

Discerning Group-Type Analysis of Crude Oils from *Oso* Platform, Nigeria, Using Vibrational Spectroscopy with Multivariate Statistics

¹Etukudo, Ifiok U.*, ¹Okop, Imeh J. , ¹Obadimu, Clement O.
¹Dept. of Chemistry, Akwalbom State University, MkpateEnin, P.M.B.1167, Uyo, Nigeria.

-----ABSTRACT-----

Operationally, twenty samples of particle-rich crude oils from *Oso*-Platform were quantitatively fractionated into saturates, aromatics, resins, and asphaltenes (SARA) by asphaltene precipitation in *n*-hexane and high-performance liquid chromatography (HPLC). The newly developed and fully automated HPLC method has a sample capacity of 0.6g of crude oil. The spectroscopic probe technique captured infrared (IR). The crude oils have been characterized by vibrational spectroscopy in the infrared (IR) region. Principal component analysis (PCA) of the data sets from IR was performed so that broad spectrum and exploratory data analysis could be conducted. Partial least-squares (PLS) regression models were built for each SARA component from IR data to predict the amount of SARA components. These models successfully fitted the experimental data from IR analysis and indicated good predictive ability for the crude oil composition. The regression models from IR were not modeled properly for aromatics and asphaltenes but were properly analyzed and modeled excellently for saturates and resin components. For SARA determination, IR spectroscopy appears to be a favorable alternative to the more time-consuming fractionation method.

KEYWORDS: Particle-rich Oils, Vibrational spectroscopy, Principal Component Analysis, Partial least-square (PLS) Regression Models

Date of Submission: 30 November 2013

Date of Publication: 05 July 2014

I.

INTRODUCTION :

In most Nigerian refineries, the chemical characteristics of feedstocks (crude oil and other chemical-based additives) introduced into the fractionating column has changed tremendously, to an extent that the disparity can be measured or determined quantitatively by a decrease of several values (points) on the API (American Petroleum Institute) gravity scale. This can be attributed to improved process technology, and understanding of crude oil system will continue to necessitate these variations in line with recent research advances in molecular level of analysis, scientific and technical innovations. With growing demand of processing heavy crude oil, the knowledge of the constituents of higher-boiling feedstocks becomes paramount.¹ From technical and practical perspectives, the problems associated with processing the heavier feedstocks which is tantamount to the chemical character and the amount of complex in higher boiling constituents. Basically, heavy crude oil is a complex mixture of hydrocarbons (saturated-SHC, aromatic-AHC), containing a small fraction of heteroatoms (N, S, O) and resins. However, the classification of crude oil is in some aspect not distinct, but has normally been restricted to the more viscous part of conventional petroleum, having an API gravity of less than 20⁰. Due to its complex nature, it is impossible to determine its individual molecular constituents, and compositional studies are usually carried out by fractionation into predefined chemical families. The selective SARA group type fractionation process separates the crude oil into different groups: saturates (S), aromatics (A), resins (R), and asphaltenes (A). Actually, this fractionation technique has remarkable applicability in combination with high-performance liquid chromatography (HPLC). Since this technique is time-consuming and requires expensive laboratory equipment, several attempts have been made in recent times to model out alternative analytical techniques.

II. LITERATURE SURVEY:

An extensive study of the chemistry and structure of crude oils have provided a thorough understanding of components obtained from the analysis.^{1,11} The discerning SARA group type fractionation process separates the crude oil into different groups: saturates (S), aromatics (A), resins (R), and asphaltenes (A).^{1,6,11,12}

It has been established that, in order to minimize the effect of oxidation, 1000cm³(1dm³) of each of the crude oils was stored in a UV-protective containers (under nitrogen atmosphere) have been extensively used in analysis.^{8,12,13,14} Statistical method used [for example partial least –squares (PLS) and principal component analysis (PCA)], normally extract most important variables from the data set and ignore irrelevant information, and depending on statistical principles. Specifically, the “ principle of parsimony”¹⁵ stipulates that a simple model (example one with fewer variables) will tend to be better at predicting a new, anticipated and previously unseen data set.^{13,14,16,17}

III. RESEARCH ELABORATION(EXPERIMENTAL PROCEDURE)

Materials and Samples Collection: Crude oils samples from Osso-Platform at the temperature of the immediate environment (ambient temperature), were collected from exploration sites on the Bight of Bonny, South- South Nigeria (extending to the coastal and ocean region of Akwa Ibom State). Twenty (20) samples of crude oils were collected from the Atlantic ocean and the Bight of Bonny to form a sample set (Osso-01 to Osso-20) of oils differing in physical properties from *conventional* petroleum to heavy or heavy particle rich oil. A lot of water was found in some samples (between 45-50wt%), this was removed before analyses. The rest of the samples were preserved in their original state due to remarkably very stable water-in-oil systems. To minimize the effect of oxidation, 1000cm³(1dm³) of each of the crude oils was stored in a UV-protective containers (under nitrogen atmosphere).^{12,13,14}

A thorough introduction to the chemistry and structure of crude oils have been established.¹ Saturates are hydrocarbons ranging from straight-chain paraffins to cycloparaffins (naphthenes), aromatic fractions includes those hydrocarbons containing one or more aromatic ring (nuclei) that can be substituted with naphthenes or paraffins. Asphaltenes are the soluble group of crude oils that easily precipitate in the presence of aliphatic solvents (specifically n-hexane), while the resins component are the soluble fractions in light alkanes but insoluble in liquid propane. Basically, asphaltenes and resins are known as large, polar, poly-nuclear molecules consisting of condensed aromatic rings and heteroatoms (N,S,O), which include nitrogen, sulphur, and oxygen. Asphaltenes and resins have remarkably spurred interest with the knowledge of the significant effect of these components on the overall performance of heavy crude oils specifically.

Discerning .Group-Type Fractionation:

(a) Precipitation of Asphaltenes: In this aspect, the precipitation output (i.e. percentage yield), of asphaltenes from crude oil in liquid hydrocarbons have been found to be greatly influenced by both the degree of dilution and contact time. Precipitation was carried out in three (3) parallel trial steps, allowed for proper mixing overnight and filtered through a 0.45µm membrane filter. Asphaltene content was thereafter determined gravimetrically after solvent removal.

(b) High Performance Liquid Chromatography (HPLC) Technique: Fractionation of Maltenes. In carrying out maltene fractionation, the HPLC was characterized by helium degassing unit, sub-controller, preparative pump, automatic sample injector, two high-pressure flow channel selection valves, dual mode UV-visible spectrophotometer detector, refractive index detector, fraction collector, and system controller. The column was packed with unbounded silica (15µm, 21.2 x 250mm), amino (10µm, 21.2 x 50mm). Dichloromethane (98%) and n-hexane (96%) was used as the mobile phase. Initial preparative separations of organic samples favour the use of unmodified silica for packing the column.¹³ In most cases, experiments on several preparative silica column indicated that dichloromethane does not exhibit the required solvent strength to elute effectively the most polar components in crude oil. Instead, an amino column, weaker in column strength was used in the multidimensional column set up. Actually, the amino precolumn and the silica preparative column can be reloaded individually. This set up reduced the time of analysis from 72 to 28 minutes. A sample of 5ml of crude oil (with low asphaltene content) in n-hexane (equivalent to 0.6g of crude oil) was injected into the 20ml/min flow of filtered and degassed n-hexane. Crude oil components travel through the columns and were retained mainly depending on polarity. Saturates having no retention on the columns, were collected from refractive index (RI) signals (3-5 minutes). After 15 minutes, all aromatics have left then amino precolumn, but some yellow-coloured aromatics have not yet eluted from the silica column. These were collected from the preparative column by a dichloromethane backflush (<25 minutes), whereas the resin fraction was desorbed from the amino precolumn by a dichloromethane backflush, 24-28 minutes after sample injection. This procedure provides the chromatogram (in figure 2b), where the three fractions are well resolved. The SAR weight fractions were determined gravimetrically after controlled evaporation of solvent in N₂ atmosphere. In order to evaluate experimental errors associated with sample preparation and fractionation, the SARA procedure was repeated for five crude oils, five (5) weeks after the original measurements. Since standard deviations are not necessarily additive,

the experimental error cannot be estimated from the individual standard deviations of the five samples. However, the variance are additives, the average standard deviations, (SD_{ev}) was estimated by calculating the pooled average for all samples from the replicate variances.

$$SD_{ev} = \sqrt{\frac{1}{n(j-1)} \sum_{i=1}^n \sum_{j=1}^j (y_{ij} - \bar{y}_i)^2} \dots\dots\dots(2a)$$

Where j is the replicate number of sample i.

VIBRATIONAL SPECTROSCOPIC MEASUREMENTS

(a) IR Spectroscopy;- The Infrared(IR) spectra were recorded in the range between 600 and 4000 cm^{-1} with a tensor 27 FT-IR (Fourier Transform- Infrared) spectrometer. The instrument is fitted with a Michelson interferometer with a N₂-cooled Mercury Cadmium Telluride(MCT) detector. One drop of the essential oil was placed onto the diamond crystal(only 4mm²) and analyzed. To improve the signal- to -noise ratio, up to 206 scans were recorded at a resolution of 4 cm^{-1} and averaged. Thereafter, the resultant IR spectra were Fourier-Transformed with a Blackman- Harris three-term apodization function and a zero-filling factor of 2. In addition to the spectra of original crude oil samples, we also have access to spectra of individual SARA fractions(undiluted). Some of these will be applied to understand the individual contributions to the overall vibrational absorption of crude oil samples used in this experiment. With every facts obtained from the field, the data acquisitions and Fourier transformations were performed using the recent OPUS software.

(b) Chemometrics and Pre-Data Processing;- Due to much observed diversity of chemical substances both in a particular crude oil system and each of the predefined SARA classes, the resulting IR spectra normally show a large number of features that can be exploited to the most effect only with the help of useful chemometrics analytical tools. It became necessary to perform preprocessing of the data prior to multivariate statistics. Actually, a number of possible methods of data handling and pre-processing. Definitely, these transformations are suitable with specific problem justification. Observed sharp peaks characterize vibrational features in the IR. But due to the complex nature of crude oil system, with hundreds of different chemical structures, some specific overlap of adjacent bands was observed. In addition, minor base-line shifts was also prominent. It is pertinent to emphasize here that when performing multivariate statistics on a set of data from spectroscopic measurements, it is preferable to use all available measured variables. Statistical method used [for example partial least -squares(PLS) and principal component analysis(PCA)], normally extract most important variables from the data set and ignore irrelevant information, and depending on statistical principles. Specifically, the “principle of parsimony”¹⁵ stipulates that a simple model(example one with fewer variables) will tend to be better at predicting a new, anticipated and previously unseen data set.¹⁶

(c) Calibration and Prediction: Principal component analysis (PCA) was carried out. This has to do with a multivariate statistical method in which the information on spectral frequencies (the original variables are projected onto a smaller number of uncorrelated variables called principal components(PCs)).¹⁸ Plotting the principal components, the researcher can view the interrelationships between different variables, detect and interpret sample patterns, similarities or disparities. In addition, partial least-squares regressions(PLS),¹⁸ were performed on each of the SARA groups with predictive purposes in mind. Statistically, PLS is a method for relating the variations in one response variable (independent variable-y variable, in this case, mass percent of solubility group) to the variations of several predictors(dependent variable- x variable , which represent IR frequencies)

(d) Data Validation and Estimation of Error: A critical consideration of full cross- validation which represents an efficient way of using a set of few samples, the samples are used for both model calibration and validation. A sample is left out from the calibration set and predicted with the specific calibration model designed from the remaining samples. The process is repeated for each sample until all samples are treated. In term of root-mean-square error of prediction(RMSEP), this is the square root of the average of squared differences between predicted (\hat{y}_i) and measured (y_i) response values of the validation objects.¹⁸ The RMSEP is also interpreted as the average prediction error, expressed in the same unit as the original independent values(variables). Root-mean-square error of calibration(RMSEC) is therefore the corresponding measure suitable for the model, calculated from the calibration objects only.

Statistically, this can be depicted as follows;-

$$RMSEP = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \dots\dots\dots(2b)$$

The RMSEP represents a fair estimate of the average prediction error, it is therefore pertinent to compare this error with field data or experimental standard deviation. In this research, replicate spectroscopic measurement

were performed on five(5) crude oils for visual inspection of replicate variance compared to inter sample variance.

VI. RESULTS AND DISCUSSIONS

Determination of crude oil composition; The result of separation of crude oil into SARA and water fractions are presented in Table 1. Selected crude oil samples(Osso-04, Osso-06, Osso-20) are high in water content. However, water determination of the filtrates from asphaltenes precipitation showed that the water content was reduced to less than 0.02 wt % (of the crude oil). Water is mainly removed during filtration, also supported by the observation of water droplets on the filter paper. Asphaltene contents are reported with mean values and standard deviations of three parallel filtrations. From observation, the insufficient sample recovery from the lighter crude oils(Characterized by low resin content) can be attributed to evaporation losses during solvent removal. Gas Chromatography(GC) analyses carried out indicated that losses of compounds in the carbon number range up to 13 were found primarily from the saturate fraction and sparingly from the aromatic fraction. Saturate(SHC) and Aromatic(AHC) mass fractions were thereafter normalized to an overall yield of 100%. The actual cut points between HPLC fractions were selected from detector signals and validated through systematic IR analyses of all SARA fractions in the samples(especially non diluted fractions). The fingerprint region of crude oil(Osso-01) fractions in figure 4 affirms our specific definition of the maltenes, as the saturates fraction is free of aromatics(AHC) and aromatic fractions is free of hetero-compounds(polar), i.e. N, S, or O. Hence, similar conclusions were drawn from spectra of the 19 remaining crude oils (Osso-02 to Osso-20), in the sample set. Based on five(5) replicates (Osso-05', Osso-08', Osso-11', Osso-15' and Osso-18'), the average experimental errors were reported as average standard deviations (\overline{SD}_{av}). Average errors for saturates and aromatics were found to be within reasonable limits. Determination of asphaltene content was associated with an average standard deviation of 0.18 wt%, which was seen as reasonable.

IR Spectroscopy

(a) Spectral Characteristics-The spectral characteristics presented the most useful information from IR spectra of crude oils in the regions of $3250-2650\text{cm}^{-1}$ and $1850-650\text{cm}^{-1}$. Spectral information within these regions was used to design regression models. The fingerprints area of crude oil Osso-01 are shown in figure 3. To really trace the individual contributions from each SARA components, IR spectra of components. General observation indicated that for the saturates fractions, CH_3 group showed characteristic C-H stretching bands at 2960 and 2870cm^{-1} , whereas CH_2 groups showed bands at 2920 and 2850cm^{-1} . The bending vibration of methyl appeared at 1380 and 1450cm^{-1} , while the vibration of CH_2 appeared at 1470cm^{-1} .¹⁶ These bands were all observed for saturate fraction of crude oil, Osso-01 in figure 3. The band resulting from the CH_2 rocking vibration, in which all CH_2 groups normally rock in phase, appeared near 720cm^{-1} for straight-chain alkanes of six(6) or more carbon atoms and at some higher frequencies for shorter hydrocarbon chains.¹⁶ The C-H stretching bands of aromatic appeared in the $3150-3050\text{cm}^{-1}$ region, whereas the in-plane and out-plane bending bands appeared in the $1300-1000\text{cm}^{-1}$ and $900-600\text{cm}^{-1}$ regions, respectively.^{1, 17} Of particular interest is the aromatic fraction in figure 3, this exhibited the group of peaks at $850, 820,$ and 740cm^{-1} and this can be attributed to the possibility of substituted aromatic ring. As observed, these latter bands are most prominent and most informative for aromatics. The aromatic fraction of crude Osso-01 showed the typical absorption at 1600cm^{-1} and carbon-carbon(C-C) stretching within the ring. Considering the aromatic fraction of crude oil Osso-01, the aromatic bands at $900-600\text{cm}^{-1}$ and 1600cm^{-1} are likely for resins and asphaltenes, and the broad and strong absorption between 1330 and 1100cm^{-1} has been assigned to sulphur(S)-oxygen absorption.¹ From figure 4, the aromaticity of the asphaltenes is predominantly more than that of the maltenes fraction. Combined resin and asphaltene fractions indicated absorption between $1000-1050\text{cm}^{-1}$, which is typical for aromatic ethers and sulphoxide functionalities. Typical resins fractions containing acids and phenols are likely to be featured by O-H and C=O stretching vibrations at $3700-3100\text{cm}^{-1}$ and $1710-1680\text{cm}^{-1}$ respectively. The C=O stretching vibration in aliphatic acids(naphthenic acids) appeared at 1705cm^{-1} , as for crude oil Osso-01 resins. To effectively explore the potential of quantitative determination and regression from IR spectroscopy, multivariate statistics, principal component analysis of the first-order-differentiated set of crude oil spectra was carried out.

Multivariate Statistics: (a) Principal Component Analysis; The score plot in first and second principal components are presented in figure 4. The crude oils with high water content are presented as extreme values. These group of samples can also be included in the PLS regression model, but there is every likelihood of resulting in problems as the crude oil components are expected to be modeled with the same frequencies as those frequencies where water exhibits strong absorption ($3750-3850\text{cm}^{-1}$, $1800-1550\text{cm}^{-1}$). Actually, samples Osso-04, Osso-06, and Osso-20 were removed from the calibration set. Visual inspection of the score plot indicated that the reproducibility of the IR analysis is very significant, as replicate variations are relatively small compared to inter sample variations. The expressed variance in the first two PCs(principal component) is remarkably very high(95%).

Partial Least-Square(PLS) Regression: Despite the fact that the distinctive bands of individual SARA fractions have been used for qualitative structural analysis, their overall additive contributions prevent the use of specific bands for accurate quantitative determination of aromatics and asphaltenes from crude oil data set. Aromatics and asphaltenes are less characterized by similar vibrational features, which may result in problematic regression. The aromatics were therefore modeled properly as long as the asphaltene content is low. Here, the contribution from asphaltenes to the total aromatic absorption is relatively small. Besides, for asphaltene-rich crude oils strong aromatic vibrational absorption, the contribution becomes significant thereby affecting the model. In addition, IR analysis of separate asphaltene fractions indicated that the vibrational characteristics of the 20 asphaltenes are not similar. The inter-sample variation complicated the regression of asphaltenes from IR spectra. The aromatic structure of resins notwithstanding, these were still modeled adequately because of the typical acidic band of 1700cm^{-1} . The spectral ranged from 1500 to 1800cm^{-1} and was used for model design. Actually, the correlation of predicted to measured amounts of resins was plotted in figure 5. Definitely, the application of the regression model from the IR spectra provided a better predictive potential (RMSEP= 1.36%, three PCs). It also required lesser principal components to really accommodate the spectral characteristics of data set. Quality parameters are presented in table 2. In essence, the saturates were modeled by the response from only 6(six) wave numbers(variables) from typical carbon-hydrogen (C-H) stretching and bending vibrations as follows; $2800-2950\text{cm}^{-1}$ and $1380-1450\text{cm}^{-1}$ regions. This indicated that IR spectra of crude oil contain remarkable information about saturates concentrated in a lesser number of wave numbers. Considerably, three principal components captured 97.4% of the total variance in the six wave numbers and explains 96.2% of the total variance in saturate content within the specified data set. It can vehemently be stated that both saturate and resin PLS models separate sample Osso-03 from the regression as an outlier. Reproduced information of the IR analysis and SARA fraction indicated that data from this sample are likely to be affected by experimental errors.

Variations in percentage composition of SARA Components:

Figures 6,7,8 and 9 depict the variations of SARA components in terms of percentage composition. Graphically, The saturates are predominant in figures 6, 7 and 8 while the resins are of low values in the sample set(figures 8 and 9). In principles, the samples contain remarkable and appreciable percentage compositions in saturates and aromatics than the resins and asphaltenes in the respective crude oil sample that was analysed.

V. CONCLUSIONS

In laconic terms, model development from IR data resulted in inadequate model suitability for the aromatic(AHC) and asphaltene contents. The anticipated and expected diversity of asphaltenes within the sample set obtained was indicated by IR analyses of pure asphaltenes. Since aromatics were modeled by almost the same spectral frequencies as asphaltenes. Definitely, saturates and resins were predicted perfectly from IR data, with RMSEP_s of 1.82 and 1.30 wt% respectively. The saturates content was predicted with the absorption using six spectral frequencies, all of them in the regions represented typical saturate frequencies. Hence three principal components really captured 97.4 and 96.2% of the total variance in saturate content with the selected data set.

VI. ACKNOWLEDGEMENT:

The researcher is thankful to Kyushu University-Japan for providing the facilities for the analyses.

REFERENCES

- [1] Ali, M. A. and Nofal, W.A. (1994) "Application of High-Performance Liquid Chromatography for Hydrocarbon Group Type Analysis of Crude Oils", Fuel Science and Technology International Journal, Volume 12, Pages 21-24.
- [2] Aske, N., Kallevik, H. and Sjöblom, J. (2001) "Determination of Saturates, Aromatic, Resin, and Asphaltene (SARA) Components in Crude Oils by means of Infrared and Near -Infrared Spectroscopy". Energy Fuels, Volume 15, Pages 1304-1306.
- [3] Boukir, A. (2001), "Subfractionation, Characterization and Photo-oxidation of Crude Oil Resins", Chemosphere, Volume 43, pages 279-282.
- [4] Chung, H. and Ku, M. S.(2000) "Comparison of Near-Infrared, Infrared and Raman Spectroscopy for Analysis of Heavy Petroleum Products", Applied Spectroscopy, volume 54, pages 239-241.
- [5] Esbesen, K. H. (2001) "Multivariate Data Analyses- In Practice", fifth edition, Aalborg University Press, Esbjerg, Denmark.
- [6] Fan, T. G. and Buckley, J. S. (2002), "Rapid and Accurate SARA Analyses of Medium Gravity Crude oils", Energy and Fuels, Volume 16, pages 1571-1573
- [7] Geladini, P. and Kowalski, R. B.(1986), "Partial Least Squares Regression: A Tutorial", Anal. Chem., Volume 58, Pages 1-3
- [8] Hammani, A, Ferworn, K. A. Nighsander, J.A.,Overa, S. and Strange, E.(1998),"Asphaltene Crude Oil Characterization: An Experimental Investigation of The Effects of Resins on Stability of Asphaltenes" Petroleum Science and Technology, Volume 16, pages 227-231.
- [9] Mullins, O. C (1990) "Asphaltene in Crude Oil: Absorbers and/or Scatters in the Near-Infrared Region", Anal. Chem., Volume 58, Pages 1-3.

- [10] Radke, M., Willsch, H; Welte, D. H (1980),” Preparative Hydrocarbon Group Type Determination by Automated Medium- Pressure Liquid Chromatography” Anal. Chem., Volume 52, Pages 406-408.
- [11] Speight, J. G.(1998), “ The Chemistry and Technology of Petroleum”, third edition, Marcel Dekker Publishers, New York.
- [12] Suotoni, J. C. and Swab, R. E(1976) “ Preparative Hydrocarbon Compound Type Analysis by High- Performance Liquid Chromatography”. Journal of Chromatographic Science, Volume 14, Pages 535-536.
- [13] Snyder, L. R; Kirkland, J. J.; Glajch, J. L. (1997) “ Practical HPLC Method Development”, second edition, John Wiley and Sons, New York.
- [14] Seasholtz, M.B. and Kowalski, B. (1993) “ The Parsimony Principle Applied to Multivariate Calibration”. Anal.Chim.Acta, Pages 227, 165.
- [15] Shaw, A. D.; Winson, M. K; Woodward A. M.; McGovern, A.C.; Davey, H.M.; Kaderbhai, N.; Broadhurst, D.; Gilbert, R.J.; Taylor, J.; Timmins, E. M.; Goodacre, R.; Kell, D.B.; Alsberg, B.K. and Rowland, J.J. (2000) “ Rapid Analysis of High-Dimensional Bioprocesses Using Multivariate Spectroscopies and Advanced Chemometrics”, Adv. BioChem. Eng/Biotech. Volume 66.
- [16] Silverstien, R. M.(1998)” Spectroscopic Identification of Organic Compounds”, sixth edition, John Wiley and Sons, New York.
- [17] Saltry, M. I. S.; Chopra, A.; Sarpal, A. S.; Jain, S. K.; Srivastava, S. P.; Bhatnagar, A. K.(1998) “ Determination of Physicochemical Properties and Carbon- Type analysis of base oils using mid- IR Spectroscopy and Partial Least- Square Regression Analysis”, Energy and Fuels, Volume 12, Pages 304-306.
- [18] Wold, S.; Esbensen, K.; Geladi, P. (1987) “ Principal Component Analysis,” Chemometrics Intellectual Laboratory Systems, Volume 2, pages 37-40.

Table 1: Experimental Compositions of *Oss* Crude Oils Separated into SARA and Water Fractions.

Crude oil origin & no.	Saturates (wt%)	Aromatics (wt%)	Resins (wt%)	Asphaltenes(wt%)		Water (Wt%)	Yield (wt%)
				mean ^a	SD _{av} ^a		

Calibration Set

Oss-01	32.5	45.5	17.6	1.3	0.2	0.6	98.8
Oss-02	36.5	34.6	13.6	1.7	0.3	0.8	90.9
Oss-03	31.5	42.2	21.9	4.5	0.2	0.2	99.9
Oss-04	36.0	30.8	22.1	3.1	0.1	9.6	99.5
Oss-05	43.0	29.7	10.4	1.2	0.1	0.3	84.5
Oss-06	26.5	38.8	20.5	2.8	0.1	12.6	98.3
Oss-07	46.9	37.9	14.0	1.7	0.0	0.1	99.6
Oss-08	33.6	38.1	16.8	12.9	0.2	0.1	99.7
Oss-09	45.0	32.9	9.8	1.2	0.1	0.2	89.2
Oss-10	45.7	38.6	11.5	0.8	0.1	0.1	96.6
Oss-11	43.6	32.9	7.2	0.6	0.0	0.4	84.8
Oss-12	33.0	42.8	11.6	3.9	0.0	0.5	91.8
Oss-13	32.0	32.5	28.5	4.3	0.2	1.2	98.6
Oss-14	53.1	30.9	8.2	1.0	0.0	0.2	93.3
Oss-15	33.4	38.6	18.2	4.6	0.2	0.1	94.8
Oss-16	37.3	42.6	14.1	3.8	0.4	0.3	98.1
Oss-17	41.5	33.3	7.0	0.2	0.0	0.1	82.0
Oss-18	26.0	41.1	21.9	10.2	0.3	0.1	99.4
Oss-19	44.3	26.3	8.1	0.2	0.0	0.5	79.4
Oss-20	22.1	46.5	11.5	5.5	0.2	4.3	92.7

Replicates

Oss-05 [†]	42.3	28.5	8.5	1.3	0.1	0.3	83.5
Oss-08 [†]	35.0	36.4	16.2	12.4	0.3	0.1	99.5
Oss-11 [†]	43.4	30.2	7.9	0.7	0.1	0.4	81.5
Oss-15 [†]	34.8	41.2	18.9	4.6	0.2	0.1	99.6
Oss-18 [†]	26.1	40.7	21.9	10.6	0.3	0.1	99.3

Average Experimental Error^bSD_{av}

1.2	0.8	0.8	0.2	0.02-0.6 ^c
-----	-----	-----	-----	-----------------------

^aAsphaltene mean values and standard deviations are estimated from three parallel precipitations.

^bAverage standard deviations were computed from the variance of replicates and the corresponding samples in the calibration set (eq 1).^c Heteroscedastic error distribution.

Table 3: Quality Parameters for the Prediction of Crude Oil Composition Using FT-IR Spectroscopy in combination with PLS Regression.

Group type	Prediction range(wt%)	PCs ^a	x-expl (%) ^b	y-expl (%) ^b	RMSEC (wt %) ^c	RMSEP (wt %) ^c	Sample outliers	Spectral Range (cm ⁻¹)
Saturates	26.2-57.3	3	97.4	96.2	1.29	1.82	3	1382-1470
Aromatics								2880-2960
Resins	7.0-28.5	3	96.5	94.6	1.06	1.30	3	1550-1750

^aNumber of principal components used in the PLS regression model. ^b Explained calibration variance in x (IR spectra) and y(experimental values). ^c Root-mean-square error of calibration/prediction.

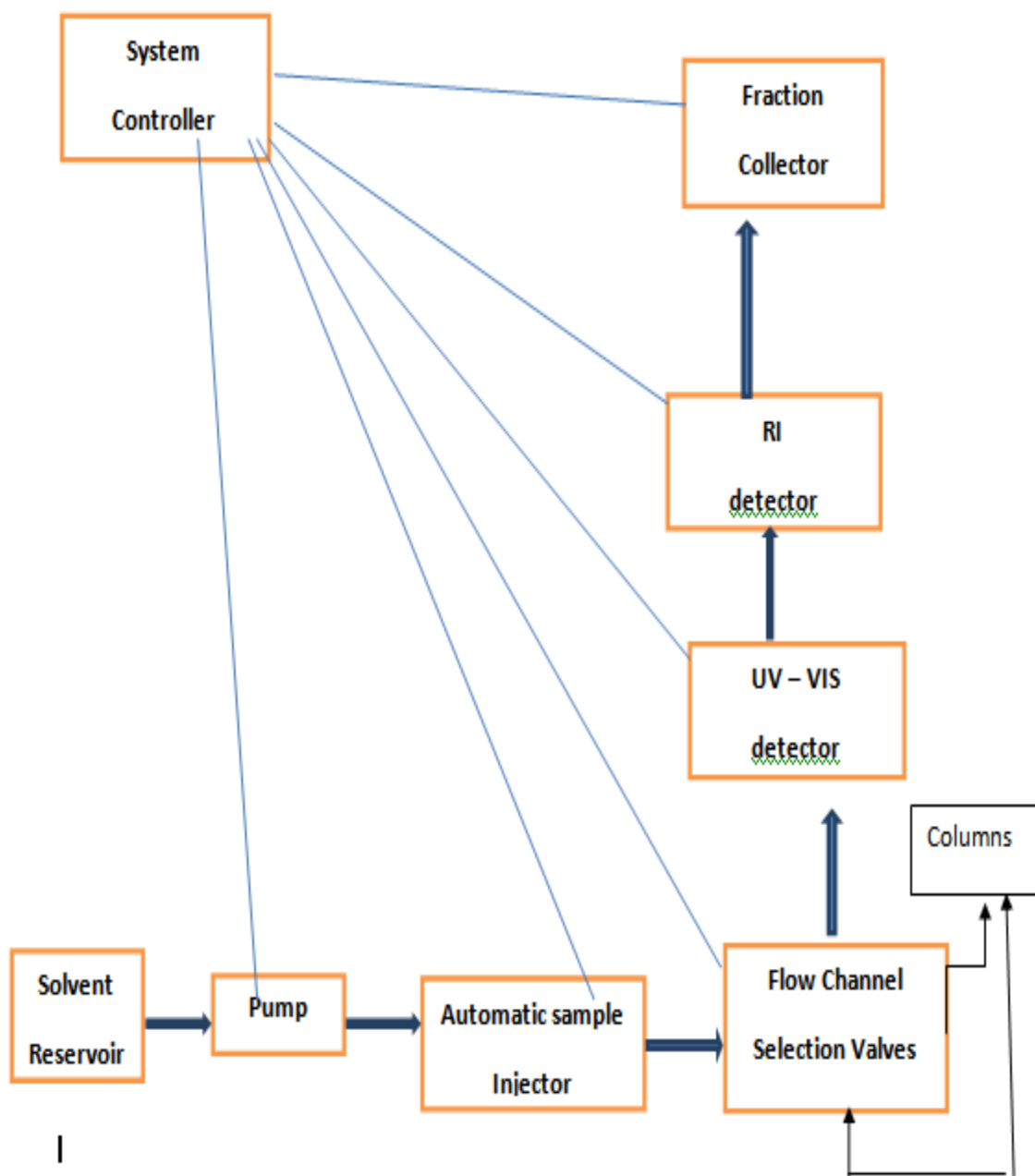


Figure 1: A Schematic representation of HPLC system with main Components

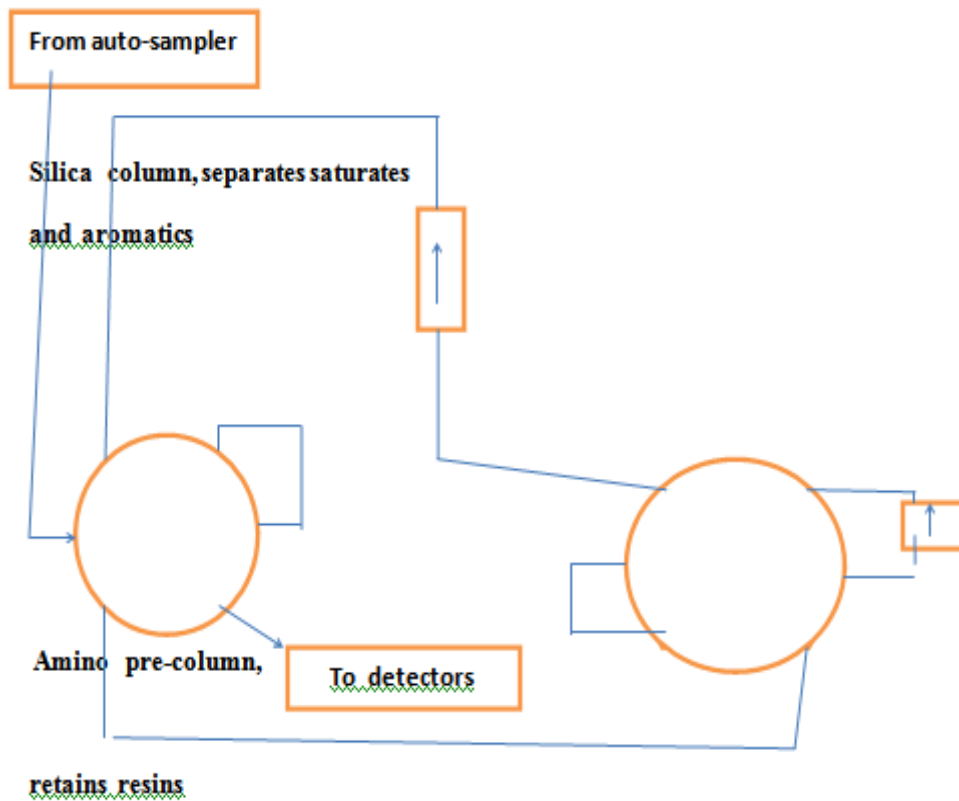


Figure 2a : Schematic representation of flow channel selection valves. Valve A controls the flow direction through columns (i.e. normal/backflush), and valve B includes or excludes the amino pre-column.

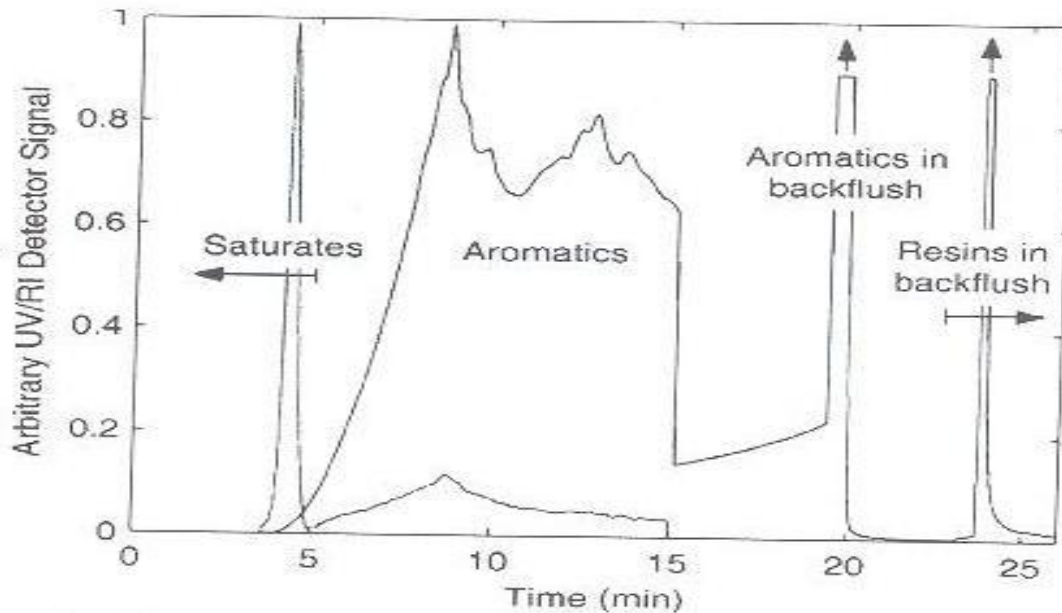


Figure 2b:A chromatogram from preparative HPLC analysis, The solid line depicts the RI signal and the broken line depicts the UV Chromatogram at 254nm.

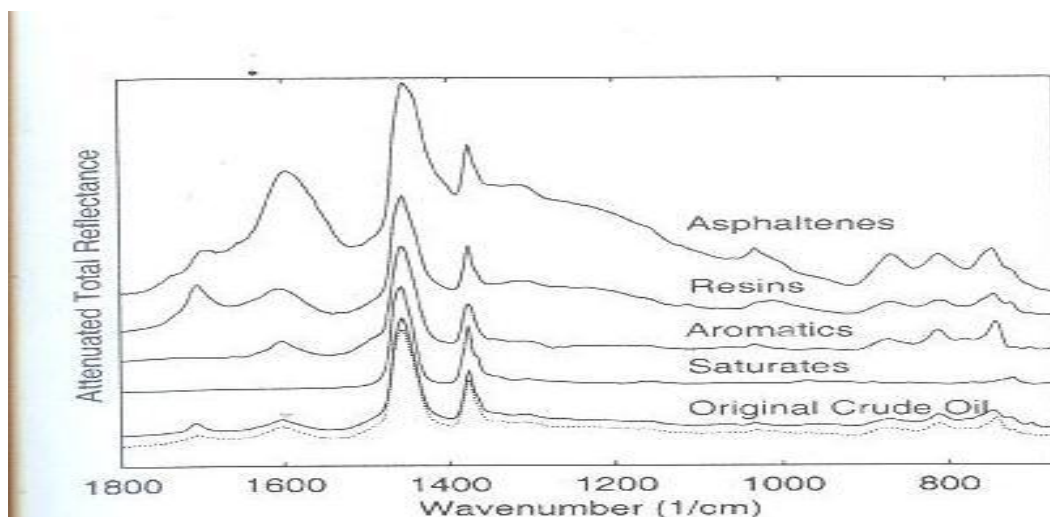


Figure 3: The IR spectra of SARA fractions from Osso -01 alongside with a spectrum of original crude from the field.

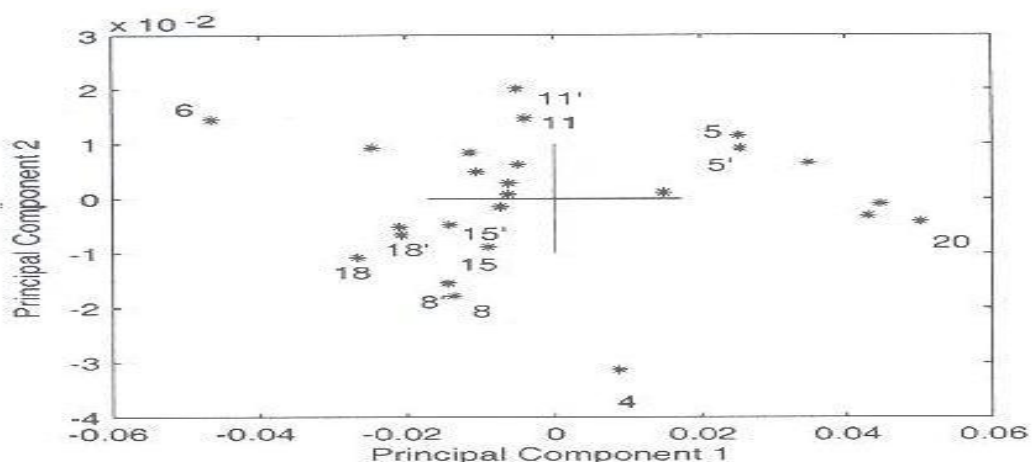


Figure 4: Specific score plot from principal component analysis of first-order-differentiated Fourier Transform-Infrared (FT-IR) spectra. Plot of replicate samples (5', 8', 11', 15' and 18') are also included.

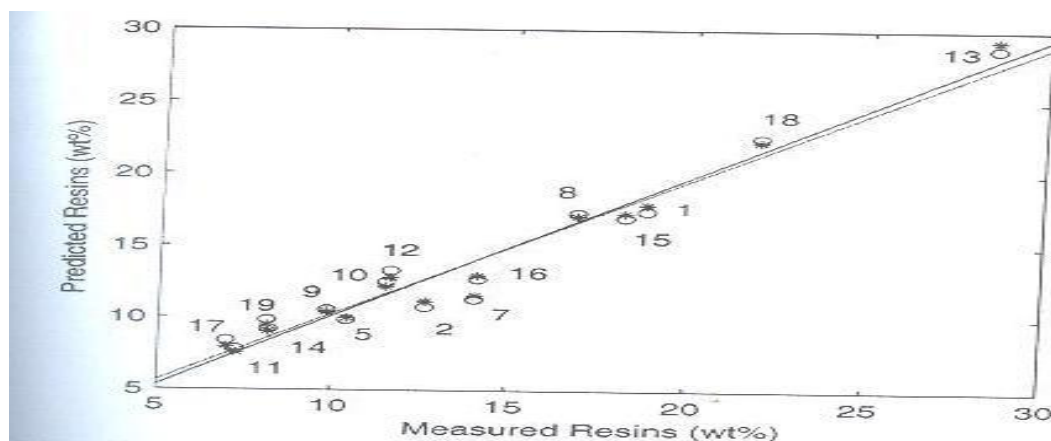


Figure 5: The correlation of resins from the PLS (Partial Least Square) models of IR Spectra to measured values.

The solid line is a linear least-squares regression calculated from the calibration set, and the broken line denote the corresponding regression calculated from validation samples.

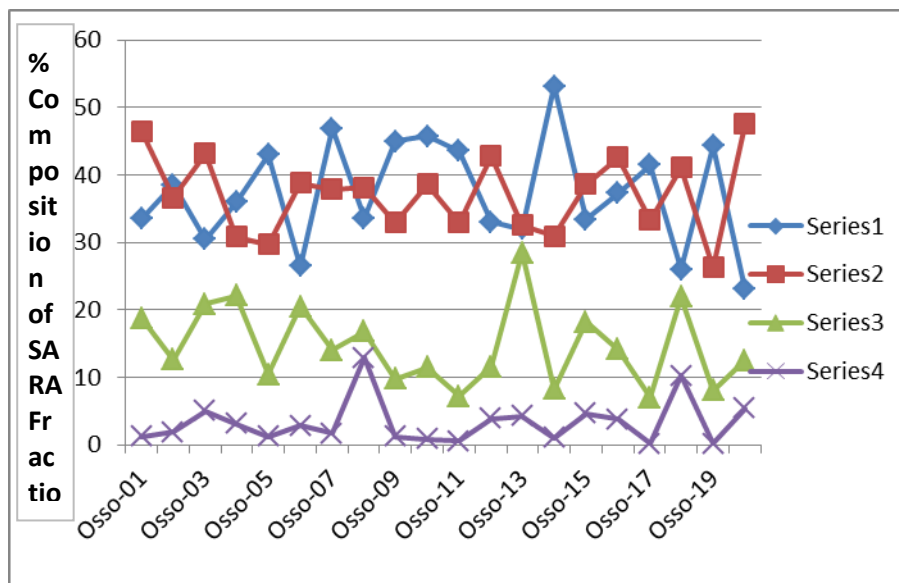
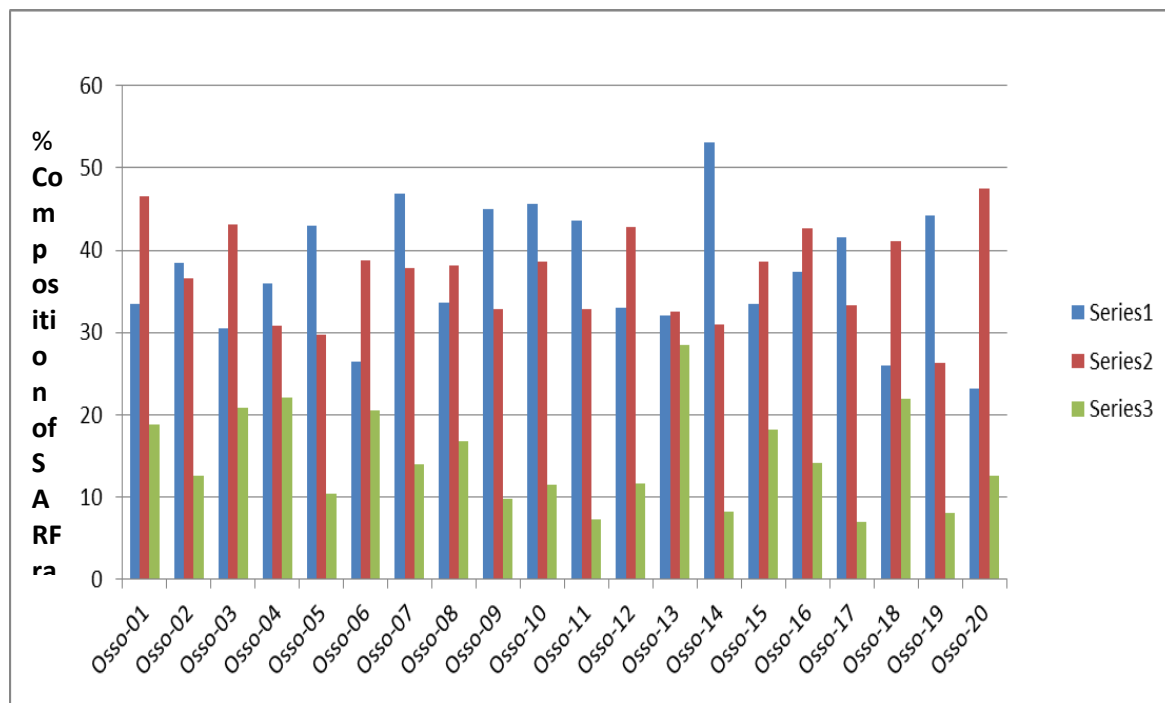


Figure 6: Variations of percentage(%) composition of SARA fractions in the calibration set.(Series 1 represent saturates, series 2 represent aromatics, series 3 represent resins, and series 4 represent asphaltenes)

Figure 7: Selective distribution of three components' percentage composition of saturates, aromatics, and resins(SAR) in the calibration set. (series 1 is saturates, series 2 represent aromatics while series 3 represent resins.)



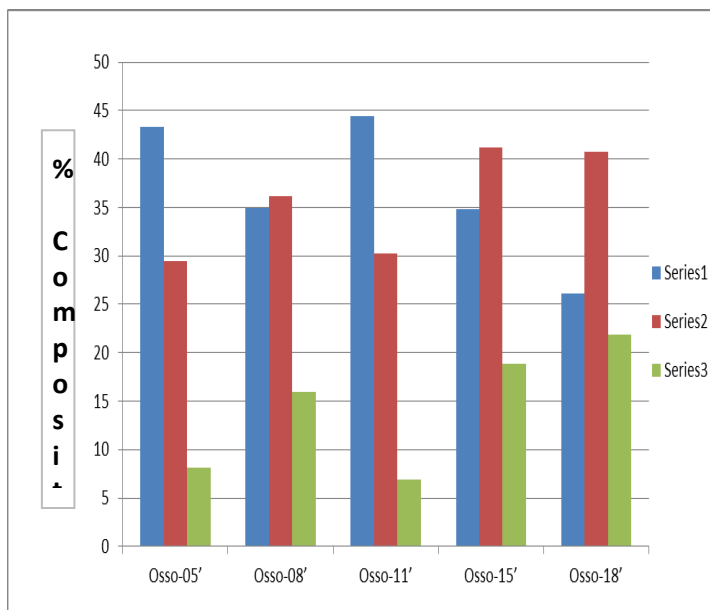


Figure 8: Distribution of percentage composition of SAR fractions in replicate samples.(series 1 depicts saturates, series 2 depicts aromatics and series 3 depicts resins)

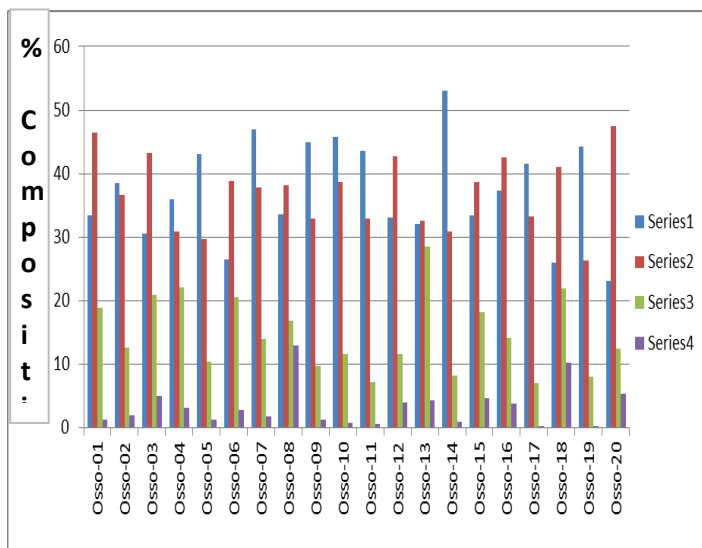


Figure 9: Grouped distribution of SARA components in calibration set. (series 1 depicts saturates, series 2 depicts aromatics, series 3 depicts resins, and series 4 represent asphaltenes)