

An Integrated Spectroscopic Approach to the Chemical
Characterization of Pyrolysis Oils

Barbara L. Hoesterey, Willem Windig, Henk L.C. Meuzelaar (Biomaterials Profiling Center); Edward M. Eyring, David M. Grant (Chemistry Department); and Ronald J. Pugmire (Fuels Engineering Department)

University of Utah
Salt Lake City, UT 84108

ABSTRACT

The hydrocarbon ("oil") fraction of a coal pyrolysis tar prepared by open column liquid chromatography (LC) was separated into 16 subfractions by a second LC run. The first 13 of these fractions were chosen for integrated spectroscopic analysis. Low voltage mass spectrometry (MS), infrared spectroscopy (IR), and proton (PMR) as well as carbon-13 nuclear magnetic resonance spectrometry (CMR) were performed on the 13 fractions. Computerized multivariate analysis procedures such as factor or discriminant analysis followed by canonical correlation techniques were used to extract the overlapping information from the analytical data. Subsequent evaluation of the integrated analytical data revealed chemical information which could not have been obtained readily from the individual spectroscopic techniques. The approach described is generally applicable to multisource analytical data on pyrolysis oils and other complex mixtures.

INTRODUCTION

Due to the extremely complex nature of pyrolysis tars obtained from recent or fossil biomass samples structural and compositional analysis of such tars poses a formidable challenge to analytical chemists. Even when armed with an arsenal of sophisticated analytical techniques a detailed qualitative analysis requires careful, laborious combination and integration of voluminous chromatographic and spectroscopic data. True quantitative analysis is generally not within reach of current analytical methodologies, especially if the tar contains nonvolatile and/or highly polar or reactive components. Although in recent years impressive advances have been made with the physical coupling of two or more chromatographic and/or spectroscopic techniques into so-called "hyphenated" methods, e.g., GC/MS, LC/MS, GC/FTIR, MS/MS, etc., true integration of the analytical data by means of multivariate analysis methods such as canonical correlation analysis is rarely ever attempted. Yet, intuitively the potential advantages and benefits of data integration methods are easily understood. With these considerations in mind the authors carried out the present feasibility study of a coal-derived pyrolytic tar using a combination of chromatographic (LC), spectroscopic (MS, IR, PMR, CMR) and chemometrics (factor, discriminant and canonical variate analysis) techniques. In order to reduce the complexity of the analytical problem to more manageable proportions, a completely distillable coal tar was selected. Moreover, polar and/or highly reactive components were removed by open column LC. Preliminary results of this integrated analytical approach will be presented here.

EXPERIMENTAL

A pyrolysis tar from a high volatile B bituminous Hiawatha seam coal (Wasatch Plateau field, Utah) was obtained from the fixed bed Wellman Galusha gasifier operated by Black, Sivalls and Bryson in Minneapolis. Open column liquid chromatography (LC) on silica gel using four solvents and solvent mixtures of increasing polarity; i.e. hexane; hexane/benzene 8/1; benzene/ether 4/1; and benzene/methanol 1/1, was used to separate the whole tar into broad compound classes as described by McClennen *et al.* (1). The hexane and hexane/benzene eluted fractions constituted

complex mixtures, principally composed of hydrocarbons. These fractions were combined and further separated by a second IC run. Fractions were eluted from the column with a nonlinear gradient beginning with 100% n-hexane and stabilizing at 10% benzene/90% hexane over a period of 30 min. Sixteen fractions were collected and weighed over a total of 40 minutes. Approximately 1 ml was taken from each sample for low voltage MS analysis. The remaining subfractions were then rotary evaporated and weights of residue were recorded. The calculated elution volumes are shown in Figure 1.

Low voltage mass spectra were run on an Extranuclear 5000-1 quadrupole mass spectrometer with Curie-point heating inlet. Low voltage mass spectra of subfractions 1-15 were obtained using 1/4 ul glass capillary probe tips as described by McClennen et al. [2]. Electron energy was set at 12 eV. Samples were scanned from m/z 20 to m/z 300. The inlet was heated to 200°C. Mass spectra were stored on an IBM 9000 computer and printed out in the form of bar plots. Examples of low voltage mass spectra are shown in Figure 2.

FTIR spectra were obtained using neat samples on NaCl (salt) disks. The instrument was a Nicolet 7000 series spectrometer, resolution 4 cm^{-1} , 200 scans, operated in the absorbance mode. Samples were scanned from 4000-600 cm^{-1} . Absorbance intensities were recorded for 20-30 peaks in each spectrum. In this way 33 wavenumber variables were obtained. Examples of FTIR spectra are shown in Figure 3.

Proton NMR spectra of the hydrocarbon subfractions dissolved in CD_2Cl_2 (with TMS) were taken using a Varian 300 superconducting instrument over the 1-10 ppm region. Integrated peak intensities for eight regions of the spectrum were tabulated for each subfraction, in addition to a table containing an overall view of the number of aliphatic, aromatic and olefinic protons present.

Carbon 13 NMR spectra of subfractions were also run in CD_2Cl_2 on the Varian SC 300 from 0-180 ppm. Peak intensities were measured using integration curves. Twenty three variables were chosen. Table 1 shows overall data from FMR and CMR.

Computerized multivariate analysis was carried out using the interactive SIGMA program package developed at the University of Utah Biomaterials Profiling Center which affords scaling, as well as factor, discriminant and canonical correlation analysis (3). Chemical components were numerically extracted using the Variance Diagram technique described by Windig et al. (4).

RESULTS AND DISCUSSION

The emphasis of this paper is on the general method of multisource data integration using Factor Analysis and Canonical Correlation Analysis. Figure 4 shows the variances calculated for the factors in each data set. For mass spectral and IR data, only eigenvalues greater than 1.0 are shown, whereas all factors were used for the FMR and CMR data. The dashed line shows eigenvalues <1.0, e.g. in the FMR, only Factor 1 had an eigenvalue > 1.0. Six factors from each data set were used for the canonical correlation analyses. Figure 3b shows the percent variance from the original factors that was represented in the subspace spanned by the canonical variate functions. Between 40% (MS) and 80% (FMR) of the original variance is represented by Canonical Variate functions 1 + 2.

Our discussion of the factor analyses presented in Figure 5 will first identify components characteristic of early eluting samples and then move on to later eluting samples. Investigation of the correlated mass peaks loading on factors 1 and 2 (Figure 5a) by means of the variance diagram method revealed 8 components. In Figure 5a component (a) (130°) represents the ion series $\text{C}_n\text{H}_{2n-1}^+$, whereas component (b) (160°) shows $\text{C}_n\text{H}_{2n}^+$ ions from monocyclics or alkenes. A large component (c) ($190-240^\circ$) contains $\text{C}_n\text{H}_{2n-1}^+$ ions (190°), $\text{C}_n\text{H}_{2n-2}^+$ ion (220°) as well as fragment ions at m/z 149, 163, 177, and 191 characteristic of terpenoid resins or other $\text{C}_n\text{H}_{2n-4}$ compounds (240°).

Aromatic compounds, such as short (C_n , $n = 1, 2, 3$) alkyl substituted benzenes occur at component (d) (280°), with longer chain (C_n , $n = 5, 6, 7$) benzenes + tetralins at 320° ; component (e). Component (f) at 0° is thought to represent $\text{C}_n\text{H}_{2n-10}^+$ series. Naphthalenes are found at component (g) (30°) and acenaphthene/biphenyl ions are present at component (h) (50° , Figure 5a). Note that

the scores in this factor space roughly describe a circle, with the exception of fraction 13, which is found near sample 10. Factor 3 (not shown) distinguishes sample 13 from the others with a component axis containing anthracene/phenanthrene moieties as well as an ion series at m/z 180, 194, 208.

Factor analysis was performed on the IR spectra of subfractions 1 to 13 using all 33 wavenumber variables. Figure 5b shows the factor score plot of the IR data on subfraction 1 to 13 in the F1 vs. F2 factor space. Samples 1-7 are very close together, implying that infrared spectroscopy does not detect much difference between these dominantly aliphatic mixtures. Analysis of the underlying correlation between variables by means of the variance diagram method showed component (a) 350° represents methyl and methylene absorptions such as 2870, 2850, 2920, 1460 and 720cm^{-1} . Component axes (b) 120° with peak 1516cm^{-1} and (c) 160° with 3050, 3015 and 1600 represented aromatic absorptions. A component axis (d) 240° , which points to subfractions 12 and 13, represents peaks 750, 2940, 1030 and 1180cm^{-1} .

Initially, we tentatively assigned 1030 and 1180cm^{-1} as C-O stretches, but further examination of infrared spectra of aromatic standards showed that these are probably CH in plane bends, e.g., 1030cm^{-1} (benzene). An interesting feature of the IR data is the peak at 2868cm^{-1} which correlates with the F1 aromatic axis, although it is believed to represent a combination of methyl and methylene stretches. Painter et al., (4) also found this behavior in IR spectra of coal macerals. The data strongly suggest a reinterpretation of this peak assignment.

Several peaks in the F1+ direction of Figure 5b can be assigned as olefin CH out-of-plane bends. These turned out to be important in the combined (canonical variate) space and will be discussed later.

The factor score plot F1 vs. F2 (91% variance) of the PMR data in Figure 5c shows a general distribution of samples forming a circle. The variance diagram of F1 vs. F2 from the proton NMR data shows that the positive F1 axis contains methyl and methylene groups attached to aliphatic (sp^3 hybridized) carbon groups, and olefinic protons. The F1 axis contains aromatic protons, split into two groups. The component axis at 200° represents methyl substituted benzenes ($\text{CH}_3 + 1$ ring aromatic), oriented toward fractions 9 and 10. The 170° rotation contains 2-ring and 3-ring aromatics and longer chain aromatic substituents (CH_2) oriented toward fractions 11-13.

Factor analysis of the CMR data gave 6 factors with eigenvalues >1.0 . The score plot of F1 vs. F2 (56% of the variance) (Figure 5d) shows that samples 1-7 appear to be similar in this dimension oriented along the negative side of F1. Components in this direction include aliphatic peaks such as at 23, 30, 32 and 38 ppm and (with weaker loadings), at 97 and 114 ppm, probably olefinic carbons. Fraction 8 is somewhat removed from fractions 1-7 but still on F1-, and therefore predominantly aliphatic in character. Fractions 9-13 are widely spread on factor F1+. A component axis at 350° represents peaks at 20, 122, 126, 131 and 135 ppm. Fractions 10 and 11 have an associated component which includes the peaks at 40 and 134 ppm. Fractions 12 and 13 have an associated component axis with the peaks at 127, 129, 132 and 142 ppm. All peaks on F1+ (except at 20 and 40 ppm) are likely aromatic carbons. The 20 and 40 ppm peaks are sp^3 hybridized carbon substituents on aromatic rings.

Canonical correlation of the factor analysis results from the MS, IR, PMR and CMR data using 6 factors from each data set gave four canonical variate functions with correlation coefficient greater than 0.90. The variance associated with the four CV functions is shown in Figure 4. Figure 6 is a score plot of CV1 vs. CV2 for the four data sets. The scores from each fraction analyzed by the four methods are connected by lines. A small polyhedron implies that the methods see the sample in a similar way. The later eluting samples (9-13) appear to group into clusters that are widely separated from one another (e.g., 9 and 10, 11 and 12, 13) whereas early eluting samples (1-7) are close together in this space. Figure 7 shows a consensus picture of the component directions from each method found in this CV space. Correlated variables consistent with an interpretation of aliphatic compounds are clustered around CV1-, near fractions 1-4. Fraction 7 appears also in this direction. A component axis of alkyl benzenes (m/z 92, 106...) from the MS data loads weakly in this CV space. From the original factor analyses it can be seen

that the mass spectral and PMR data clearly differentiated fraction 7 from the other fractions, but that other data sets grouped 7 with fractions 1-6. The PMR data showed no unique component associated with fraction 7. This says that the mass spectral picture of fraction 7 is in a sense unique, and does not appear in the CV space. A component axis corresponding to olefinic variables (IR, PMR, CMR) appears at 150°, in the direction of fractions 5 and 6. The mass spectral data shows ion series with 2 and 3 units of unsaturation, one or more of these apparently being a double bond. The positive half of CVI reveals three components, each one consistent with an assignment of aromaticity. The PMR and IR (CH in-plane bend modes) show increasing fused ring aromaticity in the ccw direction (300° to 50°). The mass spectral data identified the component at 300° with indane/tetralin, the 0° component as >C₃ alkyl substituted naphthalenes and the 50° component as acenaphthene/biphenyls. A component axis at 80° (mass spectral data only) showed peaks characteristic of alkyl anthracene/phenanthrenes. The CMR data has not been interpreted in as great a detail, but groups of aromatic peaks found in these three directions are not inconsistent with interpretations from the other data sets. A point of interest is provided by the two CMR aliphatic peaks, 20 and 40 ppm (0°), which correlate with aromatic carbons and are from alkyl substituents.

CONCLUSIONS

Valuable information was gained by correlating the four analytical techniques. For example, mass spectral peaks of samples containing 2 and 3 units of unsaturation, as determined by PMR, were shown to contain double bonds, whereas mass spectral peaks corresponding to molecules containing one unit of unsaturation were found to be cyclic. Since all the techniques showed the predominantly aromatic fractions to be very different from each other when moving to higher fused ring systems, a better understanding of spectral interpretation of aromatic hydrocarbons mixture data appears possible.

A major limitation of the present study is that only that portion of the analytical data common to all four analytical techniques was interpreted. Future studies will have to address the highly important portions of the analytical data unique for each analytical method in order to reap the full benefits of the integrated spectroscopic approach.

ACKNOWLEDGMENT

The expert assistance of Dr. Jim Liang, Mary West and Alireza Pezeshk in obtaining the analytical data, and of Dr. Liang and Dr. Charles Mayne in interpreting the results is gratefully acknowledged. Drs. George R. Hill and Larry L. Anderson are thanked for their stimulating support of this research. The work reported here is part of the research carried out at the University of Utah, under the auspices of the Consortium for Fossil Fuel Liquefaction Science funded under DOE contract RF-4-21033-86-24.

REFERENCES

1. McClellan, W.H.; Meuzelaar, H.L.C.; Metcalf, G.S. and Hill, G.S., Fuel, 62 (1983).
2. Meuzelaar, H.L.C.; Haverkamp, J.; Hileman, F.D., Tandem Mass Spectrometric Analysis (MS/MS) of Jet Fuels, Part I and II, Air Force Aero Propulsion Laboratory, Wright-Patterson Air Force Base, Ohio, AFWAL-TR-85-2047, Parts I and II, 1985.
3. Windig, W. and Meuzelaar, H.L.C., SIGMA, System for Interactive Graphics-oriented Multivariate Analysis, Proc. 34rd ASMS Conf. on Mass Spec. All. Topics, Cincinnati, 1986, pp. 64-65.
4. Windig, W. and Meuzelaar, H.L.C., Anal. Chem., 56, (1984) 2297-2303.
5. Kuehn, D.W.; Davis, A. and Painter, P.C., ACS Symp. Series, 252, Ch. 7, 100-119, 1984.

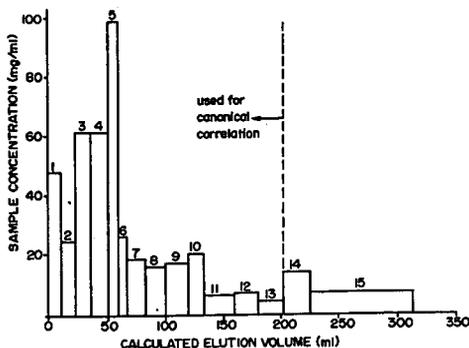


Figure 1. Reconstructed liquid "chromatogram".

TABLE I
INTEGRATED INTENSITIES OF ALIPHATIC OLEFINIC AND AROMATIC
REGIONS OF THE NMR SPECTRUM FOR SUBFRACTIONS OF THE
HYDROCARBON FRACTION OF BIAXINIA TAR

SUBFRACT.	(Proton NMR Data)			ALIPHATIC H- AROMATIC H	(Carbon 13 NMR Data)				
	ALIPHATIC (1-5 ppm)	OLEFINIC (4-6 ppm)	AROMATIC (6-9 ppm)		SUBFRACT.	ALIPHATIC	OLEFINIC	AROMATIC	ALIPHATIC C- AROMATIC C
1	1.0				1	1.0			
2	1.0				2	1.0			
3	0.959	0.041			3	0.98	.02		
4	0.948	0.052			4	0.922	.078		
5	0.952	0.038			5	0.922	.078		
6	0.931	0.049			6	0.885	.063		
7	0.905	0.043	0.032	17.40	7	0.759	.121	.052	
8	0.909	0.091	0.091	9.99	8	0.596		.404	1.48
9	0.827	0.173	0.173	4.78	9	0.513		.487	1.05
10	0.822	0.178	0.178	4.62	10	0.337		.663	0.56
11	0.789	0.211	0.211	3.74	11	0.333		.667	0.50
12	0.786	0.214	0.214	3.87	12	0.261		.739	0.33
13	0.849	0.151	0.151	5.62	13	0.211		.789	0.27

* integration from 50-150 ppm, olefinic and aromatic arbitrarily made equal.

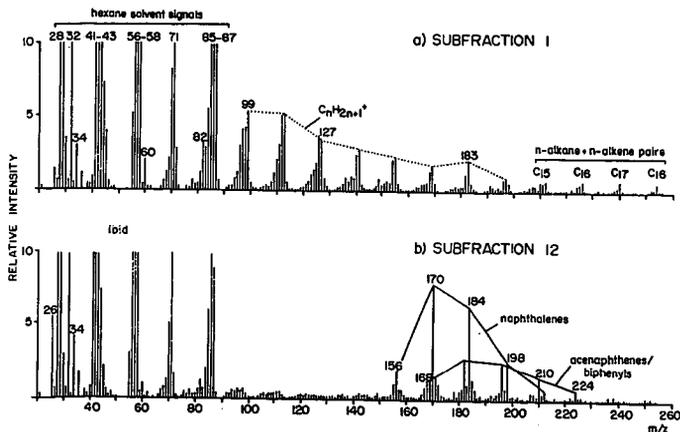


Figure 2. Low voltage mass spectra of (a) subfraction 1 and (b) subfraction 12.

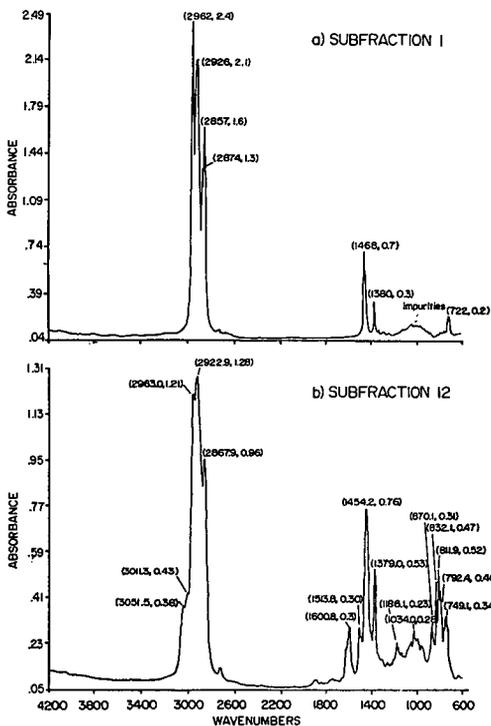


Figure 3. FTIR spectra of (a) subfraction 1 and (b) subfraction 12.

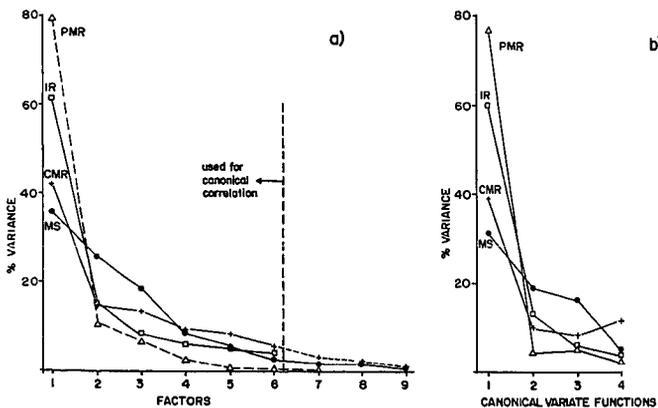


Figure 4. Percent total variance explained by (a) factors and (b) canonical variate functions.

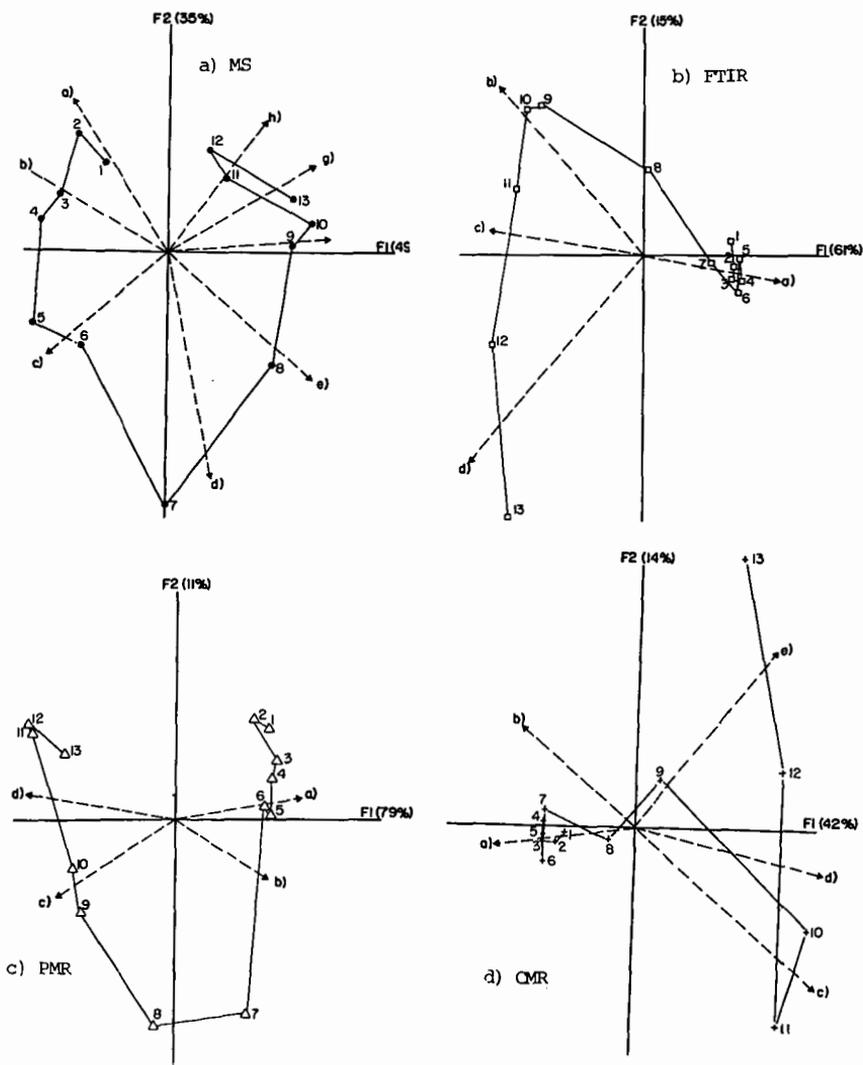


Figure 5. Factor score plots in FI/FII spaces of all four individual data sets.

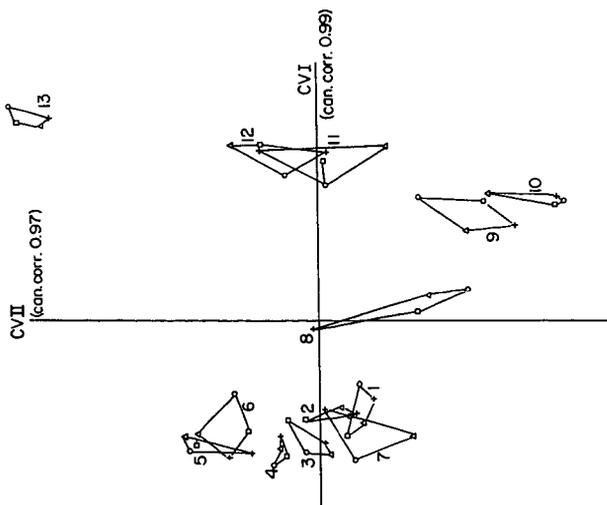


Figure 6. Combined score plots of integrated spectroscopic data in "common" CV1/CV2 space.

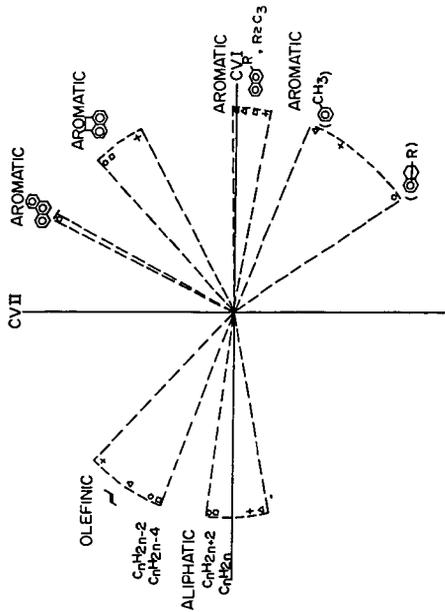


Figure 7. Combined loading plot of integrated spectroscopic data in CV1/CV2 space showing common chemical components.