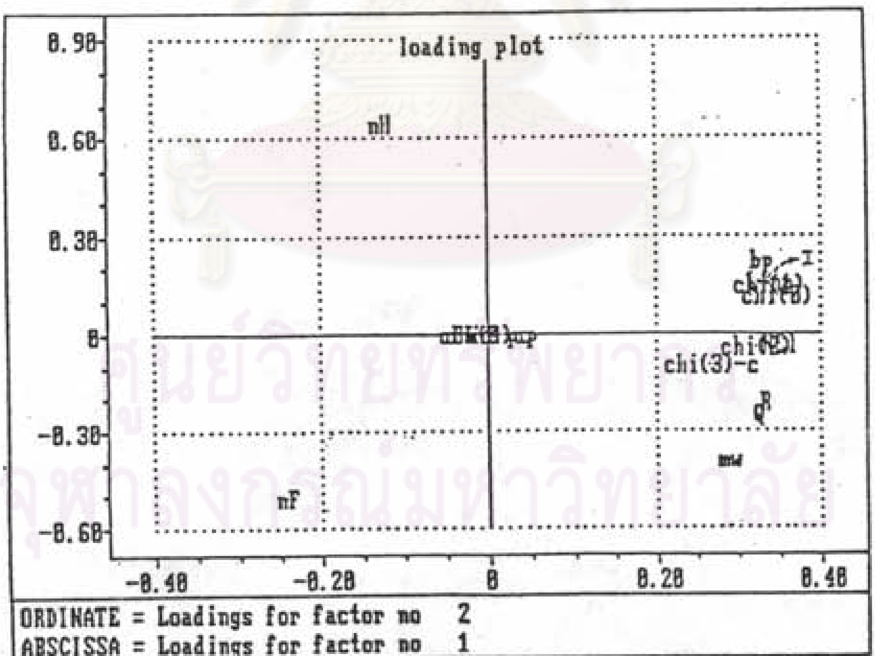
Fig. 5.21 The prediction results of all products on slightly polar column using PLS analysis



UNSCRAMBLER model : UNSC Version : 0 Nov 13 1989 14.46

Licensee : DUPHAR B.V., ANALYTICAL DEP.

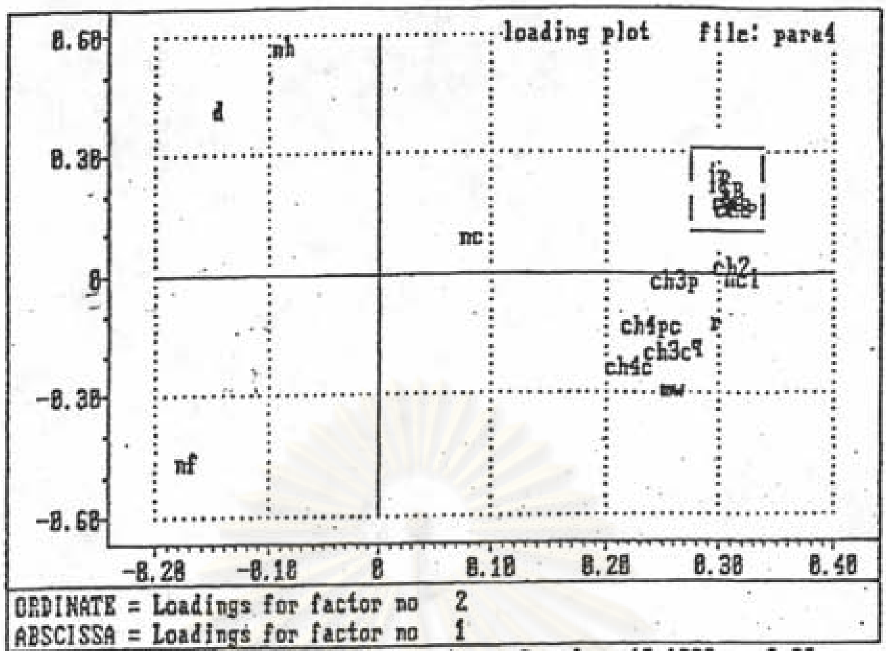
Fig. 5.22 The score plot of family 1 on apolar column using PLS analysis



UNSCRAMBLER model : UNSC Version : 0 Nov 13 1989 14.27

Licensee : DUPHAR B.V., ANALYTICAL DEP.

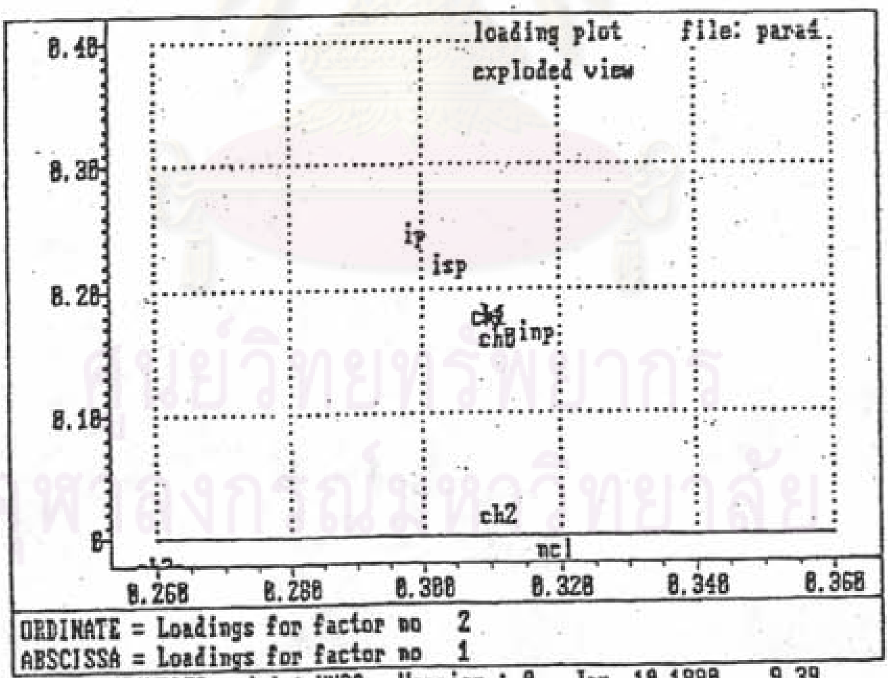
Fig. 5.23 The loading plot of family 1 on apolar column using PLS analysis



UNSCRAMBLER model : UNSC Version : 0 Jan 18 1998 9.38

Licensee : DUPHAR B.V., ANALYTICAL DEP.

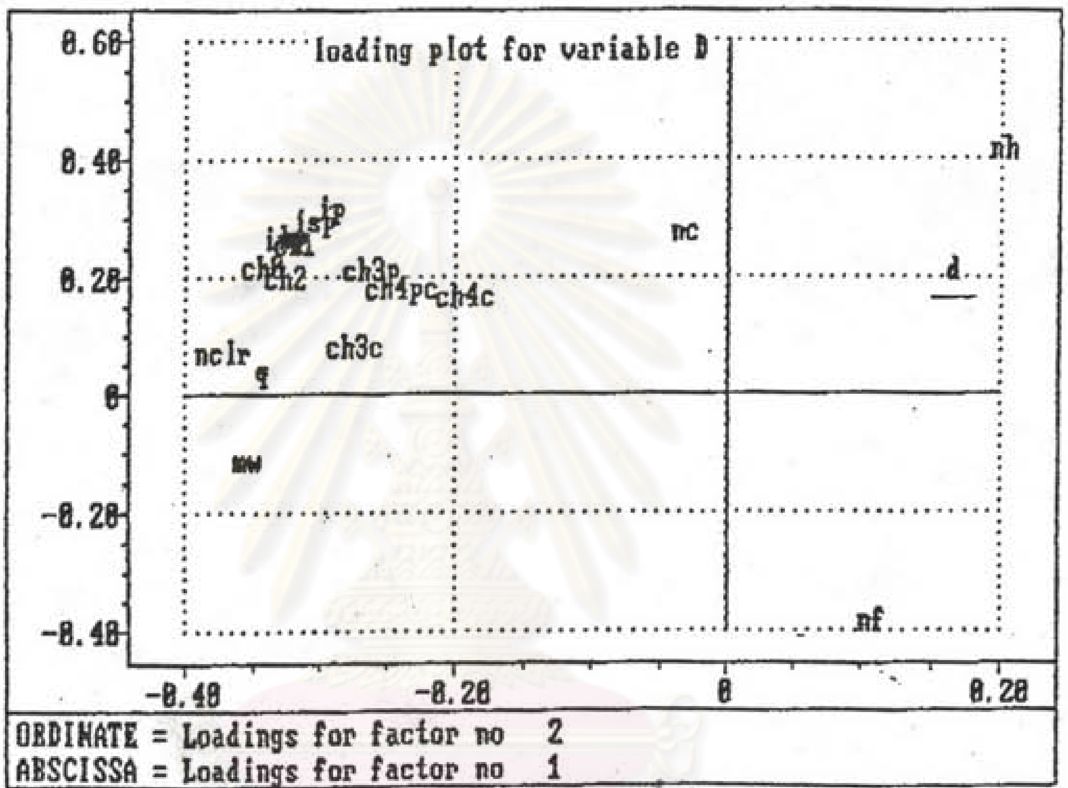
Fig. 5.24 The loading plot of retention indices and all parameters



UNSCRAMBLER model : UNSC Version : 0 Jan 18 1998 9.39

Licensee : DUPHAR B.V., ANALYTICAL DEP.

Fig. 5.25 Exploaded view of Fig. 5.24 showing
the three indices clustering



UNSCRAMBLER model : UNSC Version : 1 Jan 10 1990 16.15

Licensee : DUPHAR B.V., ANALYTICAL DEP.

Fig. 5.26 The loading plot for dependent variable D

Table 5.4 Predicted values of D by PLS analysis

Products	D-real	D-pred
10	0.000	0.111
11	0.500	0.431
12	0.500	0.526
13	0.500	0.650
20	1.100	1.081
21	1.300	1.241
22	1.400	1.422
23	1.600	1.664
30	1.800	1.662
32	2.000	2.046
40	1.900	1.634
41	1.800	1.925
110	0.000	0.973
112	m	0.186
112a	m	0.860
113	0.400	0.326
113a	m	1.020
114	0.500	0.600
115	0.800	0.853
116	m	0.973
120	1.000	1.016
121	m	0.546
122	m	0.985
123	m	1.366
123a	m	1.266
124	m	1.557
125	m	1.709
130	1.500	1.263
130a	m	1.465
131	m	1.529
131a	m	1.155
132	m	1.769
132b	m	1.541
133a	m	1.855
134	m	2.217
134a	m	2.019
140	1.400	1.643
140a	1.700	1.677
141	m	1.987
141b	m	1.369
142	2.100	2.340
142b	2.200	2.173
143	m	2.585
143a	2.300	1.831

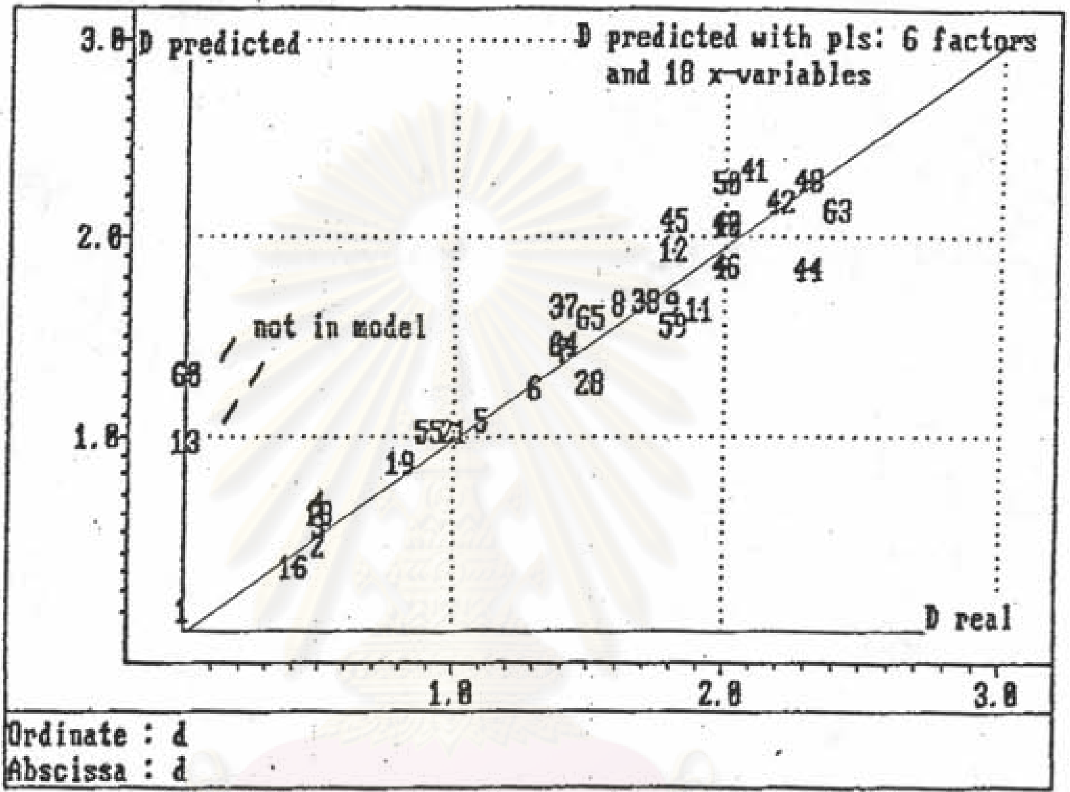
m missing data

Table 5.4 (continued)

Products	D-real	D-pred
150	1.800	2.077
150a	2.000	1.854
151a	m	2.097
152a	2.300	2.279
160	2.000	1.854
161	2.000	2.097
1110	m	0.580
1112-cis	m	0.705
1112a	m	1.243
1113	m	1.127
1120	0.900	1.014
1121	m	1.101
1122	m	1.472
1123	m	1.610
1130-cis	1.800	1.547
1130-trans	0.000	1.302
1130a	m	1.508
1131a	m	1.322
1132	2.400	2.134
1132a	1.400	1.448
1140	1.500	1.586

m missing data

Licensee : DUPHAR B.V., ANALYTICAL DEP.



UNSCRAMBLER general plot Jan 10 1990 16.29

Licensee : DUPHAR B.V., ANALYTICAL DEP.

Fig. 5.27 The prediction results of dipole moment (D) using PLS analysis

5.3 Multiparameter Regression Analysis

Multiparameter stepwise regression analysis is the last approach to analyse the data. It can provide several models according to the purposes and data available. For this work, the main purpose is to predict retention indices of known products. With this purpose in mind, two groups of models are developed for different groups of data. The first group of models use only parameters that can be calculated from the structure of the molecules (Table 5.5 to 5.7) and the second group are the models based on the parameters that can be calculated from the structure of the molecules but also including physico-chemical parameters, like boiling point and/or dipole moment (Table 5.8 to 5.10). The results show that, for every column, the models which include BP as one of the parameters give the higher percentage of variance accounted for the data and a lower standard error (considering from %variance and S.E.) than the models without BP. BP is the parameter that most accounted for the variance in the data. It means that the elution order for these compounds on the three columns considered in gas chromatography is primarily based on their boiling points.

To test the validity of the correlations, these models are used to predict the retention indices of the tested products. The results (Figure 5.28 to 5.33) confirm that the models with BP as one of the parameters can give better predictions. For the apolar column, the error between measured and predicted values is in the range of 1 to 44 index units (i.u.) with an average error of 14.1 i.u. (Table 5.11). For the slightly polar column, the error is in the range of 1 to 60 i.u. with an average error of 24.5 i.u.(Table 5.12). For the polar column, the error is in the range of 1 to 37 i.u. with an average error of 20.4 i.u. (Table 5.13). The average of the three columns is 19.7 i.u. These values of error are less than 10% of the average predicted values. So these models with BP are acceptable. For the models without BP as one of the parameters are also satisfactory with an average error of 45.9 i.u. but they will give poorer results.

From the results of stepwise regression analysis, one can also learn or understand the behavior of the system. Comparing the models of the polar columns to the models of the apolar column using the same groups of parameters (Table 5.5 to 5.10), it is found that the percentage of variance of the polar models are less than that of the apolar models. This means that there are some effects on the polar columns which are not explained by these parameters. For all the parameters used to investigate, the dipole moment is the parameter that can give the information about the effect of polarity. The results (Table 5.14 and 5.15) show that when the polarity of the column increases, this specific parameter, dipole moment (D), becomes more important. The stepwise regression method chooses D as the sixth parameter for the slightly polar column with quite insignificance (considering from the calculated t-value of D which is less than the theoretical t-value), while D is the second parameter for the polar model with high significance (the calculated t-value is higher than the theoretical t-value). This leads to the idea that the permanent dipole of the polar stationary phase can induce a dipole momemt in a neighbouring molecule effecting the dipole-interaction.

To combine the models from the three columns becoming one model, the equations based on the same parameters of the three stationary phases will be related with the polarity on the McReynold scale (P_{MC}). The common parameters chosen are BP and $^1\chi$ which gives a high percentage of variance. These equations are shown in Table 5.16.

Equations that relate coefficients in Table 5.16 and the polarity on the McReynold scale are shown in Table 5.17

From Equation (3.1) and equations in Table 5.17, the following final equation is obtained :-

$$I = 310.42+2.33\,BP+65.55\,{}^{1}\chi + [0.12+0.0008\,BP-0.044\,{}^{1}\chi]P_{MC} \quad (5.1)$$

Equation (5.1) relates the retention index to the solute properties and the polarity of the stationary phase. It should be valid to calculate the retention index of CFCs on any stationary phase. This equation is used to predict some non-available products on the same stationary phases. The results are shown in Table 5.18.

The predicted values of those products are plotted on figures showing the relationship between retention index and nF to see whether the index values lie on the line. The results (Figure 5.34 to 5.37) show that the predicted indices do not deviate from the line so much and the error is in the acceptable range. The prediction on the apolar column is more accurate than the polar columns.

The other problem which can be encountered in gas chromatography results, is the recognition of the peaks of unknown products. For this, the regression analysis can give a model that can be used to recognize these unknown peaks. Peak recognition can be based on the relationship between the retention index and ${}^{1}\chi$ (Table 5.19). From this relationship, ${}^{1}\chi$ value can be determined. All products falling out of ± 10% of the calculated ${}^{1}\chi$ value can be neglected. The remain compounds can be divided in the groups of products available in the laboratory and not available. After testing the available products, some compounds can be neglected. The ${}^{1}\chi$ value can give structural information and perhaps an idea about the family to which the product belongs. An example of this case can be seen from the result of the apolar column. The unknown peak (1112-trans) was found during experiments on family 7. This unknown peak eluted after the peak of 1112-cis product. From the relationship between the retention index and $^{1}\chi$, the $^{1}\chi$ value was 1.325. Within the range of ± 10% of the calculated value, the $^{1}\chi$ value was in the range 1.2 to 1.5. There were 14 CFC compounds with $^{1}\chi$ values in this range. They were 21, 40, 123, 123a, 132, 132a, 132c, 1112-cis, 1112-trans, 1112a, 1131-cis, 1131-trans, 1131a and 1140. From this list, only 3 products were not available. The other 11 products were examined and only 3 products were suspected because they had nearly the same indices. They were 21, 132a, 132c, 1112-cis,1112-trans and 1112a. The final result was confirmed by the GC-MS. It showed that the unknown peak was 1112. When considering the three isomers of 1112, the products 1112a, 1112-cis and 1112-trans eluted in respective order according to their increasing boiling points. Therefore it was quite certain that the unknown peak was 1112-trans.

This procedure is only possible if the product is a CFC, HCFC or HFC of the families under investigation. Of course a GC-MS will comfirm and give better results if it is available.

Table 5.5 Regression Models of Apolar Column for all Parameters except boiling point (BP) and dipole moment (D)

FAMILY	PARAMETERS	MODELS	%VARIANCE	S.E.
1	$^0\chi$		91.6	51.10
	$^0\chi, ^2\chi$		91.8	50.46
	$^0\chi, ^2\chi, ^3\chi_c$		93.6	44.64
	$^0\chi, ^2\chi, ^3\chi_c, ^1\chi$	$I = 226.7-20\ ^0\chi+98.9\ ^2\chi-65\ ^3\chi_c+137.6\ ^1\chi$	94.5	41.38
2	R		99.5	21.86
	R, $^3\chi_p$	$I = -927.3+365.25\ R+20.77\ ^3\chi_p$	99.9	11.43
3	nF		99.8	12.04
	nF, $^3\chi_p$	$I = 907.1-140.03\ nF+17.85\ ^3\chi_p$	99.9	8.03
4	$^0\chi$		99.2	19.86
	$^0\chi$, Q	$I = 681+140.6\ ^0\chi-202.1\ Q$	99.5	15.25
5	nF		89.1	65.42
	nF, $^0\chi$	$I = 3271.4-878.4\ nF-519.5\ ^0\chi$	100.0	2.95
6	$^0\chi$		90.3	41.94
	$^0\chi, ^3\chi_p$		96.5	25.08
	$^0\chi, ^3\chi_p$, nF	$I = 20.4+137.2\ ^0\chi+228.3\ ^3\chi_p+57.8\ nF$	99.7	7.15
7	$^0\chi$		96.0	36.25
	$^0\chi, ^3\chi_c$	$I = 120+126.2\ ^0\chi-32.4\ ^3\chi_c$	96.1	35.62
2-6	$^1\chi$		94.4	54.27
	$^1\chi, ^0\chi$		95.6	48.00
	$^1\chi, ^0\chi, ^3\chi_p$		96.4	43.51
	$^1\chi, ^0\chi, ^3\chi_p$, MW		96.4	43.37
	$^1\chi, ^0\chi, ^3\chi_p$, MW, Q		96.8	41.22
	$^1\chi, ^0\chi, ^3\chi_p$, MW, Q, R	$I = 287.7+55.1\ ^1\chi+9.35\ ^0\chi+23\ ^3\chi_p+0.07\ MW$ $-804\ Q+740\ R$	98.8	25.24
1-7	$^1\chi$		93.0	57.26
	$^1\chi, ^0\chi$		94.5	50.40
	$^1\chi, ^0\chi$, R		95.4	46.45
	$^1\chi, ^0\chi$, R, Q		96.1	42.60
	$^1\chi, ^0\chi$, R, Q, $^3\chi_p$	$I = 159.9+83\ ^1\chi+28.3\ ^0\chi+270.1\ R-262.1\ Q$ $+28.4\ ^3\chi_p$	96.5	40.14

Table 5.6 Regression Models of Slightly Polar Column for all Parameters except boiling point (BP) and dipole moment (D)

FAMILY	PARAMETERS	MODELS	%VARIANCE	S.E.
1	${}^{0}\chi$		83.8	79.84
	${}^{0}\chi$,nF		84.7	77.72
	${}^{0}\chi$,nF,nCl	$I = 1695-1507\,{}^{0}\chi-549.6\,nF+1341\,nCl$	98.4	25.27
2	R		99.0	33.81
	R,${}^{3}\chi_p$		99.8	14.41
	R,${}^{3}\chi_p$,Q		99.9	12.07
	R,${}^{3}\chi_p$,Q,${}^{1}\chi$	$I = 195.6+757.5\,R+33.821\,{}^{3}\chi_p-661.9\,Q+26.52\,{}^{1}\chi$	100.0	0.75
3	nF		99.9	8.52
	nF,R		99.9	7.15
	nF,R,${}^{3}\chi_p$	$I = 57-72.3\,nF+210\,R+8\,{}^{3}\chi_p$	100.0	6.26
4	${}^{0}\chi$		97.1	41.18
	${}^{0}\chi$,Q	$I = 1270+197.4\,{}^{0}\chi-441\,Q$	98.5	29.68
5	nF		81.4	101.16
	nF,${}^{0}\chi$	$I = 4778.1-1295.4\,nF-803.6\,{}^{0}\chi$	100.0	3.95
6	${}^{0}\chi$		88.5	53.79
	${}^{0}\chi$,${}^{3}\chi_p$		97.6	24.40
	${}^{0}\chi$,${}^{3}\chi_p$,nF	$I = 35.9+152.53\,{}^{0}\chi+310.7\,{}^{3}\chi_p+56.5\,nF$	99.9	6.1
7	${}^{1}\chi$	$I = 120+246.6\,{}^{1}\chi$	93.7	52.06
2-6	${}^{1}\chi$		90.1	32.05
	${}^{1}\chi$,${}^{0}\chi$		91.3	74.80
	${}^{1}\chi$,${}^{0}\chi$,${}^{3}\chi_p$		91.5	73.68
	${}^{1}\chi$,${}^{0}\chi$,${}^{3}\chi_p$,Q		91.8	72.68
	${}^{1}\chi$,${}^{0}\chi$,${}^{3}\chi_p$,Q,R		97.2	42.20
	${}^{1}\chi$,${}^{0}\chi$,${}^{3}\chi_p$,Q,R,nH	$I = 1919+182.8\,{}^{1}\chi+50.4\,{}^{0}\chi+34.2\,{}^{3}\chi_p-1440\,Q$ $+809\,R-137.5\,nH$	97.7	38.46
1-7	${}^{1}\chi$		88.3	83.15
	${}^{1}\chi$,${}^{0}\chi$		89.9	77.16
	${}^{1}\chi$,${}^{0}\chi$,${}^{3}\chi_p$		90.3	75.79
	${}^{1}\chi$,${}^{0}\chi$,${}^{3}\chi_p$,nH		90.7	73.86
	${}^{1}\chi$,${}^{0}\chi$,${}^{3}\chi_p$,nH,MW	$I = 48.4+87.5\,{}^{1}\chi+46.6\,{}^{0}\chi+35\,{}^{3}\chi_p+37.7\,nH$ $+1.24\,MW$	92.3	67.48

Table 5.7 Regression Models of Polar Column for all Parameters except boiling point (BP) and dipole moment (D)

FAMILY	PARAMETERS	MODELS	%VARIANCE	S.E.
1	${}^0\chi$		82.6	72.31
	${}^0\chi$,nF		86.6	63.58
	${}^0\chi$,nF,nCl	$I = 1557-1019\ {}^0\chi-478\ nF+1125.3\ nCl$	99.3	14.10
2	nF		99.8	14.56
	nF,${}^3\chi_p$	$I = 1084.64-164.86\ nF+14.26\ {}^3\chi_p$	100.0	6.03
3	nF		99.9	10.56
	nF,${}^3\chi_p$	$I = 1029.7-144.91\ nF+17.07\ {}^3\chi_p$	100.0	5.63
4	${}^0\chi$		97.6	36.11
	${}^0\chi$,Q	$I = 1310+190.5\ {}^0\chi-429\ Q$	99.1	21.40
5	nF		77.5	102.01
	nF,${}^0\chi$	$I = 4857-1286.2\ nF-809.9\ {}^0\chi$	99.9	5.49
6	${}^0\chi$		83.8	59.53
	${}^0\chi$,${}^3\chi_p$		96.0	29.59
	${}^0\chi$,${}^3\chi_p$,nF	$I = 125.12+143.39\ {}^0\chi+338.65\ {}^3\chi_p+69.62\ nF$	100.0	1.75
7	${}^1\chi$	$I = 194.7+227.8\ {}^1\chi$	91.9	54.85
2-6	${}^1\chi$		89.0	79.81
	${}^1\chi$,${}^0\chi$		90.1	75.92
	${}^1\chi$,${}^0\chi$,Q		90.7	73.35
	${}^1\chi$,${}^0\chi$,Q,R		96.4	45.85
	${}^1\chi$,${}^0\chi$,Q,R,${}^2\chi$		96.7	43.58
	${}^1\chi$,${}^0\chi$,Q,R,${}^2\chi$,${}^3\chi_p$	$I = 557.6+39.1\ {}^1\chi+17.6\ {}^0\chi-1390\ Q+1225\ R$ $-38.2\ {}^2\chi+33.9\ {}^3\chi_p$	97.1	41.36
1-7	${}^1\chi$		86.1	86.34
	${}^1\chi$,${}^0\chi$		87.6	81.44
	${}^1\chi$,${}^0\chi$,nH		89.0	76.92
	${}^1\chi$,${}^0\chi$,nH,MW		91.3	68.40
	${}^1\chi$,${}^0\chi$,nH,MW,${}^3\chi_p$	$I = 82.1+78.9\ {}^1\chi+43.1\ {}^0\chi+49.54\ nH+1.384\ MW$ $+35.4\ {}^3\chi_p$	91.8	66.30

Table 5.8 Regression Models of Apolar Column for all Parameters except dipole moment (D)

FAMILY	PARAMETERS	MODELS	%VARIANCE	S.E.
1	BP		98.6	21.00
	BP, $^{1}\chi$	$I = 879.8+2.616\ BP+28.5\ ^{1}\chi$	98.7	19.91
2	BP		99.8	15.87
	BP, $^{2}\chi$		99.9	8.01
	BP, $^{2}\chi$, R	$I = 999.2+4.90\ BP+19.5163\ ^{2}\chi-182.345\ R$	100.0	0.00
3	BP		100.0	6.01
	BP, R	$I = 112.0+2.641\ BP+87.2\ R$	100.0	5.10
4	BP		99.8	8.63
	BP, MW		99.9	6.92
	BP, MW, R	$I = 69.2+2.93\ BP+7.26\ MW-199.4\ R$	100.0	4.29
5	BP		97.2	33.09
	BP, $^{1}\chi$	$I = 305.29+2.2622\ BP+68.48\ ^{1}\chi$	100.0	1.87
6	BP		99.8	5.87
	BP, R	$I = 286.9+2.533\ BP+42.9\ R$	99.9	3.44
7	BP		98.7	20.46
	BP, $^{1}\chi$	$I = 314+2.087\ BP+73.7\ ^{1}\chi$	99.5	13.31
2-6	BP		99.0	23.33
	BP, nCl		99.3	18.47
	BP, nCl, Q		99.5	15.57
	BP, nCl, Q, $^{4}\chi_{PC}$	$I = 478.1+2.321\ BP+68.1\ nCl-69.8\ Q+6.77\ ^{4}\chi_{PC}$	99.7	13.48
1-7	BP		97.9	31.42
	BP, $^{1}\chi$		98.9	22.52
	BP, $^{1}\chi$, $^{4}\chi_{PC}$		99.3	18.54
	BP, $^{1}\chi$, $^{4}\chi_{PC}$, nC	$I = 374.5+2.329\ BP+54.67\ ^{1}\chi+11.64\ ^{4}\chi_{PC}-21.53nC$	99.4	16.88

Table 5.9 Regression Models of Slightly Polar Column for all Parameters except dipole moment (D)

FAMILY	PARAMETERS	MODELS	%VARIANCE	S.E.
1	BP		96.8	35.43
	BP,Q	$I = 595.5+3.934\ BP-65.7\ Q$	97.4	32.09
2	BP		99.3	29.87
	BP,MW		99.9	12.63
	BP,MW,$^{3}\chi_{P}$	$I = 7774+20.77\ BP-43.82\ MW-37.07\ ^{3}\chi_{P}$	100.0	2.51
3	BP		99.8	12.84
	BP,$^{1}\chi$	$I = 376.9+2.163\ BP+95.2\ ^{1}\chi$	99.9	7.03
4	BP		99.4	19.36
	BP,$^{3}\chi_{P}$	$I = 428.73+4.63\ BP-101.9\ ^{3}\chi_{P}$	99.8	10.59
5	BP		99.4	18.22
	BP,$^{1}\chi$		99.8	10.18
	BP,$^{1}\chi$,nF	$I = 188+3.474\ BP+108\ ^{1}\chi+63.7\ nF$	100.0	3.85
6	BP	$I = 437.97+3.4383\ BP$	99.8	6.45
7	BP		98.5	25.72
	BP,nF	$I = 465.2+3.06\ BP-20.45\ nF$	99.0	21.16
2-6	BP		98.4	31.59
	BP,Q		98.7	29.19
	BP,Q,R	$I = 497.4+2.869\ BP-539\ Q+463\ R$	99.0	25.18
1-7	BP		97.7	36.76
	BP,nF		98.2	32.23
	BP,nF,nC		98.3	31.70
	BP,nF,nC,$^{4}\chi_{PC}$		98.3	31.57
	BP,nF,nC,$^{4}\chi_{PC}$,$^{3}\chi_{C}$	$I = 510+3.506\ BP-12.57\ nF-27\ nC+10.56\ ^{4}\chi_{PC}-16.02\ ^{3}\chi_{C}$	98.4	30.40

Table 5.10 Regression Models of Polar Column for all Parameters except dipole moment (D)

FAMILY	PARAMETERS	MODELS	%VARIANCE	S.E.
1	BP		96.4	33.08
	BP,MW	I = 584.1+3.451 BP-0.98 MW	98.8	19.24
2	BP	I = 406.79+3.921 BP	100.0	3.72
3	BP		99.9	6.64
	BP,nF	I = 684+2.374 BP-52.1 nF	100.0	5.12
4	BP		99.3	19.22
	BP,$^3\chi_c$	I = 478.99+4.038 BP-23.56 $^3\chi_c$	99.7	13.52
5	BP	I = 493.27+3.6646 BP	99.8	9.48
6	BP		99.3	12.41
	BP,$^1\chi$		99.3	12.32
	BP,$^1\chi$,nF	I = 947.6+6.4 BP-316 $^1\chi$-130 nF	99.9	3.35
7	BP		98.4	24.72
	BP,Q	I = 608.5+3.436 BP-52.1 Q	98.9	20.54
2-6	BP		97.9	35.27
	BP,Q		99.0	23.53
	BP,Q,R		99.1	22.37
	BP,Q,R,$^2\chi$	I = 626.9+3.368 BP-440 Q+350 R-20.26 $^2\chi$	99.2	21.84
1-7	BP		97.4	37.26
	BP,Q		98.3	30.37
	BP,Q,nH		98.5	28.74
	BP,Q,nH,R		98.6	27.17
	BP,Q,nH,R,$^4\chi_{PC}$		98.7	26.66
	BP,Q,nH,R,$^4\chi_{PC}$,$^2\chi$	I = 565.4+3.595 BP-179 Q+8.81 nH+130.3 R +10.43 $^4\chi_{PC}$-12.82 $^2\chi$	98.8	25.68

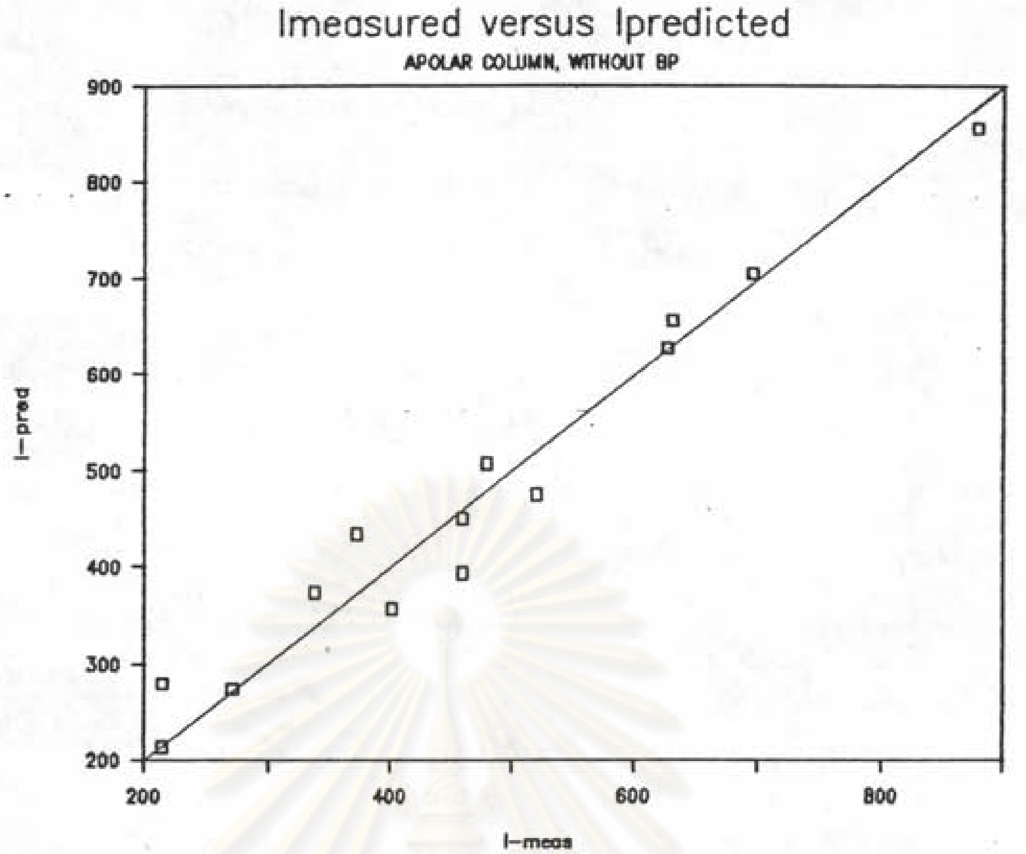


Fig. 5.28 The results of tested products on apolar column

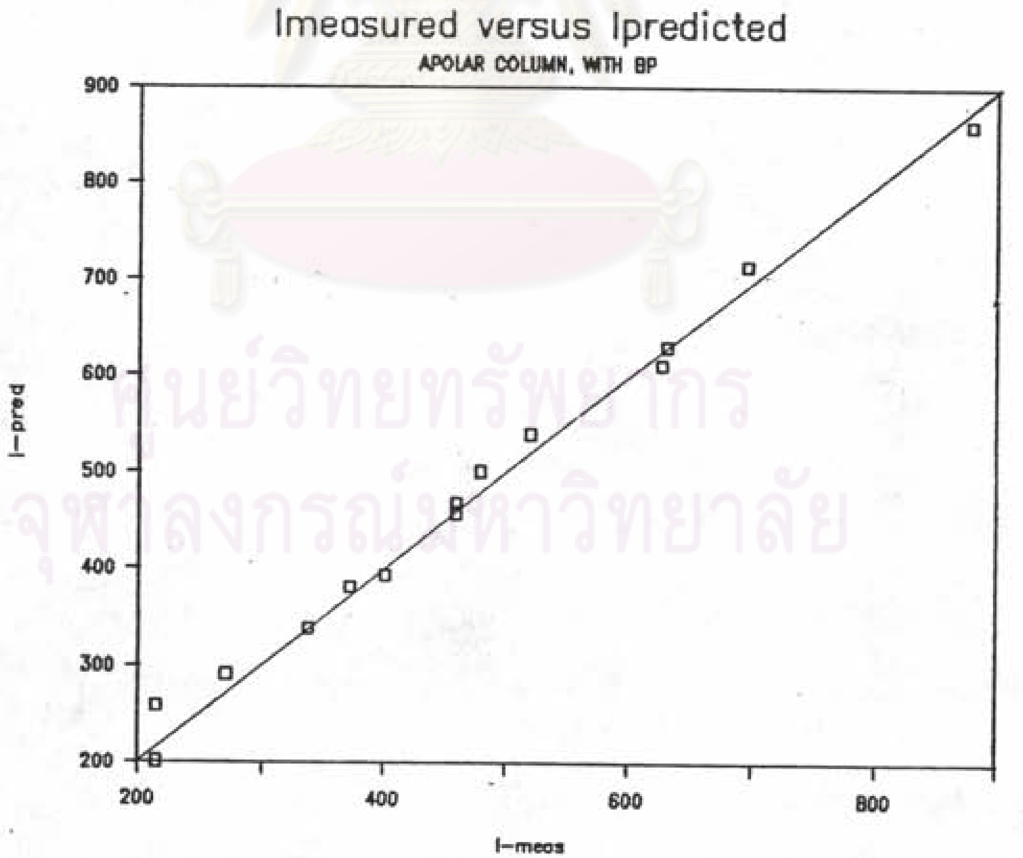


Fig. 5.29 The results of tested products on apolar column with BP

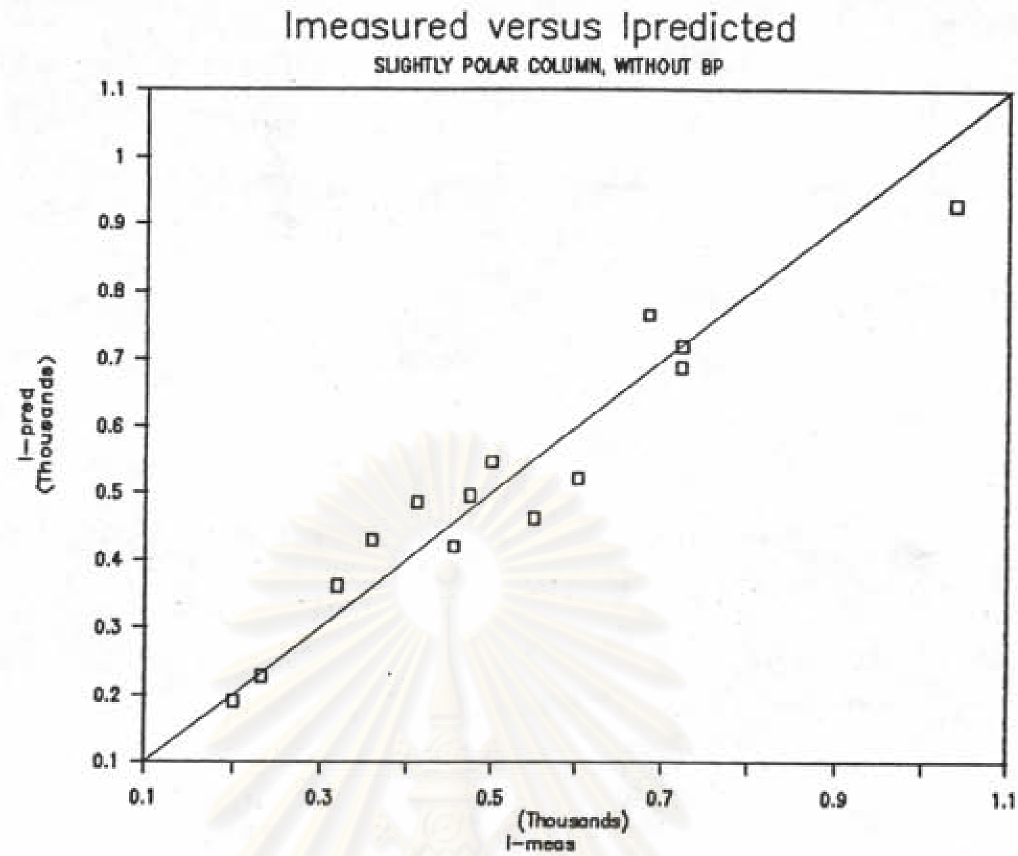


Fig.5.30 The results of tested products on slightly polar column

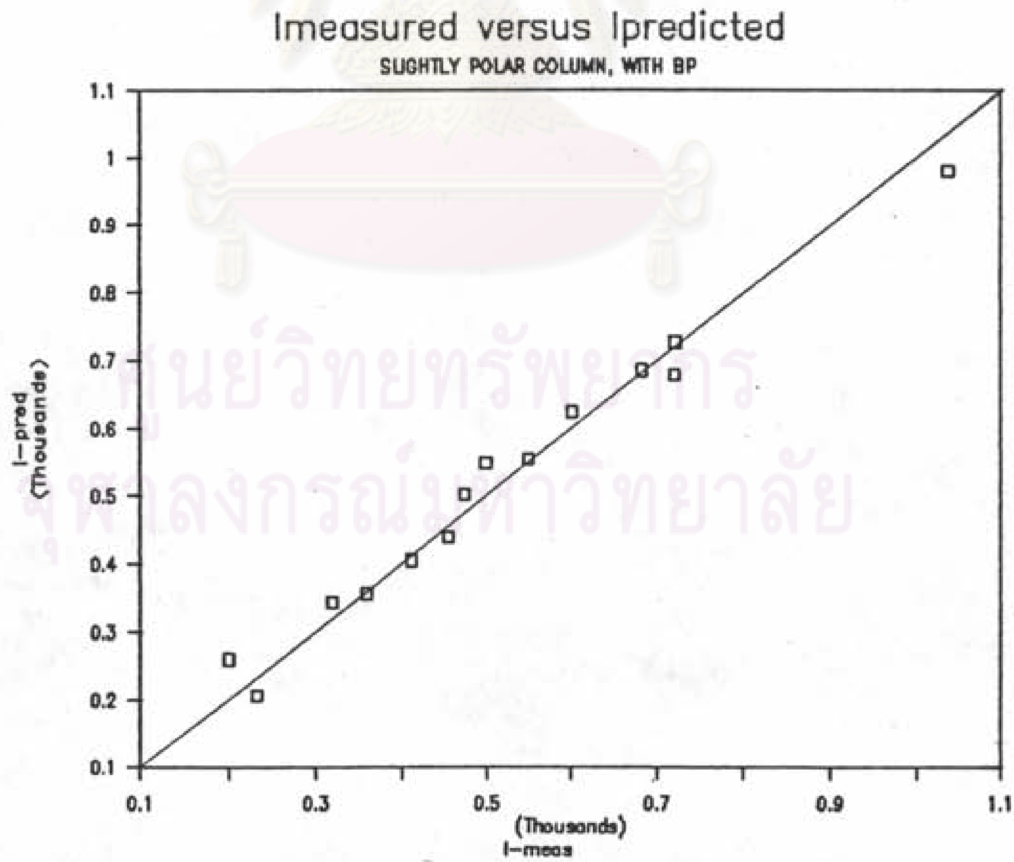


Fig.5.31 results of tested products on slightly polar column with BP

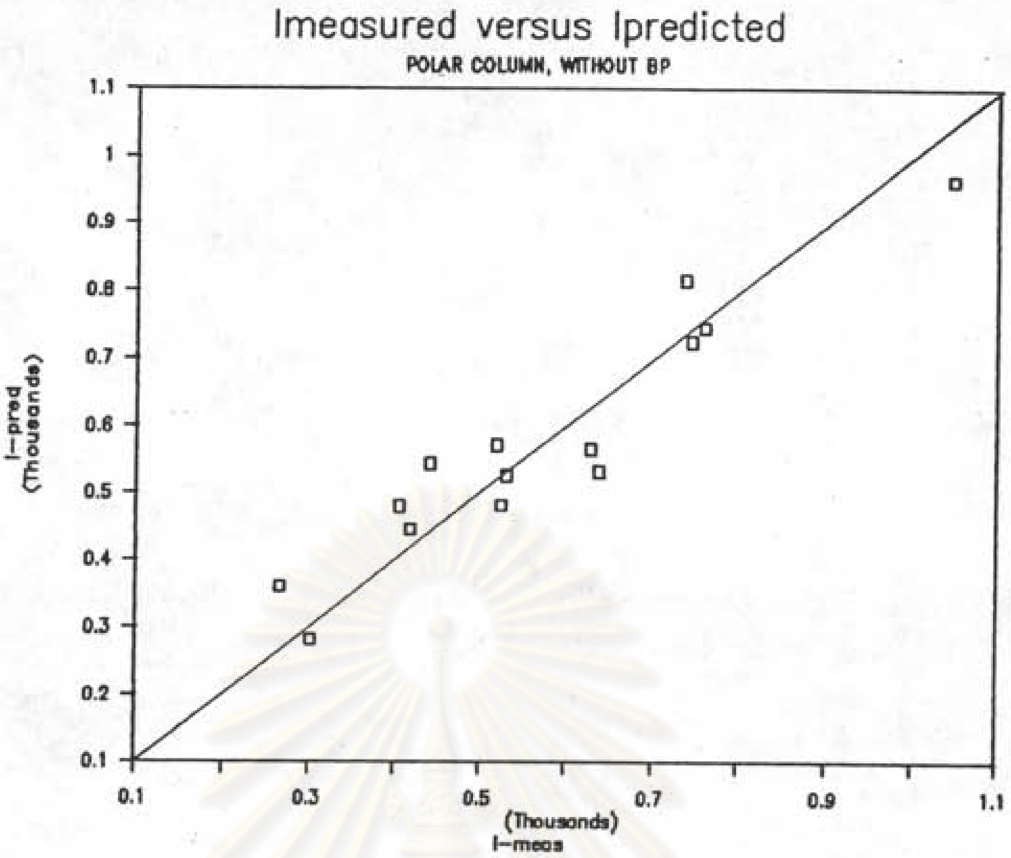


Fig. 5.32 The results of tested products on polar column

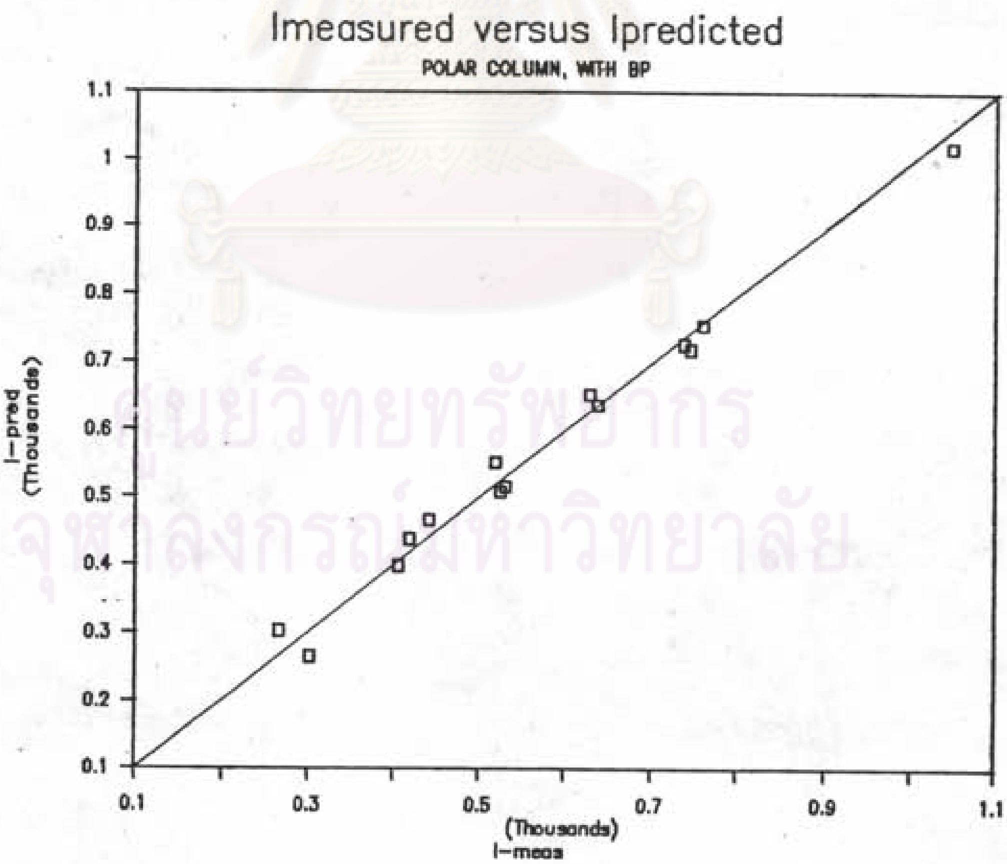


Fig. 5.33 The results of tested products on polar column with BP

Table 5.11 The predicted results of tested products for apolar column using regression models

1. $I\text{-}NP = 159.9 + 83\,{}^{1}\chi + 28.3\,{}^{0}\chi + 270.1\,R - 262.1\,Q + 28.4\,{}^{3}\chi_{p}$
2. $I\text{-}NP = 374.5 + 2.329\,BP + 54.67\,{}^{1}\chi + 11.64\,{}^{4}\chi_{pc} - 21.53\,nC$

Products	Imeas	Ipred(1)	ΔI(1)	Ipred(2)	ΔI(2)
11	479	507.6	-28.6	500.3	-21.3
30	520.7	476.2	44.5	538.8	-18.1
112a	695.7	703.8	-8.1	713.6	-17.9
115	214	278.9	-64.9	258.4	-44.4
122	626.5	626.7	-0.2	610.3	16.2
130	878.1	855.7	22.4	861.2	16.9
133a	401.1	357.4	43.7	393.4	7.7
140a	630.6	655.9	-25.3	629.4	1.2
142	459	394.7	64.3	467.7	-8.7
152a	270.5	273.5	-3	291.3	-20.8
1112-cis	459.7	450.9	8.8	458	1.7
1123	213.9	213.9	0	201.6	12.3
1131a	337	373.8	-36.8	338.4	-1.4
1140	372.2	434.2	-62	381.4	-9.2
		average	29.5	average	14.1

Table 5.12 The predicted results of tested products for slightly polar column using regression models

1. $I\text{-}SP = 48.4 + 78.5\,{}^{1}\chi + 46.6\,{}^{0}\chi + 35\,{}^{3}\chi_{p} + 37.7\,nH + 1.24\,MW$
2. $I\text{-}SP = 510 + 3.506\,BP - 12.57\,nF - 27\,nC - 10.56\,{}^{4}\chi_{pc} - 16.02\,{}^{3}\chi_{c}$

Products	Imeas	Ipred(1)	ΔI(1)	Ipred(2)	ΔI(2)
11	500	545.8	-45.8	547.9	-47.9
30	602.1	521.3	80.8	623.6	-21.5
112a	721.4	718.4	3	726.2	-4.8
115	200	191.2	8.8	259.6	-59.6
122	721.2	687	34.2	678.2	43
130	1037.6	929.4	108.2	979.2	58.4
133a	455.2	420	35.2	439.6	15.6
140a	683.4	764.6	81.2	685	-1.6
142	550.9	463.2	87.7	553.9	-3
152a	318.5	361.6	-43.1	343	-24.5
1112-cis	474.6	495.7	21.1	502.7	-28.1
1123	231.4	228.2	3.2	205.3	26.1
1131a	358.3	429.5	-71.2	356	2.3
1140	411.8	486.4	-74.6	404.9	6.9
		average	49.8	average	24.5

Table 5.13 The predicted results of tested products for polar column using regression models

1. $I\text{-}P = 82.1 + 78.9\,{}^{1}\chi + 43\,{}^{0}\chi + 49.54\ nH + 1.384\ MW + 35.4\,{}^{3}\chi_{p}$
2. $I\text{-}P = 565.4 + 3.595\ BP - 179\ Q + 8.81\ nH + 130.3\ R + 10.43\,{}^{4}\chi_{pc} - 12.82\,{}^{2}\chi$

Products	Imeas	Ipred(1)	ΔI(1)	Ipred(2)	ΔI(2)
11	520	571	-51	552.1	-32.1
30	628.2	565.2	63	652.4	-24.2
112a	760.8	745.7	15.1	755.2	5.6
115	266.7	360.6	-93.9	303.4	-36.7
122	745.2	724.4	20.8	718.4	26.8
130	1047.5	963.5	84	1019.3	28.2
133a	525	480.7	44.3	508.9	16.1
140a	737.1	815	-77.9	726.5	10.6
142	637.9	530.9	107	636.7	1.2
152a	419.4	445.4	-26	438.6	-19.2
1112-cis	530.7	525.4	5.3	516.5	14.2
1123	302.7	282.2	20.5	265.4	37.3
1131a	406.5	480.4	-73.9	398.6	7.9
1140	441.9	542.8	-100.9	467.6	-25.7
		average	56	average	20.4

Table 5.14 Regression of models with D and without D as the parameter for slightly polar column (only products with known D values)

TYPE	PARAMETERS	MODELS	%VARIANCE	S.E.
WITHOUT D	BP		97.9	39.08
	BP,nF		98.6	32.06
	BP,nF,$^4\chi_{PC}$	$I = 474.68+3.081\ BP-25.52\ nF+12.86\ ^4\chi_{PC}$	98.9	27.91
WITH D	BP		97.9	39.08
	BP,nF		98.6	32.06
	BP,nF,$^4\chi_{PC}$		98.9	27.91
	BP,nF,$^4\chi_{PC}$,Q		99.0	26.68
	BP,nF,$^4\chi_{PC}$,Q,nCl		99.0	26.23
	BP,nF,$^4\chi_{PC}$,Q,nCl,D	$I = 523.6+3.282\ BP-8.92\ nF+13.39\ ^4\chi_{PC}-59.2\ Q$ $+27.4\ nCl+16.3\ D$	99.1	25.89 (t-test of D=1.25)

** The theoretical t-test of D with degree of freedom = 20 and 95% confidence interval = 2.086

Table 5.15 Regression of models with D and without D as the parameter for the polar column (only products with known D values)

TYPE	PARAMETERS	MODELS	%VARIANCE	S.E.
WITHOUT D	BP		98.0	35.19
	BP,nF		98.6	29.47
	BP,nF,$^4\chi_{PC}$		98.9	25.73
	BP,nF,$^4\chi_{PC}$,nCl		99.3	21.27
	BP,nF,$^4\chi_{PC}$,nCl,$^4\chi_C$	$I = 552.2+3.392\ BP-17.4\ nF+52.1\ ^4\chi_{PC}-25.22\ nCl$ $-325\ ^4\chi_C$	99.4	19.62
WITH D	BP		98.0	35.19
	BP,D		98.6	29.33
	BP,D,$^3\chi_P$		99.0	25.38
	BP,D,$^3\chi_P$,nF		99.4	19.21
	BP,D,$^3\chi_P$,nF,$^4\chi_{PC}$		99.4	19.19
	BP,D,$^3\chi_P$,nF,$^4\chi_{PC}$,$^4\chi_C$	$I = 474.45+2.8926\ BP+26.18\ D+40.5\ ^3\chi_P-17.86\ nF$ $+48.4\ ^4\chi_{PC}-435\ ^4\chi_C$	99.6	15.69 (t-test of D=5.40)

** The theoretical t-test of D with degree of freedom = 20 and 95% confidence interval = 2.086

Table 5.16 Equation $I = f(BP, ^1\chi)$ on the three columns

STATIONARY PHASE	P_{MC}	EQUATION	%VARIANCE	S.E.
OV-1	222	$I = 331.36+2.32\ BP+60.72\ ^1\chi$	98.9	22.52
DB-1701	789	$I = 413.2+3.276\ BP+21.9\ ^1\chi$	97.8	36.24
DB-210	1520	$I = 486.1+3.402\ BP+2.3\ ^1\chi$	97.4	37.63

Table 5.17 Variation of the coefficients of the Equation $I=f(BP,{}^1\chi)$ with the polarity of the statioary phase

Coefficient	Equation	R
a_0	$a_0 = 310.42 + 0.12\ P_{MC}$	0.9844
a_1	$a_1 = 2.33 + 0.0008\ P_{MC}$	0.8825
a_2	$a_2 = 65.55 - 0.044\ P_{MC}$	0.9663

Table 5.18 Prediction of some non-available products by Equation

$$I = 310.42+2.33\ BP+65.55\ {}^1\chi + [0.12+0.0008\ BP-0.044\ {}^1\chi]P_{MC}$$

Products	Ipred(1)	Ipred(2)	Ipred(3)
31	407.9	430.1	458.6
111	854.6	907.1	974.4
133	446.0	474.2	510.7
1111	636.8	681.9	740.4

Ipred(1) = Predicted indices on apolar column
Ipred(2) = Predicted indices on slightly polar column
Ipred(3) = Predicted indices on polar column

Table 5.19 Equation $I = f({}^1\chi)$ on the three columns

COLUMN	EQUATION	%VARIANCE	S.E.
APOLAR	$I = 200.8+196.78\ {}^1\chi$	93.0	57.26
SLIGHTLY	$I = 228.5+213.8\ {}^1\chi$	88.3	83.15
POLAR	$I = 295.3+201.6\ {}^1\chi$	86.1	86.34

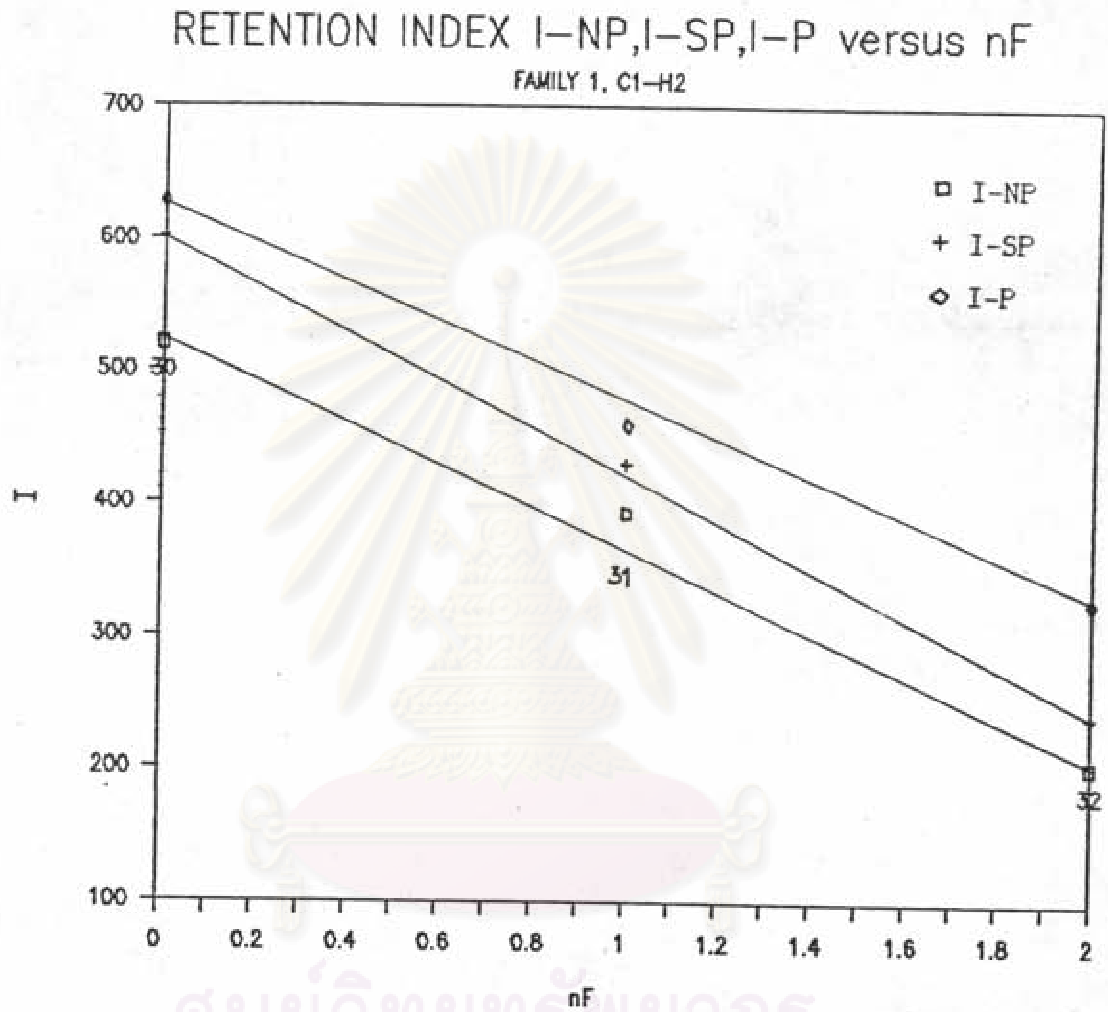


Fig. 5.34 The predicted retention indices of non-available product (31) on three columns

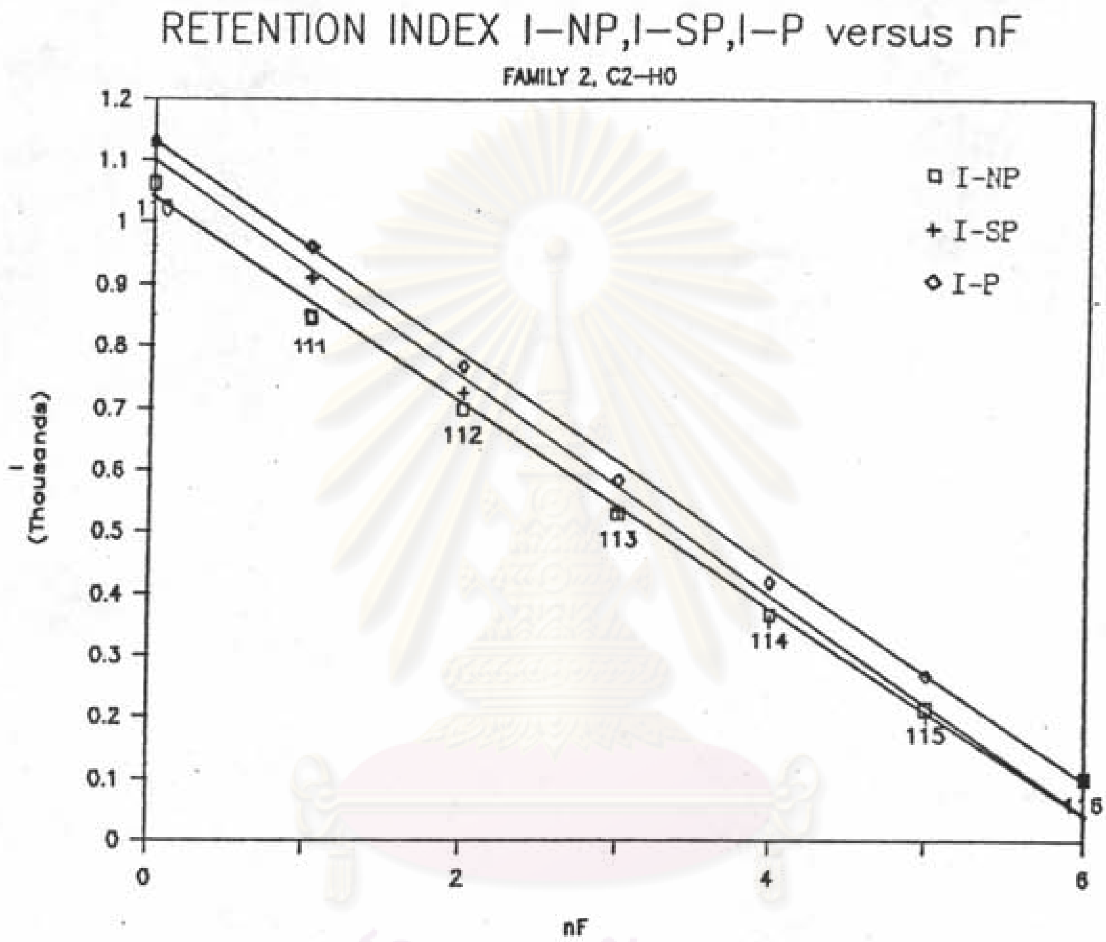


Fig. 5.35 The predicted retention indices of non-available product (111) on three columns

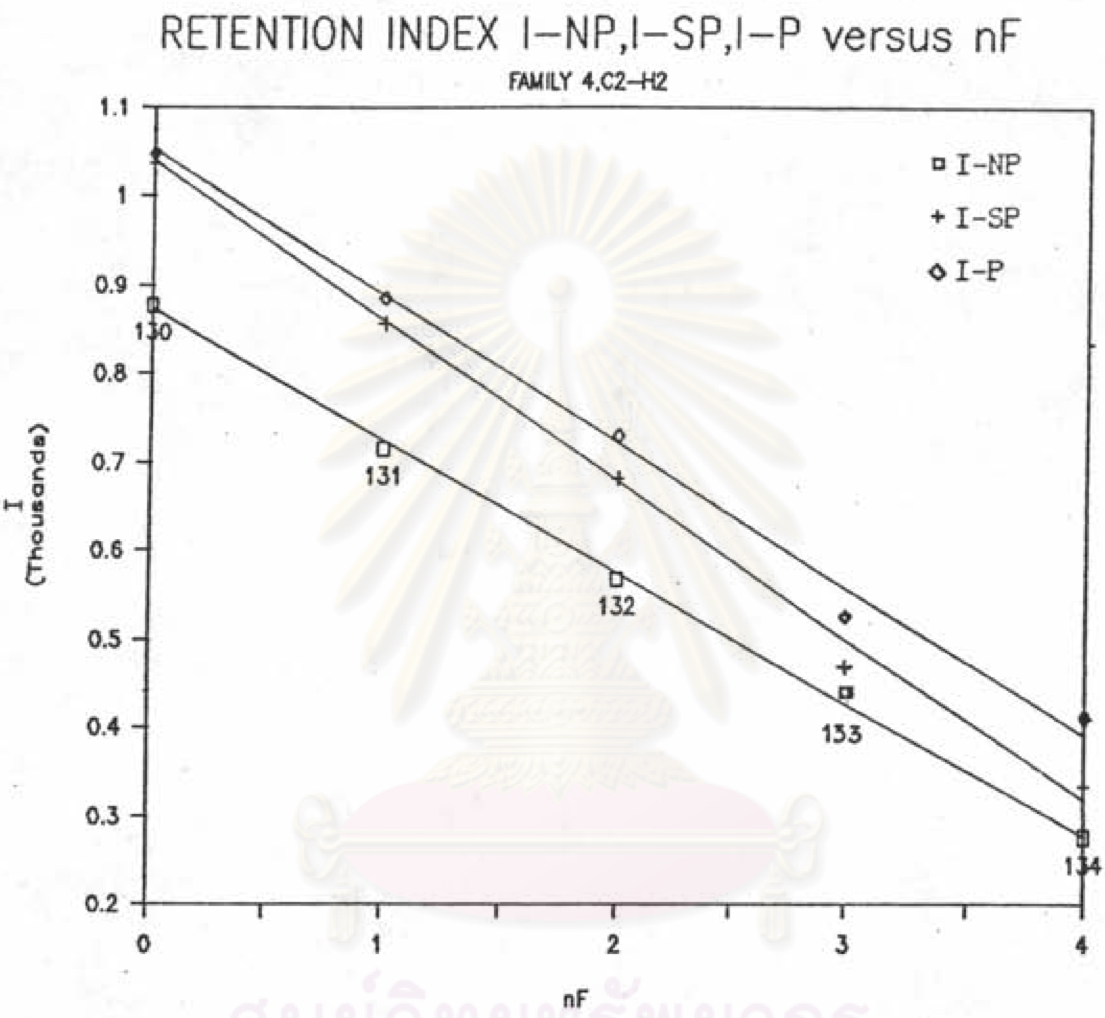


Fig. 5.36 The predicted retention indices of non-available product (133) on three columns

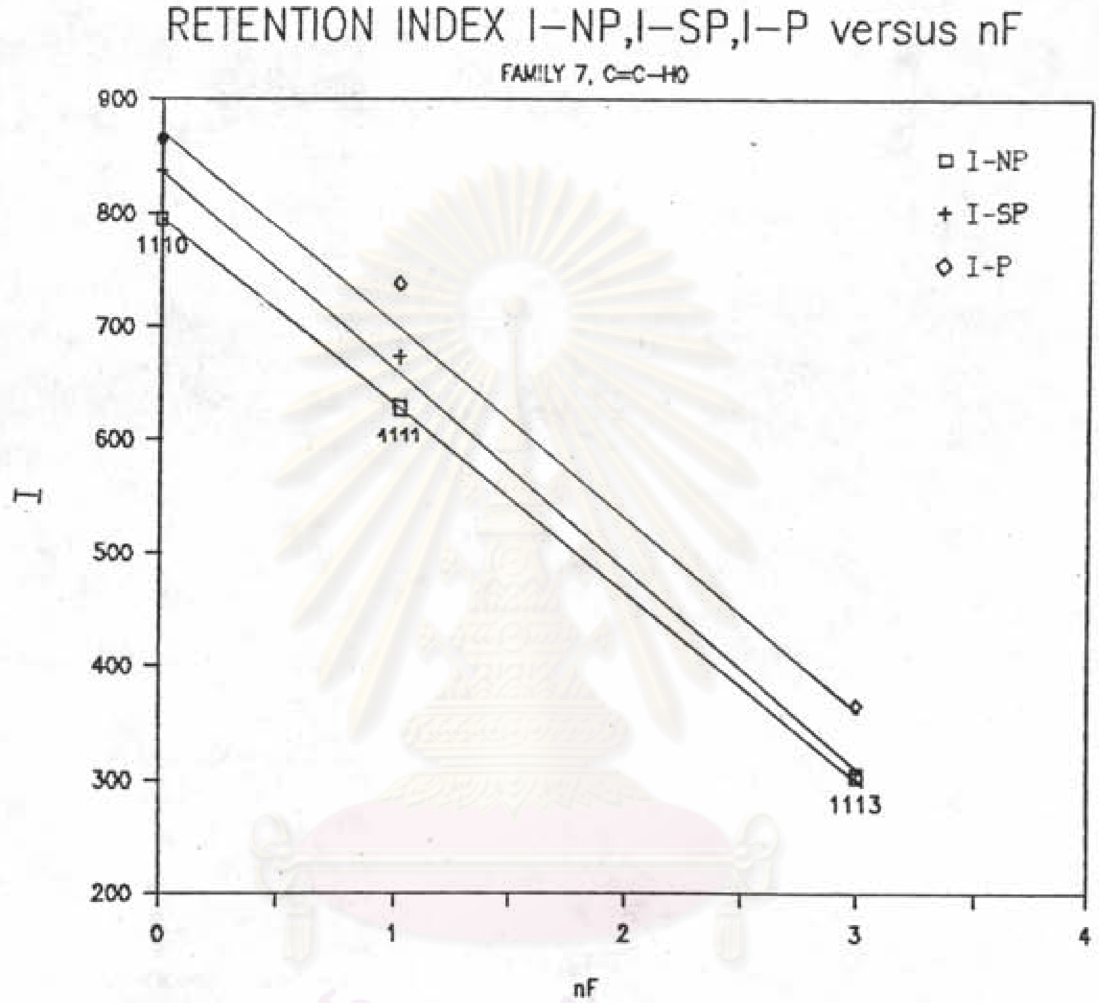


Fig. 5.37 The predicted retention indices of non-available product (1111) on three columns