REFERENCES

1. Mark, R. Schoeberl, "Ozone Depletion, Dynamics Weaken the Polar Hole," Nature, 336(12), 420-421, 1988
2. Barret, J.W., P.M. Soloman, et al., "Formation of the Antarctic Ozone Hole by the ClO Dimer Mechanism," Nature, 336(12), 455-458, 1988
3. European Chemical News, "New CFC Toxics Group Forms," September 12, 1988
4. Van den Dool, H., and P. Dec Kratz, "A Generalization of the Retention Index System Including Linear Temperature Programmed Gas-liquid Partition Chromatography," J. of Chromatography, 11, 463-471, 1963
5. Roman Kalizan, Quantitative Structure-Chromatographic Retention Relationships (Winefordner, J.D. ed.), Chemical Analysis, vol. 93, A Wiley-Interscience Publication, New York, 1987
6. Martin, A.J.P, " Some Theoretical Aspects of Partition Chromatography," Biochem. Soc. Symp., 3,4, 1950
7. Bermejo, J., J.S. Canga, O.M. Gayol, and M.D. Guillen, "Utilization of Physico-Chemical Properties and Structural Parameters for Calculating Retention Indices of Alkylbenzenes," J. of Chromatographic Science 22, 252-255 1984
8. Morishita, F., Y. Terashima, M.Ichise, and T. Kojima, "Prediction of Retention Indices from Molecular Structures of Chlorinated Alkanes," J. of Chromatographic Science, 21, 209-213, 1983
9. Bermejo, J., and M.D. Guillen, "Prediction of Kovats Retention Index of Saturated Alcohols on Stationary Phases of of Different Polarity," Anal. Chem., 59, 94-97, 1987
10. Papp, O., M.Jozan, Gy. Szasz, and I.Hermecz, "Correlation Between Molar Refraction and Gas Chromatographic Retention Data," J. of Chromatography, 403, 11-18, 1987

11. Engewald, W., U. Billing, I.Topalova, and N. Petsev, "Structure-Retention Correlations of Alkylphenols in Gas-Liquid and Gas-Solid Chromatography," J. of Chromatography, 440, 71-77, 1988

12. Saura-Calixto, F., A. Garcia-Raso, and M.A. Raso, "Study of the Applications of Magnitudes of Energy and Charge of Molecular Orbitals to GC Retention-Esters," J. of Chromatographic Science, 22, 22-26, 1984

13. Tarjan, G., Sz. Nyiredy, M. Gyor, E.R. Lombosi, M.V. Budahegyi, S.Y. Meszaros, and J.M. Takacs,"Thirtieth Anniversary of the Retention Index According to Kovats in Gas-Liquid Chromatography," J. of Chromatography, 472, 1-92, 1989

14. James, A. T., and A. J. P. Martin, "Gas-Liquid partition Chromatography : The Separation and Micro-estimation of Volatile Fatty Acids from Formic Acid to Dodecanoic Acid," Biochem. J. 50, 679, 1952

15. McReynolds, W.O., "Characterization of Some Liquid Phases," J. of Chromatographic Science, 8, 685-691, 1970

16. Evans, M.B., and J.K. Haken, "Recent Development in the Gas Chromatographic Retention Index Scheme," J. of Chromatography, 472, 93-127, 1989

17. Yasuhara, A., M. Morita, and K. Fuwa, "Temperature-programmed Retention Indices of 221 Halogenated Organic Compounds with 1-Bromoalkanes as References," J. of Chromatography 328, 35-48, 1985

18. Raymer, J., D. Wiesler, and M. Novotny, "Structure-Retention Studies of Model Ketones by Capillary Gas Chromatography, J. of Chromatography, 325, 13-22, 1985

19. Sabljic, A., "Calculation of Retention Indices by Molecular Topology : Chlorinated Alkanes," J. of Chromatohraphy 3, 1-12, 1984

20. Gorgenyi, M., Z.Fekete, and L. Seres, "Estimation and Prediction of the Retention Indices of Selected trans-Diazenes," Chromatographia, 27(11), 581-584, 1989

21. Bermejo, J.,and M.D. Guillen, "Prediction of Gas Chromatographic Retention Indices of Linear, Branched and Cyclic Alkanes from Their Physico-chemical Properties," J. of High Resolution Chromatography, 7(4), 191-195, 1984

22. Buyden, L., D.L. Massart, and P. Geerlings, "Prediction of Gas Chromatographic Retention Indexes with Topological, Physicochemical and Quantum Chemical Parameters," Anal. Chem., 55, 738-744, 1983

23. Buyden, L., D. Coomans, M. Vanbelle, D.L. Massart, and R.Van den Driessche, "Comparative Study of Topological and Linear Free Energy-Related Parameters for the Prediction of GC Retention Indices," J. of Pharmaceutical Science 72(11), 1327-1329, 1983

24. Bermejo, J., and M.D. Guillen, "A Study of Kovats Retention Indices of Aliphatic Saturated Esters and Their Retention to the Polarity of the Stationary Phase," J. of Chromatography, 318, 187-194, 1985

25. Bondi, A., "Van der Waals Volume and Radii," The Journal of Physical Chemistry, 68(3), 441-451, 1964

26. Kier, L.B., and L.H. Hall, Molecular Connectivity in Chemistry and Drug Research (G. Steven. ed.), Medical Chemistry, vol. 14, Acedemic Press, New York, 1976

27. Budahegyi, M.V.,E.R. Lombosi,S.Y. Meszaros, Sz. Nyiredy,G.Tarjan I. Timer, and J.M. Takacs, "Twenty-Fifth Anniversary of the Retention Index System in Gas-Liquid Chromatography," J. of Chromatography, 271, 213-307, 1983

28. Wold, S., P. Geladi, K. Esbensen, and J. Ohman, "Multi-Way Principal Components and PLS-Analysis," J. of Chemometrics, 1, 41-56, 1987

29. Gmehling, J., P. Rasmussen, and A. Fredenslund, "Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension 2," Ind. Eng. Chem. Process Des. Dev., 21, 118-127, 1982

30. Abrams, D.S., and J.M. Prausnitz, "Statistical Thermodynamics of Liquid Mixtures : A New Expression for the Excess Gibbs Energy of Partly or Completely Miscible System," AIChE Journal, 21(1), 116-128, 1975

31. Randic, M., "On Characterization of Molecular Branching," J. of Am. Chem. Soc., 97, 6609-6615, 1975

32. Kier, L.B., and L.H. Hall, "Derivation and Significance of Valence Molecular Connectivity," J. of Pharmaceutical Sciences, 70(6), 583-589, 1981

33. Wold, S., K. Esbensen, and P. Geladi, "Principal Component Analysis," Chemometrics and Intelligent Laboratory Systems, 2, 37-52, 1987

34. Malinowski, E.R., and D.G. Howery, Factor Analysis in Chemistry, A Wiley-Interscience Publication, New York, 1980

35. Refrigerating Engineering, 65, 49, 1957

APPENDIX

APPENDIX A

CALCULATION OF STRUCTURAL VOLUME (R)

AND SURFACE-AREA (Q) PARAMETERS

Structural volume (R) and surface area (Q) parameters, based on Van der Waals volume (V_w) and surface area (A_w), are the parameters that may be related to the ability of a solute to take part in dispersion interactions. In UNIQUAC or UNIFAC group-contribution methods, which are applied for the prediction of the activity coefficients for non-ideal mixtures, they present the differences in size and shape of the molecules in the mixture. The basic idea of this method is that the properties of the compounds are the sum of contributions made by the molecule's functional groups. So it is possible to correlate the properties of a very large number of compounds in terms of a much smaller number of parameters which characterize the contributions of individual groups. Parameters R and Q can be calculated as the sum of the group volume (R_k) and surface-area (Q_k) parameters :

$$R = \sum_{k} \nu_k R_k \qquad (A.1)$$

$$Q = \sum_{k} \nu_k Q_k \qquad (A.2)$$

where ν_k = integer number of groups of type K in the molecule

R_k, Q_k are obtained from general articles about vapor-liquid equilibrium by UNIFAC or UNIQUAC group-contribution (for this work, these values are obtained from ref.29) or are calculated directly

from the Van der Waals group volume (V_{wk}) and surface area (A_{wk}) given by Bondi (25) :-

$$R_k = V_{wk} / 15.17 \quad ; \quad V_{wk} = cm^3/mole \qquad (A.3)$$

$$Q_k = A_{wk} / (2.5*10^9) \quad ; \quad A_{wk} = cm^2/mole \qquad (A.4)$$

The normalization factors 15.17 and $2.5*10^9$ are given by Abrams and Prausnitz (30).

For example, for 1,1,2 - trichloro, 1,2,2 - trifluoroethane (CCl_2F-$CClF_2$, CFC-113), R and Q may be constructed from the contribution of CCl_2, CCl and three fluorines. So :-

$$CCl_2F\text{-}CClF_2 \quad <-------> \quad CCl_2 + CCl + 3F$$

For CCl_2 ; R_k = 1.8016 Q_k = 1.448 (ref.29)

CCl ; R_k = 1.0060 Q_k = 0.724 (ref.29)

F ; V_{wk} = 5.72 A_{wk} = $1.10*10^9$ (ref.25)

$$R_{113} = \frac{1.8016 + 1.0060 + 5.72}{15.17}$$

$$= 3.9388$$

$$Q_{113} = \frac{1.448 + 0.724 + 1.10*10^9}{2.5*10^9}$$

$$= 3.492$$

By dividing the molecule into groups, one must keep in mind that best results are obtained if the molecule is fragmented in the least numbers of subgroups.

APPENDIX B

CALCULATION OF THE CONNECTIVITY INDEX (χ)

The connectivity index (χ) is a quantitative topological parameter that translates molecular structure into a unique characteristic structural descriptor by means of the chemical graph theory. It describes the general size and shape of the molecule. The formula shows how bonds connect different atoms in the molecule. It was introduced by Randic (31) for the characterization of molecular branching and was further developed by Kier and Hall (26).

To calculate this parameter, the molecule is written down in skeletal form and each atom "i" is assigned a δ_i value. In the original-suppressed graph representation known as the Randic formulation (31), δ_i is a count of the nonhydrogen sigma-bond electrons contributed by atom i.Problems inherent in this formulation are for the unsaturated molecules, cyclic molecules and the treatment of heteroatoms in the molecules. Kier and Hall (32) went beyond the structural fact of adjacency and considered the count of the valence electrons participating in sigma, pi and lone-pair orbitals on each atoms, exclusive of bonds to hydrogen. This count, the valence delta, δ^v, is used to calculate the valence molecular connectivity index, χ^v, in order to account for the nature of the atoms and the unsaturation of bonds. According to the difference between the number of valence electrons (Z^v) and the number of hydrogen atoms attached to that atom (h_i), δ_i^v can be defined as follows :-

$$\delta_i^v = Z^v - h_i \qquad \text{(B.1)}$$

In the case of CFC compounds, δ_i^v values have to be calculated for C, Cl and F atoms.

For the C-atom : $\delta_c^v = 4 - h_i$

the Cl and F-atoms : $\delta_{Cl}^v = \delta_F^v = 7 - h_i$

This means that the δ^v value does not make a distinction between F and Cl atoms so that different molecules are not adequately described by these values. Nevertheless the contributions of each halogen increase with the atomic number. Except fluorine, the halogen are in quantum levels higher than carbon. Kier and Hall (26) reconsidered a more fundamental approach to assign valence values to the heteroatoms in the graph. The objective was to develop valence values for heteroatoms that are non-empirical in the sense of formal connectivity but that are in accord with the underlying electronic structures of the heteroatoms in the molecules. It is apparent that the electronic structure of each halogen influences many properties, in addition to connectivety. Kier and Hall (26) developed empirical values of valencies for the halogens consistent with carbon, nitrogen and oxygen using the molar refraction values of benzene substituents. They found (26) :-

$$\delta_F^v = (-)20$$
$$\delta_{Cl}^v = 0.69$$

It should be noted that the δ^v value for fluorine which is negative leads to a contribution which is substracted to the $^1\chi$ value.

The respective order of connectivity indices are calculated by means of the following equations :

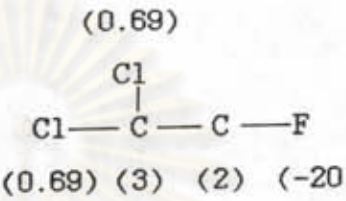
1. <u>Zero-order term</u> ($^0\chi$)

$$^0\chi = \sum_{i=1}^{n} (\delta_i^v)^{-1/2} \qquad (B.2)$$

where δ_i^v = δ^v value of the atom i

n = number of atoms in the molecule

For an example of 1,1-dichloro,2-fluoroethane ($CHCl_2$-CH_2F, HCFC-141a)



$$^0\chi = \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}} + \frac{2}{\sqrt{0.69}} - \frac{1}{\sqrt{20}}$$

$$= 3.469$$

2. First-order term ($^1\chi$)

$$^1\chi = \sum_{l=1}^{N} (\delta_i^v \delta_j^v)_l^{-1/2} \qquad (B.3)$$

where $\delta_i^v \delta_j^v$ = δ^v values of the adjacent atom i and j

N = number of single bonds in the molecule

For $CHCl_2$-CH_2F (HCFC-141a) ;

$$^1\chi = \frac{2}{\sqrt{3(0.69)}} + \frac{1}{\sqrt{3(2)}} - \frac{1}{\sqrt{2(20)}}$$

$$= 1.64$$

3. Second-order term ($^2\chi$)

$$^2\chi = \sum_{q=1}^{M} (\delta_i^v \delta_j^v \delta_k^v)_q^{-1/2} \qquad (B.4)$$

where $\delta_i^V\delta_j^V\delta_k^V$ = δ^V values of atoms of two adjacent bonds

M = number of two adjacent bonds in the molecule

For $CHCl_2$-CH_2F (HCFC-141a)

$$^2\chi = \frac{1}{\sqrt{3(0.69)(0.69)}} + \frac{2}{\sqrt{2(3)(0.69)}} - \frac{1}{\sqrt{2(3)(20)}}$$

$$= 1.728$$

4. Third-order term ($^3\chi$) : There are two possibilities, named a "path" and a "cluster" term

4.1 Path term ($^3\chi_p$)

$$^3\chi_p = \sum_{q=1}^{P} (\delta_i^V\delta_j^V\delta_k^V\delta_l^V)_q^{-1/2} \qquad (B.5)$$

where $\delta_i^V\delta_j^V\delta_k^V\delta_l^V$ = δ^V values of atoms of three adjacent bonds forming a chain

P = number of adjacent bonds forming a chain in the molecule

For $CHCl_2$-CH_2F (HCFC-141a)

$$^3\chi_p = -\frac{2}{\sqrt{(0.69)(2)(3)(20)}}$$

$$= -0.220$$

4.2 Cluster term ($^3\chi_c$)

$$^3\chi_c = \sum_{q=1}^{R} (\delta_i^V\delta_j^V\delta_k^V\delta_l^V)_q^{-1/2} \qquad (B.6)$$

where $\delta_i^v\delta_j^v\delta_k^v\delta_l^v$ = δ^v values of atoms of three adjacent bonds forming a star

R = number of three adjacent bonds forming a star in the molecule

For $CHCl_2$-CH_2F (HCFC-141a)

$$^3\chi_c = \frac{1}{\sqrt{(3)(2)(0.69)(0.69)}}$$

$$= 0.592$$

5. Fourth-order ($^4\chi$) and higher order term

Fourth-order and higher order term can be considered in path, cluster and path/cluster term. The equation can be written in general form :-

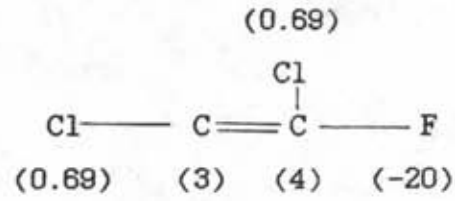
$$^m\chi = \sum_{s=1}^{n_m} \prod_{i}^{m+1} (\delta_i^v)_s^{-1/2} \qquad (B.7)$$

where $^m\chi$ = connectivity index of m^{th} order term

n_m = number of m adjacent bonds in the molecule

6. Unsaturated bonds

In calculating the connectivity index of different order for molecules containing unsaturated bonds, like double or triple bonds, the atoms that are connected with these bonds will be regarded as two or three single bonds by using the same δ^v values. For example, the first-order term for 1,2-dichloro,1-fluoroethylene (CClF=CHCl, HCFC-1121) :-



$$1 = \frac{1}{\sqrt{(0.69)(3)}} + \frac{2}{\sqrt{3(4)}} + \frac{1}{\sqrt{4(0.69)}} - \frac{1}{\sqrt{4(20)}}$$

$$= 1.763$$

APPENDIX C

PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA was first formulated in statistics by Pearson (33), who formulated the analysis as finding "lines and planes of closet fit to system of points in space" and was briefly mentioned by Fischer and Mackenzie (33) as more suitable than analysis of the variance for modelling of response data. Until now, the utility of PCA has been rediscovered in many diverse scientific field and goes under many names. In chemistry, PCA was introduced by Malinowski around 1960 under the name principal factor analysis (PFA) and a large number of chemical applications have been published (34). The summary of PCA as described by Malinowski is mentioned here.

The starting point of multivariate data analysis is a data matrix (a data table)denoted by [D]. The n row designees in the table are termed "objects" and the c column designees are termed "variables". In the technique of analysis, each data in the table is viewed as a linear sum of factros which each of them is weighted differently and can be written in the mathematic form :

$$d_{ik} = s_{i1}l_{1k} + s_{i2}l_{2k} + \ldots + s_{in}l_{nk} = \sum_{j=1}^{n} s_{ij}l_{jk} \qquad (C.1)$$

The symbol "d_{ik}" represents the data point associated with the i^{th} row designee and the k^{th} column designee of the matrix

"s_{ij}" represents the row cofactor associated with the i^{th} row designee and the j^{th} factor of the data matrix, which is called "score"

"l_{jk}" represents the column cofactor associated with the k^{th} column designee and the j^{th} factor of the data

matrix, which is called "loading"

For data modelled by equation (C.1), the data matrix can be decomposed into 2 matrices by standard rules of matrix multification.

$$\underset{\text{data matrix}}{[D]} = \underset{\text{score matrix}}{[S]} \quad \underset{\text{loading matrix}}{[L]} \qquad \text{(C.2)}$$

The score matrix, provides information about the true score of the compounds. The loading matrix will give information about the nature of the measurements.

Analysis Procedures

PCA involves the following main steps : data preparation, reproduction, target testing, abstract rotation, combination and prediction. Figure C-1 shows the sequencing of the steps and the most important information resulting from each step.

1. Data Preparation In this step, the data are selected and pretreated mathematically to obtain a data matrix best suited for analysis. The ultimate success or failure of analysis depends strongly on this step.

2. Reproduction This step provides an abstract solution using the complete models of real factors. Reproduction involves 2 procedures : obtaining the "principal" factor solution, and determining the correct number of factors using a data reproduction method. The final abstract solution is expressed as principal factor matrix. These procedures are summarized in figure C-2.

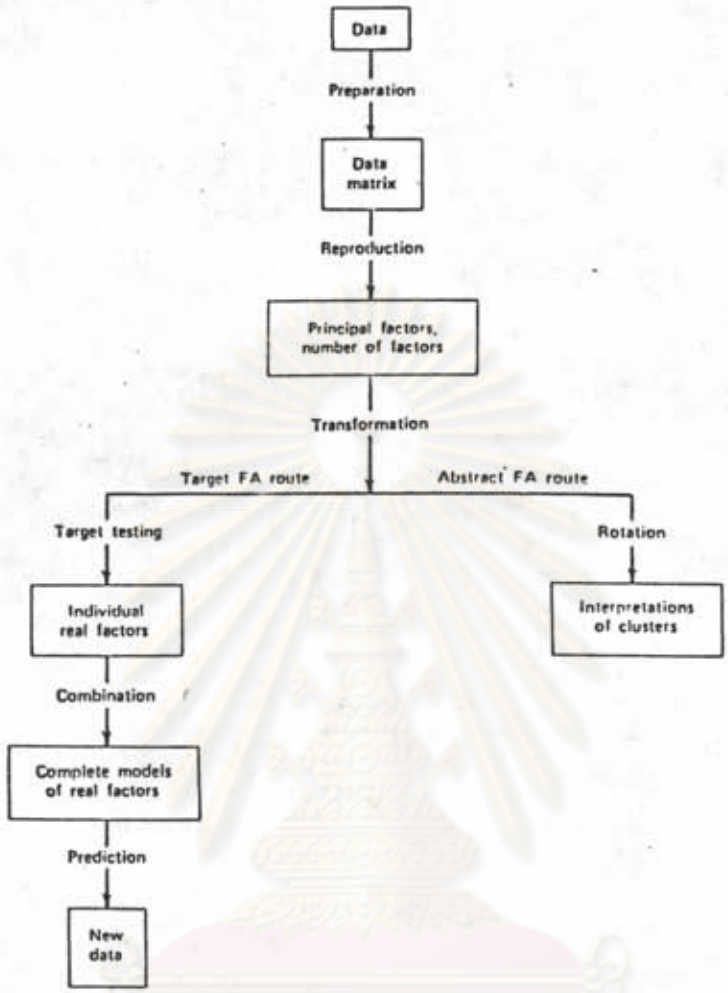


Fig. C.1 Block diagram of the main step in PCA

The procedures for calculating the abstract solution involves a mathematical method called "eigenanalysis" and uses a least-squares technique to carry out the eigenanalysis.

PCA yields an abstract solution consisting a set of "abstract eigenvectors" and an associated set of "abstract eigenvalues". Each principal eigenvector represents an abstract factor and its eigenvalue measures the relative importance of the associated eigenvector. A large eigenvalue indicates a mojor factor, whereas a very small eigenvalue indicates an unimportant factor.

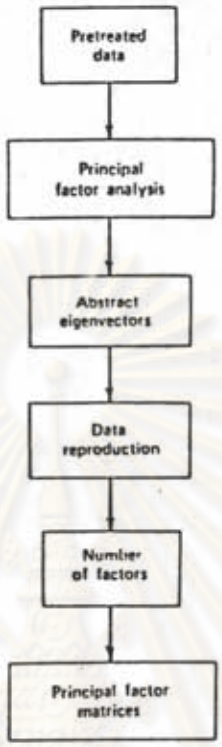


Fig. C.2 Block diagram of the reproduction step

The raw data matrix, [D], is not factor-analyzed directly. Instead, the matrix is first converted into a covariance matrix (or a correlation matrix in case of normalized data). By standard mathematical technique, this matrix is decomposed into a set of "abstact" factors when multiplied via a short-circuit route, it will reproduce the original data. These factors are called abstract because they have no real physical or chemical meaning in the present form, although they have mathematical meaning.

The covariance matrix, [Z] (or $[Z]_N$ for correlation matrix) is constructed by premultiplying the data matrix by its transpose :-

$$[Z] = [D]^T [D] \qquad (C.3)$$

This matrix is then diagonalized by finding a matrix, [Q], such that

$$[Q]^{-1}\,[Z]\,[Q] \;=\; [\lambda_j\ \delta_{jk}] \;=\; [\lambda] \qquad (C.4)$$

here $[\lambda]$ is a matrix containing nonzero elements only on the diagonal

δ_{jk} is the well-known Kronecker delta,

$$\delta_{jk} \;=\; \begin{cases} 0 & \text{if } j = k \\ 1 & \text{if } j = k \end{cases} \qquad (C.5)$$

and λ_j is an eigenvalue of the set of equations

$$[Z]\,Q_j \;=\; \lambda_j\,Q_j \qquad (C.6)$$

where Q_j is the j^{th} column of [Q]. These columns called eigenvectors, constitute a mutually orthogonal set which is usually normalized to form one orthonormal set. Hence

$$[Q]^{-1} \;=\; [Q]^T \qquad (C.7)$$

$[Q]^T$ can be identified with [L] by reasoning as follows :

$$\begin{aligned} [Q]^{-1}\,[Z]\,[Q] &= [Q]^{-1}\,[D]^T\,[D]\,[Q] \\ &= [Q]^T\,[D]^T\,[D]\,[Q] \\ &= [U]^T\,[U] \\ \text{where } [U] &= [D]\,[Q] \end{aligned} \qquad (C.8)$$

Upon rearranging equation (C.8), the data matrix [D] can be exposed as a product of 2 matrices :

$$[D] \;=\; [U]\,[Q]^{-1} \;=\; [U]\,[Q]^T \qquad (C.9)$$

By comparing this equation to equation (C.2), it can be seen that

$$[Q]^T = [L] \qquad (C.10)$$

$$\text{and} \quad [U] = [S] \qquad (C.11)$$

It means that the transpose of the matrix which diagonalizes the covariance matrix represents the column matrix. Since each row of this matrix is an eigenvector, this matrix is generally called the "eigenvector matrix".

From eq.(C.8) and eq.(C.11), the row matrix can be calculated. Having found [S] and [L], then the data matrix can be reproduced. Thus the short-circuit reproduction is readily achieved in an abstract manner.

If the data were free of experimental or theoretical error, PCA would yield exactly n eigenvectors, one for each of the controlling factors. Because of the error, PCA solutions will generate c eigenvectors, one for each of the c columns in the data matrix. However, only n of this set of c eigenvectors, associated with n largest eigenvalues, have physical meaning. From the least-squares nature of PCA, the column matrix is formed using eigenvectors in decreasing order of importance. Factors are ranked according to their ability to account for variation in the data. The first row in the abstract column matrix, representing the first factor, is associated with the largest and most important eigenvalue. The c^{th} row is the least important, being associated with the smallest eigenvalue. In determination of the correct factor size, the abstract reproduction procedure is often utilized. The first eigenvector is used to compute the reproduced data matrix. Then the second eigenvector is chosen, orthogonal to the first, and the reproduced data matrix is computed again. Each stage of

reproduction will be calculated and compared in the following way :

$$[S]_j\,[L]_j \quad = \quad [D]_j \quad \stackrel{?}{=} \quad [D] \qquad (C.12)$$

here $[S]_j$ and $[L]_j$ are abstract matrices based on the j most important eigenvectors

$[D]_j$ is the reproduced data matrix using the first j abstract factors

[D] is the original data matrix

When the correct number of factors is employed (j=n), the reproduced data matrix, $[D]_n$, should equal the original data matrix within the experimental or theoretical error.

3. Transformation Because the main objective of a general chemical problem is to gain insight into the nature of the factors, the abstract solution must be transformed into a meaningful solution. To transform the PCA solution, two distinctly different approaches are employed. The first, target testing, is a unique method for testing potential of each factor at a time and the second, abstract rotation, is a technique for transforming PCA abstract matrices into other abstract matrices.

3.1 Target Testing

Target transformation, involves target testing, serves as a mathematical bridge between abstract and real solutions. Using target testing, one can evaluate ideas concerning the nature of the factors, to develop physically significant models of data. Within target testing, three important features can be achieved ;

a. Each factor can be evaluated independently, although the other factors simultaneously influences the data. This

feature, in particular, distinguishes PCA from multiple regression analysis.

b. Complete models, involving combination of real factors,can be developed using a target combination procedure. Factor analytical procedures can be also employed to predict a new data.

c. Target factor analysis serves as both a theoretical tool and an empirical tool. It can be utilized not only to confirm theory, but also to extend and modify theoretical models. When insight into the lacking data, target testing can be used as a guide to search, term by term, for the best empirical model.

Regardless of the complexity of the problem, the potential factors will be tested individually. The procedure for target testing is summarized by

$$[S]_n T = S_{predicted} \stackrel{?}{=} S_{test} \qquad (C.13)$$

here $[S]_n$ is the row matrix from PCA based on n factors and the other three quantities are vectors. In target transformation technique, a least-squares criterion is used for finding a transformation vector that best constructs the real target vector as a linear combination of abstract cofactors. The target vector, T, results from PCA solution and the individual "target", designated by vector, S_{test}, is tested. If the test vector is a real factor, the predicted vector, $S_{predicted}$, will be similar to the test vector. If the test and predicted vectors are sufficiently dissimilar, the parameter tested is not a real factor, thus leading to rejection of the tested data.

3.2 Rotation

The second method for transforming factor analytical solution involves abstract rotation. Using this approach, the column

matrix obtained from PCA can be mathematically transformed into a new abstract matrix according to

$$[T]_{rotation}\ [L]_n = [L]_{rotated} \qquad (C.14)$$

here $[T]_{rotation}$ is the transformation matrix required to carry out the desired rotation

$[L]_n$ is the principal factor matrix obtained from reproduction step

$[L]_{rotated}$ is the abstract column matrix resulting from the rotation

Rotated solutions are easier to interpret than the principal factor solution. In particular, by bringing out "clustering" in the data, rotated factors are valuable for classifying the designees in a data matrix. If a group of designees form a "cluster",it means that they will explain the system in the same way. Because abstract factors are linear combinations of basic factors, it is difficult to relate rotated factors directly to real factors. So one can attempt to associate the abstract factor with the known basic factor. However, interpretations are always uncertain since the rotation cannot adequately isolate individual basic factor, but these basic factors can be identified using target testing.

4. Combination In the combination step, the data matrix is reproduced from real factors rather than from abstract factors. The combination procedure can be summarized by the following equations ;

$$[S]_{real}\{[T]^{-1}[L]_n\} = [S]_{real}[L]_{column} = [D]_{combine} \stackrel{?}{=} [D] \qquad (C.15)$$

here $[S]_{real}$ is a row matrix formed from a selected set of n real

factors which successfully passed the target test

$[T]^{-1}$ is the inverse of the target transformation matrix

$[L]_n$ is the principal factor matrix

$[L]_{combine}$ is a new column matrix resulted from premultiplying $[L]_n$ by $[T]^{-1}$

$[D]_{combine}$ is a combination-reproduced matrix

If $[S]_{real}$ represents all the factors in the problem, the $[D]_{combine}$ will be reasonably similar to the original data matrix, thus confirming the reliability of the tested combination of the real factors.

For successful combination tests, the factors in $[S]_{real}$ are called "key" factors. Sometimes a best set of factors is indicated from theoretical principles or can be found by examining all or selected combinations of real factors, n at a time. The set of factors which reproduces the data matrix with the least error is the best model, which is useful for predicting new data.

5. Prediction In this step, missing data on test vectors, as well as new data can be obtained using target prediction. For test vectors, the basic data do not have to be completed, missing values for test points can be left blank, called "free-floating". With a successful target test, the free-floated will be predicted automatically.

An extension of the combination-target factor analysis procedure can be used to add new rows and new columns to the original data matrix. For example, to predict a datum associated with a new row designee, x, and an original column designee, k, the basic equation can be employed :

$$d_{xk}(\text{predicted}) = \sum_{j=1}^{n} s_{xj}(\text{known})\, l_{jk}(\text{combination}) \qquad (C.16)$$

where column cofactors, l_{jk}, are taken from $[L]_{combine}$ but the values of row cofactors, s_{xj}, must be known from other sources that are independent of factor analysis.

The summary features of the main steps in factor analysis are shown in table C.1

Step	Purpose	Procedure	Result
1. Preparation	To obtain a matrix best suited for factor analysis	Data selection, data pretreatment	Complete data matrix in suitable form
2. Reproduction	To generate an abstract model	Principal factor analysis, stepwise abstract reproduction	Principal factor matrices, number of factors
3. Transformation			
Target testing	To evaluate test factors individually	Transformation into real factors	Identification of real factors
Rotation	To interpret abstract model	Transformation into new abstract matrices	Clusterings of data
4. Combination	To develop models from sets of real factors	Simultaneous transformation into a set of real factors	Key sets of real factors
5. Prediction	To calculate new data	Free-float missing points, employ key combination set	New target data, new data rows and columns

Table C.1 Features of the main steps in PCA

Example of PCA Analysis

Suppose that there exist different mixtures of two compounds with different concentrations. One can construct a data mixture in which each row concerns a certain wavelength and each column concerns a certain mixture. Each point in this mixture represents the absorbance of the mixtures at a certain wavelength :

$$D = \begin{array}{c} \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} \#1 & \#2 & \#3 & \cdots \end{array} \\ \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ A_{41} & A_{42} & A_{43} & \cdots \\ \cdots & & & \end{bmatrix} \end{array}$$

According to Van Beer's Law, the absorbance of a mixture can be written as the combination of the concentration of the compounds with their molar extinction coefficient (molar absorptivity per unit path length) :

$$A_{11} = \varepsilon_{11}C_{11} + \varepsilon_{12}C_{21} \qquad (C.17)$$

In this way the data matrix can be rewritten as :

$$D = \begin{array}{c} \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{array} \begin{array}{c} \begin{array}{ccc} \#1 & \#2 & \#3 \end{array} \\ \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \\ A_{41} & A_{42} & A_{43} \end{bmatrix} \end{array} = \begin{array}{c} \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{array} \begin{array}{c} \begin{array}{cc} comp1 & comp2 \end{array} \\ \begin{bmatrix} \varepsilon_{12} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \\ \varepsilon_{31} & \varepsilon_{32} \\ \varepsilon_{41} & \varepsilon_{41} \end{bmatrix} \end{array} \begin{array}{c} \\ comp1 \\ comp2 \end{array} \begin{array}{c} \begin{array}{ccc} \#1 & \#2 & \#3 \end{array} \\ \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \end{bmatrix} \end{array}$$

$$[D] = [S]\ [L]$$

Matrix [S] is called the row cofactor and describes the true scores of the products at the different wavelengths, that is why this matrix is also called score matrix.

Matrix [L] is called the column cofactor and describes the importance of the true score to the total results, A_{ij}. This matrix is also called the loading matrix.

Following the block diagram of Fig. C.1, the next step after data preparation would be data reproduction to find the number of factors describing the original data matrix. In this example, the number of factors equal the number of products in the mixtures, which is 2.

With rotation transformation it would be possible to search for cluster in the matrix. The score plots would show cluster of wavelengths containing analogous information because it plots the abstract scores (i.e. molar extinction coefficient) for the different wavelengths of the two principal compounds : the products. The loading plots would show the mixtures having the same composition because here the mixtures are plotted in function of the concentrations of the two compounds. It is not possible with rotation transformation to know the real factors (i.e. the absolute concentrations or molar extinction coefficients of the two products). Rotation transformation always works with abstract factors.

Nevertheless if one has an idea about the products or one of the products, one can use a vector with these products, molar extinction coefficient, for the different wavelengths as target vector. Whether the reproduced vector is similar or dissimilar depends on whether this factor is a real factor of the original data matrix. With target transformation it is possible to learn about the real factors.

In this thesis PCA will be used to find clusters of parameters with retention index containing analogous information. Similarity of parameters will be searched for with rotation transformation using the scores and the loading plots. These plots will show clusters of variables, so that outliers can easily be determined. PCA will also be used to try to predict data for products not available. Data associated with a new row (new product) can be

predicted by using the values of the column cofactors of the original data matrix taken from PCA and the values of the scores of this new data calculated independently of PCA.

APPENDIX D

FLUOROCARBON NUMBERING

Numerical codes to designate simpler fluorocarbons are based on the American Society of Refrigerating Engineers (ASRE) Standard (34) for methane, ethane and cycloalkane refrigerants. The rules for determining code numbers are as follows :

1. The first digit on the right is the number of fluorine atoms in the compounds.

2. The second digit from the right is one more than the number of hydrogen atoms in the compound.

3. The third number from the right is one less than the number of carbon atoms in the compound. When this digit is zero, it is omitted from the number.

4. The number of chlorines is found by subtracting the sum of the fluorine and hydrogen atoms from the total number of atoms which can be connected to carbon atoms.

5. For cyclic derivatives the letter C is used before the identifying number.

6. In the case of isomers, each has the same number and the most symmetrical one is indicated by the number without any lower case letter following it. As the isomers become more unsymmetrical, the letter a,b,c... etc., are appended. Symmetry is determined by adding the atomic weights of the groups attached to each carbon and subtracting one sum from another. The smaller difference the more symmetrical the product (see Table D1 for values).

7. In unsaturated compounds the number of double bonds is shown by the fourth number from the right.

Table D1 Values of the sum of substituents for several groups

Group	CH_3	CH_2F	CH_2Cl	CHF_2	CHClF	CF_3
Sum of Substituents	3	21	38	39	55	57
Group	$CHCl_2$	CF_2Cl	CCl_2F	CCl3		
Sum of Substituents	72	73	90	106		

Table D2 Examples of code numbering without isomers

Compounds	Double Bonds	No. C -1	No. H +1	No. F	FC-No.
$CHClF_2$	0	0	2	2	FC-22
CF_3CHF_2	0	1	2	5	FC-125
CF_2=CHF	1	1	2	3	FC-1123

Table D3 Examples of code numbering with isomers, FC-142's

Compounds	C1*	C2*	(C1-C2) Symmetry	Letter	FC-No.
CHF_2CH_2Cl	39	38	1		142
$CHClFCH_2F$	55	21	34	a	142a
$CClF_2CH_3$	73	3	70	b	142b

* see Table D1

APPENDIX E

EXPERIMENTAL RESULTS OF CFC COMPOUNDS

Table E1 The retention times of family 1 on the three columns

COMPONENT NAME	OV-1 t_R (min)	DB-1701 t_R (min)	DB-210 t_R (min)
METHANE	2.360	2.731	2.792
CF4 (14)	2.536	2.871	2.844
CF3C1 (13)	2.659	2.786	2.844
CHF3(23)	2.659	2.871	2.886
CH3F (41)	2.661	2.939	3.001
ETHANE	2.764	2.871	2.844
CH2F2 (32)	2.849	3.089	3.130
PROPANE	5.012	3.394	3.001
CHC1F2 (22)	5.618	4.271	3.439
CC12F2 (12)	5.991	3.651	3.238
CH3C1 (40)	7.850	5.220	4.117
BUTANE	13.394	5.435	3.535
CHC12F (21)	20.292	13.041	6.218
CC13F (11)	23.565	10.984	5.864
PENTANE	26.267	10.984	5.091
CH2C12 (30)	28.973	20.789	10.683
HEXANE	39.349	20.556	8.860
CHC13 (20)	39.628	31.125	14.491
CC14 (10)	45.401	30.017	14.969
HEPTANE	51.458	31.761	15.424

Table E2 The retention times of family 2 on the three columns

COMPONENT NAME	OV-1 t_R (min)	DB-1701 t_R (min)	DB-210 t_R (min)
METHANE	2.368	2.707	2.791
CF3-CF3 (116)	2.368	2.707	2.791
ETHANE	2.773	2.851	2.843
CC1F2-CF3 (115)	3.011	2.851	2.937
PROPANE	5.040	3.370	2.998
CC1F2-CC1F2 (114)	10.636	4.509	3.865
BUTANE	13.416	5.408	3.532
PENTANE	26.251	10.947	5.084
CF3-CC13 (113a)	30.302	14.549	8.298
CC12F-CC1F2 (113)	30.511	14.696	8.878
HEXANE	39.433	21.301	8.878
CC13-CC1F2 (112a)	51.029	34.067	20.479
CC12F-CC12F (112)	51.367	34.448	21.080
HEPTANE	51.555	31.698	15.458
OCTANE	62.659	42.710	23.749
NONANE	72.854	53.064	32.411
DECANE	82.295	62.688	40.750
CC13-CC13 (110)	87.731	73.614	50.947
UNDECANE	91.073	71.672	50.947
DODECANE	-	80.082	55.974

Table E3 The retention times of family 3 on the three columns

COMPONENT NAME	OV-1 t_R (min)	DB-1701 t_R (min)	DB-210 t_R (min)
METHANE	2.374	2.697	2.802
ETHANE	2.780	2.839	2.855
CHF2-CF3 (125)	3.004	3.046	3.056
PROPANE	5.071	3.353	3.013
CHC1F-CF3 (124)	8.817	5.793	4.190
BUTANE	13.430	5.373	3.545
CHC12-CF3 (123)	24.096	17.204	8.575
CHC1F-CC1F2 (123a)	24.439	16.342	8.932
PENTANE	26.180	10.887	5.105
HEXANE	39.407	21.212	8.932
CHC12-CC1F2 (122)	42.610	33.992	19.225
HEPTANE	51.498	31.653	15.507
CHC12-CC12F (121)	60.722	51.373	32.811
OCTANE	62.573	42.680	23.810
NONANE	72.758	53.036	32.476
CHC12-CC13 (120)	77.789	68.229	46.288
DECANE	82.197	62.676	40.815
UNDECANE	-	71.667	48.674

Table E4 The retention times of family 4 on the three columns

COMPONENT NAME	OV-1 t_R (min)	DB-1701 t_R (min)	DB-210 t_R (min)
METHANE	2.368	2.700	2.799
ETHANE	2.776	2.843	2.851
CF3-CH2F (134a)	3.817	3.465	3.298
CHF2-CHF2 (134)	4.514	4.077	3.773
PROPANE	5.051	3.362	3.006
BUTANE	13.423	5.397	3.537
CF3-CH2Cl (133a)	13.568	8.455	6.058
PENTANE	26.241	10.936	5.093
CClF2-CH2Cl (132b)	30.898	21.977	13.792
CHClF-CHClF (132)	35.053	29.811	17.914
HEXANE	39.383	21.264	8.895
CH2Cl-CCl2F (131a)	48.671	38.949	25.571
HEPTANE	51.496	31.704	15.465
CHCl2-CHClF (131)	53.227	48.585	31.216
OCTANE	62.563	42.735	23.765
CCl3-CH2Cl (130a)	65.209	55.227	37.577
CHCl2-CHCl2 (130)	70.515	66.130	44.523
NONANE	72.749	53.098	32.430
DECANE	82.185	62.743	40.770
UNDECANE	-	71.741	48.630

Table E5 The retention times of family 5 on the three columns

COMPONENT NAME	OV-1 t_R (min)	DB-1701 t_R (min)	DB-210 t_R (min)
METHANE	2.363	2.698	2.796
ETHANE	2.768	2.840	2.848
CF3-CH3 (143a)	2.963	2.988	3.133
PROPANE	5.025	3.353	3.003
CHF2-CH2F (143)	6.420	5.280	5.143
CC1F2-CH3 (142b)	8.622	5.011	4.488
BUTANE	13.407	5.368	3.357
CHF2-CH2C1 (142)	20.995	15.761	11.373
CC12F-CH3 (141b)	24.788	13.909	9.128
PENTANE	26.261	10.887	5.143
CHC1F-CH2C1 (141)	39.281	32.915	22.245
HEXANE	39.343	20.462	8.858
CC13-CH3 (140a)	43.061	29.819	18.553
HEPTANE	51.481	31.688	15.465
CHC12-CH2C1 (140)	56.164	49.612	33.078
OCTANE	62.544	42.727	23.812
NONANE	-	53.109	32.438
DECANE	-	-	40.771

Table E6 The retention times of family 6 on the three columns

COMPONENT NAME	OV-1 t_R (min)	DB-1701 t_R (min)	DB-210 t_R (min)
METHANE	2.360	2.700	2.799
ETHANE	2.765	2.842	2.852
CHF2-CH3 (152a)	4.353	3.732	3.854
CH2F-CH3 (161)	4.526	3.669	3.644
PROPANE	5.019	3.357	3.009
BUTANE	13.392	5.379	3.544
CH2Cl-CH3 (160)	16.228	9.888	6.685
CHClF-CH3 (151a)	17.444	10.035	7.453
PENTANE	26.257	10.903	5.101
CHCl2-CH2Cl (150a)	34.261	24.747	15.053
HEXANE	39.334	20.471	8.883
CH2Cl-CH2Cl (150)	42.230	34.680	23.156
HEPTANE	51.458	31.683	15.476
OCTANE	62.528	42.715	23.829

Table E7 The retention times of family 7 on the three columns

COMPONENT NAME	QV-1 t_R (min)	DB-1701 t_R (min)	DB-210 t_R (min)
METHANE	2.362	2.708	2.788
CH2=CF2 (1132a)	2.671	2.838	2.882
ETHANE	2.767	2.838	2.841
CH2=CHF (1141)	2.979	3.003	2.986
CHF=CF2 (1123)	3.079	3.003	3.016
CHF=CHF (1132)	4.900	4.094	3.769
PROPANE	5.028	3.364	2.986
CC1F=CF2 (1113)	5.339	3.482	3.355
CH2=CC1F (1131a)	8.139	4.540	3.649
CHC1=CF2 (1122)	9.643	5.049	3.979
CH2=CHC1 (1140)	11.095	6.032	4.178
BUTANE	13.427	5.318	3.534
CF2=CC12 (1112a)	19.797	8.715	5.825
CC1F=CC1F (1112-cis)	21.086	9.491	6.265
CC1F=CC1F (1112-trs)	21.324	-	-
PENTANE	26.251	10.890	5.095
CH2=CC12 (1130a)	27.344	15.091	7.883
CHC1=CC1F (1121)	30.230	17.975	9.905
CHC1=CHC1 (1130-trs)	33.442	21.104	10.801
CHC1=CHC1 (1130-cis)	38.172	28.485	16.366
HEXANE	39.332	20.447	8.870
CHC1=CC12 (1120)	49.479	36.223	20.983
HEPTANE	51.466	31.655	15.461
CC12=CC12 (1110)	61.991	46.529	29.457
OCTANE	62.537	42.683	23.793
NONANE	-	53.089	32.470

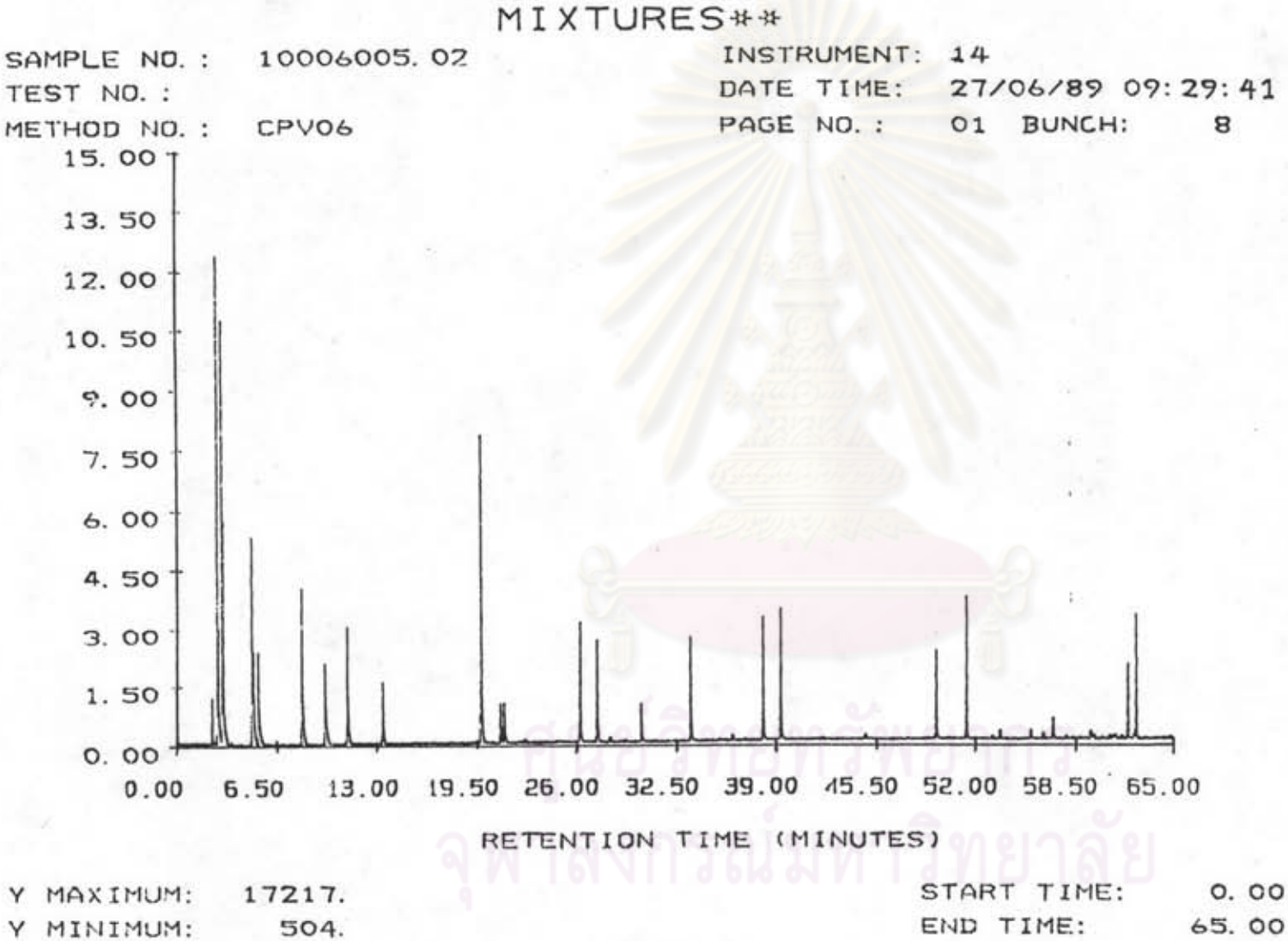


Fig. F1 The chromatogram of family 7 on apolar column

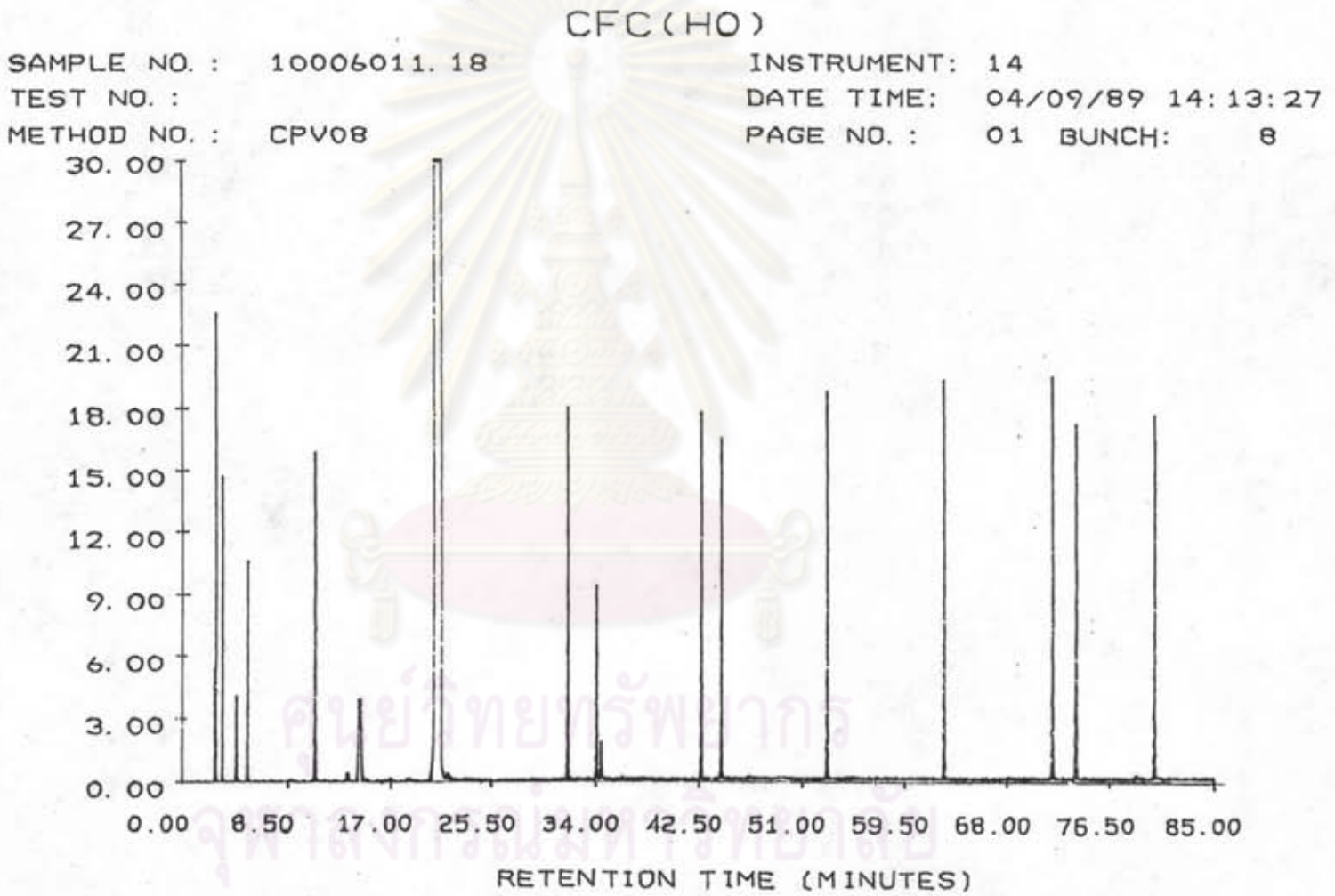


Fig. F2 The chromatogram of family 2 on slightly polar column

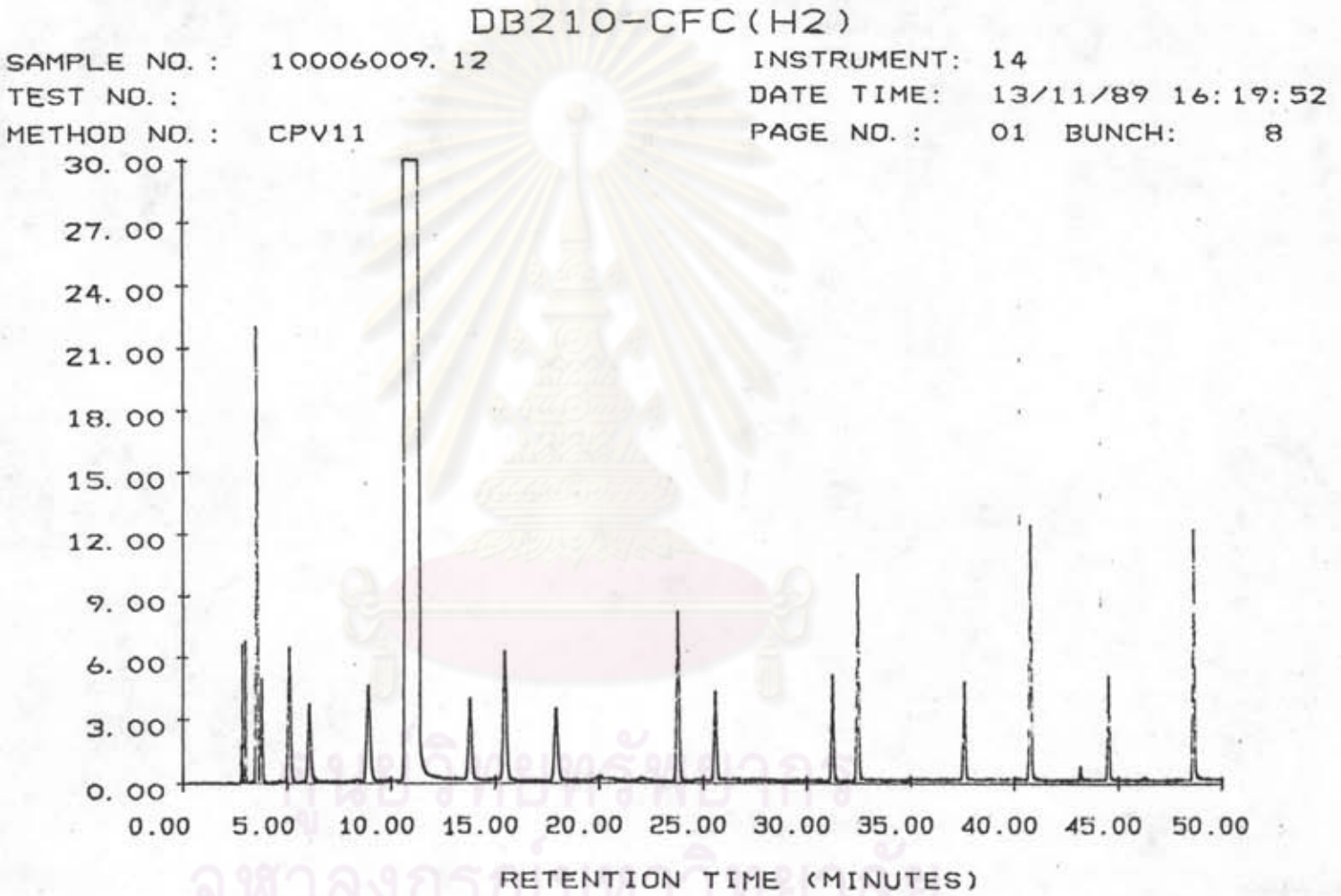


Fig. F3 The chromatogram of family 4 on polar column

APPENDIX G

SAMPLE RESULT OF PRINCIPAL COMPONENT ANALYSIS

Table. G1 The result of PCA calculation of all families on apolar column

Let the following notations be defined as follows :

NF = number of factor
EV = the eigenvalue
RE = real error
IE = imbedded error
IND = the factor indicator function
REV = reduced eigenvalue
SL = significant level
AET = the apparent error in the test vector
EDM = the error contributed by the data matrix
df = degree of freedom

****
** EIGENVALUE ANALYSIS **
****

DATA FILE: B:CFC1

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN

DATA MATRIX: NUMBER OF ROWS = 68 NUMBER OF COLUMNS = 15

SUMMARY OF EVA RESULTS USING (COLUMN) CORRELATION

NF	EV	RE	IE	IND	REV	F(calc)	%SL
1	.101771D+02	.712D-01	.184D-01	.363D-03	.998D-02	13.61	.2
2	.248183D+01	.515D-01	.188D-01	.305D-03	.265D-02	6.38	2.5
3	.888458D+00	.422D-01	.189D-01	.293D-03	.104D-02	3.41	9.0
4	.581142D+00	.341D-01	.176D-01	.282D-03	.745D-03	3.42	9.1
5	.514662D+00	.229D-01	.132D-01	.229D-03	.731D-03	6.76	2.6
6	.151143D+00	.183D-01	.116D-01	.226D-03	.240D-03	3.11	11.1
7	.857507D-01	.148D-01	.101D-01	.232D-03	.154D-03	2.71	13.9
8	.719023D-01	.100D-01	.733D-02	.205D-03	.147D-03	4.99	6.1
9	.204047D-01	.822D-02	.637D-02	.228D-03	.486D-04	2.12	19.6
10	.153309D-01	.600D-02	.490D-02	.240D-03	.433D-04	3.00	14.4
11	.626274D-02	.470D-02	.402D-02	.293D-03	.216D-04	2.02	22.9
12	.514202D-02	.205D-02	.183D-02	.227D-03	.226D-04	8.77	6.0
13	.716359D-03	.101D-02	.937D-03	.252D-03	.426D-05	5.08	15.3
14	.127930D-03	.380D-03	.367D-03	.380D-03	.116D-05	6.41	23.9
15	.979359D-05	--------	--------	--------	.181D-06		

****
** BINARY CROSS VALIDATION **
****

(Based on dividing the rows into two groups)

DATA FILE: B:CFC1

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN

NUMBER OF ROWS = 34
NUMBER OF COLUMNS = 15

NF	SUM OF SQUARES	STAND. DEV.	WOLD S.D.
1	.2826D+06	.1723D+02	.1736D+02
2	.5268D+05	.7720D+01	.7836D+01
3	.3433D+03	.6486D+00	.6634D+00
4	.2209D+03	.5435D+00	.5602D+00
5	.1726D+03	.5037D+00	.5234D+00
6	.1222D+03	.4469D+00	.4681D+00
7	.5947D+02	.3306D+00	.3491D+00
8	.3336D+02	.2647D+00	.2818D+00
9	.2783D+02	.2612D+00	.2804D+00
10	.2172D+02	.2527D+00	.2736D+00
11	.3160D+01	.1078D+00	.1177D+00
12	.1076D+01	.7261D-01	.8002D-01
13	.6642D+00	.6989D-01	.7771D-01
14	.5238D-01	.2775D-01	.3114D-01

```
**********************************************************************************
**                        COMPLETE CROSS VALIDATION                             **
**********************************************************************************

              (based on individual deletion of each data row)

DATA FILE: B:CFC1

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN


   NUMBER OF ROWS =  68
NUMBER OF COLUMNS =  15

NF    SUM OF SQUARES    STAND. DEV.      WOLD S.D.

 1       .2768D+06       .1705D+02       .1718D+02
 2       .5668D+05       .8007D+01       .8128D+01
 3       .3250D+03       .6311D+00       .6455D+00
 4       .1855D+03       .4981D+00       .5134D+00
 5       .1405D+03       .4546D+00       .4723D+00
 6       .5983D+02       .3127D+00       .3275D+00
 7       .2770D+02       .2257D+00       .2383D+00
 8       .1903D+02       .2000D+00       .2129D+00
 9       .1303D+02       .1787D+00       .1919D+00
10       .6260D+01       .1357D+00       .1469D+00
11       .1669D+01       .7834D-01       .8557D-01
12       .1209D+01       .7697D-01       .8482D-01
13       .6282D+00       .6796D-01       .7557D-01
14       .6036D-01       .2979D-01       .3343D-01

**********************************************************************************
**                          PRINT ABSTRACT FACTORS                              **
**********************************************************************************

DATA FILE: B:CFC1

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN


RESULTS BASED ON (COLUMN) CORRELATION USING   5 FACTORS

AbSTRACT ROW FACTOR MATRIX

              1             2             3             4             5

110     .125178D+01   .736819D+00   .342815D+00   .123863D+00   .172423D+00
112     .494998D+00  -.948072D-01  -.110995D+00   .228044D+00  -.141247D+00
112a    .634072D+00   .872084D-01   .128644D+00  -.154551D+00  -.102013D+00
113     .375049D+00  -.182431D+00  -.177072D-01   .842696D-01  -.144602D+00
113a    .486843D+00  -.454453D-01   .183216D+00  -.309763D+00  -.131858D+00
114     .268883D+00  -.221491D+00   .980995D-01   .587899D-01  -.103062D+00
115     .184242D+00  -.249489D+00   .190437D+00   .520635D-01  -.746873D-01
116     .113891D+00  -.271880D+00   .267983D+00   .745916D-01  -.507241D-01
120     .928115D+00   .363489D+00   .642688D-01   .681355D-01   .515669D-01
121     .593748D+00  -.716830D-03  -.135358D+00   .179229D+00  -.701806D-01
122     .432091D+00  -.121171D+00  -.563447D-01   .237158D-01  -.871836D-01
123     .312311D+00  -.195181D+00   .244324D-01  -.553263D-01  -.840942D-01
123a    .293322D+00  -.198209D+00   .200527D-01   .441062D-01  -.572161D-01
124     .205308D+00  -.231297D+00   .109447D+00   .286648D-01  -.324554D-01
125     .129934D+00  -.248483D+00   .193821D+00   .539077D-01  -.299239D-02
130     .657229D+00   .726853D-01  -.170245D+00   .124713D+00  -.631033D-02
130a    .734030D+00   .184581D+00   .569500D-02  -.130435D+00   .300772D-01
131     .465641D+00  -.820476D-01  -.133395D+00   .268291D-01  -.354416D-01
131a    .440873D+00  -.111210D+00  -.151712D+00   .737593D-01  -.356975D-01
```

132	.325654D+00	-.166859D+00	-.636591D-01	.284700D-01	-.142566D-01
132b	.321390D+00	-.170299D+00	-.477931D-01	.205223D-01	-.872451D-02
133a	.227629D+00	-.210707D+00	.403833D-01	-.401542D-02	.130078D-01
134	.143212D+00	-.223771D+00	.120906D+00	.335819D-01	.447170D-01
134a	.141593D+00	-.229248D+00	.125745D+00	.325072D-01	.459020D-01
140	.500110D+00	-.430153D-01	-.189486D+00	.191444D-01	.242138D-01
140a	.578289D+00	.687757D-01	.139047D-01	-.307601D+00	.533327D-01
141	.347262D+00	-.139270D+00	-.126796D+00	.960118D-02	.371370D-01
141b	.314606D+00	-.176771D+00	-.116796D+00	-.645799D-02	.332445D-01
142	.245042D+00	-.182139D+00	-.341612D-01	-.187396D-01	.606182D-01
142b	.219515D+00	-.198767D+00	.548750D-02	-.499594D-02	.839778D-01
143	.157125D+00	-.202995D+00	.502130D-01	.128957D-01	.924326D-01
143a	.135186D+00	-.228744D+00	.886197D-01	.619873D-02	.109444D+00
150	.374159D+00	-.111141D+00	-.185301D+00	-.114573D-01	.912271D-01
150a	.373418D+00	-.112956D+00	-.155640D+00	-.826728D-01	.851799D-01
151a	.242151D+00	-.175830D+00	-.737766D-01	-.253207D-01	.127497D+00
132a	.148077D+00	-.205574D+00	.191107D-01	-.134828D-01	.158977D+00
160	.258210D+00	-.158229D+00	-.126292D+00	-.569999D-01	.181273D+00
161	.147782D+00	-.200032D+00	-.244238D-01	-.357348D-01	.220074D+00
1110	.803806D+00	.371441D+00	.917808D-01	.171319D+00	.439032D-01
1112-cis	.278964D+00	-.144877D+00	-.237485D-01	.728042D-01	-.102776D+00
1112-trs	.279473D+00	-.144504D+00	-.244712D-01	.728700D-01	-.103107D+00
1112a	.325783D+00	-.830637D-01	.845183D-01	-.169734D+00	-.100064D+00
1113	.169516D+00	-.188701D+00	.917586D-01	.272191D-01	-.656112D-01
1120	.548310D+00	.109096D+00	-.450067D-01	.143724D-01	-.123344D-01
1121	.316392D+00	-.100246D+00	-.716000D-01	.429231D-01	-.437100D-01
1122	.191696D+00	-.161063D+00	.416247D-01	-.252408D-01	-.146859D-01
1123	.977605D-01	-.191613D+00	.128132D+00	.335298D-01	.268332D-01
1130-cis	.358910D+00	-.487015D-01	-.128843D+00	.222598D-01	.136928D-01
1130-trs	.351468D+00	-.531930D-01	-.119165D+00	.213738D-01	.179666D-01
1130a	.380959D+00	-.103342D-01	-.295308D-01	-.156626D+00	.229922D-01
1131a	.190862D+00	-.154509D+00	-.205414D-01	-.472737D-02	.447546D-01
1132	.121209D+00	-.164316D+00	.467821D-01	.104910D-01	.689664D-01
1132a	.907669D-01	-.189312D+00	.931035D-01	.348413D-02	.896130D-01
1140	.218975D+00	-.118931D+00	-.694391D-01	-.498981D-01	.985807D-01
1141	.100693D+00	-.173402D+00	.304065D-01	-.169106D-01	.141658D+00
10	.567459D+00	.854954D-01	-.420537D-01	-.250277D+00	-.180619D+00
11	.333833D+00	-.889606D-01	-.339454D-01	-.287316D-01	-.111241D+00
12	.195235D+00	-.161415D+00	.479514D-01	.365064D-01	-.649330D-01
13	.108702D+00	-.183874D+00	.153931D+00	.301334D-01	-.321726D-01
14	.444431D-01	-.218619D+00	.235408D+00	.383545D-01	-.445873D-02
20	.392854D+00	-.362034D-01	-.107653D+00	-.857205D-01	-.766405D-01
21	.244101D+00	-.115944D+00	-.249862D-01	-.109590D-01	-.273174D-01
22	.148144D+00	-.153288D+00	.650987D-01	.742992D-02	.677189D-02
23	.608944D-01	-.183355D+00	.151509D+00	.202902D-01	.367616D-01
30	.273748D+00	-.917909D-01	-.113366D+00	-.274694D-01	.151317D-01
32	.745914D-01	-.158425D+00	.791936D-01	.179435D-02	.833196D-01
40	.165067D+00	-.127925D+00	-.575020D-01	-.292756D-01	.114228D+00
41	.627729D-01	-.159407D+00	.496334D-01	-.209112D-01	.150871D+00

ABSTRACT COLUMN FACTOR MATRIX TRANSPOSED

	1	2	3	4	5
I-NP	.127386D+04	-.551876D+03	-.433390D+03	.483152D+02	.555756D+02
MW	.302387D+03	-.225872D+03	.123368D+03	.631170D+02	-.176749D+03
BP	.132291D+03	.154530D+03	-.241202D+03	.215951D+02	-.120490D+03
CHI(0)	.835814D+01	-.121224D+01	-.463295D+01	-.130576D+01	-.507129D+00
CHI(1)	.424531D+01	-.259785D+00	-.324959D+01	-.377487D+00	-.106136D+01
CHI(2)	.433279D+01	.255834D+01	-.116608D+01	-.294898D+01	-.130247D+01
CHI(3)-P	.143636D+01	.166113D+01	.944494D-01	.361949D+01	.853377D+00
CHI(3)-C	.179304D+01	.210581D+01	.185328D+01	-.334759D+01	-.411555D+00
CHI(4)-PC	.191439D+01	.358637D+01	.271879D+01	.420706D+01	.315534D+01
CHI(4)-C	.273466D+00	.448172D+00	.625200D+00	-.547006D+00	.295414D+00
R	.722930D+01	-.574193D+01	.148381D+01	.633189D+00	-.102307D+01
Q	.629375D+01	-.555239D+01	.185428D+01	.595490D+00	-.775846D+00
nF	.251497D+01	-.852447D+01	.979070D+01	.315753D+01	-.299854D+01
nCl	.573850D+01	.309545D-01	-.198445D+01	.553861D-01	-.458521D+01
nH	.306665D+01	-.577806D+01	-.447877D+01	-.296682D+01	.140869D+02

```
**********************************************************************************
**                          PRINT DATA MATRIX                                   **
**********************************************************************************

DATA FILE: B:CFC2

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN


      NUMBER OF ROWS =  10
   NUMBER OF COLUMNS =   6

                110           112           113a          161           14

I-NP        .106190D+04   .698400D+03   .530700D+03   .278100D+03   .143200D+03
MW          .236800D+03   .203800D+03   .187400D+03   .480000D+02   .880000D+02
CHI(0)      .817100D+01   .279100D+01   .391500D+01   .148300D+01  -.394000D+00
CHI(1)      .383600D+01   .301400D+01   .170700D+01   .549000D+00  -.447000D+00
CHI(2)      .607900D+01   .197700D+01   .231800D+01  -.158000D+00   .150000D+00
CHI(3)-P    .321400D+01   .117400D+01  -.601000D+00   .100000D-06   .100000D-06
CHI(3)-C    .385000D+01  -.244000D+00   .195700D+01   .100000D-06  -.220000D-01
R           .528020D+01   .435730D+01   .404610D+01   .195260D+01   .180160D+01
Q           .436800D+01   .377600D+01   .356400D+01   .182800D+01   .184000D+01
nC1         .600000D+01   .400000D+01   .300000D+01   .100000D-06   .100000D-06

              1112-trs

I-NP        .461600D+03
MW          .132900D+03
CHI(0)      .294300D+01
CHI(1)      .147200D+01
CHI(2)      .704000D+00
CHI(3)-P    .472000D+00
CHI(3)-C   -.240000D-01
R           .299670D+01
Q           .281300D+01
nC1         .200000D+01
**********************************************************************************
**                          PRINT TARGET VECTORS                                **
**********************************************************************************

DATA FILE: B:CFC2

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN


Each column is a basic row test vector.  There are   5target vectors.

                BP          CHI(4)-PC     CHI(4)-C      nF            nH

110         .18500D+03    .76840D+01    .85400D+00    .10000D-06    .10000D-06
112         .92800D+02    .96900D+00   -.16000D+00    .20000D+01    .10000D-06
113a        .45700D+02   -.58400D+00    .42400D+00    .30000D+01    .10000D-06
161        -.37100D+02    .10000D-06    .10000D-06    .10000D+01    .50000D+01
14         -.12800D+03    .10000D-06    .10000D-06    .40000D+01    .10000D-06
1112-trs    .22000D+02   -.26000D+00   -.67000D-01    .20000D+01    .10000D-06
```

```
**********************************************************************************
**                          EIGENVALUE ANALYSIS                                 **
**********************************************************************************

DATA FILE: B:CFC2

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN


DATA MATRIX:   NUMBER OF ROWS =  10    NUMBER OF COLUMNS =   6

SUMMARY OF EVA RESULTS USING (ROW) CORRELATION

NF        EV          RE          IE          IND         REV        F(calc)    %SL

 1   .890796D+01   .148D+00    .603D-01    .591D-02    .148D+00      15.63     1.1
 2   .656517D+00   .104D+00    .602D-01    .652D-02    .146D-01       2.34    20.0
 3   .314628D+00   .635D-01    .449D-01    .705D-02    .983D-02       3.09    17.7
 4   .748214D-01   .480D-01    .392D-01    .120D-01    .356D-02       1.31    37.0
 5   .411154D-01   .223D-01    .203D-01    .223D-01    .343D-02       3.46    31.4
 6   .495403D-02   --------    --------    --------    .991D-03
**********************************************************************************
**                          TARGET TESTING                                      **
**********************************************************************************

DATA FILE: B:CFC2

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN


      NUMBER OF ROWS =  10
   NUMBER OF COLUMNS =   6
   NUMBER OF FACTORS =   5

SUMMARY OF TARGET ERRORS USING   5 FACTORS BASED ON (ROW) CORRELATION

 TARGET         AET          EDM          SPOIL       F      df1   df2     %SL

BP           .1158D+01    .1093D+02       .00       .00      1     1     96.6
CHI(4)-PC    .5291D+00    .1711D+00      2.93      2.39      1     1     36.5
CHI(4)-C     .2810D-01    .2225D-01       .77       .40      1     1     64.1
nF           .4800D+00    .2657D+00      1.50       .82      1     1     53.2
nH           .2814D+01    .4042D+00      6.89     12.12      1     1     17.8

TARGET = BP

    COLUMN       TARGET VECTOR     PREDICTED VECTOR   DIFFERENCE

  1 110          .185000D+03       .185010D+03       .101541D-01
  2 112          .928000D+02       .931379D+02       .337950D+00
  3 113a         .457000D+02       .457428D+02       .428109D-01
  4 161         -.371000D+02      -.364484D+02       .651576D+00
  5 14          -.128000D+03      -.128057D+03      -.567646D-01
  6 1112-trs     .220000D+02       .211078D+02      -.892232D+00

TARGET = CHI(4)-PC

    COLUMN       TARGET VECTOR     PREDICTED VECTOR   DIFFERENCE

  1 110           .768400D+01       .767936D+01      -.464116D-02
  2 112           .969000D+00       .814532D+00      -.154468D+00
  3 113a         -.584000D+00      -.603568D+00      -.195677D-01
  4 161           .100000D-06      -.297818D+00      -.297818D+00
  5 14            .100000D-06       .259457D-01       .259456D-01
  6 1112-trs     -.260000D+00       .147815D+00       .407815D+00
```

TARGET = CHI(4)-C

COLUMN	TARGET VECTOR	PREDICTED VECTOR	DIFFERENCE
1 110	.854000D+00	.853753D+00	-.246510D-03
2 112	-.160000D+00	-.168204D+00	-.820439D-02
3 113a	.424000D+00	.422961D+00	-.103932D-02
4 161	.100000D-06	-.158182D-01	-.158183D-01
5 14	.100000D-06	.137817D-02	.137807D-02
6 1112-trs	-.670000D-01	-.453394D-01	.216606D-01

TARGET = nF

COLUMN	TARGET VECTOR	PREDICTED VECTOR	DIFFERENCE
1 110	.100000D-06	.421037D-02	.421027D-02
2 112	.200000D+01	.214013D+01	.140127D+00
3 113a	.300000D+01	.301775D+01	.177510D-01
4 161	.100000D+01	.127017D+01	.270168D+00
5 14	.400000D+01	.397646D+01	-.235367D-01
6 1112-trs	.200000D+01	.163005D+01	-.369953D+00

TARGET = nH

COLUMN	TARGET VECTOR	PREDICTED VECTOR	DIFFERENCE
1 110	.100000D-06	-.246871D-01	-.246872D-01
2 112	.100000D-06	-.821644D+00	-.821644D+00
3 113a	.100000D-06	-.104084D+00	-.104084D+00
4 161	.500000D+01	.341585D+01	-.158415D+01
5 14	.100000D-06	.138010D+00	.138010D+00
6 1112-trs	.100000D-06	.216925D+01	.216925D+01

****
** TARGET LOADINGS **
****

DATA FILE: B:CFC2

PRINCIPAL COMPONENT ANALYSIS OF ALL CFCs FOR APOLAR COLUMN

NUMBER OF ROWS = 10
NUMBER OF COLUMNS = 6
NUMBER OF FACTORS = 5

FACTOR LOADINGS BASED ON (ROW) CORRELATION

	BP	CHI(4)-PC	CHI(4)-C	nF	nH
I-NP	.369046D+01	.637402D+02	-.128417D+03	.152571D+03	.646462D+02
MW	.959601D+00	.984130D+01	-.190359D+02	.525814D+02	.735914D+01
CHI(0)	.238971D-01	.208190D+00	.253593D+01	.644706D+00	.542426D+00
CHI(1)	.185784D-01	.177783D+00	-.113149D+01	.481640D+00	.161607D+00
CHI(2)	.115159D-01	.318636D+00	.175442D+01	.408652D+00	-.521661D-01
CHI(3)-P	.312066D-02	.529886D+00	-.167714D+01	.960989D-01	.384849D-01
CHI(3)-C	.246731D-02	.130862D-01	.385647D+01	.728526D-01	.925767D-02
R	.211774D-01	.216809D+00	-.351234D+00	.112302D+01	.369436D+00
Q	.176924D-01	.188694D+00	-.410314D+00	.101970D+01	.352226D+00
nC1	.235151D-01	.261312D+00	-.419205D+00	.752194D+00	.266918D-01

ERROR ESTIMATIONS OF THE FACTOR LOADINGS

	BP	CHI(4)-PC	CHI(4)-C	nF	nH
I-NP	.282411D+00	.116048D+02	.835694D+02	.828493D+01	.112477D+02
MW	.268350D-01	.110271D+01	.794087D+01	.787244D+00	.106877D+01
CHI(0)	.458678D-02	.188480D+00	.135729D+01	.134560D+00	.182680D+00
CHI(1)	.238885D-03	.981627D-02	.706895D-01	.700804D-02	.951419D-02
CHI(2)	.564102D-03	.231801D-01	.166926D+00	.165487D-01	.224667D-01
CHI(3)-P	.802582D-03	.329797D-01	.237495D+00	.235449D-01	.319648D-01
CHI(3)-C	.128249D-03	.527001D-02	.379507D-01	.376237D-02	.510783D-02
R	.107753D-02	.442779D-01	.318857D+00	.316109D-01	.429153D-01
Q	.137728D-02	.565951D-01	.407556D+00	.404045D-01	.548535D-01
nC1	.615315D-04	.252845D-02	.182081D-01	.180512D-02	.245064D-02

DATA REPRODUCED WITH 5 BASIC FACTORS

	110	112	113a	161	14
I-NP	.106285D+04	.729928D+03	.534694D+03	.338886D+03	.137904D+03
MW	.236890D+03	.206796D+03	.187780D+03	.537759D+02	.874968D+02
CHI(0)	.818639D+01	.330305D+01	.397987D+01	.247025D+01	-.480009D+00
CHI(1)	.383680D+01	.304067D+01	.171038D+01	.600418D+00	-.451480D+00
CHI(2)	.607711D+01	.191403D+01	.231002D+01	-.279417D+00	.160577D+00
CHI(3)-P	.321669D+01	.126360D+01	-.589650D+00	.172747D+00	-.150495D-01
CHI(3)-C	.385043D+01	-.229683D+00	.195881D+01	.276041D-01	-.244051D-01
R	.528381D+01	.447759D+01	.406134D+01	.218453D+01	.178139D+01
Q	.437262D+01	.392976D+01	.358348D+01	.212444D+01	.181417D+01
nC1	.600021D+01	.400687D+01	.300087D+01	.132439D-01	-.115371D-02

	1112-trs
I-NP	.378363D+03
MW	.124991D+03
CHI(0)	.159111D+01

CHI(1)	.140139D+01
CHI(2)	.870261D+00
CHI(3)-P	.235451D+00
CHI(3)-C	-.617995D-01
R	.267911D+01
Q	.240707D+01
nC1	.198186D+01

ERRORS IN DATA REPRODUCTION USING 5 BASIC FACTORS

	110	112	113a	161	14
I-NP	.947349D+00	.315275D+02	.399389D+01	.607858D+02	-.529560D+01
MW	.900283D-01	.299580D+01	.379514D+00	.577593D+01	-.503192D+00
CHI(0)	.153858D-01	.512054D+00	.648664D-01	.987253D+00	-.860086D-01
CHI(1)	.801589D-03	.266686D-01	.337851D-02	.514173D-01	-.447938D-02
CHI(2)	-.189171D-02	-.629743D-01	-.797723D-02	-.121417D+00	.105776D-01
CHI(3)-P	.269226D-02	.895978D-01	.113502D-01	.172747D+00	-.150495D-01
CHI(3)-C	.430440D-03	.143175D-01	.181385D-02	.276040D-01	-.240500D-02
R	.361466D-02	.120293D+00	.152387D-01	.231926D+00	-.202052D-01
Q	.462002D-02	.153756D+00	.194776D-01	.296444D+00	-.258259D-01
nC1	.206873D-03	.686954D-02	.870512D-03	.132438D-01	-.115375D-02

	1112-trs
I-NP	-.832366D+02
MW	-.790924D+01
CHI(0)	-.135189D+01
CHI(1)	-.704080D-01
CHI(2)	.166261D+00
CHI(3)-P	-.236550D+00
CHI(3)-C	-.377997D-01
R	-.317587D+00
Q	-.405933D+00
nC1	-.181355D-01

RMS ERRORS FOR TFA REPRODUCED DATA ROWS

ROW		RMS ERROR
1	I-NP	.4409D+02
2	MW	.4189D+01
3	CHI(0)	.7160D+00
4	CHI(1)	.3729D-01
5	CHI(2)	.8806D-01
6	CHI(3)-P	.1253D+00
7	CHI(3)-C	.2002D-01
8	R	.1682D+00
9	Q	.2150D+00
10	nC1	.9606D-02

APPENDIX H

SAMPLE RESULT OF MULTIPARAMETER REGRESSION ANALYSIS

Table. H1 The result of regression analysis of all families for all parameters except dipole moment (D) on apolar column

***** REGRESSION ANALYSIS *****

RESPONSE VARIATE: INP
FITTED TERMS: CONSTANT, BP

*** SUMMARY OF ANALYSIS ***

	D.F.	S.S.	M.S.
REGRESSION	1	2372751.	2372750.5
RESIDUAL	51	50336.	987.0
TOTAL	52	2423087.	46597.8
CHANGE	-1	-2372751.	2372750.5

PERCENTAGE VARIANCE ACCOUNTED FOR 97.9

* MESSAGE: THE FOLLOWING UNITS HAVE LARGE RESIDUALS:

11	2.40
34	-2.96
53	2.39

* MESSAGE: THE FOLLOWING UNITS HAVE HIGH LEVERAGE:

11	0.135
17	0.104

*** ACCUMULATED ANALYSIS OF VARIANCE ***

CHANGE	D.F.	S.S.	M.S.	V.R.
+ BP	1	2372750.5	2372750.5	2404.04
RESIDUAL	51	50336.3	987.0	
TOTAL	52	2423086.7	46597.8	

*** ESTIMATES OF REGRESSION COEFFICIENTS ***

	ESTIMATE	S.E.	T
CONSTANT	392.40	4.57	85.81
BP	3.2396	0.0661	49.03

*** CORRELATIONS ***

CONSTANT	1	1.000	
BP	2	-0.331	1.000
		1	2

***** REGRESSION ANALYSIS *****

RESPONSE VARIATE: INP
FITTED TERMS: CONSTANT, BP, CH1

*** SUMMARY OF ANALYSIS ***

	D.F.	S.S.	M.S.
REGRESSION	2	2397725.	1198862.5
RESIDUAL	50	25362.	507.2
TOTAL	52	2423087.	46597.8
CHANGE	-1	-24975.	24974.6

PERCENTAGE VARIANCE ACCOUNTED FOR 98.9

* MESSAGE: THE FOLLOWING UNITS HAVE LARGE RESIDUALS:

7	2.62
11	3.24

* MESSAGE: THE FOLLOWING UNITS HAVE HIGH LEVERAGE:

11	0.135
28	0.144
34	0.195
52	0.143
53	0.147

*** ACCUMULATED ANALYSIS OF VARIANCE ***

CHANGE	D.F.	S.S.	M.S.	V.R.
+ BP	1	2372750.5	2372750.5	4677.83
+ CH1	1	24974.6	24974.6	49.24
RESIDUAL	50	25361.7	507.2	
TOTAL	52	2423086.7	46597.8	

*** ESTIMATES OF REGRESSION COEFFICIENTS ***

	ESTIMATE	S.E.	T
CONSTANT	331.36	9.30	35.65
BP	2.323	0.139	16.72
CH1	60.72	8.65	7.02

*** CORRELATIONS ***

CONSTANT	1	1.000		
BP	2	0.840	1.000	
CH1	3	-0.936	-0.940	1.000
		1	2	3

***** REGRESSION ANALYSIS *****

RESPONSE VARIATE: INP
FITTED TERMS: CONSTANT, BP, CH1, CH4PC

*** SUMMARY OF ANALYSIS ***

	D.F.	S.S.	M.S.
REGRESSION	3	2406237.	802079.0
RESIDUAL	49	16850.	343.9
TOTAL	52	2423087.	46597.8
CHANGE	-1	-8512.	8511.9

PERCENTAGE VARIANCE ACCOUNTED FOR 99.3

* MESSAGE: THE FOLLOWING UNITS HAVE LARGE RESIDUALS:

7	2.61

* MESSAGE: THE FOLLOWING UNITS HAVE HIGH LEVERAGE:

11	0.675
34	0.195

*** ACCUMULATED ANALYSIS OF VARIANCE ***

CHANGE	D.F.	S.S.	M.S.	V.R.
+ BP	1	2372750.5	2372750.5	6900.10
+ CH1	1	24974.6	24974.6	72.63
+ CH4PC	1	8511.9	8511.9	24.75
RESIDUAL	49	16849.7	343.9	
TOTAL	52	2423086.7	46597.8	

*** ESTIMATES OF REGRESSION COEFFICIENTS ***

	ESTIMATE	S.E.	T
CONSTANT	332.59	7.66	43.43
BP	2.220	0.116	19.09
CH1	58.49	7.14	8.19
CH4PC	12.27	2.47	4.98

*** CORRELATIONS ***

CONSTANT	1	1.000			
BP	2	0.820	1.000		
CH1	3	-0.935	-0.912	1.000	
CH4PC	4	0.032	-0.179	-0.063	1.000
		1	2	3	4

***** REGRESSION ANALYSIS *****

RESPONSE VARIATE: INP
FITTED TERMS: CONSTANT, BP, CH1, CH4PC, NC

*** SUMMARY OF ANALYSIS ***

	D.F.	S.S.	M.S.
REGRESSION	4	2409419.	602354.6
RESIDUAL	48	13668.	284.8
TOTAL	52	2423087.	46597.8
CHANGE	-1	-3182.	3181.5

PERCENTAGE VARIANCE ACCOUNTED FOR 99.4

* MESSAGE: THE FOLLOWING UNITS HAVE HIGH LEVERAGE:
11 0.675

*** ACCUMULATED ANALYSIS OF VARIANCE ***

CHANGE	D.F.	S.S.	M.S.	V.R.
+ BP	1	2372750.5	2372750.5	8332.62
+ CH1	1	24974.6	24974.6	87.71
+ CH4PC	1	8511.9	8511.9	29.89
+ NC	1	3181.5	3181.5	11.17
RESIDUAL	48	13668.2	284.8	
TOTAL	52	2423086.7	46597.8	

*** ESTIMATES OF REGRESSION COEFFICIENTS ***

	ESTIMATE	S.E.	T
CONSTANT	374.5	14.3	26.12
BP	2.329	0.111	21.03
CH1	54.67	6.60	8.29
CH4PC	11.64	2.25	5.17
NC	-21.53	6.44	-3.34

*** CORRELATIONS ***

CONSTANT	1	1.000				
BP	2	0.639	1.000			
CH1	3	-0.599	-0.909	1.000		
CH4PC	4	-0.057	-0.195	-0.047	1.000	
NC	5	-0.874	-0.296	0.173	0.083	1.000
		1	2	3	4	5

*** FITTED VALUES AND RESIDUALS ***

UNIT	RESPONSE	FITTED VALUE	STANDARDIZED RESIDUAL	LEVERAGE
1	650.0	662.5	-0.81	0.177
2	311.7	336.7	-1.57	0.113
3	174.1	177.7	-0.23	0.146
4	602.3	608.6	-0.41	0.154
5	453.6	442.0	0.73	0.108
6	337.3	281.5	1.61	0.103
7	174.1	140.5	2.14	0.136
8	203.8	215.2	-0.74	0.172
9	333.9	362.4	-1.81	0.128
10	174.1	157.8	1.03	0.126
11	1061.9	1061.5	0.05	0.675
12	698.4	723.6	-1.57	0.099
13	532.3	534.5	-0.13	0.032
14	530.7	524.3	0.39	0.048
15	366.8	392.2	-1.53	0.029
16	100.0	125.9	-1.62	0.105
17	953.3	942.7	0.69	0.171
18	785.3	768.6	0.89	0.058
19	483.6	465.6	1.08	0.028
20	486.4	462.0	1.47	0.033
21	344.8	332.5	0.75	0.046
22	209.8	201.5	0.52	0.105
23	826.0	813.8	0.75	0.072
24	715.6	696.6	1.17	0.076
25	676.7	664.2	0.76	0.049
26	567.1	546.8	1.24	0.064
27	535.5	524.7	0.65	0.035
28	276.5	275.4	0.07	0.148
29	245.7	261.9	-1.02	0.118
30	742.3	744.7	-0.15	0.071
31	599.5	605.9	-0.39	0.064
32	488.6	492.5	-0.23	0.029
33	342.9	357.7	-0.89	0.033
34	316.7	342.6	-1.71	0.195
35	208.6	229.5	-1.28	0.063
36	623.9	645.6	-1.32	0.054
37	561.2	571.8	-0.64	0.037
38	431.5	431.1	0.02	0.026
39	422.1	446.7	-1.50	0.049
40	276.1	275.0	0.19	0.055
41	795.1	810.5	-0.98	0.131
42	461.6	460.1	0.09	0.034
43	449.7	453.1	-0.20	0.036
44	303.7	309.5	-0.35	0.053
45	683.6	679.3	0.26	0.043
46	530.4	507.1	1.41	0.040
47	355.6	347.6	0.45	0.056
48	591.1	583.3	0.47	0.040
49	555.0	556.0	-0.06	0.049
50	505.4	521.5	-0.80	0.048
51	294.3	288.5	0.35	0.047
52	176.5	166.8	0.64	0.197
53	231.4	200.8	2.02	0.196
MEAN		466.6	0.00	0.094

APPENDIX I

EXAMPLE RESULT OF PARTIAL-LEAST SQUARES ANALYSIS

Table I1 The example result of PLS analysis of all families on the apolar and slightly polar column

Model : UNSC

Version : 0

Prediction date : Nov 15 1989 13:01

Matrix : X data para4.inp 3 datasets samengevoegd

Calibration method : PLS2

Number of factors : 5

Outlier sensitivity : 5

OUTLIER WARNINGS

	Object	Variable	Matrix	Value	Limit
1 :	110	:-------	Hi	0.560	0.522

PREDICTION RESULT

	Object	I-NP	Dev_01	I-SP	Dev_02
1:	110	1048.450	21.990	1110.817	38.421
2:	112	690.197	25.506	754.448	44.565
3:	112a	685.572	32.022	728.305	55.950
4:	113	544.781	8.636	584.136	15.089
5:	113a	539.677	36.468	556.162	63.719
6:	114	393.031	12.704	412.264	22.196
7:	115	264.362	10.993	268.702	19.207

8:	116	148.489	14.898	142.435	26.031
9:	120	944.982	9.066	1031.683	15.841
10:	121	778.356	17.348	863.241	30.311
11:	122	614.023	6.805	678.952	11.891
12:	123	469.180	15.161	515.224	26.490
13:	123a	458.221	10.390	509.969	18.154
14:	124	331.373	9.403	370.420	16.429
15:	125	216.245	10.667	250.380	18.637
16:	130	859.406	14.929	977.969	18.154
17:	130a	823.289	18.634	909.540	32.559
18:	131	693.875	5.443	798.950	9.511
19:	131a	663.877	6.563	750.552	11.467
20:	132	541.998	8.364	633.755	14.614
21:	132a	551.471	11.864	641.784	20.729
22:	132b	513.169	9.877	583.971	17.258
23:	133a	382.488	11.316	439.529	19.772
24:	134	282.959	8.744	356.333	15.278
25:	134a	267.978	9.709	326.486	16.964
26:	140	738.503	5.940	854.068	10.378
27:	140a	669.233	35.513	726.475	62.049
28:	141	588.469	8.388	697.044	14.656
29:	141b	513.254	14.713	571.650	25.706
30:	142	452.792	9.816	549.764	17.150
31:	142b	364.031	14.948	404.611	26.118
32:	143	340.492	9.267	441.434	16.192
33:	143a	239.963	15.631	269.367	27.311
34:	150	625.581	9.191	740.539	16.058
35:	150a	586.586	16.821	672.170	29.389
36:	151a	431.574	13.263	505.290	23.174
37:	152a	297.035	13.942	358.553	24.360
38:	160	446.878	18.781	513.538	32.815
39:	161	285.636	20.587	330.294	35.970
40:	10	678.212	37.651	713.553	65.786

41:	11	501.501	15.477	533.028	27.042
42:	12	329.675	17.696	336.871	30.920
43:	13	182.279	16.264	169.467	28.418
44:	14	52.974	15.097	14.263	26.378
45:	20	599.408	15.923	672.780	27.822
46:	21	421.383	16.551	471.489	28.918
47:	22	278.047	14.701	309.758	25.686
48:	23	151.485	16.615	173.495	29.029
49:	30	504.632	18.585	591.221	32.472
50:	32	218.169	16.681	275.388	29.146
51:	40	332.050	12.367	374.186	21.608
52:	41	177.008	11.046	200.783	19.300
53:	1110	808.317	30.015	880.069	52.443
54:	1112cis	452.162	16.444	491.267	28.731
55:	1112trs	453.893	16.564	494.070	28.942
56:	1112a	446.020	12.869	474.343	22.485
57:	1113	295.057	11.921	310.981	20.829
58:	1120	682.723	20.694	761.053	36.157
59:	1121	501.562	15.443	559.220	26.983
60:	1122	332.795	9.538	362.184	16.665
61:	1123	195.821	9.377	212.409	16.384
62:	1130cis	574.091	17.297	664.232	30.222
63:	1130trs	551.591	15.707	627.801	27.444
64:	1130a	525.531	9.428	578.691	16.473
65:	1131a	335.455	6.564	356.817	11.469
66:	1132	277.385	11.741	337.427	20.514
67:	1132a	167.301	8.214	155.959	14.351
68:	1140	378.339	7.224	414.509	12.623
69:	1141	201.532	8.847	205.128	15.458

AUTOBIOGRAPHY

Promporn Wattanapitakul was born on January 9, 1967 in Bangkok, Thailand. She graduated Mathayom 6 from Triam Udom Suksa school in 1984. She received her Bachelors Degree in Chemical Engineering from Chulalongkorn University, Thailand, in May 1988 and continued her Master's study at Chulalongkorn University in the same year. She was granted the degree in 1990.