

SATELLITE DETERMINATION OF API GRAVITY AND SARA COMPONENTS OF OFFSHORE PETROLEUM SEEPS

Talita Lammoglia and Carlos Roberto Souza Filho

Recebido em 3 fevereiro, 2011 / Aceito em 18 maio, 2012
Received on February 3, 2011 / Accepted on May 18, 2012

ABSTRACT. Crude oils can be classified based on their API gravity or their SARA components (i.e. saturated, aromatics, resins, and asphaltenes). This work evaluates the possibility to infer such oil's characteristics by means of reflectance spectroscopy and multispectral and hyperspectral imagery. Reflectance spectra (0.35-2.5 μm) of oils and films of oil on water were measured under laboratory conditions using a high-resolution spectroradiometer. Multivariate statistics (i.e. principal component and partial least square analysis) were used to evaluate these spectra, taking into consideration both the spectral resolution at laboratory measurements (2150 bands) and the spectral resolution offered by orbital sensors (i.e. Hyperion (220 bands) and ASTER (9 bands)). The statistical approach yielded predictive models based on the correlation of oil composition and its spectral response, allowing remote assessment of the oil quality of a particular seepage from the Campos Basin (Brazil). This oil has a known API gravity ranging between 19-22, 40-49% mass/mass of saturated, 33-25 of aromatics, 20-28 of resins and 3-1.5 of asphaltenes. The remotely retrieved values based on the spectral response of the seep are within the actual range and are remarkably similar: API gravity of 19.6 and 45.38% mass/mass of saturated, 26.91 of aromatics, 24.61 of resins and 2.14 of asphaltenes. This result indicates the potential of this methodology, first proposed in this work, for the indirect inference of API gravity and SARA composition based on remote sensing data and techniques.

Keywords: seepage, spectroscopy, remote sensing, API, SARA.

RESUMO. Petróleos podem ser classificados segundo o grau API ou frações de hidrocarbonetos (HCs) saturados, aromáticos, resinas e asfaltenos (SARA). Esse trabalho objetiva avaliar a possibilidade de qualificar óleos remotamente quanto a essas características com base em dados de espectroscopia de reflectância e de imageamento multiespectral e hiperespectral. Para tanto, amostras de óleo e de filmes de óleo sobre água foram medidas em laboratório, através do espectrorradiômetro FieldSpec FR, para caracterização de seus espectros de reflectância (0,35-2,5 μm). Esses espectros foram avaliados por estatística multivariada (análise de componentes principais e regressão por mínimos quadrados parciais), tanto na resolução ultraespectral (2150 bandas) do espectrorradiômetro, como também na resolução espectral de sensores orbitais hiperespectrais (i.e. Hyperion (220 bandas)) e multiespectrais (i.e. ASTER (9 bandas)). A aproximação estatística visou o estabelecimento de modelos preditivos baseados na relação entre as características qualitativas de óleos e as feições espectrais que apresentam. Com base nesses modelos foi possível a avaliação remota de uma exsudação na Bacia de Campos. O petróleo exsudado realisticamente tem °API entre 19-22 e % massa/massa de hidrocarbonetos variando entre 40-49 para saturados, 33-25 para aromáticos, 20-28 para resinas e 3-1,5 para asfaltenos. Os valores derivados dos modelos preditivos baseados na assinatura espectral remota da exsudação foram °API = 19,6; %m/m_{saturados} = 45,38; %m/m_{aromáticos} = 26,91; %m/m_{resinas} = 24,61; %m/m_{asfaltenos} = 2,14 (% m/m). Esses resultados mostram o potencial do método aqui desenvolvido de forma pioneira para a determinação do grau API e da %m/m de hidrocarbonetos SARA em exsudações com base em dados e técnicas de sensoriamento remoto ótico.

Palavras-chave: exsudação, espectroscopia, sensoriamento remoto, API, SARA.

INTRODUCTION AND THEORETICAL CONTEXT

Remote sensing can be used as a tool in oil exploration, as well as in environmental monitoring. Considering recent technological progress, optical data is becoming more and more accessible, as is the case, for instance, that from the Advanced Spaceborne Thermal Emission and Reflection Radiometer sensor on board the Earth satellite. In spite of the limitations of this type of data in bad weather conditions (i.e. cloud cover), they help in the segmentation of offshore oil seeps (e.g., Lorenzetti et al., 2006), and make possible qualitative inferences about the oil (e.g., Lamoglia & Souza Filho, 2011; 2012; Wettle et al., 2009). The knowledge about the seep's oil quality on the water contributes to making decisions during the exploratory process. In parallel, when dealing with environmental monitoring situations, this information can be helpful for scanning the oil source in the case of oil pipeline seeps, ship spills or, additionally, illegal discharges.

This article evaluates the possibility of achieving the oil's remote characterization by means of reflectance spectroscopy, according to the hydrocarbons' saturated, aromatic, resin and asphaltene (SARA) components. The spectral measurements generated under laboratory conditions and associated with computational models will be tested for a real seepage situation registered at the Campos Basin (Brazil).

Petroleum

Petroleum in the liquid state is an oily, flammable mixture, usually less dense than water and with color varying between black and light brown. It is basically made up of a mixture of organic chemical components (hydrocarbons-HCs) that, depending on the molecule size and environmental pressure, is found as gas or liquid. According to the American Petroleum Institute (API), all oils substantially contain, in different amounts, a wide variety of hydrocarbons, which also have different characteristics (i.e., density, color, viscosity, gas content and oxidation resistance). However, the relative amount of the individual compounds within each hydrocarbon group is, approximately, of the same order of magnitude for different oils. Analysis of any typical crude elementary oil (in weight %) includes: hydrogen (11 to 14%), carbon (83 to 87%), sulfur (0.06 to 8%), nitrogen (0.11 to 1.7%), oxygen (0.1 to 2%) and metals (up to 0.3%) (Hunt, 1979; Tissot & Welte, 1984; Thomas, 2001).

Crude oils can be characterized based on their API gravity unit adopted by the American Petroleum Institute (<http://www.api.org>). The degree is defined by: $\text{API gravity} = [(141.5/\text{specific gravity}) - 131.5]$. According to the Brazilian Petroleum Agency (ANP), ultra-

heavy crude oils have API gravity <12 ; heavy oils have API gravity <22 , medium oils have API gravity between 22 and 31, while light oils have API gravity >31 . Many times, the words "heavy" and "viscous" are interchanged to define heavy oils, because heavy oils tend to be more viscous (Thomas, 2001; Alboudwarej et al., 2006).

The hydrocarbon characterization, particularly its heavier fraction, can be accomplished through the constituents' solubility behavior in paraffinic and aromatic organic solvents, classifying the fractions as saturated, aromatic, resins and asphaltenes (SARA), expressed as percentage in weight, solubility and chromatography (Alboudwarej et al., 2006). These are *saturated hydrocarbons*, *aromatic hydrocarbons* and *resins* and *asphaltenes*. The *saturated hydrocarbons*, also known as alkanes and paraffins, comprise normal and branched chain alkanes (paraffinic), including the C_nH_{2n+2} type hydrocarbon long-chain. The *aromatic hydrocarbons* hold pure aromatic molecules, cycloalkane-aromatic (naphthene-aromatic) and, usually, cyclo compounds of aromatic sulfur. In that case, the aromatic incorporate an additional benzene ring [C_6H_6] in comparison with the saturated. They are liquid or solids under surface conditions. *Resins* and *asphaltenes*, that are quite operational definitions, are high-molecular-weight polycyclic components, comprising atoms of nitrogen, sulfur and oxygen, that are part of the petroleum's non-volatile fraction, also denominated residue. Asphaltenes are insoluble in pentane, hexane or heptane and soluble in aromatics; their molecules are very large and the hydrogen/carbon ratio is very low (0.9 to 1.2). They have a complex structure, made up of aromatic nuclei, saturated cyclic rings and aliphatic chains. They usually contain S, N and O heteroatoms. Resins are compounds with chemical structure similar to asphaltenes, however, in general, they are less polar, their molecular weight is lighter, with less aromatics. The hydrogen/carbon ratio is low (1.2 to 1.7), and they are more soluble in n-pentane [C_5H_{12}] or in n-heptane [C_7H_{16}], however insoluble in liquid propane (Zilio & Pinto, 2002).

Hydrocarbon's spectral behavior and physicochemical properties

As summarized in Winkelmann's work (2005), Carbon-Hydrogen bonds (C-H, C-H₂ and C-H₃), hydroxyl groups (O-H), double and triple bonds of asphaltenes and aromatics, carboxyl groups (C=O), ethers (C-O-C), amino groups (N-H) and other structural groups of organic chemicals exhibit characteristic fundamental vibrations that are evident in the (4,000 cm^{-1}) 2.5 μm to (1,500 cm^{-1}) 6.67 μm wavelength region. Furthermore, they

Table 1 – Hydrocarbon fundamental vibrations. Source: Cloutis (1989) and Silverstein & Webster (1998).

Compound		Type of deformation	Vibrations	
			cm ⁻¹	μm
[1]	C-H	Axial deformation of C-H (Alkanes and Aromatics)	3032	3.30
[2]	CH ₃	Methyl asymmetric axial deformation (Alkanes)	2950	3.39
[3]	CH ₂	Methylene asymmetric axial deformation (Alkanes)	2920	3.42
[4]	CH ₃	Methyl symmetric axial deformation (Alkanes)	2875	3.48
[5]	CH ₂	Methylene symmetric axial deformation	2850	3.51
[6]	C=O	Carbonyl axial deformation	1700	5.88
[7]	C	C axial deformation (aromatic ring)	1600	6.25
[8]	CH ₂ and CH ₃	Methyl asymmetric angular deformation + Methylene symmetric deformation (Alkanes)	1450	6.90
[9]	CH ₃	Methyl symmetric angular deformation (Alkanes)	1375	7.27
[10]	C-H	Angular deformation out of the plane of the C-H ring bonds (Mononuclear and Polynuclear Aromatics)	900 to 675	11.1 to 14.8
[11]	C-H	Angular deformation out of the plane of the C-H in naphthenes	862 to 735	11.6 to 13.6

exhibit overtone and combination tones in the visible infrared spectrum between 0.5 μm (20,000 cm⁻¹) and 6.67 μm (1,500 cm⁻¹), as can be seen in Table 1 and Table 2.

Table 2 – Overtones and combinations of organic molecules. Source: Cloutis (1989).

Overtones and combinations	Wavelength (μm)
2*[1]	1.65
[1]+[3]	1.68
2*[2]	1.69
[1]+[6]	1.7
2*[3]	1.71
[2]+[4]	1.72
[3]+[3]	1.73
2*[4]	1.74
2*[5]	1.75
[5]+[6]	2.2
[3]+[7]	2.21
[5]+[7]	2.25
[2]+[8]	2.27
[3]+[8]	2.29
[2]+[9], [4]+[8]	2.31
[5]+[8]	2.33
[4]+[9]	2.35
[5]+[9]	2.53

In relation to the hydrocarbon spectral behavior, the spectrum regions can be approached based on the following intervals: (i) from 11 to 25 μm the fundamental bands of the aromatic structures and heavy atom bonds are observed; (ii) between 7.6 and 11.1 μm there is a complex interaction of vibrations and the identification of structural or functional groups is not possible; (iii) between 4.0 and 6.7 μm the fundamental bands, stretching vibrations of double and triple bonds (C=O, C=C, C=N, C≡C, C≡N, etc.) are registered; (iv) between 2.5 and 4.0 μm the fundamental bands, axial deformation vibrations of light atoms (C-H, O-H, N-H, etc.) are registered, and (v) between 0.8 and 2.5 μm overtones and the fundamental bands' combination tones are observed (Cloutis, 1989; Silverstein & Webster, 1998; Winkelmann, 2005).

Cloutis (1989) concluded that the most promising interval for detection of absorption organic bands is close to 1.7 μm and between 2.2 and 2.5 μm. Furthermore, these are regions of the spectrum where there is high transmittance of the solar radiation (i.e., atmospheric windows) and, consequently, they are more interesting for remote sensing.

MATERIALS AND METHODS

Materials: UNICAMP's oil collection

This work counted on 17 hydrocarbon samples from several Brazilian producing basins that belong to the University of Cam-

Table 3 – Composition of the twelve hydrocarbon samples available for this research.

	1	2	3	4	5	6	7	8	9	10	11	12
API gravity	40.2	29.4	47.2	21.3	36.2	27.4	27.7	28.1	14.4	19.4	13	20
Viscosity at 20°C (mm ² /s)	5.821	21.23	–	185.2	–	84.87	85.72	34.45	81.9	44.93	–	–
Viscosity at 50°C (mm ² /s)	3.09	8.022	1.5	63.83	13.66	20.44	20.4	11.02	12.36	78.53	–	–
Saturated (% m/m)	79.2	50.4	85.9	49.6	81.2	72.7	68.1	55.7	42.5	40.2	35.7	44.9
Aromatics (% m/m)	13.4	28.1	14.1	28.6	6	13.9	17.6	24.3	33.1	33.3	24.6	32.1
Resins (% m/m)	7.4	19.7	<0.1	20.04	12.8	13.35	14.3	19.1	22.28	23.4	32.4	20.6
Asphaltenes (% m/m)	<0.5	1.8	<0.1	1.76	<0.5	<0.5	<0.5	0.9	2.12	3.1	7.3	2.4
Sulfur (% m/m)	0.0928	0.32	0.0062	0.47	0.0402	0.18	0.153	0.373	0.64	0.767	0.31	0.68

*Replaced by a measurement at 30°C when measurements at 20° are absent.

pinas (UNICAMP) collection. Twelve of these samples have a known composition (Table 3). For the rest of the samples, only the API gravity is available: sample 13 = API gravity 21; sample 14 = API gravity 25.6; sample 15 = API gravity 36; sample 16 = API gravity 19.2 and sample 17 = API gravity 17.

Methods

Physical (density), chemical and spectroradiometric differentiation of the Brazilian Petroleum

The 17 oil samples were measured for their spectral reflectance under different water saturations, acquisition/illumination thicknesses and geometries. The spectra were acquired using an Analytical Spectral Devices (ASD) FieldSpec[®] Pro spectroradiometer, operating over the 350-2500 nm interval, with ultraspectral resolution (>2000 bands). Measurements of pure oil samples, as well as of oil on water were made, and the basic characteristics of an offshore seep were observed. When dealing with the pure oil sample, it was spilled on a Petri plate. In order to undertake measurements of the oil films on water, the oil was deposited on sea water also on the Petri plate. The sea water samples utilized were collected at the Copacabana beach (Rio de Janeiro – RJ), and salt water was prepared at the laboratory with the *Instant Ocean Sea Salt* (Aquarium Systems Inc., Sarrebourg France). The amount of oil was enough for an oil film 127 μm thick. For both cases, the edges of the Petri plates were covered with a cardboard painted

with an amorphous paint, exempt of spectral response, avoiding any spectral interference in the oil measurements (more details in Lammoglia & Souza Filho, 2011).

Figure 1, presents the representative spectra of samples with larger (3) and smaller (11) relative amount of saturated, larger (10) and smaller (5) amount of aromatics, larger (11) and smaller (3) amount of resins and larger (11) and smaller (3) amount of asphaltenes.

Chemometrics: Interpretation and Prediction of Petroleum Qualitative Information from their ultraspectral signature

The differences among the spectra, referring to the sample's chemical composition, can hardly be pointed out without the aid of statistical or mathematical methods. Such difficulty is due to the superimposition of several individual peaks (overtones and combination bands) that compose the hydrocarbon's reflectance spectrum in the near and short wave infrared interval (NIR and SWIR). Thus, the interpretation of those spectra is quite complex. In order to deal with such difficulties, the choice was to use chemometrics (i.e multivariate statistics), centered on the Principal Component Analysis and Partial Least Square Analysis. Several published papers demonstrate the applicability of that technique for interpretation of transmittance or hydrocarbon absorbance spectra (e.g. Hidajat & Chong, 2000; Kallevik et al.,

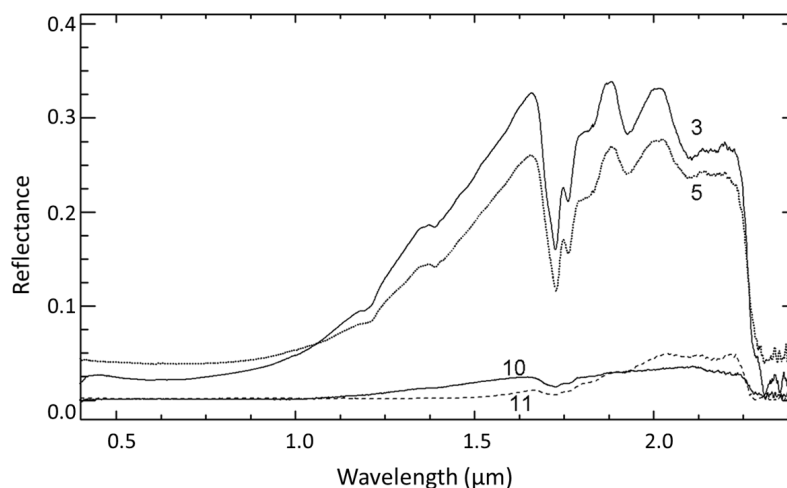


Figure 1 – Reflectance spectra of hydrocarbon samples (samples 3, 5, 10 and 11) measured at the laboratory conditions.

2000; Aske et al., 2001; Hannisdal et al., 2005; Pasquini & Bueno, 2007), but few mention the reflectance spectra and, especially, with remote sensing purposes.

Partial least square analysis (MQP) is a technique combining characteristics of the principal component analysis with multiple regression, aiming to relate response variations (Y variables) with variations of the prediction parameters (X variables). This regression type correlates components of the X variables, in this case the reflectance values, that are also relevant for the predicted (Y) variables (i.e. API gravity or SARA composition). In this process, are found Scores and Loadings that simultaneously decompose X and Y with the premise of explaining in its full extent the covariance among these variables. Afterwards, a regression between the latent variables and the response is carried out. In order to avoid the regression overfitting, it is important to find the best number of latent variables, which is usually made using a procedure called cross validation, in which the minimum prediction error is determined (Aske et al. 2001; The Unscrambler® 2012; LAQQA 2012).

The purpose of using this technique is calculation of the unknown response value, for instance, API gravity, using a regression model. In addition, it is appropriate for correlated and/or redundant data, such as spectral data, and for non trivial relationships, such as reflectance spectrum and API gravity or SARA composition.

The computer processing, using the The Unscrambler® software, makes it possible to accomplish the regression for each dependent variable per turn (i.e. saturated, aromatic, resins and asphaltenes separately). In this case, called MQP1, the Y matrix is a column vector. However, it is also possible that all dependent

variables be simultaneously calculated, taking advantage of the correlations and collinearity among the Y variables (SARA).

Errors of the calibration and prediction models may be expressed by RMSE (Root Mean Squared Error). The prediction error (RMSEP) is calculated through:

$$\sqrt{\sum (y_{ip} - y_{ir})^2 \div n} \quad (1)$$

where y_{ip} = value of the Y parameter estimated for sample i ; y_{ir} = reference value for that value and n = number of used samples. This error, that presents the same unit of the used original parameters, can be understood as the average of the prediction errors. Root Mean Squared Error of Calibration (RMSEC) is the measurement of the adjustment and quality of the model based only on the calibration samples' values. Furthermore, the model quality, as well as its predictive capacity, can be evaluated through the R^2 , which represents the square of the coefficient between the predicted and measured values. The R^2 values vary between 0 and 1, desirable values being closer to 1 (Geladi & Kowalski, 1986; Beebe & Kowalski, 1987; The Unscrambler® 2012).

Once registered, the reflectance spectra of the oils measured at the laboratory were resampled for the spectral resolution of orbital sensors, aiming to empirically evaluate the possibility of characterizing the seeps based on data furnished by these sensors. The spectral response value for each interval (i.e. band) to be resampled is calculated based on a Gaussian distribution, centered at the maximum distribution point of each band of the sensor. For resampling, the spectral arrays of the following operating sensors were selected: (i) the Hyperion hyperspectral sensor, on board the Earth Observation-1 satellite (EO-1). This sensor has

Table 4 – Parameters and results of saturated, aromatic, resins and asphaltene prediction derived from the regression by MQP for one dependent variable per turn (MQP1).

		Spectral interval	Number of samples	PCs	RMSEC (calibration)	RMSEP (prediction)	R-square C	R-square P
Pure oil	API	380-2450 nm	51	11	0.312	0.597	0.999	0.997
	Saturated	380-2450 nm	34	11	0.828	2.663	0.997	0.979
	Aromatics	380-2450 nm	32	8	0.674	0.848	0.994	0.990
	Resins	380-2450 nm	31	8	1.089	1.707	0.982	0.957
	Asphaltenes	380-2450 nm	33	12	0.112	0.506	0.997	0.933
Pure oil resampled for Hyperion	API	400-2450 nm	51		1.200	2.576	0.987	0.942
	Saturated	400-2450 nm	34	10	2.335	4.055	0.981	0.945
	Aromatics	400-2450 nm	33	11	1.207	2.325	0.980	0.925
	Resins	400-2450 nm	33	7	1.764	2.452	0.952	0.904
	Asphaltenes	400-2450 nm	36	10	0.266	0.405	0.980	0.958
Pure oil resampled for ASTER	API	9 bands	51		2.653	3.542	0.919	0.839
	Saturated	9 bands	35	9	3.114	4.170	0.963	0.945
	Aromatics	9 bands	31	9	1.871	2.160	0.936	0.936
	Resins	9 bands	32	9	1.401	2.442	0.971	0.934
	Asphaltenes	9 bands	31	7	0.590	0.761	0.916	0.910
Film of oil on water	API	400-2350 nm	51	12	2.502	2.842	0.928	0.912
	Saturated	400-2350 nm	35	12	4.083	4.866	0.940	0.907
	Aromatics	400-2350 nm	34	8	1.812	2.830	0.954	0.902
	Resins	400-2350 nm	34	8	1.820	2.509	0.945	0.907
	Asphaltenes	400-2350 nm	36	7	0.847	1.146	0.745	0.568
Film of oil on water resampled for Hyperion	API	400-2400 nm	51		2.670	3.068	0.919	0.906
	Saturated	400-2400 nm	36	7	4.198	4.922	0.936	0.914
	Aromatics	400-2400 nm	35	6	2.550	3.356	0.911	0.841
	Resins	400-2400 nm	36	10	2.330	2.795	0.907	0.894
	Asphaltenes	400-2400 nm	35	6	0.853	1.159	0.745	0.560
Film of oil on water resampled for ASTER	API	9 bands	51		3.015	3.729	0.894	0.845
	Saturated	9 bands	32	9	5.150	6.662	0.899	0.826
	Aromatics	9 bands	29	9	2.892	3.966	0.847	0.719
	Resins	9 bands	34	9	2.365	3.265	0.905	0.854
	Asphaltenes	9 bands	34	9	0.617	0.691	0.581	0.462

220 contiguous spectral bands between 400 and 2500 nm (visible (VIS), near infrared (NIR), and short wave infrared (SWIR)), with band width of 10 nm (Pearlman et al., 2000); (ii) the multi-spectral Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) sensor, on board the EARTH platform, which offers a total of 14 bands, among them nine in the VIS-NIR-SWIR (Abrams et al., 2002).

The spectra set was processed both by the regression algorithms by MQP for a dependent variable per turn (MQP1), and for determination of all dependent variables simultaneously (MQP2). In Table 4 the parameters used for processing separately each one of the variables are exposed (saturated, aromatic, resins and

asphaltenes). The counterpart is presented in Table 5, showing the parameters of the concomitant analysis of all variables. For each hydrocarbon sample, triplicates of the spectral response were used.

RESULTS

Results obtained by MQP1 and MQP2 are presented, respectively, in Tables 4 and 5. The results are similar, indicating that the best option between the two cases should be individually evaluated. Values of the prediction error are smaller for models generated with complete spectra. Furthermore, degradation of the models is observed (i.e. increase of the predictive error), as there is a

Table 5 – Parameters and results of saturated, aromatic, resins and asphaltene prediction for regression by MQP for all simultaneously dependent variables (MQP2).

		Spectral interval	Number of samples	PCs	RMSEC (calibration)	RMSEP (prediction)	R-square C	R-square P
Pure oil	API	380-2450 nm	32	14	0.604	1.32	0.996	0.983
	Saturated				1.138	2.442	0.995	0.978
	Aromatics				1.854	2.849	0.988	0.955
	Resins				0.818	1.136	0.989	0.944
	Asphaltenes				0.299	0.378	0.979	0.971
Oil on water	API	400-2350 nm	34	10	2.640	2.910	0.926	0.914
	Saturated				4.131	4.801	0.939	0.924
	Aromatics				3.418	3.944	0.852	0.794
	Resins				2.423	2.754	0.899	0.588
	Asphaltenes				1.162	1.288	0.520	0.423
Pure oil resampled for Hyperion	API	406-2400 nm	35	10	1.05	1.874	0.990	0.970
	Saturated				1.371	3.017	0.994	0.970
	Aromatics				1.576	2.644	0.966	0.909
	Resins				1.123	1.827	0.982	0.953
	Asphaltenes				0.307	0.433	0.976	0.942
Pure oil resampled for ASTER	API	9 bands	29	9	2.35	3.327	0.951	0.904
	Saturated				2.033	3.463	0.984	0.956
	Aromatics				1.448	2.123	0.959	0.922
	Resins				1.764	3.253	0.957	0.845
	Asphaltenes				0.523	0.781	0.936	0.829
Oil on water resampled for Hyperion	API	406-2400 nm	35	10	1.843	2.872	0.966	0.931
	Saturated				4.100	4.804	0.940	0.924
	Aromatics				3.365	3.828	0.856	0.827
	Resins				2.448	2.857	0.897	0.886
	Asphaltenes				1.162	1.315	0.520	0.425
Oil on water resampled for ASTER	API	9 bands	34	9	2.519	3.172	0.938	0.893
	Saturated				5.042	6.497	0.912	0.852
	Aromatics				3.569	4.552	0.835	0.744
	Resins				2.703	3.830	0.888	0.769
	Asphaltenes				1.169	1.666	0.631	0.241

decrease of the spectra resolution due to the resampling process for the sensors selected herein.

When simulating the spectra of pure oil resampled for the Hyperion and ASTER sensors, a degradation of the predictive models is noticed, as pointed out by the RMSE increase. However, the results indicate that data from such sensors make possible estimates for the components evaluated herein. For the oil on water case, although the models are less robust, when compared with the pure oil models, there is also the possibility of inferences about the hydrocarbon composition.

The variable loadings in the regression models by MQP indicate that, for models generated with pure oil spectra, the most

important variables concentrate between 1600 and 2400 nm. Such evaluation agrees with the visual analysis of the spectral responses from the HCs, once their main diagnosis features are concentrated in this interval. The loadings of the models generated with the reflectance spectra of oil films on water indicate that the region between 1000 and 1400 is also relevant.

These results are comparable to those reported by Aske et al. (2001), which were obtained with absorbance spectroscopy from 18 samples for the interval between 780 to 2500 nm. When researching, Aske et al. (2001) found, respectively, RMSEC and RMSEP of 1.13 and 2.78 for saturated, 0.85 and 2.39 for aromatic, 0.54 and 1.14 for resins, and 0.29 and 0.98 for asphaltenes.

When dealing with the API gravity prediction, the results found in this work are comparable with those found by Pasquini & Bueno (2007), where, based on the absorbance spectra of 122 samples (1000 to 2500 nm), the authors found RMSEP between 0.24 and 0.43.

CASE STUDY: APPLICATION TO REAL SITUATION

With the objective of verifying the functionality of the estimates generated from laboratory samples in a real and operational situation, a study was carried out on a seep that occurred in November 2004 in the Campos Basin, which was registered by the ASTER sensor, as pointed out by previous works (Lorenzetti et al., 2006; Bentz et al., 2007) (Fig. 2).

The Campos Basin's hydrocarbons have API gravity varying between 17-37 and they represent a mixture of biodegraded and non-biodegraded oils. The study area, when plotted on a map with the Campos Basin reservoirs (Fig. 2), is located on the Marlim Sul giant field, which contains oil with API gravity ranging between 17 and 24 and the majority of the oil produced has the API gravity between 19 and 22. The Marlim Sul oil composition ranges between 40-49% mass/mass of saturated, 33-25% m/m of aromatics, 20-28% m/m of resins, 3-1.5% m/m of asphaltenes (Mello et al., 1994; Borin, 2003; CETESB, 2003; Lima et al., 2007; Gama, 2008).

ASTER Data – characteristics and preprocessing

Sensor ASTER has three independent imaging subsystems that cover the following regions of the electromagnetic spectrum: VIS-NIR – three bands with 15-meter spatial resolution, SWIR – 6 bands with 60-meter spatial resolution, and TIR – with 90-meter spatial resolution. Each scene covers approximately 60 × 60 km (Abrams et al., 2002). The ASTER sensor's SWIR bands are invariably affected by the cross-talk effect (Iwasaki et al., 2001). This effect was corrected with the software developed by Iwasaki & Tonooka (2005). In order to correct the atmospheric effects, the MODTRAN4 model of radioactive transfer was used, implemented in the Atmospheric Correction Now (ACORN) (Imspec, 2001) software.

Interpretation and prediction of petroleum qualitative information as a function of their multispectral signature (ASTER)

When searching for the seep's representative spectra, the option was to extract them as mean spectra from small sub-regions within the seep's limits and registered in the ASTER image's

pixels converted into reflectance within the VNIR-SWIR interval (nine bands). Thus, mean spectra from 10 small sub-regions of the feature were selected. Each sub-region comprises from 3,000 to 6,000 pixels. Figure 3, shows an attempt of comparing these spectra amongst themselves, with three representative spectra extracted from pixels of the ASTER scene positioned in the following regions (i) near the initial point of the seep, in other words, up to 4 km from the source, (ii) at an intermediate distance from the source, in other words, between 4 and 23 km distant from the source, and (iii) distant from the source, in other words, a distance greater than 23 km from the initial point of the seep.

The spectra registered at the laboratory together with those derived from the scene were normalized by the removal of continuous (Fig. 3b). Such normalization comprises removal of the albedo variation from the spectral curve by a polynomial that describes it, leading to the absolute enhancement of the spectral features that are diagnostic of hydrocarbons, facilitating comparisons among spectra (e.g. Van der Meer, 2004).

Considering the normalized spectra, a regression model was generated by MQP, based on spectral measures from several angles. A total of 68 spectra, representative of 12 oil on water samples were used, with the premise of containing at least three spectra representative of each oil type. Once processing had begun, outliers were detected and excluded.

The regression models by MQP were also individually produced for each variable (MQP1) and for all variables simultaneously (MQP 2), aiming at characterization of the seeped oil in a real situation. Results are presented in Table 6. As can be observed in the table, data from the ASTER sensor were enough for the Campos Basin's seep remote qualitative characterization. The standard deviation values can be reduced using more samples when building the models. The estimates generated by the predictable model are comparable to the real values of the oil produced in the Marlim Sul field. Thus, the results obtained in this study case confirm that the proposed methodology was successful, ratifying the possibility of obtaining remotely inferred API gravity and SARA composition.

DISCUSSION

Only 12 samples were used for the regression models by MQP, except in the MQP1 case for the API gravity determination, that counted with 5 additional samples, totaling 17. The universal application of those models in any area, should ideally comprise a great variety of known hydrocarbons. In this context, the number of samples used for the accomplishment of this work represents only a portion of the hydrocarbon diversity. As an extreme case,

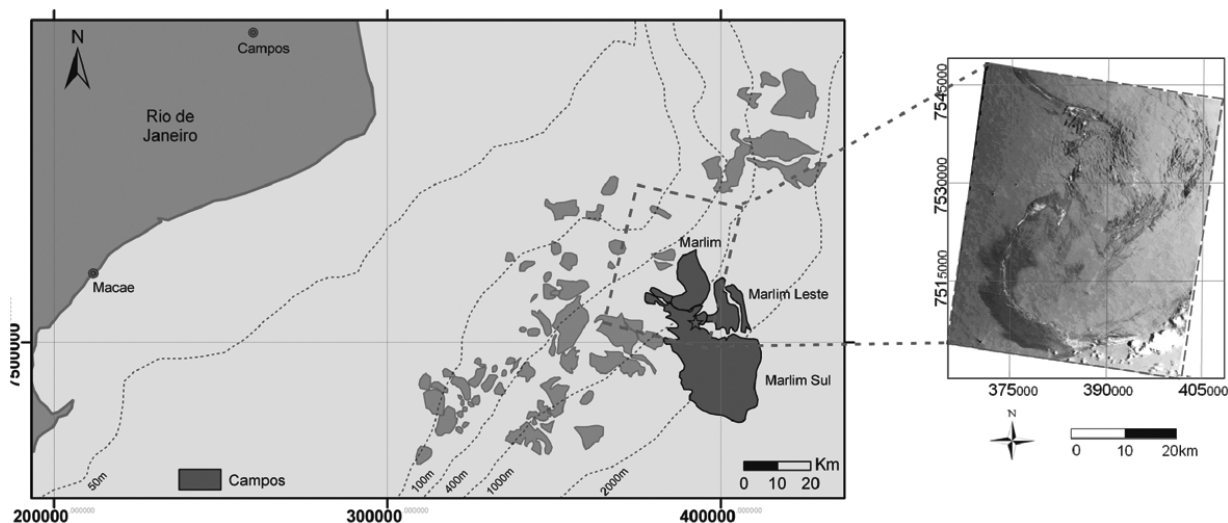


Figure 2 – Study Area: Campos Basin – on the State of Rio de Janeiro coast. The star indicates the seep focal point. The broken rectangle indicates the area comprised by the ASTER scene (November 2004) that registered the seep (base map adapted from Milani & Araújo, 2003).

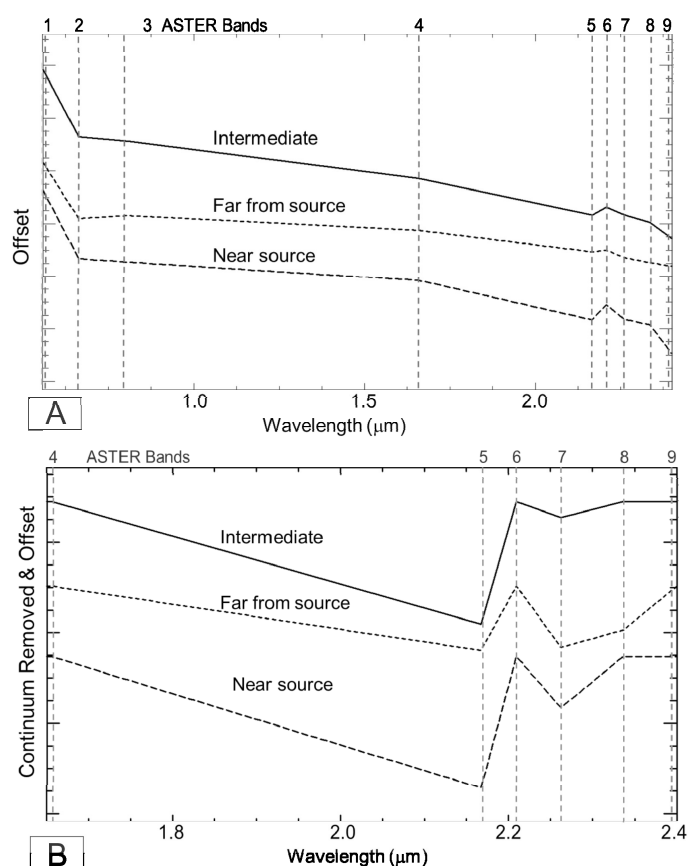


Figure 3 – Spectra extracted from pixels of the ASTER scene related with the hydrocarbon seeps from the Campos Basin. In (a) and (b) three representative spectra extracted from pixels (i) near the seep source, (ii) at an intermediate distance in relation to the source, and (iii) distant from the source. In (a) complete spectra are presented and in (b) the same spectra are presented normalized by the removal of continuous in the NIR-VNIR interval.

the difficulty to determine the % (m/m) for asphaltenes was noticed, once samples used in this research present a smaller diversity of such component.

Table 6 – Results of the characterization of the oil seep in Campos basin through MQP regression. Representative values of the hydrocarbons produced at Marlim Sul are also presented.

	MQP 1	MQP 2	Example of real values from the Marlim Sul Field
API	19.6	18.61	19 to 22
Deviation	1.9	2.8	
Saturated	45.38	43.97	44
Deviation	8.41	9.80	
Aromatics	26.91	29.20	33
Deviation	4.85	6.22	
Resins	24.61	24.07	22
Deviation	2.60	4.77	
Asphaltenes	2.14	2.78	1.5
Deviation	0.53	1.25	

The presented models, generated from laboratory reflectance data (2150 bands) are comparable to those reported in similar works, but based on absorbance data (i.e. Aske et al., 2001). The method was also functional considering measures based on thin films of oil on sea water. This confirms the method potential for characterization and estimation of oil samples, in accordance to determinations commonly made by remote sensing of the Earth resources (i.e., those that do not measure the target's absorbance).

Models based on simulated spectra for the Hyperion (22 bands) and ASTER (9 bands) orbital sensors, although subsampled in relation to spectra measured at the laboratory within the VNIR-SWIR interval, were equally successful in the characterization of the investigated oils, which also demonstrates the potential operational advantage of this methodology.

Taking into consideration the calibration or forecast errors (RMSE), it may be seen that models generated from films of oil on water spectral reflectance curves are less robust, and the same happens when curves are resampled for the orbital sensors' spectral resolution. However, in spite of the robustness loss, models are sensitive to the spectral patterns and they make possible inferences about the hydrocarbon types and characteristics.

In general, models generated from MQP regression for one dependent variable per turn (MQP1) show smaller calibration and forecast errors than those obtained simultaneously for all variables (MQP2). However, in this work, it was demonstrated that both models make possible remote inferences about the seep composition.

Results obtained for the Campos Basin case study, based on ASTER data, show how successful the method was for remote qualification of several chemical and physical attributes of the seeped oil. However, it should be emphasized that the model was generated with hydrocarbons from Brazilian producing basins, resulting in a better calibration for the study area. For other areas, the suggestion is to use a larger number of samples for model creation, and sensors with higher spectral resolution or with more bands positioned in the spectral regions where the hydrocarbon responses are diagnostic and more accentuated.

CONCLUSIONS

This work permitted the API gravity and SARA composition characterization of hydrocarbon samples from their spectral reflectance curves obtained at the laboratory conditions and processed by statistical methods. Based on the methodology developed and described herein, spectral reflectance data from samples of pure oil and samples of films of oil on water, resampled or not for the Hyperion and ASTER sensors' spectral resolution, made possible the API gravity estimate, as well as the hydrocarbon's percentage (% m/m) of saturated, aromatic, resins and asphaltenes in the samples. Thus, the laboratory experiments and simulations for orbital sensors indicate the possibility of using this methodology for remote inferences of the seep's composition – a fact without precedents in remote sensing research.

Data processing indicated that regression by MQP for one dependent variable (MQP1) individually presents smaller associated errors when compared with the same regression with all variables simultaneously processed (MQP2). However, in the case study, results obtained from the MQP2 model present more similarity with real value examples from the Marlim oil field. Thus, the use of model MPQ2 is advised for making decisions, since its processing is more agile and capable of evaluating simultaneously all variables.

Results obtained studying a real case in Campos Basin based on ASTER data, were considered quite credible, since they made possible the remote characterization of an oil seep, indicating that the methodology was successful. The potential of this methodology may be equally favorable for oil exploration and environmental monitoring, allowing the oil type and status recognition excluding an extensive on-site verification.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Wilson José Oliveira of Petrobras for help with data assemblage, and Prof. Dr. Jarbas J.R.

Rohwedder who assisted with the use of the Unscrambler software. CRSF acknowledges the CNPq for the Productivity in Research Scholarship.

REFERENCES

- ABRAMS M, HOOK S & RAMACHANDRAN B. 2002. ASTER User Handbook. Version 2. Jet Propulsion Laboratory. 135 pp.
- ALBOUDWAREJ JH, FELIX J & TAYLOR S. 2006. Highlighting Heavy Oil. *Oilfield Review*, 34–53. Available on <http://www.slb.com.br>. Accessed on June 01, 2012.
- ASKE N, KALLEVIK H & SJÖBLOM J. 2001. Determination of Saturate, Aromatic, Resin, and Asphaltenic (SARA) Components in Crude Oils by Means of Infrared and Near-Infrared Spectroscopy. *Energy & Fuel*, (15): 1304–1312.
- BEEBE KR & KOWALSKI BR. 1987. An introduction to Multivariate Calibration and Analysis. *Anal Chem.*, 59(17): 1007A–1017A.
- BENTZ CB, LORENZZETTI JA, KAMPEL M, POLITANO AT & LUCCA EVL. 2007. Contribuição de dados ASTER, CBERS, R99/SIPAM e OrbISAR-1 para monitoramento oceânico – Resultados do Projeto FITOSAT. In: *Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, SC, Brasil, INPE*, p. 3755–3762.
- BORIN JE. 2003. Estudo Geoquímico de óleo do Campo de Marlim, Baía de Campos, Síntese de Biomarcadores Esteranos Aromáticos. M.Sc. dissertation. Instituto de Química, Universidade Estadual de Campinas, SP, 139 pp.
- CETESB (COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO). 2003. Relatório Técnico – Atendimento 178/03. Operação TEBAR VII – Píer Sul. Available on <http://www.cetesb.sp.gov.br/emergencia/acidentes/vazamento/publicacoes/13.pdf>. Accessed on June 01, 2012.
- CLOUTIS EA. 1989. Spectra Reflectance Properties of Hydrocarbons: Remote Sensing Implications. *Science*, 245: 165–168. Washington: American Association for the Advancement of Science.
- GAMA R. 2008. Caracterização Geoquímica de Óleos Biodegradados dos Campos de Marlim Sul e Roncador, na Baía de Campos. M.Sc. dissertation, Universidade Federal do Rio de Janeiro, RJ, 188p.
- GELADI P & KOWALSKI BR. 1986. Partial Least-Square Regression: A Tutorial. *Anal. Chim. Acta*, 185: 1–17.
- HANNISDAL A, HEMMINGSEN PV & SJÖBLOM J. 2005. Group Type Analysis of Heavy Crude Oils Using Vibrational Spectroscopy in Combination with Multivariate Analysis. *Ind. Eng. Chem. Res.*, 44: 1349–1357.
- HIDAJAT K & CHONG SM. 2000. Quality Characterization of crude oils by partial least square calibration of NIR spectral profiles. *J. Near Infrared Spectrosc.*, 8: 53–59.
- HUNT GR. 1979. *Petroleum Geochemistry and Geology*. W.H. Freeman and Company, San Francisco, 617 pp.
- IMSPEC. 2001. *ACORN User's Guide*. Analytical Imaging and Geophysics. Boulder, CO, USA, 64 pp.
- IWASAKI A, FUJISADA H, AKAO H, SHINDOU O & AKAGI S. 2001. Enhancement of spectral separation performance for ASTER/SWIR. *Proceedings of SPIE, The International Society for Optical Engineering*, 4486: 4-2-50.
- IWASAKI A & TONOOKA H. 2005. Validation of crosstalk correction algorithm for ASTER/SAWIR. *IEEE T Geosci Remote Sensing*, 43(12): 2747–2751.
- KALLEVIK H, KVALHEIM OM & SJÖBLOM J. 2000. Quantitative Determination of Asphaltenes and Resins in Solution by Means of Near-Infrared Spectroscopy: Correlations to Emulsion Stability. *J. Colloid Interf. Sci.*, 225: 494–504.
- LAQQA (LABORATÓRIO DE QUIMIOMETRIA EM QUÍMICA ANALÍTICA). 2012. Available on <http://laqqa.iqm.unicamp.br/>. Accessed on June 01, 2012.
- LIMA AA, FARAH MA & RAJAGOPAL K. 2007. A Viscosidade De Óleo Cru. In: *4o PDPETRO (ABPG), Campinas, SP 3.1.0174-1–1-3.1.0174-1–8*.
- LAMMOGLIA T & SOUZA FILHO CR. 2011. Spectroscopic characterization of oils yielded from Brazilian offshore basins: Potential applications of remote sensing *Rem. Sen. of Environment*, Volume 115, 10(17): 2525–2535.
- LAMMOGLIA T & SOUZA FILHO CR. 2012. Mapping and characterization of the API gravity of offshore hydrocarbon seepages using multispectral ASTER data. *Rem. Sen. of Environment*, 123: 381–389.
- LORENZZETTI JA, KAMPEL M, BENTZ CM & TORRES Jr AR. 2006. A Meso-scale Brazil current frontal eddy: Observations by ASAR, RADARSAT-1 complemented with visible and infrared sensors, *in situ* data and numerical modeling. In: *SEASAR, Frascati, Italy, SP-613*.
- MELLO MR, KOUTSOUKOS EAM, MOHRIAK WU, BACOCOLI G. 1994. Selected petroleum systems in Brazil, In: *MAGOON LB & DOW WG (Eds.). The petroleum system – From source to trap, AAPG Memoir 60*, p. 499–512.
- MILANI EJ & ARAÚJO LM. 2003. Recursos Minerais Energéticos: Petróleo. In: *BIZZI LA, SCHOBENHAUS C, VIDOTTI RM & GONÇALVES JH (Eds.). Geologia, Tectônica e Recursos Minerais do Brasil (pp. 541-576), CPRM, Brasília, BR*.
- PASQUINI C & BUENO AF. 2007. Characterization of petroleum using near-infrared spectroscopy: Quantitative modeling for the true boiling point curve and specific gravity. *Fuel*, 86: 1927–1934.
- PEARLMAN J, SEGAL C, LIAO L., CARMAN S, FOLKMAN M, BROWNE B & UNGAR S. 2000. *Development and Operations of the EO-1 Hyperion Imaging Spectrometer*. 11p. Available on

- <http://eo1.gsfc.nasa.gov/miscPages/TechForumOther/Yperion%20SPIE%20Publication.pdf>. Accessed on Dec. 01, 2007.
- SILVERSTEIN RM & WEBSTER FX. 1998. Infrared Spectroscopy In Spectrometric Identification of Organic Compounds. New York, Ed. Wiley. pp. 70–105.
- THE UNSCRAMBLER. 2012. Available on <http://www.camo.com/>. Accessed on June 01, 2012.
- THOMAS EJ. 2001. O Petróleo In: Fundamentos de Engenharia de Petróleo. Editora Interciência 2ed., pp. 1–13.
- TISSOT BP & WELTE DH. 1984. Petroleum formation and occurrence, Heidelberg, 699p.
- VAN DER MEER F. 2004. Analysis of spectral absorption features in hyperspectral imagery. *Int. J. Earth Observ.*, 5(1): 55–68.
- WETTLE M, DANIEL PJ, LOGAN GA & THANKAPPAN M. 2009. Assessing the effect of hydrocarbon oil type and thickness on a remote sensing signal: A sensitivity study based on the optical properties of two different oil types and the HYMAP and Quickbird sensors. *Remote Sens. Environ.*, 113: 2000–2010.
- WINKELMANN KH. 2005. On the applicability of imaging spectrometry for the detection and investigation of contaminated sites with particular consideration given to the detection of fuel hydrocarbon contaminants in soil. Doctore thesis. Universidade de Cottbus – Alemanha, 236 p.
- ZÍLIO EL & PINTO UB. 2002. Identificação e Distribuição dos Principais Grupos de Compostos Presentes nos Petróleos Brasileiros. *Boletim Técnico. PETROBRAS*, Rio de Janeiro, 45(1): 21–25.

NOTES ABOUT THE AUTHORS

Talita Lammoglia. Ph.D. (2010) and B.Sc. (2006) in Geology (Unicamp). International fellowship at the Institute of Geology of the University of Missouri/Columbia (USA) (2005). Worked with mineral exploration (Yamana Gold Inc.) and modeling of oil reservoirs (Schlumberger Serviços de Petróleo). Awarded twice with the National Prize of Technology in Petroleum Exploration of Petrobras (2006-undergraduate student/2008-master's degree) and Prize of Better Work of Scientific Initiation at the XIII Brazilian Symposium of Remote Sensing, in Florianópolis (2007). Since 2008 works for Petróleo Brasileiro S.A. – Petrobras, in the field of applied geology to oil exploration.

Carlos Roberto de Souza Filho graduated on Geological Engineer (Universidade Federal de Ouro Preto, 1988). He holds a Master degree in Metalogenesis (Universidade Estadual de Campinas – UNICAMP, 1991) and a Ph.D. on Earth Sciences/Remote Sensing (The Open University, England, 1995). At UNICAMP, he became an Assistant Professor in 1997, an Associate Professor in 2002 and achieved the Full-Professorship in 2008. He is currently the General Coordinator of the Graduate Research Programs at the Geosciences Institute and the leader of the Research Team on Geotechnologies. He is a Level 1A Researcher at the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Member of the São Paulo Academy of Sciences. He is responsible for several scientific laboratories, including the Reflectance Spectroscopy lab, the FTIR spectroscopy lab and a suite of Image Processing labs at UNICAMP.