5.4 Task 3: Profiling of Major Components by GC-FID and GC-MS

Prior to chemical analysis, the illicit heroin of each individual case was homogenized and ground into a fine powder. Approximately 80 – 100 mg of the sample was subjected to qualitative and quantitative determination of the eight target compounds by gas chromatography-mass spectrometry (GC-MS) and gas chromatography-flame ionization detection (GC-FID) respectively. This task is divided into six subtasks: 1) GC-FID optimization and method validation, 2) evaluation of the robustness of the GC-FID method by statistical analysis, 3) statistical validation of GC-FID data using 8 simulated links, 4) GC-MS method validation, 5) analysis and statistical classification of the case samples for sample-to-sample comparison at the source level using opium-based alkaloids, and 6) development of a novel statistical approach for sample classification.

5.4.1 GC-FID Method Optimization and Validation

GC-FID was chosen for quantitative analysis because it showed better sensitivity over GC-MS in this study. However, the latter is indispensable as it is also vital to confirm the target compounds present in the street samples. Both methods were optimized and validated accordingly to meet the profiling requirements of the Malaysian enforcement laboratory.

In this subtask, the eight major components (three adulterants: paracetamol (PC), caffeine (CF) and dextromethorphan (DM) and five opium-based alkaloids: codeine (CD), morphine (MP), acetylcodeine (AC), 6-monoacetylmorphine (MM) and heroin (HR)/diamorphine) were chosen mainly because they were often detectable in the heroin samples seized in Malaysia. The opium-based alkaloids are useful to estimate the source (as well as the production batch) of the samples since the natural alkaloids such as codeine, morphine and its simple derivatives are directly related to its source. In

contrast, the adulterants not only serve to characterize the local heroin seizures, they are also characteristic of the countries that dilute/cut the samples. As the samples were highly cut, papaverine and noscapine were normally absent and if present, were treated as minor rather than major components. The choice of 2,2,2 triphenyl acetophenone as the internal standard (IS) was made based on the previous work reported by the Australian Government National Measurement Institute (2007).

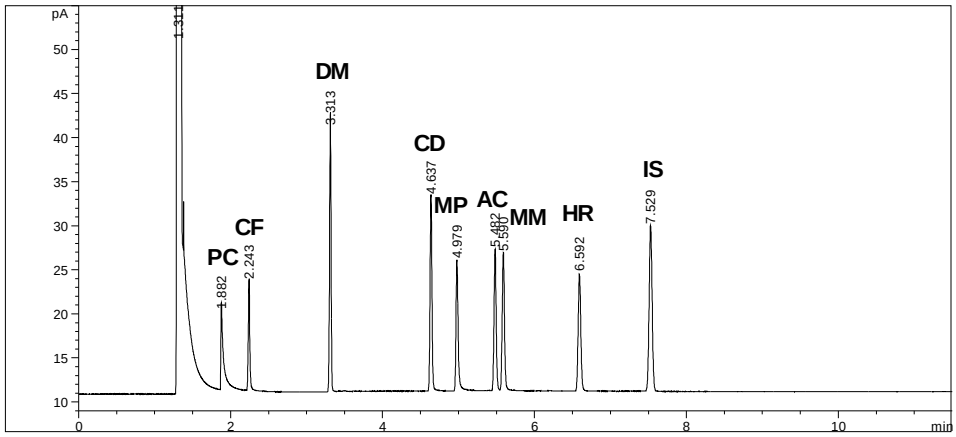
5.4.1.1 Choice of GC Capillary Column and Selectivity

A slightly polar stationary phase is widely accepted for general use and in drug analysis, the use of a 5% polar column is routine and therefore it was chosen in this study. The GC conditions were first optimized using two types of slightly polar GC capillary columns, namely J&W HP Ultra 2 (25 m x 200 µm x 0.33 µm) and J&W HP-5 (30 m x 250 µm x 0.25 µm). Based on the best separation and peak height achieved by both the columns, the parameters as shown in Table 5.7 were finally obtained.

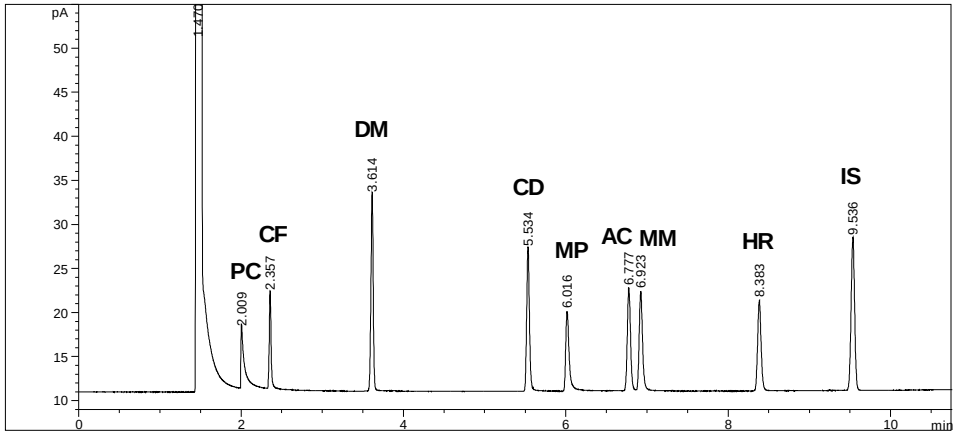
Table 5.7: GC-FID parameters for quantitative determination of eight target compounds

Condition	Option 1	Option 2
Column:	J&W HP Ultra 2 (5% phenyl 95% methyl siloxane)	J&W HP-5 (5% phenyl 95% methyl siloxane)
Dimensions:	Length: 25 m i.d.: 200 µm Film thickness: 0.33 µm	Length: 30 m i.d.: 250 µm Film thickness: 0.25 µm
Carrier gas:	Helium	Helium
Pressure:	239.1 kPa	134.7 kPa
Total flow:	44.2 mL/min	38.7 mL/min
Injection volume:	1 µL	1 µL
Split ratio:	40 : 1	40.5 : 1
Flow rate:	1.0 mL/min	0.9 mL/min
Injector temp.:	290 ^{o}C	280 ^{o}C
Temp. programming:	240 ^{o}C hold 1 min, ramp at 12 ^{o}C/min to 270 ^{o}C and hold 8 min.	250 ^{o}C hold 4 min, ramp at 7^{o}C/min to 260 ^{o}C and hold 2 min, ramp at 6^{o}C/min to 280^{o}C.
Detector temp.:	290^{o}C	280^{o}C
H_2 flow:	30 mL/min	30 mL/min
Air flow:	300 mL/min	300 mL/min
He makeup flow:	25 mL/min	25 mL/min
Total run time:	< 12 min	< 11 min

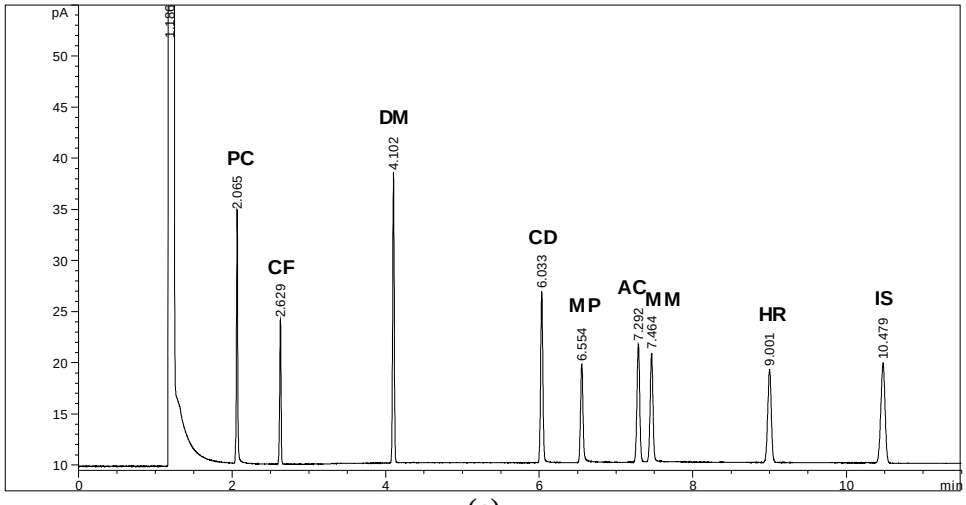
An initial investigation performed on these two GC capillary columns demonstrated different resolutions. The components were first separated using the parameters in Option 1 and the HP-5 column. All the major peaks were eluted quickly; however separation between acetylcodeine and 6-monoacetylmorphine was poor with tailing observed in the paracetamol peak (Figure 5.22(a)). Using the conditions in Option 2, the chromatogram in Figure 5.22(b) illustrates that all the compounds are well resolved but peak tailing is still evident in the paracetamol peak. Further optimization of the split ratio, flow rate and temperature did not resolve this problem. This suggested sample overloading and a thicker film was required in the column.



(a)



(b)



(c)

Figure 5.22: Chromatographic selectivity shown by two GC columns (Different outcomes are obtained with (a) HP-5 using parameters in Option 1; (b) HP-5 in Option 2 and (c) HP Ultra 2 in Option 1; separation is improved between PC and CF in (c); all components were at ~0.1 mg/mL; PC = paracetamol; CF = caffeine; DM = dextromethorphan; CD = codeine; MP = morphine; AC = acetylcodeine; MM = 6-monoacetylmorphine, HR = heroin; IS = internal standard, 2,2,2-triphenyl acetophenone)

Changing the column to the HP Ultra 2 with Option 1 resulted in good selectivity and better peak resolution without significant peak tailing even though a higher ramping rate was used and this is illustrated in Figure 5.22(c). In addition, the retention time for the first three compounds was longer, thus promoting better separation. A slightly higher injector temperature (290°C) was also used as the thicker film within the column was able to retain a larger amount of paracetamol and dextromethorphan, improving the sensitivity. However, 3-monoacetylmorphine co-eluted with 6-monoacetylmorphine, the quantitative results obtained with the calibration curve of 6-monoacetylmorphine are thus used to represent both the components in this study.

For quantitative determination of the eight major target components in the illicit heroin, Option 1 with the HP Ultra 2 column was selected and was used for the subsequent validation.

5.4.1.2 Solvent Studies

Methanol and chloroform are two commonly used solvents in drug analysis. The combination of these two solvents provides an ideal medium for both polar and non-polar compounds present in the sample matrix to be extracted. Hence, both of these solvents were considered. With 100% chloroform as the dissolving solvent, dissolution of the salts of codeine, morphine and 6-monoacetylmorphine was often a problem and at higher concentrations of the drug, a colloidal solution or suspensions were observed in this solvent. To compare the efficiency of methanol, chloroform and their combinations, each component was prepared at approximately 0.1 mg/mL which resulted in a clear solution for each solvent type. Two separate sets of analyses with 6 replicate injections of each mixture were performed on different days using Option 1. Table 5.8 summarizes the RSDs of all the compounds expressed as peak area ratios

(relative to the IS) obtained from the two sets of analyses. All combinations demonstrated good performance (RSD < 2%) apart from the 100% chloroform in which inconsistent readings for morphine were significant, which could be due to its relatively high polarity (Medina, 1989) and very low solubility of the compound in the solvent. Since the presence of methanol can cause transacetylation of paracetamol and morphine as well as facilitating the decomposition of heroin/diamorphine (Huizer & Poortman, 1989; Zelkowics *et al.*, 2005), it was decided to use 1:9 methanol:chloroform as the dissolving solvent for this study. In addition, the minimal use of methanol was required to facilitate the dissolution of more polar compounds.

Table 5.8: RSD (%) of area ratios (peak relative to IS) for eight target compounds, each at approximately 0.1 mg/mL in different solvent combinations (n = 6)

Solvent[a]	PC	CF	DM	CD	MP	AC	MM	HR	Mean[b]	Precision[c]
M	[d]0.37	0.50	0.79	0.41	0.90	0.34	0.52	0.20	0.51	0.23
(100%)	[e](0.23)	(0.29)	(0.30)	(0.11)	(0.59)	(0.13)	(0.27)	(0.16)	(0.26)	(0.15)
M:C	0.18	0.33	0.18	0.16	0.41	0.07	0.29	0.14	0.22	0.11
(9:1)	(0.24)	(0.22)	(0.29)	(0.27)	(1.37)	(0.28)	(0.42)	(0.26)	(0.42)	(0.39)
M:C	0.13	0.21	0.21	0.17	0.49	0.14	0.14	0.17	0.21	0.12
(7:3)	(0.27)	(0.21)	(0.13)	(0.22)	(0.73)	(0.18)	(0.33)	(0.22)	(0.29)	(0.19)
M:C	0.23	0.16	0.12	0.23	1.87	0.16	0.49	0.28	0.44	0.59
(5:5)	(0.25)	(0.31)	(0.16)	(0.16)	(0.46)	(0.14)	(0.22)	(0.08)	(0.22)	(0.12)
M:C	0.20	0.35	0.17	0.12	0.70	0.10	0.23	0.17	0.26	0.19
(3:7)	(0.08)	(0.27)	(0.11)	(0.06)	(0.56)	(0.89)	(0.15)	(0.11)	(0.28)	(0.30)
M:C	0.42	0.37	0.22	0.20	0.72	0.15	0.26	0.96	0.41	0.29
(1:9)	(0.41)	(0.36)	(0.19)	(0.21)	(0.48)	(0.37)	(0.21)	(0.26)	(0.31)	(0.11)
C	1.00	0.23	0.23	0.20	8.36	0.15	0.49	0.16	1.35	2.85
(100%)	(0.42)	(0.62)	(0.17)	(0.12)	(6.31)	(0.36)	(0.28)	(0.15)	(1.05)	(2.13)

[a]M = methanol; C = chloroform.
[b]Mean = Average of each row.
[c]Precision = Standard deviation of each row.
[d]Series 1
[e]Series 2

5.4.1.3 Repeatability and Reproducibility

With 1:9 methanol:chloroform, a mixed standard containing the eight target analytes at their respective working concentrations and the IS (0.18 mg/mL) were prepared to evaluate intra-day, inter-hour and inter-day precision using the chromatographic conditions in Option 1. According to UNODC (2009c), the acceptable criteria for the precision should be within 15% for seized materials using GC.

For the intra-day precision (repeatability), the RSD was computed for each compound from ten injections. Table 5.9 shows that the peak areas of all the analytes and the IS achieved excellent repeatability (RSD < 1.5%).

Similarly, the RSDs of the peak areas and peak area ratios (relative to IS) were computed from the repeated injections performed over a 28-hour period (Table 5.9). For the inter-hour precision, large deviations were observed in the peak areas. This can be attributed to solvent evaporation that causes an increase in the analyte per unit volume (the peak area) over time. These errors were corrected when all the target analytes were normalized to the IS. The inter-hour precision based on area ratios achieved an RSD < 3%, indicating sufficient system stability for high throughput analysis. This favorable outcome enables samples to be prepared and left unattended on a GC sample tray for overnight analysis.

The inter-day precision (reproducibility) is a good means of evaluating system stability and other uncontrolled variables over time (e.g. system drift). Table 5.9 displays that all major peak area ratios (relative to the IS) showed an RSD < 5% except for morphine (RSD < 9%). The relatively poor performance of morphine is most likely due to its relatively low solubility in the presence of chloroform. For sample-to-sample comparison, all the major compounds were sufficiently reproducible and stable. However, morphine has to be analyzed separately with a more suitable solvent if a more rigorous analysis is required.

Table 5.9: RSD (%) of peak areas and area ratios (peak relative to IS) for a standard mixture injected on the same day, within 28 hours and on ten different days

		PC	CF	DM	CD	MP	AC	MM	HR	IS
Intra-day precision (n = 10)	**Peak area**	0.92	0.85	0.83	0.82	1.45	0.94	0.80	0.81	1.21
	Area ratio	1.09	1.27	1.26	1.29	1.46	1.28	1.22	1.25	-
Inter-hour precision (n = 10)	**Peak area**	9.75	10.09	9.63	9.81	8.62	10.41	9.23	9.65	9.81
	Area ratio	0.51	0.72	0.81	0.90	2.90	1.00	1.29	0.70	-
Inter-day precision (n = 10)	**Peak area**	3.62	3.17	1.86	2.05	7.18	2.00	1.87	1.75	2.04
	Area ratio	4.98	2.56	2.42	2.74	8.72	1.96	3.08	2.32	-

5.4.1.4 Linearity and Limit of Detection (LOD)

Linearity defines the working range within which the analyte should fall. In the illicit heroin, each analyte has its self-defined range that varies from those of other analytes present in the same sample. Using a mixed standard confers an advantage of mimicking the sample matrix. Eight individual calibration curves (area ratio versus known concentration) were constructed from a series concentrations covering 50 – 150% of the working range and their results are shown in Table 5.10 and Appendix 9. As indicated by the correlation coefficient, $r^2 \geq 0.9997$ and the linearity index within ± 5%, each linear range of interest was sufficiently good for routine quantitative determination using Option 1. In line with the work done by Anastos *et al.* (2005a), the negative y-intercept values obtained in this study may be the suppression effect arising from the mixture. This effect however did not affect the analytical results.

Table 5.10: Linearity of the study range and LOD of the instrument

Major component	Concentration range covered (mg/mL)	Linearity function	Correlation coefficient, r^2	RSD for area ratio[a] (%)	Linearity index[b] (%)	LOD (mg/mL)
Paracetamol	0.04 – 0.80	y = 3.44x – 0.026	0.9998	0.17 – 0.69	0.98 – 1.05	0.0021
Caffeine	0.50 – 10.0	y = 2.11x – 0.121	0.9998	0.05 – 0.45	0.98 – 1.05	0.0006
Dextromethorphan	0.02 – 0.30	y = 5.30x – 0.012	0.9997	0.10 – 0.55	0.98 – 1.05	0.0004
Codeine	0.01 – 0.20	y = 4.27x – 0.005	0.9998	0.13 – 0.54	0.99 – 1.02	0.0005
Morphine	0.01 – 0.20	y = 3.36x – 0.010	0.9997	0.65 – 1.91	0.98 – 1.03	0.0061
Acetylcodeine	0.05 – 0.10	y = 4.37x – 0.022	0.9998	0.10 – 0.56	0.99 – 1.03	0.0008
6-Monoacetylmorphine	0.05 – 1.00	y = 3.95x – 0.032	0.9997	0.24 – 0.74	0.98 – 1.03	0.0019
Heroin	0.05 – 1.00	y = 3.73x – 0.016	0.9999	0.08 – 0.55	0.99 – 1.02	0.0013

[a]The range of RSD values indicates the precision for each of the eight concentration levels calculated from 6 injections. Calculation is based on peak relative to the IS.
[b]Percentage difference between the known standard concentration and calculated concentration from the experimental data using the equation obtained from the least squares fit.
Note: All alkaloids show very good fits within ±5% linearity indices.

The lowest detected concentration was used to determine the limit of detection (LOD) (Appendix 10). The instrument was found to be sufficiently sensitive to detect the analytes with the LODs between 0.4 μg/mL and 6.1 μg/mL. In cases where the concentrations of the major components are too close to the LODs, the weight of the sample has to be increased for analysis.

5.4.1.5 Accuracy as Measured by Recovery

For prosecution purposes, the absolute amount of the drug present in the sample is of major concern. As the GC method of this study was also designed for routine analysis, recovery studies are necessary to evaluate the level of accuracy that the method can offer. The mean recoveries for the eight target analytes determined with Option 1 are displayed in Table 5.11. As the typical accuracy of the recovery for major components is expected to be 98 – 102% (Chan, Lam, Lee & Zhang, 2004) and since all the analytes in this study showed mean recoveries between 99% and 102%, thus the method is acceptable for quantification.

Table 5.11: Recovery (%) for eight target compounds

Major component	Low	Medium	High	Mean	Standard deviation
Paracetamol	99.90	99.45	100.74	100.03	0.66
Caffeine	101.82	98.66	98.30	99.59	1.94
Dextromethorphan	99.60	99.68	99.76	99.68	0.08
Codeine	98.30	99.01	99.77	99.03	0.73
Morphine	99.74	101.70	101.94	101.13	1.21
Acetylcodeine	100.37	99.25	101.34	100.32	1.04
6-Monoacetylmorphine	98.79	100.66	101.82	100.42	1.53
Heroin	100.34	101.37	101.53	101.08	0.65

5.4.1.6 Dissolution Vessels and Sample Weight Test

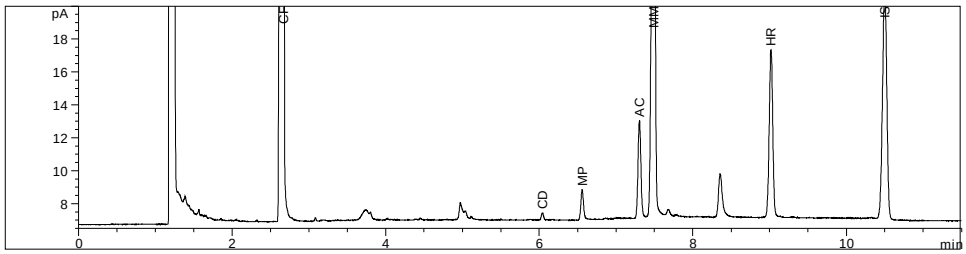
The use of plastic ware in organic analysis is avoided since it leaches the phthalate plasticizers. When the quantitative work is set for profiling purposes, the absolute concentration becomes less critical and hence a disposable plastic vessel may be used for sample dissolution. The suitability of polypropylene centrifuge tubes (PT) and glass volumetric flasks (VF) for sample preparation was examined using three randomly chosen case samples (marked 'A', 'B' and 'C'). The GC reading for each component is expressed as a percentage analyte. At each weight level prepared in the PT and VF, the total percentage analyte representing the sum of the eight components in the vessel was calculated (Table 5.12). The small standard deviation in each case reveals that the weight used did not have much effect on the analysis. However, the PTs showed a lower total percentage analyte content compared to the corresponding VFs (especially in Samples B and C), suggesting a loss of analytes when using the plastic vessels. This preliminary investigation indicates that VF must be employed in order to minimize the loss of analytes.

Table 5.12: Total % analyte content ± standard deviation for three samples prepared in PTs and VFs

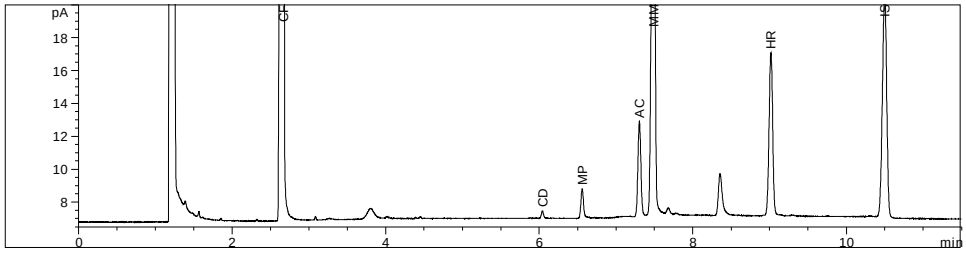
Weight (mg)	Sample A		Sample B		Sample C	
	PT	VF	PT	VF	PT	VF
50	88.4	90.0	83.6	86.1	92.5	94.2
60	89.3	89.3	85.5	87.5	92.2	94.9
70	88.6	88.4	86.5	89.2	91.7	94.0
80	88.6	87.5	81.9	85.2	92.5	93.6
Mean	**88.7**	**88.8**	**84.4**	**87.0**	**92.2**	**94.2**
Standard Deviation	**0.4**	**1.1**	**2.1**	**1.7**	**0.4**	**0.5**

In addition, artifacts were observed in the samples extracted in the PTs and these became significant when the aliquots were analyzed on day 4 after cold storage. These

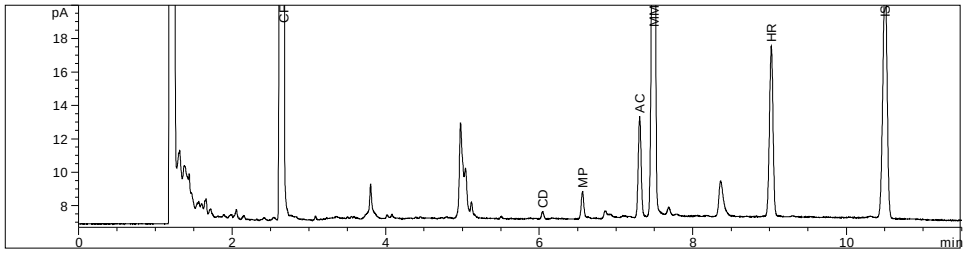
artifacts had a greater impact on the elution times of paracetamol and dextromethorphan. In contrast, the chromatograms of the samples prepared in the VFs (Figures 5.23(b) and 5.23(d)) show relatively constant baselines. It is therefore established that the use of plastic vessels must be avoided in organic analysis.



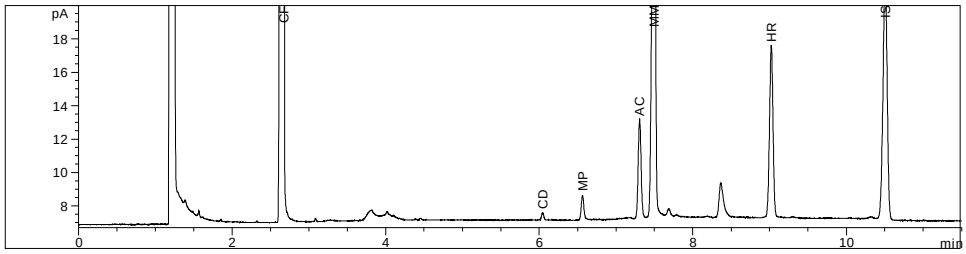
(a)



(b)



(c)



(d)

Figure 5.23: Chromatograms of a sample prepared in PTs and VFs (Chromatograms in (a) and (b) are samples in PT and VF respectively analyzed on the day of sample preparation; chromatograms in (c) and (d) are samples in PT and VF respectively analyzed on day 4. Artifacts are significant in (a) and (c))

From the six analyses at each concentration level, the sample weight range of 50 – 80 mg resulted in an RSD < 3%. Lower weights however may run a risk of having undetectable and unquantifiable analytes (such as codeine and morphine). Therefore, it was decided to use a higher sample weight (> 70 mg) as a starting weight. The weight was adjusted accordingly for components that are present at very low or extremely high concentrations.

5.4.1.7 Sample Stability

A sample containing relatively high analyte contents was diluted to yield a low content sample to assess sample stability. Table 5.13 shows the cumulative RSD of the peak area ratios of the eight major components. Discrepancies in the peak area ratios became greater as the analysis time increased. Both the samples performed very well within the first week (RSD < 4%). The performance gradually degraded over the subsequent weeks. Paracetamol and morphine were more affected as the samples aged, suggesting that samples should be analyzed within a week.

Table 5.13: Cumulative RSD (%) of area ratios of eight major components obtained within a month

Week	Sample	PC	CF	DM	CD	MP	AC	MM	HR
0	**Low**	1.74	0.44	0.11	0.16	0.40	0.15	0.12	0.12
(n = 3)	**High**	0.32	0.27	0.47	0.43	0.33	0.53	0.45	0.65
1	**Low**	1.84	1.48	1.02	1.04	3.28	0.96	1.38	0.73
(n = 6)	**High**	1.22	0.45	0.66	0.79	1.44	0.76	1.24	0.67
2	**Low**	6.05	1.41	1.28	1.66	2.86	1.32	2.52	1.40
(n = 9)	**High**	4.99	1.03	1.61	2.16	3.12	1.78	2.56	1.96
3	**Low**	13.17	3.24	3.53	7.07	17.60	4.36	7.44	3.49
(n = 12)	**High**	12.02	3.43	3.71	5.49	8.85	4.30	5.91	3.71
4	**Low**	14.42	3.85	3.89	6.75	15.65	4.36	7.45	3.63
(n = 15)	**High**	13.82	3.92	4.17	5.79	8.65	4.56	6.37	4.16

5.4.1.8 Capability of the Method for Sample Classification

A valid method should be rugged and capable of withstanding environmental changes and other uncontrolled factors. To assess the method capability, ten samples were prepared according to Table 5.14 while analysis was performed using Option 1.

Table 5.14: Description of case samples considered for analysis

Sample	Description[a]
1 - 2	5 samples (40 mg, 50 mg, 60 mg, 70 mg and 80 mg) prepared and analyzed on day 1,
1 - 3	4 samples (50 mg, 60 mg, 70 mg, and 80 mg) prepared and analyzed on day 2.
1, 4 - 10	1 sample (80 mg) prepared and analyzed on day 3 1 diluted sample (from day 3) prepared and analyzed on day 4

[a]Analyses were performed as such due to the limited availability of the amount of the sample.

For sample-to-sample comparison at the distribution/street level, the GC readings were transformed to percentage analyte (called the variable-*i*) and each of the eight compounds was standardized (the variable-*i* divided by the standard deviation of that variable-*i*) prior to hierarchical cluster analysis (HCA) using the single linkage method with Euclidean distance measure as a tentative statistical technique. According to Figure 5.24, all ten target clusters were obtained at a similarity level of 92.59 illustrating that the method was capable of providing reproducible results irrespective of the inter-day variations.

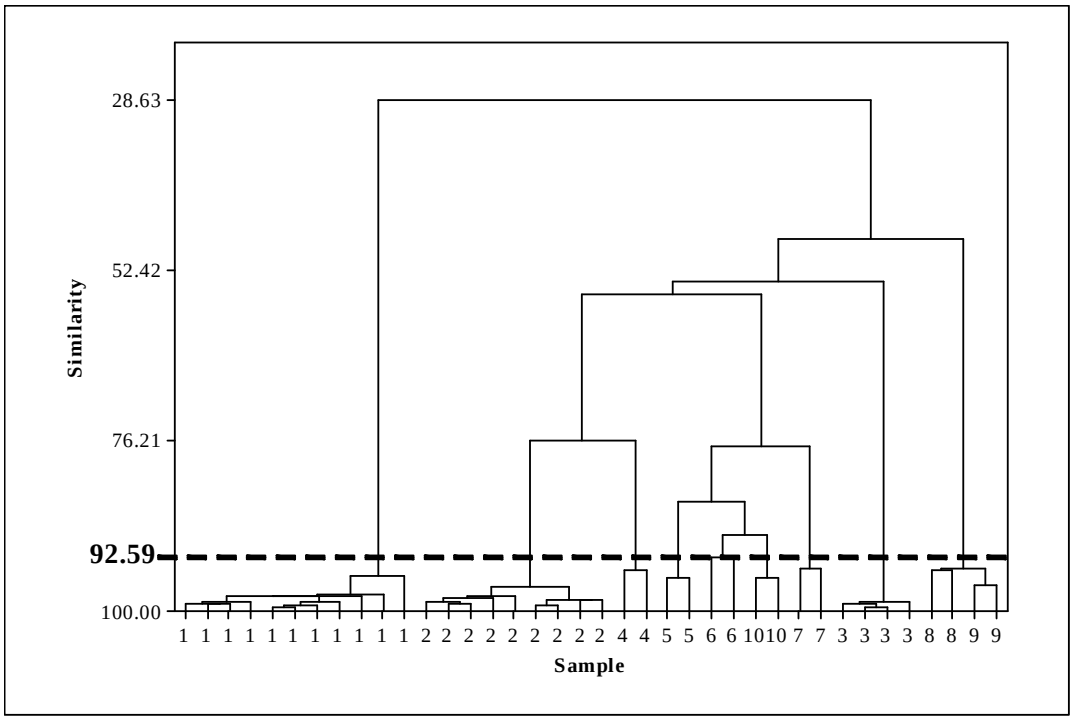


Figure 5.24: A dendogram expressed in similarity showing the relationships between ten samples analyzed under different conditions using Single linkage and Euclidean distance (Standardized % analytes are used to evaluate the capability of the method to cluster similar samples at the distribution level)

As the ratios of the opium-based alkaloids can minimize the adulterant effects and experimental errors (Zhang, D. *et al.*, 2004; Zelkowics *et al*., 2005), it is thus useful to assess the sample relationships that existed before they were cut. The GC readings in mg/mL were normalized to achieve five tentative ratios: AC/HR, AC/MM, AC/(MM+HR), CD/(MP+MM+HR) and AC/(MP+ MM+HR) without particularly addressing the decomposition effects at this stage (these ratios are optimized in Section 5.4.3 Statistical Validation Using Simulated Heroin Links). As such, the GC method can be reliably assessed without much help from the statistical correction. The selected quotients were only used to assess intra-sample variation by evaluating the relationships between the known related samples (e.g. relationships between data points of Sample 2) rather than inter-sample variation since the histories of the ten samples were uncertain. The multivariate data were decomposed by PCA in the correlation mode. Figure 5.25

illustrates that the samples within each individual cluster analyzed under different conditions show insignificant differences in the analytical results except for Sample 1. This could be due to the absence of heroin/diamorphine in Sample 1. Samples 3 and 4 prepared and analyzed at different concentrations with different calibration curves illustrated extremely close agreement in their respective clusters. The GC system proved rugged and very insensitive to experimental changes. Slight separations between the units within the individual clusters are theoretically due to random errors. In general, this method was able to cluster all the samples into their respective groups correctly.

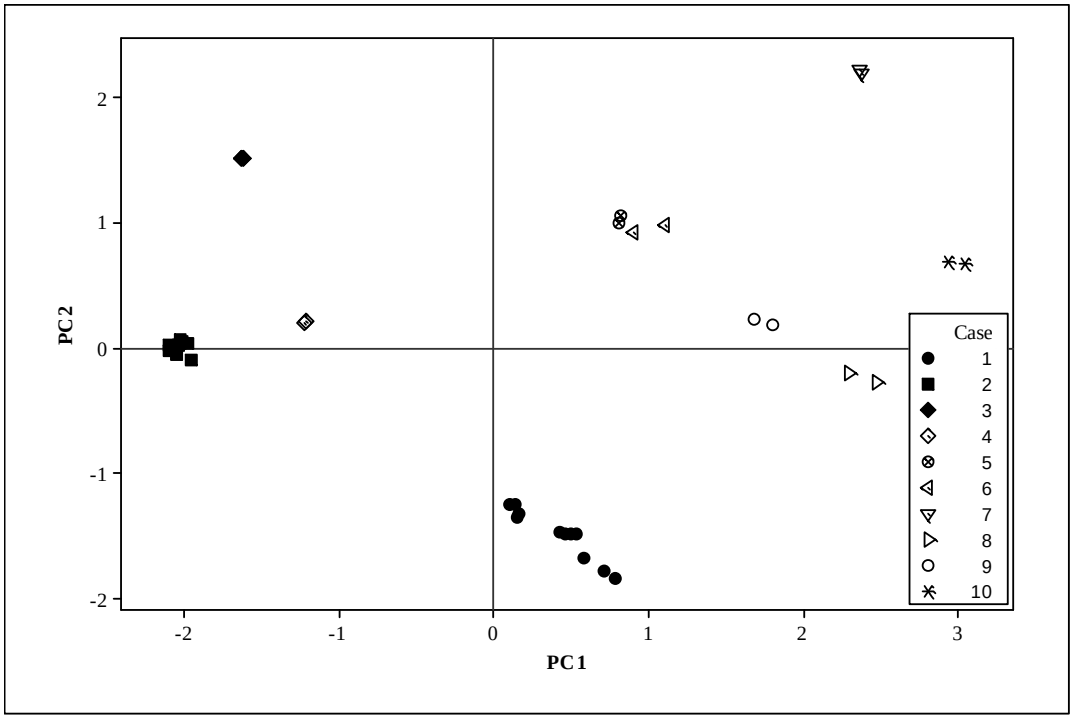


Figure 5.25: A score plot of PCA decomposed GC-FID data in correlation mode of ten samples analyzed under different conditions, $\%V_1$ = 55.6 % and $\%V_2$ = 26.4% (Five ratios of normalized alkaloids are used to evaluate the capability of the method to cluster similar samples at the source/manufacturing level; samples 3 and 4 show overlapped symbols, indicating extremely close agreement in their clusters)

5.4.1.9 Summary

Quantification of the eight major components in illicit heroin is feasible with GC–FID using a thicker film of an HP Ultra 2 column and the chromatographic conditions specified in Option 1. The method was partially validated by the selected aspects and it was found to be simple, accurate, precise and sensitive. To enhance the reliability of the results, it is proposed that the illicit heroin sample should be weighed in excess of 70 mg to ensure that the analytes are detectable and quantifiable. Glass volumetric flasks should also be used to minimize artifact formation. In addition, all prepared samples must be analyzed within a week. Statistical analysis of ten groups of heroin samples indicated a good capability of the method to link the related samples correctly.

5.4.2 Evaluation of the Statistical Robustness of the Optimized GC-FID Method in Sample Classification by Pattern Analysis

In this subtask, pattern analysis is proposed to evaluate the robustness of the optimized GC-FID method by utilizing the data of the five opium-based alkaloids present in the illicit heroin analyzed with the validated method. The three adulterants were disregarded as they do not contribute information to the sample origin. A valid method (Option 1) and a poor method (Option 2) which respectively serves as a positive control method (PCM) and a negative control method (NCM) were employed for statistical evaluation. The GC conditions of these two methods have been specified in Table 5.7. A total of 43 illicit heroin samples of unknown origins were analyzed using both methods by GC-FID. The robustness of both the methods was assessed by decomposing the GC data into two components using PCA. The statistical procedure eventually demonstrates how method robustness is achieved based on data distribution patterns. Hypothetically, a robust method (represented by the PCM) will show its unchanged pattern in data distribution whereas the method with poor robustness (represented by the NCM) will show significant pattern distortion.

5.4.2.1 Performance of the Methods

The optimized GC-FID conditions (Option 1 or the PCM) have shown excellent performance in Section 5.4.1 GC-FID Method Optimization and Validation. This method was further assessed statistically against a second method (Option 2 or the NCM) to assure that the GC data are free of measurable errors. Five critical validation aspects were also performed on Option 2 and the validation results are summarized in Table 5.15.

Table 5.15: Validation results for Option 2 (NCM)

Compound	Intra-day precision[a] RSD (n = 6)	Inter-day precision[b] RSD (n = 6)	Concentration range covered[c] (mg/mL)	Correlation coefficient, r^2	Recovery %
Codeine	1.16	1.56	0.02 – 0.05	0.9805	101.03
Morphine	3.02	5.29	0.05 – 0.11	0.9601	105.26
Acetylcodeine	0.39	0.85	0.05 – 0.11	0.9750	101.67
6-Monoacetyl-morphine	0.77	1.98	0.15 – 0.34	0.9742	106.51
Heroin	0.56	1.11	0.20 – 0.47	0.9802	98.06

[a]Intra-day precision (repeatability) was assessed using a single mixture containing five opium-based alkaloids at the working concentrations. The mixture was injected six times consecutively.
[b]Inter-day precision (reproducibility) was assessed as for repeatability but six injections were performed on different days.
[c]From the series of dilutions, only the consecutive points that give the best r^2 value for the regression were used to define the working range. The range obtained is however narrow and not linear.
Note: All GC data are expressed as peak area relative to the IS (area ratio).

According to Table 5.15, Option 2 showed repeatable and reproducible results for the five opium-based alkaloids. However, the linear working range of this method was significantly narrower than that of Option 1. The non-linear responses with r^2 values between 0.96 and 0.98 posed a problem with Option 2. Such r^2 values obtained for the linearity in this subtask are unacceptable because the analytes cannot be accurately quantified within the working ranges. At certain concentrations, Option 2 tended to exhibit high or low readings that gave rise to the poor r^2 values. Also, this method indicated excessive recovery for morphine and 6-monoacetylmorphine. In summary, the performance of Option 2 is unacceptable (with the above-mentioned undesirable linear ranges, linearity and recoveries) and therefore was chosen as the negative control (NCM) to illustrate its poor robustness in terms of pattern distortion. The target method, using Option 1 on the other hand was compared with Option 2 to reveal its robustness as an acceptable method (or the PCM) in establishing sample relationships through PCA.

5.4.2.2 Evaluation of Statistical Robustness using PCA

43 case samples were analyzed using Option 1 and Option 2 in duplicate. They were selected because these samples contained all the target peaks. Hence, the presence of all the target peaks will help minimize statistical errors arising from zero-values (absence of peaks) during statistical treatment. The GC data were obtained in two forms. The first type is by using the 'peak area relative to the IS' through which errors due to inconsistent split ratios and evaporative losses can be greatly minimized. The second type is to take the concentration in mg/mL based on the one-point calibration performed daily using area ratio (peak area relative to the IS). In addition to the errors addressed by the first method, the second method could also compensate for any unknown variables that would affect the GC system. In other words, the system and the GC data were corrected once it was calibrated using chemical standards.

To assess the relationships between samples of unknown origins, data normalization was performed on the GC data targeting five major opium-based alkaloids found in the illicit heroin in order to minimize environmental and statistical errors as well as the cutting effects due to the presence of adulterants. However, as related samples (or samples of a known common source) were not involved in this subtask, all samples were treated as independent units. Hence, inter-sample variation must be taken into consideration during sample classification. The data were again tentatively optimized according to CD/(MM+HR), MP/(MM+HR), AC/(MM+HR), CD/(MP+MM+HR), AC/(MP+MM+HR) and (CD+AC)/(MP+MM+HR) to collectively compensate for the effects of decomposition, adulteration and analytical errors. As heroin/diamorphine is not stable, the decomposition of heroin to 6-monoacetylmorphine and sometimes further deacetylation to morphine must also be taken into consideration. As a result, the sum of the morphine contents was employed as a denominator to form the quotients. The quotients also have a merit of overcoming the decomposition effects

and thus the sample relationships can be assessed on the basis of their possible origins. In this regard, the data obtained from Option 1 and Option 2 can be compared when all the possible factors have been taken into consideration. These pretreated data were decomposed by PCA to show the relationships between the samples.

Figures 5.26(a) and 5.26(b) respectively show the decomposed outcomes for the data obtained in 'peak area relative to the IS' and concentration calibrated by standards using Option 1. The score plots display that the general patterns of the data distribution seem to be vertically rotated as the GC data were interpreted in the two different forms. However, the relationships (in terms of distance) between the data points are relatively in close agreement irrespective of whether the system was calibrated by the standards. For example, Sample 65 is close to Samples 50 and 111 and it maintains a relatively long distance with Sample 277 in both plots. Insignificant differences in distance are also present in the score plots. In this case, the negligible deviations of the patterns in Figure 5.26(a) from Figure 5.26(b) are assumed to be the small random errors that should be corrected using chemical standards on a daily basis. In this regard, the data distribution in Figure 5.26(b) corrected by the daily calibration is thus more reliable than that in Figure 5.26(a) because the daily errors have been corrected by the chemical standards.

In order to view both score plots on the same scale, standardization is an ideal pretreatment method to transform the data into a standard type of readings for comparison. To this aim, each normalized variable was treated as the variable-*i* and the individual variables-*i* were then divided by the standard deviation computed from that variable-*i*. When the standardized data were decomposed by PCA in Figure 5.26(c) and 5.26(d), the distribution patterns and the relationships between the samples revealed in the new score plots are generally similar. Hence, the data distribution in the normalized form or in the standardized form of the normalized data is consistent with one general

relational pattern irrespective of whether the GC data were corrected by calibration or not. In this subtask, the comparison between the normalized data and the standardized form of normalized data is not discussed. The patterns in these forms are different because of the statistical pretreatment and not the analytical procedures. Further discussion regarding the statistical pretreatment is emphasized in Section 5.4.3 Statistical Validation Using Simulated Heroin Links.

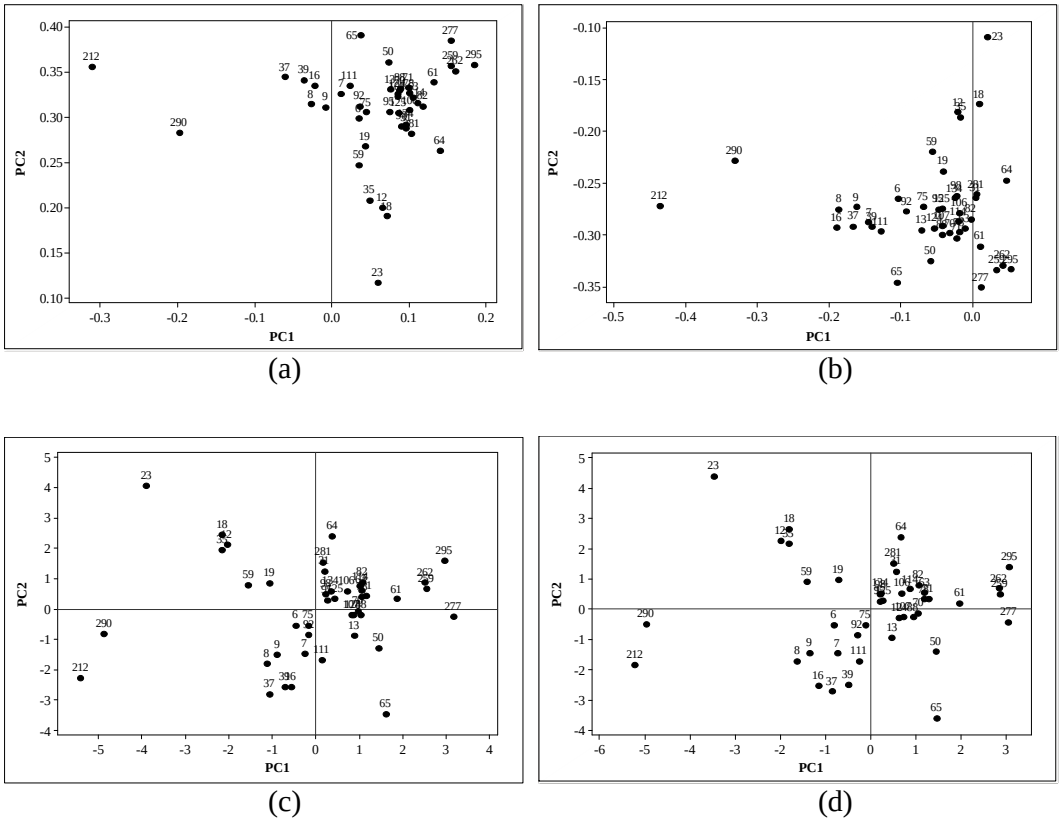


Figure 5.26: Score plots of 43 data points obtained from Option 1 and decomposed by PCA (The distributions of the data of (a) peak to the IS ($\%V_1$ = 73.0%, $\%V_2$ = 25.4%) and (b) concentration in mg/mL ($\%V_1$ = 80.0%, $\%V_2$ = 19.1%), are obtained by decomposing the data in covariance mode. The distributions of the data of (c) peak to the IS with standardization ($\%V_1$ = 55.5%, $\%V_2$ = 40.5%) and (d) concentration in mg/mL with standardization ($\%V_1$ = 55.3%, $\%V_2$ = 41.4%), are obtained by decomposing the data in correlation mode. The labeling is the original case assignment. The patterns in (a) and (b) are due to the difference in the types of data readings. Using standardization, the patterns in (c) and (d) become very similar and the relationships between data points are not affected)

In summary, the optimized method with Option 1 is statistically robust and free of any significant analytical errors. The distribution and relationships depicted in the plots were used as a standard map against which the Option 2 method was compared.

Measurable pattern distortions are observed in the data obtained with Option 2. From Figures 5.27(a) and 5.27(b), the data in the normalized form show different sample relationships between uncalibrated and calibrated readings. When the data were standardized and decomposed by PCA, the patterns in Figures 5.27(c) and 5.27(d) are still not in close agreement. The calibrated data distribution pattern showed some distortion as compared to the uncalibrated data distribution pattern when the daily calibration was considered for analysis. This infers that the analytical errors occurring in Option 2 are significant even though they were corrected by the daily calibration. As a result, different interpretations (in terms of data relationships) could be generated from the poor method. Sample relationships become very dependent on the types of readings processed by the PCA. Hence, methods with poor robustness will reveal pattern distortion such as that demonstrated in the poor method/NCM when different types of readings are used.

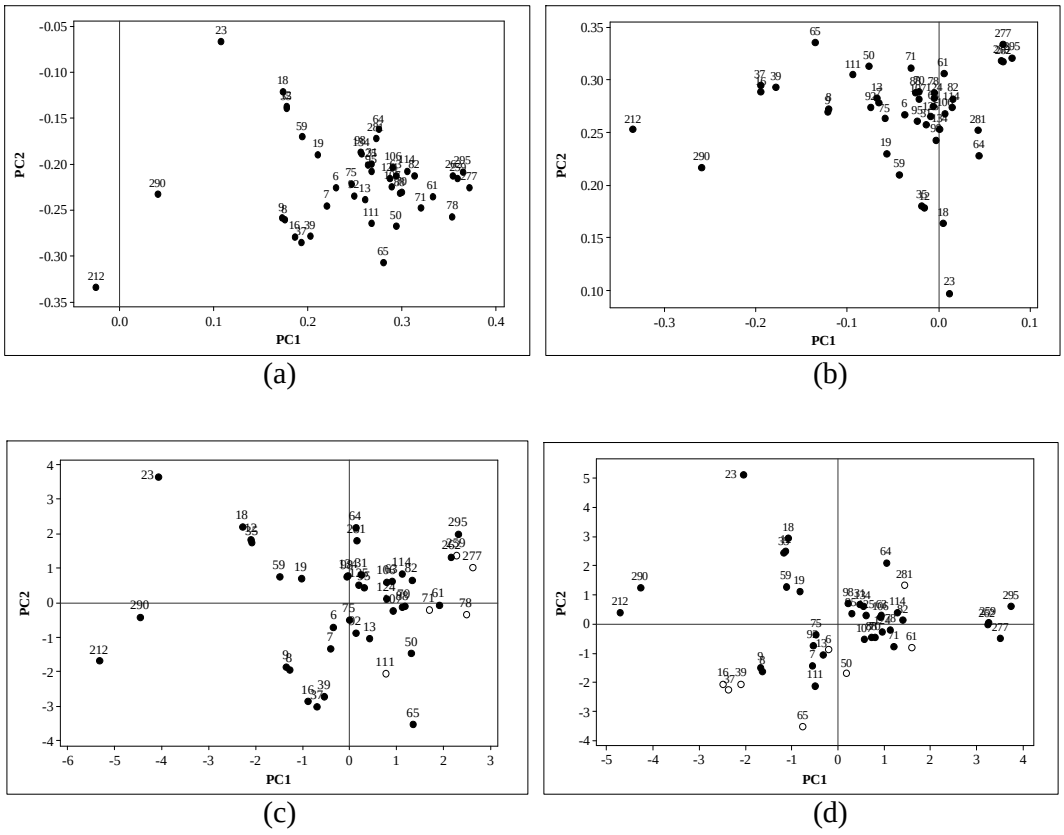


Figure 5.27: Score plots of 43 data points obtained from Option 2 and decomposed by PCA (The distributions of the data of (a) peak to the IS ($\%V_1$ = 71.4%, $\%V_2$ = 26.2%) and (b) concentration in mg/mL ($\%V_1$ = 76.7%, $\%V_2$ = 22.2%), are obtained by decomposing the data in covariance mode. The distributions of the data of (c) peak to the IS with standardization ($\%V_1$ = 53.7%, $\%V_2$ = 41.8%) and (d) concentration in mg/mL with standardization ($\%V_1$ = 54.9%, $\%V_2$ = 41.8%), are obtained by decomposing the data in correlation mode. Using standardization, the patterns in (c) and (d) are not in close agreement)

A comparison between Figures 5.26 and 5.27 shows the extent to which the data differ in both systems. For the standardized data, measurable distortion impacts were experienced by the data points in bright circles when the score plots in Figures 5.27(c) and (d) are compared against the corresponding score plots in Figures 5.26(c) and (d). Figure 5.27(c) demonstrates some discrepancy in the upper right portion as compared to Figure 5.26(c). Furthermore, Option 2 resulted in loosely packed data. A severe shift was observed in sample 78 which was displaced from the cluster in the middle area. Although the possible errors in Option 2 were corrected by the daily calibration, the

data points in Figure 5.27(d) are somewhat inconsistent with those in Figure 5.26(d). Significant shifts were observed in Samples 50, 61 and 65 and the general pattern in the lower left portion of the score plot. For example, the cluster comprising Samples 16, 37 and 39 should be closely related to Samples 7 and 111 but the former have been incorrectly separated from the latter in Figure 5.27(d). Taking Figure 5.26(d) as the most reliable pattern, it can be concluded that the data from Option 2 generally show distorted parts in their pattern and altered relationships. Misinterpretation can arise if this unreliable method is used for establishing the sample relationships.

5.4.2.3 Summary

In summary, an analytically valid method should also perform statistically well in data interpretation. Option 1 was shown to be analytically and statistically robust irrespective of how the data were manipulated (e.g. calibrated versus uncalibrated data). Option 2 on the other hand is indeed a poor method and therefore not robust. It illustrated that analytical errors associated with the method will lead to data misinterpretation.

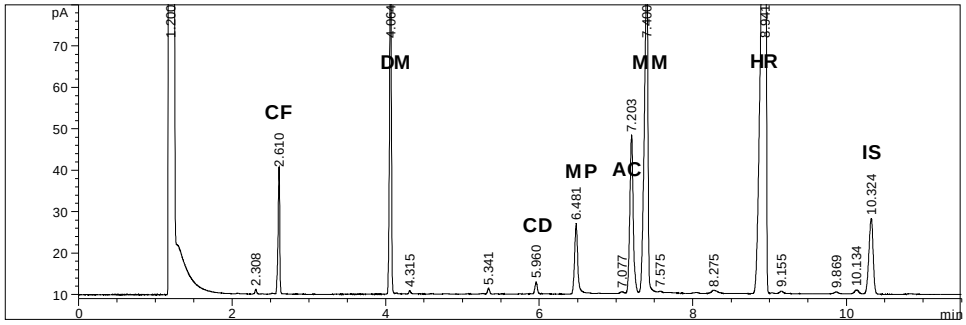
5.4.3 Statistical Validation Using Simulated Heroin Links

After the analytical method (Option 1) has been demonstrated to be functioning well as predicted with the samples under this study, statistical validation was performed on the GC data to find the suitable opium-based alkaloid quotients, the best pretreatment method as well as the ideal statistical classification techniques for unsupervised pattern recognition. Therefore, a simulated dataset containing eight links with known sample histories were prepared from 8 pre-cut samples (A1, A2, B1, B2, C1, C2, D1 and E1) to simulate the chemical composition of the local heroin samples. Each link/batch contained 9 samples (1 pre-cut, 4 uncolored post-cuts and 4 colored post-cuts). In addition, each sample was prepared in 3 individual aliquots which were then analyzed in triplicate. The natural intra-sample and intra-batch variation as well as the inter-batch variation contained in the dataset were used for statistical validation. An ideal statistical technique should be able to minimize the intra-sample and intra-batch variation and maximize the inter-batch variation for sample classification.

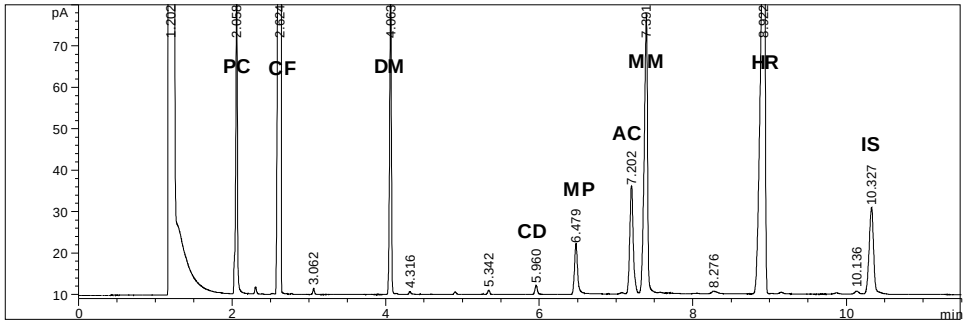
5.4.3.1 Simulated Dataset

A total of 8 simulated links of illicit heroin were prepared by diluting/cutting the high purity heroin with adulterants/cutting agents. By taking into consideration the dealer's interest in coloring the heroin samples, half of the post-cut samples were mixed with ten food coloring agents while the other half were left uncolored. These colors were chosen because they served to cover the red, green, orange and yellow appearance of the street heroin commonly encountered in Malaysia. The cutting process inevitably changes the absolute abundance of the opium-based alkaloids but the change usually takes place in a proportional manner. Figure 5.28 shows the change in the content before and after the cutting agents were added to a pre-cut sample. The GC chromatograms also reveal that the unknown background impurities originally present

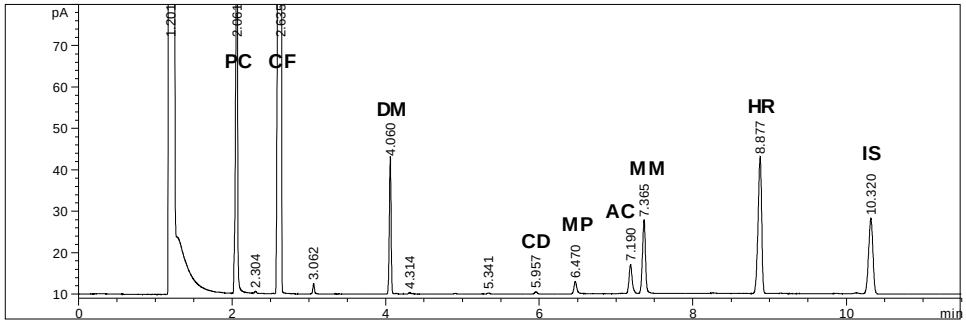
in the pre-cut sample tend to be attenuated by the cutting agents but the opium-based alkaloids remain detectable and are reduced proportionately in a constant ratio. When the samples were largely cut, most of the unknown impurities became undetectable with the GC-FID (Figure 5.28(c)). This suggests the feasibility of the five chosen opium-based alkaloids instead of other insignificant impurities for the profiling of heroin in this task.



(a)



(b)



(c)

Figure 5.28: Chromatograms of pre-cut and post-cut samples of B2 (The chromatograms respectively represent (a) pre-cut sample (b) 50% pre-cut sample with 50% cutting agents (c) 12.5% pre-cut sample with 87.5% cutting agents)

The simulated chemical composition is represented by the box-and-whisker plot in Figure 5.29 which summarizes the percent composition of the major components in the pre-cut and post-cut samples. Outliers indicated by asterisks represent the contents of the pre-cut samples. Those represented by the whisky-boxes are the street contents of the 64 post-cut samples. The opium-based alkaloid contents were lowered in these samples as their respective pre-cut samples were diluted by the cutting agents. Upon cutting, the post-cut samples containing 0.54 – 22.26% heroin base were obtained for the simulated dataset. About 75% of the samples contained < 10% heroin base and these contents were close to those of the real case samples and hence ideal for statistical validation. In general, the dataset was suitably simulated to represent the compositions of the local heroin samples.

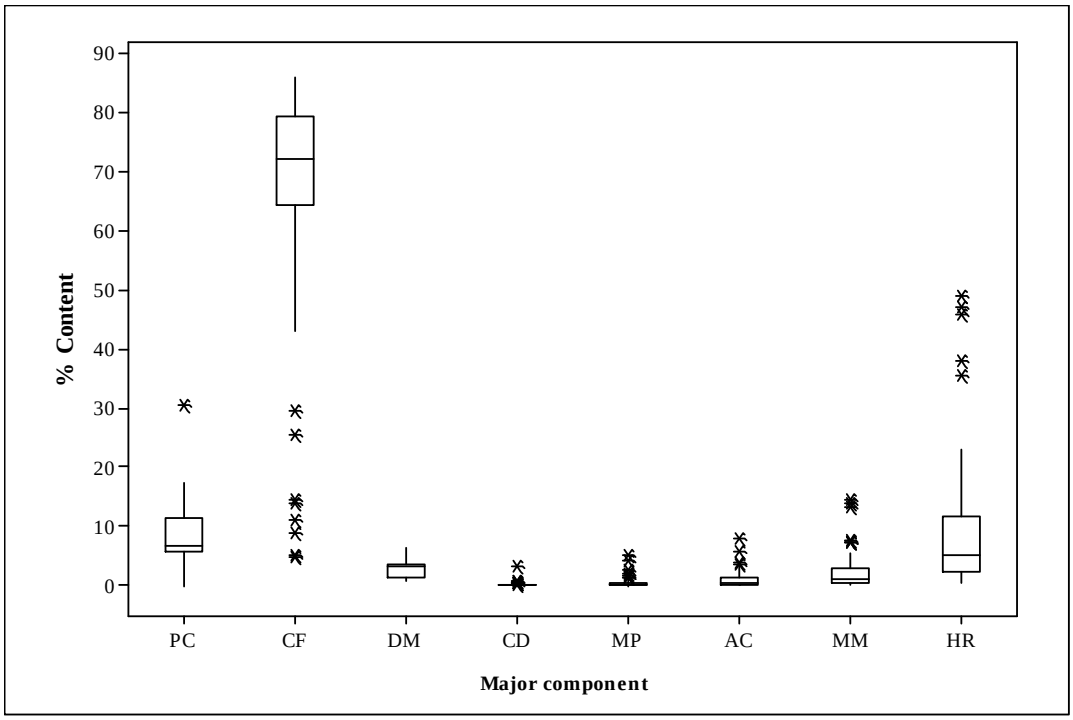


Figure 5.29: Boxplots showing the % contents of eight major components encountered in 8 pre-cut samples, 32 post-cut (uncolored) and 32 post-cut (colored) samples (The '*' is the content of the pre-cut sample)

5.4.3.2 Cutting Efficiency

In the post-cut samples, the composition obtained is sometimes not in accordance with the intended composition. To ensure the simulated dataset contain the intended composition, the cutting efficiency that serves as a gross measure of how well the cutting process has been performed was evaluated. This is best described by the regression line between the measured content against the percentage of the pre-cut sample added. A good cutting process should decrease the alkaloid content of interest in a linear pattern. To measure this, morphine, monoacetylmorphine or heroin/diamorphine is not a good indicator since their contents are easily affected by decomposition. Low amounts of codeine may not be quantifiable in cases where the samples were highly cut. Hence, acetylcodeine was chosen to evaluate the cutting efficiency by plotting the recovered % acetylcodeine quantified versus the theoretical % of pre-cut sample added. The r^2 value in Table 5.16 is a good measure to assess how close the recovered and theoretical contents are. In general, the cutting was done reasonably well since all the 8 links achieved an $r^2 > 0.99$ and the alkaloid contents decreased in the intended linear fashion when more cutting agents were added. From the equation, the negative y-intercept suggests losses of analytes that are probably due to heating.

Table 5.16: Cutting efficiency shown by the regression line between % acetycodeine quantified versus % pre-cut sample added

Link	Linearity function	Correlation coefficient, r^2
A1	y = 0.0020x – 0.0067	0.9985
A2	y = 0.0133x – 0.0261	0.9974
B1	y = 0.0271x + 0.0378	0.9939
B2	y = 0.0431x – 0.1243	0.9990
C1	y = 0.0112x – 0.0076	0.9999
C2	y = 0.0102x – 0.0282	0.9996
D1	y = 0.0068x – 0.0007	0.9979
E1	y = 0.0523x – 0.1357	0.9982

5.4.3.3 Compositional Variation

At the distribution/street level, the absolute % contents of the major components are the ideal parameters to distinguish between post-cut samples. When a pre-cut sample undergoes the same cutting process in a single cutting line, it will usually give a similar composition for the derivative post-cut samples. However, discrepancy sometimes occurs. According to Table 5.17, Batch 1 and Batch 2 being from two different pre-cut samples must show different compositions despite being cut with 75% of mixture X and hence they belonged to two different cutting lines. However, slight deviations are observed in the batch. For example, caffeine in B1U3-2 and B1U3-3 differs by 1.82%. In fact, the extent to which the analyte contents in the heroin samples from a similar batch differ relies on 1) the cutting process and 2) whether the samples are homogenized before they are packed. In a clandestine laboratory, except for the single action of stirring during the cutting process to ensure even distribution of the components, the heroin packer would scoop out the heroin granules into individual packages without further homogenizing the whole of the post-cut heroin sample. This subtask also simulated the similar operation by not homogenizing the post-cut sample before sampling was done. As a result, some deviation in the compositions of the same post-cut samples was obtained in the simulated dataset.

Table 5.17: Compositional comparison of post-cut samples

Batch	ID.	Cutting agent	PPS (%)	PC	CF	DM	CD	MP	AC	MM	HR
1	A1U3-1	X	25	5.90	67.93	3.94	0.01	0.05	0.75	0.81	11.08
	A1U3-2			5.94	66.66	3.96	0.01	0.04	0.75	0.81	11.16
2	B1U3-2	X	25	5.25	67.70	3.58	0.10	0.68	0.87	3.52	9.69
	B1U3-3			5.35	65.88	3.64	0.11	0.72	0.89	3.64	9.96

The extent to which these similar samples deviate (the intra-sample variation) is the natural variation associated with the cutting process. This was investigated by

examining the RSD for each component contained in three laboratory samples taken from the random parts of a post-cut sample. Then, the average of the RSDs of all the post-cut samples belonging to the same pre-cut sample (PPS) ratio was calculated. Table 5.18 summarizes the maximum, minimum and average RSDs calculated from n-number of post-cut samples in the same proportion of the pre-cut sample (PPS) ratio. The results show that all the components tended to exhibit a wide range of RSDs, indicating that the samples taken from a single post-cut sample varied irrespective of the cutting ratios and colors. The dyes added did not have significant effects on the composition. As the sampling for analysis was done randomly through scooping, hence the lower RSDs shown in the minimum suggests proper mixing as well as even distribution of the components in the sample during cutting. Improper mixing and uneven distribution resulted in higher RSDs. Besides the factor of mixing, it is expected that some variation in morphine, monoacetylmorphines and heroin could be the result of hydrolytic decomposition in the presence of water added and heating since these two factors often render instability of the morphine-based alkaloids (UNODC, 2005). These unpredictable factors collectively give rise to the natural variation called the intra-sample variation (referring to the variation within a post-cut sample) as well as the intra-batch variation (referring to the variation within the same link). According to Table 5.18, the intra-sample variation observed from the cutting process show an RSD < 11% for all the components except for morphine (RSD < 19 %).

The natural variation is random. It could occur in samples from two similar cutting processes prepared on the same day. This was evident in the four individual samples prepared in the same cutting processes which showed an RSD < 16% in their composition (Table 5.19). This shows that the random parts of a post-cut sample varied in their composition although the mixing process was carried out satisfactorily well during cutting.

Table 5.18: Intra-sample RSD (%) based on the analyte content for n-number[a] of uncolored and colored post-cut samples

PPS (%)			RSD (%)															
			PC		CF		DM		CD		MP		AC		MM		HR	
			U	C	U	C	U	C	U	C	U	C	U	C	U	C	U	C
		Max	4.95	5.95	1.52	2.04	4.42	5.62	-	-	-	-	4.69	5.99	5.82	6.13	4.34	6.44
2.5	**(n=3)**	**Mean**	**3.56**	**3.58**	**0.68**	**1.35**	**3.47**	**3.82**	**1.52[1]**	**4.30[1]**	**9.31[1]**	**6.24[1]**	**3.43**	**5.06**	**4.34**	**5.19**	**3.18**	**4.65**
		Min	1.01	2.04	0.24	0.70	1.57	1.31	-	-	-	-	1.47	3.54	2.74	3.35	1.32	1.87
		Max	8.15	5.34	2.36	1.85	6.93	6.99	5.80[3]	9.36[3]	12.42[5]	15.99[5]	7.35	6.59	7.75	5.89	7.11	5.26
5	**(n=6)**	**Mean**	**4.33**	**3.50**	**1.34**	**1.06**	**4.05**	**3.63**	**5.37[3]**	**5.89[3]**	**9.00[5]**	**7.61[5]**	**4.79**	**3.36**	**5.52**	**3.81**	**4.64**	**3.13**
		Min	0.97	1.14	0.48	0.44	1.66	1.26	4.72[3]	3.65[3]	4.98[5]	3.43[5]	2.63	0.85	3.16	1.58	2.43	0.81
		Max	4.67	3.71	1.21	2.12	5.07	3.78	4.31[2]	3.49[2]	5.08	11.03	5.76	4.68	4.36	6.10	5.58	4.58
7.5	**(n=3)**	**Mean**	**2.88**	**2.08**	**0.79**	**1.24**	**3.51**	**2.10**	**5.10[2]**	**2.65[2]**	**4.14**	**7.11**	**3.67**	**2.63**	**3.69**	**3.72**	**3.51**	**2.31**
		Min	0.50	0.98	0.54	0.75	0.85	1.02	5.89[2]	1.80[2]	2.42	1.27	0.90	1.34	0.93	2.24	0.46	0.74
		Max	8.25	5.85	2.24	2.73	8.24	5.98	7.73[6]	7.43[6]	18.45	5.98	7.79	6.75	7.33	10.22	7.77	6.00
12.5	**(n=7)**	**Mean**	**3.71**	**3.85**	**1.18**	**1.26**	**3.87**	**3.82**	**6.25[6]**	**4.70[6]**	**8.25**	**4.78**	**4.58**	**4.02**	**4.97**	**4.62**	**4.47**	**3.77**
		Min	1.15	2.79	0.45	0.74	1.54	2.83	5.08[6]	2.51[6]	5.13	2.48	2.75	2.66	1.54	2.78	2.69	2.68
		Max	6.49	7.53	2.73	2.36	8.32	7.13	8.49[8]	7.08[8]	10.16	7.96	8.38	6.76	6.52	5.66	8.64	7.11
25	**(n=9)**	**Mean**	**3.30**	**3.59**	**1.70**	**1.56**	**4.30**	**3.50**	**5.23[8]**	**3.82[8]**	**4.72**	**4.33**	**4.21**	**3.32**	**3.61**	**2.76**	**4.22**	**3.57**
		Min	1.14	0.30	0.91	0.54	1.12	0.40	1.91[8]	0.64[8]	1.28	1.41	1.44	0.55	1.80	0.38	1.72	1.42
		Max	3.23	4.83	2.13	3.56	2.42	5.02	3.08	6.21	3.39	5.66	1.91	5.17	2.78	5.40	2.80	5.19
50	**(n=4)**	**Mean**	**1.92**	**2.77**	**1.29**	**1.40**	**1.71**	**2.73**	**2.51**	**3.72**	**1.89**	**2.96**	**1.52**	**2.76**	**1.79**	**2.60**	**1.71**	**2.64**
		Min	0.72	1.14	0.60	0.56	0.84	0.96	1.73	1.97	0.92	1.77	0.91	1.11	0.94	1.30	1.02	0.94

[a]The n-value indicates the number of post-cut samples included in consideration for the max, mean and min values in each PPS category.
U = Uncolored, C = Colored.
Superscript = Number of post-cut samples considered after excluding the sample(s) with undetected analyte.

Table 5.19: Intra-sample RSD (%) based on the analyte content for two similar cutting processes

Sample ID	PPS (%)	RSD (%)															
		PC		CF		DM		CD		MP		AC		MM		HR	
		U	C	U	C	U	C	U	C	U	C	U	C	U	C	U	C
A2U1	**5**	3.81	3.39	0.66	0.84	4.04	3.37	N.D.	N.D.	9.55	15.90	4.44	4.44	5.97	5.30	4.25	4.23
A2U2		8.15	2.34	2.36	1.83	6.93	1.54	N.D.	N.D.	12.42	6.99	7.35	1.19	7.75	3.29	7.11	1.47
C1U3	**25**	3.04	7.53	2.50	2.36	8.32	7.13	7.50	7.08	8.59	7.48	8.26	8.38	4.36	5.66	8.64	7.11
C1U4		6.49	2.08	1.07	1.55	8.30	1.98	8.49	2.82	10.16	2.90	6.76	1.60	7.18	1.56	8.19	1.15

U = Uncolored.
C = Colored.
N.D. = Not detected.

5.4.3.4 Evaluation of Pretreatment Methods

As indicated by the compositional variation, the obtained intra-sample variability renders some degree of dissimilarity between similar/related samples, and this must be minimized through statistical treatment. The dataset was found suitable for statistical validation because it contained the preferred ranges of analyte contents and also a reasonable range of compositional variations in the samples. With this simulated dataset (containing 216 data points obtained from the total individual aliquots analyzed), statistical validation was performed in two steps, namely 1) data pretreatment and 2) assessment of the linkage methods and distance measures. As all the GC data were obtained in concentration units (mg/mL) which is equivalent to peak area relative to the IS, so the differential effects arising from the instrument were readily minimized. To assess the origin of the samples, all data should be further normalized to minimize the cutting effects as well as the weight difference associated with weighing. Two initial sets of normalized data were feasible. The first method denoted as N_{sum}: each alkaloid is normalized to the sum of alkaloids. The N_{sum} data of five individual alkaloids over the sum (namely, CD/Sum, MP/Sum, AC/Sum, MM/Sum and HR/Sum) were further pretreated with standardization and fourth root according to the categories in Table 5.20. Subsequently, the five parameters in each pretreated category were directly decomposed by PCA (in covariance mode) into the first three components in order to ascertain the category that was able to group the related samples according to the known links. The covariance mode is important to ensure the PCA decomposes the data according to the pretreatment method. In contrast, if the correlation mode is set, all the pretreated data will automatically be standardized. Therefore this mode was not chosen for this subtask. The standardization step was performed separately prior to the PCA.

Table 5.20: Pretreatment methods for GC-FID data (N_{sum} and $N_{selected}$ respectively are the individual variable-*i*, SD_i = standard deviation of that variable-*i*)

Normalization	Pretreatment	Abbreviation	Formula
N_{sum}	Standardization	$N_{sum} + S$	$\frac{N_{sum\,i}}{SD_i}$
	Fourth root	$N_{sum} + 4R$	$\sqrt[4]{N_{sum\,i}}$
	Standardization + fourth root	$N_{sum} + S + 4R$	$\sqrt[4]{\frac{N_{sum\,i}}{SD_i}}$
	Fourth root + standardization	$N_{sum} + 4R + S$	$\frac{\sqrt[4]{N_{sum\,i}}}{SD_i}$
$N_{selected}$	Standardization	$N_{selected} + S$	$\frac{N_{selected\,i}}{SD_i}$
	Fourth root	$N_{selected} + 4R$	$\sqrt[4]{N_{selected\,i}}$
	Standardization + fourth root	$N_{selected} + S + 4R$	$\sqrt[4]{\frac{N_{selected\,i}}{SD_i}}$
	Fourth root + standardization	$N_{selected} + 4R + S$	$\frac{\sqrt[4]{N_{selected\,i}}}{SD_i}$

In PCA, ideal grouping should comprise seven distinct groups derived from the seven unrelated pre-cut samples (Groups C1 and C2 should be closely packed in a cluster because they were only different in adulterants but possess the same alkaloidal ratios). However, none of the pretreated N_{sum} data was able to group the samples according to the links into distinct groups. The failure of N_{sum} and its pretreatment is probably due to the compositional change in morphine, monoacetylmorphines and heroin arising from hydrolysis. Therefore, a second normalization method denoted as $N_{selected}$ was employed: a selected alkaloid or a combination of alkaloids is normalized to a single or a combination of alkaloids. Various $N_{selected}$ quotients were formed based on three criteria: 1) the quotient value should avoid 'zero' arising from codeine and/or morphine, 2) simple quotients such as MM/HR are critical to highly stable samples whose composition is not greatly altered and 3) complex quotients such as

(CD+MP+MM+HR)/AC will help compensate for decomposition effects for unstable samples in which the alkaloid composition is altered to a measurable extent. The quotients were further pretreated according to Table 5.20 and decomposed by PCA (in covariance mode) into the first three components. Based on repeated testing using PCA, successful grouping was achieved by eleven $N_{selected}$ quotients with standardization ($N_{selected}$ + S) through the PCA (Figure 5.30). These $N_{selected}$ quotients included AC/HR, AC/MM, AC/(MM+HR), AC/(MP+MM+HR), MM/HR, (CD+MP)/ (MM+HR), HR/MM, (CD+AC)/(MP+MM+HR), (CD+MP+MM+HR)/AC, HR/(CD+MP+AC+MM) and (MP+MM+HR)/(CD+AC). The success of these quotients could be due to the relatively sufficient variabilities of the eleven $N_{selected}$ + S parameters to allow for the separation of the unrelated samples to the extent without jeopardizing the relationships between the related samples (Appendix 11).

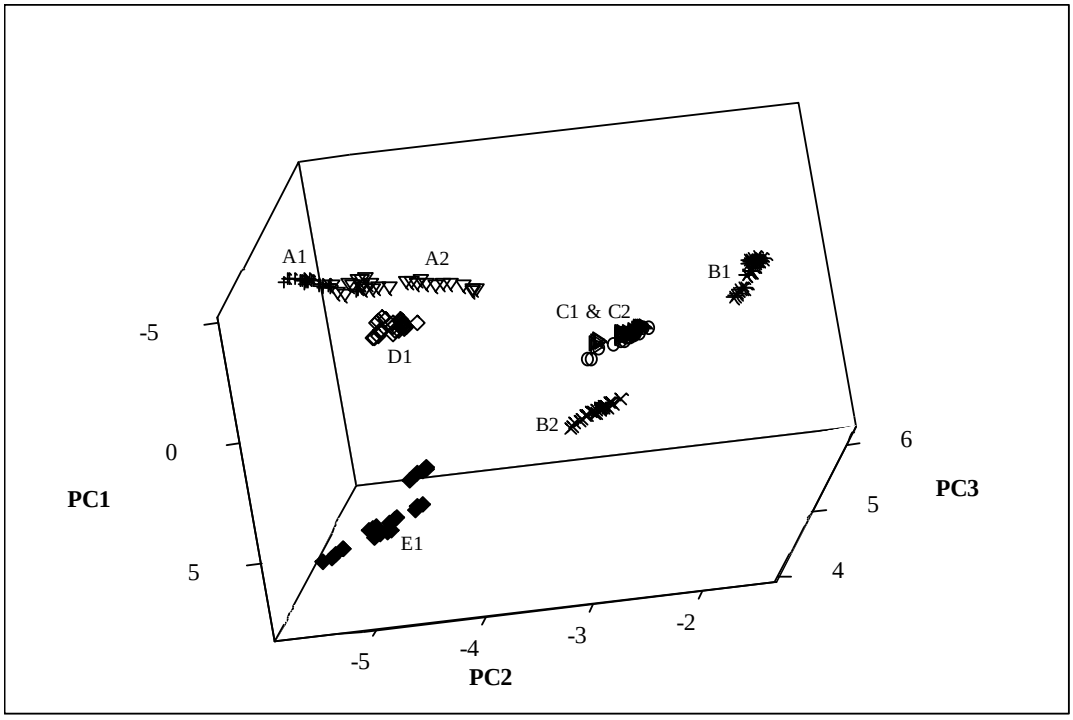


Figure 5.30: A score plot representing 11 $N_{selected}$ + S parameters of 216 data points decomposed by PCA in covariance mode into three dimensions, $\%V_1$ = 80.5%, $\%V_2$ = 16.8% and $\%V_3$ = 1.7% (Units of A1 and A2 are packed within their distinct groups although the groups are close together)

As part of the statistical validation, it is also crucial to examine the contributions of the selected parameters to the principal components in terms of loadings. According to Table 5.21, the loadings relating the eleven standardized parameters to the first three components are the correlations of the variables with the factors. The contributions of 6 $N_{selected}$ + S parameters (labeled 1, 3 – 6, 8) with loadings > 0.3 were found to be associated with the first component as compared to the rest. Almost all the $N_{selected}$ + S parameters showed low correlations in the second component. Relatively high loadings (> 0.4) were associated with three $N_{selected}$ + S parameters (labeled 5, 9, 11) in the third component. On the other hand, 99% of the total variabilities of these parameters were retained in the first three components for sample clustering (Figure 5.30).

Table 5.21: Loadings of the first three principal components of 11 $N_{selected}$ + S data of 216 simulated samples

Label	$N_{selected}$ + S parameter	PC1	PC2	PC3
1	AC/HR	0.318	-0.206	0.320
2	AC/MM	-0.166	-0.638	-0.110
3	AC/(MM+HR)	0.320	-0.224	-0.002
4	AC/(MP+MM+HR)	0.318	-0.233	-0.036
5	MM/HR	0.322	0.029	0.578
6	(CD+MP)/(MM+HR)	0.327	-0.047	0.195
7	HR/MM	-0.263	-0.457	0.120
8	(CD+AC)/(MP+MM+HR)	0.318	-0.223	0.053
9	(CD+MP+MM+HR)/AC	-0.318	0.133	0.537
10	HR/(CD+MP+AC+MM)	-0.278	-0.408	0.093
11	(MP+MM+HR)/(CD+AC)	-0.329	0.027	0.445

The suitability of the eleven $N_{selected}$ + S parameters was further confirmed by discriminant analysis (DA) which assigns unknown samples based on the characteristics of known samples. From the three random samples of each post-cut sample, one of them was segregated as a blind test unknown sample with the remaining as the known samples for training the DA model. Therefore, 144 training samples were used to classify 72 blind test samples using DA. Table 5.22 shows that the training and test

samples were correctly grouped under their respective links except some samples in Groups C1 and C2. As with PCA, these two groups are in fact a single group since their samples shared the same alkaloidal ratios. Hence, the sample units of these groups are flexible and acceptable to be clustered in either C1 or C2. As $N_{selected}$ + S showed promising grouping results in both the PCA and DA, this pretreatment method was chosen for the evaluation of linkage methods and distance measures by HCA.

Table 5.22: Summary of classification with cross-validation for 216 simulated samples

Put into Group	A1	A2	B1	B2	C1	C2	D1	E1
A1	**18** (9)							
A2		**18** (9)						
B1			**18** (9)					
B2				**18** (9)				
C1					**17** (9)	**10** (5)		
C2					**1**	**8** (4)		
D1							**18** (9)	
E1								**18** (9)
Total N	**18**	**18**	**18**	**18**	**18**	**18**	**18**	**18**
N correct	**18**	**18**	**18**	**18**	**17**	**8**	**18**	**18**
Proportion	**1.000**	**1.000**	**1.000**	**1.000**	**0.944**	**0.444**	**1.000**	**1.000**
N = 144		**N Correct = 133**					**Proportion Correct = 0.924**	

Note 1: Linear discriminant function was used.
Note 2: Figure in bold = Figure for training set.
Note 3: Parenthesis = Number of blind test samples assigned.

5.4.3.5 Evaluation of Linkages and Distance Measures

With the aid of HCA, seven linkage methods and five distance measures available in the Minitab 15 software were assessed using $N_{selected}$ + S data. A good statistical technique must fulfill three criteria. First, for linked samples, it must have none or the least number of erroneously clustered sample units. Second, the sample units in each cluster should be closely packed. This can be estimated by the maximum intra-group distance, d (the distance value at which the last linkage is located) within

the same group. This distance is an indication as to how farther apart the linked sample units are located within the group. The smaller the d value, the closer the sample units tend to be. Third, the groups should be distinctly separated. Inter-group separation is grossly measured by the maximum inter-group distance, D that is the distance between the two final clusters. This measure is an indication as to how farther apart all the data points are scattered. Larger D values indicate better separation. The second and third criteria can be evaluated by the following formula, d_m:

$$\text{Modified distance index, } d_m = \frac{\text{average d}}{\text{D}} \text{ X } 100$$

whereby average d is the mean of the d values of all the known groups

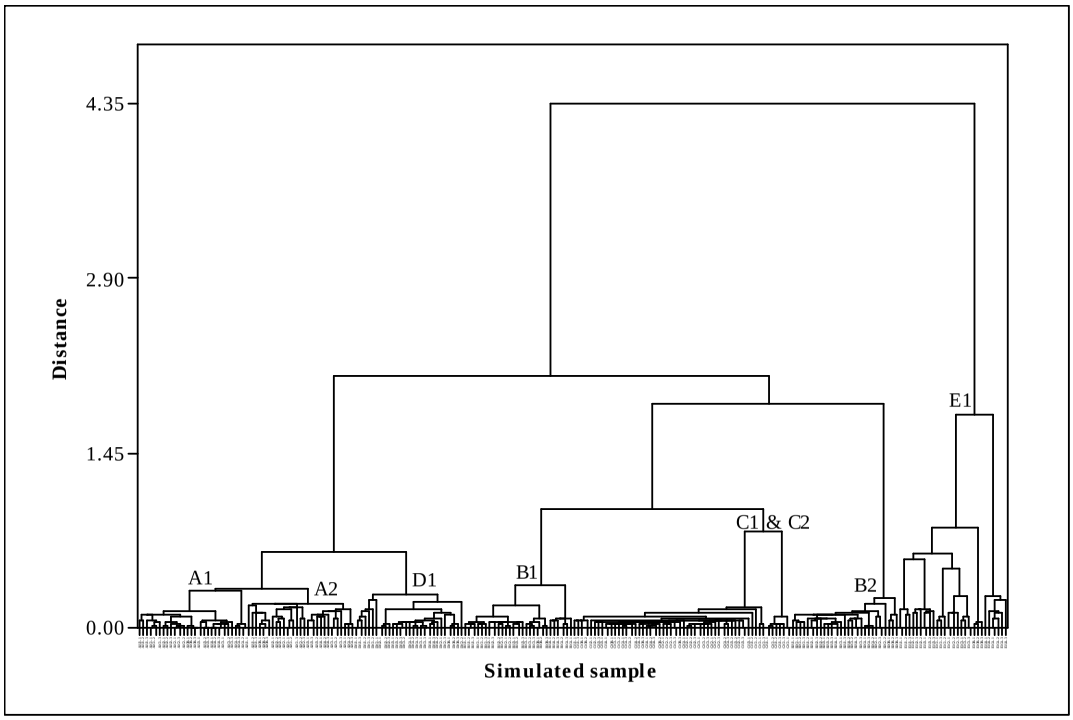
A good clustering method will demonstrate a low value of d_m. This also means that it will illustrate closely packed linked samples and widely separated clusters (unlinked samples) on a dendogram.

The clustering performance of the linkages and distance measures is presented in Table 5.23. The pair of Euclidean and Pearson as well as the pair of Squared Euclidean and Squared Pearson achieved the same performance in the clustering. The use of one can supplant the use of the other in its pair. Based on the number of mistakes recorded in Table 5.23, only seven linkage-distance combinations were able to cluster all the linked sample units (with zero erroneously clustered units or mistaken units) into their respective groups. They were Single-Euclidean, Ward-Euclidean, Ward-Manhattan, Single-Pearson, Ward-Pearson, Single-Squared Euclidean and Single-Squared Pearson. Other combinations were not ‘sensitive’ to differentiate between certain sample units of A1, A2 and D1 on the dendogram, rendering closely packed data points coming from these groups in close proximity.

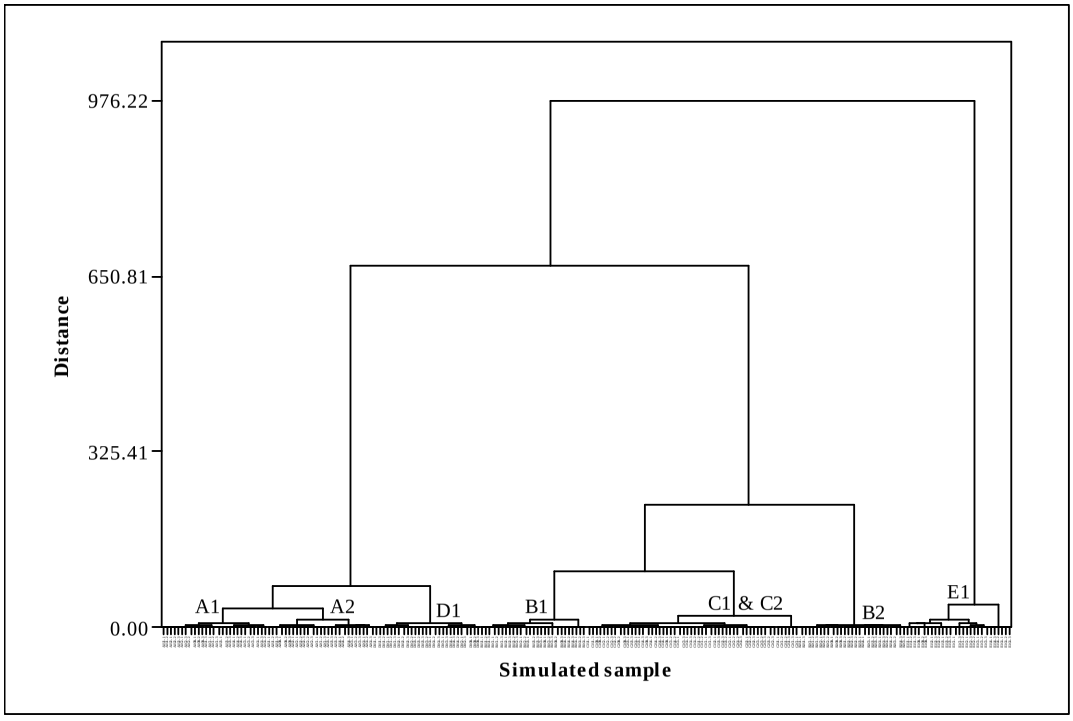
Table 5.23: Number of samples erroneously clustered and the d_m value in parenthesis obtained with 216 simulated samples analyzed by HCA

Method		Distance measure				
		Euclidean	**Manhattan**	**Pearson**	**Squared Euclidean**	**Squared Pearson**
Linkage method	**Average**	15 (9.96)	15 (7.41)	15 (9.96)	15 (1.81)	15 (1.81)
	Centroid	15 (10.22)	16 (7.07)	15 (10.22)	15 (1.82)	15 (1.82)
	Complete	27 (9.04)	18 (7.12)	27 (9.04)	27 (1.52)	27 (1.52)
	McQuitty	15 (9.91)	11 (7.14)	15 (9.91)	15 (1.84)	15 (1.84)
	Median	13 (10.14)	15 (6.93)	13 (10.14)	15 (1.89)	15 (1.89)
	Single	0 (13.00)	3 (9.10)	0 (13.00)	0 (3.13)	0 (3.13)
	Ward	0 (2.35)	0 (1.56)	0 (2.35)	12 (0.37)	12 (0.37)

The values of d_m shown in Table 5.23 indicate that the single linkage seems to give a higher d_m value with each distance measure as compared to other linkage methods and hence offering poorer separation. In contrast, the ward linkage is the best technique with its constantly low d_m values with all the distance measures. Of the seven linkage-distance combinations which fulfilled the first criterion, the d_m values demonstrate that Ward-Manhattan is the best method for clustering since it has the least d_m value. Figure 5.31 illustrates how the d_m value affects the clustering outcome. Ward linkage in Figure 5.31(b) shows densely packed linked samples within the individual groups separated by larger distances compared to the Single linkage in Figure 5.31(a) wherein the unlinked samples are closely packed within a close proximity. In the latter case, distinct groups will not be apparent if all the data points are very close despite different groups. Single-Squared Euclidean and Single-Squared Pearson likewise also achieved the same performance as Ward-Manhattan. However, these two combinations did not have low d_m values and sometimes inter-group distances were very close to show distinction or separation.



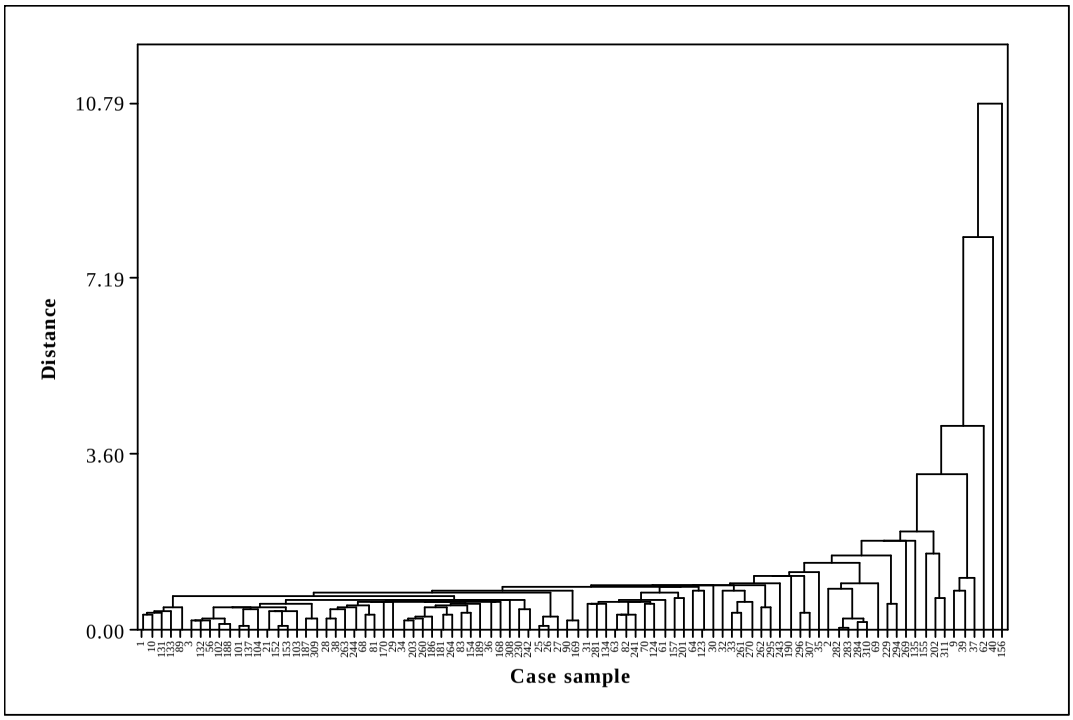
(a)



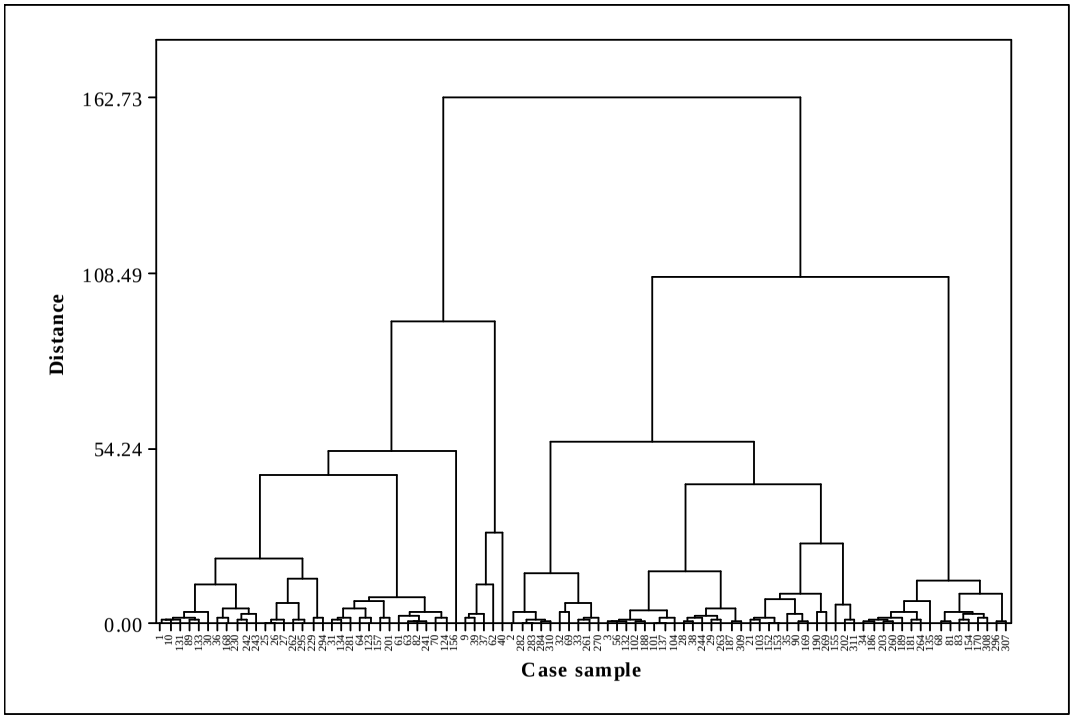
(b)

Figure 5.31: Dendograms expressed in distance showing the distances between clusters according to (a) Single linkage and Euclidean distance and (b) Ward linkage and Manhattan distance

The suitability of the seven combinations was further evaluated by 90 random case samples of unknown origins. Single-Euclidean (Figure 5.32(a)), Single-Pearson, Single-Squared Euclidean and Single-Squared Pearson demonstrate incremental relations for the case samples. As previously stated, distinct groups are not apparent because the distance relationships between the samples are very close. The statistical linking technique tends to weave all these 'intimate' units into a single large group. Hence, the clustering outcomes may be less meaningful if locating specific groups is the main objective. These combinations are more suitable for samples of known origins where only the relationships between the samples are investigated while their specific groups can be derived from the case histories. Ward-Euclidean, Ward-Manhattan (Figure 5.32(b)) and Ward-Pearson suggest several distinct groups on the dendogram. These techniques are useful when the origin of the samples is unknown and the number of clusters is of interest. As Ward-Manhattan demonstrated a good separation power in terms of the distance, it is decided to employ this technique as a preferred combination of linkage and distance measure for unsupervised pattern recognition for the real case samples.



(a)



(b)

Figure 5.32: Dendograms expressed in distance showing the distance relationships between 90 random case samples using (a) Single linkage and Euclidean distance and (b) Ward linkage and Manhattan distance

5.4.3.6 Summary

In this subtask, 8 distribution links of post-cut samples were simulated. Mixing the cutting agents with the pre-cut samples during the cutting process did not necessarily yield homogeneous post-cut samples. Compositional variability did occur within the sample irrespective of the cutting ratios and colors. However, this did not significantly influence the alkaloidal ratios which were used to establish links between the related samples. The combination of the Ward linkage and Manhattan distance was found to be suitable to cluster all the samples into their respective links using $N_{selected}$ + S pretreated GC data obtained with Option 1. This technique was finally chosen for the classification of real case samples via unsupervised pattern recognition.

5.4.4 GC-MS Method Optimization and Validation

GC-MS is inevitable in chemical analysis as it serves as a confirmatory test for the identification of the target compounds present in an unknown sample. Similar to GC-FID, GC-MS was also validated to ensure that it is suitable for qualitative analysis.

5.4.4.1 Optimization of GC-MS Parameters

As demonstrated in Section 5.4.1.1 Choice of Column and Selectivity, the HP-5 capillary column coupled with the Option 1 GC conditions is sufficient for the separation of the eight major components although the peak shapes are poor for quantitative analysis. However, as qualitative analysis is only concerned with the identities of the compounds, this column was chosen since it can offer a relatively shorter analysis time when the conditions of Option 1 were applied. Moreover, a different column (HP-5) used for peak identification could further assure the reliability of the results obtained with the HP Ultra 2 column used in quantitative analysis. Due to the use of a thinner film in this subtask, the GC conditions were also further optimized and the final conditions are summarized in Table 5.24.

After testing with several local heroin samples, it was found that components that were quantifiable by the FID at 1 µL injection volume were not detectable by the MS. Therefore, a higher volume of injection (3 µL) and a smaller split ratio (10:1) were required to ensure the small target peaks such as codeine and morphine are detectable in the MS. Figure 5.33 and Appendix 12 respectively show a typical total ion chromatogram (TIC) and their respective mass spectra for all the eight target components.

Table 5.24: GC-MS parameters for qualitative determination of eight target compounds

Condition	Setting
Column:	J&W HP-5 (5% phenyl 95% methyl siloxane)
Dimensions:	Length: 30 m i.d.: 250 µm Film thickness: 0.25 µm
Carrier gas:	Helium
Injection volume:	3 µL
Split ratio:	10 : 1
Flow rate:	1.0 mL/min
Injector temp.:	290 ^{o}C
Temp. programming:	240 ^{o}C hold for 1 min, ramp at 12 ^{o}C/min to 270 ^{o}C and hold for 8 min.
Transfer line:	290 ^{o}C
MS	Scan mode (40 – 450 m/z at 70eV)
Total run time	< 9 min

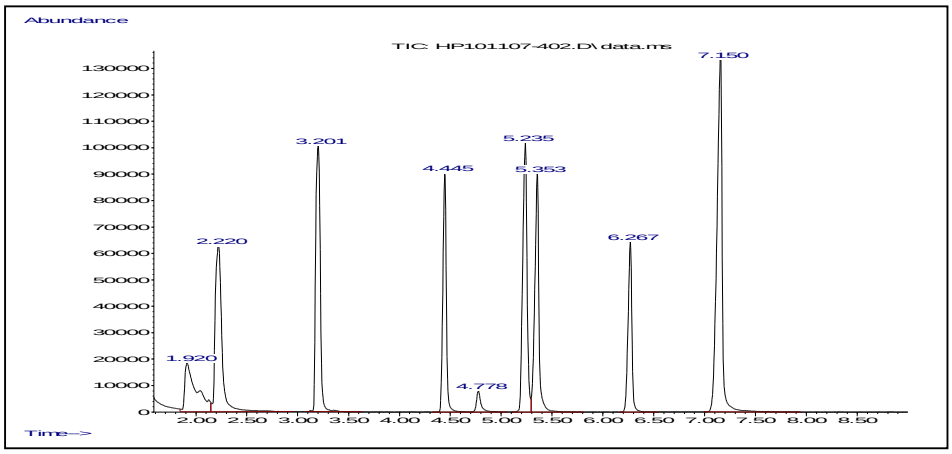


Figure 5.33: Reconstructed TIC of a mixture of standards at equal concentrations analyzed by GC-MS (The respective peaks are PC (retention time, RT = 1.92 min), CF (RT = 2.22 min), DM (RT = 3.20 min), CD (RT = 4.44 min), MP (RT = 4.77 min), AC (RT = 5.23 min), MM (RT = 5.35 min), HR (RT = 6.26 min) and IS (RT = 7.15 min))

5.4.4.2 Specificity and Precision of the Results

Specificity was examined using ten individual illicit heroin samples which were respectively spiked with the eight target analytes and IS. All the ten samples showed 100% positive results for the eight target analytes. The findings have fulfilled the

acceptable criterion of 'not more than one sample in five showing false negatives' as stipulated in the manual issued by UNODC (2009c).

5.4.4.3 Precision of Retention Time

GC-MS plays a crucial part in ascertaining the identity of a compound present in a sample. Apart from the confirmative results given in the form of mass spectra, retention time is another important measure to confirm the identity of the analyte. It is therefore vital to ensure the retention time (RT) for each target component is repeatable. As the GC column may sometimes be cut leading to a shorter RT, so a better parameter in terms of relative retention time or RRT (RT of the analyte relative to that of the IS) was employed to measure the precision of the RT.

The RSD was calculated for each RRT from ten repeated injections of a standard mixture. The RRTs of all the eight target analytes achieved an RSD < 0.2% (Table 5.25), indicating sufficient stability of the instrument. The precision of the RRT was further examined using ten individual spiked samples under the influence of sample matrix. Table 5.25 shows that all analytes maintained an RSD < 0.8% except for paracetamol (RSD = 2.49%). The slight shift in the RRT was due to the difference in the analyte concentration. However, the relatively poor RRT of paracetamol could be due to the interference of the sample matrix. The overall findings however show satisfactory precision for the RRTs of all the target analytes.

Table 5.25: RSD (%) of RRT obtained from a mixed standard and ten different spiked samples

	PC	**CF**	**DM**	**CD**	**MP**	**AC**	**MM**	**HR**
Standard mixture (n = 10)	0.19	0.12	0.06	0.03	0.04	0.02	0.03	0.01
Different samples (n = 10)	2.49	0.20	0.14	0.15	0.69	0.11	0.53	0.79

5.4.4.4 Limit of Detection (LOD)

The LOD of the instrument was determined from the lowest detected analyte concentration using the same formula as shown in Appendix 10. The three times signal-to-noise ratio for all the analytes were found to be at 2.5 – 7.9 µg/mL. When this range is compared with that obtained with the GC-FID (Table 5.26), the MS is found to be less sensitive than the FID. However, this range meets the requirement for the purpose of qualitative analysis.

Table 5.26: A comparison between the LODs obtained with GC-MS and GC-FID

Major component	LOD (mg/mL)	
	GC-MS	GC-FID
Paracetamol	0.0041	0.0021
Caffeine	0.0025	0.0006
Dextromethorphan	0.0032	0.0004
Codeine	0.0044	0.0005
Morphine	0.0079	0.0061
Acetylcodeine	0.0034	0.0008
6-monoacetylmorphine	0.0047	0.0019
Heroin	0.0032	0.0013

5.4.4.5 Summary

The GC-MS was validated and the technique was found to be sufficiently good for qualitative analysis. The instrument showed 100% accuracy in detecting the target analytes with repeatable RRTs. With this method, a sample can be analyzed in less than 8 – 9 min since a relatively thinner film of the HP-5 column promotes a faster elution time.

5.4.5 Analysis of Heroin Case Samples

With the analytically and statistically sound GC methods, the 311 case samples were analyzed both qualitatively and quantitatively using the optimized methods. The validated statistical procedure was also applied to these samples to estimate the relationships between the samples.

5.4.5.1 Qualitative Analysis by GC-MS

The eight major components present in the case samples were confirmed by the GC-MS and the results were congruent with those of the GC-FID. Additionally, other excipients were also simultaneously identified by the MS. From these street samples, three compounds were tentatively identified to have the characteristics (e.g. molecular weight, MW and fingerprint spectra) of N-phenyl acetamide, aminophylline and chloroquine (Figures 5.34 and 5.35). These compounds were not quantified because the laboratory lacked the chemical standards for further investigation. However, these compounds will be the future focus for heroin profiling.

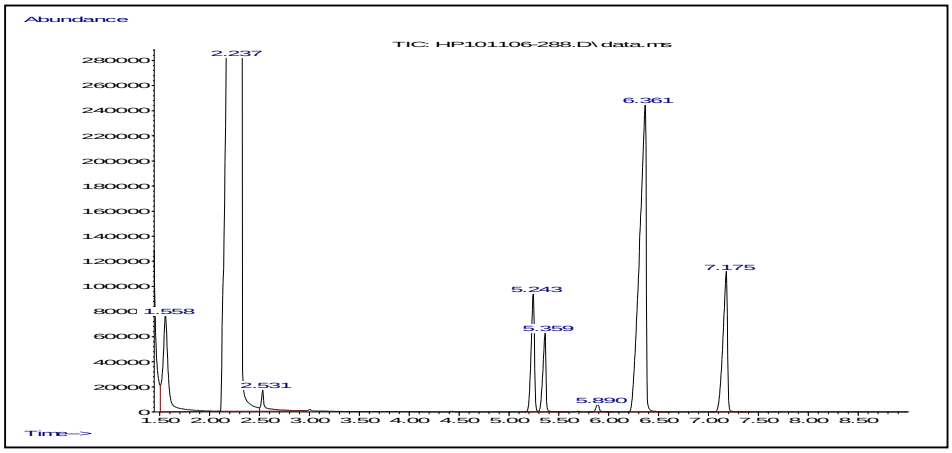
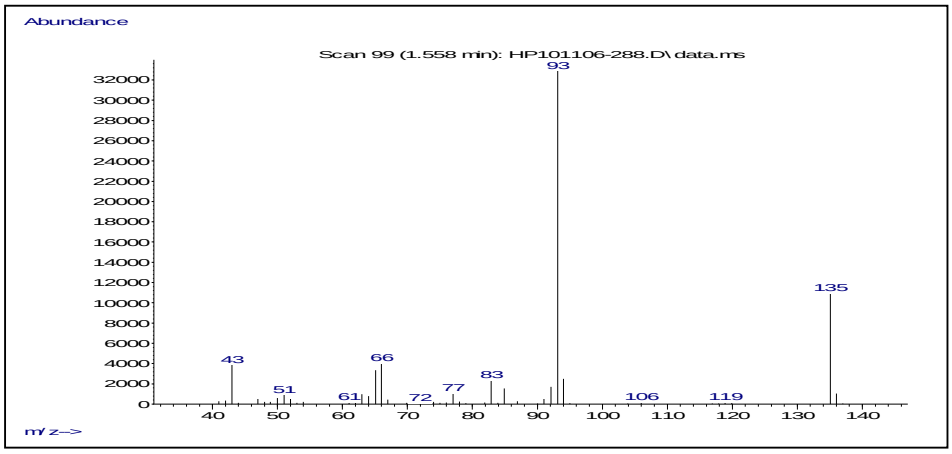
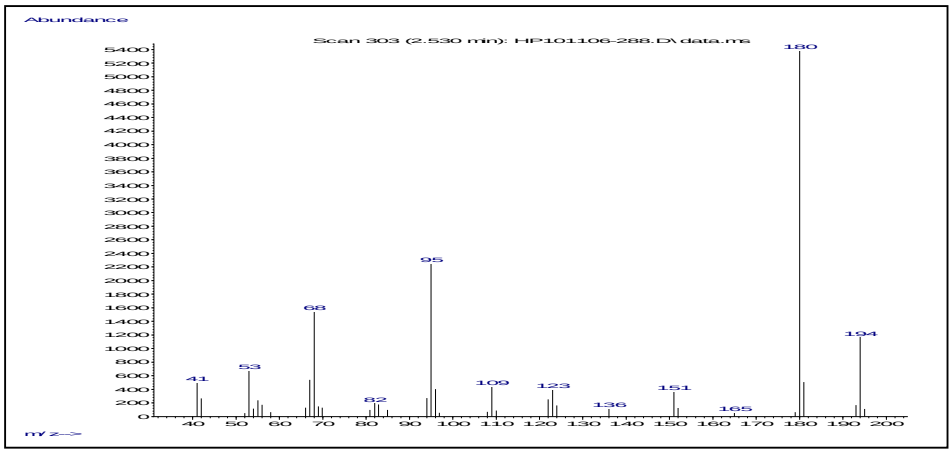


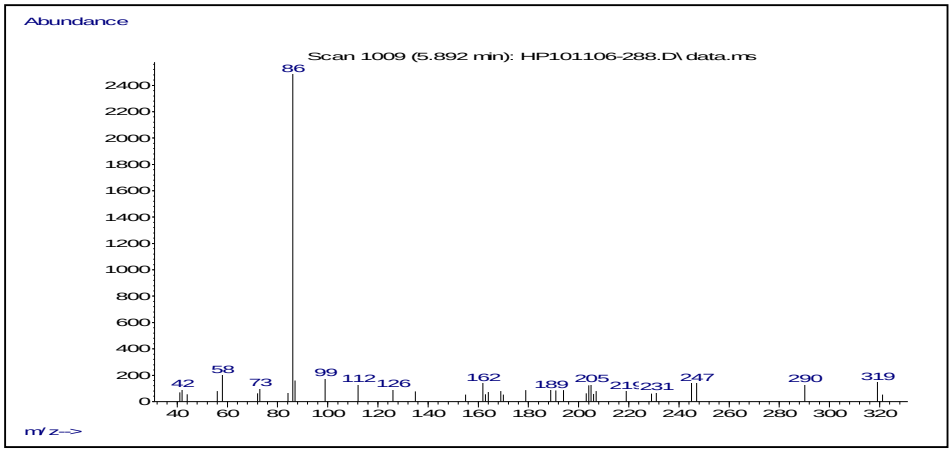
Figure 5.34: Reconstructed TIC of a case sample analyzed by GC-MS (The respective peaks are acetamide (N-phenyl) (RT = 1.55 min), caffeine (RT = 2.23 min), aminophylline (RT = 2.53 min), acetylcodeine (RT = 5.24 min), 6-monoacetylmorphine (RT = 5.35 min), chloroquine (RT = 5.89 min), heroin (RT = 6.36 min) and IS (RT = 7.17 min))



(a)



(b)



(c)

Figure 5.35: Three other major components tentatively identified by GC-MS (They are (a) acetamide (N-phenyl), MW =135; (b) aminophylline, MW = 180; (c) chloroquine, MW = 319. The retention time will shift depending on the concentration of the compound)

From the qualitative data, all the case samples were found to have been cut with caffeine. Its high prevalence was due to the fact that caffeine is easily obtained in bulk and therefore was used as the major diluent. Furthermore, as caffeine can cause heroin to evaporate at a lower temperature and thus its addition is suited for quick smoking and inhaling of heroin (UNODC, 2009, June 22). In some cases, paracetamol was chosen as another major cutting agent and thus decreasing the use of caffeine. Its mild analgesic properties and bitter taste may help mask heroin of poor quality (UNODC, 2009, June 22). Dextromethorphan was rarely used as it is gradually becoming a drug of abuse rather than an adulterant in Malaysia (however this assumption is only applicable to in-country cutting). In fact, the presence of chloroquine has been a signature of Malaysian seizures. This compound could be doubled as an anti-malarial agent since the sharing of needles is prevalent among drug addicts. In particular, this widely available compound is inexpensive and does not alter the effects of heroin or influence the way in which it is consumed (UNODC, 2009, June 22). Aminophylline could be a degradation product from caffeine after prolonged storage.

In most cases, the five opium-based alkaloids were concurrently present in the samples with varying peak intensities. The production process inevitably results in these opium-based alkaloid impurities in detectable amounts although they were highly cut. Two cases were detected to be fake heroin samples, in which the opium-based alkaloids were totally absent. The samples only contained caffeine and a combination of caffeine and paracetamol.

5.4.5.2 Quantitative analysis by GC-FID

During quantitative analysis, the stability of the instrument was confirmed by a control standard inserted between the samples. The GC data of a sequence of samples were deemed acceptable when the percent errors shown by the control sample placed

after each sequence did not exceed the maximum limits. The percent error recorded for the system was within ±10% for morphine and ±5% for the remaining seven compounds, indicative of sufficient stability of the system and hence the results are reliable. A larger percentage error was permitted for morphine due to its low solubility in the solvents chosen. In this study, the quantification of low amounts of codeine and morphine posed a problem. Increasing the sample weight did not necessarily increase the content level to the linearity range. Therefore, further increase in sample weight was not performed and the reported findings were based on the detection at the specified sample weight.

The percentages of the target analyte contents in the 311 illicit samples are presented as box-and-whisker plots in Figure 5.36 and the calculation included zero-values (absence). Table 5.27 on the other hand presents the statistics summarized from the case samples excluding the zero values. The boxplots are very useful to identify outliers and provide a good preliminary overview of the data. As paracetamol was a less common cutting agent, cases with high amounts of this compound displayed on the plot may indicate different batches at the street level. Similarly, low contents of caffeine imply this phenomenon. In the 18 paracetamol-caffeine cut samples, a relatively strong negative relationship between paracetamol and caffeine was obtained (coefficient of correlation, r^2 = -0.801). Hence, the increase in the use of one cutting agent usually leads to the decrease of the other. For cases involving dextromethorphan, the compound present was only in very small amounts. Another reason for its less usage is that high doses of this compound will cause acute toxicity (Chung, 2005). In addition, the low preference towards dextromethorphan could be related to the fact that this compound is now being monitored closely in Malaysia after the commercial dextromethorphan was widely abused. Hence for this reason it has greatly reduced the accessibility. For the opium-based alkaloids, namely codeine, morphine and acetylcodeine, they represent the

major impurities frequently inherited in the final products despite how carefully the processing was performed. As the samples were significantly cut with diluents, these impurities were therefore present in lower amounts as compared to heroin/diamorphine when they eventually reached the street.

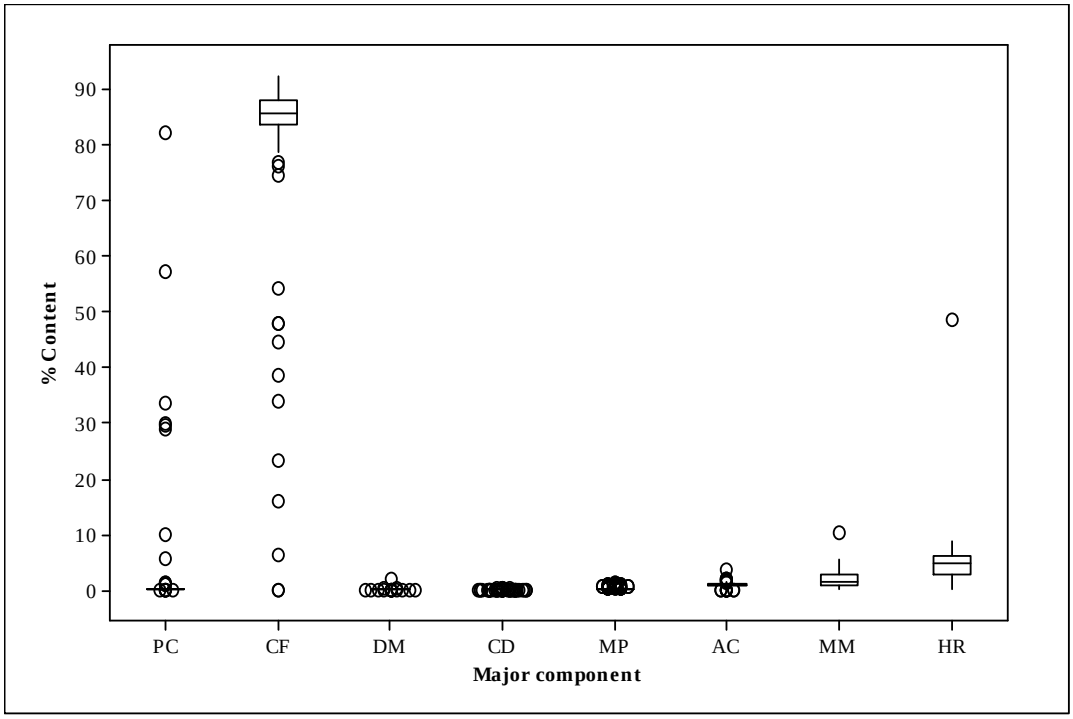


Figure 5.36: Boxplots showing the % analytes of eight target components in 311 illicit heroin case samples including zero values (absence) (The range of analyte contents also fall within the range created in the simulated samples in Figure 5.29)

Table 5.27: Statistical parameters for the % analytes of eight target components in the heroin case samples, excluding zero values (absence)

Major component	Frequency	Mean ± SD	Median	Range
Paracetamol	18 (6%)	15.78 ± 23.68	1.41	0.04 – 82.33
Caffeine	311 (100%)	83.82 ± 11.55	85.69	0.03 – 92.53
Dextromethorphan	16 (5%)	0.24 ± 0.53	0.03	0.02 – 2.13
Codeine	283 (91%)	0.07 ± 0.08	0.04	0.01 – 0.38
Morphine	290 (93%)	0.24 ± 0.30	0.11	0.01 – 1.37
Acetylcodeine	309 (99%)	0.86 ± 0.32	0.85	0.05 – 3.73
6-Monoacetylmorphine	309 (99%)	1.99 ± 1.34	1.57	0.29 – 10.47
Heroin	308 (99%)	4.57 ± 3.21	4.70	0.04 – 48.89

The relationships between these opium-based alkaloids are depicted by the coefficients of correlation, r^2 presented in Table 5.28. A strong positive relationship between codeine and morphine attests that both often co-exist and the presence of one may lead to the presence of the other. Very interesting relationships exist between heroin, codeine, morphine and acetylcodeine. When more heroin is produced through acetylation, other opium-based alkaloids such as codeine and morphine are also simultaneously acetylated. The decrease in codeine and morphine (or the negative r^2 values respectively at -0.214 and -0.224) that leads to the increase in acetylcodeine (r^2 = 0.655) is indicative of this phenomenon. Two plausible reasons are responsible for the negative correlations between morphine, 6-monoacetylmorphine (also 3-monoacetylmorphine, the quantified 6-monoacetylmorphine represents both the 6-monoacetylmorphine and 3-monoacetylmorphine) and heroin. First, morphine is acetylated to form 3-monoacetylmorphine and eventually to heroin during processing. Second, the unstable heroin is deacetylated to 6-monoacetylmorphine and sometimes hydrolyzed to morphine under improper storage conditions. As a result, negative relationships were evident between heroin and morphine as well as heroin and 6-monoacetylmorphine. Other explanations for the associative relationships between these opium-based alkaloids can be found in the studies by Small and Lutz (1932); Moore and Klein (1978); Huizer (1983a).

Table 5.28: Pearson correlation, r^2 of five opium-based alkaloids

	Codeine	**Morphine**	**Acetylcodeine**	**6-Monoacetylmorphine**
Morphine	0.921			
Acetylcodeine	0.054	-0.087		
6-Monoacetylmorphine	0.622	0.672	0.118	
Heroin	-0.214	-0.224	0.655	-0.126

5.4.5.3 Classification of Heroin Case Samples by Major Opium-based Alkaloids

It is more meaningful to express the opium-based alkaloids in ratios to study the relationships between the samples at the pre-cut level. Eleven parameters including AC/HR, AC/MM, AC/(MM+HR), AC/(MP+MM+HR), MM/HR, (CD+MP)/(MM+HR), HR/MM, (CD+AC)/(MP+MM+HR), (CD+MP+MM+HR)/AC, HR/(CD+MP+AC+MM) and (MP+MM+HR)/(CD+AC) determined in Section 5.4.3 Statistical Validation Using Simulated Heroin Links were used for unsupervised pattern recognition. All GC data in mg/mL were normalized according to the quotients followed by standardization (each quotient is divided by the standard deviation of the quotient variable) prior to PCA and HCA. For the HCA, the combination of Ward linkage and Manhattan distance was employed to construct a dendogram. Figures 5.37 and 5.38 show the relationships between 309 case samples (excluding the fake samples) seized from the four geographical locations of interest.

Two stretches of dense areas forming a 'V' shape were obtained in a 3-dimensional score plot (Figure 5.37). This is highly indicative of the heroin samples being processed from at least two sources. As the use of the opium-based alkaloids is useful to predict geographical origins, hence the street heroin samples considered in this study could be imported from at least two different origins. Atypical samples away from the 'V' shape including Samples 22, 40, 91, 135, 140 and 156 may be classified as exceptional cases.

On the other hand, the dendogram also indicates that many samples are closely related in several small clusters at the similarity level approaching 100%. These related samples could have shared similar sources. As the similarity level decreases to -158.48, five major groups are obtained and each contains about an equal number of case samples. Two final groups are obtained before they are joined at the similarity level = -675.43. These two clusters are assumed to be very dissimilar.

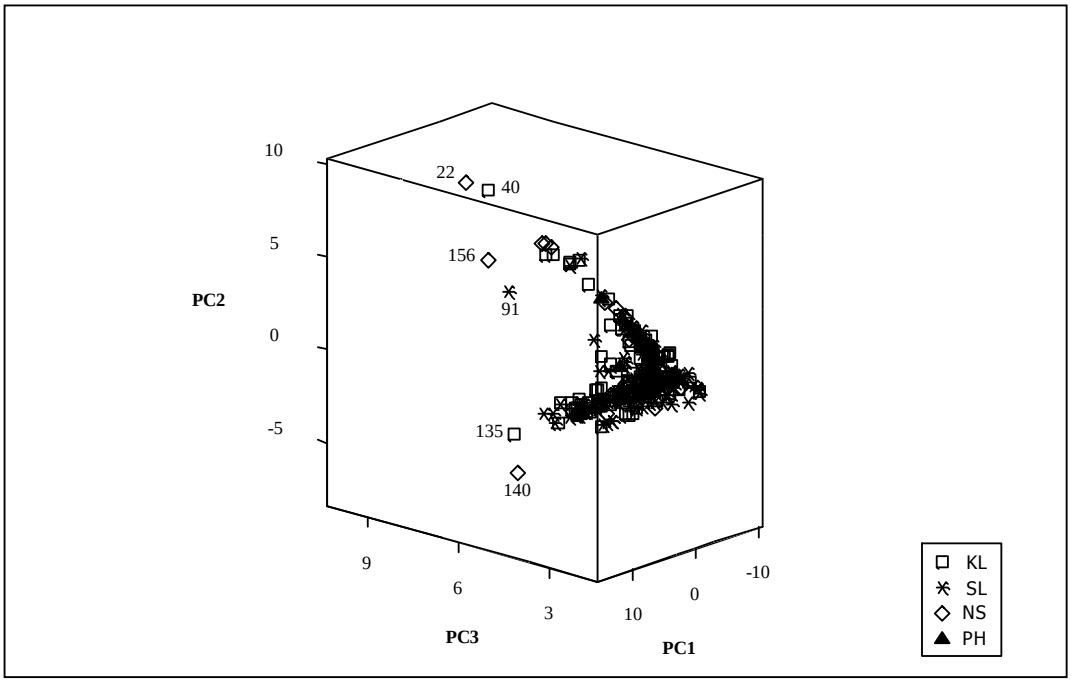


Figure 5.37: A score plot representing $N_{selected}$ + S parameters obtained from 309 cases analyzed by PCA in covariance mode, $\%V_1$ = 41.2 %, $\%V_2$ = 37.2% and $\%V_3$ = 12.1%

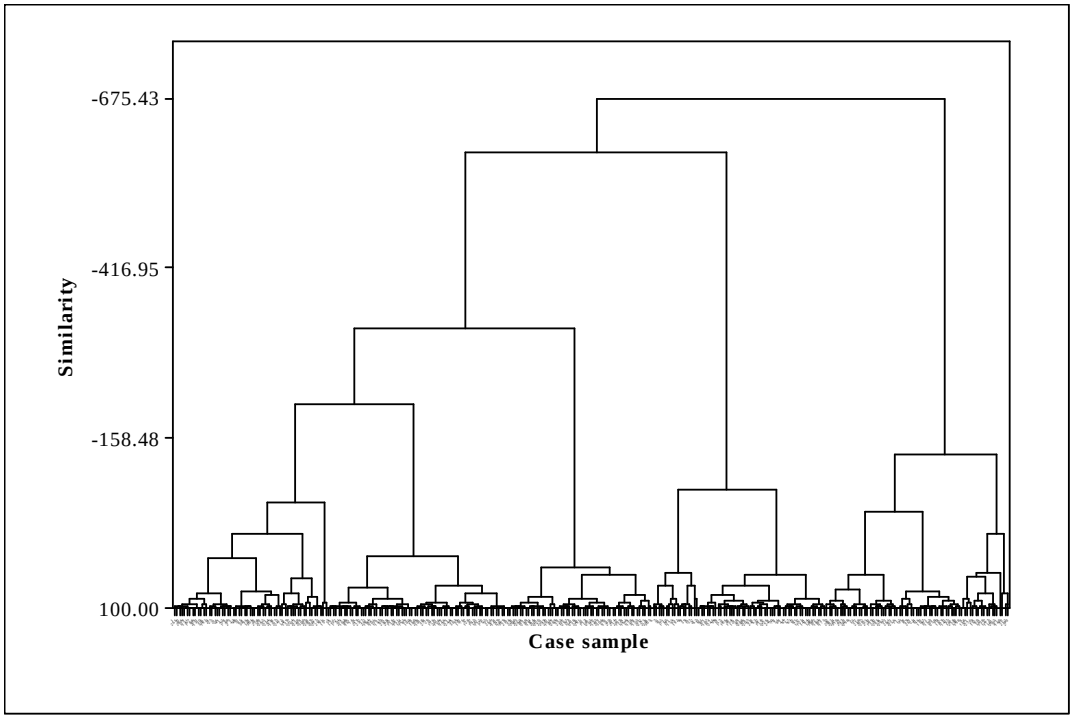


Figure 5.38: A dendogram expressed in similarity showing the relationships between 309 case samples using Ward linkage and Manhattan distance

Very often heroin shipments are adulterated and diluted before they leave the source country. Specifically, caffeine, paracetamol, dextromethorphan and chloroquine are adulterants added to heroin originating from Afghanistan. As a result, the presence of these compounds in Malaysian heroin may not fully represent a local in-country adulteration pattern; instead it could be indicative of the drug's actual origin.

5.4.5.4 Summary

Forensic characterization of the seized illicit heroin cases was carried out. The chemical data showed that the samples were largely cut with caffeine, yielding very low amounts of opium-based alkaloids. This however did not hamper the profiling work. By employing the opium-based alkaloids for sample classification, HCA (Ward-Manhattan) indicated that many samples were indeed closely related at high similarity levels. Large clusters were also obtained at very dissimilar levels. PCA on the other hand suggested two major stretches of clusters on the score plot. For unsupervised pattern recognition, it can be broadly concluded at least two major sources were responsible for the heroin case samples seized in Malaysia.

5.4.6 Additional Study: Development of a Novel Chemometric Procedure (Tetrahedrons) for Sample Classification

Developing a new chemometric procedure for sample classification opens up another opportunity to evaluate the relationships between samples. In this additional study, a tetrahedron model was developed to illustrate a graphical signature for the heroin sample and to define the statistical dissimilarity between the samples. Nine random case samples were used as a pilot batch. As only four parameters are permitted for a tetrahedron, CD/(MM+HR), AC/(MM+HR), CD/(MP+MM+HR) and AC/(MP+MM+HR) out of the six parameters used in Section 5.4.2 Evaluation of the Statistical Robustness of the Optimized GC-FID Method in Sample Classification by Pattern Analysis were chosen to develop this novel approach. Due to limited parameters allowed in this model, they were chosen so that 1) all the five opium-based alkaloids were included in the quotients and 2) the cutting and decomposition effects were considered.

This approach aims to provide a simplified chemometric procedure for real-time graph plotting. Hence, standardization that calculates the standard deviation from the whole dataset was not performed. In other words, one can construct a tetrahedron using the selected quotients generated from a single sample regardless of the characteristics of other samples.

5.4.6.1 Development of the Tetrahedron Model

i) Construction of a Tetrahedron

Construction of a tetrahedron can be done using the graphical construction feature ‘radar’ available in the Microsoft Excel software. A graphical working area (graphical plane) is defined by four extending arms marked ‘a’, ‘b’, ‘c’ and ‘d’ respectively. The four parameters of a heroin case sample are arbitrarily assigned to

these arms and four lines connecting the data points on the arms are constructed. A tetrahedron such as that in Figure 5.39 is obtained. This tetrahedron is the graphical signature for that heroin sample.

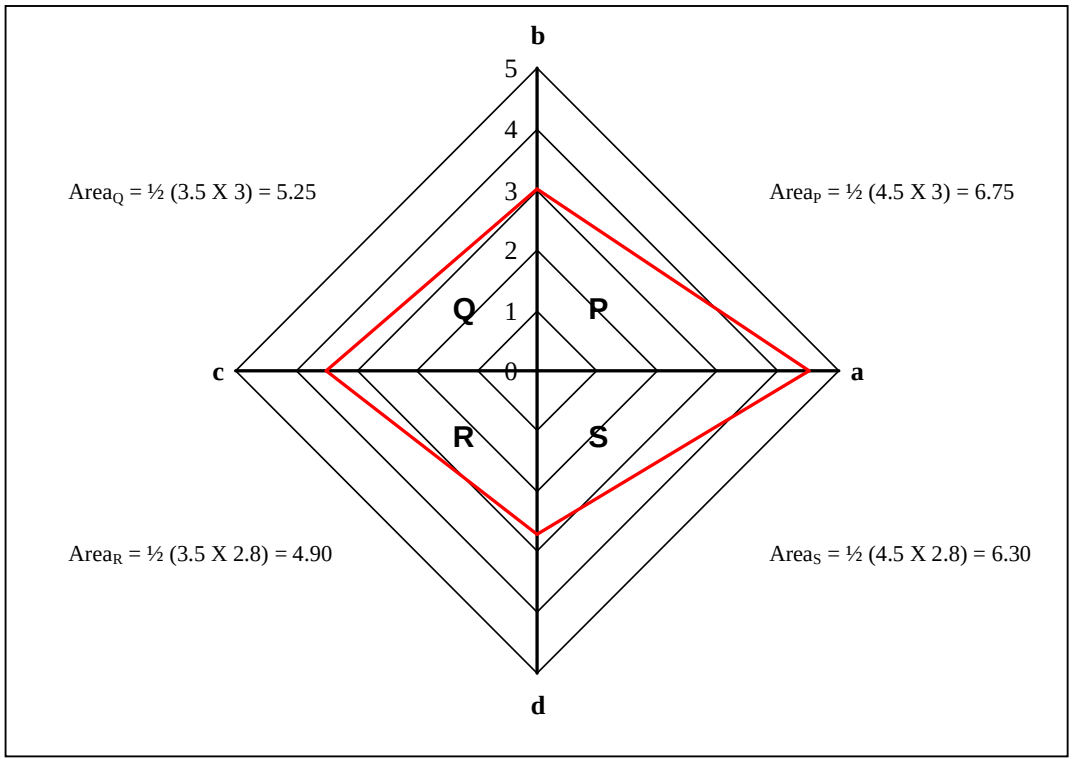


Figure 5.39: Construction of a tetrahedron. P, Q, R and S are the plane areas to be used for computation

Different assignments of the parameters to the arms will give different patterns. This concept is similar to the different clustering outcomes given off by HCA if different linkage methods and distance measures are used. The possible tetrahedral variants as illustrated in Figure 5.40 show that three basic configurations (BC) give rise to nine other configurations, forming a total of 12 possible tetrahedrons with different assignments of parameters A, B, C and D on the four arms. To minimize confusion, the analyst should focus on one particular type of tetrahedron for comparison throughout the whole process. Ideally, the suitable assignments should be validated using samples of known origins.

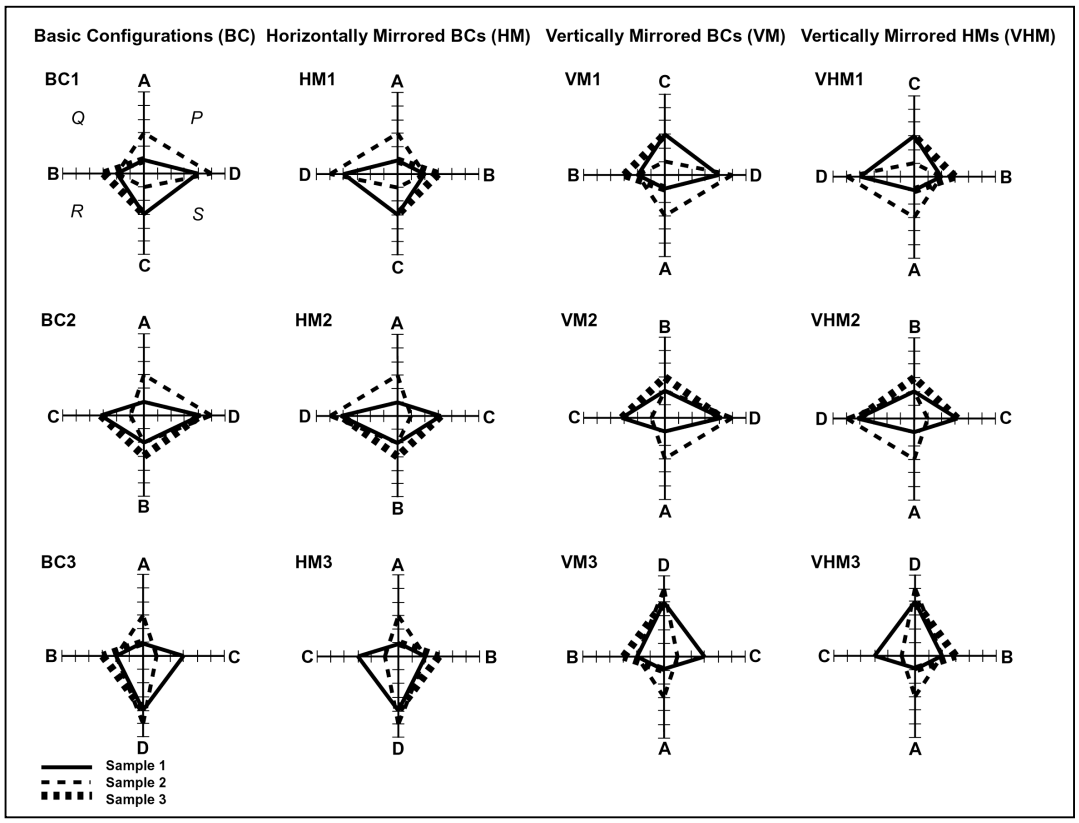


Figure 5.40: Possible tetrahedral patterns constructed from different assignments of parameters to the arms

In cases where further computations are not required, a graphical plane consisting of three-arm, five-arm and others can be employed to form their respective polyhedrons. By utilizing the arms repeatedly, the more complex polyhedron would therefore produce a more unique pattern. Figure 5.41 clearly explicates the complexities imparted by four types of polyhedron (the most complex being the octahedron).

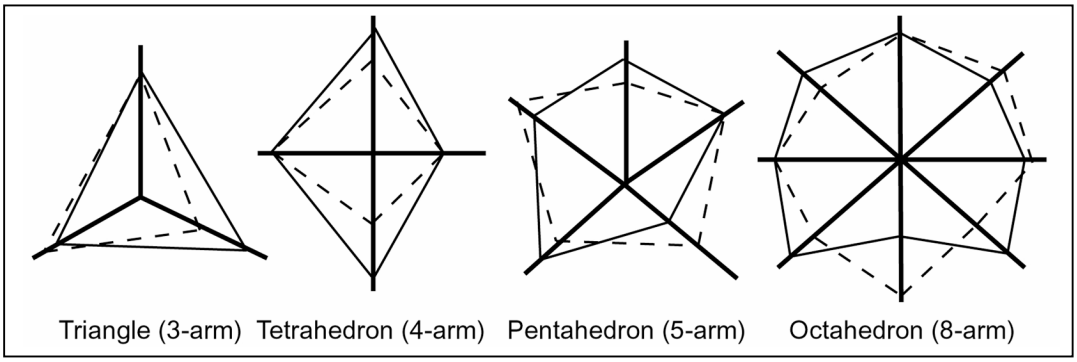


Figure 5.41: Complexities imparted by four types of polyhedrons (Black and dotted lines indicate two different signatures)

ii) Computation

Each area of the four separate planes P, Q, R and S that constitute a single tetrahedron can be computed using a simple mathematical formula for a triangle:

$$\text{Area, } A = \tfrac{1}{2} (arm_1 \text{ X } arm_2)$$

Examples of calculation are shown in Figure 5.39.

The computed areas of two tetrahedrons can be compared to determine the degree of dissimilarity. In this case, P, Q, R and S are treated as four individual variables in a matrix. The sum of differences expressed as positive values is finally used to determine the degree of dissimilarity in terms of area. The following equations show a simple bivariate analysis of two tetrahedral areas:

$$\text{Differences} = \begin{pmatrix} A_{P1} & A_{Q1} \\ A_{R1} & A_{S1} \end{pmatrix} - \begin{pmatrix} A_{P2} & A_{Q2} \\ A_{R2} & A_{S2} \end{pmatrix} = \begin{pmatrix} A_{P} & A_{Q} \\ A_{R} & A_{S} \end{pmatrix}$$

$$\text{Sum of differences} = \text{Degree of dissimilarity} = [A_P] + [A_Q] + [A_R] + [A_S]$$

whereby [ ] *is a positive value*

If the product of the degree of dissimilarity is low, it indicates that there is a strong association between two samples. In contrast, high values suggest very weak relationships.

As discussed previously, for a 4-parameter combination, it can be represented by 12 types of tetrahedrons. When two samples are constructed on the 12 different tetrahedron types, the sum of differences (degree of dissimilarity) calculated from these tetrahedron types will be different. As all the mirrored tetrahedrons in Figure 5.40 are derived from the three BCs, computing the sum of differences will only result in three sets of matrices out of the 12 types of tetrahedrons. For three samples represented tetrahedrally by A-B-C-D in unit (sample 1: 1-2-3-4; sample 2: 3-2-1-5; sample 3: 1-3-3-4) in Figure 5.40, their degrees of dissimilarity expressed by the three BCs are presented in Table 5.29.

Table 5.29: Area differences or the degrees of dissimilarity expressed by BC1, BC2 and BC3

	Plane P			Plane Q			Plane R			Plane S			Degree of dissimilarity		
	Sample 1	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**
							Basic configuration 1 (BC1)								
Sample 1	0.0			0.0			0.0			0.0			0		
Sample 2	2.0	0.0		2.0	0.0		3.5	0.0		5.5	0.0		13	0	
Sample 3	0.5	1.5	0.0	1.5	3.5	0.0	0.0	3.5	0.0	0.0	5.5	0.0	2	14	0
							Basic configuration 2 (BC2)								
Sample 1	0.0			0.0			0.0			0.0			0		
Sample 2	0.0	0.0		2.0	0.0		1.0	0.0		5.5	0.0		9	0	
Sample 3	0.0	0.0	0.0	1.5	3.5	0.0	2.0	1.0	0.0	0.0	5.5	0.0	4	10	0
							Basic configuration 3 (BC3)								
Sample 1	0.0			0.0			0.0			0.0			0		
Sample 2	2.0	0.0		1.0	0.0		3.5	0.0		0.0	0.0		7	0	
Sample 3	0.5	1.5	0.0	2.0	1.0	0.0	0.0	3.5	0.0	0.0	0.0	0.0	3	6	0

All the three matrices suggest that Samples 1 and 3 have the lowest degree of dissimilarity and hence are related. BC1 suggests that the pair of Samples 1 and 2 has a lower dissimilarity level than the pair of Samples 2 and 3. However, BC3 suggests the reverse. Therefore, the analyst should verify for the ideal tetrahedron type before the degree of dissimilarity in the real dataset is assessed. Again, this concept is similar to HCA in which different similarity indices are suggested if different linkage methods and distance measures are used for the same dataset.

It must also be ensured that no undesirable balance-off in the area difference by having perpendicular triangles in the same plane. Figure 5.42 clearly explains this phenomenon. Figures 5.42(a) and 5.42(b) show the ideal formation of triangles on the same plane, and the area difference can reliably represent the dissimilarity between the two triangles. Figures 5.42(c) and 5.42(d) however have two crossed triangles on the plane. Computing the area difference between the two crossed triangles is prone to errors. Having such triangles will show this balance-off, resulting in no area difference although they are indeed different.

Perpendicular triangles are observed the planes P and S respectively in BC2 and BC3 given in Table 5.29. These BCs can cause false negatives (false zeroes) and are not suitable for computation but for pattern analysis. Practically, such avoidance is uncontrollable especially when a very large dataset is involved. Certainly, this is similar to other well-established statistical tools, where falsities can only be estimated rather than deliberately controlled.

For other types of polyhedron where straightforward calculation of the area is not possible, special software or a more complex manual calculation involving trigonometry may be required. As a result, the tetrahedron is the rational model as it offers a moderately complex visual signature and eases further statistical calculation.

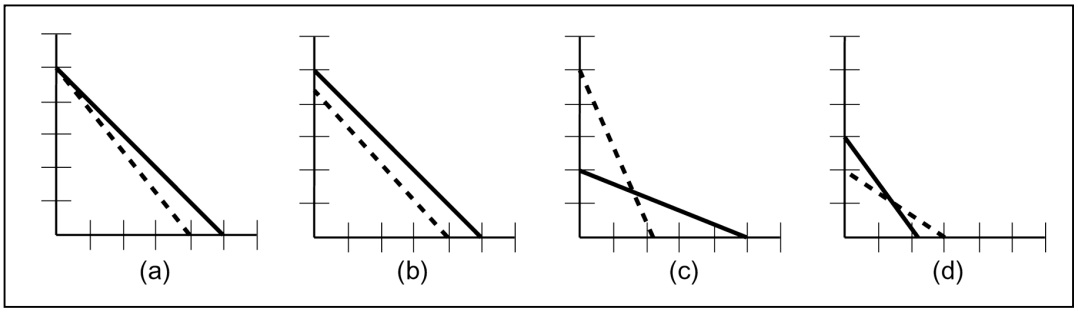


Figure 5.42: (a) and (b) show the ideal formation of triangles on the same plane for computation. (c) and (d) show false negatives arising from balance-off between two perpendicularly formed triangles on the same plane

iii) Advantages of the Tetrahedron Method

The central tenet of using graphical illustrations in profiling is based on a common saying that a picture is worth a thousand words. The tetrahedral representation of each heroin sample/batch/case offers the following advantages:

i. Statistically treated chemical data can be visually presented just as other forms of representations such as chromatograms and spectra.

ii. Interpretation becomes much easier. Differences between two case samples can be explicitly interpreted by superimposition. Multiple comparisons are possible with the aid of tetrahedrons.

iii. Degree of dissimilarity can be derived mathematically.

iv. The arms can be extended to allow for a larger data range for comparison using the same graphical plane.

v. Construction and computation are straightforward. Special statistical software is not required. Hence, minimal or no training is required.

5.4.6.2 Chemometric Procedure Using Tetrahedrons

The four pre-selected parameters derived from the chemical data of heroin were arbitrarily assigned to the four arms. Nine tetrahedrons representing 9 illicit heroin samples were constructed using the same graphical plane with the scales of arms

determined by the dataset. Each sample gave rise to a unique tetrahedron based on the four selected parameters presented in Table 5.30. To analyze the degree of dissimilarity of these samples, the 9 tetrahedrons are superimposed to produce a single profile (Figure 5.43).

Table 5.30: Parameters of 9 heroin samples

Case No.	CD/(MM+HR)	AC/(MM+HR)	AC/(MP+MM+HR)	CD/(MP+MM+HR)
8	0.053374	0.105294	0.082913	0.042029
9	0.051191	0.109190	0.087612	0.041075
19	0.031316	0.116982	0.103996	0.027840
61	0.038360	0.165881	0.150492	0.034801
63	0.034948	0.154094	0.137987	0.031295
82	0.032989	0.150574	0.136249	0.029851
88	0.041841	0.149852	0.130659	0.036481
98	0.035955	0.132186	0.118565	0.032250
106	0.035228	0.143477	0.128321	0.031506

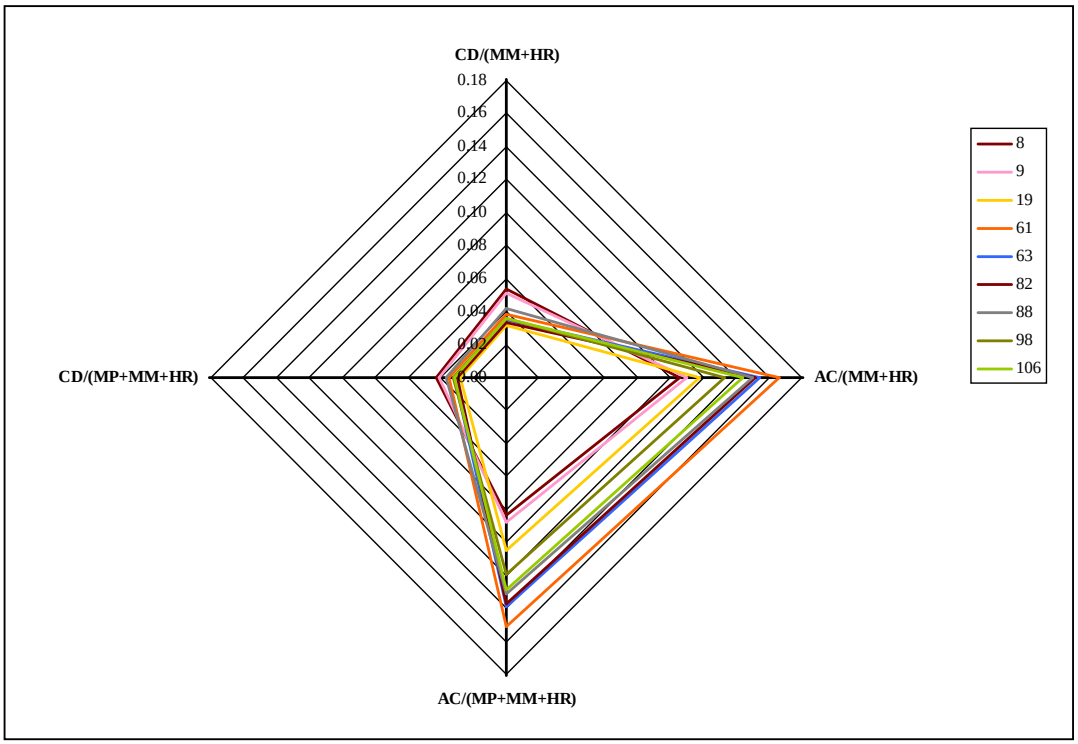


Figure 5.43: Tetrahedrons of 9 heroin samples

Based on their visual agreement in the shape of tetrahedron, Sample 61 is very much different from the rest. This also means that its chemical characteristics are very much different from other case samples. In addition, other tetrahedral signatures also suggest that the pairs of Samples 8-9, 63-82 and 88-106 are apparently related in terms of their chemical features.

5.4.6.3 Computation

Sometimes, timely decisions must be made to determine if the case samples show some degree of dissimilarity. This can be done immediately either manually or with the Excel spreadsheet. The four areas of each tetrahedron are first computed after which the differences in areas between the tetrahedrons are calculated. Such differences are summed up and compared against each other to determine the degree of dissimilarity. The degrees of dissimilarity of the 9 heroin samples calculated through this approach are summarized in Table 5.31.

Table 5.31: Differences in the total areas/degrees of dissimilarity between 9 samples

Sample	8	9	19	61	63	82	88	98
9	0.00056							
19	0.00368	0.00323						
61	0.00982	0.00929	0.00915					
63	0.00738	0.00681	0.00623	0.00292				
82	0.00714	0.00658	0.00547	0.00368	0.00076			
88	0.00675	0.00622	0.00627	0.00307	0.00172	0.00174		
98	0.00462	0.00406	0.00291	0.00625	0.00339	0.00274	0.00337	
106	0.00597	0.00541	0.00451	0.00464	0.00174	0.00117	0.00176	0.00165

According to Table 5.31, Samples 8 and 9 are the closest pair with its lowest degree of dissimilarity. Similarly, Samples 63 and 82 form another pair with the second lowest degree of dissimilarity. Then, Samples 82 and 106 are the third closest pair. On

the other hand, Sample 61 has distant relationships with Samples 8, 9 and 19 but it has some similarity to Samples 63 and 88.

The tetrahedron method is a valid model to simplify the chemometric concepts. When the same data were analyzed by PCA, the general interpretations given in the previous paragraph are in accord with those of the PCA (Figure 5.44). The estimation of sample relationships using this novel model can thus be applied to the real samples as an alternative chemometric procedure for drug classification.

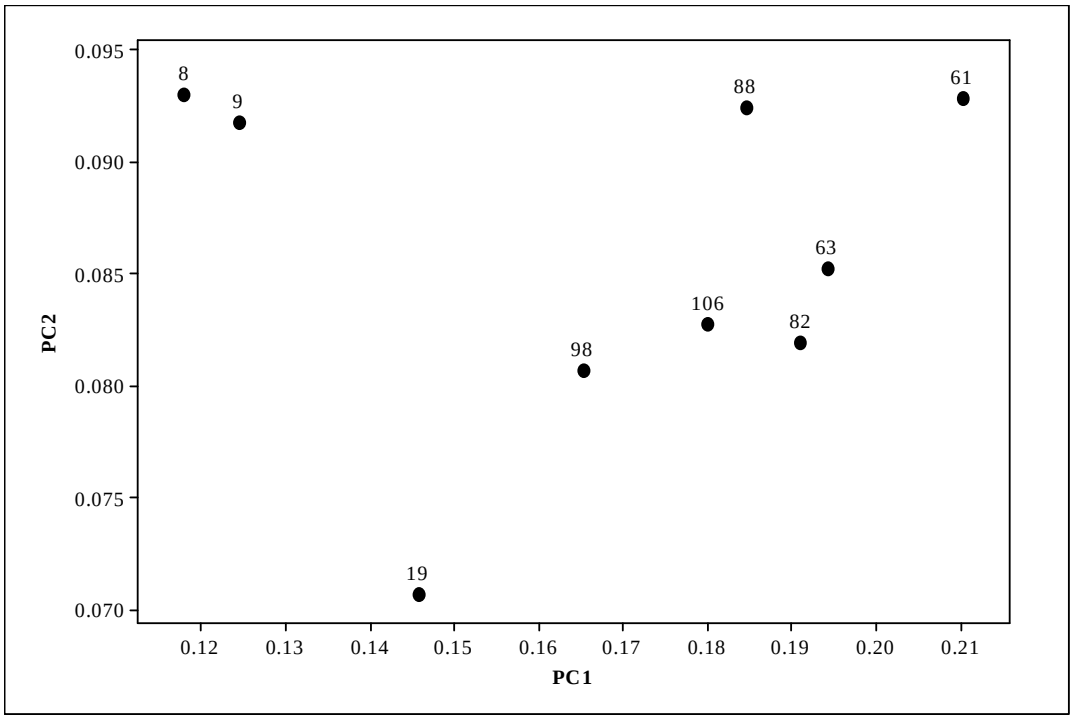


Figure 5.44: A score plot showing the relationships between 9 case samples decomposed by PCA, $\%V_1 = 94.70\%$, $\%V_2 = 5.20\%$

5.4.6.4 Summary

The tetrahedron model is able to give a moderately complex graphical representation for a case sample. With this novel approach, the identity of a seized heroin sample becomes very much straightforward and comprehensible when pretreated data are represented graphically. The degree of dissimilarity can also be derived directly

from the graph by visual comparison or through mathematical computation. Pilot testing was performed on this model. It was found that the sample relationships between 9 samples depicted from the tetrahedron model are similar to the relationships suggested by the PCA. Hence, this approach could be used as an alternative to simplify the chemometric procedure for gross estimation.