Characterization of Marine Oil Spills: Estimation of Saturates, Aromatics, Resins, and Asphaltenes using Radar Derived Physical Properties

Durell S. Desmond[a,*], *Aura Diaz*[a], *Nasima Chorfa*[a], *Nhu Duong*[a], *Feiyue Wang*[a], *Jens Ehn*[a], *Eric Collins*[a], *Gary A. Stern*[a], and *Dustin Isleifson*[a]

[a] Centre for Earth Observation Science, Department of Environment and Geography, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

[*]Corresponding author. E-mail address: umdesmod@myumanitoba.ca

Abstract

Oil spill response measures (bioremediation, in-situ burning, and mechanical recovery) depend on the composition of the oil spilled; therefore, timely knowledge of an oil's composition (e.g., Saturates-Aromatics-Resins-Asphaltenes, SARA[*]) in the event of a spill is crucial. Remote sensing (RS) is a rapid technique, especially for spills in remote regions such as the Arctic, and has been used to estimate oil's physical properties (e.g., density). Here, we demonstrate that RS can also be used to estimate the oil's chemical composition. This study demonstrates how the SARA[*] composition of oil spills can be obtained from radar derived properties through both inversion and machine learning (ML) techniques using dual or single inputs (e.g., density, dynamic viscosity). We measured the SARA[*] composition and dynamic viscosity of 11 petroleum oils (crude and fuel) ranging in density between 0.8 and 1.0 g/mL. The SARA[*] fractions of the petroleum oils were estimated using single predictor (density) inversion models and both single (density) and dual predictor (density and dynamic viscosity) ML models. Linear correlation between SARA[*] fractions measured in the

laboratory and SARA* fractions estimated using ML dual predictor and single predictor models yielded R^2-values ranging from 0.57 to 0.94 and 0.58 to 0.87, respectively. The ML models yielded more reliable values for the SARA* fractions than the single predictor inversion models (R^2 ranging from 0.05 to 0.86). These findings provide the foundational framework for estimating the SARA* fractions of oil using oil physical properties through radar remote sensing.

Keywords: Petroleum hydrocarbons, Oil composition, Oil physical properties, Inverse modelling, Machine learning modelling, Oil spill mitigation

Highlights:

- Inversion and machine learning methods can estimate an oil's SARA* composition
- The density of oil can be used to estimate the SARA* contents of oil
- Machine learning provides more reliable estimates compared to inversion
- The use of two predictors slightly improves the accuracy of estimation

1 Introduction

Major compound groups of petroleum oils include alkanes, cycloalkanes (including steranes and terpanes), aromatics (including polycyclic aromatic hydrocarbons, PAHs), and heterocyclic compounds that contain elements of nitrogen, sulphur, and oxygen (NSOs) [1]. These compound groups are often categorized into four major fractions (SARA*): Saturates (S, encompassing alkanes and cycloalkanes), Aromatics (A, including PAHs), Resins (R), and Asphaltenes (A*, both of which encompass NSOs with varying polarity and aromaticity) [2]. Past research has shown that $S < A \leq R \leq A^*$ in terms of density [3-6], dielectric constant [4-5, 7-9], and thermal conductivity [10-15]. Therefore, there is an inherent connection

between these compound groups' chemistry and physical properties. API Tech Data Book 10 [10] states that an oil's density can be described by the weighted average of its constituents. Assuming that the independent concentration of individual compounds adequately describes the physical properties of a complex mixture, the density (or related property) of bulk oil should be governed by the abundance of saturates, aromatics, resins, and asphaltenes.

Previous research has shown that bulk oil density is directly linked to other physical properties such as the oil's dielectric constant [16-17], refractive index [18], thermal conductivity [19], dynamic viscosity [10,20], and API gravity [21]. Moreover, regression [18, 22-24], machine learning [24], and genetic programming [24] have been used to relate the physical properties of bulk oil (e.g., density, refractive index, API gravity) to its chemical fractions in which knowledge of the SARA* percentages can be used to estimate the oil's physical properties. However, fewer studies have investigated the use of an oil's physical properties (e.g., density, viscosity) to estimate its composition [6, 25-26]. Friisø et al. [6] developed quantitative formulae for estimating the asphaltene content in crude oil based on measured conductivities and dielectric constants of various solutions of asphaltene, resin, and total interfacially active fractions. This procedure was complex and resulted in large uncertainties (up to 93% difference and an average difference of 18%). Sulaimon et al. [25] used regression analyses (n=310) to create linear equations between the respective SARA* fractions and that of density and dynamic viscosity, in which density and dynamic viscosity were used as both single and dual predictors. The dual predictor correlations were found to be more accurate and reliable than the single predictor correlations [25] and were found to best estimate saturates (coefficient of determination, R^2=0.61) compared to aromatics

(R^2=0.24), resins (R^2=0.39), and asphaltenes (R^2=0.13) when comparing 216 crude oils [26]. Kulkarni et al. [26] achieved better estimations of SARA* using machine learning based on 216 crude oils in which density and dynamic viscosity were used as dual predictors. The accuracy of this method for predicting the SARA* fractions was determined to be R^2=0.84, R^2=0.54, R^2=0.78, and R^2=0.74, respectively.

Recent studies have demonstrated the potential of remotely inferring the physical properties (e.g., dielectric constant and density) of spilled oil with the use of radar remote sensing [27-28]. Given the relationship between an oil's chemistry and its physical properties, it may be possible to estimate the SARA* content of spilled oil remotely. Understanding the SARA* composition of spilled oil in the ocean is important as it can help guide an appropriate mitigation response, as various mitigation procedures (e.g., bioremediation, in-situ burning, and mechanical recovery) depend on the composition of the oil [1, 29-33]. For example, biodegradation is known to preferentially degrade saturates > aromatics > resins ≥ asphaltenes [30-31]. Therefore, a lighter oil would be more susceptible to biodegradation. In-situ burning requires a fuel source provided by the vaporization of oil. A particular oil's ignitability and potential to perform a steady-state burn will depend on the prior composition of the oil and its degree of weathering (i.e., lighter and more volatile or heavier) [32-33]. Lastly, mechanical recovery often consists of using skimmers to recover oil from the water surface. Heavier oils with increased viscosity and reduced flow present operational challenges to the efficiency and effectiveness of skimmers [1, 29].

In this study, we demonstrate that the oil's chemical composition (Saturates-Aromatics-Resins-Asphaltenes, SARA*) can be determined from the physical properties (i.e.,

density and dynamic viscosity) using inversion and machine learning models. This research seeks to aid in oil spill response measures (bioremediation, in-situ burning, and mechanical recovery) by providing the framework for a simple, quick, and inexpensive means of estimating the SARA* fractions of petroleum oils from physical properties that can be derived from radar remote sensing. Our research builds on previous work (e.g., Chamkalani [18] and Kulkarni et al. [26]) by demonstrating the potential of optimized inversion and machine learning models to estimate the SARA* fractions of oil (crude and fuel) using its physical properties (i.e., density and dynamic viscosity) as dual and single inputs.

2 Materials and Methods

2.1 Sample Information

The 11 oil samples included 6 crude oils, 4 diesel fuels, and 1 bunker fuel. Bunker Fuel C was provided by the University of Calgary. Diesel Fuel (Red River Co-op) was acquired from a Red River Co-op gas station. Diesel Fuel (Churchill) was provided by the community of Churchill. Regular Diesel Fuel was acquired from an undisclosed gas station. Marine Diesel Fuel (Amundsen) originated from the Canadian Coast Guard Ship (CCGS) Amundsen during the 2018 cruise within Hudson Bay [34]. Tundra Crude was provided by Tundra Oil & Gas Partnership. Lastly, Manitoba Virden Crude, Hibernia Crude, Hebron Crude, Conventional Heavy Crude, and Cold Lake Crude were obtained as part of the Multi-Partner Research Initiative (MPRI). The solvents used were gas chromatography grade, including pentane (99.9%), dichloromethane (DCM, 99.9%), and toluene (99%), all from Fisher Chemical™. Other chemicals included hydrochloric acid (36.5-38.0%), silica gel (100-200

Mesh particle size), alumina basic (60-325 Mesh), Fuller´s earth (~100 Mesh), and sodium sulfate anhydrous (≥99.0%), all of which were from Millipore Sigma Canada Ltd.

2.2 Laboratory Measurements

The measurement of the SARA* fractions of the oils followed a procedure modified from ASTM D2007-19 [35] and ASTMD6560-17 [36] as shown in Figure 1. The extraction process involved separating the components based on their polarity using a combination of gravimetric technique and clay-alumina/silica chromatography. Each sample was first acidified with HCl and then diluted in pentane. Afterwards, the samples were passed over sodium sulphate in a funnel to filter out pentane-insoluble asphaltenes and remove any residual water. Each filtrate was collected in a round bottom flask filled with Fuller's earth to absorb the resins, but not the aromatic and saturated hydrocarbons. The asphaltenes were recovered by rinsing the sodium sulphate with dichloromethane and the resins by rinsing the clay with dichloromethane. Next, pentane containing both aromatic and saturated hydrocarbons was poured onto an alumina/silica column. The saturates passed through the column into the pentane while the aromatics were adsorbed in the alumina/silica gel. The aromatics were collected using dichloromethane to desorb and elute from the gel. The mass percentage of the residual hydrocarbon compound mixtures were determined after removal of the solvent using a Buchi Rotovapor® R-300. The extraction process was repeated if the difference between two replicate results exceeded 2%.

2.3 Density and Viscosity Measurements

The density (Fig. 2) of each oil product was determined at 20 °C using an Anton Paar SVM 3001 Viscometer following the protocols of ASTM (ASTM D341-20e1; ASTM D7042-21a; ASTM D4052) [37-39]. The dynamic viscosity of each oil product was determined at a range of temperatures according to Figure 3. This instrument is capable of measuring density and viscosity (dynamic and kinematic) across a wide temperature range at atmospheric pressure. In this study, a temperature range between –20 and 60 °C was chosen for the various samples based on initial observations at room, fridge, and freezer temperatures. Extrapolations of measured viscosity values to lower temperatures were conducted by the instrument in accordance with the ASTM protocols (ASTM D341-20e1) [37]. Before sample measurements, calibration was run with certified reference standards supplied by Anton Paar (APN7.5 and APS3), which have known density and viscosity (dynamic and kinematic) values at the desired temperatures. 5 mL syringes were used to inject the respective oil samples and standards into the instrument at each desired temperature in which both density and viscosity cells were filled in a single cycle, and measurements were carried out simultaneously. Results are the average values of 2 out of 3 repeated measurements showing the smallest difference (in accordance with the Anton Paar instrument protocols). The relative deviation of the viscosity is 0.35% of the difference between two measured values, and that of the density is ±0.0001 g/cm³ (Anton Paar).

2.4 Inversion and Machine Learning Modelling Scheme

To assess the capability of obtaining SARA* fractions from density (ρ), laboratory measurements (see sections 2.2 and 2.3) were used to derive a single predictor inversion

model (see Equation 1), an optimization-based calculation applying a cost function (see Equation 2), and machine-learning models.

$$\rho_{(SARA^*)} = 1.033587 - 0.00184\,S - 0.00204\,A - 0.00268\,R + 0.005009\,A^* \quad (1)$$

$$Cost\ Function = \frac{|\rho(exp) - \rho(est)|^2}{\rho(exp)^2} \quad (2)$$

Eq. (1) exhibited a significance F value of 0.0017. Its multiple R, R^2, adjusted R^2, and standard error are 0.96, 0.92, 0.87, and 0.021, respectively. A compositional table of SARA* fractions was calculated from Eq. (1). The compositional table contained all possible combinations of SARA* within known ranges based on our observed data (S: 44 to 90%, A: 2 to 17%, R: 6 to 24%, A*: 0 to 26%), where each fraction increases in steps of one and their total adds to one hundred. The cost function equation (Eq. 2) was applied to the SARA* fractions values in the compositional table. The SARA* fractions with the best optimization were chosen as the best estimate for comparison with the experimental data. A second compositional table was then computed using the multiple linear regression derived by Chamkalani [18],

$$\rho_{(SARA^*)} = -13.26466 + 0.1402635\,S + 0.1390173\,A + 0.1655588\,R^{0.9731603} + 0.1014975\,As^{1.121967} \quad (3).$$

Eq. (2) was likewise applied to the SARA* fractions values in this compositional table.

The machine-learning regression was performed using the fitrnet function, part of the Statistics and Machine Learning Toolbox in MATLAB (R2024b). The fitrnet function was run to obtain two predictor machine-learning regression models (single and dual predictor models) that incorporated all SARA* fractions. The laboratory data (density, viscosity) were used as training data for the fitrnet function. The predictor model was, then, used to estimate the SARA* fractions for a given density (single predictor) and density and dynamic viscosity

(dual predictors) at 20 °C. The fractions were added to verify that they summed to one hundred. An optimizing parameter (part of fitrnet) was also applied, utilizing the standardize and ReLU activation functions.

Tabulated results for each SARA* estimation method are presented in Table 1 and are compared with the laboratory results. All calculations were performed using MATLAB 2024b. Two-sample t-tests assuming unequal variance were also conducted to calculate two-tailed p-values for data comparisons between the different models.

3 Results and Discussion

Two single predictor inversion models (based on Eqs. (1), (2), and (3)) and two machine learning (ML) models (single and dual predictors) were used to estimate the SARA* fractions of 11 oils for comparison with the experimentally determined fractions. The inversion and single predictor ML models use density (20°C) as the sole predictor, whereas the ML dual predictor model uses both density and dynamic viscosity (20°C) as inputs.

Figure 4 shows the averaged absolute difference between laboratory measurements and estimated SARA* fractions. The laboratory measurements and estimated SARA* percentages are presented in Table 1. The estimation of SARA* using the dual predictor ML model yielded more accurate results compared to the estimation of SARA* using a single predictor in both the inversion and ML models (Figure 4). Between the two inversion models, the model based on the multiple linear regression developed from the data herein (Eq. (1)) performed better than the inversion model based on the data from Chamkalani [18] (Eq. (3)) as indicated by the lower averaged absolute difference (Figure 4) and overall higher R^2

(shown in the bottom right corner Figure 5). For all models, asphaltenes exhibit the highest R^2 and saturates exhibit the lowest (shown in the bottom right corner Figures 5). However, all estimated SARA* fractions were not found to be significantly different (two tailed p-value > 0.05; Table 2).

The linear correlation between the laboratory measurements of SARA* fractions and the estimated SARA* fractions using the optimized inversion and ML models are shown in Figure 5. The R^2-values for the estimated SARA* fractions based on the two optimized inversion models are smaller than the ML predictions. The dual predictor ML model yielded more accurate results than the single predictor ML model, with R^2-values of 0.80, 0.57, 0.78, and 0.94 for the SARA* fractions, respectively.

Despite the smaller dataset herein compared to previous studies (i.e., Kulkarni et al. [26] and Chamkalani [18]), our inversion and ML models yielded reasonably good results in terms of accuracy (Figures 4 and 5). Indeed, the R^2-values (0.80, 0.57, 0.78, and 0.94; Figure 5) for the dual predictor ML model are comparable to the R^2-values (0.84, 0.54, 0.78, 0.74) reported by Kulkarni et al. [26] who also used a dual predictor ML model based on density and dynamic viscosity. The optimized based calculation of the estimated SARA* fractions using our inversion single predictor models Eqs. (1 and 3) provided similar results (see Figures 4 and 5), supporting the validity of our multiple linear regression model. This study is the first to present and compare optimized inversion models and a single predictor ML model. Hence, the inversion and ML models presented here provide the foundational framework for estimating the SARA* fractions of oil using oil physical properties.

To demonstrate how chemical composition of spilled oil can be determined from the physical properties of oil, we estimated the SARA* fractions from radar-derived oil density using the optimized inversion single predictor model based on Eqs. (2) and (3) and radar derived oil density reported by Desmond et al. [27]. We chose this model given that oil density is more attainable than dynamic viscosity and that optimized inversion models can be easily implemented in spreadsheets (e.g., Microsoft Excel). Eq. (3) was used over Eq. (1) as it was not developed using the oils detailed in Section 2.1. Desmond et al. [27] successfully obtained the density and dielectric constant of spilled diesel fuel (Red River Co-op) on seawater using radar inversion. Based on the retrieved density (0.86 g/mL) and single predictor inversion, the estimated SARA* fractions of the spilled diesel fuel (weathered oil) are 62, 17, 6, and 15%, respectively. The SARA* fractions for the fresh oil (derived herein) using the same model are 72, 11, 13, and 4%, respectively (Table 1). Comparing the SARA* fractions, we observe a lower saturate content and a higher aromatic content for the weathered oil, indicative of evaporation of the more volatile saturates, thereby increasing the oil's density. This simulation shows the potential for tracking the characteristics of spilled oil through remote sensing using the radar derived physical properties (e.g., oil density) to infer the oil's chemical composition, providing essential information for oil response measures.

Lastly, given the relationship between an oil's density and dynamic viscosity (exponential at 20°, R^2=0.9979 herein) [40], dielectric constant (linear at -40° to 60°, R^2=0.9616) [27], and thermal conductivity [41], future work could focus on deriving a second property synthetically for use as a dual predictor in inversion and ML models for determining SARA* fractions.

4 Conclusion

Knowing the chemical composition of spilled oil is vital for the implementation of a rapid and adequate response measure (bioremediation, in-situ burning, and mechanical recovery). This study demonstrated that the chemical composition of petroleum oils can be estimated from the physical properties (i.e., density and dynamic viscosity) using inversion and machine learning models. From laboratory measurements of density, dynamic viscosity, and SARA* fractions of 11 petroleum oils (4 diesel fuels, 6 crude oils, and 1 bunker fuel), we developed optimized inversion (single predictor) and machine learning models (single and dual predictor). The single predictor optimized inversion model obtained from the laboratory measurements of the 11 petroleum oils of this study yielded slightly more reliable results of the chemical composition (SARA*) compared to the single predictor optimized inversion model based on Chamkalani. Both machine learning models performed comparably. Overall, the machine learning models produced the most reliable estimates of the oil's chemical composition.

Our findings provide the foundational framework (i.e., proof of concept) for a simple, quick, and inexpensive means of estimating the SARA* fractions of petroleum oils from radar derived oil properties (e.g., density). Therefore, it is recommended that future work train a widely available machine learning model based on a large dataset of oils in which density or both density and dynamic viscosity be used as single or dual inputs to estimate the SARA* fractions.

Acknowledgements

This work was supported by Natural Sciences and Engineering Research Council (NSERC) of Canada, and the University of Manitoba. This work was supported by and is a contribution to GENICE II, Genome Canada, Research Manitoba, and Multi-Partner Research Initiative (MPRI). We thank Katarzyna Polcwiartek (University of Manitoba) for her assistance with data processing support.

References:

[1] AMAP Assessment 2007: oil and gas activities in the arctic — effects and potential effects. *Arctic Monitoring & Assessment Programme*, 2010, Oslo, vols. 1 and 2.

[2] Lee K (chair), Boufadel M, Chen B, et al. Expert Panel Report on the Behaviour and Environmental Impacts of Crude Oil Released into Aqueous Environments. Royal Society of Canada, Ottawa, 2015. ON. ISBN: 978-1-928140-02-3

[3] Dymond J H, Malhotra R. "Densities of n-alkanes and their mixtures at elevated pressures." *Int. J. Thermophys.,* 1987, 8: 541–555. https://doi.org/10.1007/BF00503641

[4] Panuganti S R, Wang F, Chapman W G, et al. A Simple Method for Estimation of Dielectric Constants and Polarizabilities of Nonpolar and Slightly Polar Hydrocarbons." *Int. J. Thermophys.*, 2016, 37: 75. https://doi.org/10.1007/s10765-016-2075-8

[5] Alvarez J O, Jacobi D, Althaus S, et al. Dielectric characterization of geochemical properties of liquid hydrocarbons from 25°C to 125 °C. *Fuel*, 2021, 288: 119679.

[6] Friisø T, Schildberg Y, Rambeau O, et al. COMPLEX PERMITTIVITY OF CRUDE OILS AND SOLUTIONS OF HEAVY CRUDE OIL FRACTIONS. *Journal of Dispersion Science and Technology*, 1998, 19(1): 93-126. https://doi.org/10.1080/01932699808913163.

[7] Sen A D, Anicich V G, Arakelian T. Dielectric constant of liquid alkanes and hydrocarbon mixtures. *Journal of Physics D: Applied Physics,* 1992, 25: 516-521.

[8] Van Dyck C, Marks T J, Ratner M A. Chain Length Dependence of the Dielectric Constant and Polarizability in Conjugated Organic Thin Films. *ACS Nano*, 2017, 11: 5970–5981. https://doi.org/10.1021/acsnano.7b01807

[9] Alvarez J O. Dielectric Properties of Aromatic Components of Crude Oil. *Energy Fuels*, 2020, 34(1): 270–277.

[10] API Tech Data Book 10. Density. In: API Technical Data Book, 10th ed., 2016. The American Petroleum Institute and EPCON International.

[11] Elam S K, Tokura I, Saito K, et al. Thermal conductivity of crude oils. *Exp. Therm. Fluid Sci.*, 1989, 2: 1–6. https://doi.org/10.1016/0894-1777(89)90043-5

[12] Jamieson D T, Irving J B, Tudhope J S. Prediction of the thermal conductivity of petroleum products. *Wear*, 1975, 33: 75–83. https://doi.org/10.1016/0043-1648(75)90225-2

[13] Jamieson D T, Cartwright G. Thermal Conductivity of Cycloaliphatic Liquids, in: Thermal Conductivity 16. Springer US, Boston, MA, 1983, pp. 435–438. https://doi.org/10.1007/978-1-4684-4265-6_38

[14] Lashkarbolooki M, Hezave A Z, Bayat M. Correlating thermal conductivity of pure hydrocarbons and aromatics via perceptron artificial neural network (PANN) method. *Chinese J. Chem. Eng.*, 2017, 25: 547–554. https://doi.org/10.1016/J.CJCHE.2016.08.025

[15] Lei Q, Hou Y-C, Lin R. A new correlation for thermal conductivity of liquids. *Chem. Eng. Sci.*, 1997, 52: 1243–1251. https://doi.org/10.1016/S0009-2509(96)00481-2

[16] Desmond D S, Neusitzer T D, Lemes M, et al. Examining the Physical Interactions of Corn Oil (Medium Crude Oil Surrogate) in Sea Ice and its Potential for Chemical Partitioning within an Arctic Environment. Proceedings of the Forty-second AMOP Technical Seminar, Environment and Climate Change Canada, Ottawa, ON, Canada, 2019, pp. 66-95.

[17] Desmond D S, Saltymakova D, Neusitzer T D, et al. Oil Behavior in Sea Ice: Changes in Chemical Composition and Resultant Effect on Sea Ice Dielectrics. *Marine Pollution Bulletin*, 2019, 142: 216-233.

[18] Chamkalani A. Correlations between SARA fractions, density, and RI to investigate the stability of asphaltene. *ISRN Anal. Chem. (I)*, 2012, 219276: 1-6, https://doi.org/10.5402/2012/219276.

[19] Aboul-Seoud A, Moharam H M. A simple thermal conductivity-temperature correlation for undefined petroleum and coal liquid fractions. *Chem. Eng. Res. Des.*, 1999, 77: 248–252. https://doi.org/10.1205/026387699526043.

[20] Stratiev D, Shishkova I, Dinkov R, et al. Prediction of petroleum viscosity from molecular weight and density. *Fuel*, 2023, 331: 125679. https://doi.org/10.1016/j.fuel.2022.

[21] API gravity: Definition, calculation, and converter. The Engineering ToolBox. (Accessed 2025, July 29). https://www.engineeringtoolbox.com/api-gravity-d_1212.html

[22] Fan T, Buckley J S. Rapid and accurate SARA analysis of medium gravity crude oils. *Energy Fuels*, 2002, 16(6): 1571-1575. https://doi.org/10.1021/ef0201228

[23] Fan T, Wang J, Buckley J S. Evaluating crude oils by SARA analysis. Proceedings of the SPE/DOE Improved Oil Recovery Symposium, SPE 75228, Tulsa, Okla, USA, 2002.

[24] Goel P, Saurabh K, Patil-Shinde V, et al. Prediction of API values of crude oils by use of saturates/aromatics/resins/asphaltenes analysis: computational-intelligence-based models. SPE J., 2017, 22(3): 817-853. https://doi.org/10.2118/184391-pa

[25] Sulaimon A A, Rajan A/L, Qasim H, et al. Developing new correlations for asphaltene deposition involving SARA fractions and colloidal instability index. *J. Petrol. Sci. Eng.*, 2023, 220: 111143. https://doi.org/10.1016/j.petrol.2022.111143

[26] Kulkarni A D, Khurpade P D, Nandi, S. Estimation of SARA composition of crudes purely from density and viscosity using machine learning based models. *Petroleum*, 2024, 10: 620-630.

[27] Desmond D S, Zabihi Mayvanm M, Chorfa N, et al. Towards a framework for oil classification during oil spills in Arctic open water using microwave remote sensing. *Journal of Sea Research,* 2026, 211: 102698. https://doi.org/10.1016/j.seares.2026.102698

[28] Desmond D, Zabihi Mayvan M, Lashkari S, et al. Discrimination of Oil Types and Thickness during Oil Spills on Newly Formed Sea Ice using C-band Radar. IGARSS 2026 - 2026 IEEE International Geoscience and Remote Sensing Symposium, Washington DC, DC, USA, 2026, pp. 585-588, doi: 10.1109/IGARSS60577.2026.11730638

[29] JIP (Joint Industry Program): Oil spill contingency for Arctic and ice-infested waters. SINTEF, 2006.

[30] Stout S A, Wang Z. Chemical fingerprinting methods and factors affecting petroleum fingerprints in the environment, in Standard Handbook Oil Spill Environmental Forensics: Fingerprinting and Source Identification, Chap. 3 Stout S and Wang Z (2nd ed.), Elsevier/Academic Press, 2016.

[31] Prince R C, Walters C C. Biodegradation of oil hydrocarbons and its implications for source identification, in Standard Handbook Oil Spill Environmental Forensics: Fingerprinting and Source Identification, Chap. 19, Stout S and Wang Z (2nd ed.), Elsevier/Academic Press, 2016.

[32] Wilkinson J, Beegle-Krause C J, Evers K-U, et al. Oil spill response capabilities and technologies for ice-covered Arctic marine waters: A review of recent developments and established practices. *Ambio*, 2017, 46(Suppl. 3): S423–S441.

[33] Potter S, Buist I. In situ burning for oil spills in Arctic waters—State-of-the-art and future research needs. In Oil spill response—A global perspective, eds. Davidson W F, Lee K, and Cogswell A. Springer, 2008, 23–39.

[34] Hudson Bay System Study (BaySys). Centre for Earth Observation Science | University of Manitoba, 2024. https://umanitoba.ca/earth-observation-science/research/hudson-bay-system-study-baysys (accessed 2026-09-10)

[35] ASTM D2007-19. Standard Test Method for Characteristic Groups in Rubber Extender and Processing Oils and Other Petroleum-Derived Oils by the Clay-Gel Absorption Chromatographic Method. 2024. https://www.astm.org/d2007-19.html (accessed 2024-02-13)

[36] ASTM D6560-17. Standard Test Method for Determination of Asphaltenes (Heptane Insolubles) in Crude Petroleum and Petroleum Products. 2022. https://www.astm.org/d6560-17.html (accessed 2024-02-13)

[37] ASTM D341-20e1. Standard Practice for Viscosity-Temperature Equations and Charts for Liquid Petroleum or Hydrocarbon Products. 2020. https://compass.astm.org/document/?contentCode=ASTM%7CD2007-19%7Cen-US (accessed 2024-02-13).

[38] ASTM D7042-21a. Standard Test Method for Dynamic Viscosity and Density of Liquids by Stabinger Viscometer (and the Calculation of Kinematic Viscosity). 2021. https://www.astm.org/d7042-21a.html (accessed 2024-02-23)

[39] ASTM D4052. Standard Test Method for Density, Relative Density, and API Gravity of Liquids by Digital Density Meter -- eLearning Course. 2022. https://www.astm.org/astm-tpt-430.html (accessed 2024-02-23)

[40] Dynamic, Absolute, and Kinematic Viscosity – Definitions & Conversions. The Engineering ToolBox. (accessed 2025, July 29). https://www.engineeringtoolbox.com/dynamic-absolute-kinematic-viscosity-d_412.html

[41] Aboul-Seoud A., Moharam H M. A simple thermal conductivity-temperature correlation for undefined petroleum and coal liquid fractions. *Chem. Eng. Res. Des.* 77, 1999, 248–252. https://doi.org/10.1205/026387699526043

Figures

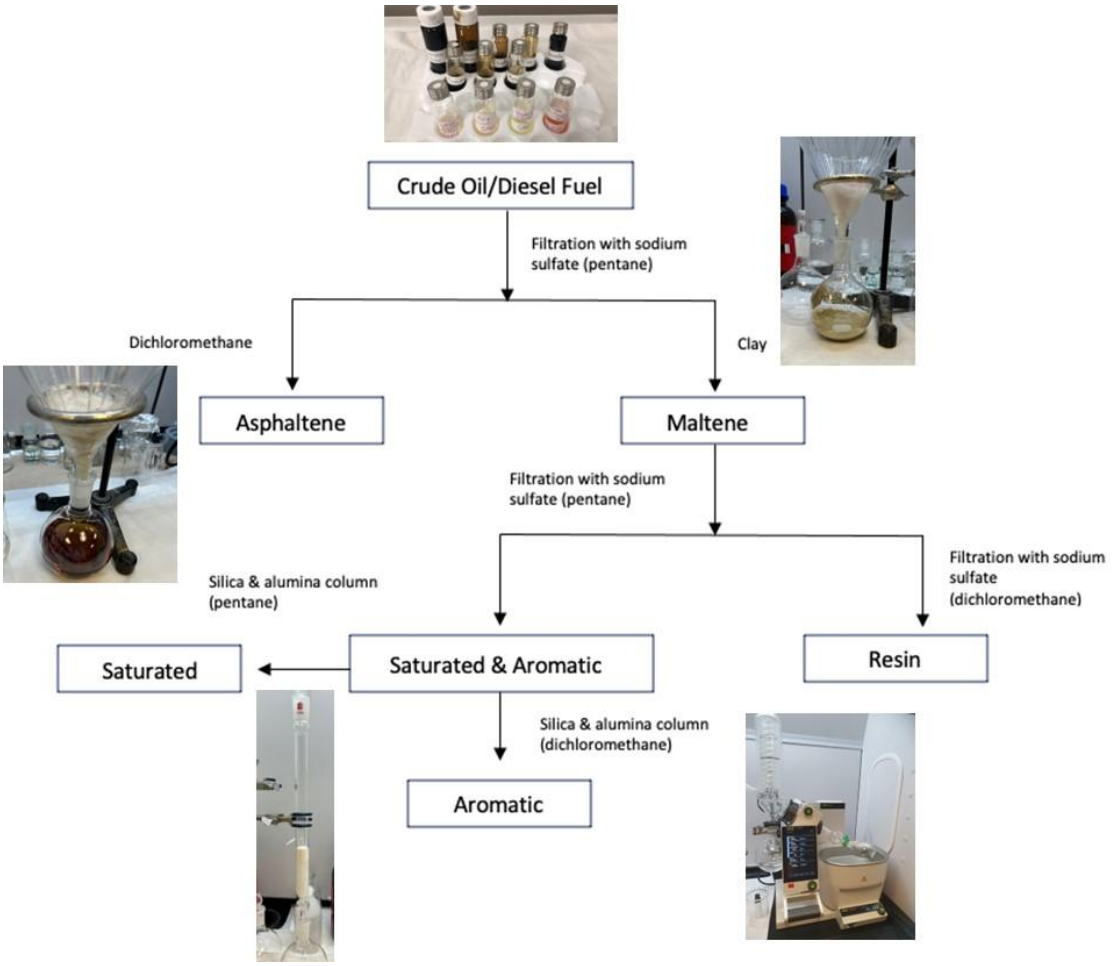

Figure 1. SARA* extraction experimental setup.

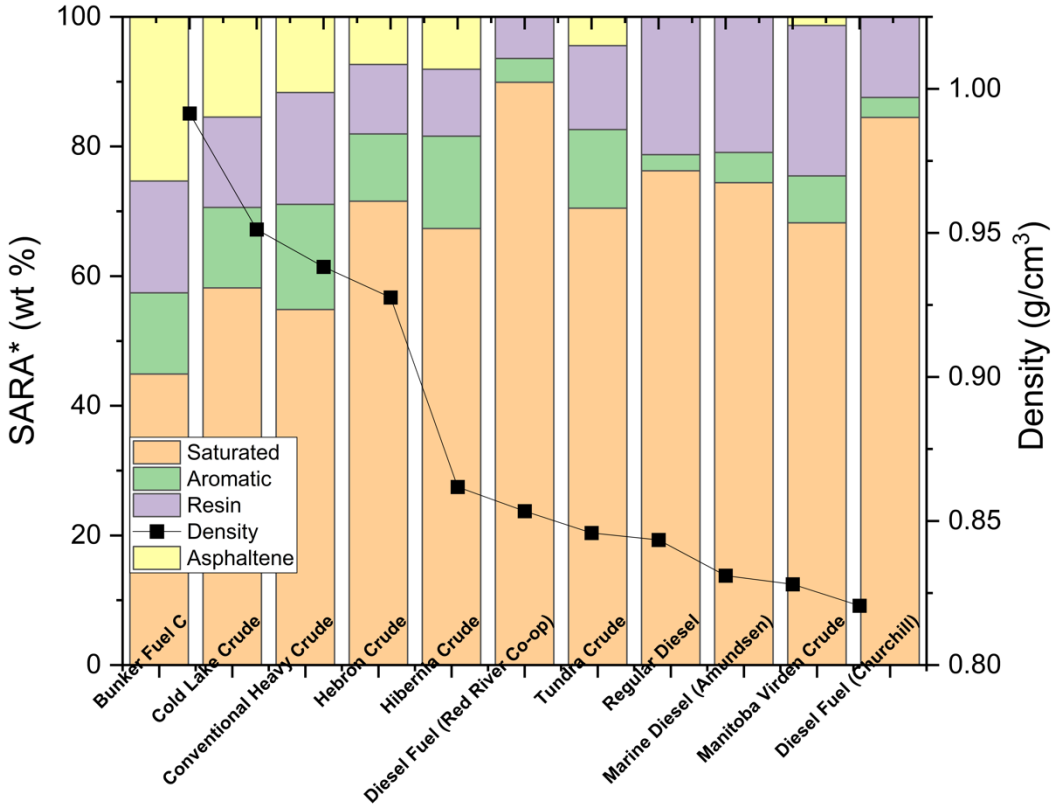

Figure 2. Relative weight percent of Saturates-Aromatics-Resins-Asphaltenes (SARA*) content for 11 oils plotted against their respective density at 20 °C.

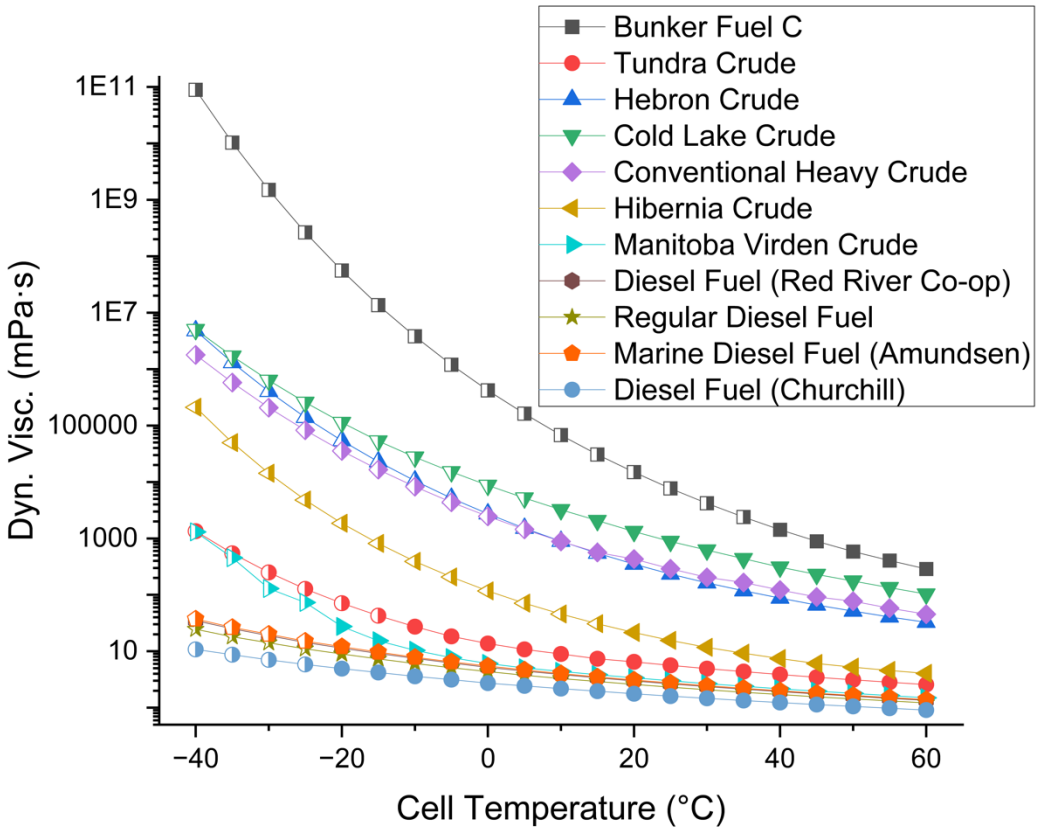

Figure 3. Viscosity (mPa·s) of 11 oils as a function of temperature (°C). Half hollowed data points indicate extrapolation (see Section 2.3 for details).

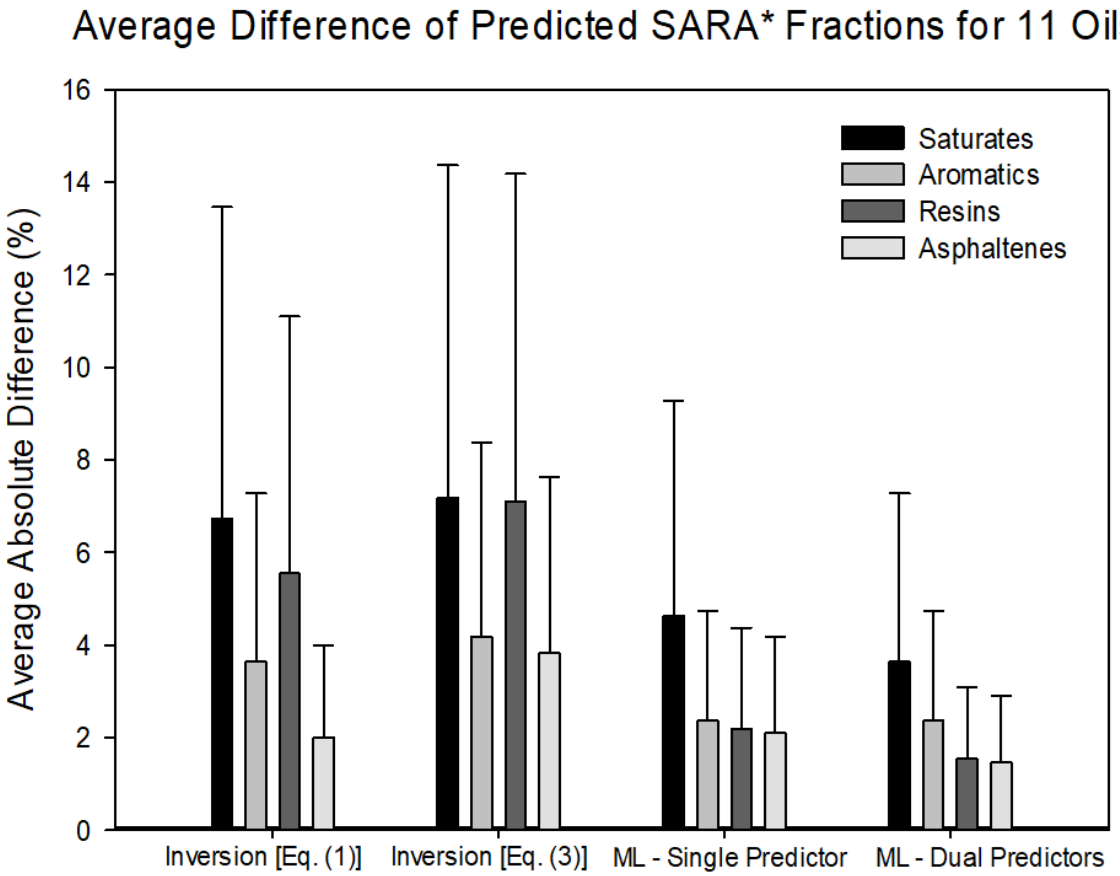

Figure 4. Averaged absolute difference (%) between estimated and experimental Saturates, Aromatics, Resins, and Asphaltenes (SARA*) for 11 oils. ML – Machine Learning; Single predictor – density (20°C); Dual predictors – density and dynamic viscosity (20°C).

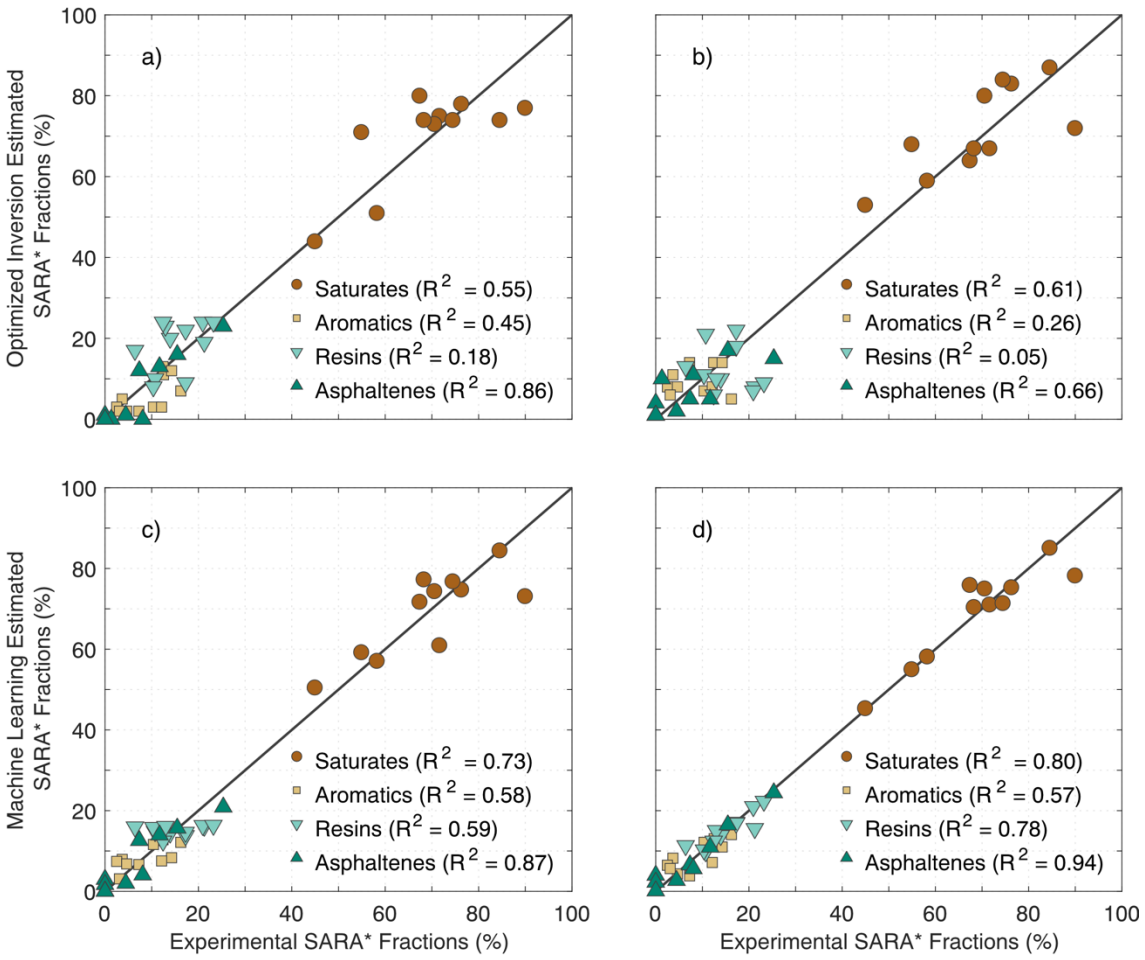

Figure 5. Scatter plots between laboratory measurements and estimated SARA* fractions. Results are shown for optimized inversion Eq. (1) (panel a), optimized inversion Eq. (3) (panel b), single predictor machine learning (panel c), and dual predictor machine learning (panel d). The solid black line represents the one-to-one fit.

Tables

Table 1. Estimated SARA* fractions using inversion and machine learning methods

Sample	S (%)	A (%)	R (%)	A* (%)
Bunker Fuel C (Experimental)	45	13	17	25
Bunker Fuel C (Inversion Eq. (1))	44	11	22	23
Bunker Fuel C (Inversion Eq. (3))	53	14	18	15
Bunker Fuel C (Machine Learning – Single Predictor)	45	13	20	22
Bunker Fuel C (Machine Learning – Dual Predictor)	45	14	16	25
Cold Lake Crude (Experimental)	58	12	14	15
Cold Lake Crude (Inversion Eq. (1))	51	13	20	16
Cold Lake Crude (Inversion Eq. (3))	59	14	10	17
Cold Lake Crude (Machine Learning – Single Predictor)	57	13	14	16
Cold Lake Crude (Machine Learning – Dual Predictor)	59	12	14	17
Conventional Heavy Crude (Experimental)	55	16	17	12
Conventional Heavy Crude (Inversion Eq. (1))	71	7	9	13
Conventional Heavy Crude (Inversion Eq. (3))	68	5	22	5
Conventional Heavy Crude (Machine Learning – Single Predictor)	61	13	12	13
Conventional Heavy Crude (Machine Learning – Dual Predictor)	57	14	17	11
Hebron Crude (Experimental)	72	10	11	7
Hebron Crude (Inversion Eq. (1))	75	3	10	12
Hebron Crude (Inversion Eq. (3))	67	7	21	5
Hebron Crude (Machine Learning – Single Predictor)	65	13	11	12
Hebron Crude (Machine Learning – Dual Predictor)	68	11	12	8
Hibernia Crude (Experimental)	67	14	10	8
Hibernia Crude (Inversion Eq. (1))	80	12	8	0
Hibernia Crude (Inversion Eq. (3))	64	14	11	11
Hibernia Crude (Machine Learning – Single Predictor)	78	10	9	2
Hibernia Crude (Machine Learning – Dual Predictor)	77	11	8	5
Diesel Fuel (Red River Co-op) (Experimental)	90	4	6	0
Diesel Fuel (Red River Co-op) (Inversion Eq. (1))	77	5	17	1
Diesel Fuel (Red River Co-op) (Inversion Eq. (3))	72	11	13	4
Diesel Fuel (Red River Co-op) (Machine Learning – Single Predictor)	76	9	13	2
Diesel Fuel (Red River Co-op) (Machine Learning – Dual Predictor)	77	10	10	4
Tundra Crude (Experimental)	70	12	13	4
Tundra Crude (Inversion Eq. (1))	73	3	23	1
Tundra Crude (Inversion Eq. (3))	80	8	10	2
Tundra Crude (Machine Learning – Single Predictor)	75	8	15	2
Tundra Crude (Machine Learning – Dual Predictor)	76	7	14	2

Regular Diesel Fuel (Experimental)	76	3	21	0
Regular Diesel Fuel (Inversion Eq. (1))	78	3	19	0
Regular Diesel Fuel (Inversion Eq. (3))	83	8	8	1
Regular Diesel Fuel (Machine Learning – Single Predictor)	75	7	16	2
Regular Diesel Fuel (Machine Learning – Dual Predictor)	76	7	15	1
Marine Diesel Fuel (Amundsen) (Experimental)	74	5	21	0
Marine Diesel Fuel (Amundsen) (Inversion Eq. (1))	74	2	24	0
Marine Diesel Fuel (Amundsen) (Inversion Eq. (3))	84	8	7	1
Marine Diesel Fuel (Machine Learning – Single Predictor)	72	5	21	1
Marine Diesel Fuel (Machine Learning – Dual Predictor)	72	6	21	0
Manitoba Virden Crude (Experimental)	68	7	23	1
Manitoba Virden Crude (Inversion Eq. (1))	74	2	24	0
Manitoba Virden Crude (Inversion Eq. (3))	67	14	9	10
Manitoba Virden Crude (Machine Learning – Single Predictor)	72	5	22	1
Manitoba Virden Crude (Machine Learning – Dual Predictor)	69	6	23	0
Diesel Fuel (Churchill) (Experimental)	84	3	12	0
Diesel Fuel (Churchill) (Inversion Eq. (1))	74	2	24	0
Diesel Fuel (Churchill) (Inversion Eq. (3))	87	6	6	1
Diesel Fuel (Churchill) (Machine Learning – Single Predictor)	84	3	12	0
Diesel Fuel (Churchill) (Machine Learning – Dual Predictor)	85	1	14	1

S – Saturates; A – Aromatics; R – Resins; A – Asphaltenes; Single predictor refers to density (20 °C); Dual predictors refer to density and dynamic viscosity (20 °C); SARA* values rounded to whole numbers.

Table 2. Two tailed p-values between the experimentally and theoretically determined saturate, aromatic, resin, and asphaltene percentages based on Table 1

Sample	**P-Values**
Saturates (Exp. vs Inversion Eq. (1))	0.84
Saturates (Exp. vs Inversion Eq. (3))	0.78
Saturates (Exp. vs ML – Single Predictor)	0.99
Saturates (Exp. vs ML – Dual Predictors)	0.97
Aromatics (Exp. vs Inversion Eq. (1))	0.11
Aromatics (Exp. vs Inversion Eq. (3))	0.41
Aromatics (Exp. vs ML – Single Predictor)	1.00
Aromatics (Exp. vs ML – Dual Predictors)	1.00
Resins (Exp. vs Inversion Eq. (1))	0.22
Resins (Exp. vs Inversion Eq. (3))	0.23
Resins (Exp. vs ML – Single Predictor)	1.00
Resins (Exp. vs ML – Dual Predictors)	0.97
Asphaltenes (Exp. vs Inversion Eq. (1))	0.88
Asphaltenes (Exp. vs Inversion Eq. (3))	0.90
Asphaltenes (Exp. vs ML – Single Predictor)	0.98
Asphaltenes (Exp. vs ML – Dual Predictors)	0.96

*Exp. – Experimental; ML – Machine Learning; Single Predictor – Density (20°C); Dual Predictors – Density and Dynamic Viscosity (20°C)