APPLICATION OF TRACE ANALYSIS AND CHEMOMETRICS TO ENVIRONMENTAL PROBLEMS

KENNETH JAMES CALDICOTT

A THESIS SUBMITTED TO THE COUNCIL FOR NATIONAL ACADEMIC AWARDS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

November 1991

The Polytechnic of Wales in collaboration with the

Water Research Centre

DECLARATION

This thesis has not been, nor is currently, submitted for the award of any other degree or similar qualification.

K.J.Caldicott

APPLICATIONS OF TRACE ANALYSIS AND CHEMOMETRICS TO ENVIRONMENTAL PROBLEMS

K.J.CALDICOTT

ABSTRACT

The original aim of this investigation was to determine the factors affecting the treatability of coloured waters found in peaty areas. This involved the extraction of the humic acid thought to be causing the problem, followed by the analysis of the products of controlled degradation. It was found that extraction of the humic acid caused some changes to occur in its structure and that degradation, however controlled, was not reproducible. It was finally discovered that the treatability of the coloured waters was controlled by the treatment process rather than by the water itself.

The project continued with an investigation into polyaromatic hydrocarbons in a South Gwent Valley. A procedure was determined which involved extraction, clean-up and HPLC analysis. The analytical system was validated and applied to a series of soil samples. The resulting data were subjected to a series of chemometric techniques.

Factor analysis of the data set indicated that there were two major profiles present. This, and cluster analysis, showed that the Afon Llwyd is a major source of polycyclic hydrocarbons along its flood plain. The analysis of sub-groups of the data set showed that on ground not affected by the river only one source was describing the data. The sub-group containing just samples on the flood plain suggested the possibility of a third factor. This group was in the vicinity of a toxic waste incinerator.

The final section attempted to investigate nitrated polyaromatic hydrocarbons in the same area. Whilst an extremely sensitive analytical procedure was developed it was found that the sub parts per billion levels of the compounds of interest were swamped by co-extracted components.

ACKNOWLEDGEMENTS

I would like to thank Dr.Ken Burton, my director of studies, for all the advice and encouragement that he has given me throughout the period of this work. Thanks are also due to my other supervisor Dr.Alex Waggott. Other academic thanks go to Dr.Graham Nickless who rekindled the spark.

I also thank Drs.Terry Goodburn and John Garner who kindly allowed me the "out of hours" use of the Analytical Laboratory at Parke-Davis Research.

I must also mention my wife, Shirley, and my sons, Andrew and Alex, who endured my absences.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

MULTI-VARIATE STATISTICS

1.1 Literature survey

1.1.1 Introduction

Multi-variate statistics have long been used by psychologists and social scientists. During the 1960's clinical chemists began to investigate their use.

In 1971 Kowalski and Bender (1) published a description of pattern recognition showing how an obscure property of a collection of objects can be detected and / or predicted using indirect measurements made on the objects.

Also in 1971 the Swedish chemist, S.Wold, first coined the word "chemometrics" when submitting a grant application. Chemometrics has been defined (2) as the chemical discipline that uses mathematical, statistical and other methods employing formal logic:-

a) to design or select optimal measurement procedures and experiments

b) to provide maximum relevant chemical information by analyzing chemical data.

A good example of experimental design is simplex optimization of chromatographic separations. This has been well developed by Berridge (3) who defined a "Chromatographic Response Function", which is a numerical way to define a chromatographic separation, in order to allow a mathematical procedure to be used.

Of particular interest is the use of multi-variate statistics in pattern recognition and principal component and factor analysis. The use of micro-computers in recent years has resulted in a much wider application of these techniques than was possible previously. Traditionally, analytical decisions have been made on the basis of individual parameters e.g a sample may be rejected if a single result is outside of some preset specification. Modern techniques, however, may produce not one but a series of values. The multi-variate approach may be used to sort data into groups of like members. In pattern recognition there are two fundamental sub-divisions which, for historic reasons, are known as unsupervised and supervised learning.

In the former group, no assumptions are made about the number of classes to be expected. The raw data are examined and the use of the procedure results in a classification. The experimentalist must then use his judgement about its validity.

In supervised learning systems the knowledge of a training set of data is used to attempt to classify unknowns. In effect the boundary conditions of a class are sought. An unknown can be classified if its data falls within these boundaries.

1.1.2 Unsupervised Learning Methods

Probably the most frequently used of the unsupervised learning methods is cluster analysis. Massart et al (4) used a clustering technique to categorize iron meteorites by their trace element patterns.

Hopke (5) also used cluster analysis to interpret a data set of ninety samples, each analyzed for eighteen elements, taken from the Boston city area. The clustering technique yielded two distinct clusters. One contained sites in the inner city area and the other from the more suburban areas. Within the two main clusters were several smaller ones which represented the influence of local sources on the data.

In another publication (6) he used the same technique to investigate the chemical and physical parameters describing the nature of grab samples taken from a lake in New York state. Fifteen elements were determined from each sample and cluster analysis indicated four distinct groups. One group consisted of sites near the centre of the lake and the other three were sites near the edge, Of the three shore-line clusters one was found to be mainly sandy material, one was finely deposited material from waterways entering the lake and the third was of an intermediate nature.

Brill et al (12) used both supervised and unsupervised methods based on the ARTHUR chemometrics package. They used the hydrocarbon profiles of fire-ants to differentiate between colonies. The LLM (linear learning machine) approach gave an average efficiency of 99%, MULTI (multi-category classification test) was 91% efficient. The KNN (K nearest neighbour) procedure was 88% efficient and the SIMCA approach 81%. All of these were considered to be extremely good as the same breed of ants was being studied and the only variation was in the position of the colony.

In 1986 Thrane and Gunderson (17) used "Fuzzy C-Varieties" pattern recognition to cluster the data produced on Polyaromatic Hydrocarbon (PAH) contamination as measured by a series of air monitors. The fuzzy c-varieties algorithm is a form of unsupervised clustering technique which allows partial membership of several clusters. They were able to correlate the PAH data into two clusters, one arising from traffic and the other from wood burning.

1.1.3 Supervised Learning Methods

One of the most commonly reported supervised learning methods is "Soft Independent Modelling of Class Analogy" (SIMCA). This was developed by Wold and is available as a PC program.

Kvalheim et al (7) have used the technique to distinguish between pristine and polluted mussels on the basis of capillary GC profiles of naturally occurring components of muscle and gonad tissue.

Jellum et al (8) examined 105 peaks from GC analysis of 16 brain biopsies. They hoped to use SIMCA to distinguish between three classes; normal, brain tumour present and hypofysis tumour present. No test set was used and the classification was checked using the test set as the training set. Fourteen of the sixteen samples were found to be classified correctly. The two mis-classified results were from different types of tumour, thus the distinction between tumour and non-tumour was 100% correct.

SIMCA has also been used to investigate the data produced by Mecke and Noack (9). A small data set was prepared consisting of seven variables from the infra-red and ultra-violet spectra measured on sixteen α,β-unsaturated carbonyl compounds. The data set consisted of six objects with known trans-configuration and three with known cis-configuration. SIMCA was used to determine whether the seven remaining objects belonged to the trans- or cis- classes or whether they belonged to a class of their own.

Onuska et al (16) used a principal component multivariable statistical method based on the SIMCA-3B algorithm to investigate the capillary GC analysis of PCB's. They were able to provide a means of classifying and identifying PCB homologue mixtures in environmental samples.

Duewar et al (10) used SIMCA to classify oil spillages, according to their source, on the basis of trace element profile. They found that, although SIMCA was not completely successful, it compared well with other techniques.

Cohen et al (11) used a supervised pattern recognition technique which permits the fitting of non-linear surfaces to each category using non-statistical methods. They used Cohen orthogonal polynomials to investigate HPLC data on the analysis of acids present in ascitic fluid, a substance which accumulates in the abdomen of patients with liver disorders. They were able to predict which patients had developed spontaneous bacterial peritonitis.

1.1.4 Principal Components / Factor Analysis

The most important use of principal components and factor analysis is to represent n-dimensional data in a smaller number of dimensions, usually two or three.

In 1978 Alpert and Hopke (13) tested the theory by investigating synthetic data sets. They prepared a series of twenty samples containing known masses of five standard materials, each having known elemental composition. Neutron activation analysis was used to determine the concentration of thirty seven elements in each sample. Factor analysis was applied to the data and correctly indicated the number of factors as five. The known elemental compositions of the five standards were used as test vectors for target testing. Results were good for elements in high concentrations but poor for trace components.

Factor analysis has also been applied to the interpretation of environmental data. Hopke (14) used the technique to aid in the interpretation of thirty two physical and chemical measurements made on seventy nine sediment samples from Lake Chautaugua in New York state. He was able to find five factors associated with the data. Analysis of these factors and their loadings led to the prediction of the processes acting upon the sediment to produce the data.

Smeyers-Vebeke et al (15) took air samples weekly for three years from a measuring station in the Netherlands and analyzed them by capillary GC. The analyses yielded nearly three hundred components, thirty five of which were found in

every sample They showed that wind direction was a major contributing factor to the type of pollution found as the sources of pollution were well spread and could be identified by the nature of the compounds present.

Gaarenstroom et al (18) measured the concentrations of twenty four chemical species present in solids collected from eleven points in Tucson, Arizona at various times during a one year period. Factor analysis was used to investigate species-time data matrices for each location. Between five and eight factors were found to be sufficient to reproduce the data for the various locations. The pollution sources were found to be soil, non-local aerosol and automotive exhaust.

Love et al (19) used factor analysis to investigate a matrix of fluorescence lifetime data as a function of solvents and solutes. Two factors were found to be associated with the data. Two properties of the solute, identity and number of carbon atoms attached to the exocyclic nitrogen, were found to reproduce the experimental data within experimental error. No combination of solvent properties were found which successfully reproduced the data set.

Woodruff (20) used the technique to determine the number of components under a chromatographic peak. They analyzed the MS data produced from GC-MS. In the absence of noise they found that the limit of detection of an impurity under the main component was ca. 1% but when noise was added the limit of detection increased to 3-10%.

1.1.5 Optimization of Chromatography

Many workers have been investigating procedures for the optimization of chromatographic separations since the inception of the technique. The application of micro-computers has aided the process and this has been described recently by Glajch and Snyder (56).

There can be various goals aimed at during the optimization process and the result required for the particular separation should be set before any work is begun. Typical aims are:-

a) To achieve separation with minimum cost

b) To achieve separation in the shortest possible time

c) To produce a system which gives baseline resolution in a given time.

Snyder (57) has described the solvent selectivity triangle. This is based on the approach that solvent selectivity is controlled by three main effects: the ability of the organic solvent to interact with the solute as a proton donor (basic), as a proton acceptor (acidic), or as a dipole. For reverse-phase HPLC water is the diluent solvent. This effectively precludes the use of many, potentially useful solvents due to their lack of miscibility. The best remaining choices are methanol (proton acceptor), acetonitrile (proton donor) and tetrahydrofuran (dipole).

Each of the organic modifiers, mixed with the correct amount of water to give similar retention times for the solutes of interest, can be considered as a "mixture variable" that can

influence selectivity. As the sum of the modifiers must total 100% there are only two independent variables among the three mixture variables. The mixture design is, therefore, a triangle.

The Window Diagram Technique was developed by Purnell and Laub (58). Originally this was applied to gas-solid chromatography but has since been expanded to fit liquid chromatography. It is based on the empirical, linear relationship between log (capacity factor) and mobile phase concentration for certain organic modifiers. It is especially useful as it offers a means of predicting the region of global optimization. Chromatographic systems can often produce multiple optima due to changes in elution order of components. Window techniques have the advantage of graphically displaying all the possible separations.

Berridge (59) has reviewed the simplex optimization of high performance liquid chromatography. The sequential simplex method was first proposed by Spendley et al (60) on 1962. A simplex is a geometric figure described by a number of points equal to one more than the number of variables being investigated. Thus a simplex with two variables is a triangle and that for three variables is a tetrahedron. As the number of variables is increased then the simplex is described in a higher number of dimensions. This cannot be visualised easily but the mathematics of the procedure do not increase greatly in complexity and can easily be handled by a computer.

The object of a sequential simplex is to force the

simplexes to move away from regions of poor response toward the area of optimum response. To optimize two variables a single experiment is performed at each of three points to determine the response at each position. The next and subsequent steps are made according to a set of rules:-

1) A move away from the worst response is made after each experiment.

2) A move is made into the adjacent simplex which is obtained by rejecting the point with the worst response and replacing it with a point obtained by a reflection through the mid-point of the remaining face.

3) If the new point has the worst response in the new simplex, do not re-apply rule 2 but instead reject the next-to-worst point in the new simplex.

When the optimum has been located the rules force the simplex to circle.

There are several drawbacks to the basic system and Nedder and Mead (61) made modifications to the procedure to allow for expansion and contraction of the simplex. This allows for the changing nature of the response contour surface.

1.2 Basic Theory

1.2.1 Soft Independent Modelling of Class Analogies (SIMCA)

The basic idea of SIMCA is that multi-variate data observed on a group of similar objects are well described by a simple empirical model. The idea began with BrØnsted (21) and was developed by Wold (22). The method is based on the mathematical description of each class, independent of other classes, by means of factor and principal components analysis. The classes are modelled in the training process as "confidence boxes" based on the component values and the residual standard deviations of the class. Recognition is carried out by determining the class of the box in which the point of the object of unknown class falls. The object is regarded as an outlier if its point is outside the boxes.

The data analysis is made in two stages. First, data from objects of 'known' classes are used to develop a principal components model of each class. Information about the relevance of the variables, outlying objects etc.is also obtained in this first phase. Secondly, the unassigned objects are classified. This is done by fitting each class principal component model to each object data vector. The objects are assigned to a class if the degree of fit between the object data vector and the corresponding class model is sufficiently good.

The mathematical background is based on multi-variate Taylor expansions. Provided that variables $\underline{i}$ are related to the similarity between the objects, then the data X_{ik} can be

approximated by the PC model with A product terms (components).

$$X_{ik}=\alpha_i+\sum_{a=1}^{a=A}\beta_{ia}\theta_{ak}+\epsilon_{ik} \qquad \text{Equation 1}$$

or in matrix notation

$$X = A + B.T + E$$

The parameters α, β and θ describe the systematic part of X. The PC model separates the variables i and the objects k into separate parameters α and β for the former and θ for the latter.

The parameters $_{ak}$ describe the place of the object k in the class. The residuals $_{ik}$ describe the 'random' part of X. These residuals have the standard deviation S_0.

$$S_0^2=\sum_{i=1}^{i=m}\sum_{k=1}^{k=m}\epsilon_{ik}^2/(M-A)(n-A-1) \qquad \text{Equ 2}$$

S_0 measures the typical distance between the class model and an object belonging to the class. The residuals may also be used to detect outliers, anomalies etc in the training set.

The second part of the SIMCA approach is to classify unknown objects into a test-set by their degrees of fit. This is done by fitting the data X_{ip} from each object to each of the

class models by multiple linear regression.

$$|X_{ip}-\alpha_i^q| = \sum_{a=1}^{a=A_q} t_{ap}\beta_{ia}+e_{ip} \qquad \text{equ 3}$$

The degree of fit is measured by residual standard deviation $S_{p(q)}$ with $M - A_q$ degrees of freedom.

$$S_p(q)^2 = \sum_{i=1}^{i=m} e_{ip}^2/(M-A_q) \qquad \text{equ 4}$$

If an object has been part of a class training set, then the standard deviation is too low by a factor of:-

$$\phi^2 = \frac{n_q}{(n_q-A_q-1)} \qquad \text{equ 5}$$

Hence it is corrected by multiplication by ϕ

Object p is now classified as probably belonging to class q if $S_{p(q)}$ is not much larger than the residual standard deviation of the class $S_{o(q)}$. The significance can be determined by means of an approximate F-test.

The above mathematics are available as a personal computer

program "SIMCA-3B" written in BASIC by S.Wold, Emea University.

1.2.2 Factor Analysis

In factor analysis, each element of a two-dimensional data matrix is viewed as a linear sum of factors, each factor having a different weighting.

In general terms, consider a data matrix [D] consisting of r rows and c columns.

$$[D]=\begin{bmatrix} d_{11} & d_{12}.....d_1c \\ d_{21} & d_{22}.....d_2c \\ d_{r1} & d_{r2}.....d_{rc} \end{bmatrix}$$

The objective of factor analysis is to obtain a mathematical "abstract" solution such that each point in [D] is expressed as a linear sum of product terms. The number of terms in the product sum n is called the number of factors. Solutions are sought of the form:-

$$d_{ik}=r_{ij}c_{jk}$$

The r and c terms are called co-factors, the row co-factors being termed scores and the column co-factors known as loadings. For the j^{th} factor in the sum, r_{ij} is associated with the i^{th} row designee of [D] and c_{jk} is associated with the k^{th} column designee.

The data can be decomposed into two matrices:-

$$[\ D\ ] = [\ R\ ]_{abstract}\ [\ C\ ]_{abstract}$$

where

$$[R]_{abstract} = \begin{bmatrix} & r_{11} & r_{12} \ldots\ldots r_{1n} \\ factor & r_{21} & r_{22} \ldots\ldots r_{2n} \\ & r_{r1} & r_{r2} \ldots\ldots r_{rn} \end{bmatrix}$$

$$[C]_{abstract} = \begin{bmatrix} & c_{11} & c_{12} \ldots\ldots c_{1c} \\ column & c_{21} & c_{22} \ldots\ldots c_{2c} \\ & c_{n1} & c_{n2} \ldots\ldots c_{nc} \end{bmatrix}$$

These are abstract matrices as the solution is purely mathematical and does not have any physical meaning. There are several methods available to assist in the determination of n the number of factors associated with the data and these will be discussed in greater detail in section 3.6.2. Thus, this step gives an estimate of the complexity of the data space.

The second objective of factor analysis is to convert the abstract solution into a real one. This is done mathematically by seeking a transformation matrix [T]. Post multiplying

$[R]_{abstract}$ by $[T]$ and pre-multiplying $[C]_{abstract}$ by $[T]^{-1}$, the abstract matrix can be expressed as:-

$$[D] = \{ [R]_{abstract} [T] \} \{ [T]^{-1} [C]_{abstract} \}$$

$$= [R]_{transformed} [C]_{transformed}$$

A simple example of this is described by Malinowski (142). It involves the absorbances of four different mixtures of the same absorbing compounds measured at five wavelengths. Factor analysis will produce an estimate of the number of factors present and an abstract solution. When the correct transformation matrix is found then [R] transformed will have columns which correspond to the absorbances of the pure components, essentially plotting the spectra of each. Similarly, each row of [C] transformed corresponds to the concentrations of one of the components in each of the four mixtures.

A computer program called TARGET, written by E.R.Malinowski originally for the Apple II microcomputer but translated by Hopke for use on an IBM, produces abstract factor analysis results and gives several means of predicting the number of factors. The program can be further used to test suspected parameters individually as possible factors.

1.3 Demonstration of the SIMCA approach

1.3.1 Introduction

Water samples were taken from a series of streams and UV data obtained. The SIMCA system was applied to the multivariate data in order to differentiate between the water sources.

1.3.2 Sampling

Three streams were sampled. Two of the streams were approximately 100 metres apart in the Pontstickle area of South Wales and the third was in Upper Cwmbran, Gwent. Each sample set consisted of a series of samplings from points along the stream.

1.3.3 Analysis

In the laboratory, 5.0 ml of each sample was mixed with 1.0 ml of 0.1M pH 7.5 buffer. The spectrum of each solution was obtained from 500 to 220 nm in 10mm cuvettes against a blank of distilled water (5.0ml) and buffer (1.0ml). The spectra were produced using a Perkin-Elmer Lambda 15 spectrophotometer. Absorbance readings were noted at 20nm intervals and were used to construct the data matrices shown in tables 1.3.1 to 1.3.3. A spreadsheet was used to reconstruct the average spectrum from each set (Fig.1.3.1).

Table 1.3.1

λnm	1	2	3	4	5	6	7	8	9
500	.001	.004	.002	.001	.003	.001	.003	.002	.002
480	.002	.005	.003	.001	.003	.002	.003	.002	.003
460	.002	.006	.003	.002	.003	.002	.004	.003	.004
440	.003	.008	.004	.003	.004	.003	.005	.003	.005
420	.005	.011	.006	.005	.005	.004	.006	.005	.007
400	.007	.015	.008	.007	.007	.007	.009	.007	.010
380	.011	.023	.012	.010	.011	.010	.012	.010	.015
360	.016	.031	.016	.015	.016	.014	.016	.015	.022
340	.022	.042	.023	.021	.023	.021	.022	.021	.030
320	.030	.054	.031	.030	.031	.028	.030	.029	.040
300	.040	.067	.041	.039	.041	.037	.040	.039	.054
280	.052	.088	.055	.051	.054	.049	.052	.051	.073
260	.065	.107	.068	.065	.068	.062	.066	.064	.091
240	.084	.135	.087	.082	.085	.079	.084	.081	.114
220	.150	.238	.154	.144	.156	.137	.152	.140	.178

Spectral Data from Pontstickle #1

Table 1.3.2

λnm	1	2	3	4	5	6	7	8	9
500	.002	.007	.002	.002	.002	.003	.002	.003	.002
480	.003	.009	.003	.003	.003	.004	.003	.004	.003
460	.004	.011	.003	.003	.004	.004	.003	.004	.004
440	.005	.014	.004	.004	.005	.006	.004	.006	.005
420	.007	.019	.006	.006	.007	.008	.006	.008	.007
400	.009	.027	.009	.009	.009	.010	.009	.012	.010
380	.013	.038	.013	.013	.013	.015	.014	.018	.015
360	.021	.055	.019	.019	.019	.021	.020	.025	.022
340	.030	.072	.028	.027	.027	.029	.028	.034	.030
320	.039	.089	.037	.037	.036	.039	.038	.045	.040
300	.053	.109	.050	.049	.048	.052	.050	.060	.054
280	.071	.145	.066	.065	.064	.069	.068	.079	.073
260	.088	.175	.083	.083	.082	.088	.085	.097	.091
240	.111	.221	.105	.104	.102	.110	.107	.122	.114
220	.182	.360	.162	.165	.157	.170	.168	.189	.178

Spectral Data from Pontstickle # 2

Table 1.3.3

λnm	1	2	3	4	5	6	7	8	9	10
500	.005	.004	.003	.004	.005	.005	.007	.003	.008	.005
480	.006	.004	.003	.004	.005	.006	.007	.004	.009	.006
460	.006	.005	.004	.005	.006	.006	.008	.005	.010	.007
440	.008	.006	.005	.006	.007	.007	.009	.006	.012	.008
420	.009	.008	.006	.007	.008	.009	.011	.007	.014	.010
400	.011	.010	.008	.008	.010	.010	.013	.009	.016	.012
380	.014	.013	.011	.011	.012	.014	.017	.012	.020	.014
360	.019	.017	.015	.015	.017	.018	.022	.015	.024	.018
340	.024	.023	.019	.020	.022	.022	.028	.020	.031	.023
320	.029	.027	.024	.025	.026	.027	.034	.026	.038	.029
300	.035	.035	.030	.030	.032	.033	.041	.033	.049	.038
280	.046	.046	.038	.039	.042	.042	.053	.044	.063	.049
260	.058	.057	.047	.049	.052	.052	.065	.053	.075	.059
240	.074	.074	.063	.064	.069	.077	.083	.069	.095	.077
220	.282	.288	.281	.279	.285	.430	.301	.303	.360	.325

Spectral Data from Cwmbran Stream

1.3.4 SIMCA Analysis

Two SIMCA runs were performed on the data. The first compared the results from the two Pontstickle streams. Figure 1.3.2 shows the data space reduced to two. Sample #11, which is sample #2 of series 2, appears to be an outlier. Examination of the raw data shows that this sample has a higher overall absorbance than the others in the set but that the data had not been entered incorrectly. The result was therefore allowed to stand. Outliers such as this can compress the remaining data so that classifications cannot be seen readily. However, as there was no valid reason to reject this result it was felt that it was a more stringent test of the system.

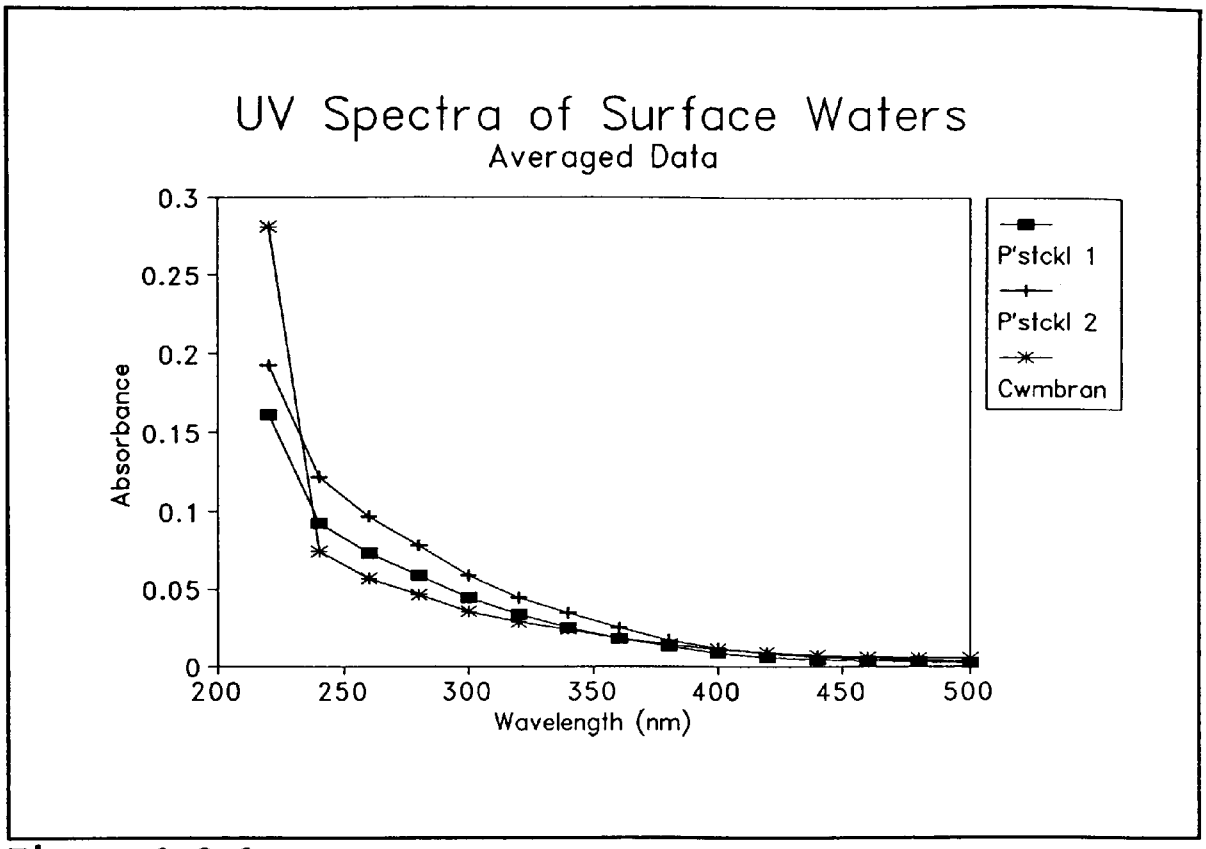
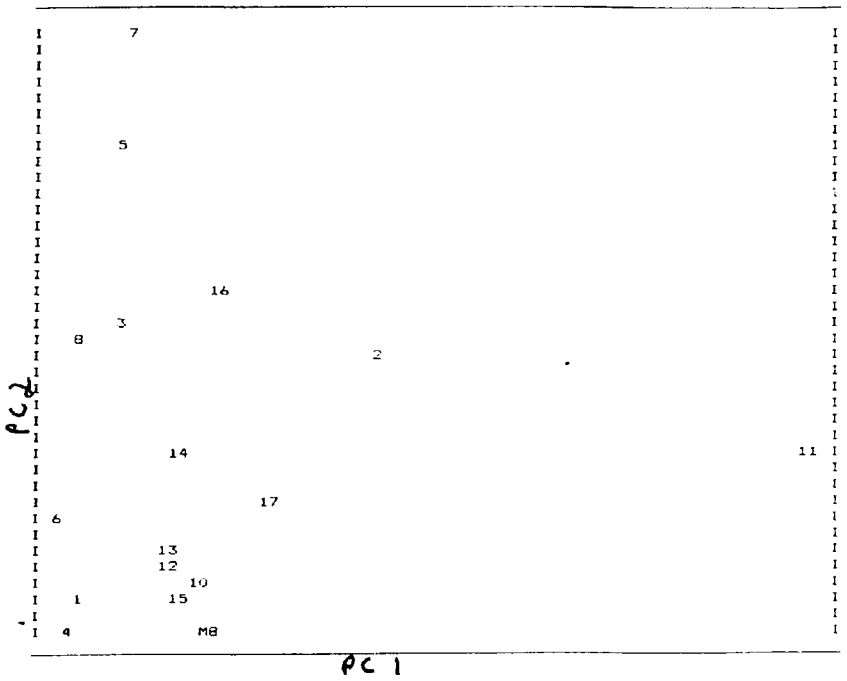


Figure 1.3.1



Pontstickle 1 and 2 Reduced to Factor Space of 2

Figure 1.3.2

The principal components were determined using the SIMCA program:-

Series #1

No.	SD before	CSV/SD	Q	Av.mod.power
1	1.000	0.317	0.913	0.7436
2	0.304	0.756	0.904	0.8455
3	0.179	1.634	0.892	-

Series #2

No.	SD before	CSV/SD	Q	Av.mod.power
1	1.000	0.112	0.913	0.9086
2	0.103	1.200	0.904	-

The component is considered significant if:-

a) CSV/SD < Q

b) CSV/SD > Q but < 1.0 and 4 vars < 0.95

c) CSV/SD > Q but < 1.0 and 3 vars < Q

The objects were fitted to their classes. In this first set the two series are very similar. This was to be expected from the geographical proximity of the two streams investigated. However, the SIMCA approach has yielded a definite clustering of the data points as seen in Figure 1.3.3

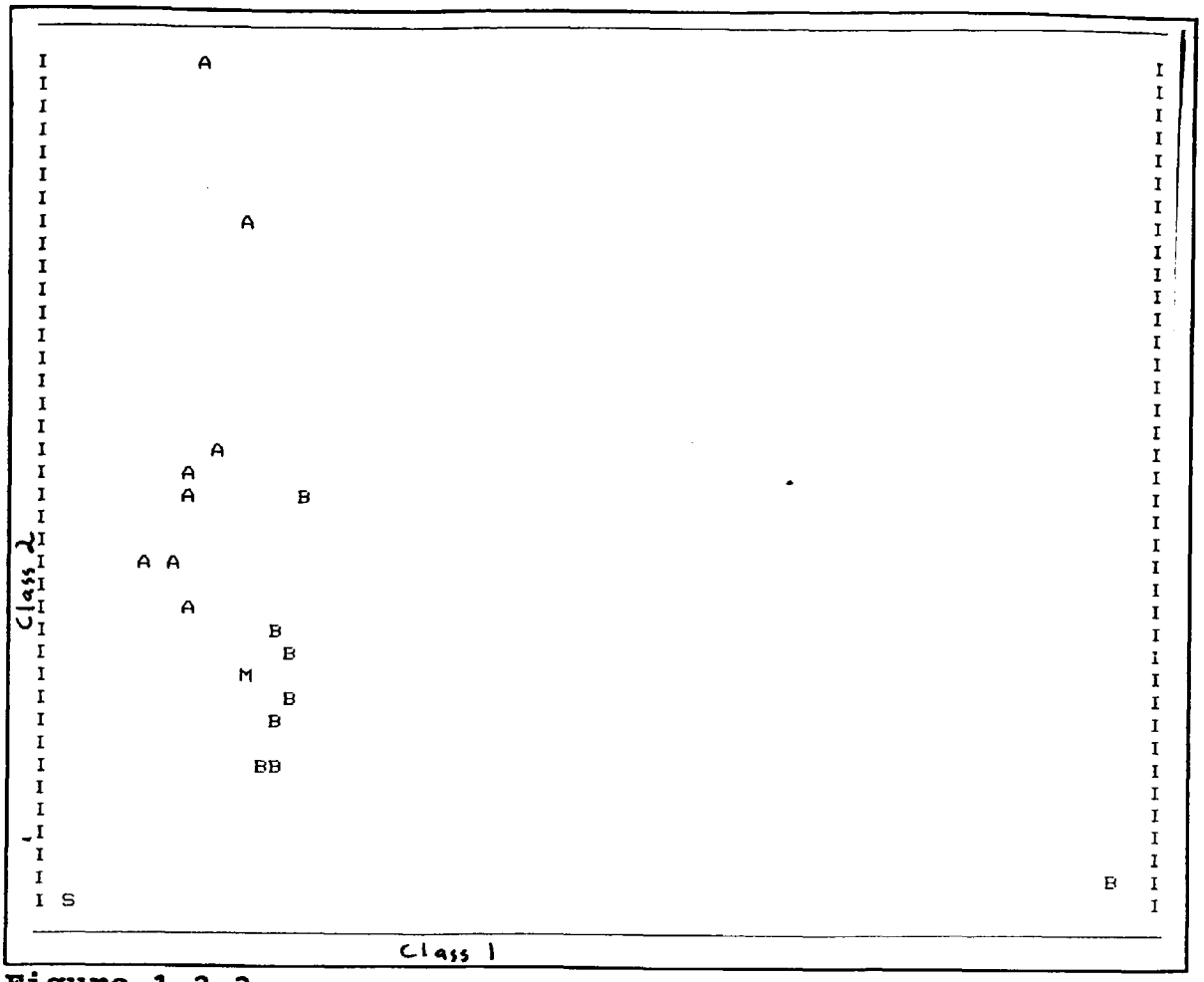


Figure 1.3.3

Class fitted, Pontstickle 1 and 2

The second SIMCA run compared the data from Pontstickle stream #1 and the Upper Cwmbran stream. Figure 1.3.4 shows the preliminary two dimensional plot and clustering can already be seen clearly.

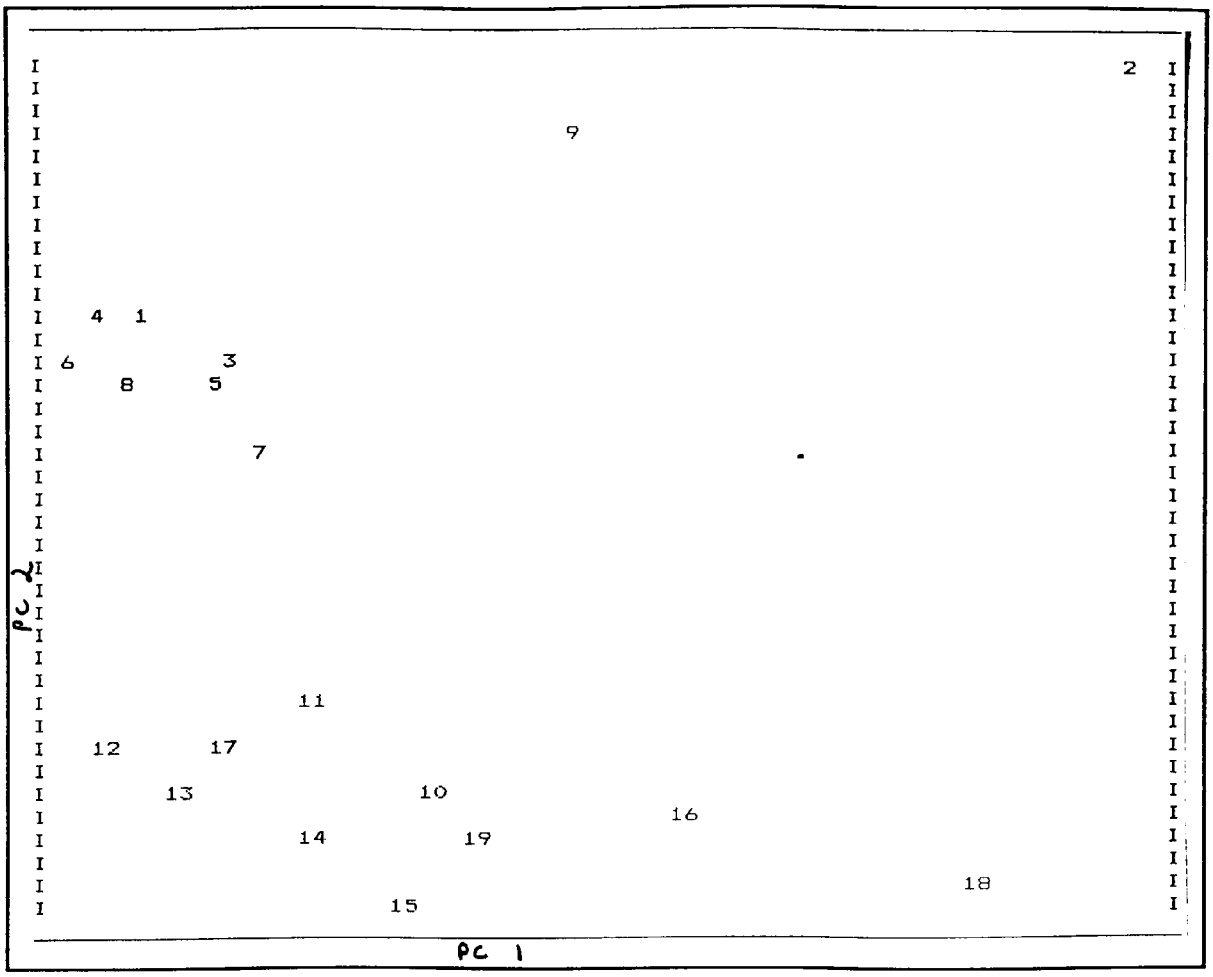


Figure 1.3.4

Pontstickle 1 / Cwmbran Reduced to Factor Space of 2

Principal components were produced as before.

Series 1

No.	SD before	CSV/SD	Q	Av.mod.power
1	1.000	0.313	0.913	0.7461
2	0.300	0.740	0.904	0.8469
3	0.177	1.124	0.892	-

Series 2

No.	SD before	CSV/SD	Q	Av.mod.power
1	1.000	0.331	0.919	0.7893
2	0.338	1.683	0.911	-

On fitting the objects to their classes and plotting the results, excellent distinction may be seen (Fig 1.4.5)

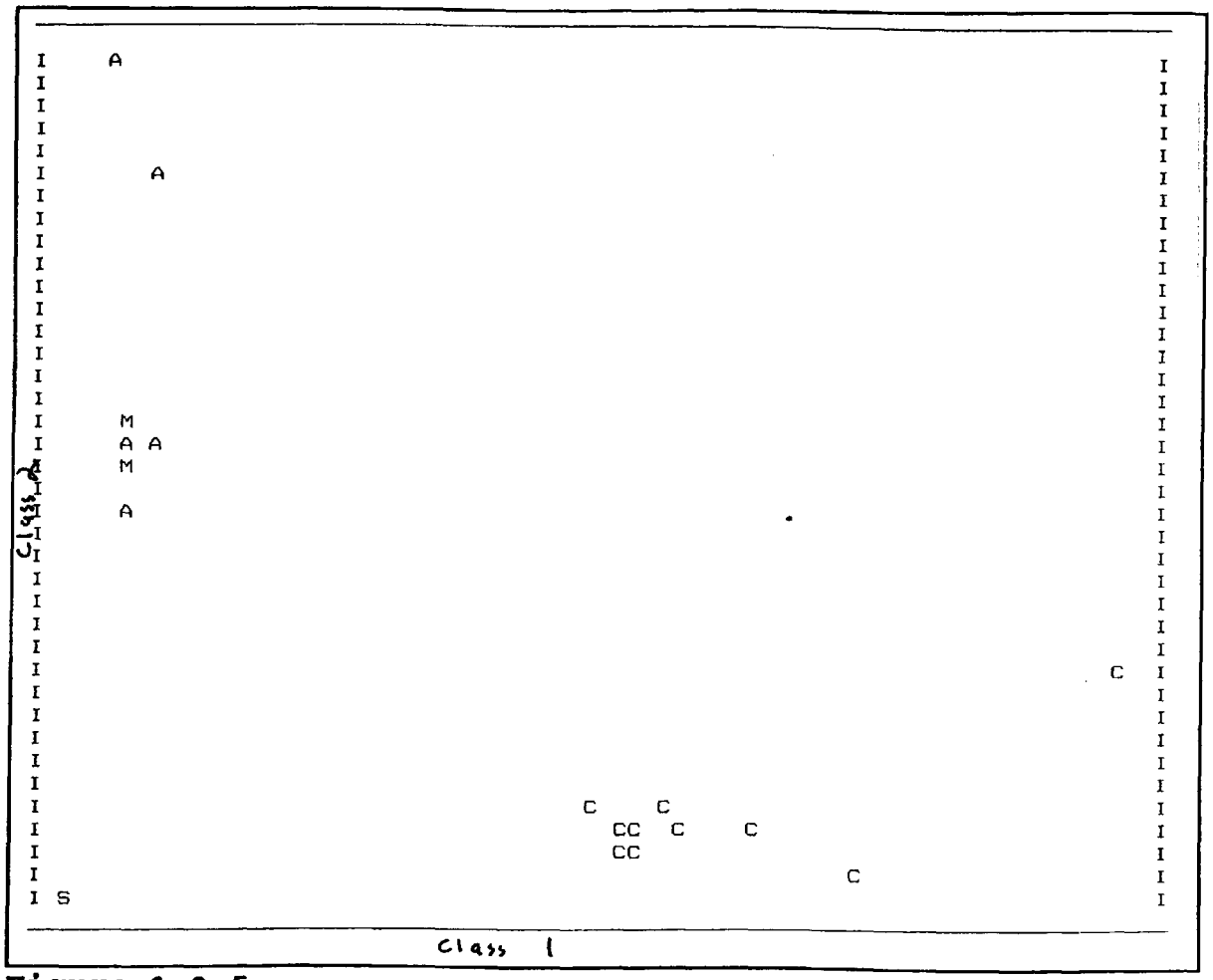


Figure 1.3.5

Class fitted, Pontstickle 1 / Cwmbran

1.3.5 Discussion

The consideration of a single variable e.g. absorbance at a single wavelength, would not have yielded any information as to the source of the water sample. However, the use of SIMCA to manage multi-variate data has been demonstrated to separate waters even from adjacent sources.

The UV absorbance of the water samples was probably due to humic and fulvic substances produced from the peaty soils found in the area. That the procedure could distinguish between two, closely adjacent, streams was quite impressive. There are several possible reasons for the slight variation in humic / fulvic components:-

1) The most obvious is that the two streams may have different sources further up the mountain. If this were the case then it is possible that they pass over different types or ages of peat. If one stream passed over more high-moor areas then the fulvic fraction would dominate over the humic.

It would also be possible that different salts could be dissolved during their trip down the mountain side and the presence of these could affect the absorbance due to the organic compounds by the formation of complexes. The presence of colloidal iron has also been shown to affect the UV spectrum of water and to interfere in the UV analysis of humic acids as the spectra are very similar.

2) If the flow rate of the two streams was different, then the uptake of the organic compounds could be affected. The production of humic acid is also affected purely by the

physical effects in the stream.

3) The level of bacteria could be different, possibly caused by shading or by the presence of dead animals, and one route in the production of humous is degradation of vegetation by bacterial action.

SECTION 2

An Investigation into the Natural Colouring Matter in Water

2.1 Introduction

Surface waters originating in peaty areas may be rich in humic substances which impart a marked yellow colour to the water. This colour is not always removed by conventional water treatment processes (23). The difficulties take the form either of an apparently untreatable fraction of the colour present or of the disproportionate usage of chemical coagulant for the successful removal of the colour.

An attempt was to be made to characterize and classify water samples from different sources and to predict whether or not difficulties would be found in the treatment of that water.

2.2 Literature Survey

2.2.1 Humic Substances

Dissolved organic matter (DOM) is a ubiquitous and dynamic component of lakes, rivers and oceans (24). In lakes this pool of DOM is dominated by dissolved humic substances which can account for up to 80% of the dissolved organic carbon (25).

The terminology used to describe natural organic matter in water is ambiguous. The term "humic substances" does not imply a distinction between humic and fulvic acids. According to Oden (26) fulvic acid is the fraction of an alkaline soil extract that does not precipitate on acidification. The precipitate is defined as humic acid.

The synthesis of humus is based on plant residues. Figure 2.1 suggests how the humification process might be

considered to take place (27).

Humification Process

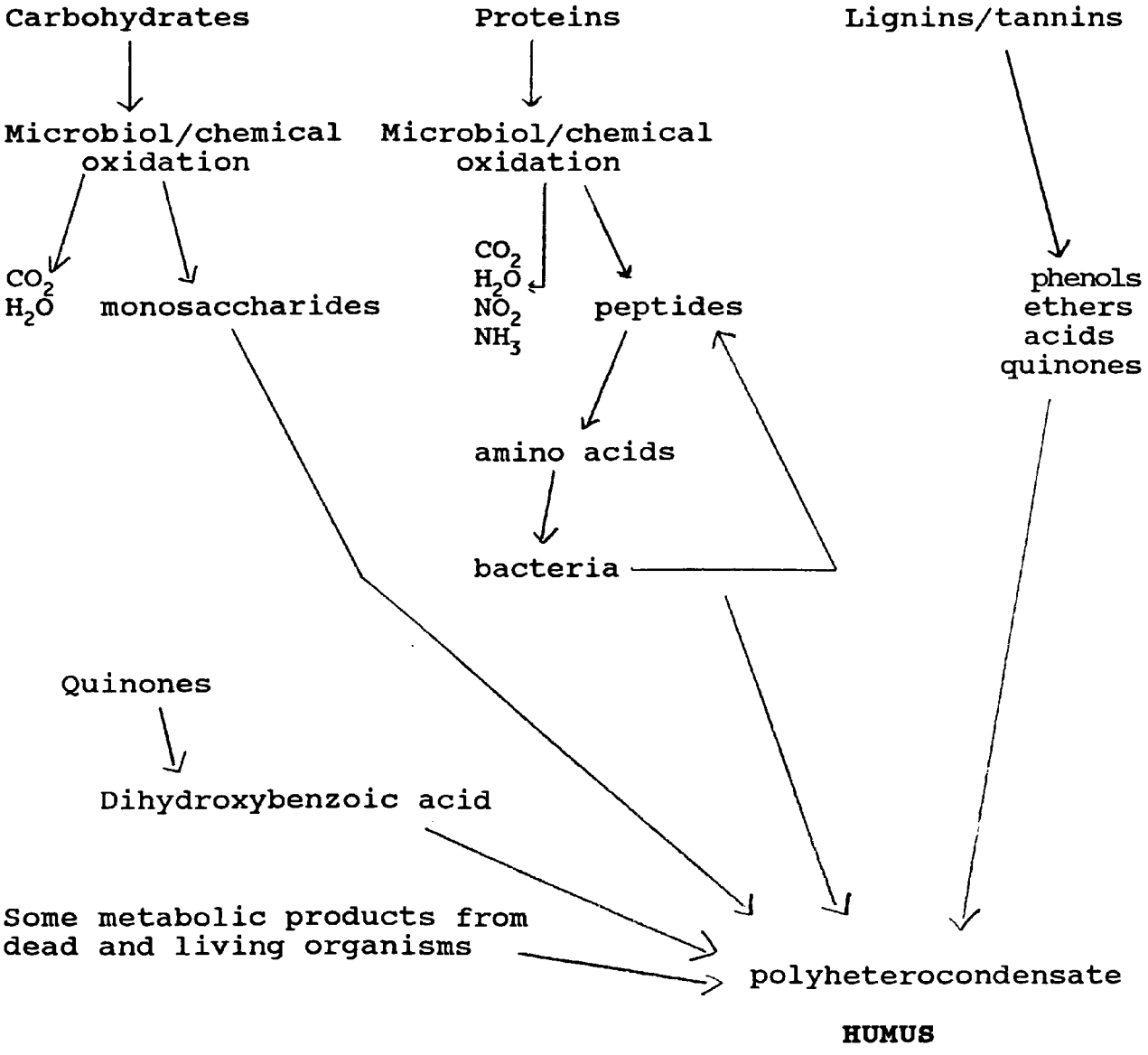


Figure 2.1

The lignin is degraded by micro-organisms to give phenolic substances which are further degraded by many other organisms. Other units may undergo enzymatic or auto-oxidative reaction to form highly reactive radicals or hydroxybenzoquinones which link with other phenolic units, peptides and amino acids to form the large humic molecules.

Anderson (28) takes this a stage further with his hypothesis in which the newly formed humic acids are envisaged as having a relatively high aliphatic content, much of it in the form of side-chains, with a predominantly aromatic core.

Fulvic acids, although primarily considered to be humic acid precursors, may be humic acid degradation products. The humin fraction, or non-extractable humus, is thought to include particulate organic matter entrapped within mineral aggregates and resistant aromatic compounds.

The production of these molecules is a very complicated process and can obviously be affected by such things as the amount and type of vegetable matter available e.g. the decomposition of easily contented vegetation such as moss and heather is slow, whereas the residues of more demanding plants are more easily attacked by micro-organisms.

The humification process is slow. The age of this soil organic matter is reported to be 50-3000 years. Radiocarbon dating has shown that the alkaline-extractable fraction of the organic matter in soil is, on average, 50-250 years old whereas

the unextractable part is considerably older, perhaps up to 2000 years.

As the formation of humic substances is such a complicated and continuous process it is obvious that its composition will vary from place to place. For example, in high-moor, peaty areas the formation of fulvic acids will dominate over other humic acids whereas in low-moor areas the reverse is true.

Gjessing (29) has described how humic substances may bond to clays and other materials. It has also been shown by Aitken et al ((27) that chemical changes in the humic substances can be caused purely by the physical conditions in flowing water. Other changes may be caused by increased microbiological activity due to increase in light or oxygen.

Humus, therefore, is not a definable organic compound and it is unlikely that its composition will be clarified in the near future (30).

2.2.2 Extraction Procedures

Several procedures for extracting humic substances from water have been reported. Black and Christman (31) pre-concentrated samples, either by vacuum distillation or by freeze-drying. They fractionated the resulting concentrates using a series of solvents.

Aho and Lehto (32) used Sephadex G-100 for gel permeation extraction of humic substances. They reported variable results depending upon pH, ionic strength, content of humic acid and

other reasons. Hine and Bursill (33) confirmed these findings.

Vissar (34) used ultra-centrifugation to remove particulate matter followed by ultrafiltration. The humic acid fraction was precipitated by adjusting the pH to 1.5 and repeating the centrifugation. The humic acids were purified by repeating the cycle of precipitation at pH 1.5 followed by dissolution at pH 8.0 until the supernatant at pH 1.5 was colourless. Final extraction of the fulvic acid fraction was from Amberlite XAD-2 using pH 8.4 buffer.

Isolation using diethylaminoethylcellulose was used by Miles et al (35). Desorption was effected using 0.1M sodium hydroxide solution.

Blondeau and Kalinowski (36) used hydrophobic interaction chromatography to fractionate humic substances. They solubilized the humic compound using 0.01M Tris-HCl buffer at pH 7.0. The sample was applied to the column in a solution 3M in sodium chloride. The fractionation was carried out by decreasing the concentration of sodium chloride in the eluting solvent to nil, then by raising the pH by the addition of 0.2M sodium hydroxide. They believed that fractionation was related to CH, CH_2 and CH_3 groups suspected to be the active sites for hydrophobic adsorption at the HA surface.

2.2.3 Analysis of Humic Substances

The analysis of humic acid has been investigated by many different techniques and for several reasons.

Its determination has been carried out by Wilson (37) who

used an oxidative method and by Marino and Ingle (38) using chemiluminescence. An electrochemical procedure was applied by Caminoli et al (39) and Almgren et al (40) utilised fluorescence measurements. However, the most common procedure remains spectrophotometry by such as Martin and Pierce (41).

One problem with spectrophotometric procedures using UV-visible is that the presence of iron, either as colloidal iron hydroxide or as an iron-humic complex, causes interference due to similarities in their spectra. Sholkovitch et at (42) tackled the determination of iron content by precipitation of the humic acid but it has been found that the iron can co-precipitate. Carpenter and Smith (43) developed a simple and rapid method utilising two absorbance measurements, one on an untreated sample aliquot and one on a sample treated to enhance iron absorbance.

Desbene and Delamar (48) used X-ray photoelectron spectroscopy (XPS) to determine the different chemical forms of carbon and sulphur existing in humous samples. They found that the elemental analysis did not show important differences between samples taken from different locations. When XPS estimates of elemental concentrations were compared with standard procedures, agreement for sulphur was obtained. However, the results for nitrogen and oxygen differed by a factor of two to three. They suggested that the surface composition of the humic acids were different from their mean composition. In the samples investigated they found three different forms of carbon, one form of nitrogen and two of

sulphur. The carbon forms were thought to be:-

a) aliphatic and aromatic

b) C - O

c) C = O

The nitrogens were slightly positively charged and the two sulphurs were neutral and positively charged (sulphones, sulphates etc.)

In 1982 Visser (51) investigated the acid functional group content of humic substances. He looked at equivalent weights and the number of carboxyl and phenolic hydroxyl groups. He was able to show that, although aquatic humic compounds have several properties in which they are quite distinct from soil humus, their acidic functional groups differ very little.

As the structure of humic acid is amorphous a great deal of work has been done to attempt to determine the sub-units present. The main thrust of this has been to degrade the humic acid and to identify the degradation products. The degradation procedures can be broadly summarised as:-

a) reduction

b) oxidation

c) hydrolysis

d) thermal

e) biological

Of these techniques reductive degradation has not been found to be particularly successful (44, 45, 46). This is probably due to humic substances containing considerable amounts of oxygen.

Liao et al (47) used both potassium permanganate oxidation

and sodium hydroxide hydrolysis to degrade humic materials. The resulting mixtures were methylated and analyzed by GC/MS to identify the components present. They were able to identify six groups of compounds:-

a) benzene carboxylic acids

b) furan carboxylic acids

c) aliphatic monobasic acids

d) aliphatic dibasic acids

e) aliphatic tribasic acids

f) carboxyphenyl glyoxylic acids

The identified products amounted to about 25% by weight of starting materials.

Schnitzer and Skinner (49) oxidised humic substances using peracetic acid at 40°C for eight days. The resultant products were identified using mass spectrometry and micro infra red techniques. They identified three major groups of products:-

a) phenolic acids

b) benzene carboxylic acids

c) aliphatic acids

With respect to total yields of oxidation products this procedure compared well with the permanganate process at 100°C but was far more time consuming.

Ruggiero et al (50) also used oxidation by peracetic acid. They investigated the products of oxidation using ^1H NMR. They showed the presence of $-(CH_2)_n-CH_3$ ($n > 6$) fragments belonging to n-alkanes and / or n-fatty acids physically adsorbed onto the macromolecule structure. Oxidative

degradation caused partial cleavage of aromatic rings.

Neyroud and Schnitzer (52) fractionated fulvic acid and treated the fraction with a variety of techniques including permanganate oxidation and sodium hydroxide hydrolysis. They found the major decomposition products to be benzenecarboxylic and phenolic acids, together with smaller amounts of aliphatic fatty acids. They also suggested that some of the fatty acids esterified to the OH groups of the phenolic acids.

Jackson et al (53) refluxed humic substances with phenol together with p-toluene sulphonic acid catalyst for twenty four hours. The resulting mixture of components was methylated and individual compounds identified by GC-MS, NMR, IR and elemental analysis. They found a series of substituted xanthenes, benzofuran and several bis-methoxyphenyl compounds.

Meuzelaar et al (54) used pyrolysis-mass spectral analysis to investigate humic acids. All the humic acids tested yielded complex but highly similar spectra. Prominent, homologous ion series were: sulphides, pyrroles, benzenes, phenols, methoxyphenols and dimethoxyphenols. Similar results were found by deHaan et al (55).

2.3 Development and Application of Analytical Methods

2.3.1 Overview

As described in section 2.2.1, humic acids are extremely complicated substances. No attempt, therefore,was to be made to elucidate any structures but rather differences between compounds from various sources were to be investigated. The approach taken was:-

1) Determination of a reproducible, quantitative extraction and concentration procedure.

2) Determination of a reproducible and controlled degradation procedure.

3) Development of a high performance liquid chromatographic separation of the degradation products.

4) Application of multi-variate statistics and pattern recognition techniques to the results.

2.3.2 Extraction

For the purposes of the current work, a rapid and quantitative extraction protocol was required. Solid Phase Extraction (SPE) columns are widely available. They contain chromatographic materials similar to those used for HPLC. The particle size, however, is larger and less well controlled so that high pressures are not required for elution. The most often used columns, together with typical extracted compounds, are listed below:-

a) Reversed phase:

Octadecyl bonded for drugs, priority pollutants, essential oils etc.

Octyl bonded for moderately polar compounds

Phenyl bonded for non-polar compounds

b) Normal phase:

Silica for adsorption extraction of polar compounds e.g. alcohols, ketones, phenols etc

Cyano and diol bonded for different selectivity

c) Ion exchange:

Amino for weak anion exchange e.g. amino acids

Quaternary amine for strong anion exchange e.g. nucleic acids, antibiotics

Aromatic sulphonic acid for strong cation exchange e.g.water soluble vitamins.

A quaternary amine system was selected for the extraction of humic acids. The details of the extraction procedure are described in section 2.3.3 (recovery).

2.3.3 Recovery Studies

Preliminary work was done using purchased humic acid sodium salt (Aldrich Chemical Company, Gillingham, UK). A series of experiments was carried out to determine the UV / visible spectra and linearity of response in a series of aqueous solvents.

A Shimadzu UV 240 scanning spectrophotometer was used to

obtain the spectra between 700 and 200 nm of a series of dilutions in 0.1M phosphate buffer at pH 7.0, 0.1M sodium hydroxide and deionised water. Five test solutions were used ranging from 30 to 70 mg/lt. Typical spectra are shown below in figures 2.3.1 to 2.3.3.

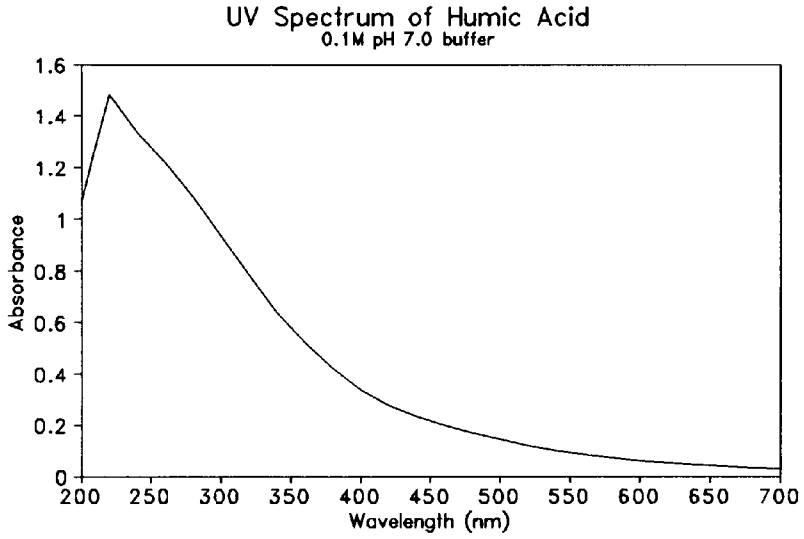


Figure 2.3.1

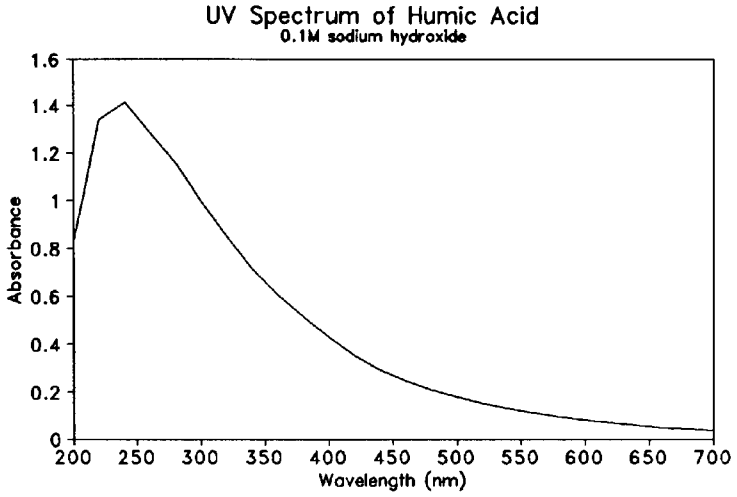


Figure 2.3.2

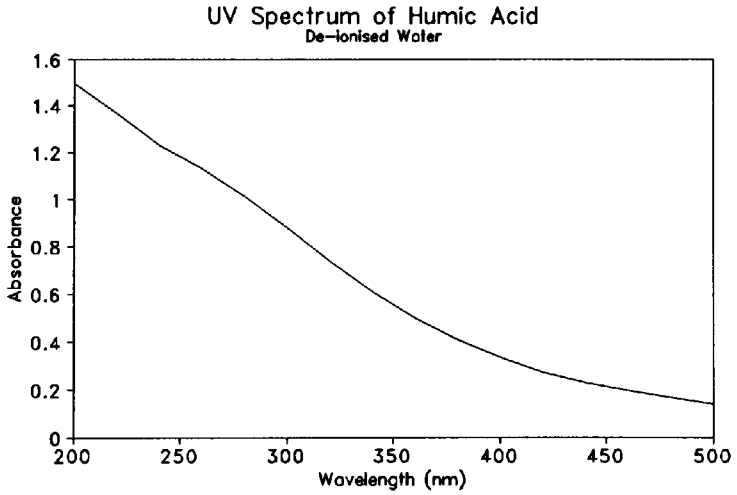


Figure 2.3.3

Linearities were found to be good across the range tested. The regression plots of absorbance at 300 nm against concentration for each solvent are shown below in figures 2.3.4 to 2.3.6. The excellent straight lines produced were typical for all wavelengths

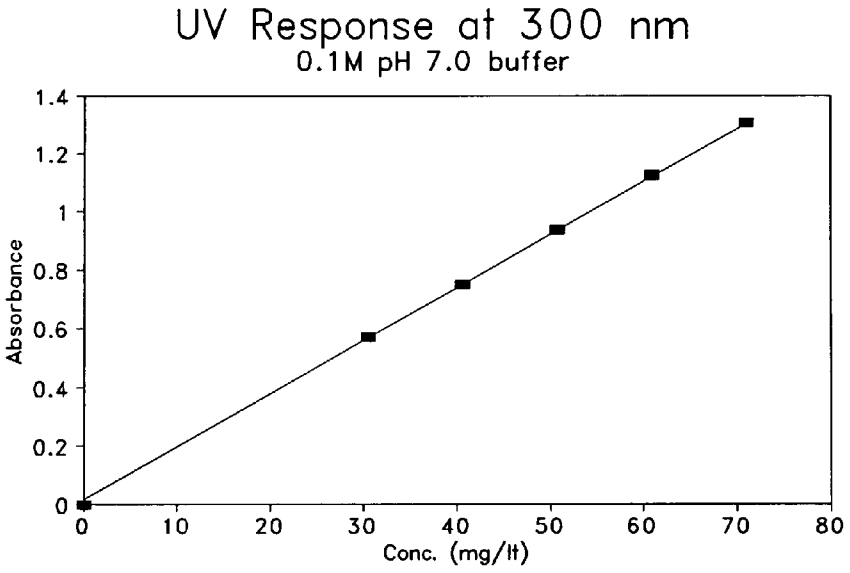


Figure 2.3.4

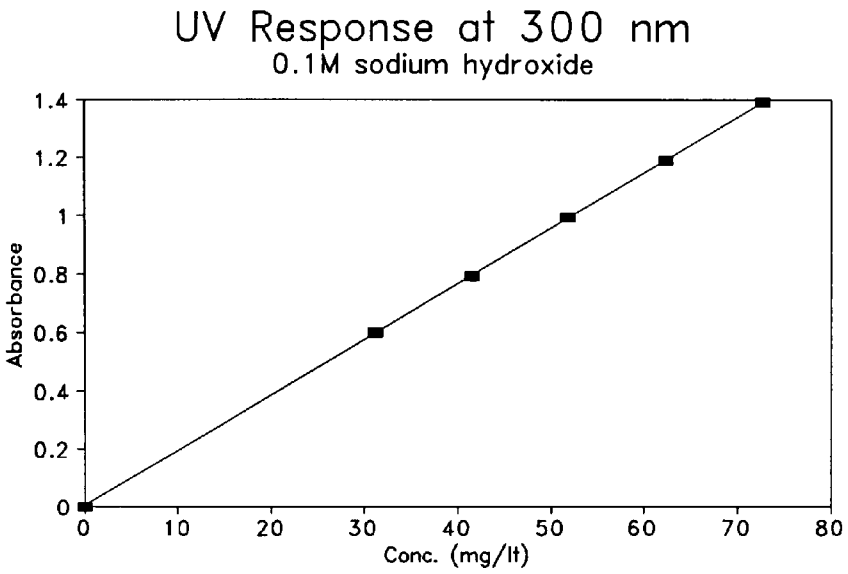


Figure 2.3.5

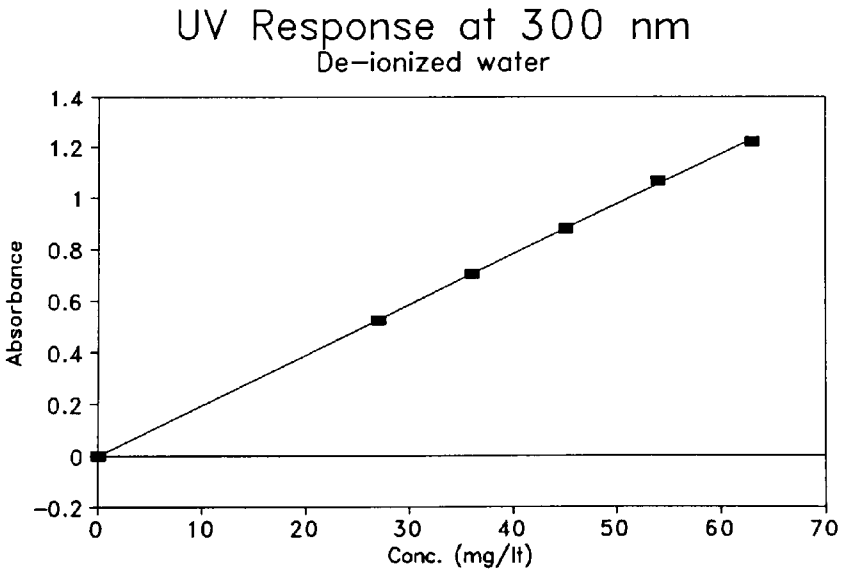


Figure 2.3.6

Having established that a linear response curve exists for UV absorption of humic acid, the recovery from an extraction column was investigated.

Free humic acid was prepared from the sodium salt by dissolution in water, acidification with hydrochloric acid and filtration. The product was dried under vacuum at 60°C.

The spectrum of a solution of humic acid in 0.01M pH 7.6 buffer was obtained between 500 and 200 nm in 10mm cuvettes. 0.01M pH 7.6 buffer was used as a reference.

A Bond-Elut SAX (quaternary amine 500mg) column was mounted on a Vac-elut manifold and washed with one column volume of methanol followed by one of water. Humic acid solution in 0.01M pH 7.6 buffer (25.0ml) was drawn through the column with the aid of vacuum. A brown ring of humic acid was clearly visible at the top of the column.

The column was washed with one volume of water and the humic acid eluted with 10.0ml of 1.0M sodium hydroxide solution. The eluent was diluted tenfold with 0.01M pH 7.6 buffer for spectrum determination.

The pre- and post- extraction spectra are shown below in Figure 2.3.7

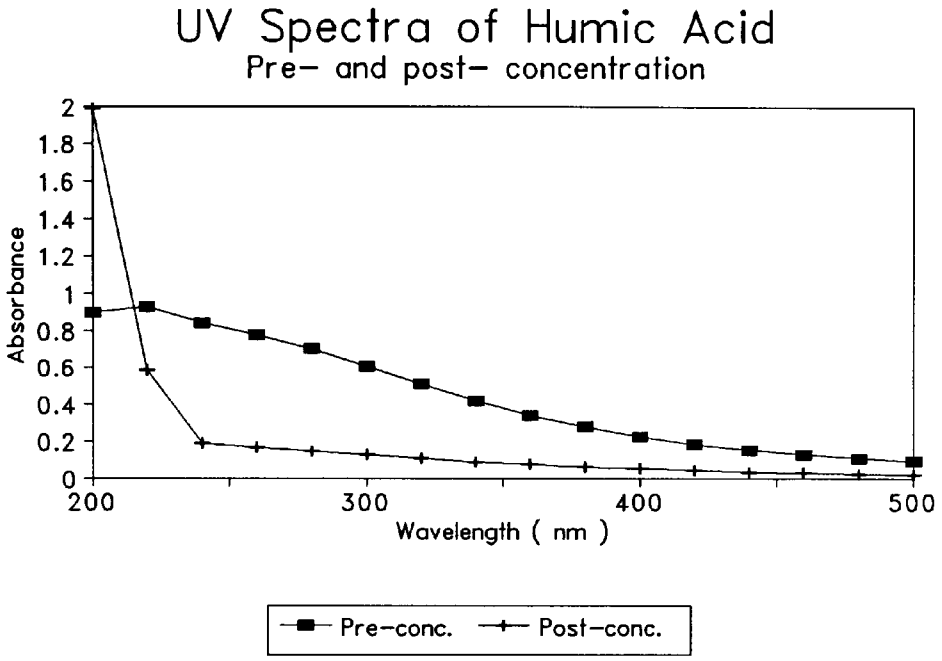


Figure 2.3.7

The absorbance at 300nm was used to estimate recovery, which was found to be ca.84%. This compares well with that obtained using diethylaminocellulose. However, major differences were evident in the spectra produced. It was apparent that some change or decomposition of the humic acid had occurred.

The experiment was repeated using 0.5M sodium hydroxide as eluent. The spectra are shown in Fig. 2.3.8

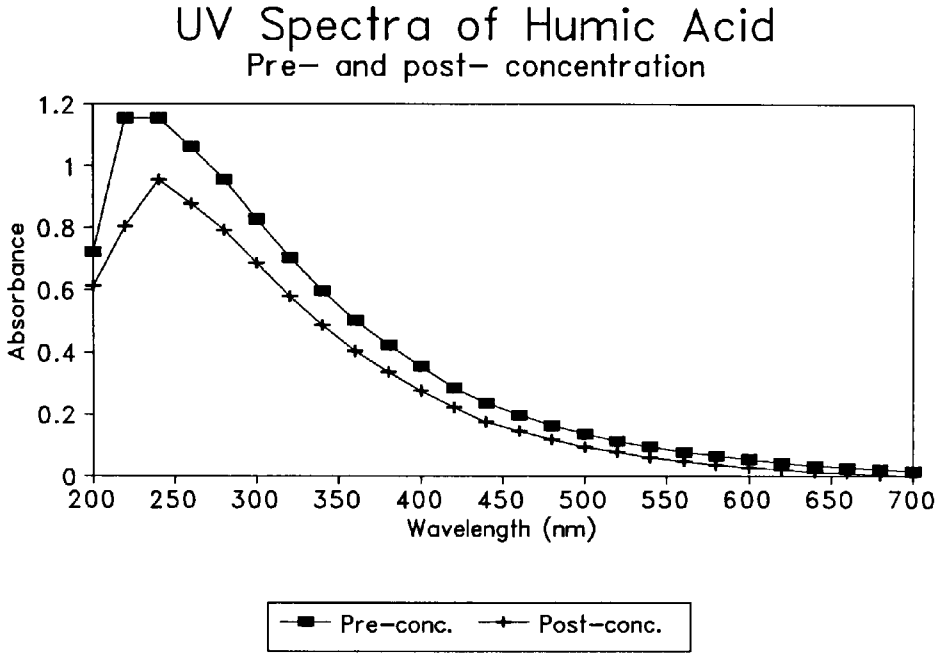


Figure 2.3.8

Whilst the spectra are similar in this case, it was noted that considerable brown colouration was left on the extraction column, indicating incomplete elution using the weaker eluent.

Personal communication from the Water Research Centre (62) has indicated that these findings have been confirmed elsewhere. This will be expanded upon in the discussion section.

2.3.4 Degradation Experiments

As discussed earlier, humic and fulvic acids are extremely complex materials. Degradation of the complex molecules to give mixtures of more simple components would seem to be a means of producing data for use in pattern recognition procedures. The oxidation procedure described by Liao et al (47) was chosen as a convenient process.

In order to study the degradation process a commercial humic acid was obtained. This was well mixed before work started. The oxidation procedure was as follows:-

Humic acid (0.5 g) was suspended in water (50ml) and sodium carbonate added to effect solution. The final pH was 10.3. Potassium permanganate (0.5g) was added and the mixture heated to reflux under an atmosphere of nitrogen. Further potassium permanganate was added to the mixture, as necessary, to maintain a purple colour.

After two hours the mixture was cooled to room temperature and the pH adjusted to 12.0 with dilute sodium hydroxide solution. The mixture was centrifuged at 6000rpm for fifteen minutes and the supernatant transferred to a beaker where it was acidified with dilute sulphuric acid. The mixture was warmed to 65°C and the excess permanganate decomposed with sodium oxalate solution added dropwise until the purple colour was lost.

Dilute sodium hydroxide solution was added to adjust the pH to 12.0 and the mixture was centrifuged. The supernatant was removed and 1.0ml acidified with phosphoric acid, mixed with

tetra-butyl ammonium hydroxide. 20μl was injected into an HPLC system using the following conditions:-

Column:	100 x 4.6mm Techsphere ODS 3μm
Mobile-phase	Acetonitrile / water / 0.1M TBAH
	Solvent A: 5 / 90 / 5
	Solvent B: 95 / 0 / 5
Gradient:	Linear 20 - 70% B during 60 minutes
Flow-rate:	1.5 mls/min
Detection:	UV at 235 nm

The chromatogram produced is shown below in Fig. 2.3.9. The peaks were tentatively identified as benzene carboxylic acids.

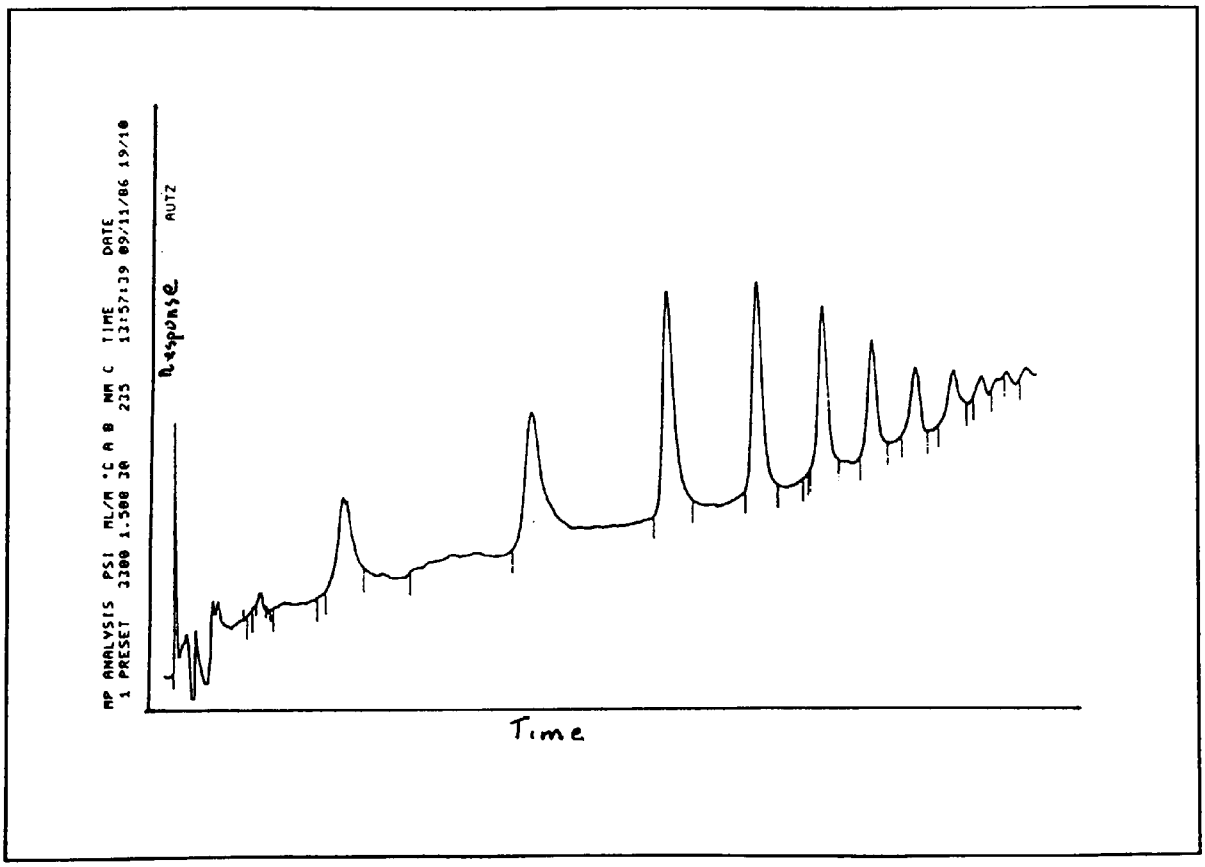


Fig. 2.3.9

Degraded Humic Acid

As the chromatogram was reasonably complex, simplex optimization was used to develop the best separation. In order to use this procedure some sort of numerical estimate of the overall separation is required. Berridge (63) developed the chromatography response function (CRF) shown below

$$CRF = \Sigma Ri + n^a - b[Ta - Tn] - c[T_0 - T_1]$$

where n = number of peaks detected

T_1= Retention time of first peak

T_0= Specified minimum retention for first peak

Ta= Maximum desired retention for last peak

Tn= Retention time of last detected peak

a,b,c = operator selectable weightings

Ri = Resolution between adjacent peak pairs

$$Resolution = \frac{[T_i - T_{(i-1)}]}{[0.5(W_i + W_{(i-1)})]}$$

where W = peak width at base. This can be approximated as 1.7 x width at half-height for ease of measurement.

A program was written in GWBASIC to calculate CRF for a separation and is shown in Fig. 2.3.10. Berridge's program for calculation of simplex results is also shown.

CRF calculator

```
) REM**CRF calculator**
) REM** Ref. J.C.Berridge**
5 REM**Techniques for Automated Optimization of HPLC Separations**
) REM**K.J.Caldicott**
5 CLS
) INPUT"Number of peaks detected";N
) DIM T(N),W(N),H(N),R(N),RES(N),SUB(N)
) FOR I=1 TO N
) INPUT"Retention time of peak ";T(I)
) INPUT"width at half-height";H(I):W(I)=H(I)*1.7
) NEXT I
)0 REM**Screen list plus error trap**
10 CLS
20 IF N>15 THEN 200
25 FOR I=1 TO N
30 GOSUB 1000
35 NEXT I
40 GOTO 300
)0 FOR I=1 TO 15
10 GOSUB 1000
20 NEXT I
30 GOTO 300
40 FOR I=16 TO N
50 GOSUB 1000
60 NEXT I
70 GOTO 300
00 INPUT"Any errors ";A$
10 IF A$="n" THEN 350
20 GOSUB 2000
25 GOTO 100
50 TOTAL=0:NUM=N*N:LONG=(15-R(N)):SHORT=(2-R(1))
60 FOR I=2 TO N
70 RES(I)=2*(T(I)-T(I-1))/(W(I)+W(I-1))
80 SUB(I)=RES(I)
90 TOTAL=TOTAL+SUB(I)
00 NEXT I
10 TOTAL=TOTAL+NUM-LONG-SHORT
60 LPRINT"PEAK NO.","RETENTION","    W (0.5)","    W(calc)","  RESOLUTION"
70 LPRINT"1                 ";
71 LPRINT USING"££.£££          ";T(1),H(1),W(1);
72 LPRINT"       N/A"
80 FOR I=2 TO N
90 LPRINT USING"££.£££          "; I,T(I),H(I),W(I),RES(I)
00 NEXT I
10 LPRINT:LPRINT
20 LPRINT "    CRF is  ";TOTAL
30 STOP
000 PRINT I,T(I),H(I)
010 RETURN
000 INPUT"Line number ";L
010 INPUT"Retention,w0.5 ";T(L),H(L):W(L)=H(L)*1.7
020 PRINT"Correct";A$
030 IF A$="n" THEN 2000
040 RETURN
```

Fig. 2.3.10

```
      SAMP ID : 5/5
BASIC: SIMPLEX
 990 REM **  VERSION 2  FEBRUARY 1981   ***
1000 CLEAR
1010 PRINT "OPTIMISATION OF VARIABLES BY SIMPLEX METHOD."
1020 PRINT "WITH EXPANSION AND CONTRACTION"
1030 PRINT "      PREPARED BY J.C.BERRIDGE."
1040 PRINT "              PFIZER CENTRAL RESEARCH. U.K."
2000 PRINT "DO YOU WANT INSTRUCTIONS?"
2005 INPUT "ANSWER Y (YES) OR N (NO)       " S$
2010 PRINT "    .........WAIT........."
2020 IF S$= "N" THEN GOTO 3000
2030 PRINT
2035 CLEAR
2040 PRINT "THIS PROGRAM NEEDS 2 TO 10 VARIABLES"
2050 PRINT "IF YOU HAVE N VARIABLES, "
2060 PRINT "YOU MUST CONDUCT N+1 EXPERIMENTS TO DETERMINE"
2070 PRINT "HOW THE RESULTS ARE AFFECTED BY EACH VARIABLE."
2090 PRINT "      IF THE COORDINATE FOR A VARIABLE IS "
2100 PRINT "OUTSIDE THE BOUNDARY FOR THAT VARIABLE, DO NOT "
2110 PRINT "CARRY OUT AN EXPERIMENT BUT INSTEAD ASSIGN "
2120 PRINT "A VERY UNDESIREABLE RESPONSE."
2124 REM SOFTWARE DELAY FOR SCREEN  **************
2125 FOR X=1TO 10 STEP .1
2126 Y=LOG(X)
2127 NEXT X
2129 CLEAR
2130 PRINT "      WHEN AN OPTIMUM HAS BEEN FOUND,"
2140 PRINT "THE SIMPLEX WILL CIRCLE - COORDINATES WILL BE "
2150 PRINT "APPROXIMATELY REPEATED."
2160 PRINT "      YOU MAY FIND A LOCAL OPTIMUM, BUT, IF"
2170 PRINT "YOU START AGAIN FROM A VERY DIFFERENT LOCATION"
2180 PRINT "AND END UP AT THE SAME OPTIMUM, YOU CAN BE MORE"
2190 PRINT "CONFIDENT THAT AN OVERALL OPTIMUM HAS BEEN FOUND."
2210 PRINT
2220 PRINT "ARE YOU HAPPY TO GO ON? "
2225 INPUT "ANSWER Y (YES) OR N (NO) "Z$
2240 IF Z$ ="Y" THEN 3000
2245 PRINT "READ DEMING AND MORGAN,ANAL.CHEM 45 NO3 1973."
2250 STOP
3000 REM START   **********
3010 DIM P(13,11)
3020 DIM A(50,50)
3025 DIM R(50)
3030 PRINT
3040 W=0
3041 Z9=0
3050 E=0
3060 CLEAR
3070 PRINT
3080 PRINT "IS THIS A NEW SERIES (N) OR A "
3090 INPUT "CONTINUATION (C) " N$
3100 IF N$="N" THEN 4000
3110 PRINT "YOU SHOULD HAVE RESULTS FROM EARLIER EXPERIMENTS."
3140 INPUT "IS THIS SO (Y/N)" R$
3150 IF R$="Y" THEN 3500
3160 IF R$="N" THEN 3210
3170 TOP
3180 PRINT
3190 PRINT "INVALID RESPONSE"
3200 GOTO 3110
3210 PRINT "SORRY-YOU WILL HAVE TO START AGAIN"
3215 FOR N=1 TO 200
3216 NEXT N
3220 GOTO 4015
3500 PRINT "PLEASE ENTER THE RESULTS YOU HAVE "
3510 GOTO 4015
```

```
4000 REM NEW SERIES **************
4010 B1=1
4015 CLEAR
4020 PRINT
4030 INPUT "HOW MANY VARIABLES IN THIS SIMPLEX " V
4050 PRINT "YOU HAVE SELECTED ";V; " VARIABLES. "
4060 IF V>10 THEN 4070
4063 IF V<2 THEN 4075
4065 GOTO 4080
4070 PRINT "BUT MAX. OF 10 "
4072 GOTO 4020
4075 PRINT "BUT MIN.OF 2 "
4077 GOTO 4020
4080 FOR N=1 TO 300
4082 NEXT N
4085 CLEAR
4090 INPUT "ENTER THE OPTIMUM RESPONSE YOU HOPE TO ACHIEVE."N
4100 FOR N=1 TO 300
4105 NEXT N
4110 E1=V+2
4115 REM ARRAY OF COORDINATES AND RESULTS**************
4120 FOR I=1TO V+1
4125 CLEAR
4130 PRINT
4140 PRINT "EXPERIMENT NO.  ";1;
4150 PRINT "------------"
4160 FOR J=1TO V
4170 PRINT
4180 PRINT "ENTER VARIABLE NO.  ";J;
4190 INPUT P(I,J)
4195 A(I,J)=P(I,J)
4200 NEXT J
4210 PRINT
4220 PRINT "ENTER RESULTS"
4230 INPUT P(I,V+1)
4235 R(I)=P(I,V+1)
4240 NEXT I
4250 B=1
4260 GOSUB 8000
4270 W=W2
4280 GOSUB 9000
5000 REM PRINT NEW VARIABLES   **************
5010 CLEAR
5020 PRINT "FOR EXPERIMENT NO. ";E1; TAB(40); "USE :"
5030 PRINT
5040 I =V+2
5050 FOR J=1 TO V
5060 PRINT "FOR VARIABLE NO. ";J; TAB(40);" :  ";P(I,J)
5080 NEXT J
5100 IF B1=3 THEN LET B1=1
5110 E1=E1+1
5140 P(V+2,V+1) =1.000000E+37
5200 PRINT
5210 INPUT "ARE THESE VALID CONDITIONS-ANSWER Y OR N "Z$
5220 FOR N=1 TO 100
5222 NEXT N
5225 CLEAR
5230 IF Z$="Y" THEN 6010
5240 PRINT
5250 PRINT "ENTER A VERY POOR RESPONSE VALUE. "
5260 GOTO 6070
```

```
6000 REM INPUT NEXT RESULTS  ************
6010 PRINT
6020 PRINT "DO YOU WANT TO STOP NOW SO THAT YOU CAN CARRY "
6025 INPUT "OUT THE EXPERIMENT (Y/N) " B$
6030 IF B$="N" THEN 6070
6035 PRINT "       THE EXPERIMENT NUMBERS, VARIABLES AND "
6040 PRINT "WILL NOW BE PRINTED ON THE PLOTTER."
6042 INPUT "IS THE PLOTTER FREE (Y/N) " A$
6044 IF A$ ="Y" THEN7000
6045 PRINT "WAIT UNTIL THE PLOTTER IS FREE"
6050 FOR N=1TO 300
6055 NEXT N
6060 PRINT
6068 PRINT "NOW, ";
6069 GOTO 6042
6070 PRINT "FOR EXPERIMENT NO. ";E1-1;TAB(40);"VARIABLES WERE:"
6079 PRINT
6081 PRINT
6090 I=V+2
6100 FOR J=1TO V
6110 PRINT "VARIABLE NO: ";J;TAB(40); ": ";P(I,J)
6111 A(E1-1,J)=P(I,J)
6120 PRINT
6130 NEXT J
6140 B=1
6150 PRINT "ENTER YOUR RESULTS"
6160 INPUT P(V+2,V+1)
6161 X=E1-1
6162 R(X)=P(V+2,V+1)
6170 IF B1=5 THEN GOTO 6280
6180 P5=ABS(M-P(V+2,V+1))
6190 IF P5>=W3  IF P5<W4  IF B1=1 THEN GOTO 6290
6200 IF P5>=W5    IF B1=1  THEN LET B= -.5
6210 IF P5<W5   IF P5>W4   IF B1=1  THEN LET B= .5
6220 IF B<1 THEN LET B1=5
6230 IF P5<W3 THEN LET B1=B1+1
6240 IF P5<W3    IF B1=2  THEN LET B=2
6250 IF P5<W3  IF B1=3  THEN GOTO 6280
6260 IF B1=2   IF B=1  THEN GOTO 8500
6270 GOTO 4280
6280 B1=1
6290 FOR J=1TO V+1
6300 P(W,J)=P(V+2,J)
6310 NEXT J
6320 GOTO 4260
7000 REM PRINT VERTEX VALUES ***********************************
7005 I=1
7007 FOR I=1 TO E1-2
7010 PRINT #P, "EXPERIMENT NO.  ",I
7020 FOR J=1 TO V
7030 PRINT #P, "VARIABLE NO. "J,A(I,J)
7040 NEXT J
7050 PRINT #P, "RESULT ",R(I)
7060 NEXT I
7070 PRINT #P," NEW VARIABLES ARE......."
7080 FOR J= 1 TO V
7090 PRINT #P, "VARIABLE NO. " J,".....",P(V+2,J)
7100 NEXT J
7110 GOTO 6079
```

```
8000 REM WORST CASE  ***************
8010 W2=0
8020 FOR K=1 TO 2
8030 X1=0
8035 W1=0
8040 W3=1.0000000E+38
8050 J= V+1
8060 FOR I=1TO V+1
8070 IF I=W2 THEN GOTO 8120
8080 X=M-P(I,J)
8090 IF ABS(X) >X1 THEN LET W1=I
8100 IF ABS(X) >X1 THEN LET X1=ABS(X)
8110 IF ABS(X) <W3 THEN LET W3=ABS(X)
8120 NEXT I
8130 IF W2=0 THEN LET W5= X1
8140 W4=X1
8150 IF W2=0 THEN LET W2=W1
8160 NEXT K
8170 RETURN
8500 REM USE LAST VERTEX   **********
8510 FOR J=1 TO V+1
8520 P(W,J) =P(V+3,J)
8530 NEXT J
8540 B1=1
8550 GOTO 4260
9000 REM REFLECTED VERTEX   ***********
9010 IF B<>2 THEN GOTO 9050
9020 FOR J=1 TO V+1
9030 P(V+3,J) =P(V+2,J)
9040 NEXT J
9050 FOR J=1 TO V
9060 P0=0
9070 FOR I=1 TO V+1
9080 IF I =W THEN GOTO 9100
9090 P0= P0 + P(I,J)
9100 NEXT I
9110 P0= P0/V
9120 P(V+2,J) = P0 + B*(P0-P(W,J))
9130 NEXT J
9140 RETURN
9150 END
```

The values for the operator selectable weightings in the CRF calculation were:-

a = 2 b = c = 1 Ta = 15 T_0 = 2

A series of carboxylic acids was selected as likely decomposition products of oxidation or hydrolysis of humic acid. The acids chosen were:

2-furoic acid

3,4-furandicarboxylic acid

2,5-dimethylbenzoic acid

2,5-dihydroxybenzoic acid

1,3,5-benzenetricarboxylic acid

1,2,3-benzenetricarboxylic acid

1,2,4-benzenetricarboxylic acid

4-ethylbenzoic acid

1,2,4,5-benzenetetracarboxylic acid

1,4-benzenedicarboxylic acid

1,2-benzenedicarboxylic acid

benzoic acid

The basic HPLC conditions were:-

Column: 125 x 4mm Superspher ODS 4μm (Merck)

Mobile-phase: A = Acetonitrile / 0.02M pH 2.7 buffer 5 / 95

B = Acetonitrile / 0.02M pH 2.7 buffer 70 / 30

Flow-rate: 2.0 mls / min

Detection: UV at 230 nm

The variables used in the simplex were:-

A = starting % B

B = length of preliminary isocratic segment

C = final % B

D = time of gradient segment

The CRF reached a maximum after 16 experiments with

A = 0.9 B = 4.4 C = 60.5 D = 6.3

The optimized separation is shown below in Fig. 2.3.11

Optimized HPLC Separation

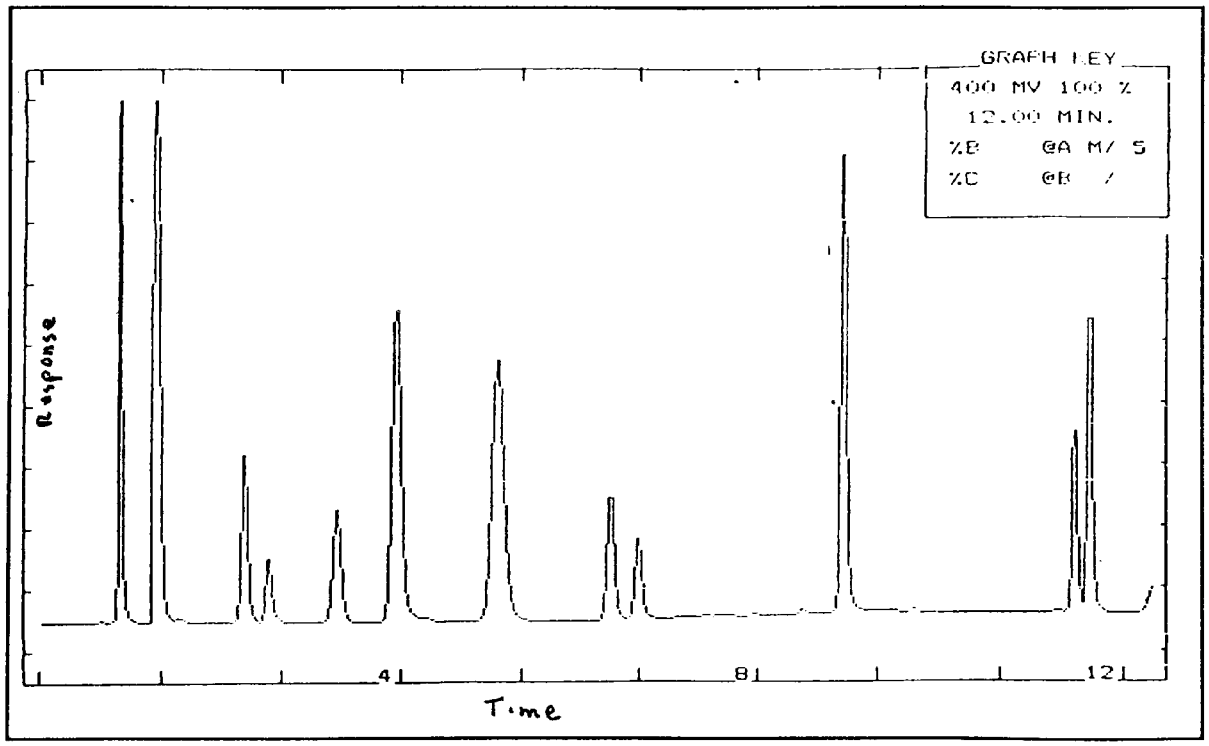


Fig. 2.3.11

The optimized separation was used to analyze the results of many attempts at oxidation using tightly controlled parameters. A typical chromatogram is shown in Fig. 2.3.12.

Similar chromatograms were produced when hydrolysed samples were injected into the system. Unfortunately this degradation system was equally irreproducible.

Typical Chromatogram after Oxidation

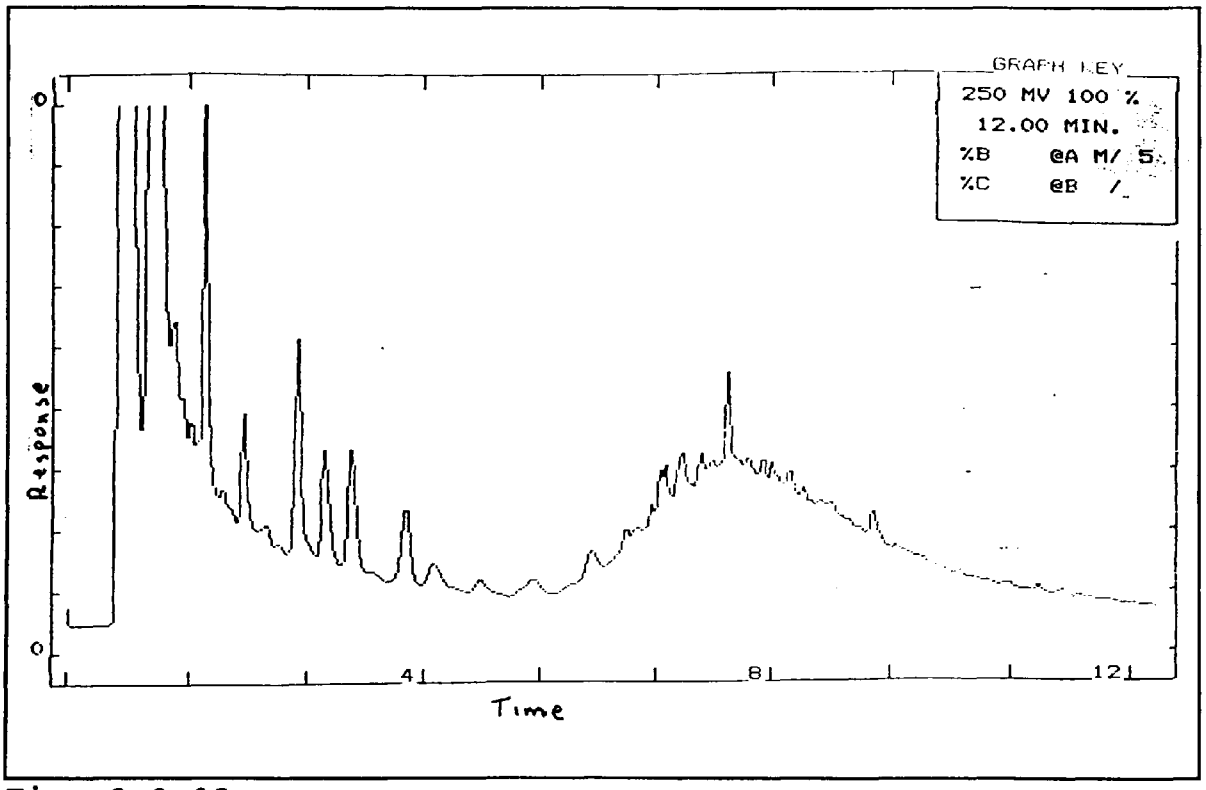


Fig. 2.3.12

2.3.5 Discussion

During the investigation of recovery and extraction it was noticed that, using 1.0M sodium hydroxide, recovery agreed with previous workers but that some decomposition or change in the humic acid had occurred.

Using a weaker solution of sodium hydroxide less decomposition was noted but not all of the brown material eluted from the column.

These findings have been seen by Connor (64). He used infra-red spectrophotometry to confirm that there were differences between alkali extracted colour and that produced by ultra-filtration. The major differences that he found were:-

1. Increased unsaturation by dehydrogenation, alkene and alkyne and also formation of amide groups.
2. Ether formation
3. Primary amine formation
4. Formation of carboxyl groups
5. Increased conjugation of alkene and carbonyl groups
6. Silicate formation caused by the dissolution of clay minerals.

The overall aim of this project was to find a way to classify coloured waters in order to predict the results of their treatment. As the colour in peaty waters is mainly caused by humic and fulvic acids an elegant approach would have been controlled degradation and analysis followed by pattern recognition techniques. A great deal of time was spent in

trying to find reproducible oxidation and hydrolysis procedures using a purchased "standard" humic acid. Unfortunately, whilst degradation occurred and the products of that degradation were in general agreement with other workers, the reactions were not reproducible. This, of course, prevented the production of any results for pattern recognition analysis.

As the conclusion was being reached that degradation work was going to be unrealistic a further report was received from the Water Research Centre. This suggested that the reasoning for this project may not have been realistic. In their report, Connor and Stiff (65) found that coloured waters, thought previously to be untreatable, were able to be decolorised when treated under laboratory conditions. They have found that:-

1. Extraction and determination of trace humic materials from water was found not to be reproducible on a quantitative basis using either current or modified methods.
2. The problem of "untreatable" peaty waters does not, in fact, exist. The treatability relates to the operating conditions of the process rather than to the variable nature of the trace organic colouring matter in the water.

Their conclusion was that for any coagulant, dose pH is crucial. They reasoned that under practical conditions it would be reasonable to assume that coagulant and water could come into contact at other than the optimum pH. This could give rise to colour enhancement (complex formation) which resists treatment.

SECTION THREE

Analysis of Polycyclic Aromatic Hydrocarbons in a South Gwent Valley

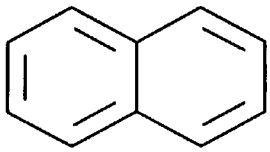
3.1.1 Introduction

Polycyclic aromatic hydrocarbons are recognised by the American Environmental Protection Agency as priority pollutants. This section aims to study the occurrence of these compounds in a semi-rural area of Gwent which includes the location of a large, toxic-waste incinerator.

3.1.2 Polycyclic Aromatic Hydrocarbons

Polynuclear hydrocarbons can be divided into two groups, those in which the rings are isolated e.g. biphenyl, diphenylmethane etc. and those in which two or more rings are fused together in the ortho position e.g.naphthalene, anthracene etc. This study is only involved with the second class.

Naphthalene is the lowest molecular weight member of the group. In 1866 Erlenmeyer proposed the symmetrical formula shown below. Graebe proved in 1869 that it did consist of two benzene rings fused together in the ortho position.



naphthalene

The highest molecular weight member of the class could be considered to be graphite, an allotropic form of carbon. However, only the environmentally mobile members of the group are considered in this study. These range from naphthalene (MW 128.16) to coronene (MW 300.36). Within this range is a considerable number of different compounds varying in the number and position of aromatic rings and substituents.

These compounds have been studied for over one hundred years and therefore many of them have been given unsystematic names. For example, naphthylene and pyrene reflect their coal-tar origin and coronene relates to its structural shape. The International Union of Pure and Applied Chemistry (IUPAC) has proposed a system of nomenclature which is generally used. The system uses the following rules:-

1. The structural diagram is constructed to present the greatest number of rings as possible in a horizontal row.
2. Horizontal and vertical axes are drawn through the centre of the horizontal row and the molecule is orientated to place the maximum number of rings in the upper right quadrant and the minimum number in the lower left.
3. Carbon atoms are numbered in a clockwise direction starting with the carbon atom that is not part of another ring and is in the most counter-clockwise position of the uppermost ring furthest to the right. Carbon atoms common to two rings are not numbered.

4. Ring faces (except those that cover two faces) are lettered alphabetically with the side between carbons one and two being labelled "a". Alphabetic order is then continued clockwise around the molecule.
5. In naming a compound formed by the addition of a component, numbers and letters are enclosed in brackets and are placed immediately after the name of the added component to describe the point of attachment or where a ring is fused to the face of a molecule. If a ring is fused to more than one face this is indicated using the appropriate letters to denote the faces involved.

3.1.3 Physical and Chemical Properties

All PAH's are solids at ambient temperatures. Their resonance energies are less than would be expected from the addition of their component benzene rings. Hence they are less aromatic than benzene and so will be more reactive.

Anthracene, for example, is very reactive in the 9,10 positions. Electrophilic reactions to form σ-complexes at the 1- or 2- position leave a "naphthalene fragment" with loss of 96.3 kJ $mole^{-1}$, whereas the σ-complex at position 9 leaves the two benzene rings intact and so the loss is 50.3 kJ $mole^{-1}$. Hence for PAH's with three or more rings, the σ-complex that contains the largest number of isolated benzene rings will be the most stable.

Most chemical reactions are by electrophilic substitution rather than by addition, e.g. naphthalene is very easily halogenated. It is also nitrated by nitric acid at room temperature to give 1-nitronaphthalene. At higher temperatures mixtures of 1,5 and 1,8 dinitronaphthalenes are formed.

Of importance during their analysis is that they are succeptable to photo-oxidation. This has been observed by Hellmann (66) who investigated the change in fluorescence intensity with time of PAH's adsorbed onto thin layer chromatography plates.

In air, PAH's are liable to reaction with nitrogen oxides and with ozone. In particular, the reaction with nitrogen oxides is liable to be very important as the nitrated PAH's are considerably more carcinogenic than the parent compounds.

All polyaromatic hydrocarbons have characteristic ultra violet / visible absorption spectra. Table 3.1 shows the absorption maxima of some typical PAH's. When substituents are present in the molecule the bands are usually shifted to longer wavelengths but the pattern is usually characteristic of the substituent group and its position in the nucleus.

Hydrocarbon	λmax nm	λmax nm	λmax nm
Benzene	184	204	254
Naphthalene	220	275	312
Phenanthrene	252	293	330
Anthracene	253	375	
Naphthacene	278	474	
Pentacene	310	580	

Table 3.1

3.1.4 Toxicity

Soot and coal tars were first suspected to be carcinogenic in the late eighteenth century when Sir Percival Pott observed that many of his patients who had cancer of the scrotum were chimney sweeps. Since then PAH's have been intensively scrutinised and the group contains some of the most studied chemical carcinogens. Many of the PAH's that are carcinogenic are derived from a benz(a)anthracene skeleton. Anthracene itself is not carcinogenic but benz(a)anthracene appears to have weak carcinogenicity (67). Addition of another benzene ring in select positions results in compounds with powerful carcinogenicity, such as dibenz(a,h)anthracene or benzo(a)pyrene. Indeed, benzo(a)pyrene is the most hazardous member of the group.

According to Barlow and Sullivan (68), benzo(a)pyrene itself is not a carcinogen but its metabolites, which bond to protein and DNA, are. Sims (69) found that in rat liver homogenates benzopyrene hydroxylase converts benzo(a)pyrene to 3-hydroxy benzo(a)pyrene, 1,2-dihydro-1,2-dihydroxy benzo(a)pyrene and 9,10-dihydro-9,10-dihydroxy benzo(a)pyrene.

A report by the International Agency for Research on Cancer (70) stated that 50 to 100 ppm of benzo(a)pyrene administered to mice in diet for 122 to 197 days produced stomach tumours in 70% of them. 250 ppm produced tumours in 100% of mice after 30 days. A single, oral administration of 100mg to nine rats produced mammary tumours in eight of them.

3.1.5 Formation of Polyaromatic Hydrocarbons

Most of the PAH's occur in coal-tar and can be extracted from this medium. They can be formed in any hydrocarbon combustion process and may be released from oil spills. The less efficient the combustion process the higher the PAH emission factor is likely to be. The major sources are stationary such as heat and power generation, refuse burning, industrial activities etc. (143)

It is believed that two distinct reaction steps are involved in the formation of PAH's by combustion. These are pyrolysis and pyrosynthesis. At high temperatures organic compounds can be partially cracked to smaller molecular fragments, mainly radicals. These recombine to yield the larger and more stable PAH molecules.

Possible natural sources of PAH's are forest fires, volcanoes and biosynthesis by algae and bacteria. Blummer and Youngblood (71) attributed the widespread distribution of PAH's in recent sediments to forest fires.

Ilnitsky et al (72) estimated that 12-14 tonnes of benzo(a)pyrene is produced annually from volcanoes. The Environmental Protection Agency has estimated (73) that 250 tonnes of benzo(a)pyrene are released into the atmosphere by the burning of fossil fuels in the USA, again, per annum.

3.1.6 Environmental Fate

When released to air PAH's may be subject to direct photolysis. They may also be removed by reaction with ozone and

nitrous oxides.

If released to water then, due to their non-polar nature, they will strongly bond to sediments and particulate matter. They will also bioconcentrate in organisms unable to metabolize them. PAH's are not expected to hydrolyse or significantly evaporate from soils and surfaces.

3.1.7 Permissable Concentrations

The US Environmental Protection Agency (EPA) has catalogued PAH's as one of the sixty five priority toxic pollutants (74). They found that there are insufficient data to propose a criterion for the protection of freshwater or of salt-water aquatic life. For the protection of human life they recommend a zero concentration. An additional lifetime cancer risk of 1 in 100,000 is posed by a concentration of 0.028μg/lt.

The 1970 WHO European Standards for Drinking Water recommends a PAH concentration not to exceed 0.2μg/lt. This is based on composite analysis of six PAH's; fluoranthene, benzo(a)pyrene, benzo(k)fluoranthene, benzo(b)fluoranthene and indeno(1,2,3-c,d)pyrene.

Microbial reclamation was used to restore land for use by general industry after a gas works had been dismantled. Bewley (75) reported a target "clean" concentration of 10mg/Kg however independent analysis by the Lancashire County Analyst indicated a mean PAH concentration of 148 +/- 13 mg/Kg.

3.1.8 Analysis of Polyaromatic Hydrocarbons

PAH's have been determined by a variety of techniques in a wide range of environmental samples. In 1953 Wedgewood and Cooper (76) investigated PAH's in industrial effluents and sewage. They extracted the PAH's into chloroform and separated them from most interfering compounds using an alumina column and fractional elution with cyclohexane. The final analysis and identification was by UV spectrophotometry.

Since that time rapid advances have been taken place in instrumental chromatographic methods. The first recorded use of gas chromatography was in 1964 (77). The technique has been used regularly since then and, with the addition of mass spectroscopy, has proved invaluable for the identification of compounds.

Tong et al (78) extracted PAH's from air particulate samples using dimethylsulphoxide. They needed to concentrate the extract, which would be difficult by normal means due to the high boiling point of the solvent. Instead they used semi-preparative HPLC to separate the DMSO extract into five fractions. A silica column was used with a gradient elution programme with n-hexane, dichloromethane and acetonitrile. The fractions were concentrated and subjected to both GC-MS and high resolution GC. They identified 108 compounds with molecular weights from 142 (2-methylnaphthalene) to 300 (coronene).

Santoni and Mandon (79) extracted PAH's from river water into hexane and from river sediment by ultrasonic extraction

into DMSO followed by extraction into hexane. Chromatography was on silica and final analysis by low temperature fluorometry. They could determine a single PAH in a mixture of six.

Mellone et al (80) used laser induced fluorescence in a graphite furnace as a screening method for PAH's. It was found to be suitable for screening crude oil, petroleum products and solid materials without sample pre-treatment.

Funk et al (81) extracted PAH's from water using cyclohexane. The extracts were evaporated to dryness redissolved in methanol and applied as a band to caffeine impregnated HPTLC plates. The plates were developed with iso-propyl ether / hexane (4:1). Detection was by fluorescence and recoveries between 90 and 110 %.

Baumeister et al (82) cleaned up toluene extracts of PAH's using cyanopropyl modified silica extraction columns. They eluted non-polar interferants with light petroleum and the PAH's with benzene / light petroleum (1:3). Final determination was by temperature programmed capillary GC. Recoveries were 90 +/-10% for a range of PAH's.

Kicinski et al (83) also used solid phase extraction columns to clean up PAH extracts. Soil samples were extracted by sonication with toluene. The columns contained C-18 silica and interfering compounds were removed by washing with water / propan-2-ol (9:1). PAH's were eluted with dichloromethane. They found better than 90% recovery for all compounds investigated.

Fernandez et al (84) extracted sediment by ultrasonication with dichloromethane. The resulting extracts were cleaned up with a column packed with both silica and alumina. The PAH fraction was eluted with 20% dichloromethane in hexane. This fraction was further cleaned up using gel permeation chromatography on Bio-beads SX-3 and SX-12. Final analysis was by capillary GC. Recoveries were between 52 and 78%.

Jinno and Niimi (85) were able to separate ten PAH's using supercritical fluid chromatography (SFC) with carbon dioxide mobile phase. Blilie and Greibrokk (86) also used SFC but used organic modifiers to the carbon dioxide mobile phase. This had the effect of reducing the retention and improving the peak shape for PAH's.

Muel and Saguem (87) determined 27 PAH's in Parisian atmospheric particulate matter. The air samples were passed through glass wool and the retained particulates extracted with cold chloroform utilising sonication. They performed a preliminary separation on a semi-preparative column of ODS silica. The same column was used for the analytical separation with increased sensitivity and selectivity being provided by the use of fluorescence detection.

In summary, there are three stages involved in the determination of PAH's in environmental samples. The first stage involves the extraction of the organic compounds from the matrix. This is usually accomplished using non-polar solvents either by Soxhlet extraction or with the aid of ultra-sonication.

As PAH's are invariably accompanied by a range of other organic compounds, both natural and pollutant, a clean-up procedure is usually required. This has been done by TLC and by classical column chromatography. However in recent years there has been a move towards "solid-phase extraction columns". A wide range of packing materials is available ranging from the usual silica to ODS and other bonded phases. This allows the clean-up procedure to be tailored to suit both analyte and matrix.

The final stage involves the final separation and detection. Although attempts have been made to use SFC and low-temperature fluorescence, these have had only limited success to date. The majority of separations have been obtained using capillary GC with flame-ionisation detection. Mass spectroscopy has also been used for detection with the added advantage of some identification possibility, although it cannot differentiate between isomers.

More recently, HPLC has been used as the final chromatographic step. When combined with fluorescence detection this gives excellent sensitivity. The main drawback with fluorescence, however, is that it must be optimized for each PAH of interest in order to maximise its response. A similar problem exists with UV detection, which is also at least one order of magnitude less sensitive that fluorometry.

UV diode array detectors offer an answer to the variation of absorption maxima between compounds. The most recent generation of these instruments show a sensitivity matching

that of the most sensitive single wavelength detectors.

3.1.9 High Performance Liquid Chromatography

High Performance Liquid Chromatography (HPLC) is possibly the analytical technique most widely used to-day. It developed from the classical low pressure, open-column procedures when instrument manufacturers produced high pressure, low dead-volume equipment with sensitive detectors. Liquid chromatographic theory is not very different from that of GC, which has been well understood since the early 1960's. This theoretical knowledge led to the development of the first high performance packings for HPLC. The technique was first introduced to a wide audience in 1969 at the Fifth Annual Symposium on Advances in Chromatography (88).

The two main separation techniques used in HPLC are adsorption, or normal phase, and reverse-phase. Adsorption chromatography typically utilises a silica column with a non-polar eluent. Using these conditions non-polar compounds elute more quickly than more polar components. This type of system has been used for the separation of PAH's (89) but not as frequently as reversed-phase procedures. The majority of separations reported in the literature utilise reverse-phase techniques.

Reverse-phase packings consist of a micro-particulate silica backbone with non-polar groups bonded to it. The most common bonded phase consists of an octadecyl chain bonded to the silica. Other common packings include octyl, phenyl and cyano. Retention for a given solute increases with chain length

of the packing which has been demonstrated by Karch et al (91). Hemetsberger et al (90) have shown a possible relationship between k' values and the % concentration of organic coverage, independent of chain length.

The k' value is known as the capacity factor and is a very convenient parameter in liquid chromatography. It is defined as the ratio of the quantity of solute in the stationary phase to that in the mobile phase and will remain constant for a given solute/column/mobile-phase system regardless of flow-rate or column length. The factor may be conveniently calculated by

$$k' = \frac{t_r - t_0}{t_0}$$

where t_r is retention time of solute

t_0 is the elution time of an unretained peak

For reverse-phase chromatography the mobile-phase, or eluent, consists of a mixture of water and organic modifier. The organic solvent is usually acetonitrile, methanol or, less frequently, tetrahydrofuran. Increase in the amount of organic modifier reduces solute retention.

Changes in pH of the mobile phase will have a significant effect on the retention and selectivity of ionized or ionizable compounds since charged molecules are distributed preferentially into the aqueous phase. To effect retention of charged compounds either the ionization must be suppressed by buffering the mobile phase or some form of ion-pairing must be

utilised. Typical ion-pairing reagents are octyl sulphonic acid, buffered to pH 7.5, which will pair with bases such as amines, or tetra-n-butyl ammonium hydroxide which is used to pair with organic acids.

As PAH's are non-ionizable compounds their retention is controlled by amount and nature of organic modifier. The series of environmentally important PAH's cover a wide range of molecular weights. An isocratic separation i.e. one using a single mobile phase, will either fail to separate early eluting peaks or will take a long time with concomitant peak broadening and loss of sensitivity. This problem can be overcome by the use of "gradient elution". This requires the use of two mobile phases. One is weak in organic modifier and the other strong. The two solvents are blended prior to introduction to the column. High pressure systems use two pumps whose flow-rates are controlled electronically by a "gradient former". Low pressure systems use proportioning valves to give the same effect.

Kirkland (92) has demonstrated the improved chromatography obtained by the use of gradient elution. He separated eleven PAH's ranging from benzene to benzo(a)pyrene using methanol / water on a C-18 modified silica. Using an isocratic system of 87.5% methanol the separation took fifteen minutes and significant peak broadening took place. Using a simple linear gradient of 62.5 to 100 % methanol at 6% per minute the same separation only took seven minutes and gave rise to much sharper peaks.

3.1.10 Validation of Analytical Procedures

The validation of the analytical procedure is an essential step in the development of methods for the determination of any analyte. Validation may be described as the demonstration that the procedure does what it is purported to do. Validation requirements can vary with the needs of the analytical technique and can include such sections as:-

1) Linearity: This demonstrates how the response of the detection system varies with change of concentration of the analyte. If the response is linear across the range tested then a single point calibration may be used. If the response is non-linear then a calibration curve, using a range of standards, will need to be constructed.
2) Sensitivity: This is defined as the slope of the analytical calibration function. Normally, the greater the slope then the greater the sensitivity.
3) Limit of detection: This is defined by IUPAC (94) as "the limit of detection, expressed as the concentration c_l or the quantity q_l, is derived from the smallest measure x_l that can be detected with reasonable certainty for a given analytical procedure."

Assuming normal distribution of errors, a detection limit can be calculated as:

$$L_d = \mu_{bl} + k_d \sigma_{bl}$$

Where:

L_d = detection limit

μ_{bl} = signal from blank

k_d = constant

σ_{bl} = std dev of blank measurement

There are a number of interpretations of this equation but one which is convenient sets k_d equal to 3 for detection limit and equal to 6 for limit of determination.

At low concentrations σ_{bl} will be very similar to σ_{sam}. Thus if a blank chromatogram is subtracted from that of the standard then a realistic estimation of limit of determination is

$$L_d = 6 \times \sigma_{ref}$$

where σ_{ref} is the standard deviation of a low point on the calibration curve for the analyte of interest.

3.1.11 The Investigation

The investigation of polyaromatic hydrocarbons can be broken down into three sections:-

1) **The development and validation of analytical procedures.** The extraction and clean-up steps were closely based on existing, well established techniques. HPLC was chosen as the final analytical step. This was partly due to the availability of a very sensitive diode array detector. This allowed the use of simple UV detection which could be optimized for each analyte of interest. An added bonus was

the ability to confirm the identity of the peaks detected by comparison with standard spectra.

2) **Survey work.** A series of soil samples were taken in an area of approximately 4 Km^2 NNE of Cwmbran, Gwent. The area is predominantly rural but contains several, significant industrial concerns, at least one of which could be a potential source of PAH's. A total of twenty three sites were sampled and a large number of potential sources included e.g.road-sides, oil-fired central heating, industrial sites etc.

3) **Multi-variate analysis.** A series of techniques was applied to attempt to determine any relationship between individual PAH content and sample site. If any relationships appeared then emission sources may be investigated.

3.2 DEVELOPMENT OF ANALYTICAL METHODS

3.2.1 Introduction

An analytical system encompassing extraction, clean-up and separation with quantitation was developed to investigate the levels of twelve PAH's present in soil samples collected from a semi-rural area near Cwmbran, Gwent.

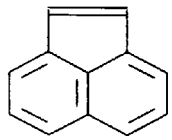
The extraction procedure utilised dichloromethane solvent with the aid of ultra-sonication. Clean-up was attained using solid-phase extraction based on the procedure described by Kicinski et al (83).

Identification and quantitation was completed using a validated HPLC procedure with external standards. Identification was proved by spectral comparison utilising a diode array detector.

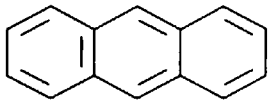
Kicinski found recoveries of between 70 and 90% for the PAH's of interest. These were assumed to be constant so no corrections were made. This was considered valid as absolute values were not the issue in this study, rather a comparison of levels found.

3.2.2 Compounds Investigated

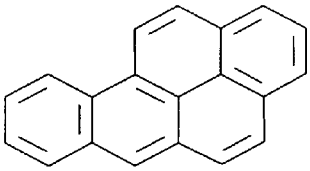
A series of purified reference standards was purchased from Aldrich Chemical Company (Gillingham, Dorset). The compounds and their structures are shown below.



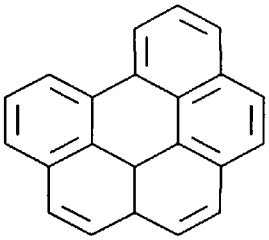
Acenaphthylene



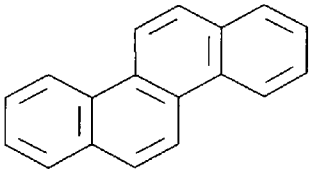
Anthracene



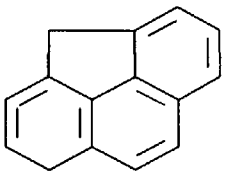
Benzo(a)pyrene



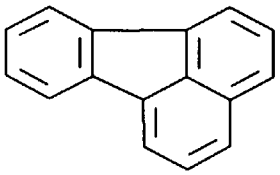
Benzo(ghi)perylene



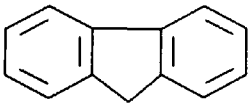
Chrysene



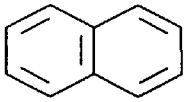
4H-cyclopenta(def)phenanthrene



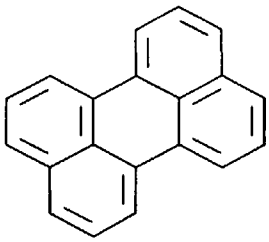
Fluoranthene



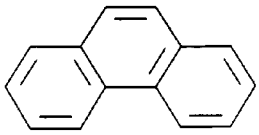
Fluorene



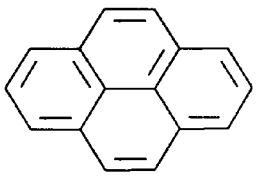
Naphthalene



Perylene



Phenanthrene



Pyrene

3.2.3 Extraction

The soil samples were all stored in amber-glass bottles at 4°C until required for analysis at which time they were dried under vacuum at room temperature using an air bleed. The

samples were protected from light during the drying process, which lasted for several days.

The non-polar compounds, including PAH's, within the soils were extracted using the following technique: Dried soil was ground using a pestle and mortar. Stones and vegetation were removed. 5.0 grammes of sample was transferred to a 50ml centrifuge tube, 15ml of dichloromethane added and the mixture agitated. The tube was transferred to an ultrasonic bath and sonicated for 15 minutes with mechanical agitation every 5 minutes. On completion, the mixture was centrifuged for two minutes and the supernatant aspirated into a beaker. The process was repeated a further four times, bulking the dichloromethane extracts. The combined extracts were filtered through a 9.0 cm Whatman GF/A filter and evaporated to dryness at room temperature under a stream of nitrogen and protected from light.

3.2.4 Clean-up

Clean-up was performed using solid-phase extraction techniques. The method used was that reported in the Baker Applications Guide (93). The procedure, as reported, refers to extraction of PAH's from water but can be equally well applied to the extracted samples obtained from soils.

The sample residue was dissolved in 10 ml of propan-2-ol using gentle warmth and sonication. The resulting mixture was diluted to 100 ml with distilled water.

An octadecyl bonded silica column was used (500mg) from

Analytichem International (Jones Chromatography, Hengoed, Mid.Glam). The column was conditioned with one column volume of methanol followed by two column volumes of 10% propan-2-ol in water. After the methanol addition the column was not allowed to become dry until all the sample had been applied. It should be noted that propan-2-ol was added to the aqueous solutions to prevent adsorption of PAH's to glass or plastic surfaces.

The prepared sample was aspirated through the column, using vacuum, at a rate of approximately 6 to 7 mls/min. When all of the sample solution had been added the beaker was rinsed with 5 ml of 10% propan-2-ol and the washings aspirated through the column.

The column was sucked dry under vacuum for 15 minutes and finally the retained, non-polar compounds were eluted, using two 500μl aliquots of dichloromethane, into a 2ml volumetric flask. The solvent was removed at room temperature under a stream of nitrogen and the residue redissolved in 0.5 ml of propan-2-ol. 0.5 ml of methanol was added followed by 0.6 ml of water and finally the mixture was diluted to volume with methanol. Samples were transferred to amber glass autosampler vials for HPLC analysis. If insoluble particles were present then mixture was first filtered through a Gelman Acrodisk to protect the HPLC system.

3.2.5 Development of HPLC Procedure

An HPLC procedure was required to separate the twelve PAH's described in 3.2.2. Much of the work reported in the

literature has utilised reversed-phase systems consisting of a C-18 column with acetonitrile / water mobile phase, so this type of system was tried first. The best conditions were found to be as follows:-

Column: 125 x 4 mm Superspher RP-18. 4μm (E.Merck)

Solvent A: Acetonitrile / water 10 / 90

Solvent B: Acetonitrile / water 90 / 10

Gradient:

Time	%B
0 mins	65
8	65
12	100 (linear gradient)

Flow: 1.5 mls/min

Detection: Diode array with monitoring at 250 nm

Injection: 50 μl loop

Good separation of the twelve PAH's was obtained within a twenty minute run-time therefore validation work was started. Unfortunately, during this period, a severe shortage of acetonitrile came about and it was necessary to make a fresh start using methanol as the organic modifier. Although acetonitrile is normally preferred for HPLC due to its UV transparency and low viscosity, in this case a better separation was found to be possible using methanol. The final conditions were:-

Pre-column: 4 x 4 mm Lichrosorb RP-18 5μm (E.Merck)

Column: 124 x 4 mm Superspher RP-18 4μm (E.Merck)

Solvent A: Methanol / water 50 / 50

Solvent B: Methanol / water 95 / 5

Gradient:

Time(mins)	% B
0	60
20	80
23	80
24	90
29	90
30	60

Flow-rate: 1.75 mls / min

Detection: Diode array monitored at 250 nm

Injection: 50μl

The hardware consisted of two CM3200 pumps, an SM5000 diode array detector, a PROMIS II autosampler and a Thermochrom data handling system. All from LDC Analytical, Stone, Staffs.

The system described above produced a standard chromatogram for the twelve PAH's as shown in Fig.3.2.1

The isogram obtained from the same system is shown in Fig.3.2.2. It demonstrates the variation in UV spectra obtained for the series of compounds under investigation. The individual spectra were obtained post-run and are shown below in Figs.3.2.3 to 3.2.14. They all show maxima at different wavelengths however the use of just five detection wavelengths will give near optimum response. The final wavelengths chosen were 220, 230, 240, 250 and 265 nm. The sensitivity for each analyte will be demonstrated during the validation process.

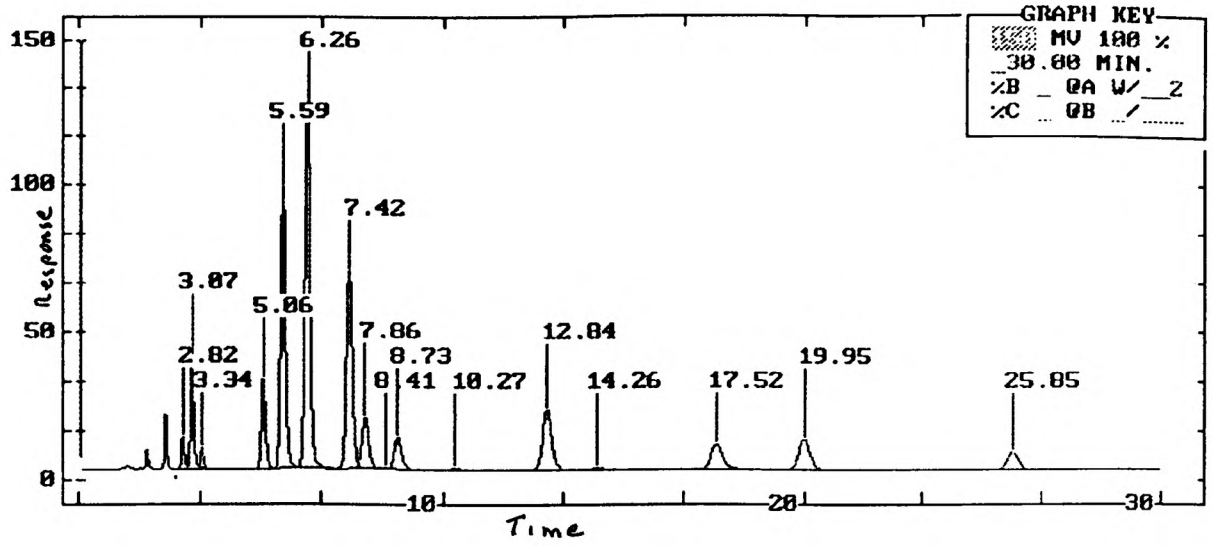


Fig.3.2.1

Chromatogram of Standards at 250nm

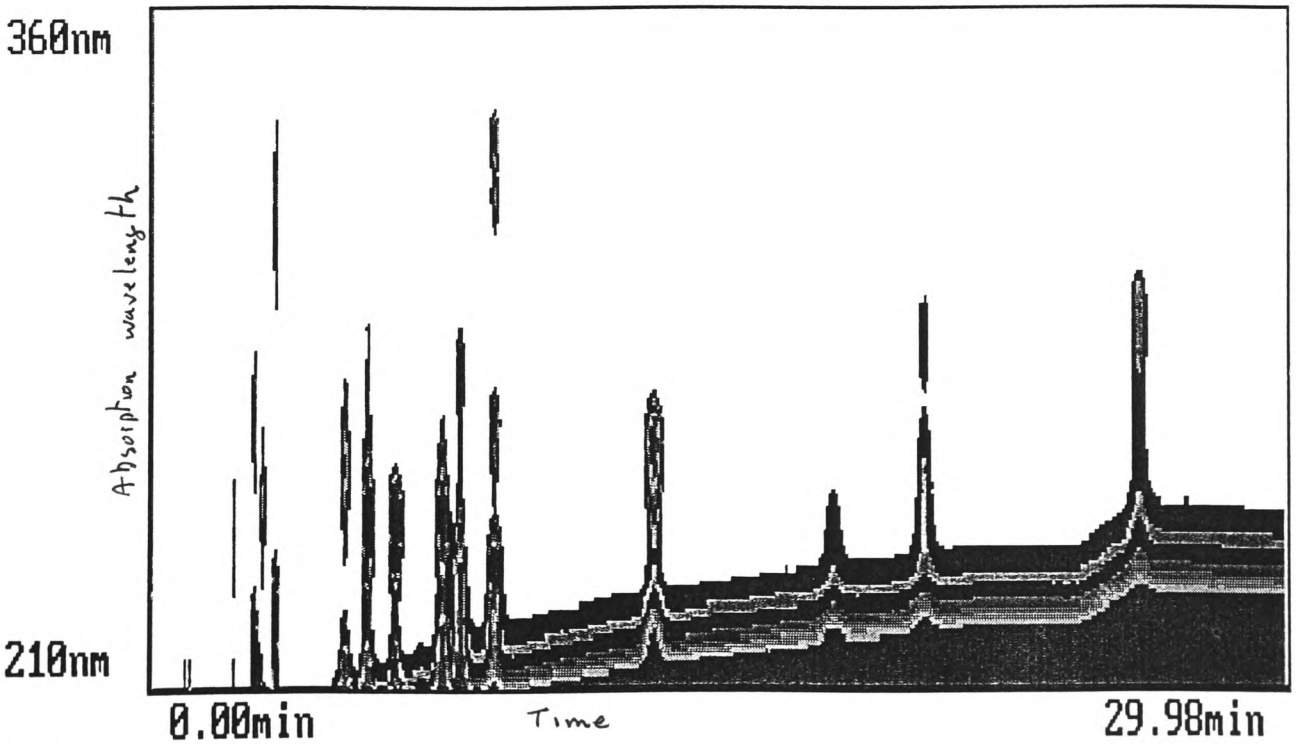


Fig. 3.2.2

Isogram of Standards

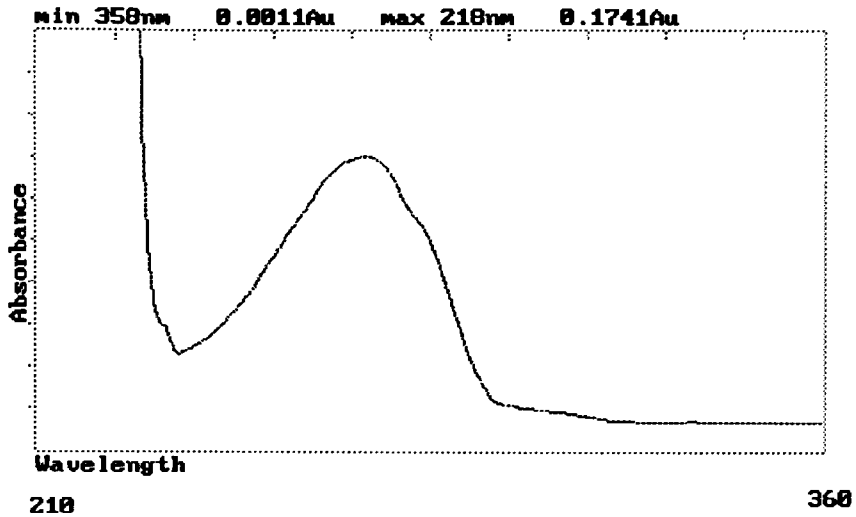


Fig.3.2.3

UV Spectrum of Naphthalene

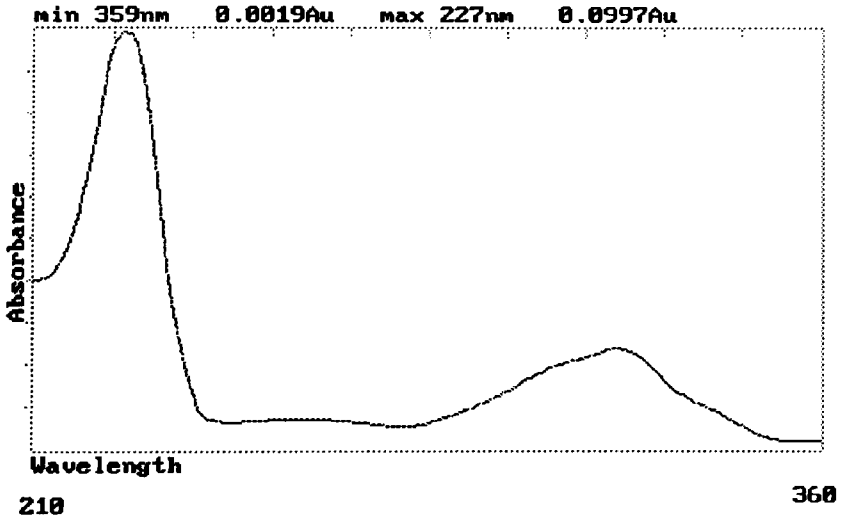


Fig.3.2.4

UV Spectrum of Acenaphthylene

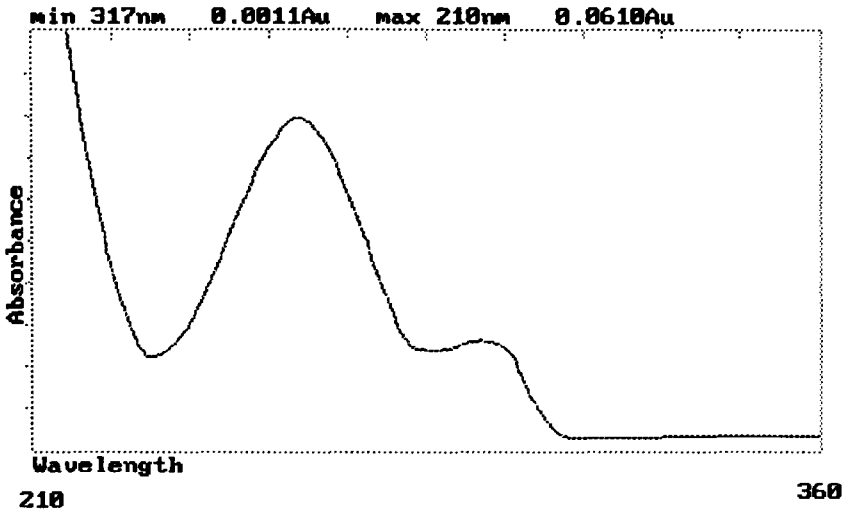


Fig.3.2.5

UV Spectrum of Fluorene

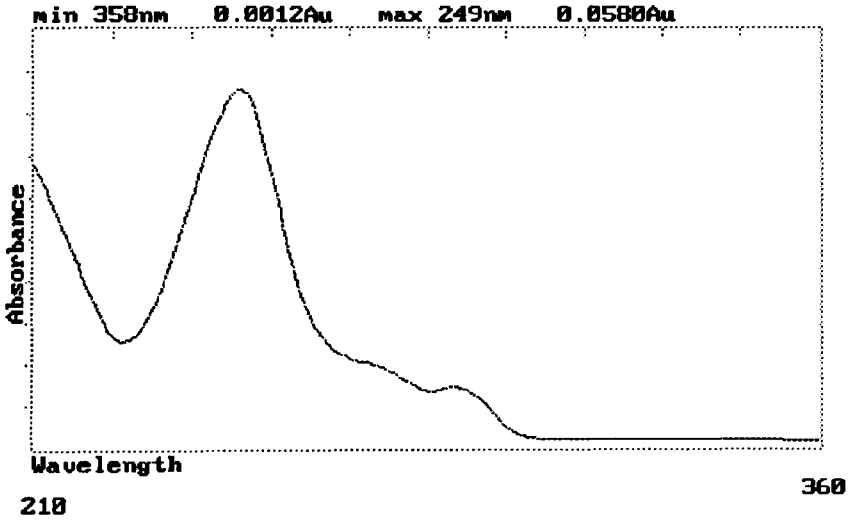


Fig.3.2.6

UV Spectrum of Phenanthrene

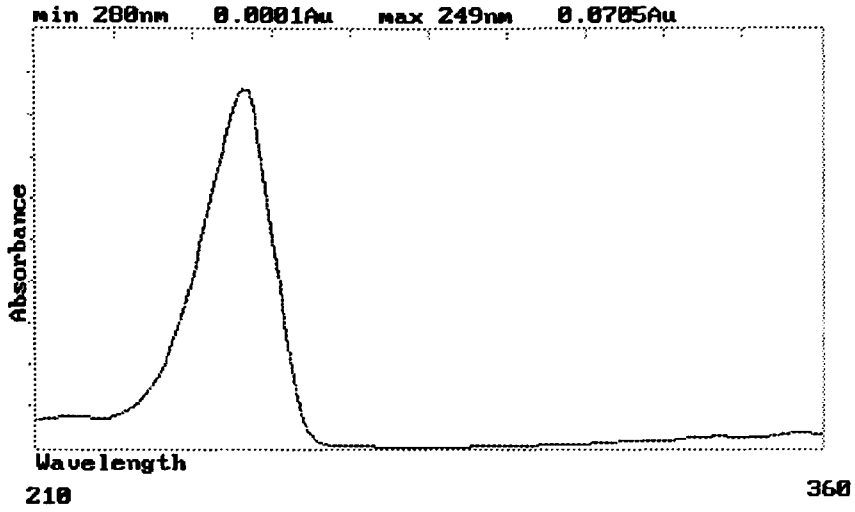


Fig.3.2.7

UV Spectrum of Anthracene

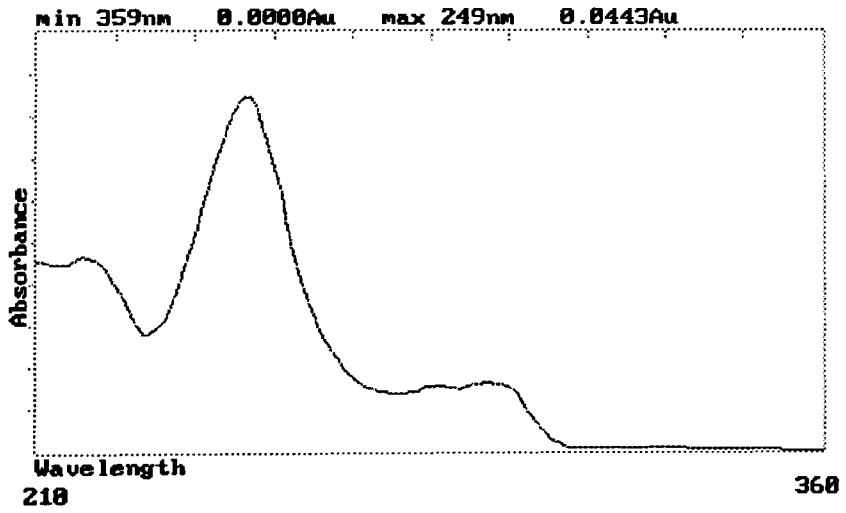


Fig.3.2.8

UV Spectrum of 4-H-cyclopenta(def)phenanthrene

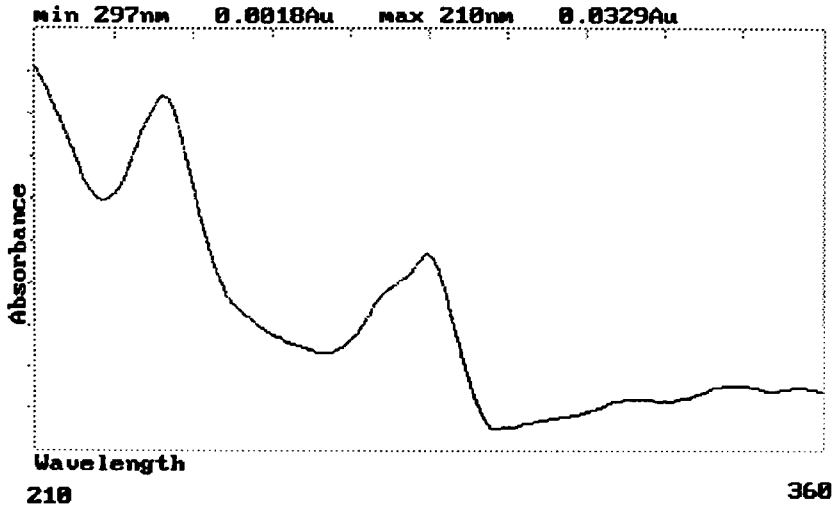


Fig.3.2.9

UV Spectrum of Fluoranthene

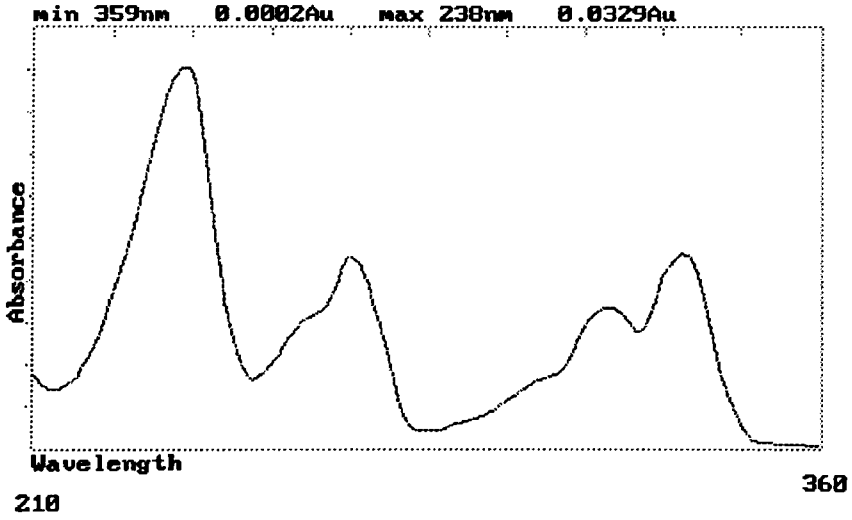


Fig.3.2.10

UV Spectrum of Pyrene

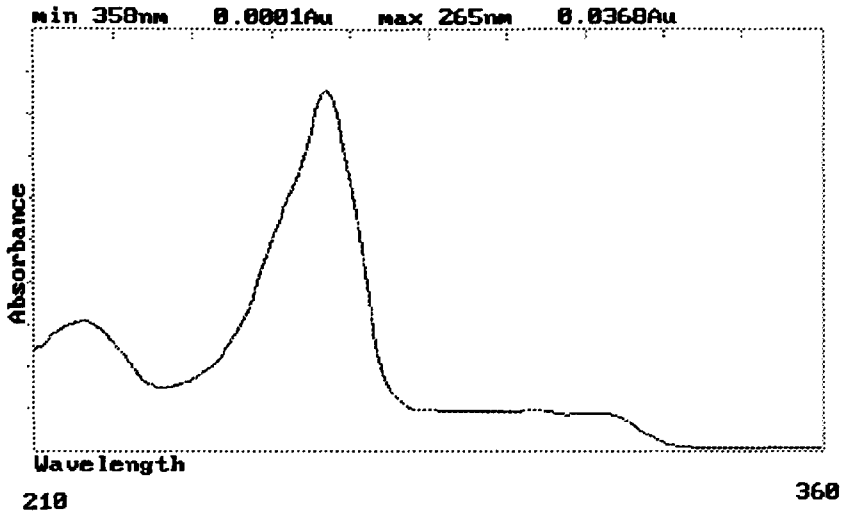


Fig.3.2.11

UV Spectrum of Chrysene

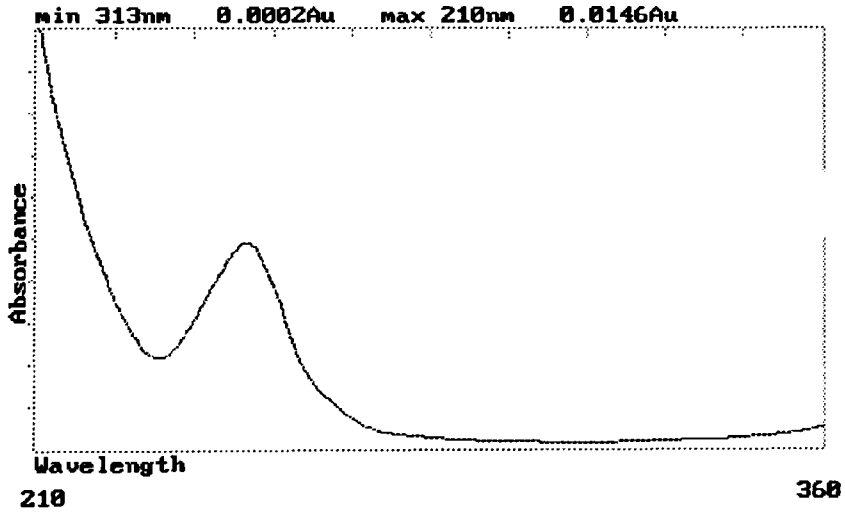


Fig.3.2.12

UV Spectrum of Perylene

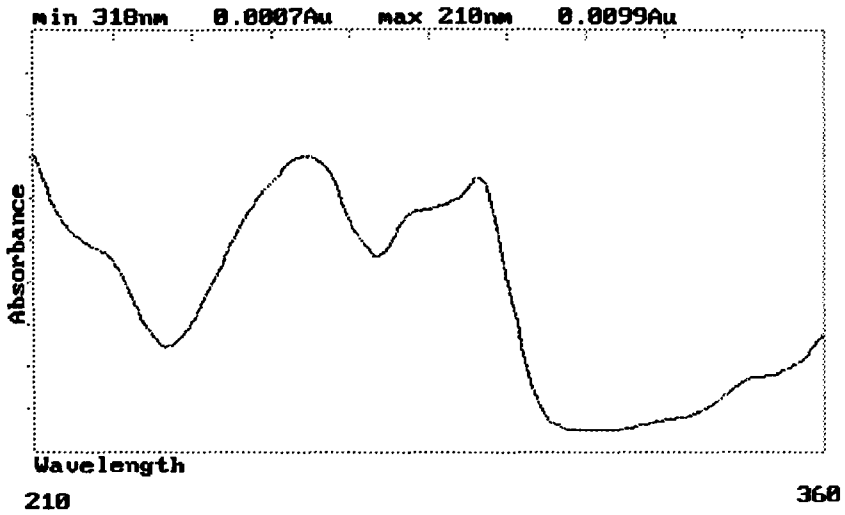


Fig.3.2.13

UV Spectrum of Benzo(a)pyrene

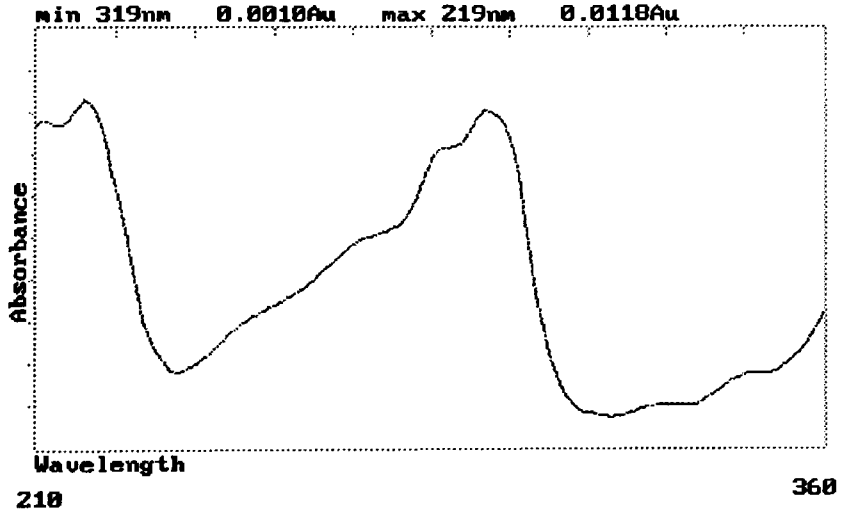


Fig.3.2.14

UV Spectrum of Benzo(ghi)perylene

3.3 VALIDATION OF ANALYTICAL METHODS

3.3.1 Linearity and Sensitivity

Approximately 5mg of each PAH standard was accurately weighed and transferred to a 100ml dark-glass, volumetric flask where they were dissolved in 10ml of dichloromethane. The resulting solution was diluted to volume with methanol. Aliquots of stock solution were further diluted with 80% methanol as shown below, giving rise to seven-point linearity plots.

Each injection could be investigated post-run and chromatograms were produced at five wavelengths. The first injection was of 80% methanol in order to produce a "gradient blank" which could also be integrated post-run. The blanks could be subtracted from the samples electronically to give true sample chromatograms free from "ghost peaks" which often occur during gradient elution. This process also served to produce a flat base-line.

The regression parameters were calculated for each wavelength and that showing the greatest slope was chosen as the detection wavelength for that compound. The regression plot for the optimum wavelength for each compound are shown in Figs.3.3.1 to 3.3.12.

One problem with linear regression is that points near to the origin can be significantly removed from the best-fit line but do not significantly effect the correlation coefficient. In order to show that the response is linear across the range

tested, a second plot, known as a response plot, is made. The X-axis is the weight of analyte on column and the Y-axis is area count / weight injected i.e. response. For linearity this should be a horizontal line within given precision limits. The plots are shown in Figs.3.3.13 to 3.3.24

Sample weights				Dilutions	
Compound	wt.(mg)				
Naphthalene	5.42				
Acenaphthylene	5.64				
Fluorene	5.22		No.	Aliquot	Volume
Phenanthrene	4.83				
Anthracene	5.81		1	0.5	250.0
4-H-CPPhenanthrene	5.30		2	0.5	100.0
Fluoranthene	4.85		3	1.0	50.0
Pyrene	4.50		4	1.0	25.0
Chrysene	3.96		5	1.0	10.0
Perylene	3.85		6	1.0	5.0
Benzo(a)pyrene	3.63		7	2.0	5.0
Benzo(ghi)perylene	4.25				

3.3.2 Results

a) Naphthalene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	5.42	8146	552	nd	252	680
2	13.55	20354	1484	328	565	1403
3	54.2	80899	5779	1307	2396	4837
4	108.4	159122	11051	2611	4752	9626
5	271	392901	26884	6481	14666	36530
6	542	749283	51234	15959	29403	53566
7	1084	over	97835	26048	47828	107641
	I'cept	5678	1045	250	759	-377
	Slope	1384	90	25	45	100
	Corr	0.9994	0.9991	0.9869	0.9867	0.9999

b) Acenaphthylene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	5.64	2707	3600	420	nd	275
2	14.1	6170	9013	970	471	560
3	56.4	25143	36033	3927	1694	2244
4	112.8	49403	71357	7888	3661	4550
5	282	125236	180460	19771	9631	11335
6	564	251417	357926	39300	19125	22864
7	1128	502812	698704	77676	38222	45633
	I'cept	-228	2067	82	-13	-10
	Slope	446	621	69	34	40
	Corr	1.0000	0.9998	0.9999	1.0000	1.0000

c) Fluorene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	5.22	1116	255	536	1210	1339
2	13.05	2702	697	1334	2897	2671
3	52.2	10929	2706	5250	11465	14343
4	104.4	22275	5715	10255	22953	28361
5	261	56831	14589	26847	58565	72310
6	522	114310	29266	53656	117697	146120
7	1044	227622	58262	107057	239928	293790
	I'cept	-203	-71	-94	-182	-535
	Slope	218	56	103	225	282
	Corr	1.0000	1.0000	1.0000	1.0000	1.0000

d) Phenanthrene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	4.83	1480	1064	2839	3793	882
2	12.08	3163	2522	6651	9460	2598
3	48.32	8598	10072	26984	38280	10442
4	96.6	25479	19866	53532	75828	20760
5	241.5	64924	50759	136272	193098	52996
6	483	130527	102239	27570	389662	106622
7	966	259163	203147	544117	774017	213588
	I'cept	-940	-64	-36	-371	-307
	Slope	270	211	564	803	221
	Corr	0.9997	1.0000	1.0000	1.0000	1.0000

e)Anthracene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	5.81	684	1134	4103	895	nd
2	14.53	1742	2777	10586	20435	326
3	58.1	6751	11297	41879	81579	1161
4	116.2	13376	22405	83078	160928	2345
5	290.5	33547	26726	210270	405413	5649
6	581	67794	114005	419236	800724	11240
7	1162	135654	227814	815898	1521494	21980
	I'cept	-91	-143	2324	9826	124
	Slope	117	196	704	1315	19
	Corr	1.0000	1.0000	0.9998	0.9993	0.9999

f)4-H-cyclopenta(def)phenanthrene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	5.3	1865	1171	2181	3345	734
2	13.25	4608	2777	5382	8376	2533
3	53	18219	11002	21195	33360	9783
4	106	36399	22092	42305	66700	19420
5	265	93135	56291	108592	170929	50491
6	530	187852	113769	218298	344383	101316
7	1060	375132	226541	433763	704006	201760
	I'cept	-444	-208	-216	-2558	-246
	Slope	354	214	410	664	191
	Corr	1.0000	1.0000	1.0000	0.9999	1.0000

g) Fluoranthene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	4.85	1717	2212	1747	865	370
2	12.13	4247	5316	4179	2140	1652
3	48.5	17064	21422	16793	8618	6445
4	97	34632	43998	33391	17638	12710
5	242.5	90117	111404	89821	47015	34568
6	485	181386	228220	178188	95036	70598
7	970	363305	454197	351906	174436	141934
	I'cept	-802	-920	-170	1032	-714
	Slope	375	469	364	182	147
	Corr	1.0000	1.0000	0.9999	0.9980	0.9999

h) Pyrene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	4.5	687	2022	2934	520	818
2	11.25	1782	5046	7335	1547	2950
3	45	6826	19981	29221	5916	11372
4	90	13548	40228	58018	11593	22433
5	225	36373	103745	150482	31665	59668
6	450	73593	209968	300183	64038	119673
7	900	145134	416836	584322	127098	237111
	I'cept	-226	-505	1117	-344	-322
	Slope	162	464	652	142	264
	Corr	0.9999	1.0000	0.9998	0.9999	0.9999

i)Chrysene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	3.96	1080	1015	618	1463	3608
2	9.9	3590	2175	2116	4015	10008
3	39.6	13391	7419	7921	15910	40245
4	79.2	26551	14638	15405	31143	79333
5	198	6844	37342	39282	79723	202479
6	396	136959	74739	80220	161381	404520
7	792	271225	151275	160244	317072	781260
	I'cept	73	-135	-258	136	2481
	Slope	343	191	203	402	990
	Corr	1.0000	0.9999	1.0000	0.9999	0.9997

j)Perylene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	3.85	1500	nd	889	1444	nd
2	9.63	2493	1024	1703	3097	804
3	38.5	12194	5355	7958	13056	3226
4	77	23872	10671	16777	25833	6312
5	192.5	60440	27296	40122	66001	16312
6	385	119127	55290	81073	135965	31376
7	770	236391	109592	161049	267319	62822
	I'cept	360	-194	78	-256	155
	Slope	307	143	209	348	81
	Corr	1.0000	1.0000	1.0000	0.9999	0.9999

k) Benzo(a)pyrene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
1	3.63	nd	4697	1717	1511	1044
2	9.08	nd	10924	3078	2956	3557
3	36.3	12617	14659	7566	12395	14063
4	72.6	21072	20904	15013	2210	28845
5	181.5	60172	30742	38760	62409	73934
6	363	115830	61859	68152	126052	149218
7	726	207847	112411	16356	250060	298383
	I'cept	4811	7879	2767	-112	-633
	Slope	285	144	173	345	412
	Corr	0.9957	0.9955	0.9970	1.0000	1.0000

l) Benzo(ghi)perylene

No.	wt.on			Area counts		
	col(ng)	220	230	240	250	265
21	4.25	2232	766	455	558	318
2	10.63	4967	1402	791	1243	1864
3	42.5	19100	7296	4493	6465	8413
4	85	37116	14121	11215	12462	16866
5	212.5	95312	35968	27299	32023	43688
6	425	197693	72276	45700	64912	88469
7	850	393357	145105	92718	131826	177659
	I'cept	-891	-221	853	-498	-636
	Slope	464	171	108	155	210
	Corr	0.9999	1.0000	0.9974	0.9999	1.0000

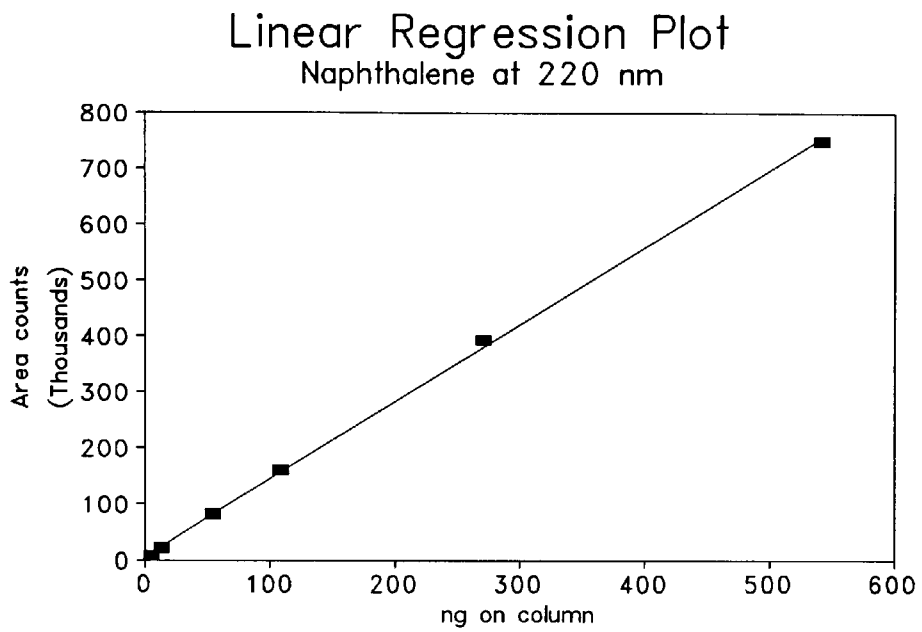


Fig.3.3.1

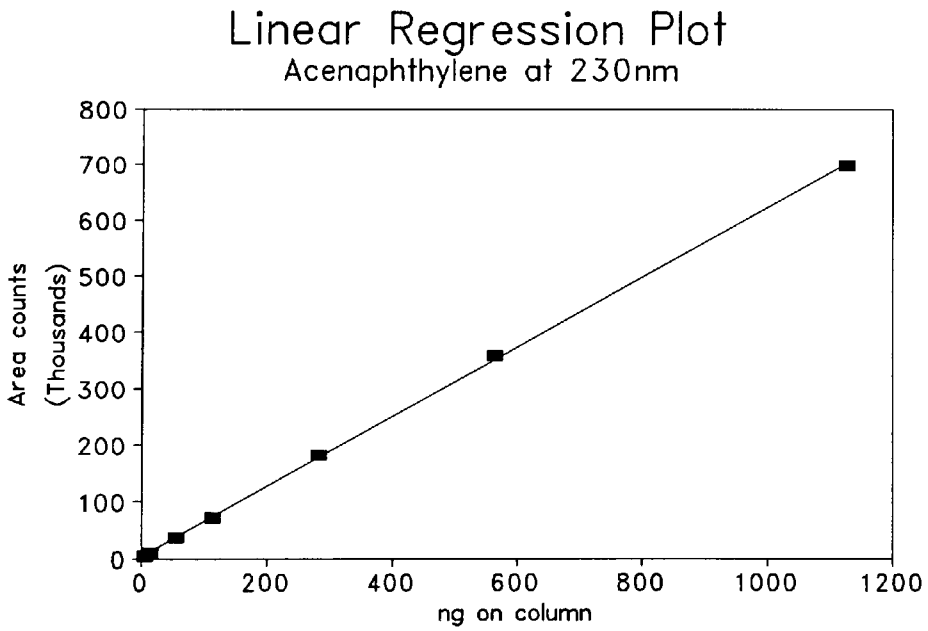


Fig.3.3.2

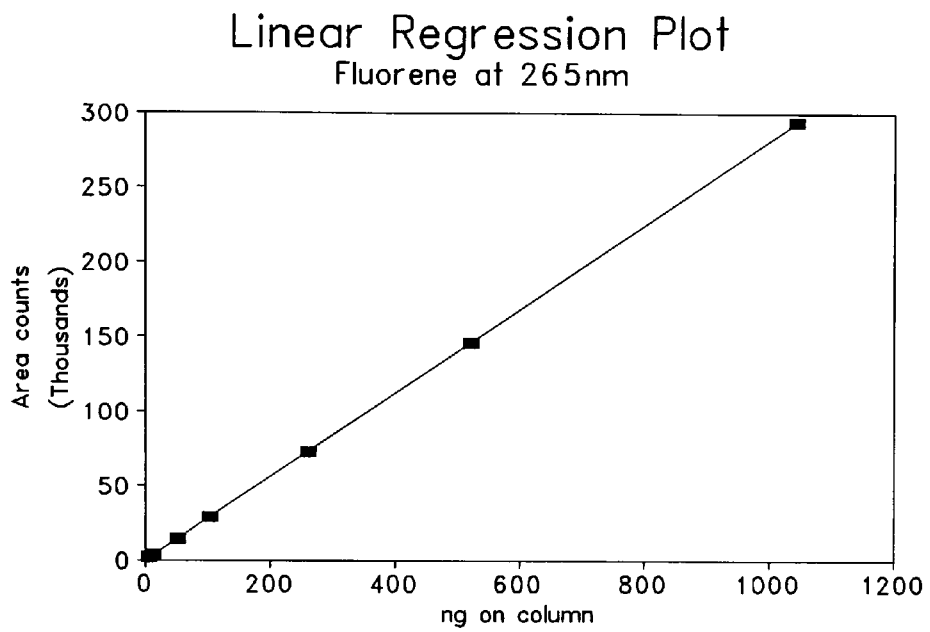


Fig.3.3.3

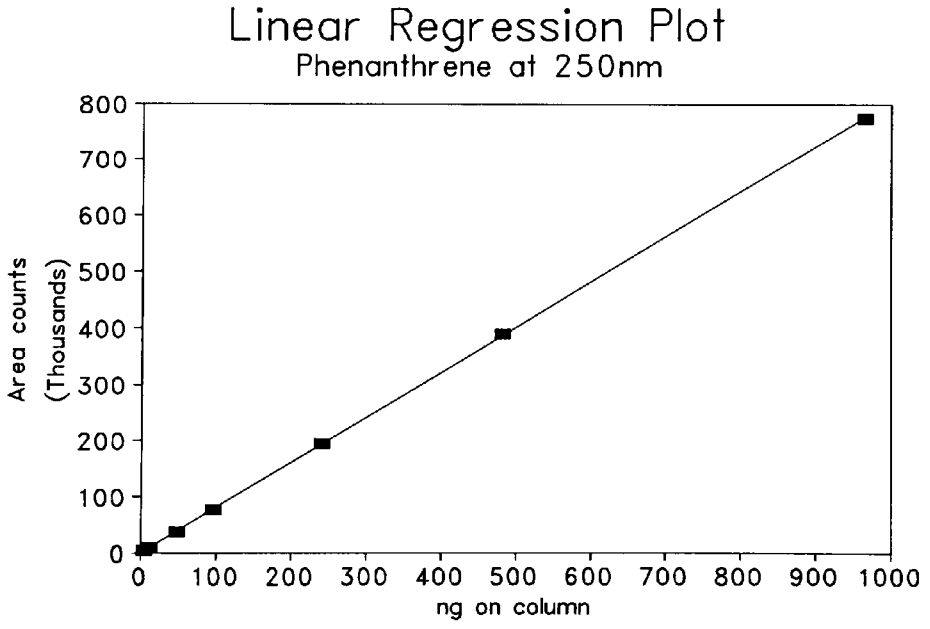


Fig.3.3.4

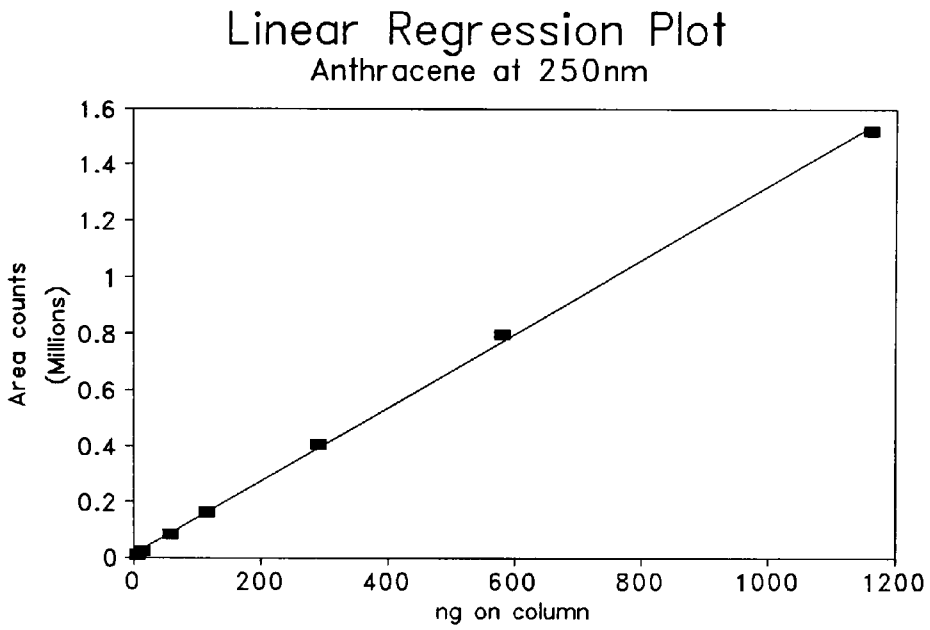


Fig.3.3.5

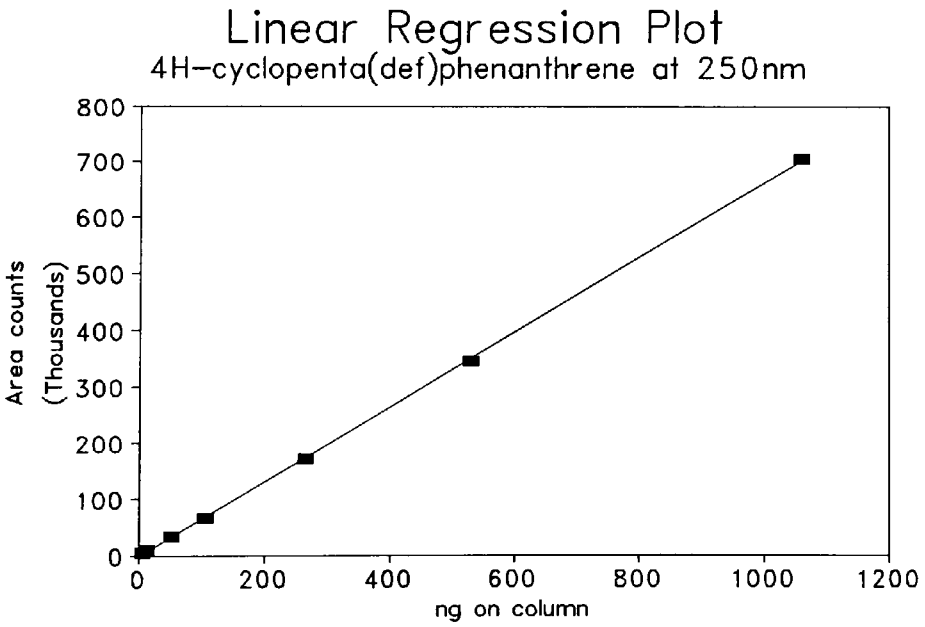


Fig.3.3.6

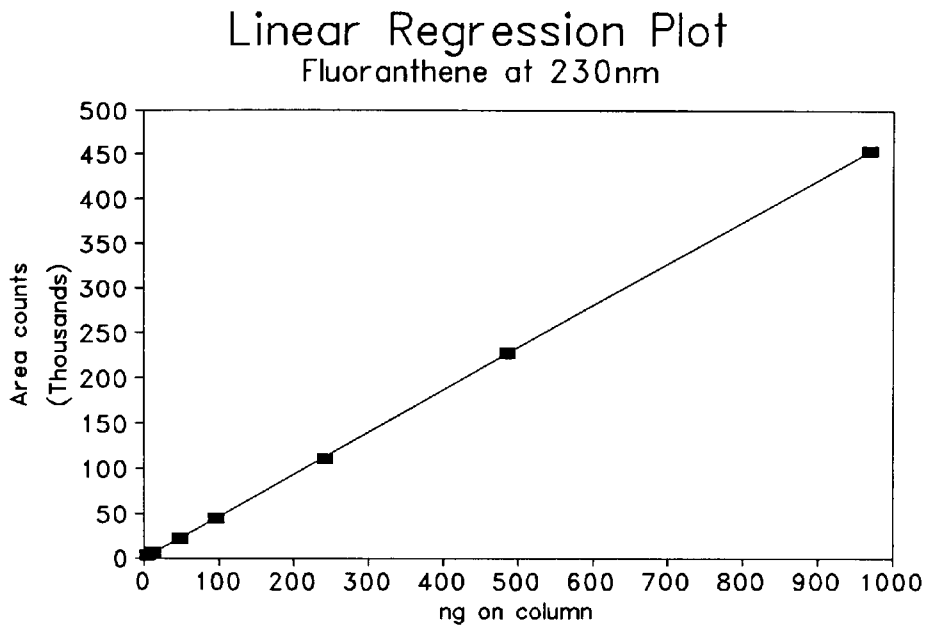


Fig.3.3.7

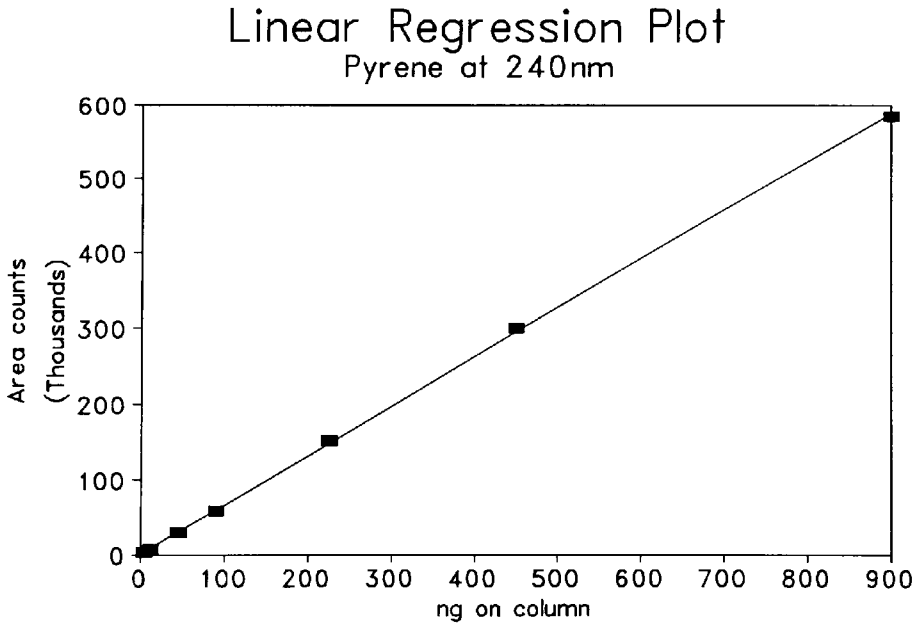


Fig.3.3.8

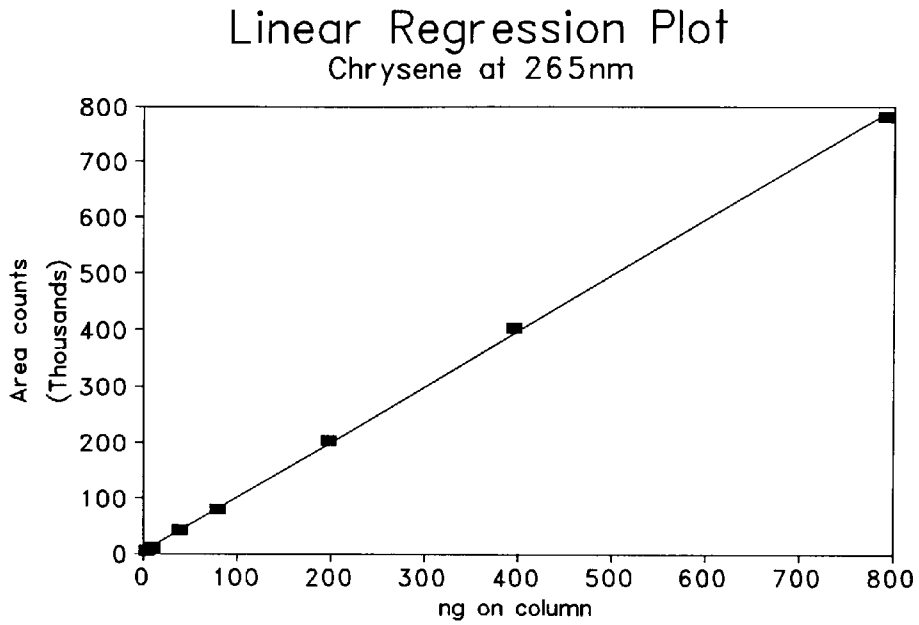


Fig.3.3.9

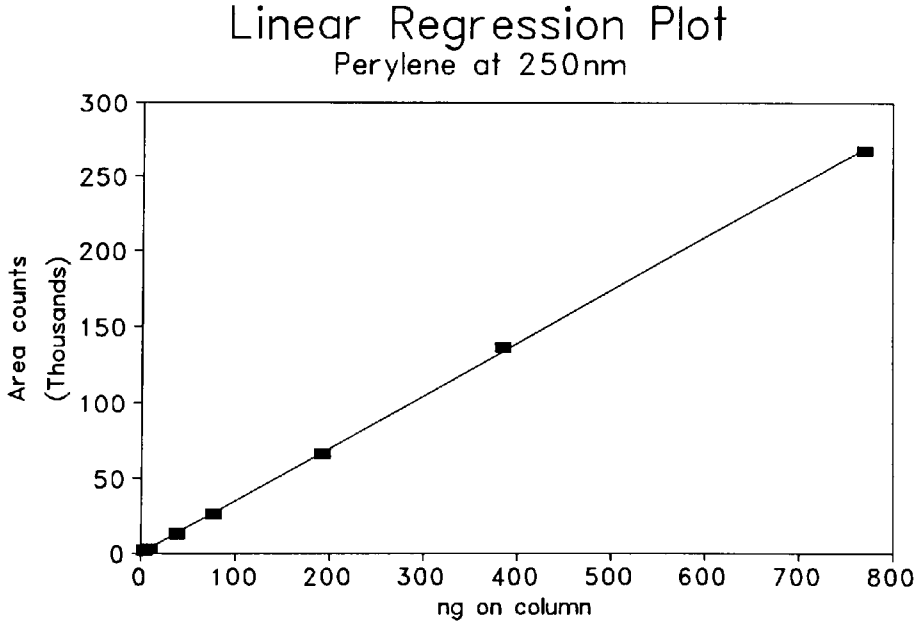


Fig.3.3.10

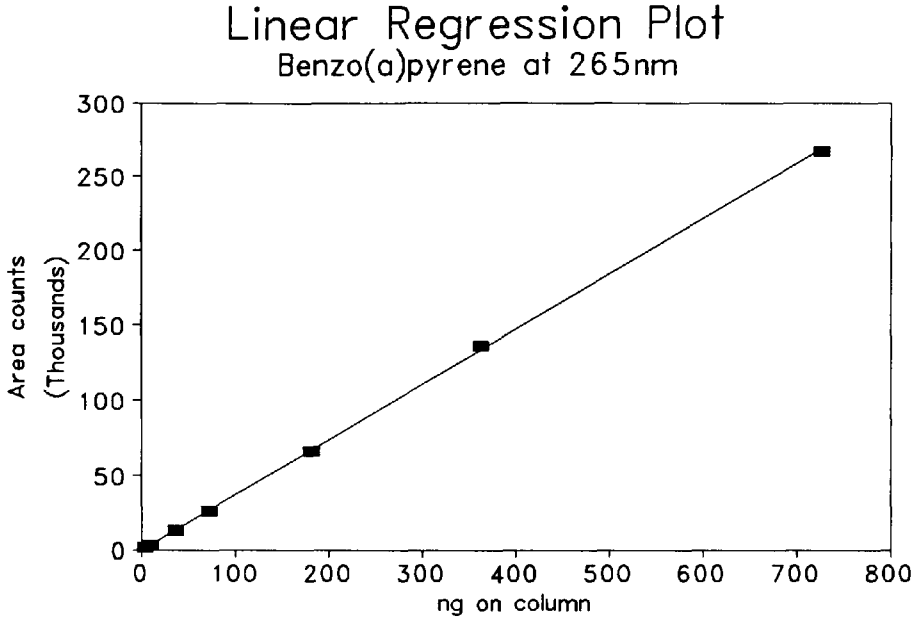


Fig.3.3.11

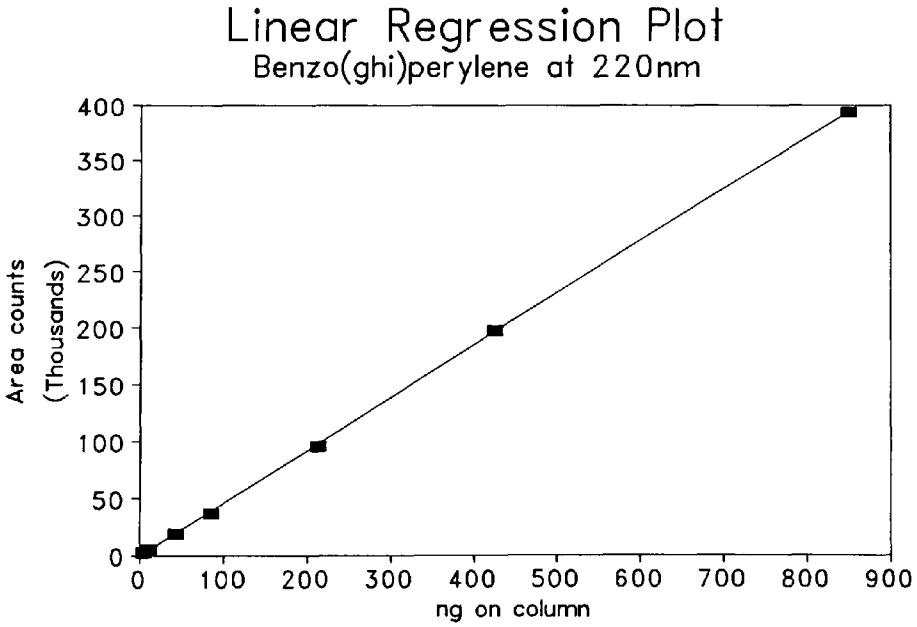


Fig.3.3.12

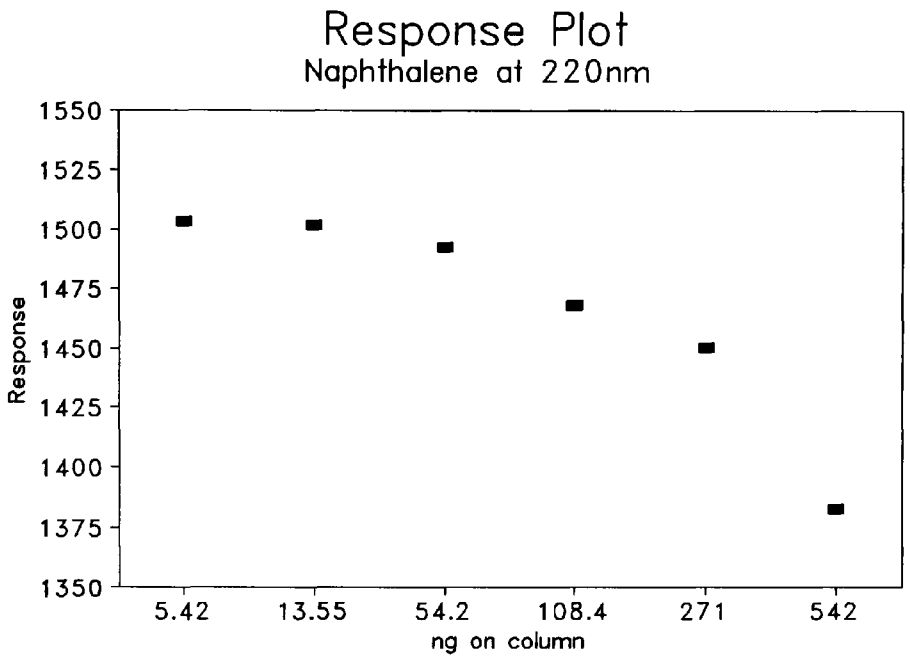


Fig.3.3.13

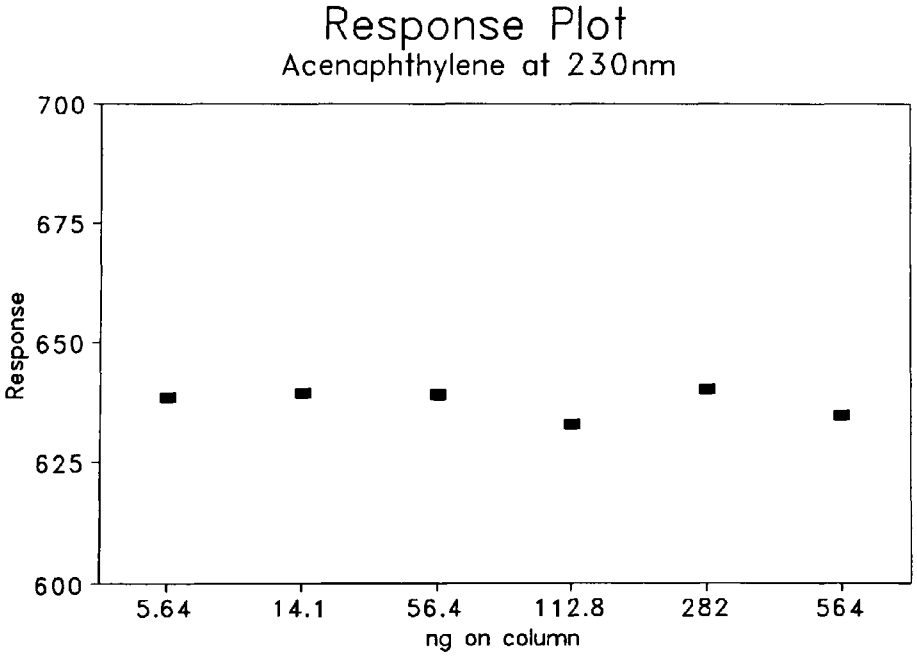


Fig.3.3.14

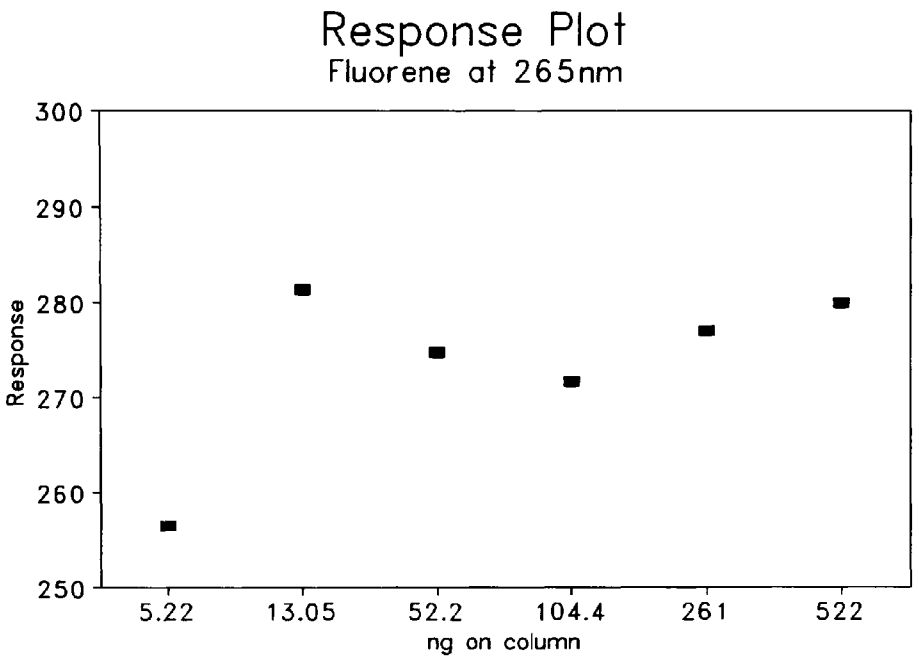


Fig.3.3.15

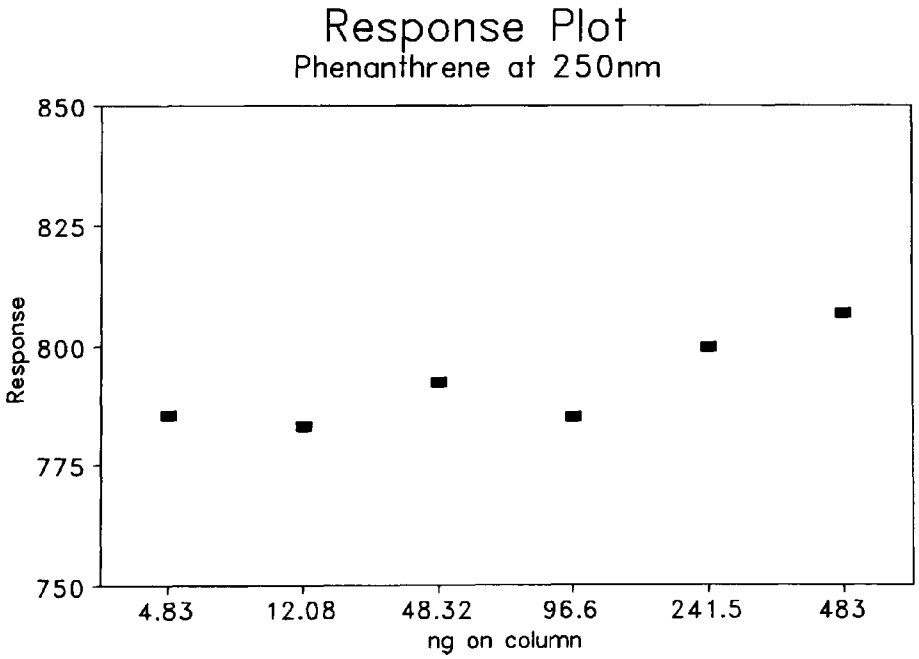


Fig.3.3.16

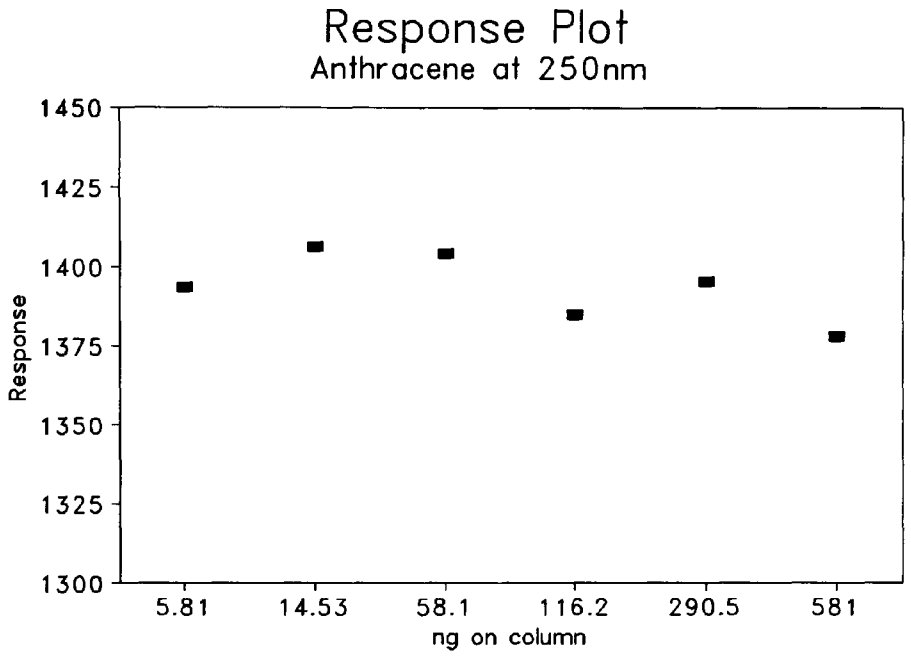


Fig.3.3.17

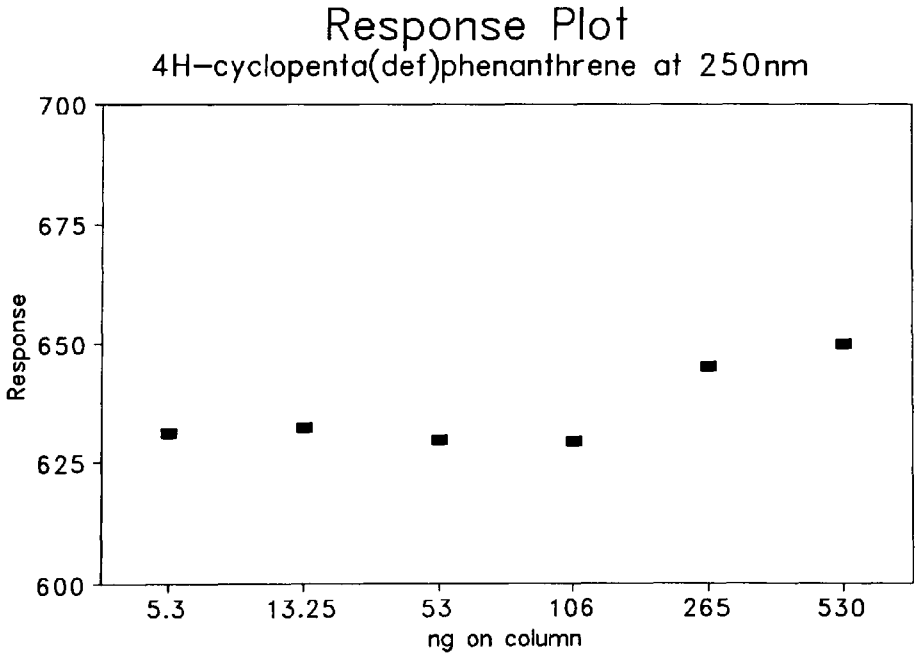


Fig.3.3.18

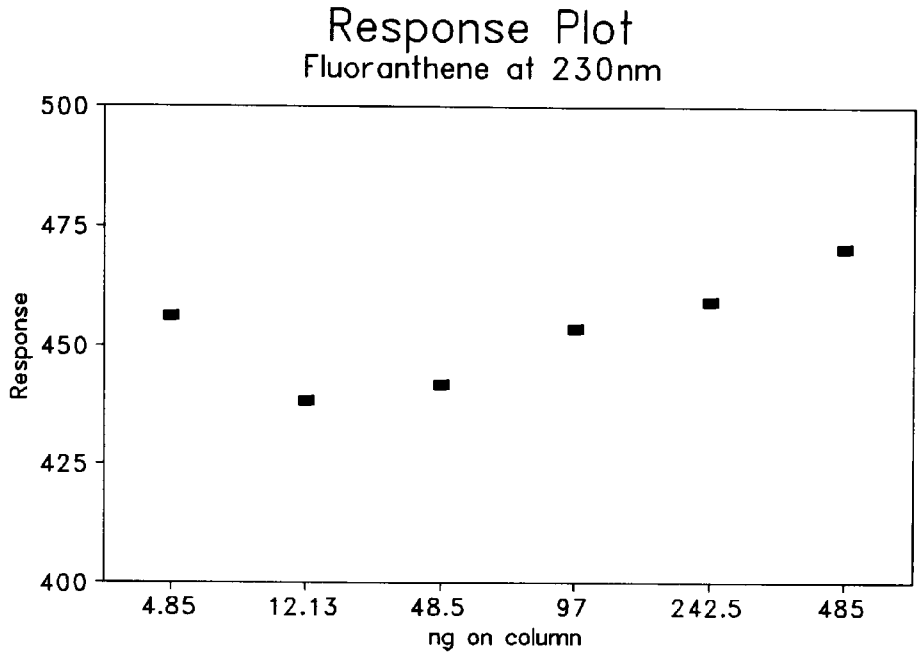


Fig.3.3.19

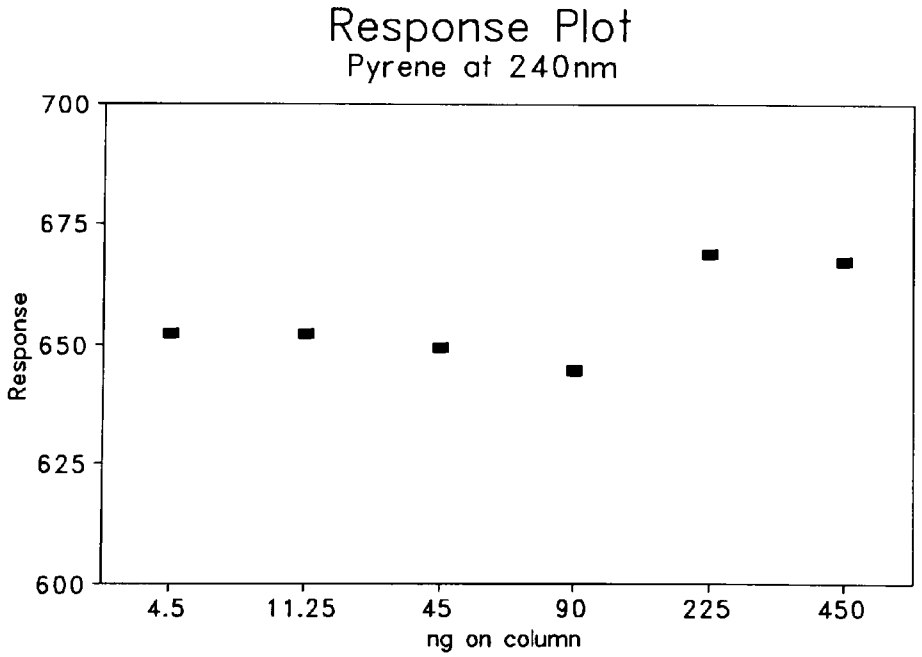


Fig.3.3.20

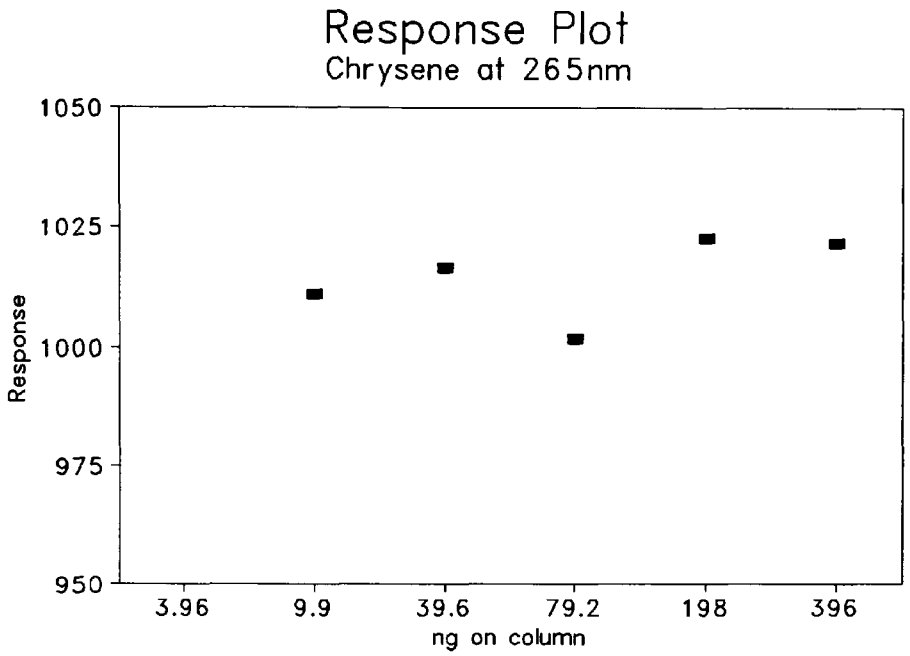


Fig.3.3.21

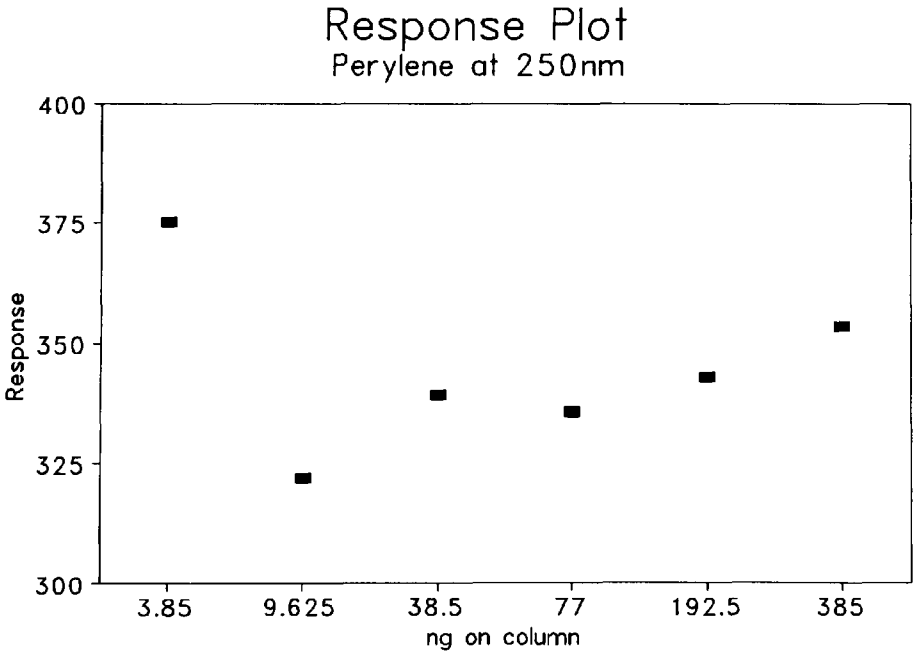


Fig.3.3.22

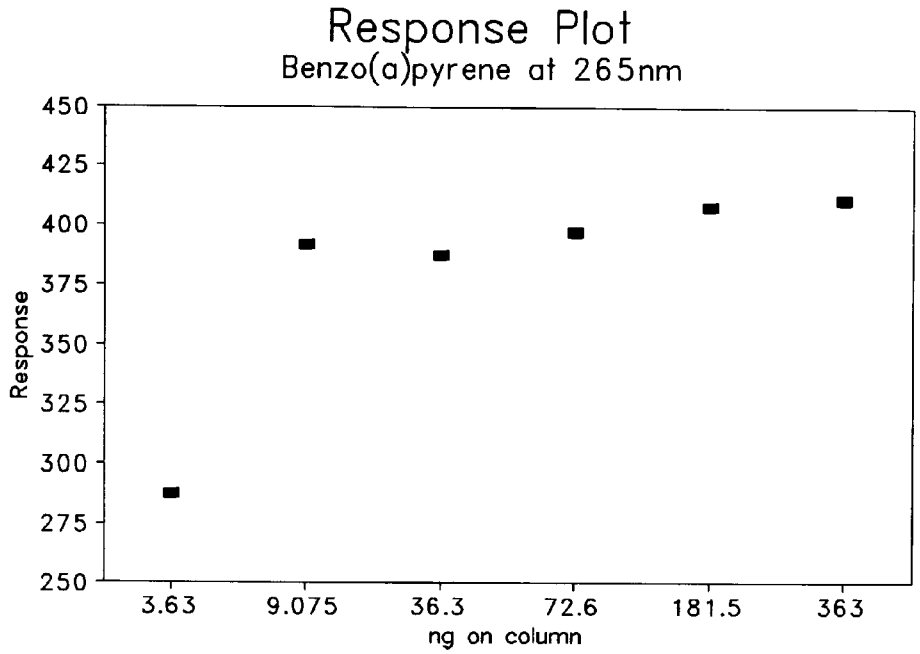


Fig.3.3.23

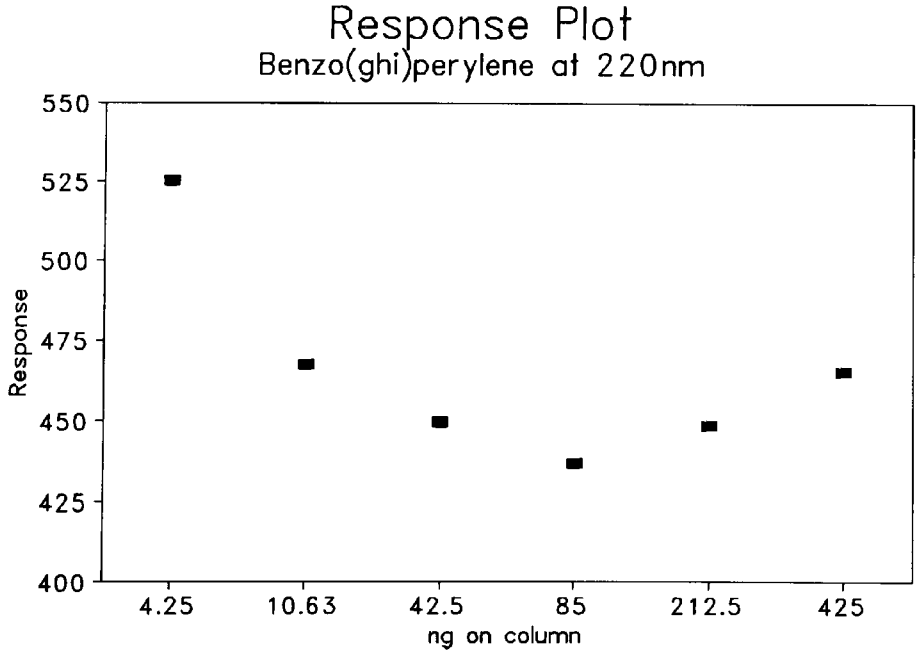


Fig.3.3.24

3.3.3 Precision and Limit of Determination

Linearity sample number 3 containing approximately 50ng / 50µl of each analyte was injected six times and data reduction carried out post run. The RSD (relative standard deviation) was calculated for each component at its optimum wavelength and the limit of determination estimated as six times this figure expressed as a fraction of the analyte injected.

i.e.

For a five gramme soil sample reduced to 2.0 ml and a 50µl injection

$$LOD \epsilon ppb = \frac{6 \times RSD \times wt.injected \times 2 \times 1000 \times 10^9}{100 \times 5 \times 10^9 \times 50}$$

The results obtained are shown in the table below

Precision Study

Code	wt			Area	count			Mean	RSD	LOD
PAH 1	54.2	79111	78427	78563	79100	78449	78442	78682	0.42	11.0
PAH 2	56.4	35329	35374	35360	35728	35436	35475	35450	0.41	11.2
PAH 3	52.2	14112	14002	14040	14178	14074	14183	14098	0.52	13.1
PAH 4	48.3	37654	37540	37601	37931	37699	37691	37686	0.36	8.2
PAH 5	58.1	80321	79588	79654	80414	79929	79985	79982	0.42	11.7
PAH 6	53.0	33215	33378	33053	33354	33174	33075	33208	0.41	10.5
PAH 7	48.5	21320	21620	21113	21184	21087	21037	21227	1.02	23.7
PAH 8	45.0	28812	28793	28531	29097	28722	28648	28767	0.67	14.4
PAH 9	39.6	39328	39297	39361	39887	39631	39707	39535	0.61	11.6
PAH 10	38.5	12895	12815	12853	12981	12898	12891	12889	0.43	7.9
PAH 11	36.3	14409	14251	14298	14421	14443	14461	14381	0.59	10.3
PAH 12	42.5	18456	18201	18285	18480	18577	18560	18427	0.82	16.8

Code	Compound	Wavelength
PAH 1	Naphthalene	220
PAH 2	Acenaphthylene	230
PAH 3	Fluorene	265
PAH 4	Phenanthrene	250
PAH 5	Anthracene	250
PAH 6	4-H-cyclopenta(def)phenanthrene	250
PAH 7	Fluoranthene	230
PAH 8	Pyrene	240
PAH 9	Chrysene	265
PAH 10	Perylene	250
PAH 11	Benzo(a)pyrene	265
PAH 12	Benzo(ghi)perylene	220

3.3.4 Discussion

The HPLC analysis for polyaromatic hydrocarbons has been validated. Use of a diode array detector has allowed for the optimization of detection wavelength for each component of interest. In practice this has meant reprocessing the chromatogram at five different wavelengths to detect twelve components.

Excellent linearity was achieved for all of the compounds investigated. The response plots showed typical deviations from linearity at very low or very high concentrations. This is often found in quantitative applications of HPLC and is due not only to linearity of response but also to such things as difficulty in accurately integrating a small peak, baseline noise etc.

The limit of determination for the chromatography was very low, ranging up to 25 ppb for the system used. This could have been improved even more if a larger sample size had been used. However 25ppb was more than sufficient for the application.

Recovery studies were not carried out. This was considered valid as a standard extraction and clean-up procedure was being followed. Additionally, the absolute values of contaminant were not required. It was the ratio which was of interest in this study.

3.4 The Survey

3.4.1 Sampling sites

Multiple samples were taken from twelve sites which were contained within an area bounded by the following grid references: ST287 992, ST320 999, ST287 970, ST320 970. This is an area NNE of Cwmbran, Gwent which is predominantly rural but also contains a selection of industrial concerns including a steelworks, an automotive parts engineering factory and a large, toxic waste incinerator. The area also has a river running through it. The samples were taken so as to investigate effects from any of the potential major sources. The actual sites are described below with their local names.

Site 1

Pentwyn Farm. A rural site using oil for heating. Two samples taken near farmhouse.

Site 2

A4042 roundabout at Croesyceiliog. A semi-rural site but on a busy through road. Approximately two kilometres from the incineration plant and 200 metres from a crematorium.

Site 3

Field opposite Pear Tree Cottage. A semi-rural site, cottage uses oil for heating, 200 metres from incinerator. Site is subject to occasional flooding from Afon Llwyd.

Site 4

Roadside opposite Lucas Girling. This company produces car components. The site is alongside a busy main road and near to the Afon Llwyd. Two samples taken.

Site 5

Private garden, New Inn village. Approximately 300 metres from incinerator.

Site 6

Pontyfelin Industrial Estate. This site is a quiet roadside between Midwest Litho, a printing firm, and Panteg steelworks.

Site 7

Pasture land approximately 200 metres from incinerator near a busy main road.

Site 8

Busy roadside about 400 metres from both steelworks and incinerator.

Site 9

Private garden, New Inn village. Possibly affected by coal or coke. Two samples taken.

Site 10

Omitted.

Site 11

Pontyfelin grazing land. 100 metres from incinerator, affected by Afon Llwyd. Two samples taken.

Sites 12 to 20

Garden of Pontyfelin House, 150 metres from incinerator. Affected by Afon Llwyd both by flooding and high water table. Fourteen samples taken.

Site 21

Omitted

Site 22

Meadow 150 metres from incinerator. Very occasional flooding from Afon Llwyd but high water table. Two samples taken.

Site 23

Polo ground field. Fairly high ground near Afon Llwyd but not affected by it. 300 metres from incinerator.

3.4.2 Sampling Technique

A bulb planter was used to take the soil samples up to a depth of 10 cms. Any large stones or plant debris were removed and the sample transferred to a hexane washed, amber glass bottle.

3.4.3 Sample analysis

Samples were analyzed in duplicate using the procedure described in section 3.2. As explained previously, detection was by means of a diode array detector. Each chromatographic run could be reintegrated using any number of detection wavelengths. In practice it was found that five were sufficient for the optimized detection of the twelve PAH's investigated.

The samples, after extraction and clean-up, were run in

batches of three or four with each batch being bracketed by standard solutions. For each day's series of runs a reagent blank injection was also made. This was also reprocessed post-run and the results subtracted from those obtained with the analytical samples in order to remove artifacts caused by gradient and ghosting effects. Also to reduce possible artifacts, high purity solvents were used throughout ie propan-2-ol and dichloromethane (Fison's AR grade), Methanol 205 (Rhone Poulenc HPLC grade with extra purification to reduce UV absorption) and water (BDH HPLC grade)

A second advantage gained from using a diode array detector is the ability to produce spectra from sample chromatograms and compare them with those produced from the standards. This gives an excellent means of confirming the identities of peaks.

3.5 Results

All the following results are reported as ppm dry weight. The PAH codes are as follows:-

PAH 1	Naphthalene
PAH 2	Acenaphthylene
PAH 3	Fluorene
PAH 4	Phenanthrene
PAH 5	Anthracene
PAH 6	4-H-cyclopenta(def)phenanthrene
PAH 7	Fluoranthene
PAH 8	Pyrene
PAH 9	Chrysene
PAH 10	Perylene
PAH 11	Benz(a)pyrene
PAH 12	Benz(ghi)perylene

			SITE			
PAH	1A	1B	2	3	4A	4B
PAH 1	0.02	0.02	0.05	0.03	0.09	0.12
PAH 2	0.07	0.12	0.15	0.25	0.02	0.02
PAH 3	0.12	0.14	0.59	0.40	0.84	0.91
PAH 4	0.02	0.02	0.33	0.08	1.24	1.06
PAH 5	0.02	0.02	0.08	0.11	0.13	0.16
PAH 6	0.02	0.02	0.18	0.02	0.26	0.28
PAH 7	0.30	0.33	2.96	1.00	3.43	3.78
PAH 8	0.31	0.27	1.97	0.80	2.69	2.93
PAH 9	0.58	0.34	1.92	0.94	2.26	2.69
PAH 10	0.54	0.43	1.68	2.57	1.93	2.48
PAH 11	0.23	0.16	0.65	0.41	0.84	1.07
PAH 12	0.09	0.08	0.21	0.16	0.49	0.66
Total	2.32	1.95	10.77	6.77	14.22	16.16
	All results recorded in ppm dry weight					

			SITE			
PAH	5	6	7	8	9A	9B
PAH 1	0.03	0.02	0.02	0.03	0.08	0.05
PAH 2	0.11	0.16	0.08	0.11	0.11	0.08
PAH 3	0.14	0.10	0.02	0.09	0.83	1.21
PAH 4	0.06	0.05	0.02	0.02	0.17	0.30
PAH 5	0.03	0.06	0.02	0.02	0.30	0.27
PAH 6	0.03	0.02	0.02	0.03	0.08	0.06
PAH 7	0.51	0.35	0.10	0.25	1.61	1.54
PAH 8	0.43	0.29	0.09	0.21	1.22	1.24
PAH 9	0.51	0.31	0.12	0.28	1.41	1.26
PAH 10	0.50	0.32	0.20	0.35	0.87	0.81
PAH 11	0.25	0.14	0.02	0.15	0.36	0.33
PAH 12	0.11	0.50	0.11	0.04	0.12	0.14
Total	2.71	2.32	0.82	1.58	7.16	7.29
	All results recorded in ppm dry weight					

			SITE		
PAH	11A	11B	12A	12B	13A
PAH 1	0.02	0.02	0.05	0.03	0.12
PAH 2	0.20	0.16	0.05	0.08	0.02
PAH 3	0.82	0.71	1.30	1.24	0.55
PAH 4	0.24	0.28	0.02	0.27	1.13
PAH 5	0.02	0.02	0.23	0.30	0.24
PAH 6	0.08	0.04	0.07	0.08	0.13
PAH 7	2.42	1.77	1.76	2.64	3.09
PAH 8	1.93	1.41	1.35	1.83	2.46
PAH 9	1.98	1.56	1.78	2.03	1.04
PAH 10	3.01	2.37	1.45	1.65	0.64
PAH 11	1.25	0.96	0.55	0.55	0.31
PAH 12	0.42	0.45	0.41	0.32	0.36
Total	12.39	9.75	9.02	11.02	10.09
All results recorded in ppm dry weight					

		SITE			
PAH	13B	14A	14B	15A	15B
PAH 1	0.10	0.05	0.06	0.05	0.05
PAH 2	0.02	0.02	0.02	0.08	0.07
PAH 3	1.17	1.21	0.83	0.89	0.93
PAH 4	1.09	0.73	0.66	0.02	0.02
PAH 5	0.21	0.21	0.22	0.32	0.28
PAH 6	0.15	0.12	0.10	0.02	0.05
PAH 7	2.95	2.64	2.36	2.20	1.85
PAH 8	2.33	2.02	1.78	1.70	1.43
PAH 9	1.67	1.62	1.66	1.85	1.83
PAH 10	0.89	1.08	0.63	1.66	1.76
PAH 11	0.40	0.59	0.72	0.98	0.49
PAH 12	0.56	0.53	0.83	0.06	0.10
Total	11.54	10.82	9.87	9.83	8.86
All results recorded in ppm dry weight					

		SITE			
PAH	16A	17A	18A	19A	19B
PAH 1	0.06	0.03	0.02	0.03	0.04
PAH 2	0.04	0.02	0.15	0.17	0.28
PAH 3	1.31	1.60	0.73	1.28	1.37
PAH 4	0.91	2.76	1.19	2.16	1.52
PAH 5	0.27	1.00	0.29	0.27	0.50
PAH 6	0.02	0.02	0.22	0.30	0.37
PAH 7	3.98	7.83	3.43	4.53	4.85
PAH 8	2.88	5.05	2.08	3.12	3.12
PAH 9	2.44	2.17	1.92	2.04	2.15
PAH 10	1.69	0.68	1.62	0.66	0.29
PAH 11	0.75	0.34	1.25	0.37	0.33
PAH 12	0.80	0.02	0.56	0.54	0.21
Total	15.15	21.52	13.46	15.47	15.03
All results recorded in ppm dry weight					

		SITE			
PAH	20A	22A	22B	23A	
PAH 1	0.14	0.04	0.02	0.02	
PAH 2	0.02	0.32	0.32	0.06	
PAH 3	4.19	1.40	1.17	0.02	
PAH 4	4.10	1.36	1.29	0.06	
PAH 5	1.75	0.41	0.54	0.02	
PAH 6	0.86	0.25	0.24	0.02	
PAH 7	10.53	4.06	3.74	0.18	
PAH 8	6.65	2.44	2.31	0.14	
PAH 9	3.93	1.74	1.54	0.14	
PAH 10	2.35	1.13	0.36	0.14	
PAH 11	0.97	0.33	0.31	0.07	
PAH 12	0.98	0.02	0.02	0.02	
Total	36.47	13.50	11.86	0.89	
All results recorded in ppm dry weight					

The use of data produced from a diode array detector can be shown by its application to sample 20A which showed a particularly high level of PAH contamination. Figure 3.5.1 shows a simple comparison of chromatograms using a detection wavelength of 250 nm. The upper trace is from the reference mixture and the lower from sample 20A. This type of figure could be produced from any system and suggests that the main peaks in the sample chromatogram match several in the reference. The match is by retention time only and, although it is a good guide, it is by no means confirmation.

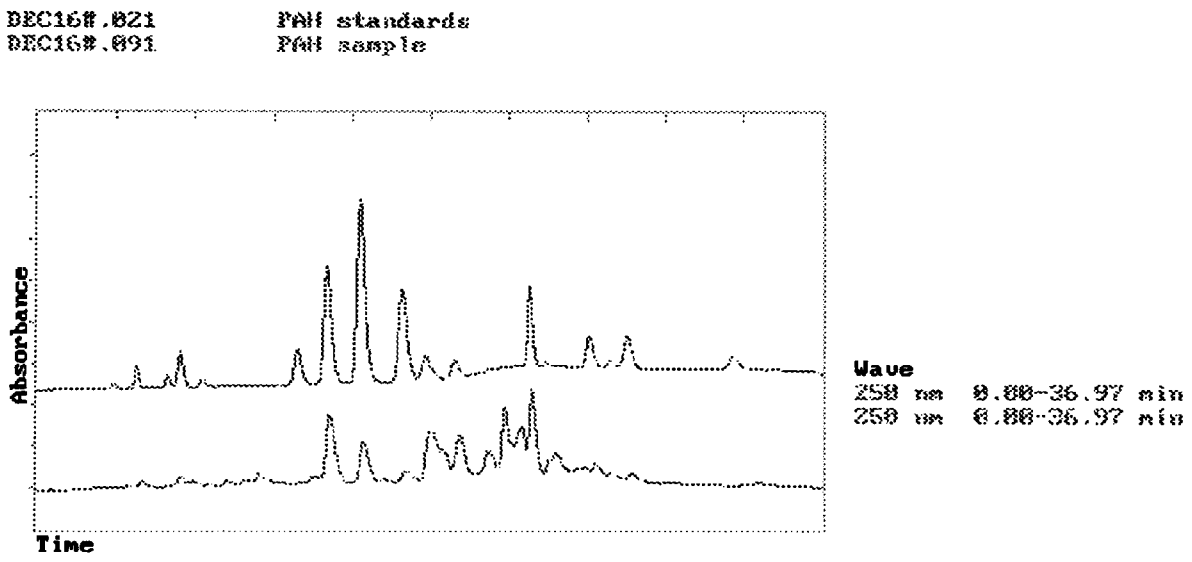


Figure 3.5.1

Chromatogram comparison at 250 nm

Reference and sample 20A

Figure 3.5.2 is the isogram produced from the sample. This is a two dimensional representation of a three dimensional

array. The x-axis represents time, the y-axis is detection wavelength and absorption is represented by depth of colour, translated here into shades of grey. Aromaticity is suggested by high absorbance at high wavelength.

There are two main uses for isogram displays. One is to determine a single, optimum detection wavelength for a group of components and the second is to demonstrate any co-elution. This may be confirmed by the appearance of non-geometrical shapes

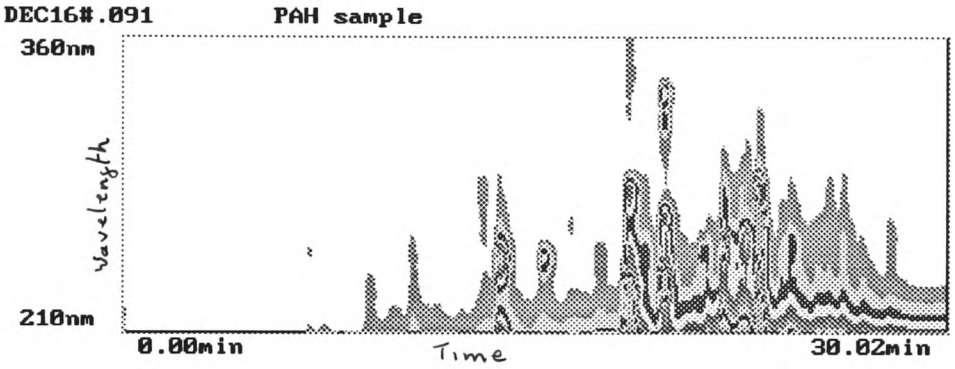


Fig.3.5.2

Isogram from sample 20A

The final confirmation is shown in figures 3.5.3 to 3.5.8 where the spectra for components eluting at equivalent times in the sample and reference chromatograms are compared.

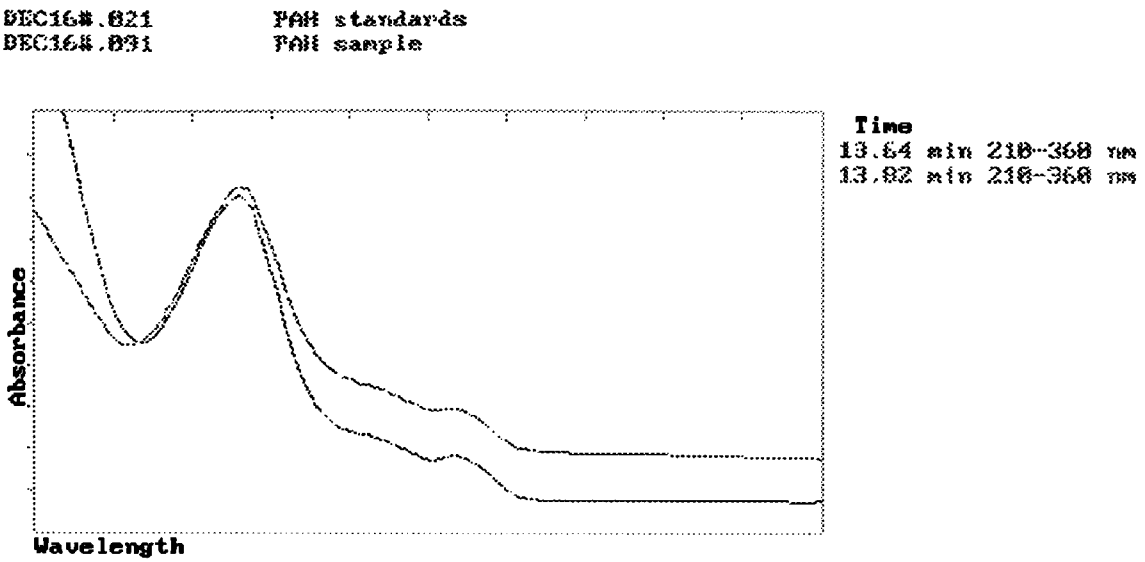


Fig.3.5.3

Spectra of Peaks at 13.7 mins

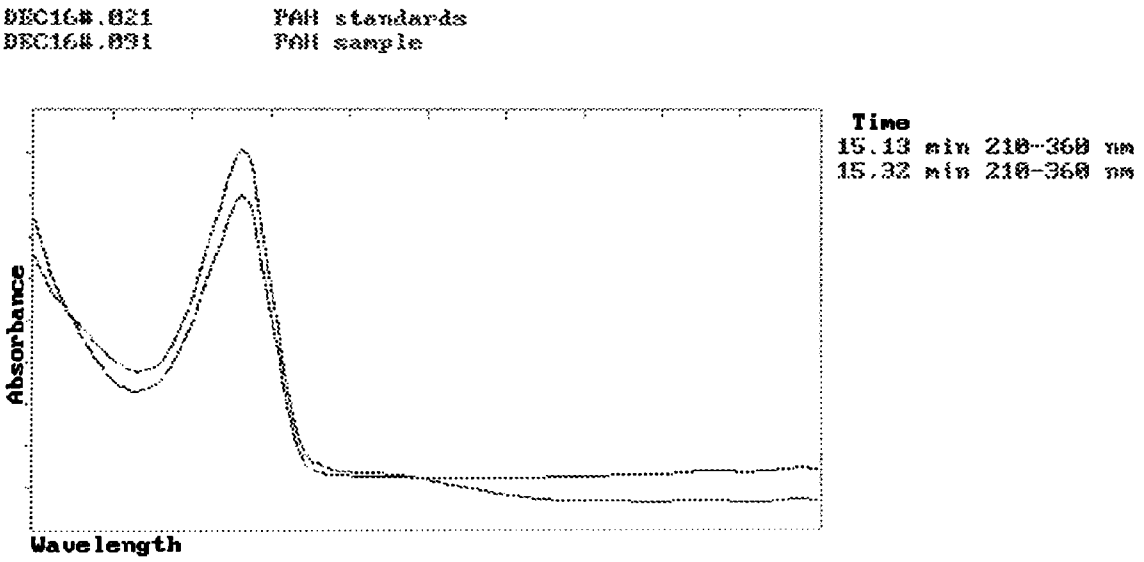


Fig.3.5.4

Spectra of Peaks at 15.2 mins

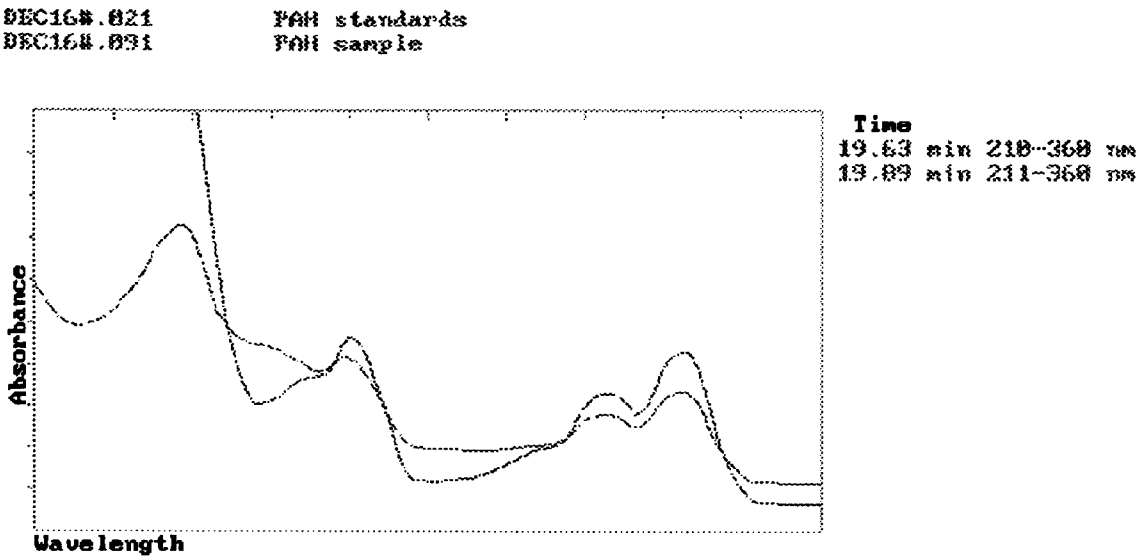


Fig.3.5.5

Spectra of Peaks at 19.7 mins

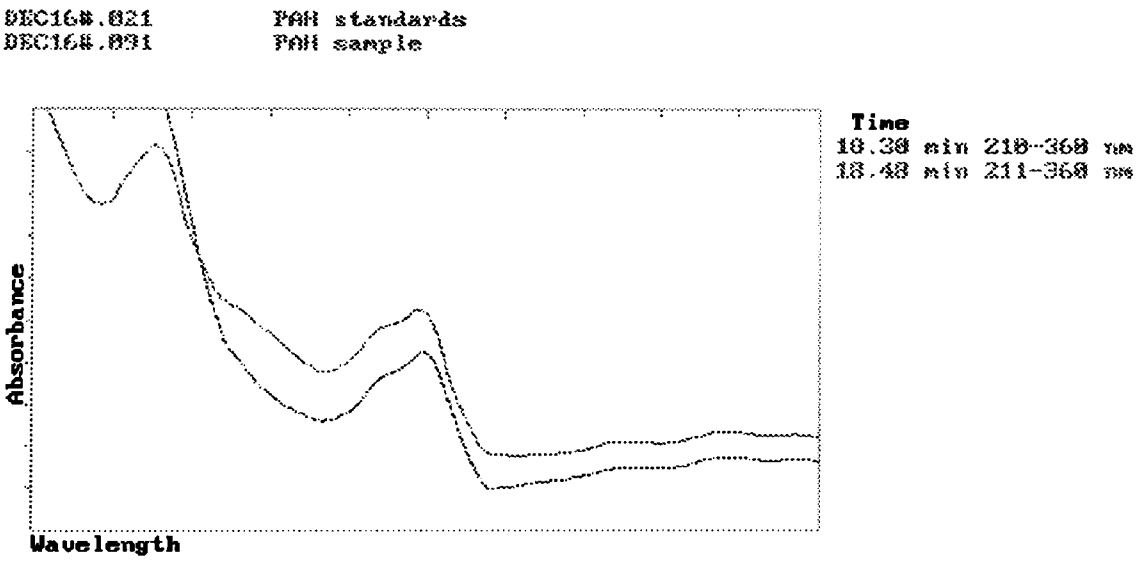


Fig.3.5.6

Spectra of Peaks at 18.4 mins

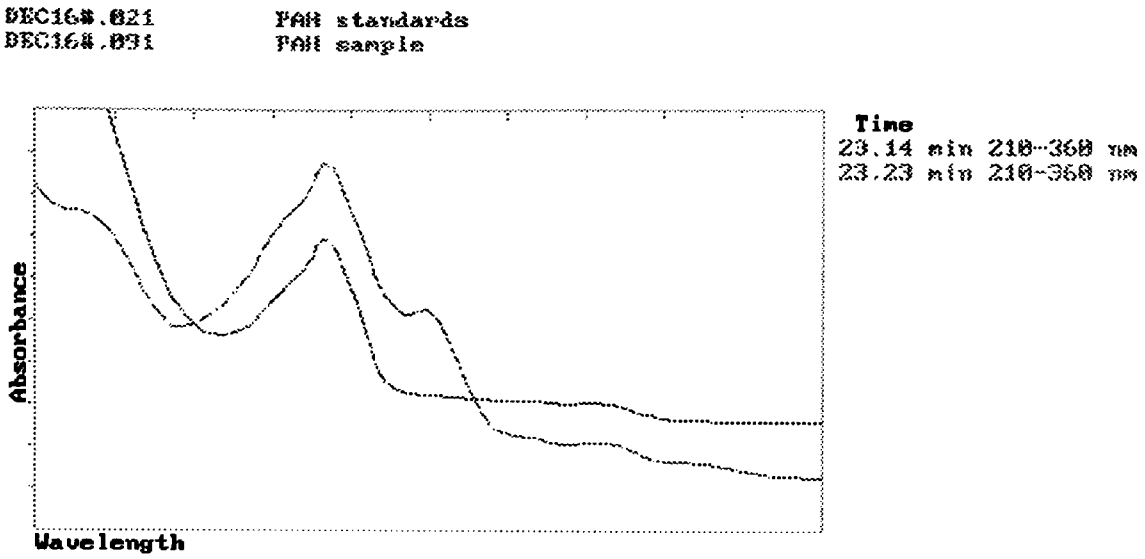


Fig.3.5.7

Spectra of Peaks at 23.2 mins

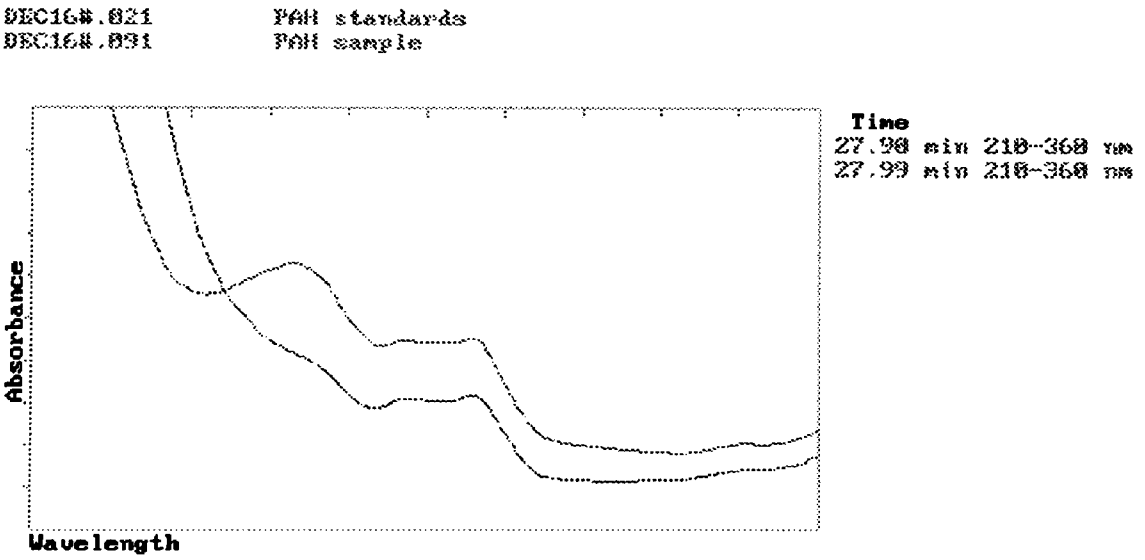


Fig.3.5.8

Spectra of Peaks at 28.0 mins

The components confirmed by the spectra shown above are: phenanthrene, anthracene, pyrene, fluoranthene, chrysene and benzo(a)pyrene.

The spectrum for the peak at 23.2 minutes in the sample suggests chrysene contaminated by a co-eluting component. Examination of the spectrum suggests that the impurity has an absorbance maximum at a higher wavelength than chrysene itself, thus the use of 265 nm as a detection wavelength minimises the effect of the unknown component.

3.6 Statistical Analysis

3.6.1 Introduction

The data reported in the previous section were subjected to statistical analysis using the SPSS/PC+ system (Statistical Package for Social Science - PC version 4) (95). This is a soft-ware package which runs on a standard IBM XT computer.

The data were analyzed by two multivariate techniques

3.6.2 Factor Analysis

Factor analysis is a technique which seeks to identify a relatively small number of factors that can be used to represent relationships among sets of interrelated variables. The analysis proceeds in four major steps:-

3.6.2.1 The correlation index is computed for all variables. Unrelated variables can be identified from this matrix and the appropriateness of the factor model can be determined. This is obviously important as to continue the analysis with unsuitable data will produce meaningless results.

Statistical techniques used to predict the usefulness of factor analysis include Bartlett's test of sphericity. This is a general test often used during analysis of variance (ANOVA).

The parameter proposed is

$$M = \sum_{i=1}^{p} (n_i-1) \ln s^2 - \sum_{i=1}^{p} (n_i-1) \ln s_i^2$$

where s_i^2 is the variance of the procedure i with n determinations and $(n_i - 1)$ degrees of freedom and

$$s^2 = \frac{\sum_{i-1}^{p} (n_i-1)\, s_i^2}{\sum_{i=1}^{p} (n_i-1)}$$

If the null hypothesis is true (the variances are equal), M follows a χ^2 distribution with p-1 degrees of freedom.

A second test is the partial correlation coefficient. If variables share common factors, the partial correlation coefficients between pairs of variables should be small when the linear effects of other variables are omitted. The partial correlations are then estimates of the correlations between the unique factors and should be close to zero when the factor analysis assumptions are met.

The negative of the partial correlation coefficient is called the anti-image correlation. If the proportion of large coefficients in this matrix is high then the use of factor analysis should be reconsidered.

The Kaiser-Meyer-Olkin measure of sampling adequacy is an index for comparing the magnitudes of the observed correlation coefficients to the magnitude of the partial correlation coefficients.

$$KMO = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum a_{ij}^2}$$

where r_{ij} is the simple correlation coefficient between variables i and j, and a_{ij} is the partial correlation coefficient between the same variables. The values of KMO can be summarized:-

0.90's	marvellous
0.80's	meritorious
0.70's	meddling
0.60's	mediocre
0.50's	miserable
<0.5	unacceptable

3.6.2.2 The second step is factor extraction. This procedure determines the number of factors necessary to represent the data. Principal components analysis is used here i.e. linear combinations of the observed variables are formed. The first principal component is the combination that accounts for the largest variance in the sample. The second principal component accounts for the second largest amount of variance and is uncorrelated with the first. Successive components explain

progressively smaller portions of the total sample variance.

It is possible to produce, mathematically, as many principal components as there are variables. This, of course, does not reduce the complexity of the multi-variate system and several of these components may well be due to noise or other errors. For a typical chemical application such as that described by Kankare (96), who determined the individual absorption spectra of a complex mixture of chloro-bismuth ions, the errors may be due to noise produced during the measurement of the spectra.

Several procedures have been postulated for the determination of the correct number of components. The default value used in SPSS/PC is that only factors with an eigenvalue greater than one (i.e. variance greater than one) should be included. This is based on the procedure proposed by Kaiser (102). A plot of eigenvalue against component number is called a Scree plot as the gradual tailing resembles the rubble formed at the foot of a mountain. As proposed by Cattell (101), the scree begins at a point representing the true number of factors.

A residual standard deviation method can also be used. The RSD is defined as

$$r(c-n)\,(RSD)^2 = \sum_{i=j}^{r} \sum_{j=n+1}^{c} (\sigma_{ij})^2$$

where $\underline{c}$ are the secondary eigenvectors and $\underline{n}$ are the primary.

The RSD is calculated first on the basis of only one factor (i.e. the eigenvector with the largest eigenvalue). The RSD is then compared with the estimated experimental error. If the RSD is greater than the experimental error then it can be deduced that more than one factor is involved. If this is the case then the process is repeated, adding factors until the RSD is approximately equal to the experimental error.

Bartlett (97) proposed a chi-squared criterion for use when the standard deviation is not constant through the data matrix. It has the disadvantage that a reasonably accurate estimate of the error for each data point must be known.

$$\chi_n^2 = \sum_{i=1}^{r} \sum_{k=1}^{c} \frac{(d_{ik}-d_{ik}^{\alpha})^2}{\sigma^2{}_{ik}}$$

σ_{ik} is the standard deviation associated with measurable $\underline{d}_{ik}$ $\underline{d}_{ik}^{\alpha}$ is the value of the corresponding point regenerated from factor analysis using the $\underline{n}$ largest eigenvalues and the sum is taken over all experimental points. For each set of eigenvectors X_n^2 is compared to its expected value given by

$$\chi_n^2(expected) = (i-n)(c-n)$$

As before, the number of factors is increased until $\chi_{\underline{n2}}$ is less than its expected value.

Hugus and El-Awady (98) showed that the standard error in an eigenvalue is related to the standard deviations of the data points. The dimensionality of the factor space is taken to be

the number of eigenvalues that have values greater that their respective standard error.

Deuwar et al (99) investigated the prediction methods that were based on accurate knowledge of potential errors. They concluded that none of the procedures was completely satisfactory when used alone. The various criteria, examined together, afford a better guide.

The imbedded error function of Malinowski (100) does not rely upon any error estimate to determine the number of factors in the data matrix.

$$IE = \left[\frac{n \sum_{j=n+1}^{c} \lambda_j^0}{rc(c-n)} \right]^{1/2}$$

It is a function of the secondary eigenvalues, the number of rows and columns in the data matrix and the number of factors.

The IE should decrease as more primary eigenvectors are used in data reproduction. Once the secondary or error eigenvectors are included then the IE should increase.

In the same paper (100) Malinowski discussed the Factor Indicator Function. This appears to be more sensitive than the imbedded error function in its ability to detect the correct number of factors. It is defined as

$$IND = \frac{RE}{(c-n)^2}$$

It is composed of the same variables as the IE and again reaches a minimum when the correct number of factors are employed. RE relates to the real error in the data matrix and can be calculated by

$$RE = \left[\frac{\sum_{j=n+1}^{c} n\lambda_j^0}{r(c-n)} \right]^{1/2}$$

3.6.2.3 The third step is rotation. This is a procedure to transform the factors to make them more amenable to interpretation. All of the algorithms used to carry out rotation may be described as either orthogonal or oblique. Orthogonal rotations preserve the angular relationships between the original sets of eigenvectors obtained from the factor analysis. The technique results in a set of uncorrelated factors. Oblique rotations do not retain the angles between eigenvectors and may lead to a set of correlated factors.

Within the two main classes are a series of different methods:-

Quartimax rotation is an orthogonal method. Utilizing a two-dimensional example, the procedure rotates the orthogonal axes so that as one axis approaches a data point the projection (i.e. the loading) of the point on the axis increases. Obviously its loading on the other axis decreases. The Quartimax procedure seeks a set of orthogonal axes that groups the points in clusters around each of those axes. The procedure

was described by Harman (103) who achieved the rotation by minimizing the Quartimax function where

$$Q = \sum_{j=1}^{n} \sum_{k=1}^{c} \lambda_j^2 \gamma_{jk}^4$$

λ_j is the j^{th} eigenvalue

γ_{jk} is the loading of the k^{th} data column vector on the j^{th} axis after the rotation has been completed.

The sum is taken over all c data columns and over n eigenvectors which are required to span the factor space. A disadvantage of this technique is that the first factor tends to be overloaded which produces one large, general factor and many smaller, subsidiary factors.

The Varimax method minimizes the total variance V of the squared loadings. Kaiser's version (104) is the most popular because of its ability to yield the same clusters regardless of the size of the data matrix.

$$V = \sum_{j=1}^{n} \left[\frac{1}{C} \sum_{k=1}^{c} (\lambda_j \gamma_{jk}^2)^2 - \frac{1}{C^2} \left(\sum_{k=1}^{c} \lambda_j \gamma_{jk}^2 \right)^2 \right]$$

Varimax attempts to limit the number of variables that have high loadings on a factor, which should enhance the interpretability of the factors.

The Equimax system used by SPSS/PC is a combination of both Quartimax and Varimax.

Oblimax is an example of oblique rotation where the number of high and low loadings on a given axis are increased by decreasing the loadings in the middle range. Saunders (105) demonstrated that this can be done by maximizing the kurtosis function K

$$K = \frac{\sum_{j=1}^{n}\sum_{k=1}^{c}\lambda_j^2\gamma_{jk}^4}{\left(\sum_{j=1}^{n}\sum_{k=1}^{c}\lambda_j\gamma_{jk}^2\right)^2}$$

Another type of transformation is Target Factor Analysis as described by Malinowski (142). This technique produces real factors which can then be used for further predictions. It also allows "target testing" which can be used to test potential factors individually. A computer program is available for this procedure (146)

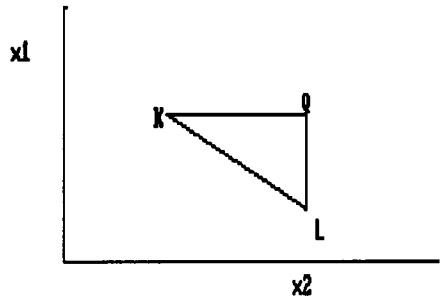
3.6.2.4 The fourth step of the analysis involves computing the scores for each factor. The resultant scores can be used in a variety of further analyses e.g. plots of factor scores for pairs of factors can be useful for detecting unusual observations.

3.6.3 Cluster Analysis

Cluster analysis may be defined as the search for relatively homogeneous groups of objects. In cluster analysis it is not necessary to know group membership in advance. Discriminant analysis, on the other hand, uses prior knowledge

to identify data sets into which unknown samples may be placed later. An example of this is SIMCA which was described earlier in 1.2.1.

The goal of cluster analysis ,then, is to identify homogeneous groups or clusters. If the data collected are controlled by just two variables then a visual study of a bi-variate plot will show any tendency towards clustering. To construct these groupings some measure of similarity between data points is needed. This may be provided by the Euclidean distance. The geometrical meaning of this value is most easily understood if pictured in two dimensions.



The D_{kl} is the hypotenuse of the triangle KLQ and

$$D_{kl} = \sqrt{(x_{k1}-x_{l1})^2 + (x_{k2}-x_{l2})^2}$$

i.e.

$$D_{kl} = \sqrt{\sum_{j=1}^{2} (x_{kj} - x_{lj})^2}$$

and for m variables

$$D_{kl} = \sqrt{\sum_{j=1}^{m} (x_{kj} - x_{lj})^2}$$

A problem arising from the use of Euclidean distance is the scale effect i.e.small values of a parameter have less effect than larger. This can be resolved by scaling the data. This is most often done by using the z-transform (also called autoscaling)

$$z_{ij} = \frac{a_{ij} - \bar{a}_j}{s_j}$$

where a_{ij} is the value for object i of variable j

$\bar{a}_j$ is the mean for variable j

s_j is the standard deviation for variable j

3.6.3.1 Hierarchical Clustering

There is a wide variety of techniques available for hierarchical clustering. The two used most often are average linkage and single linkage.

Single linkage is also known as "nearest neighbour". In

this technique the first two cases combined are those that have the smallest distance between them. The distance between the new cluster and individual cases is then computed as the minimum distance between an individual case and a case in the cluster.

For average linkage the similarities between the new cluster and the others are obtained by averaging the similarities of the first two with the others.

3.6.4 Factor Analysis of the Data

The data produced in section 3.5 were transformed into z-forms for further analysis. The full data set is summarized below in table 3.6.1, both before and after transformation.

Variable	Mean	Std.dev.	Min	Max
PAH1	0.05	0.03	0.02	0.14
PAH2	0.11	0.09	0.02	0.32
PAH3	0.91	0.77	0.02	4.19
PAH4	0.75	0.94	0.02	4.10
PAH5	0.27	0.34	0.02	1.75
PAH6	0.13	0.17	0.02	0.86
PAH7	2.67	2.26	0.02	10.53
PAH8	1.89	1.43	0.08	6.65
PAH9	1.54	0.85	0.12	3.93
PAH10	1.18	0.81	0.14	3.01
PAH11	0.52	0.34	0.02	1.25
PAH12	0.32	0.27	0.02	0.98
Z1	0.00	1.00	-0.82	2.71
Z2	0.00	1.00	-0.99	2.37
Z3	0.00	1.00	-1.15	4.25
Z4	0.00	1.00	-0.78	3.57
Z5	0.00	1.00	-0.72	4.32
Z6	0.00	1.00	-0.67	4.29
Z7	0.00	1.00	-1.17	3.48
Z8	0.00	1.00	-1.26	3.33
Z9	0.00	1.00	-1.67	2.81
Z10	0.00	1.00	-1.29	2.25
Z11	0.00	1.00	-1.46	2.13
Z12	0.00	1.00	-1.09	2.42

Summarized PAH Data Set
Table 3.6.1

The data were subjected to factor analysis using the SPSSPC program. The first section involved the computation of the correlation matrix for all variables. This matrix is reproduced below in table 3.6.2

				Correlation Matrix								
	PAH1	PAH2	PAH3	PAH4	PAH5	PAH6	PAH7	PAH8	PAH9	PAH10	PAH11	PAH12
PAH1	1.00											
PAH2	-0.50	1.00										
PAH3	0.55	-0.10	1.00									
PAH4	0.48	-0.06	0.82	1.00								
PAH5	0.45	-0.08	0.91	0.86	1.00							
PAH6	0.57	-0.07	0.80	0.81	0.75	1.00						
PAH7	0.51	-0.06	0.89	0.95	0.89	0.78	1.00					
PAH8	0.58	-0.13	0.88	0.93	0.86	0.76	0.99	1.00				
PAH9	0.58	-0.12	0.85	0.71	0.68	0.71	0.85	0.88	1.00			
PAH10	0.30	-0.02	0.35	0.11	0.13	0.24	0.29	0.33	0.61	1.00		
PAH11	0.31	-0.12	0.39	0.23	0.19	0.34	0.39	0.44	0.70	0.82	1.00	
PAH12	0.55	-0.38	0.50	0.44	0.28	0.50	0.45	0.49	0.60	0.41	0.59	1.00

Correlation Matrix for Full Data Set
Table 3.6.2

The correlation matrix shows poor correlation for variable PAH2 (acenaphthylene). In addition, the program reported that the matrix was ill-conditioned for factor processing.

During the analytical chemistry it was found that the first two components of interest, i.e. naphthalene and acenaphthylene, were generally in very low concentration and were difficult to quantitate due to interferences. Similarly PAH12 (benzo(ghi)perylene) was in low concentration and ,due to its late elution, appeared as a wide, flat peak which was also difficult to quantify.

The correlation matric for PAH's 3 to 11 was therefore produced and is shown below in table 3.6.3

			Correlation Matrix						
	PAH3	PAH4	PAH5	PAH6	PAH7	PAH8	PAH9	PAH10	PAH11
PAH3	1.00								
PAH4	0.82	1.00							
PAH5	0.91	0.86	1.00						
PAH6	0.80	0.81	0.75	1.00					
PAH7	0.89	0.95	0.89	0.78	1.00				
PAH8	0.88	0.93	0.86	0.76	0.99	1.00			
PAH9	0.85	0.71	0.68	0.71	0.85	0.88	1.00		
PAH10	0.35	0.11	0.13	0.25	0.29	0.33	0.61	1.00	
PAH11	0.39	0.23	0.19	0.34	0.39	0.44	0.70	0.82	1.00

Correlation Matrix for PAH's 3-11
Table 3.6.3

Using the reduced data set, much better correlations are obtained with over 85% being better than 0.3.

The Kaiser-Meyer-Olkin measure of sampling adequacy is 0.78 which is defined as middling to meritorious. A similar statistic appears on the diagonal of the anti-image correlation matrix shown below in table 3.6.4. All of the measures are about 0.70 which indicates that the matrix is amenable to factor analysis.

			Correlation Matrix						
	PAH3	PAH4	PAH5	PAH6	PAH7	PAH8	PAH9	PAH10	PAH11
PAH3	0.79								
PAH4	-0.21	0.76							
PAH5	-0.77	0.32	0.74						
PAH6	-0.13	-0.66	-0.08	0.82					
PAH7	0.34	-0.36	-0.62	0.03	0.81				
PAH8	-0.08	-0.36	0.21	0.41	-0.69	0.84			
PAH9	-0.58	0.67	0.55	-0.41	-0.31	-0.34	0.73		
PAH10	-0.16	0.24	0.10	-0.09	0.01	-0.13	0.06	0.76	
PAH11	0.34	-0.19	-0.15	0.04	0.09	0.10	-0.49	-0.58	0.71

Anti-image Correlation Matrix for PAH's 3-11
Table 3.6.4

The next step in the process is to determine the number of factors. The program default value is that only factors with an eigenvalue greater than 1 should be included. The initial statistics are shown below in table 3.6.5. It is apparent from these figures that two factors are sufficient to describe the data.

The scree plot of the initial statistics is shown in figure 3.6.1. The suggestion is that the true number of factors is indicated by the start of the scree

		Factor Analysis				
Variable	Commun		Factor	Eigen value	% of var	Cum %
PAH3	1.00		1	6.39	71.0	71.0
PAH4	1.00		2	1.71	19.0	90.0
PAH5	1.00		3	0.31	3.5	93.5
PAH6	1.00		4	0.26	2.9	96.4
PAH7	1.00		5	0.15	1.7	98.1
PAH8	1.00		6	0.12	1.3	99.4
PAH9	1.00		7	0.04	0.4	99.8
PAH10	1.00		8	0.01	0.1	99.9
PAH11	1.00		9	0.01	0.1	100.0

Factor Analysis of PAH's 3-11
Table 3.6.5

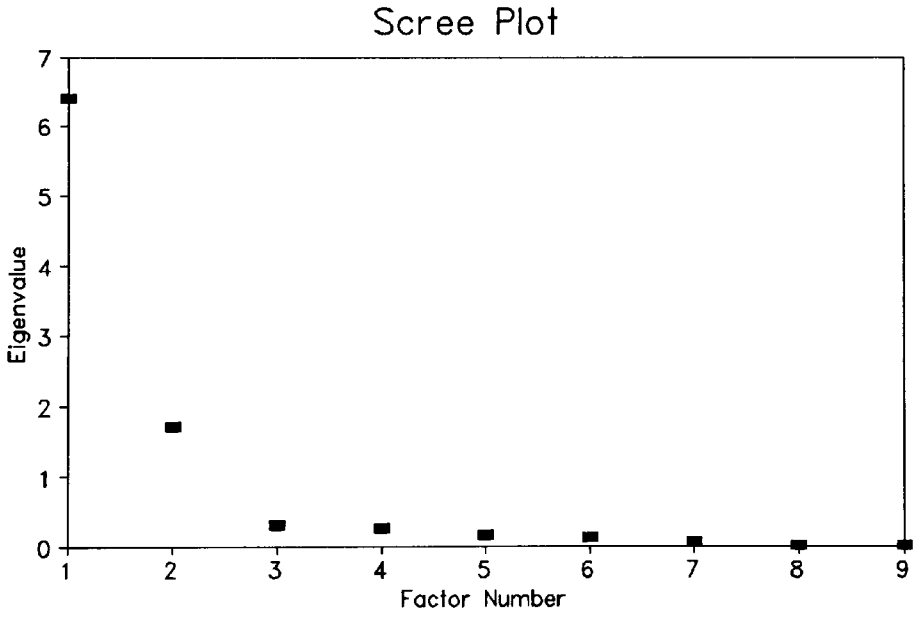


Figure 3.6.1

From an analytical chemistry point of view, a variance of about 10% would be expected due to error, which is typical for analysis at the ppm level. This lends further credence to the belief that only two factors are involved. The factor matrix and final statistics are shown below in tables 3.6.6 and 3.6.7. The communality of the variables are reduced as only two factors are now used to describe the data set.

	Factor matrix	
	FACTOR 1	FACTOR 2
PAH3	0.94	-0.10
PAH4	0.90	-0.34
PAH5	0.88	-0.34
PAH6	0.85	-0.15
PAH7	0.97	-0.15
PAH8	0.97	-0.09
PAH9	0.92	0.29
PAH10	0.44	0.84
PAH11	0.53	0.79

Factor Matrix for PAH's 3-11
Table 3.6.6

		Factor Analysis				
Variable	Commun		Factor	Eigen value	% of var	Cum %
PAH3	89.00		1	6.39	71.0	71.0
PAH4	0.93		2	1.71	19.0	90.0
PAH5	0.89					
PAH6	0.74					
PAH7	0.96					
PAH8	0.95					
PAH9	0.93					
PAH10	0.90					
PAH11	0.91					

Factor Analysis of PAH's 3-11, Final Statistics
Table 3.6.7

Communalities can range from 0 to 1, with 0 indicating that the common factors explain none of the variance and 1 indicating that all of the variance is explained by them. The variance that is not explained by the common factors is attributed to the unique factor and is called the uniqueness of the variable.

The factor matrix can be used to estimate correlations between the variables. In general, if factors are orthogonal then the estimated correlation coefficient for variables i and j is

$$r_{ij} = \sum_{f=1}^{k} r_{fi} r_{fj}$$

where k is the number of common factors and r_{fi} is the correlation between the fth factor and the ith variable. The difference between the observed correlation, as shown in table 3.6.2, and the estimated is known as a residual.

The estimated correlation coefficients and the residuals are shown below in table 3.6.8. The residuals are listed above the diagonal and the estimated correlation coefficients below. The values on the diagonal are the communalities.

			Reproduced Correlation Matrix						
	PAH3	PAH4	PAH5	PAH6	PAH7	PAH8	PAH9	PAH10	PAH11
PAH3	.89*	-0.06	0.05	-0.01	-0.04	-0.04	0.01	0.02	-0.03
PAH4	0.88	.93*	-0.05	-0.01	0.02	0.03	-0.03	0.00	0.02
PAH5	0.86	0.91	.89*	-0.05	-0.01	-0.03	-0.03	0.03	-0.01
PAH6	0.81	0.82	0.80	.74*	-0.07	-0.08	-0.03	0.00	0.01
PAH7	0.93	0.93	0.91	0.85	.96*	0.04	0.01	-0.01	0.00
PAH8	0.92	0.91	0.89	0.84	0.95	.94*	0.02	-0.01	0.00
PAH9	0.84	0.73	0.71	0.74	0.85	0.86	.93*	-0.04	-0.01
PAH10	0.33	0.11	0.10	0.24	0.30	0.35	0.65	.90*	-0.08
PAH11	0.42	0.21	0.20	0.33	0.39	0.44	0.72	0.90	.91*

Reproduced Correlation Matrix for PAH's 3-11
Table 3.6.8

The next step in the process is rotation, which attempts to transform the initial matrix into one that is easier to interpret. Three different procedures are used here; varimax, quartimax and equamax. The results are plotted below in figures 3.6.2 to 3.6.5. The best results are obtained from the equamax rotation which is a compromise between simplification of factors and variables. The final co-ordinates after rotation are shown in table 3.6.9. The figures suggest that factor#1 is associated mainly with PAH's 3 to 9 and factor#2 with PAH's 9 to 11.

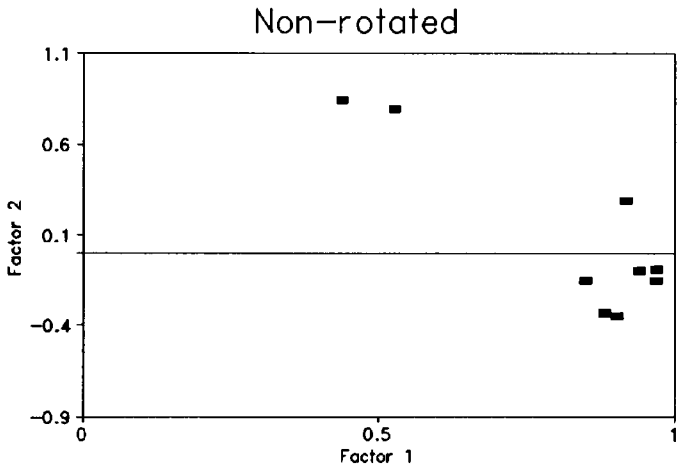


Figure 3.6.2

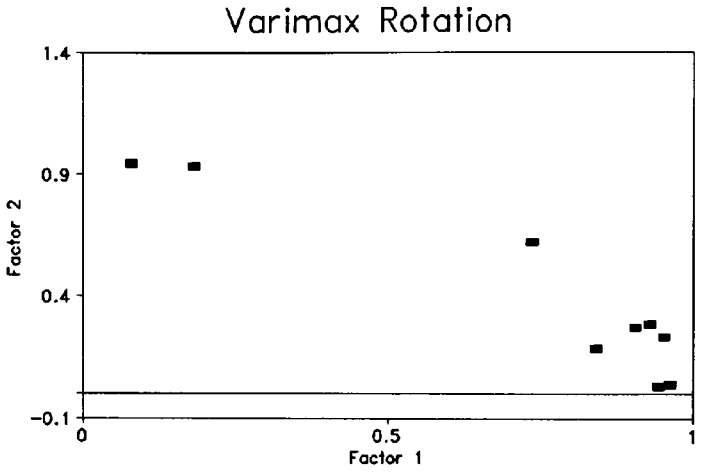


Figure 3.6.3

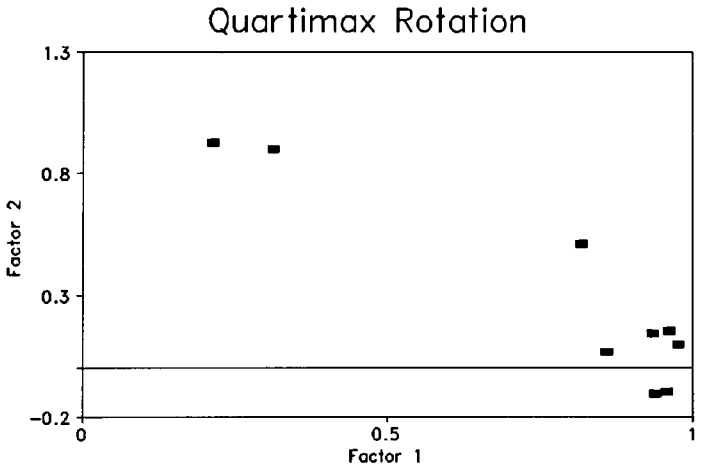


Figure 3.6.4

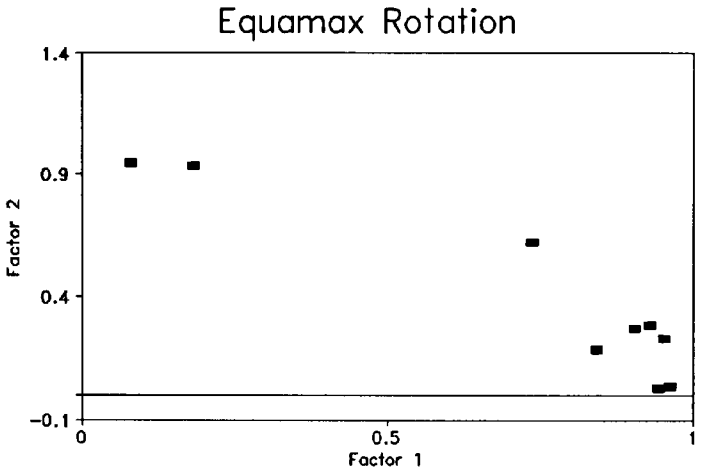
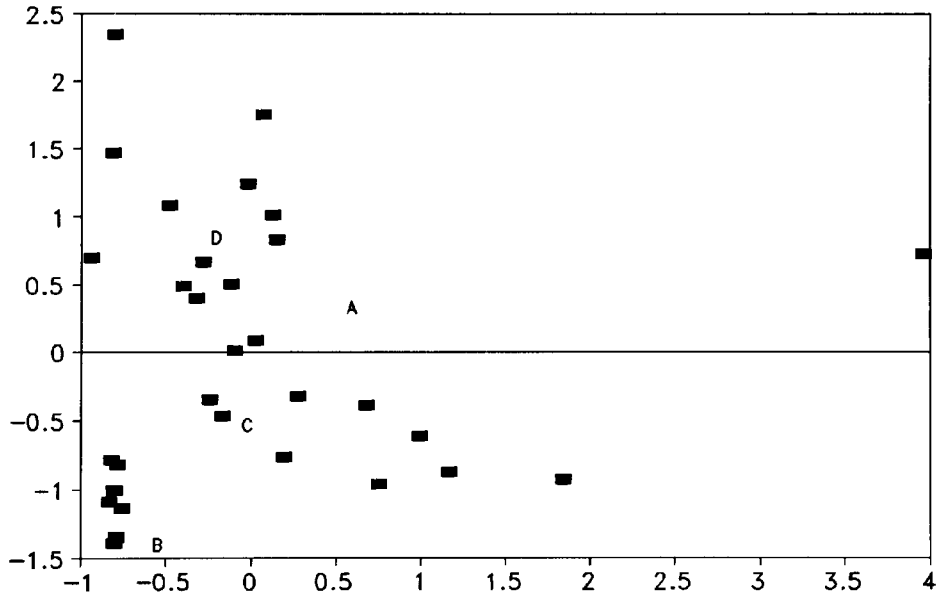


Figure 3.6.5

Equamax Rotated Factors		
	Fact 1	Fact 2
PAH3	0.906	0.272
PAH4	0.962	0.037
PAH5	0.943	0.029
PAH6	0.842	0.187
PAH7	0.953	0.232
PAH8	0.930	0.287
PAH9	0.737	0.623
PAH10	0.079	0.946
PAH11	0.182	0.935

Table 3.6.9

The final part of the factor analysis is the production of principal components plots for all of the sample sites. The plot is shown below in Figure 3.6.6. There is a clear distinction between the two sets, all of the sites encompassed by boundary "A" (with two exceptions) are affected by the Afon Llwyd. Site "C" was a private garden but further investigation showed that the owner had treated the soil with ashes from a coal fire and with coal dust. Site "D" was near to a crematorium and thus could be subjected to PIC (products of incomplete combustion). The sites in section"B" were unaffected by the river.



Principal components plot

Fig 3.6.6

The factor analysis results for the whole data set strongly suggest that two factors are responsible for the PAH

profile of the area and that one of these factors is related to the Afon Llwyd. Upstream of the sampling area this river runs through parts of Gwent that were heavily mined for coal in the recent past. It is interesting to note that one site not affected by the river but appearing to be clustered with the river sites was actually found to have been directly affected with coal.

In order to further investigate the river / non-river theory, the data were subjected to cluster analysis.

3.6.5 Cluster Analysis

Hierarchical cluster analysis may be used for combining objects into groups. Euclidean distances are used to measure the distances between objects and a single linkage or nearest-neighbour method is used to actually cluster the data. Table 3.6.10 shows the membership of clusters from 2 to 9.

It is immediately apparent that sample sites 24 and 28 are outliers. Inspection of the original data set shows that both of these sites had exceptionally high values for PAH7 (fluoranthene). If these locations are ignored then a 2-cluster result groups sites 1,2,7,8,9,10 and 31 as one cluster and all the others as the second. This is exactly the same result as was obtained from the principal components plot shown earlier. It still groups the coal-treated garden and the crematorium with the river sites.

		Number of Clusters						
Site	9	8	7	6	5	4	3	2
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1
3	2	2	2	2	2	2	1	1
4	3	3	3	3	3	2	1	1
5	4	4	4	4	2	2	1	1
6	4	4	4	4	2	2	1	1
7	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1
11	2	2	2	2	2	2	1	1
12	2	2	2	2	2	2	1	1
13	5	2	2	2	2	2	1	1
14	5	2	2	2	2	2	1	1
15	2	2	2	2	2	2	1	1
16	2	2	2	2	2	2	1	1
17	2	2	2	2	2	2	1	1
18	2	2	2	2	2	2	1	1
19	2	2	2	2	2	2	1	1
20	2	2	2	2	2	2	1	1
21	5	2	2	2	2	2	1	1
22	2	2	2	2	2	2	1	1
23	6	5	2	2	2	2	1	1
24	7	6	5	5	4	3	2	1
25	8	7	6	4	2	2	1	1
26	2	2	2	2	2	2	1	1
27	2	2	2	2	2	2	1	1
28	9	8	7	6	5	4	3	2
29	2	2	2	2	2	2	1	1
30	2	2	2	2	2	2	1	1
31	1	1	1	1	1	1	1	1

Production of Clusters from PAH Data
Table 3.6.10

3.6.6 Factor Analysis of Subsets of the Original Data

As a result of the strong confirmation of the existence of two factors describing the entire data set, and the fact that they were due to identifiable sources, it was decided to investigate subsets of the original data using factor analysis.

3.6.6.1 Non-river affected sites

The non-river affected database consisted of nine sites described by twelve variables. As before, PAH's 1,2 and 12 were omitted from the analysis as they were the greatest source of potential errors. The correlation matrix is shown below in table 3.6.11 and demonstrates good correlation between variables. The factor analysis is shown in table 3.6.12 and suggests that just one factor is responsible for 91.6% of the variance within the data. The other 8.4% is almost certainly due to analytical errors. It appears, therefore, that one factor is responsible for the PAH profile of sites removed from the affect of the Afon Llwyd.

			Correlation Matrix						
	PAH3	PAH4	PAH5	PAH6	PAH7	PAH8	PAH9	PAH10	PAH11
PAH3	1.00								
PAH4	0.99	1.00							
PAH5	0.95	0.90	1.00						
PAH6	0.90	0.82	0.98	1.00					
PAH7	0.95	0.91	0.98	0.95	1.00				
PAH8	0.97	0.92	0.98	0.95	0.99	1.00			
PAH9	0.93	0.84	0.95	0.94	0.97	0.98	1.00		
PAH10	0.87	0.75	0.86	0.85	0.91	0.93	0.97	1.00	
PAH11	0.80	0.71	0.79	0.77	0.86	0.89	0.93	0.97	1.00

Correlation Matrix for Non-River Sites
Table 3.6.11

		Factor Analysis				
Variable	Commun		Factor	Eigen value	% of var	Cum %
PAH3	1.00		1	8.25	91.6	91.6
PAH4	1.00		2	0.48	5.3	96.9
PAH5	1.00		3	0.21	2.3	99.2
PAH6	1.00		4	0.04	0.4	99.6
PAH7	1.00		5	0.02	0.2	99.8
PAH8	1.00		6	0.06	0.1	99.9
PAH9	1.00		7	0.01	0.1	100.0
PAH10	1.00		8	0.00	0.0	100.0
PAH11	1.00		9	0.00	0.0	100.0

Factor Analysis for Non-River Sites
Table 3.6.12

3.6.6.2 River affected sites

The river affected sites consisted of a matrix of twenty sites described by twelve PAH's. Again, PAH's 1,2 and 12 were omitted from the analysis.

The correlation matrix is shown in table 3.6.13 and suggests that the correlation between PAH's 10 and 11 (perylene and benzo(a)pyrene) and the others in the set is low. Examination of the raw data shows that the amounts of these two components found in all samples was relatively low.

Continuation of the factor analysis produces the eigenvalues shown in table 3.6.14. The SPSSPC default procedure suggests that two factors are responsible for the data, as was found for the original set. However, these two factors only explain 87.3% of the variance in the data set. The remaining 12.7% due to analytical error is, arguably, too much and it is possible that a third factor could be present. The scree plot in figure 3.6.7 also suggests the possibility of a third factor.

There is a theoretical possibility of a third source of PAH's in this data set as all of the river affected sites are also near a large, toxic waste incinerator. It would be interesting to investigate further sites that were above the flood plain so as to remove the major factor.

			Correlation Matrix						
	PAH3	PAH4	PAH5	PAH6	PAH7	PAH8	PAH9	PAH10	PAH11
PAH3	1.00								
PAH4	0.79	1.00							
PAH5	0.91	0.85	1.00						
PAH6	0.80	0.78	0.71	1.00					
PAH7	0.86	0.96	0.92	0.74	1.00				
PAH8	0.84	0.95	0.88	0.73	0.99	1.00			
PAH9	0.82	0.67	0.69	0.73	0.77	0.79	1.00		
PAH10	0.03	-0.23	-0.12	0.01	-0.14	-0.11	0.27	1.00	
PAH11	0.04	-0.10	-0.07	0.12	-0.02	0.01	0.41	0.71	1.00

Correlation Matrix for River Affected Sites
Table 3.6.13

		Factor Analysis				
Variable	Commun		Factor	Eigen value	% of var	Cum %
PAH3	1.00		1	5.93	65.9	65.9
PAH4	1.00		2	1.93	21.4	87.3
PAH5	1.00		3	0.36	4.0	91.3
PAH6	1.00		4	0.35	3.8	95.1
PAH7	1.00		5	0.22	2.5	97.6
PAH8	1.00		6	0.16	1.8	99.4
PAH9	1.00		7	0.04	0.5	99.8
PAH10	1.00		8	0.01	0.1	99.9
PAH11	1.00		9	0.00	0.1	100.0

Factor Analysis for River Affected Sites
Table 3.6.14

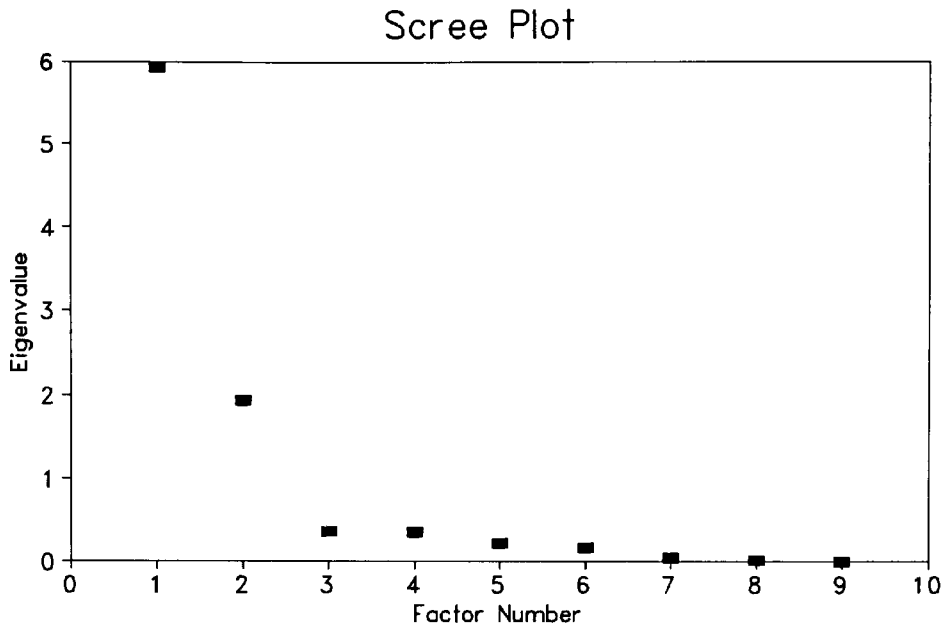


Figure 3.6.7

3.6.7 Profile Analysis

As a final test the PAH profiles of several sites were plotted against each other. Sites 1 and 7 (non-river), sites 19 and 23 (river) and sites 1 and 23 (both) were compared. Least squares linear regression was used to provide the best fits to the data, the results of which are shown below.

Sites	X coeff	Const	R^2
1 and 7	0.910	0.05	0.846
1 and 23	3.692	0.71	0.376
23 and 17	1.538	-0.16	0.983

It is evident that the non-river sites are similar and that the river sites are remarkably similar. However the different sites show negligible correlation.

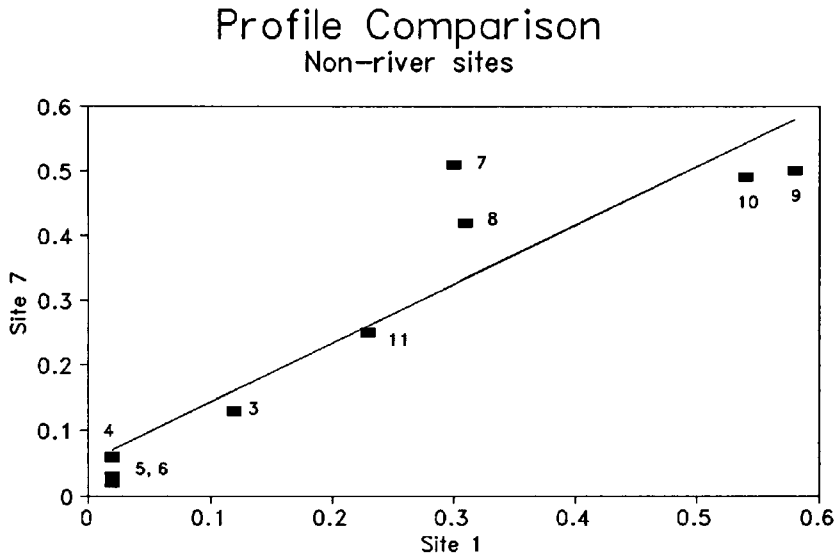


Figure 3.6.8

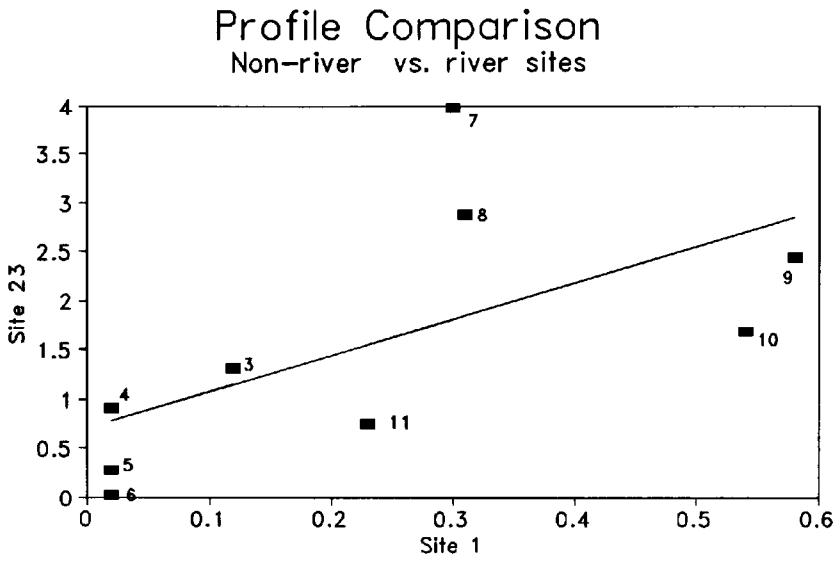


Figure 3.6.9

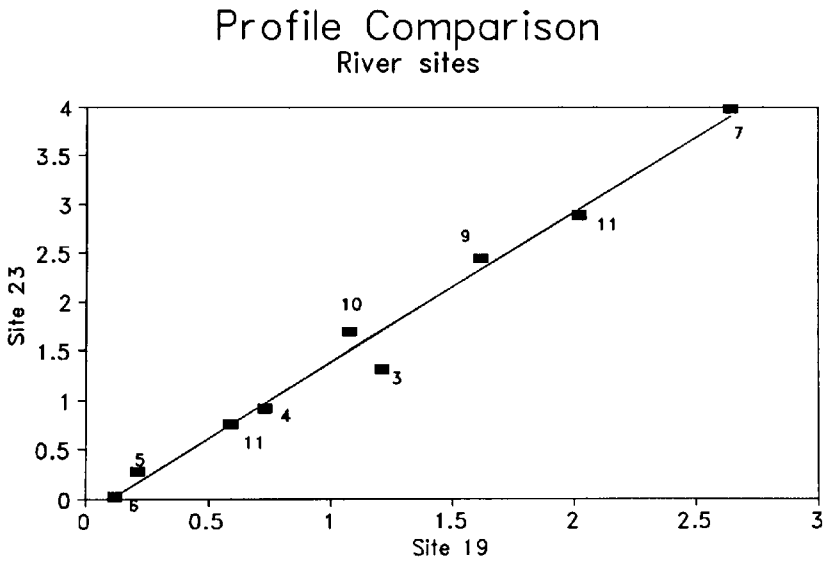


Figure 3.6.10

3.6.8 Discussion of Multi-variate Data Reduction

Several procedures were applied to the data set and all suggested that there were two sources which produced the PAH profile for the area examined. There is very strong evidence that the Afon Llwyd, which flows through the area, is the major source of the PAH's found in its vicinity. Not only is this evident from the multivariate data but also its effect can be seen by the much higher quantitative results found in sites affected by flooding or by a high water table.

Further evidence is found when the data set is split into subsections. Principal components analysis of sites not affected by the river shows that only one component is present. This could be explained by general "background" PAH's in the atmosphere. These are well mixed before deposition and subsequent adsorption to soil particles. Although there may be many sources, they are well diluted. It may be possible to investigate individual sources by analysis of air samples and further chemometric work.

Data analysis of the river affected sites again showed that two components were present although there was a reasonable suggestion that a third factor could be here. This is not totally unexpected as all the sites were within about 1 kilometre of a large toxic waste incinerator. PAH's have been identified as products of incomplete combustion in stack gases by Trenholm (106) and in bottom ash and scrubber effluent by Boegel (107). If the profile of the PAH's emitted as products of incomplete combustion is similar to the general background, which is also due mainly to burning processes, then it will be

difficult to differentiate and find them. Again, direct air sampling may be of some interest or possibly the investigation of sites in the area which are raised of the flood plain.

However, the main source of the PAH's was shown to be the Afon Llwyd. Again this is not surprising. The river's source is in the Blaenavon area of Gwent and it then flows south through Pontypool and Cwmbran before finally joining the Usk at Caerleon. In the recent past the areas around Blaenavon and Pontypool ie upstream of the sampling area, were extensively mined for coal. Local people have reported that, many years ago, there were days when the river "ran black". It is reasonable to assume that this was the original source of the pollution.

Two anomalous results arose from the multi-variate work on the full data set. It was found that site 2 and site 9 were clustered with the river sites although they were both well removed from this source of pollution. Site 9 can be explained as it was found that the owner of the property and treated his garden with coal products. It could therefore be expected to show the same profile as the river affected sites. Site 2 is more difficult to explain. The site is at the side of a busy through road so is likely to be affected by PAH's produced from vehicle emissions. It is also near to a crematorium and is thus liable to contamination by products of incomplete combustion.

Edwards (108) has suggested that the typical total PAH level in soils is in the range 1 to 10 $\mu g/Kg$ (1 - 10 ppb), resulting from plant synthesis and natural fires. On this basis even the sample taken from a relatively remote farm have been

significantly contaminated by atmospheric deposition. Further investigation may be warranted to discover if the pollution in the area sampled is generally higher than elsewhere in Wales and, if so, what the cause may be.

3.6.9 Use of Chemometrics in Environmental Analysis

Chemometrics are, potentially, a very powerful tool for analysis of analytical data. However, it must be remembered that a problem should be formulated before the mathematical methods are applied. This, of course, applies to all statistical procedures.

The success of the technique, when applied to environmental problems, has been quite varied to-date. Hopke (5) (6) used the technique to analyze data from Boston and New York (soil and sediment respectively) and he was able to find clusters although they were fairly gross. For Boston he found two major clusters which corresponded to inner and outer city sites, although there was a possibility of smaller clusters contained within them. For the New York lake the clusters were based on the composition of the sediment tested.

Similar results were obtained by Hoban (143) who examined sites in the Severn estuary for contamination by PAH's. He found relatively large amounts of these compounds and applied a wide range of techniques to the data. His work showed that only one factor was affecting the data and that it was the organic content of the sediment. This was felt to be surprising at the time as there are many potential sources of PAH's flowing into the estuary. He pointed out that the estuary

profile may not reflect the source profile as it may have been subject to modifying influences such as microbial or photolytic degradation. The actual amount of PAH adsorbed onto the sediment would depend upon its partition coefficient rather than its source.

It should also be noted that a great deal of tidal mixing occurs in the estuary and this would produce an amorphous mix of PAH's leading to partition coefficient being the only mechanism affecting adsorption. This problem will also exist for air samples when gross pollution sources are not present. This can be envisaged such that if there are several major pollution sources affecting an area then there is a good chance that the amount of pollution emitted from each site will vary from day to day, leading to a variation in the ration of pollutants. If daily samples are taken then the differences will be made apparent by chemometric analysis. However, there will always be a general background of well-mixed pollutants such as vehicle exhaust fumes and the general atmospheric loading. The origin of these pollutants is unlikely to be determined by multi-variate techniques.

Soil sampling may give rise to data which can be subjected to chemometric analysis but again will probably require fairly gross sources of contamination to be effective. In this case the sampling would not be one spot analyzed over a period of time but rather a number of spots sampled at various distances from potential pollution sources. If there were several sources and sample sites chosen such that "pre-mixing" was not a problem then it may be possible to produce useful data.

Pardo et al(144) analyzed soil samples around a lead smelter located in a semi-rural area in Spain. They determined various heavy metals in the soils and applied ANOVA and principal components analysis to the resulting data. Their work showed a significant correlation between metals and wind direction.

A similar study was carried out by Einax et al (145) who looked at soil samples in the vicinity of a metallurgical plant and a chemical plant in Germany. They applied cluster and factor analysis and were able to distinguish between areas near to each plant and unpolluted sites.

The obvious difference between these later reports and earlier, less successful, work is that they are based on inorganic, rather than organic, pollutants. This reinforces the idea of partition coefficient being a dominant factor in organic, environmental analysis.

Another problem with analysis of environmental samples for organic components is that they are present in very small quantities, typically sub parts per million levels, and are often likely to be masked by many other components. This leads to considerable variance due to analytical error which can be of the order of 10%. Thus, in principal components analysis, a factor appearing to explain 10% of the variance may actually be explaining just the analytical error and not a real, environmental factor.

The work reported in this section of the thesis found the presence of two factors which were well defined by the chemometric techniques used. The obvious explanation for this

is that there was no pre-mixing possible before adsorption. One source was general air-borne contamination and the other was the water-borne, river source.

The two anomalous sites could also be explained by direct contamination of the land, one by application of coal products to a garden and the other by extremely close proximity to a crematorium.

The "non-river' sites showed just one source which was the well-mixed, atmospheric loading. This was analogous to the conditions found in the estuary.

The sites close to the toxic-waste incinerator suggested the presence of a third factor. This could possibly have been due to the incinerator as emissions from this source would only have limited mixing with the general background before deposition.

In conclusion, whilst chemometric techniques are very powerful the technique should be applied with care and the results used to make predictions fitting the known facts.

SECTION FOUR

Investigation into the Presence of Nitrated Polyaromatic Hydrocarbons in a South Gwent Valley

4.1.1 Introduction

During section three the presence of high levels of PAH's was demonstrated. There is a steel works approximately 1Km from the main sampling area and this factory often produces clouds of NO_x from its pickling line. It is feasible that these gases may react with the PAH's to form their nitro derivatives.

4.1.2 Nitrated Polyaromatic hydrocarbons

Generally, substitution reactions in multi-ring systems can be quite complicated. The position of the first substitution may be obtained from a consideration of the nature of the intermediate carbonium ion, the one with the largest number of resonating structures being the more stable.

Using naphthalene as an example, substitution in the 1-position gives rise to four resonating structures whereas 2-substitution gives only two as shown below in fig.4.1

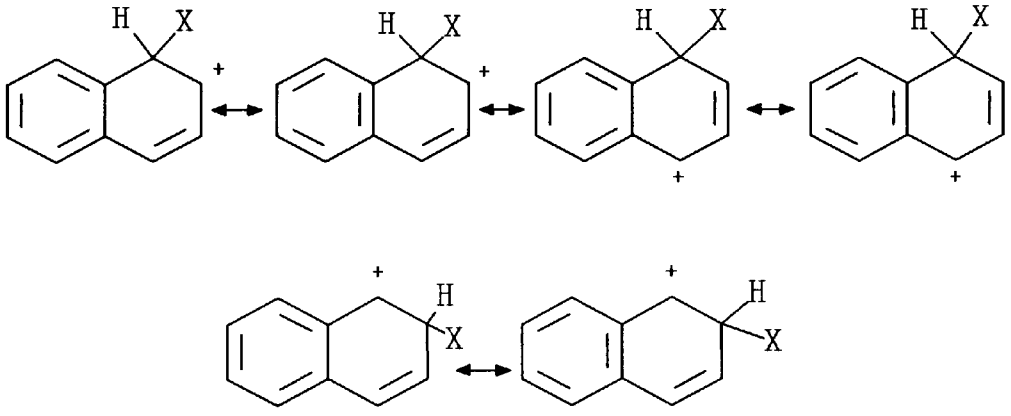


Figure 4.1

Introduction of a second substituent can give rise to homonuclear substitution, where the second group enters the same ring as the first, or to heteronuclear substitution where

it enters another ring.

Thus the first nitration product of naphthalene is 1-nitronaphthalene and more vigorous conditions give rise to a mixture of 1,5 and 1,8 dinitronaphthalenes.

In the case of anthracene, 9-nitroanthracene and 9,10 dinitroanthracene may be formed. However, nitration with aqueous nitric acid produces the oxidation product anthraquinone.

Nitration of phenanthrene gives rise to a mixture of three mono-nitro isomers, the 3-nitro form predominating.

4.1.3 Environmental Formation

In 1978 Pitts et al (109) reported the nitration of benzo(a)pyrene and perylene on exposure to simulated atmospheres containing 1ppm of NO_2 and traces of nitric acid. The studies were carried out with the parent PAH deposited on the surface of glass fibre filters. Pitts pointed out that the reaction may be different if the PAH is adsorbed on the surface of airborne particles such as soot or fly-ash. His experiments were also carried out in the dark. It has been shown by several groups of workers that photolysis is one fate of nitro-PAH's. For example in 1966 Chapman et al (110) showed that, upon photolysis, 9-nitroanthracene forms 9,10-anthraquinone in solution. Pitts (109) found the same reaction to occur on silica gel. Pierce and Katz also found polycyclic quinone in polluted air in Toronto, Canada (111).

Other major sources of nitro-PAH's have been reported, Scheutzle et al (112) found the compounds in diesel exhaust, Jager (113) found the in air particulates and Fitch et al (114)

investigated them in carbon black.

4.1.4 Toxicity

Many of the nitro-PAH's are active in standardized, short-term tests for mutagens and carcinogens, as reported by Rosenkranz et al (115). Wei (116) has also found that they are carcinogens in a number of mammalian species. Ohgaki et al (117) have found that nitro and dinitropyrenes can induce tumours at the injection site in experimental animals. Nitropyrenes were found by Howard et al (118) to cause in-vitro cell transformations in human fibroblasts.

Nitro-PAH's adhering to particulate matter can be inhaled into the lungs and the possibility exists of swallowing sputum containing particulates removed during inhalation. Kinouchi (119) postulated that particulates and food containing nitro-PAH's would reach the lower intestine where they could be reduced by normal intestinal bacteria to produce amino-PAH's.

It was a study of the carcigenicity of atmospheric organic matter that led to Pitts' paper (109) where he showed that nitro-PAH's could be formed by the reaction of nitrogen dioxide. The carcigenicity of the organic matter had been thought to be due to benzo(a)pyrene and the other PAH's (120). However, several investigators, including Grimmer (121), have noted that the carcigenicity of both urban air and exhaust particulates could not be accounted for solely on the basis of the PAH's found. Thus there was considerable "excess" carcigenicity, presumable due to unidentified compounds present. Pitts (122) has suggested that the nitro-PAH's, present in the atmosphere at extremely low levels, make a

significant contribution to this "excess" carcigenicity.

The biological activity of a nitro-PAH appears to depend upon its molecular structure. For example, 2-nitro naphthalene has been found to cause cancer in animals whereas the 1-nitro isomer does not, as was found by Poirier (123)

4.1.5 Analysis of Nitro-PAH's

Due to the extremely low levels of these compounds found in the environment, very sensitive analytical methods must be employed for their determination.

Combustion emissions are a good example of the extreme complexity of the nitro-PAH mixtures encountered in environmental samples. Paputa-Peck et al (124) detected more than 100 nitro-PAH's in a diesel exhaust particulate extract. They used capillary Gas Chromatography for the analysis and this is generally accepted as the technique with the greatest potential for resolution and detection of these compounds.

High chromatographic resolution is required as the nitro-PAH's are found in the environment together with other pollutants of similar polarity. It is therefore not an easy clean-up as the normal techniques eg chromatography on silica, will extract and concentrate all compounds of similar polarity more or less equally.

Analytical development has been further hampered by the lack of pure analytical reference standards and by the lack of "target" nitro-PAH analytes such as the fourteen priority pollutant PAH's as selected by the US Environmental Protection Agency.

The first reported use of GC for the analysis of nitro-

PAH's was reported by Streitwieser and Fahey (125) who analyzed the products resulting from the nitration of fluoranthene and naphthalene. They used a 10ft column packed with Chromosorb coated with silicone rubber and operated at 330°C.

Draper (126) used a capillary GC column of 30m x 0.32mm with a chemically bonded 0.25μm SPB-5 silicone phase to examine diesel exhaust particulates, detection was by electron capture. He extracted the sample with dichloromethane and followed this up by chromatography, first on a silica "Sep-pak" cartridge and finally by HPLC on a silica column.

In 1985, Korhonen and Lind (127) investigated the GC separation of a mixture of PAH's and nitro-PAH's. They utilized an SE-30 quartz capillary column and flame ionization detection. They were not working on environmental samples but were developing a system of identification of nitro-PAH's by means of relative retention indices.

White et al (128) also investigated the retention characteristics of nitro-PAH's on an SE-52 fused capillary column. Again, flame ionization detection was used as environmental samples were not being examined.

Ramdahl and Urdal (129) used a 30m DB-5 capillary column for the analysis. They used both electron capture and negative ion mass spectrometry for detection and found that the mass spectrometer was better in terms of resolution. Using single ion monitoring they found a detection limit of 1pg injected on column.

Robbat (130) commented on the lack of available nitro-PAH standards. He also used capillary GC for their analysis on an

SE-52 stationary phase. He then applied a multi-variate approach in order to predict retention index. He used the concept of molecular connectivity, which is a description of molecular structure based on a count of skeletal atom groupings weighted by the degree of branching. Stepwise linear regression produced an equation between retention index and molecular connectivity. A six-descriptor equation encoded sufficient information about nitro-PAH's that 19 out of 23 compounds in the data set were predicted within +/- 2.1 units of their observed value.

Robbat (131) also evaluated a chemiluminescence detector, again for use with capillary GC. The detector works as follows; after GC separation the nitro-PAH enters a pyrolysis chamber at 1000°C and decomposes to form nitrosyl radicals (NO•) and other products. The products are drawn by vacuum into a cold trap where most of the remaining fragments are frozen out. The nitrosyl radical continues on to a reaction chamber under reduced pressure where it reacts with ozone to produce a characteristic infra-red chemiluminescence. The process is highly selective due to the characteristics of the cold trap, photomultiplier tube and the red filter employed. The detection range was found to be linear between 100 ng and 50 pg.

Tomkins et al (132) analyzed diesel exhaust particulates. They extracted the organic compounds by ultrasonication with toluene solvent. Clean-up was on an aminosilane column and final analysis by packed column GC (3m 3% OV-17) with detection by a thermal energy analyzer. They found a detection limit of 0.3ng at a S/N of ca.2.

Jager (133) separated the nitro-PAH's using thin layer chromatography. Detection was by fluorescence quenching. He used two separation systems, one consisting of silica gel with cyclohexane - chloroform (1:1) and the second of cellulose with N,N dimethylformamide - water (2:3). The plates were developed and sprayed with a reducing agent. The non-fluorescent nitro compounds were reduced to fluorescent amines. The resultant spots were examined under ultraviolet light. After drying, the plates were sprayed with a quenching reagent and the changes in fluorescence observed under a 365nm light source. The detection limits were 1ng for 1-nitropyrene, 100ng for 1-nitrochrysene and 500ng for 1-nitronaphthalene.

Chou (134) purified isomers of nitro-PAH's in order to asses their biological significance. He used a Pirkle column packed with R-N-(3,5 dinitrobenzoyl)phenylglycine covalently bonded to aminopropylsilanized silica particles. The elution solvent was a mixture of ethanol, acetonitrile and hexane.

The technique of chiral chromatography is usually used to resolve optically active enantiomers, but in this case it separated 1- and 2-nitrobenzo(a)pyrene and also 1- and 3-nitrobenzo(a)pyrene.

Normal phase HPLC was used by Nielsen (135) to isolate PAH's and their nitro derivatives from complex samples. He extracted samples of airborne particulate matter, ultrasonically, with dichloromethane. The concentrated extract was chromatographed on a silica column with n-hexane:benzene mobile phase. The fractions consisting of PAH and nitro-PAH were collected, concentrated and final analysis was carried out

by GC. He tested the HPLC system with standard compounds and found that aliphatic hydrocarbons eluted first, followed by the PAH fraction which also contained traces of oxaarenes and thiaarenes. The mononitro-PAH fraction eluted just after the PAH's and could have contained alkoxy-PAH's. Other nitrogen containing compounds had higher retention times.

MacCrehan et al (136) investigated three different detection approaches for the analysis of diesel particulate extracts. The basic chromatographic system consisted of a reversed-phase octadecylsilane column with gradient systems of acetonitrile or methanol and a buffer. Differential pulse detection gave a detection limit of 260ng/mL for 1-nitropyrene. Amperometric detection had a detection limit of 3.0ng/mL and fluorescence detection, after reduction to the corresponding amine, 0.5ng/mL.

4.1.6 Extraction and Clean-up

Nitro-PAH's are usually present in trace amounts in the presence of much larger quantities of interfering substances, thus clean-up is a necessary step in their analysis. Most of the work reported to date has been carried out diesel particulates. Extraction has been very similar to that reported for PAH's, viz. ultrasonication or Soxhlet extraction using dichloromethane or toluene solvent.

Low pressure liquid chromatography is often used to fractionate large samples due to its low cost. Yu et al (137) used an alumina column with a step-wise gradient of hexane, toluene, chloroform and methanol to fractionate a diesel exhaust particulate sample. The nitro PAH's eluted in the

toluene and chloroform fractions. Final analysis was by GC with a nitrogen specific detector.

Newton et al (138) used medium pressure chromatography, again to fractionate diesel particulate extracts. A silica column was used and elution solvents were dichloromethane in hexane, dichloromethane, acetonitrile and methanol. The nitro-PAH's eluted in the early fractions.

Oehme et al (139) used HPLC to extract the nitro-PAH's from extracts of atmospheric aerosols. They used Lichrosorb silica with elution by cyclohexane, dichloromethane and acetonitrile.

D'Agustino (140) developed an unusual separation system. The concentrated extract was passed through a short silica "Sep-pak" cartridge using dichloromethane eluent. The portion eluting between 1.0 and 4.5 ml was collected. This fraction was free of the most polar compounds. The fraction was re-concentrated and chromatographed on a Sephadex LH-20 column, eluting with methanol / dichloromethane (1:1). The PAH's and nitro-PAH's eluted as a narrow band between 15 and 25 ml. This fraction was concentrated to dryness, redissolved in dichloromethane / hexane and separated using normal phase HPLC on amino or cyano bonded silica with a mobile phase of 0.5% propan-2-ol in hexane.

Cambell and Lee(141) demonstrated a separation using silica chromatography combined with potassium borohydride reduction of the nitro group to the corresponding amine and further derivatization to the pentafluoropropylamide. The derivatized sample was analyzed by capillary GC using electron

capture, nitrogen sensitive and flame ionization detectors.

4.2.1 Development of Analytical Procedures

The proposed samples were known to contain relatively high levels of PAH's, of the order of 20 ppm. If nitro-PAH's were present it was likely that they would be so in tens of ppb. After extraction the next step would be the separation of the two classes of compound. It was decided to do this by reduction of the nitro-compounds to the corresponding amine. These amines would be considerably more polar than the PAH's so should be separable using a reversed-phase cartridge system. The final chromatography would be on the reduced samples so could be HPLC with conditions tailored for amino compounds.

4.2.2 Extraction

The extraction system employed the same procedure as that used for the PAH analyses except that larger sample weights were taken.

The soil samples had all been stored at 4°C in amber glass bottles after they had been dried, under vacuum, at room temperature using an air bleed. The samples were protected from light at all times.

The dried soils were ground using a pestle and mortar. Stones and vegetation were removed and approximately 15 grammes, accurately weighed, were transferred to a 50ml centrifuge tube. Dichloromethane (15mls) was added and the mixture agitated. The tube was suspended in an ultrasonic bath and sonicated for 15 minutes with mechanical agitation every five minutes. On completion the mixture was centrifuged for two minutes and the supernatant aspirated into a beaker. The

process was repeated a further four times and the combined dichloromethane extracts filtered through a Whatman GF/A filter. The filtrate was evaporated to dryness at room temperature, under a stream of nitrogen and protected from light.

4.2.3 Clean-up

Clean-up was performed by reduction followed by solid-phase extraction.

The extracted residue was mixed with methanol (1ml) and water (2ml) was added. Further additions were made of iron powder (ca.50mg) and hydrochloric acid (1ml). The reduction was allowed to continue for 15 minutes at room temperature and protected from light.

In the meantime an octadecyl bonded silica column (500mg) from Analytichem International (Jones Chromatography, Hengoed, Mid.Glam.) was prepared. It was conditioned by flushing with one column volume of methanol and two of water. After the methanol addition the column was not allowed to become dry until after the addition of the sample.

The reduced sample mixture was applied to the column and the vial washed out with water (5ml) which was used to flush the column. The reduced compounds were eluted using 2 x 500μl aliquots of acetonitrile:phosphoric acid:water (85:1:14). The eluate was collected and the volume adjusted to 1.0 ml in a volumetric flask.

Two other reduction techniques were investigated. The first used tin with hydrochloric acid. This would have been convenient as any excess tin could have been dissolved in

sodium hydroxide and therefore not caused any problem with blocking the extraction tube. In practice, the procedure was not effective. It is thought that the tin / amine complex was too stable for reaction with sodium hydroxide.

A second system used palladium on charcoal as a catalyst for reduction by triethylamine and formic acid. Again, this did not appear to be satisfactory as the reaction did not give 100% yields.

4.2.4 HPLC Procedure

The chromatography was run on the amino-PAH's, thus a mobile phase was used which was modified to retain charged analytes. A system which is regularly used for basic compounds utilizes an octadecylsilyl bonded column and a mobile phase modified by the addition of triethylamine and buffered to pH 3.0.

The final system used was:-

Column: 125 x 4 mm Superspher ODS 4μm (E.Merck)

Mobile phase: Acetonitrile / 0.5% TEA at pH 3.0

Flow-rate: 2.0 mls / min

Detection: 230 nm

Injection: 100μl loop

The hardware used consisted of two CM3200 pumps, an SM5000 diode array detector, a PROMIS II autosampler and a thermochrom data handling system (all from LDC Analytical, Stone, Staffs).

4.2.5 Validation Studies

A validation study was carried out to demonstrate that the reduction and solid phase extraction did, in fact, work. The

procedure was also used to give an indication of linearity and limit of detection.

As discussed previously, there are very few reference nitro-PAH's available. Accordingly, five compounds were chosen as models. These were (in order of elution after reduction):-

2-nitronaphthalene
1-nitronaphthalene
3-nitrobiphenyl
2-nitrobiphenyl
1-nitropyrene

Approximately 3mg of each compound was accurately weighed and transferred to a 100ml volumetric flask. Methanol (40ml) was added and the mixture sonicated to effect dissolution. The resulting solution was diluted to volume with water.

A 2.0 ml aliquot was transferred to a vial, iron powder (ca.200mg) and hydrochloric acid (2ml) were added and the reduction allowed to proceed for 15 minutes at room temperature and protected from light.

A C18 Bond-elute column was prepared as described in section 4.2.3 and the contents of the reduction vial transferred using water (5ml) to wash out the vial and flush the column.

The amino compounds were eluted from the column using acetonitrile/phosphoric acid/water (85/1/14). The eluate was collected and diluted to 50.0 ml using 40% acetonitrile. A series of aliquots were further diluted to 20.0 ml with the same solvent, as shown in table 4.1. Sixfold replicate injections were made at each level and area counts determined.

Compound	wt (mg)							
1-nitronaphthalene	2.98							
2-nitronaphthalene	3.27							
2-nitrobiphenyl	3.12							
3-nitrobiphenyl	2.76							
1-nitropyrene	3.42							
Dilutions	0.5	1.0	2.0	3.0	4.0	5.0	20.0	

Table 4.1

The results and regression plots are shown below in tables 4.2 to 4.6 and figures 4.4 to 4.8. Also shown below are:-

a) A comparison of reduced and unreduced compounds under the same chromatographic conditions (Fig 4.1)

b) The chromatogram of the reduced compounds (Fig 4.2)

In figure 4.1 the lower trace is from the reduced compounds and the upper from the parent nitro-PAH. The order of elution was unchanged but retention time was considerably reduced indicating the increased polarity of the reaction products.

The upper trace only shows three peaks, 2-nitrobiphenyl and 1-nitropyrene were retained by the column under these conditions which had been set for the determination of the amino-PAH's.

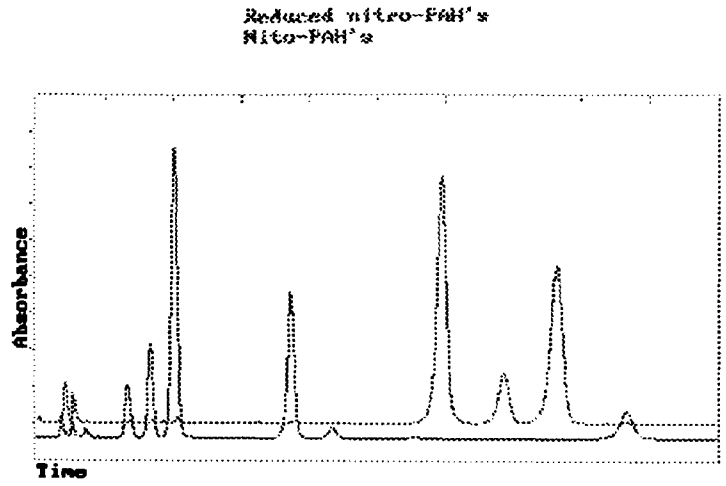


Figure 4.1

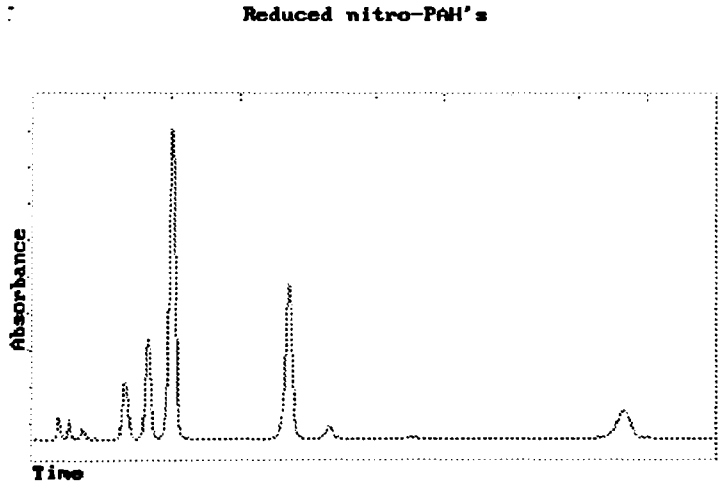


Figure 4.2

2-nitronaphthalene								
ng/ml			Area	counts			Mean	RSD
32.7	624	620	634	661	603	601	624	3.55
65.4	1170	1142	1198	1148	1174	1284	1186	4.39
130.8	2338	2289	2267	2298	2230	2304	2288	1.60
196.2	3347	3388	3349	3344	3427	3637	3415	3.32
261.6	4486	4473	4478	4458	4710	4404	4502	2.36
327.0	5641	5635	5659	5637	5697	5676	5658	0.44
1308.0	22758	23086	23085	22936	22993	24149	23168	2.14

Table 4.2

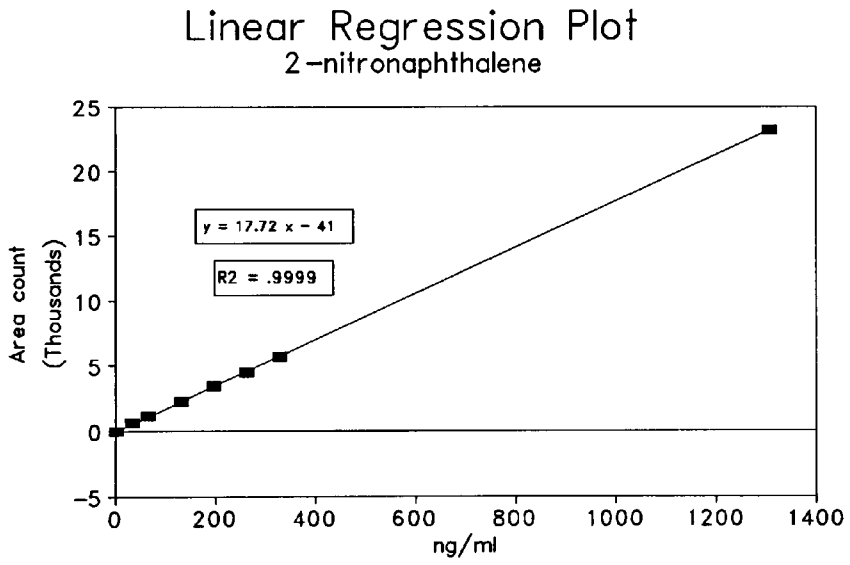


Figure 4.3

1-nitronaphthalene								
ng/ml			Area	counts			Mean	RSD
29.8	219	241	248	218	225	212	227	6.26
59.6	374	368	404	381	370	346	374	5.05
119.2	776	755	685	722	720	698	726	4.71
178.8	1011	1079	1053	1078	1091	1016	1055	3.25
238.4	1361	1351	1329	1391	1381	1389	1367	1.79
298.0	1691	1595	1622	1586	1606	1589	1615	2.45
1192.0	6858	6935	6960	6859	6908	6983	6917	0.75

Table 4.3

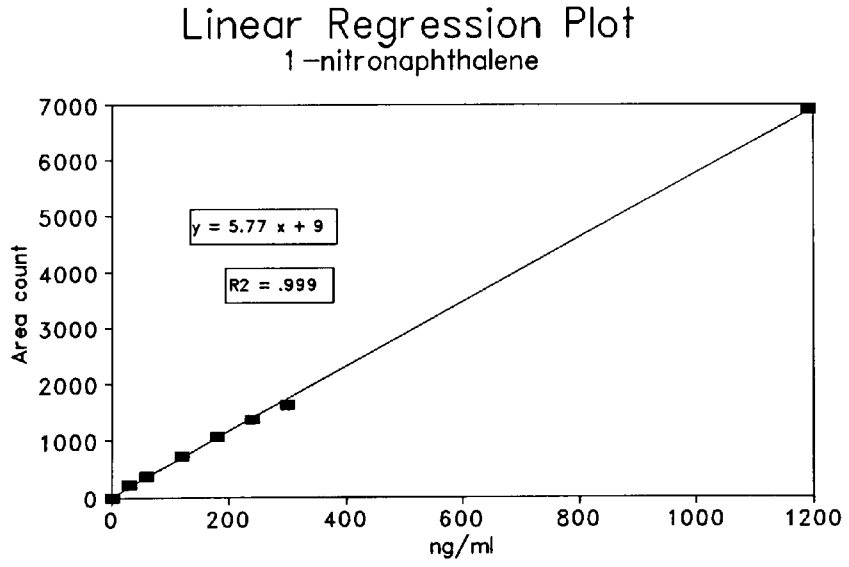


Figure 4.4

3-nitrobiphenyl								
ng/ml			Area	counts			Mean	RSD
27.6	1004	988	1013	1071	961	955	999	4.23
55.2	1908	1909	1907	1889	1842	1885	1890	1.36
110.4	3785	3793	3724	3746	3769	3776	3766	0.69
165.6	5514	5613	5576	5607	5567	5625	5584	0.73
220.8	7393	7350	7375	7388	7407	7351	7377	0.31
276.0	9273	9091	9081	9063	9063	9052	9104	0.92
1104.0	37365	37323	37386	37247	37336	37307	37327	0.13

Table 4.4

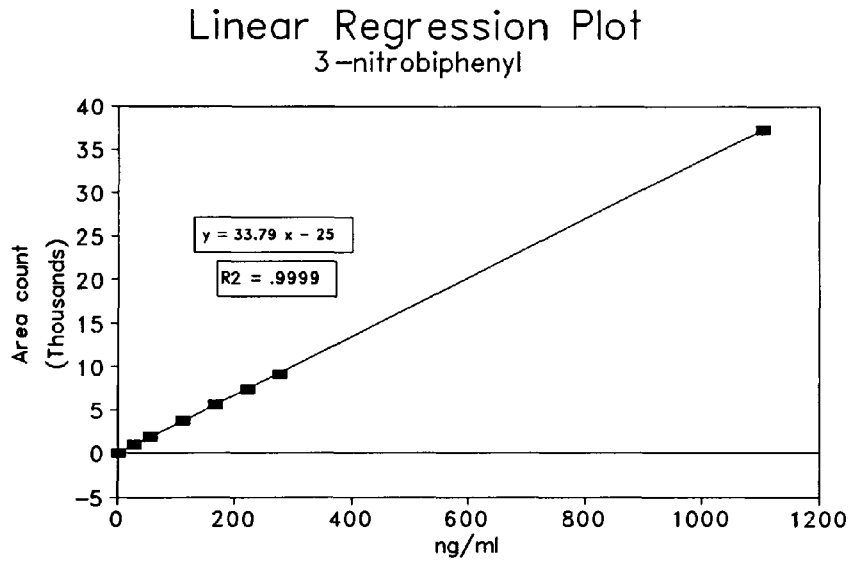


Figure 4.5

2-nitrobiphenyl								
ng/ml			Area	counts			Mean	RSD
31.2	616	654	677	691	616	700	659	5.58
62.4	1281	1307	1275	1325	1243	1335	1294	2.66
124.8	2758	2591	2638	2701	2634	2778	2683	2.78
187.2	4123	4215	4070	4123	4062	4111	4117	1.33
249.6	5494	5371	5439	5403	5465	5428	5433	0.80
312.0	6798	6732	6855	6732	6735	6674	6754	0.93
1248.0	27829	27958	28113	27964	27921	27928	27952	0.33

Table 4.5

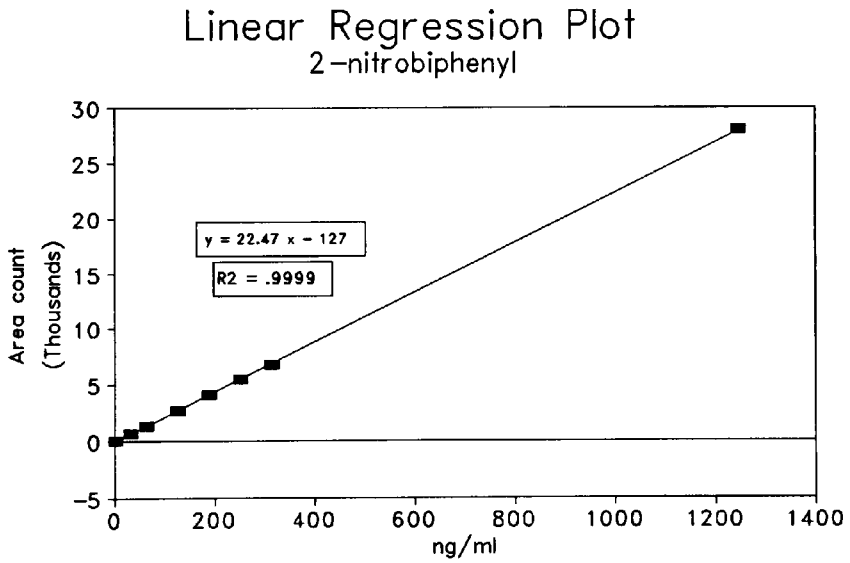


Figure 4.6

1-nitropyrene								
ng/ml			Area	counts			Mean	RSD
34.2	1486	1364	1531	1330	1337	1325	1396	6.43
68.4	2468	2319	2436	2262	2584	2375	2407	4.76
136.8	4679	4082	4351	4081	4621	4360	4362	5.84
205.2	9830	9973	9975	9839	9851	9793	9877	0.79
273.6	13663	13551	13242	13405	13530	13937	13555	1.74
342.0	17034	16939	16436	16975	16348	16948	16780	1.81
1368.0	69713	70980	69640	69546	69970	69382	69872	0.83

Table 4.6

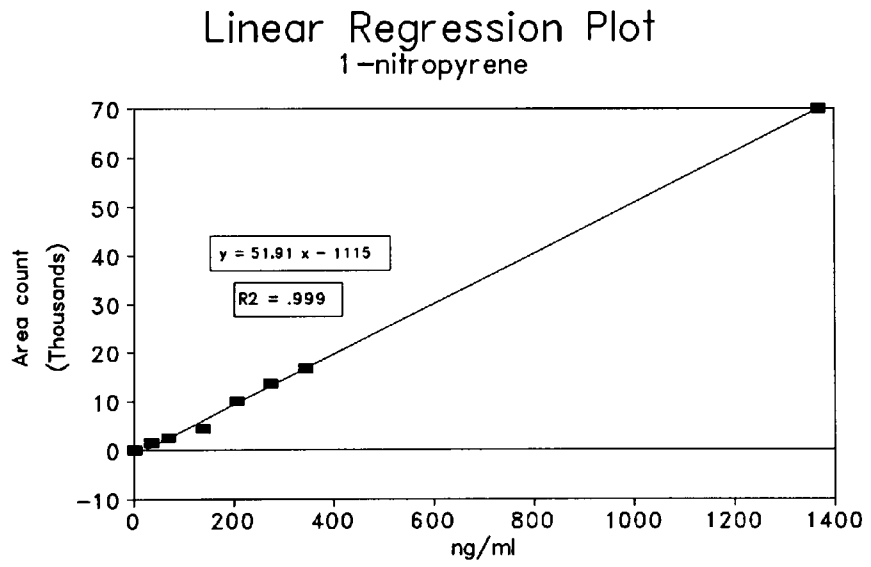


Figure 4.7

As discussed in section 3, the limit of determination may be calculated using the following equation:-

$$LOD = \frac{6 \times R \times C}{100 \times W}$$

where R = RSD of the lowest calibration point

C = Concentration of lowest calibration point (ng/ml)

W = Proposed sample weight (g)

This gives a determination limit in ng/gm (ppb). Using a proposed sample weight of 15 grammes and a final extract volume of 1.0 ml, the determination limits are:-

2-nitronaphthalene	0.5ppb
1-nitronaphthalene	0.8ppb
3-nitrobiphenyl	0.5ppb
2-nitrobiphenyl	0.7ppb
1-nitropyrene	0.9ppb

Using a series of model compounds it has been demonstrated that the proposed reduction and extraction procedure is viable. When combined with HPLC, response was found to be linear over a fortyfold range with good precision. Limits of determination were found to be less than 1ppb.

4.2.6 Application to Environmental Samples

The procedure as described above was applied to three of the soil samples taken for PAH analysis as discussed in section three. The three were chosen to provide as wide a cross-section as possible in a small number of samples. The three chosen were:-

a) Pentwyn Farm. This site showed a low level of PAH contamination and was unlikely to be affected by NO_x .

b) Roadside between Pontyhydrun Wood and Panteg Steelworks. Low level of PAH's but likely to be affected by NO_x. Also likely to be affected by diesel exhaust fumes.

c) Pontyfelin House. High level of PAH's, possibly affected by NO_x.

The chromatograms produced are shown below in figures 4.8 to 4.10. In all cases the upper trace is that from the reduced sample and the lower from an untreated sample.

The two chromatograms from the Pentwyn Farm sample (Fig 4.8) are virtually superimposable. They show a generally low level of contamination and no extra peaks are apparent in the trace from the reduced sample.

Figure 4.9 shows the pair of chromatograms from the roadside near the steel-works. Both reduced and untreated chromatograms show considerable contamination with early eluting, ie more polar, components. There appear to be some extra peaks early in the "reduced" chromatogram, which are probably due to amino-PAH's. However, there are so many other polar compounds present it is impossible to be sure.

Figure 4.10 shows the chromatograms from the garden of Pontyfelin House. This site is badly contaminated by water-borne PAH's and the relative amounts of more polar compounds are quite low. However, there do not appear to be any extra peaks in the "reduced" chromatogram.

The use of more advanced soft-ware having the capability of normalization then subtraction of the chromatograms would have made the representation of the results at this point more clear. Sufficient detail was apparent from inspection of the original graphic as seen on the computer display to make the assumptions discussed above.

Pentwyn Farm Sample

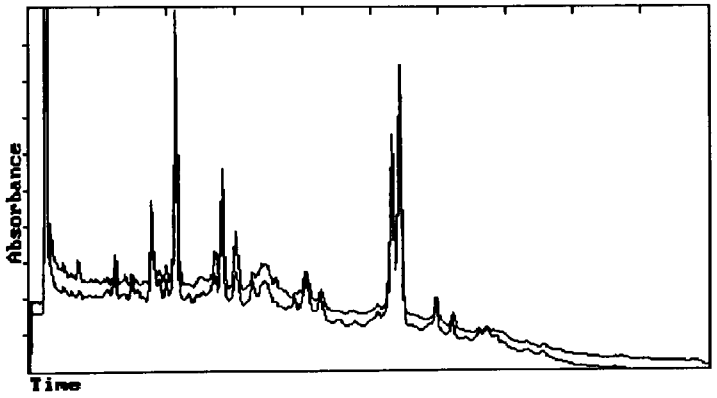


Figure 4.8

Roadside near Steelworks

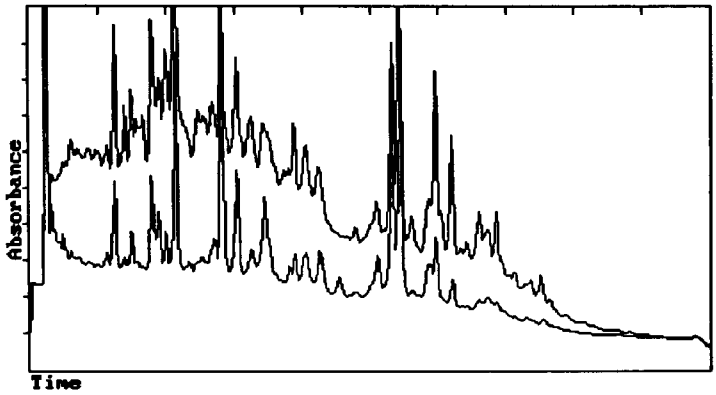


Figure 4.9

Pontyfelin House

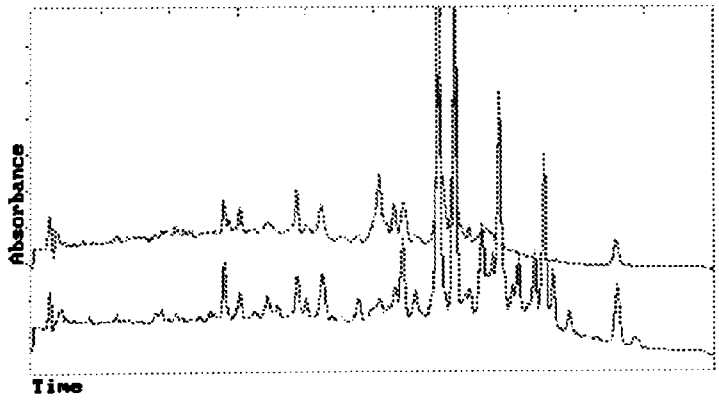


Figure 4.10

4.3 Investigation into the Formation of Nitro-PAH in Aqueous Media.

Of the environmental samples examined above, that from Pontyfelin House had the highest PAH level and was also affected by NO_x. However, there were no nitro-PAH's found, even given the lack of sensitivity of the method used. The following experiment was designed to investigate the formation of nitro-PAH's from their parent compounds in an aqueous / sediment environment.

4.3.1 Experimental

Nine model PAH's were chosen as typical environmental contaminants. They were:-

1) Anthracene
2) Benzo(a)pyrene
3) Benzo(e)acephenanthrylene
4) Carbazole
5) Chrysene
6) 4H-cyclopenta(def)phenanthrene
7) Naphthalene
8) Phenanthrene
9) Pyrene

A few mgs of each compound were transferred to each of five reaction vials together with about 1 gramme of silica, used to mimic sediment conditions. Each set of nine compounds was treated with 5mls of a different solution:-

1) 40% acetonitrile
2) 0.01% nitric acid
3) 50% nitric acid
4) Nitric acid / sulphuric acid
5) Hydrogen peroxide (20 vol)

The samples were allowed to stand for six weeks in the dark at room temperature, after which time they were neutralized, diluted with 40% acetonitrile and subjected to HPLC analysis.

It should be noted that the use of dark conditions did not mimic the real environmental situation however, it was felt that it would simulate the optimum conditions required to prevent any degradation of nitro compounds formed.

The HPLC system used consisted of the following:-

Column: 125 x 4mm Superspher ODS 4μm (E.Merck)

Mobile phase: A = acetonitrile / water 10 / 90

B = acetonitrile / water 90 / 10

Gradient: See below

Flow-rate: 2.0 mls / min

Detection: See below

In order to optimize the separations for the various mixtures, different gradients and detection wavelengths were required. These are shown below:-

Method #	Init % B	Final % B	Detection	Sample #
1	35	80	250 nm	1, 8
2	35	80	230 nm	4, 9
3	35	80	270 nm	7
4	35	100	265 nm	5
5	35	100	250 nm	6
6	50	100	250 nm	2, 3

All gradients were linear and over ten minutes. The area % results are summarized in the tables below (tables 4.8 to 4.16).Only the major peaks have been included.

At 0.01% nitric acid there appears to be no reaction with

PAH's. Reaction with the "normal" nitration mixture of nitric and sulphuric acids gives a mixed product of up to nine components. In all cases the retention time of all the nitrated products are less than that of the parent PAH. This indicates that the products are more polar than the parent, as would be expected from a nitration reaction.

The reaction with 50% nitric acid was different in each case.

a) Anthracene

Three new compounds produced, with retention times matching three of the products produced by conventional nitration.

b) Benzo(a)pyrene

Five new compounds, only one of which matches the retention time from a conventional nitration. All appear to be more polar than the parent.

c) Benzo(e)acephenanthrylene

No reaction

d) Carbazole

Six new compounds, none of which match the retention times of a conventional nitration. All but one elute later than the parent, suggesting that they are even less polar. One of the new compounds has a retention time matching that of an oxidation product.

e) Chrysene

Very little reaction with only one, more polar, compound being formed. No correlation with conventional nitration or oxidation.

f) 4H-cyclopenta(def)phenanthrene

Again, very little reaction with only three minor products, all slightly more polar than the parent, but with no correlation between conventional nitration or oxidation.

g) Naphthalene

Very little reaction with only one less polar compound being formed. Again no correlation with conventional nitration or oxidation.

h) Phenanthrene

Two, more polar compounds formed with no correlation to conventional nitration.

i) Pyrene

Only one product , with very similar polarity to the parent. No relation to conventional nitration or oxidation.

C'n			Area % at time				
	3.8	5.5	5.8	6.4	6.7	9.0	10.0
A	0.0	0.0	0.0	0.0	0.0	0.0	100.0
B	0.0	0.0	0.0	0.0	0.0	0.0	100.0
C	0.0	0.0	0.0	53.9	8.8	11.6	12.6
D	2.3	3.1	63.9	11.6	2.8	10.3	0.0
E	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Nitration Products of Anthracene
Table 4.8

C'n				Area % at time							
	3.1	3.6	4.0	4.4	5.4	6.2	6.6	7.3	9.2	9.4	11.2
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
B	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
C	0.0	0.0	0.0	0.0	0.0	53.9	8.8	6.6	11.6	12.6	0.0
D	8.0	6.5	5.3	9.5	29.0	13.9	0.0	0.0	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0

Nitration Products of Benzo(a)pyrene
Table 4.9

C'n				Area % at time						
	3.7	4.2	5.4	6.3	6.9	7.3	7.7	7.9	10.2	11.0
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
B	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
D	20.8	5.7	4.0	10.1	11.4	8.6	12.8	4.2	4.3	0.0
E	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0

Nitration Products of Benzo(e)acephenanthrylene
Table 4.10

C'n			Area % at time							
	3.1	5.6	5.9	6.5	7.0	7.2	9.5	9.9	11.1	14.4
A	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
B	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C	0.0	18.5	0.0	17.0	0.0	0.0	5.5	11.9	25.3	8.6
D	4.5	0.0	0.0	0.0	0.0	87.0	0.0	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Nitration Products of Carbazole
Table 4.11

C'n			Area % at time							
	5.0	6.3	6.6	6.7	7.2	7.3	8.4	8.6	9.6	10.2
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
B	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.2	86.0
D	6.1	4.2	9.9	35.5	16.0	4.4	0.0	0.0	0.0	5.9
E	0.0	0.0	0.0	0.0	0.0	0.0	91.5	5.3	0.0	0.0

Nitration Products of Chrysene
Table 4.12

C'n			Area % at time						
	5.4	5.6	5.8	6.0	6.5	8.1	8.3	8.4	8.5
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
B	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
C	0.0	0.0	0.0	0.0	0.0	4.1	2.9	3.3	86.8
D	36.5	9.6	7.9	0.0	23.1	0.0	0.0	0.0	0.0
E	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0

Nitration Products of 4H-cyclopenta(def)phenanthrene
Table 4.13

C'n			Area % at time				
	2.7	4.7	5.4	5.6	6.0	6.8	12.5
A	0.0	0.0	0.0	100.0	0.0	0.0	0.0
B	0.0	0.0	0.0	100.0	0.0	0.0	5.0
C	0.0	0.0	0.0	90.3	6.6	0.0	0.0
D	5.6	54.0	31.1	6.2	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Nitration Products of Naphthalene
Table 4.14

C'n			Area % at time							
	3.6	4.5	4.9	6.7	7.0	7.2	7.3	8.9	9.3	9.5
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
B	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	29.5	8.1	58.0
D	18.4	10.9	7.0	30.2	5.9	5.5	6.0	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Nitration Products of Phenanthrene
Table 4.15

C'n			Area % at time						
	2.8	3.7	4.9	5.6	5.8	7.3	10.5	11.2	11.4
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
B	0.0	0.0	0.0	0.0	0.0	0.0	5.0	95.0	0.0
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.0	0.0
D	6.1	17.9	5.6	7.7	17.9	12.2	0.0	18.5	61.4
E	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0

Nitration Products of Pyrene
Table 4.16

4.4 Summary and Discussion

A rapid extraction and clean-up procedure was devised which utilized chemical reduction and separation on a solid-phase extraction column. This would remove all but the most polar PAH's from environmental samples. The method was shown to give a good linear response and to have good precision. Using model compounds, a limit of determination of less than 1ppb (1μg/Kg) could be obtained.

Unfortunately, when applied to "real" environmental soil samples, it was found that there were many other polar compounds eluting early in the chromatogram. The comparison of chromatograms before and after reduction should have shown the presence of any nitro-compounds as they would have disappeared from the "reduced" chromatogram and would have been replaced by an early eluting peak.

Due to the large number of polar, interfering compounds it was found to be very difficult to judge if any real differences were being seen. Certainly the remote farm-house site did not seem to contain any nitro-PAH's, which was not really surprising.

The roadside sample, which was affected by several possible sources of nitro-PAH's, did suggest the presence of several of these compounds at low levels. The lack of reference materials made their identification impossible.

The third sample, containing very large quantities of PAH's, did not appear to contain any of their nitrated products. Section three of this work suggested that the high

level of PAH's in the area of the flood-plain of the Afon Llwyd was due to contamination from the river. This being the case, the PAH's had not been subject to atmospheric nitration by NO_x. The final experiment demonstrated that nitration by extremely dilute nitric acid was not a feasible option, so the result found was not totally unexpected.

These results are not totally conclusive due to the large number of interfering components found at the ppb level. The speed of the extraction and clean-up made this approach very useful. A change in the final analytical step might still make the extraction procedure viable for future work. One possibility would be HPLC with fluorescence detection. This would provide sufficient sensitivity and probably selectivity. The problem is that insufficient reference standards are available, thus it would be very difficult to determine optimum excitation and emission wavelengths. The production of calibration curves would also be very difficult, if not impossible. The parent nitro-PAH's do not fluoresce but the amino-PAH's do. If the nitro-PAH standards were to become available then the approach used here for validation could be used.

A better approach, however, would seem to be the use of capillary GC with either nitrogen specific or ion-trap detection. This has been used successfully in the analysis of diesel particulates and should still give reasonable results with the much lower levels of nitro-PAH's likely to be found in soil samples.

5. Summary and Conclusions

The original aim of this project was an investigation into the treatability of coloured waters as produced in peaty areas. The colour was thought to be due to the presence of humic and fulvic acids and the decolorisation by treatment plants was not always successful.

Humic acid is a generic term for the complex, amorphous compound formed by degradation of vegetable matter. Its structure can vary considerably with location, time of year or even the physical condition in flowing water. It seemed possible that some characteristic of the structure could cause the failure of the treatment process. Due to the amorphous nature of the compound direct analysis by spectroscopic techniques was not thought to be a feasible approach.

The planned approach, then, was to extract the humic substances from the water and to characterize them numerically in order to produce a data matrix for multi-variate statistical analysis. The characterization was to be by controlled degradation which should have produced a series of acids which could be separated and quantified by High Performance Liquid Chromatography. There was to be no attempt to determine the absolute structure of the humic acids but the differing decomposition products should have varied with change in the overall structure.

The extraction procedure employed was a fairly typical literature method using an ODS solid phase extraction column followed by elution with dilute sodium hydroxide. Investigation of this system demonstrated that some change was occurring to

the humic compounds which resulted in a change in its UV spectrum. Not withstanding this problem, which should have been surmountable, investigations began into both controlled oxidation and hydrolysis of a "standard" humic acid, again using literature procedures. Both processes gave the mixture of acids as expected but, no matter how tightly controlled, the reaction was not reproducible. A great deal of effort went into the investigation of these reactions and attempts to make them reproducible enough for statistical work. This was found to be impossible.

As the above conclusion was being reached, independent workers at the Water Research Centre confirmed the findings of decomposition of humic acid during extraction. Fairly soon after this they also reported that the original problem, that of the decolorisation of peaty water, was not a function of the water but rather of the treatment process itself.

At this point a change of direction was obviously necessary. It was decided that the statistical techniques learned as part of the original project could be used to investigate pollution in a small area of Gwent. Polyaromatic hydrocarbons are a widespread pollutant across the country and it was decided to investigate their presence in the locality described.

An analytical procedure was developed which utilized solid phase extraction and clean-up followed by HPLC analysis using a diode array detector. This system was chosen so as to optimize detection wavelength for sensitivity. It had the added advantage that peaks found could be identified by their spectra

as well as by retention time. A gradient elution process was required which can give rise to problems due to "ghost" peaks. This was overcome by running an injection blank during each series of determinations and electronically subtracting its chromatogram from that produced during the sample run. The system was validated for linearity of response and estimates were made of limits of determination.

A series of soil samples was analyzed from a range of sites and the results were submitted to a selection of chemometric techniques. Factor analysis of the entire data set indicated that two components were responsible for the variance of the data that was not explained by analytical error. A plot of factor scores suggested two clusters which was confirmed by cluster analysis. The statistics showed clearly which sites were associated with each cluster and further investigation indicated that a river, the Afon Llwyd, was responsible for the considerable PAH contamination on its flood plain.

Statistical work on subsets of the data demonstrated that, on ground unaffected by the river, only one component was present. This was thought to be due to general background levels due to vehicle emissions, natural fires etc. These are well mixed in the atmosphere before deposition and adsorption to soil particles.

The analysis of only river affected sites showed that there was possibly a third factor present. Although this could have been due to analytical error it seemed to be slightly too high. Interestingly, the river sites were also in the vicinity of a large, toxic-waste disposal company.

The use of the these powerful statistical techniques demonstrated that their application must be chosen with care and with great thought as to the chemistry of the process under investigation. Soil sampling with the vague idea of identifying industrial contamination must be treated with caution. Pre-mixing of the pollutants before deposition could lead to a process which is controlled more by the soil than by the pollutant. Air sampling on a time basis may give more information of this kind. That this investigation produced useful information was due to the sources being physically different i.e. atmospheric and aqueous deposition.

Further work in this area could include more sampling downstream; the levels of PAH found were extremely high and in the Croesyceiliog area the flood plain is used for growing vegetables. Possible uptake by plants could be checked.

Another approach could be the use of target factor analysis. This may be of limited use on soil samples due to decomposition of the analytes and the variation in their partition coefficients on the soil particles. However, its application to air samples could be of interest. Its main advantage is that potential pollution vectors can be tested as possible targets. This means that sources from industrial sites or vehicle exhausts, for example, can be tested as potential components.

The final section looked at possible contamination by nitro-PAH's. A very sensitive analytical procedure was developed involving derivatization of the nitro compounds, to enable their separation from the parent entity by solid phase

extraction, followed by HPLC analysis. The system was validated by reference to model compounds. When applied to "real" samples, however, it was found that interferences from co-extracted organic components virtually swamped the response from the derivatized nitro compounds.

The section went on to investigate possible production routes of the nitro-PAH's and concluded that the previously reported atmospheric production was far more likely than production in aqueous medium.

There is potential for further work in this area but HPLC is not the method of choice for the final analytical step. Capillary GC seems to present a better option, particularly if combined with a more specific detector. There still remains a major problem in that PAH's can be nitrated to give a mixture of products. A second problem is that there is a lack of reference compounds in this area. Possibly a team approach would be of interest with organic chemists investigating reactions and producing pure standards and analysts measuring their presence in the environment. Again, high volume air samplers might be a better option than soil analysis.

REFERENCES

1. B.R.Kowalski and C.F.Bender JACS 94:16 (1972) 5632
2. D.L.Massart, B.G.M.Vandeginste, S.N.Deming, Y.Michotte and L.Kaufman "Chemometrics: A textbook" Elsevier 1988 ISBN 0-444-42660-4
3. J.C.Berridge J.Chromatogr. 244 (1982) 1
4. D.L.Massart, L.Kaufman and K.H.Esbensen Anal.Chem. 54 (1982) 911
5. P.K.Hopke Annal.New York Acad.Sci. 103 (1980)
6. P.K.Hopke, D.F.Ruppert, P.R.Clute, W.T.Metzger, D.F.Crowley J.Radional.Chem. 29 (1976) 39
7. O.M.Kvalheim, K.Oygard, O.Grahl-Nielsen Anal.Chim.Acta 150 (1983) 145
8. E.Jellum, I.Bjornson, R.Nesbakken, E.Johansson, S.Wold J.Chromatogr. 217 (1981) 231
9. R.Mecke and K.Noack Chem.Ber. 93 (1960) 210
10. D.L.Deuwer, B.D.Kowalski and T.F.Schatzki Anal.Chem. 47 (1975) 1573
11. M.E.Cohen, D.L.Hudson, L.T.Mann, J.VandenBogaerde and N.Gitlin J.Chromatogr. 384 (1987) 145
12. J.H.Brill, T.Mar, H.T.Mayfield and W.Bertsch Analyst 111 (1986) 711
13. D.J.Alpert and P.K.Hopke Proc.Conf.Quality Assurance Environ.Meas. Denver Colorado (1978) 204
14. P.K.Hopke, J.Environ.Sci.Health A11 6 (1976) 376

15. J.Smeyers-Verbeke, J.C.Den Hartog, W.H.Dekker, D.Coomans, L.Baydens and D.L.Massart Atmos.Environ. 18 (1984) 2741

16. F.I.Onuska, A.Mudroch, S.Davies HRC CC J.High Resolut.Chromatogr. Chromatogr.Commun. 8.11 (Nov 1985) 747

17. K.E.Thrane and R.W.Gunderson Anal.Chim.Acta. 191 (1986) 309

18. P.D.Gaarenstroom, S.D.Perone, J.L.Moyers Environ.Sci.Technol. 11 (1977) 795

19. L.J.C.Love, P.C.Tway, L.M.Upton Anal.Chem. 52 (1980) 311

20. H.B.Woodruff, P.C.Tway, L.J.C.Love Anal.Chem. 53 (1981) 81

21. J.N.BrØnsted, K.J.Pederson Z.Physikal.Chem. 108 (1924) 185

22. S.Wold Pattern Recognition 8 (1976) 127

23. K.J.Connor, The Characterisation and Chemical Treatment of Soluble Colour in Water. Water Research Centre Processes, December 1986

24. M.Maybeck in "Flux of Organic Carbon by Rivers to the Ocean" report of a workshop, Woods Hole, Mass, 21-25 Sept 1980 (Dept.of Energy, Washington DC 1981) p.219

25. Steinberg and Muenster, Chap 5 of Humic Substances in Soil, Sediment and Water. Geochemistry, Isolation and Characterisation by Aiken, McKnight, Wershaw and MacCarthy. Published by Wiley Interscience.

26. K.Oden Kolloid Chem.Beihefte 11 (1919) 75

27. Chap.1 of ref. 25

28. D.W.Anderson J.Soil.Sci. 30 (1979) 77

29. E.T.Gjessing Phys. and Chem. Chararcteristics of Aquatic Humus. Ann Arbor Science Publishers Inc, 1976

30. G.C.Choudry Humic Substances, structural, a l , photophysical, photochemical and free radical aspects and interactions with environmental chemicals. Gordon and Brown Science Publishers

31. A.P.Black, R.F.Christman J.Amer.Water Works Ass. 55(7) (1963) 897

32. J.Aho and O.Lehto Archive fur Hydrobiol 101 (1984) 21

33. P.T.Hine and D.B.Bursill Water Research 18(11) (1984) 1461

34. S.A.Vissar Environ.Sci.Technol. 17(7) (1983) 412

35. C.J.Miles, J.T.Tuschall Jr., P.L.Brezonik Anal.Chem 55 (1983) 410

36. R.Blondeau and E.Kalinowski J.Chromatogr. 351 (1986) 585

37. A.L.Wilson J.Appl.Chem.,9 (1959) 510

38. D.F.Marino and J.D.Ingle Jr., Anal.Chim.Acta, 124 (1981) 23

39. A.Cominoli, J.Bufflek, W.Haerdi J.Electroanal.Chem., 110 (1980) 259

40. T.Almgren, B.Josefsson and G.Nyquist Anal.Chim.Acta. 78 (1975) 411

41. D.F.Martin and R.H.Pierce Jr., Environ.Lett. 1 (1971) 49

42. E.R.Sholkovitz, E.A.Boyle and N.B.Price, Earth Planet.Sci.Lett., 40 (1978) 130

43. P.D.Carpenter and J.D.Smith Anal.Chim.Acta. 159 (1984) 299

44. R.V.Cheshire, P.A.Cranwell, C.P.Falshaw, S.J.Floyd and R.D.Haworth, Tetrahedron 23 (1907) 1669

45. E.H.Hansen, M.Schnitzer, Soil.Sci.Soc.Amer.Proc., 33 (1969) 29

46. D.H.Stuermer, G.R.Harvey, Marine Chem. 6 (1978) 55

47. W.Liao, R.F.Christman, J.D.Johnson, D.S.Millington and J.R.Haas, Environ.Sci.Technol. 16 (1982) 403

48. P.L.Debene and M Delamar Anal.Letters 19(21 & 22) (1986) 2131

49. M.Schnitzer and S.I.M.Skinner Can.J.Chem. 22 (1974) 1072

50. P.Ruggiero, F.S.Interesse, L.Cassidei and O.Scracovelli, Geochimica and Cosmochimica Acta 44 (1980) 603

51. S.A.Visser J.Environ.Sci.Health A17(6) (1982) 767

52. J.A.Neyroud and M.Schnitzer Can.J.Chem. 52 (1975) 4123

53. M.P.Jackson, R.S.Swift, A.M.Posner and J.R.Knox Soil.Science 114(1) (1972) 75

54. H.L.C.Meuzelaar, K.Haider, B.R.Nagar and J.P.Martin Geoderma 17 (1977) 239

55. H.de Haan, T.deBoer and G.Halma Freshwater Biol. 9 (1979) 315

56. J.L.Clajch and L.R.Snyder Computer Assisted Method Development for High Performance Liquid Chromatography Elsevier, New York 1990 ISBN 0-444-88748-2

57. L.R.Snyder J.Chromatogr.Sci. 16 (1983) 223

58. R.J.Laub and J.H.Purnell J.Chromatogr. 112 (1975) 71

59. J.C.Berridge J.Chromatogr. 485 (1989) 3

60. W.Spendley, G.D.Hext and F.R.Hinsworth Technometrics 4 (1962) 441

61. J.A.Nelder and R.Mead Comput.J. 7 (1965) 308

62. A.Waggott, Water Research Centre, Stevenage, Herts Personal communication 1988

63. J.C.Berridge Techniques for Automated Optimization of HPLC separations.
John Wiley and Sons, London 1985

64. K.J.Connor Water Research Centre Processes "The Characterization and Chemical Treatment of Soluble Colour in Water, Part 1" December 1986

65. K.J.Connor and M.J.Stiff Water Research Centre Processes "The Characterization and Chemical Treatment of Soluble Colour in Water, Part 3" May 1987

66. M.Hellmann Fresnius.Z.Anal.Chem. 295 (1979) 24

67. Casarett and Doull's Toxicology, The Basic Science of Poisons 2nd Ed. Edited by J.Doull, C.D.Klaassen and M.O.Amdar. MacMillan Publishing Company, New York 1980. ISBN 0-02-330040-X

68. S.M.Barlow and F.M.Sullivan, Reproductive Hazards of Industrial Chemicals. Academic Press, London 1982 ISBN 0-12-078960-4

69. P.Sims Biochemical Pharmacology 19 (1970) 285

70. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Man: Certain Polycyclic Aromatic Hydrocarbons and Heterocyclic Compounds. (1973) Volume 3 Lyon, France

71. M.Blumer and W.W.Youngblood Science 53 (1975) 188

72. P.A.Ilnitsky, G.A.Belitsky and L.M.Shabad Cancer Lett (Amsterdam) 1 (1976) 291

73. P.C.Siebert, C.A.Craig and E.B.Coffey Final Report EPA Control No. 68-02-2836, Environmental Protection Agency (1978)

74. U.S. Environmental Protection Agency, Polynuclear Aromatic Hydrocarbons: Ambient Water Quality Criteria, Washington DC (1980)

75. R.J.F.Bewley Chem.Ind. (London) (1989) 778

76. P.Wedgewood and R.L.Cooper Analyst 78 (1953) 170

77. A.Liberti, G.P.Cartoni and V.Cantat J.Chromatogr. 15 (1964) 141

78. H.Y.Tong. J.D.Centen, F.W.Karasek, E.Jellum, P.Helland J.Chromatogr. 324 (1985) 373

79. B.Santoni and C.Mandon Analusis 9(6) (1981) 259

80. A.Mellone, B.W.Smith and J.D.W.Winefordner Talanta 37(1) (1990) 111

81. W.Funk, G.Donnevert, B.Schuch, V.Gluck and J.Becker J.Planar Chromatogr.-Mod.TLC 2(4) (1989) 317

82. W.Baumeister, B.Zens, O.Vierle Fresenius Z.Anal.Chem. 333(7) (1989) 710

83. H.G.Kicinski, S.Adamek and A.Kettrup Chromatographia 28(3/4) (1989) 203

84. P.Fernandez, C.Porte, D.Barcelo, J.M.Bayona and J.Albarges J.Chromatogr. 456(1) (1988) 155

85. K.Jinno and S.Niimi J.Chromatogr. 455 (1988) 29

86. A.L.Blilie and T Gremibrokk Anal.Chem. 57 (1985) 2239

87. B.Muel and S.Saguem Int.J.Environ.Anal.Chem. 19(2) (1985) 111

88. A.Zlatkis ed. Advances in Chromatography,1969 Preston Technical Abstracts Co. 1969

89. I.Halasz, R.Endele, and J.Asshauer J.Chromatogr.112 (1975) 37

90. H.Hemetsberger, W.Maasfeld and H.Ricken Chromatographia 9 (1976) 303

91. K.Karch, I.Sebastian and I.Halasz J.Chromatogr. 122 (1976) 3

92. J.J.Kirkland, Chromatographia 8 (1975) 661

93. Baker-10 SPE Applications Guide Vol.2, Published by J.T.Baker Chemical Company 1984

94. IUPAC Spectrochim.Acta 33B (1978) 242

95. SPSS UK Ltd., SPSS House, 5 London St., Chertsey, Surrey

96. J.J.Kankare Anal.Chem. 42 (1970) 1322

97. M.S.Bartlett Br.J.Psychol.Stat. Sect 3 (1950) 77

98. Z.Z.Hugus Jr., A.A El-Awady J.Phys.Chem. 75 (1971) 2954

99. D.L.Deuwar, B.R.Kowalski and J.L.Fasching Anal.Chem. 48 (1976) 2002

100. E.R.Malinowski Anal.Chem. 49 (1977) 612

101. R.B.Cattell Multivariate Behav.Res. 1 (1966) 245

102. HF.Kaiser Educ.Psych.Meas. 20 (1960) 141

103. H.H.Harman Modern Factor Analysis 3rd ed. rev., University of Chicago Press, Chicago 1967

104. H.F.Kaiser Psychometrika 23 (1958) 187

105. D.R.Saunders Psychometrika 25 (1960) 199

106. A.Trenholm and C.C.Lee Analysis of PIC and tota mass emissions from an incinerator In: Land Disposal, Remedial Action, Incineration and Treatment of Hazardous Waste, Proceedings of the Thirteenth Annual Research Symposium, USEPA, Hazardous Waste Engineering Laboratory, Cincinnati OH EPA 600/9-86/022 August 1986

107. J.Boegel Assessment of Residues from Incineration In: as 106 but EPA/600/9-87/015 July 1987

108. N.T.Edwards J.Environ.Qual. 12 427 (1983)

109. J.N.Pitts Jr., K.A.VanCauwenberghe, D.Grosjean, J.P.Schmid, D.R.Fitz, W.L.Belser Jr., G.B.Knudson, P.M.Hynds Science 202 515 (1978)

110. D.L.Chapman, D.C.Heckert, J.W.Reasoner, S.P.Thackaberry J.Am.Chem.Soc. 88 5550 (1966)

111. R.C.Pearce, M.Katz Environ.Sci.Technol. 10 45 (1976)

112. D.Scheutzle, F.S-C.Lee, T.Prater, S.Tejade in "Proceedings of the 10th Annual Symposium on the Analytical Chemistry of Pollutants" Gordon and Breach Science Publishers, New York 1980 pp193

113. J.Jager J.Chromatogr. 152 575 (1978)

114. W.L.Fitch, E.T.Everhart, D.H.Smith Anal.Chem. 50 2122 (1978)

115. H.S.Rosenkranz, R.Mermelstein Mutal.Res. 114 217 (1983)

116. E.T.Wei, H.P.Shu Am.J.Public Health 73 1085 (1983)

117. H.Ohgaki, C.Negishi, K.Wakabayashi, K.Kusama, S.Sata,T.Sugimura Carcinogenesis 5 583 (1984)

118. P.C.Howard, J.A.Gerrard, G.E.Milo,P.P.Fu, F.A.Beland, F.F.Kadluber Carcinogenesis 4 353 (1983)

119. T.Kinouchi, Y.Manabe, K.Wakisaka, Y.Ohniski Microbiol. Immunol.26 993 (1982)

120. Committee on Biologic Effects of Atmospheric Pollutants, Particulate Polycyclic Organic Matter. National Acadamy of Sciences, Washington DC 1972

121. G.Grimmer, Int.Agency Res.Cancer Sci.Publ. 16 29 (1977)

122. J.N.Pitts Jr., Phil.Trans.R.Soc.London 290 551 (1979)

123. L.A.Poirier, E.K.Weisburger J.Nat.Cancer Inst. 62 833 (1979)

124. H.C.Paputa-Peck, R.S.Marano, D.Scheutzle, T.L.Riley, C.V.Hampton, T.J.Prater, L.M.Skewes, T.E.Jensen, P.H.Ruehle, L.C.Bosch, W.P.Duncan Anal.Chem 55 1946 (1983)

125. A Stratweiser Jr., R.C.Fahey J.Org.Chem 27 2352 (1962)

126. W.M.Draper Chemosphere 15(4) 437 (1986)

127. I.O.O.Korhonen, M.A.Lind, J.Chromatogr. 332 71 (1985)

128. C.M.White, A.Robbat Jr., R.M.Hoes Chromatographia 17(11) 605 (1983)

129. T.Ramdahl, K.Urdal Anal.Chem. 54 2256 (1982)

130. A.Robbat Jr., N.P.Corso, P.J.Doherty, D.Marshall Anal.Chem. 58 2072 (1986)

131. A.Robbat Jr., N.P.Corso, P.J.Doherty, M.H.Wolf Anal.Chem. 58 2078 (1986)

132. B.A.Tomkins, R.S.Brazell, M.E.Roth, V.H.Ostrum Anal.Chem. 56 781 (1984)

133. J.Jager J.Chromatogr. 152 575 (1978)

134. M.W.Chou, J.High.Res.Chrom. and Chrom.Commun. 8 304 (1985)

135. T.Nielsen Anal.Chem. 55 286 (1983)

136. W.A.MacCrehan, W.E.May, S.D.Yang, B.A.Benner Jr. Anal.Chem. 60 194 (1988)

137. W.C.Yu, D.H.Fine, K.S.Chiu, K.Biemann Anal.Chem. 56 1158 (1984)

138. D.L.Newton, M.D.Erickson, K.B.Tomer, E.D.Pellizzari, P.Gentry, R.B.Zweidinger Environ.Sci.Technol. 16 206 (1982)

139. M.Oehme, S.Mano, H.Stray HRC and CC 5 417 (1982)

140. P.A.D'Agostino, D.R.Narine, B.E.McCarry, M.A.Quilliam in M.Cooke and A.J.Dennis (eds) "Polynuclear Aromatic Hydrocarbons: Formation and Metabolism " Battelle Press, Columbus OH (1983)

141. R.M.Cambell, M.L.Lee Anal.Chem. 56 1026 (1984)

142. E.R.Malinowski and D.G.Howery "Factor Analysis in Chemistry" Wiley Interscience, New York 1980 ISBN 0 471 05881-5

143. T.J.Hoban Ph.D. Thesis, Polytechnic of Wales, 1985

144. R.Pardo, E.Barrado, Y.Castrillejo J.Chemometrics 3 315 (1988)

145. J.Einax, K.Oswald, K.Danzer Fresenius J.Anal.Chem. 336 394 (1990)

146. E.R.Malinowski, D.G.Howery, P.H.Weiner, J.R.Soroka,, P.T.Funke, R.S.Selzer, A.Levinstone, "FACTANAL", Program 320, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Ind. 1976