UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Geociências

SAEID ASADZADEH

CHARACTERIZATION AND COMPARATIVE STUDY OF ONSHORE HYDROCARBON SEEPAGES AND PETROLEUM PLAYS IN CENTRAL IRAN AND SOUTHEASTERN BRAZIL USING SPECTRAL REMOTE SENSING TECHNIQUES

CARACTERIZAÇÃO E ESTUDO COMPARATIVO DE EXSUDAÇÕES DE HIDROCARBONETOS E PLAYS PETROLÍFEROS EM BACIAS TERRESTRES DAS REGIÕES CENTRAL DO IRÃ E SUDESTE DO BRASIL USANDO SENSORIAMENTO REMOTO ESPECTRAL

CAMPINAS
2018

SAEID ASADZADEH

CHARACTERIZATION AND COMPARATIVE STUDY OF ONSHORE HYDROCARBON SEEPAGES AND PETROLEUM PLAYS IN CENTRAL IRAN AND SOUTHEASTERN BRAZIL USING SPECTRAL REMOTE SENSING TECHNIQUES

CARACTERIZAÇÃO E ESTUDO COMPARATIVO DE EXSUDAÇÕES DE HIDROCARBONETOS E PLAYS PETROLÍFEROS EM BACIAS TERRESTRES DAS REGIÕES CENTRAL DO IRÃ E SUDESTE DO BRASIL USANDO SENSORIAMENTO REMOTO ESPECTRAL

THESIS PRESENTED TO THE INSTITUTE OF GEOSCIENCES OF THE UNIVERSITY OF CAMPINAS TO OBTAIN THE DEGREE OF DOCTOR OF PHILOSOPHY IN SCIENCES IN THE AREA OF GEOLOGY AND NATURAL RESOURCES.

TESE APRESENTADA AO INSTITUTO DE GEOCIÊNCIAS DA UNIVERSIDADE ESTADUAL DE CAMPINAS PARA OBTENÇÃO DO TÍTULO DE DOUTOR(A) EM CIÊNCIAS NA ÁREA DE GEOLOGIA E RECURSOS NATURAIS.

ORIENTADOR: PROF. CARLOS ROBERTO DE SOUZA FILHO

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO/TESE DEFENDIDA PELO ALUNO SAEID ASADZADEH E ORIENTADA PELO PROF. DR. CARLOS ROBERTO DE SOUZA FILHO

CAMPINAS

2018

Agência(s) de fomento e nº(s) de processo(s): FAPESP, 2015/06663-7
ORCID: https://orcid.org/0000-0001-9043-075

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Geociências
Marta dos Santos - CRB 8/5892

As11c Asadzadeh, Saeid, 1978-
Characterization and comparative study of onshore hydrocarbon seepages and petroleum plays in central Iran and southeastern Brazil using spectral remote sensing techniques / Saeid Asadzadeh. – Campinas, SP : [s.n.], 2018.

Orientador: Carlos Roberto de Souza Filho.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Geociências.

1. Exsudação. 2. Petróleo - Prospecção. 3. Plataforma continental. 4. Sensoriamento remoto. 5. Análise Espectral. I. Souza Filho, Carlos Roberto, 1965-. II. Universidade Estadual de Campinas. Instituto de Geociências. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Caracterização e estudo comparativo de exsudações de hidrocarbonetos e plays petrolíferos em bacias terrestres das regiões central do Irã e sudeste do Brasil usando sensoriamento remoto espectral
Palavras-chave em inglês:
Seepage
Petroleum Prospecting
Continental shelf
Remote sensing
Spectral Analysis
Área de concentração: Geologia e Recursos Naturais
Titulação: Doutor em Geociências
Banca examinadora:
Carlos Roberto de Souza Filho [Orientador]
Lenio Soares Galvao
Fernando Pellon de Miranda
Diego Fernando Ducart
Emilson Pereira Leite
Data de defesa: 04-04-2018
Programa de Pós-Graduação: Geociências

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE GEOCIÊNCIAS

AUTOR: Saeid Asadzadeh

Characterization and Comparative Study of Onshore Hydrocarbon Seepages and Petroleum Plays in Central Iran and Southeastern Brazil Using Spectral Remote Sensing Techniques

ORIENTADOR: Prof. Dr. Carlos Roberto de Souza Filho

Aprovado em: 04 / 04 / 2018

EXAMINADORES:

Prof. Dr. Carlos Roberto de Souza Filho - Presidente

Prof. Dr. Diego Fernando Ducart

Prof. Dr. Emilson Pereira Leite

Prof. Dr. Fernando Pellon de Miranda

Prof. Dr. Lênio Soares Galvão

A Ata de Defesa assinada pelos membros da Comissão Examinadora, consta no processo de vida acadêmica do aluno.

Campinas, 04 de abril de 2018.

BIOGRAPHY

Saeid Asadzadeh received the B.Sc. degree in mineral exploration from Sahand University of Technology, Tabriz, Iran, in 2002, the M.Sc. degree in mineral exploration from AmirKabir University of Technology, Tehran, Iran, in 2005, and the Ph.D. degree in geologic remote sensing from the University of Campinas, Campinas-SP, Brazil, in 2018. Between 2005 and 2013, he was involved in numerous exploration activities, and was instrumental in the discovery of several mineral deposits including three previously unknown porphyry Cu-Au mineralizations in magmatic belts of Iran. During his Ph.D., he received two international awards for innovation in petroleum remote sensing from the Geological Remote Sensing Group (UK; 2016) and the American Association of Petroleum Geologists (USA; 2017). His research interests include multi-, and hyperspectral imaging techniques for mineral and hydrocarbon exploration, geologic remote sensing, environmental monitoring, spectroscopy of minerals and petroleum, and spectral analysis. His current research as a postdoc fellow in the Institute of Geosciences, University of Campinas is 'petroleum seepage characterization by multiple wavelengths (VNIR–SWIR–LWIR) spectral remote sensing technology'.

Dedicated to my beloved wife (Roya), my son (Saiman), and my parents

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Carlos Roberto de Souza Filho (Beto), for his time, motivation, patience, and friendship. His continuous support during my research made the four years of living in Brazil enjoyable.

I would specifically like to thank Marcos Nanni for acquiring ground-based hyperspectral data, and Alessandro Batazelli and Javad Bagheri for their companionship in the field. Ricardo Perobelli and Celio Pasquini are thanked respectively for providing me the access to XRD and sisuCHEMA laboratories. Special thanks go to Soheil Pajotan (Hezareh Sevom Co) for his help with the storage and then shipment of the samples to Unicamp. During my research, I received assistance and hints from several professors and colleagues in Brazil and abroad including Jacinta Enzweiler, Rômulo Angélica, Benoit Rivard, Raymond Kokaly, Thomas Cudahy, Aline Belila, Emanuel Hernandez, Edward Cloutis, Joao Paulo Pitombeira, Rebbeca Scafutto, Rosa Pabón, Gustavo Melo, and Luiza Mendes to name but a few. Thank you all for your help.

I thankfully acknowledge the FAPESP, CNPq, and CAPES for research grants, and the exploration directorate of the National Iranian Oil Company for supporting the fieldworks in Iran. I also appreciate the numerous individuals who have supported us personally throughout the years of living in Brazil. In specific, I am thankful to Rene J. Martins and Catharina Nucci who were always keen to help us in difficulties. Many Brazilian friends helped me overcome the language barrier and thought me the basics of the Portuguese language. In particular would like to mention and thank Marcela Paschoal, Raisa Fagundes, Joao Motta, Rosa Pabón, and Samia Passarella.

I must express my profound gratitude to my wife, Roya, for her inspiration and enthusiastic support of my academic pursuits. Lastly, I would like to thank my son, who had the patience to wait until the defense of my thesis to then being born!

ABSTRACT

The objective of this research was to explore for the signatures of seeping hydrocarbons on the surface using spectral remote sensing technology. Two comprehensive literature reviews were produced on the capacity and potential of the technique for direct and indirect seepage detection. Using these theoretical guides, RS was applied to investigate two distinctive test sites located in Iran and Brazil known to retain active microseepage systems and bituminous outcrops, respectively. The Alborz oilfield, located near the city of Qom in Iran, is buried under Oligocene sediments of the Upper-Red Formation. The Anhembi oil-sands, whose outcrops occur near the town of Anhembi, on the eastern edge of the Paraná Basin in Brazil, includes bitumen accumulations in Triassic sandstones of the Pirambóia Formation. The work on the Qom area integrated evidence from (i) petrographic, spectroscopic, and geochemical studies in the laboratory, (ii) outcrop investigations in the field, and (iii) broad-scale anomaly mapping via orbital remote sensing data. Novel mineralogical (in terms of abundance, composition, and crystallinity) and geochemical indicators for seepage characterization were revealed and a classification scheme for microseepage-induced alterations was introduced. The study indicated that active microseepage systems occur in large parts of the lithofacies in the Qom area, implying that the extent of the petroleum reservoir is much larger than previously thought. The Anhembi area was sensed by a multitude of sensors in the VNIR–SWIR range (400–2500nm), at different scales, including WorldView-3 multispectral instrument and AisaFENIX airborne and ground-based hyperspectral imaging system. Spectroscopic characterization of hand samples was accomplished in the laboratory using both imaging (sisuCHEMA) and nonimaging (ASD) instruments. The study demonstrated that a multi-scale spectroscopic approach could provide a complete picture of the variations in the content and composition of bitumen and associated alteration mineralogy. The oil signature, especially the one centered at 2300 nm, proved to be consistent and comparable among observations, and capable of estimating the bitumen content of oil-sands at all imaging scales. Furthermore, WorldView-3 satellite data were proved capable of detecting petroleum by resolving the absorption feature centered at 1700 nm.

Keywords: Seepage, Petroleum Prospecting, Onshore Basin, Remote Sensing, Spectral Analysis.

RESUMO

O objetivo desta pesquisa foi explorar as assinaturas de exsudações de hidrocarbonetos na superfície terrestre usando a tecnologia de detecção remota espectral. Duas revisões abrangentes da literatura foram produzidas sobre a capacidade e potencial da tecnologia para a detecção direta e indireta de exsudações. Com base nessa fundamentação teórica, o sensoriamento remoto foi aplicado para investigar dois locais de teste distintos localizados no Irã e no Brasil, conhecidos por conter sistemas de exsudação ativos e afloramentos betuminosos, respectivamente. O campo petrolífero de Alborz, localizado perto da cidade de Qom, no Irã, encontra-se encerrado por sedimentos oligocênicos da Upper-Red Formation. As areias betuminosas de Anhembi, cujos afloramentos-tipo ocorrem próximos à cidade de Anhembi, na borda leste da Bacia do Paraná, no Brasil, incluem acumulações de betume nos arenitos Triássicos da Formação Pirambóia. O trabalho na área de Qom integrou evidências de (i) estudos petrográficos, espectroscópicos e geoquímicos em laboratório, (ii) investigações de afloramentos no campo, e (iii) mapeamento de anomalias espectrais em larga escala via dados de sensoriamento remoto orbital. Novos indicadores mineralógicos (em termos de abundância, composição e cristalinidade) e geoquímicos para caracterização da micro-exsudações foram revelados e um novo esquema de classificação para alterações induzidas por micro-exsudação foi introduzido. O estudo indicou que sistemas ativos de exsudação ocorrem em grande parte das litofácies na área de Qom, implicando que a extensão do reservatório de petróleo é muito maior do que o estimado originalmente. A área de Anhembi foi detectada por uma série de sensores na faixa VNIR–SWIR (400–2500 nm), em diferentes escalas, incluindo o sistema multiespectral orbital WorldView-3 e o escanner imageador hiperespectral AisaFENIX, nos modos aerotransportado e terrestre. A caracterização espectroscópica das amostras coletdas no campo foi realizada em laboratório, utilizando sistemas imageadores (sisuCHEMA) e não-imageadores (ASD). O estudo demonstrou que uma abordagem espectroscópica multi-escala pode fornecer uma imagem completa das variações no conteúdo e composição do betume e da mineralogia associada à alteração. A assinatura espectral do óleo, especialmente aquela centrada em 2300 nm, mostrou-se consistente e comparável entre as observações e capaz de estimar o teor de betume das areias betuminosas em todas as escalas de imageamento testadas. Além disso, os dados do satélite WorldView-3 provaram ser capazes de detectar petróleo resolvendo a característica de absorção centrada em 1700 nm.

Palavras-chave: Exsudação, Petróleo Prospecção, Bacia Continental, Sensoriamento Remoto, Análise Espectral**.**

TABLE OF CONTENTS

Chapter 1

Research summary

The objective of this research was to explore for the surficial signatures of seeping hydrocarbons (HCs) using spectral remote sensing technology. In petroleum geochemistry, the manifestation of liquid/gaseous petroleum that is clearly visible to the naked eye is called 'macroseepage' (or simply seepage) and the trace amounts of invisible HCs (C_1–C_5) in soils and sediments is termed 'microseepage'. Whereas an oily seepage could be targeted directly using the spectral signatures of leaking HCs centered at 2300 and 1700 nm wavelengths, a microseepage should be explored indirectly via the mineralogical (and other physio-chemical) footprints of migrating HCs on the strata overlying an accumulation. In this way, remote sensing technique is employed to characterize the depletion, enrichment, or transformation of diagenetic minerals such as clays, carbonates, iron oxides/oxyhydroxides, and sulfides/sulphates in the system using the diagnostic absorption features of the targets in the visible-near infrared (VNIR) and the shortwave infrared (SWIR) wavelengths. The outcomes are then interpreted in the geologic context of the area aiming to interrelate the anomalies to posible subsurface petroleum accumulations. In this research, we intended to investigate both the direct and indirect spectral detection techniques over two distinctive test sites in Iran and Brazil. The first study area is located near the city of Qom, some 100 km to the south of Tehran, Iran. This arid to semi-arid area consists of the Alborz oilfield in a structural trap sealed by a very thick evaporitic sequence of the late Miocene and buried by Oligocene sediments of the Upper-Red Formation (URF). The Anhembi test site, which is located near the town of Anhembi, on the eastern edge of the Paraná Basin, Brazil, conform to an exhumed HC reservoir and includes bitumen accumulations in the early Triassic sandstones of the Pirambóia Formation.

The research was firstly grounded on a comprehensive literature review of macro-, and microseepage approached with remote sensing, which involved overviewing hundreds of scientific articles on the subject and tabulating the relevant case studies. The work benefited from illustrative products generated over two study areas located in the Ventura Basin, State of California, USA and the Tucano Basin, State of Bahia, Brazil, known to host distinctive macro-, and microseepage systems, respectively. This review article is provided in Chapter 2. Furthermore, our attempts to better understand different sorts of spectral processing techniques available to geologic remote sensing gave rise to a second review paper on the subject outlined in Chapter 3.

In the Qom area, our research characterized the footprints of HC microseepage on Upper Red Formation (URF) lithofacies and provided insights into the chemical and physical processes responsible for diagenetic alterations. This was accomplished by integrating evidence from (i) outcrop-scale fieldwork and sampling, (ii) lab-based spectroscopy and analytical geochemistry, and (iii) reservoir-scale alteration mapping using orbital remote sensing data.

In the URF outcrop-scale studies, we developed a good knowledge about the color variations and patterns of diagenetic minerals of the strata and collected a large suite of samples for follow-up laboratory analysis. In the Lab, the samples were measured spectrally using a FieldSpec-4 spectrometer. While performing spectral analysis, we found the current commercial software packages (i.e. the TSG or PRISM) insufficient, and thus developed our own methodology and software

for spectral analysis using the Interactive Data Language (IDL) program. The developed package (>2000 lines of code), named Automated Absorption-based Mineral Spectral Analyzer (AMISA), optimizes the number of bands utilized for curve fitting and retrieves the spectral parameters (e.g. wavelength position and depth) of a given absorption feature within an array of spectra automatically, with no need for user interference. Some technical aspects of the embedded algorithms are described in Chapter 4. AMISA was deployed to analyze the spectral data collected from both study cases in Iran and Brazil. It was particularly utilized to characterize the microseepage-induced diagenetic changes over Alborz petroleum reservoir. The details of the spectroscopic analysis are explained in Chapters 5.

Based on field observation and spectroscopic studies, we nominated ∼75 specimens for various geochemical analysis including Inductively Coupled Plasma-Mass Spectrometry (ICP-MS; n = 22), X-Ray Fluorescence (XRF; n = 35), X-Ray Powder Diffraction (XRD; n = 41), and stable isotope geochemistry (n = 15). Further 25 samples were also prepared for petrographic studies. To give a clear picture of the changes induced in the URF, we contrasted the geochemical, petrographical, and mineralogical aspects of the HC-affected samples with unaltered counterparts from a nearby control area. The results provided novel mineralogical indicators for remote sensing microseepage mapping. Ultimately, the readily available multispectral datasets from ASTER and Sentinel-2 instruments were used to outline the footprints of the microseepage systems. Our study indicated that active microseepage systems occur in large parts of the area implying that the extent of the petroleum reservoir is much larger than previously thought. Chapter 6 presents the findings of these studies.

Following the successful launch of WorldView-3 (WV-3) satellite in orbit, we set out to discover its potentials for direct HC detection aiming to extend the capability of hyperspectral data to orbital remote sensing of global coverage. By using simulated imaging and non-imaging datasets from oil-bearing sediments/samples in the lab/field, we demonstrated that the instrument is capable of detecting oil via its SWIR bands. The results of this study are reported in chapter 7. Following this achievement, DigitalGlobe, the major provider of high-resolution multispectral imageries worldwide, agreed to image our Anhembi test site with WV-3, using the 16 VNIR-SWIR bands sensor arrangment. Fresh data acquired in February 2016 was delivered to us to complement the experiment with a real case study.

In the Anhembi area, the aim of the research was twofold: (i) to delineate the diagenetic alteration mineralogy associated with bitumen emplacement in sandstone beds, and (ii) to detect and quantify the bitumen content of oil-sands by spectral remote sensing technology at multiple operational scales. Accordingly, besides WV-3 data already collected, the oil-sand outcrops/walls were imaged by an AisaFENIX hyperspectral imaging system from the air onboard an airplane and on the ground by mounting the camera on a tripod. During the follow-up fieldworks, we also collected representative samples (n=20) from both impregnated and oil-devoid outcrops of the Pirambóia Formation. In the laboratory, these samples were spectroscopically studied using a FieldSpec-4 spectrometer. A handful of the samples were also scanned using a sisuCHEMA hyperspectral imaging system. Our study indicated that both HC features centered at 1700 and 2300 nm are capable of determining the Total Bitumen Content (TBC) of the samples/outcrops. Even though the feature at 1700 nm yielded higher correlation with the TBC of the sample, unlike the feature at 2300 nm, it was adversely affected by within scene illumination variations and was prone to atmospheric contaminations and confusion with dry vegetation. This multi-scale hyperspectral data acquisition

and processing experiment indicated that HC signatures, especially the one centered at 2300 nm, is consistent and comparable among scales, and upon employing proper calibration data and analysis approach, is capable of estimating the bitumen content of oil-sands at all imaging scales. This case study is covered in Chapter 8.

To further explore the capability of current multispectral sensors comprising ASTER, Sentinel-2, and Landsat-OLI for direct HC detection, we repeated the simulation experiments by resampling the datasets to the response of the noted sensors. We indicated that besides WV-3, ASTER data is also capable of resolving HC feature centered at 2300 nm. However, WV-3 was shown to produce better performance. The results of this study are given in the appendix-2. The concluding remarks and some hints for future studies are provided in Chapter 9.

Overall, this research led to four publications in top remote sensing and geology journals, two submitted manuscripts (Chapter 5 and 8), seven presentations (talk/poster) at national/international conferences, two international awards from Geological Remote Sensing Group (UK) and the AAPG foundation (USA), and a software package (AMISA) for spectral analysis. An additional manuscript (chapter 6) is currently under preparation for submission to the AAPG Bulletin.

Chapter 2

Spectral remote sensing for onshore seepage characterization: A critical overview[1]

In this chapter, we overview the application of spectral remote sensing data collected by multi-, and hyperspectral instruments in the visible-near infrared (VNIR), short-wave infrared (SWIR), and longwave infrared (LWIR) wavelengths for characterization of seepage systems as an exploration indicator of subsurface hydrocarbon (HC) accumulations. Two seepage systems namely macro-, and microseepage are recognized. A macroseepage is defined as visible indications of oil and gas on the surface and in the air detectable directly by a remote sensing approach. A microseepage is defined as invisible traces of light HCs in soils and sediments that are detectable by its secondary footprints in the strata, hence an indirect remote sensing target. Based on these broad categories, firstly, a comprehensive set of well-described and reliable remote sensing case studies available in the literature are thoroughly reviewed and then systematically assessed as regards the methodological shortcomings and scantiness in data gathering, processing, and interpretation. The work subsequently attempts to go through seminal papers published on microseepage concept and interrelated geochemical and geophysical techniques, exhumed HC reservoirs, lab-based spectroscopic analysis of petroleum and other related disciplines from a remote sensing standpoint. The aim is to enrich the discussion and highlight the still unexplored capabilities of this technique in accomplishing exploration objectives using the concept of seepage system. Aspects of seepage phenomenon in environmental pollution and uncertainties associated with their role in global warming are also underlined. This work benefits from illustrative products generated over two study areas located in the Ventura Basin, the State of California, USA and the Tucano Basin, State of Bahia, Brazil known to host distinctive macro-, and microseepage systems, respectively. In conclusion, we recommend further research over a diverse range of seepage systems and advocate for a mature conceptual model for microseepage phenomenon.

2.1. Introduction

A large portion of hydrocarbon (HC) traps is not perfectly sealed and thus, their accumulations leak to the surface over time. When the surface manifestation of oil and gas is clearly visible to naked eye, it is termed as *macroseepage*, whereas the traces of invisible light HCs in near-surface soils and sedimentary rocks (sediments henceforth), which are only detectable by analytical methods and careful geochemical sampling, is called *microseepage* (Horvitz, 1985; Tedesco, 1995). Historically, seepage and HC accumulations have been tied together and, as a result, a large number of the world's oil and gas fields have been explored by drilling in the immediate area of a seep (Hunt, 1996; Yergin, 1992). In modern exploration programs, *macroseeps* are typically regarded as direct clues for the existence of mature source rock(s) and a compelling evidence for the formation of a petroleum system in a given sedimentary basin (Magoon and Beaumont, 1999; Schumacher, 2010), whereas *microseeps*, which are argued to occur in a near vertical fashion over an accumulation, are employed as a targeting tool for petroleum exploration.

Recent investigations have also revealed that seeps are a potent source of methane (with ethane and propane) greenhouse gasses to the atmosphere. It has been estimated that in the natural

[1] A version of this chapter is published in the Earth-Science Reviews journal.

methane budget, seeps are the second most important source of emissions after wetlands. The estimates also reveal that onshore seepages are a more significant emitter of CH_4 than their offshore counterparts (Etiope, 2015; Etiope and Ciccioli, 2009; Etiope and Klusman, 2010; Etiope et al., 2008).

Over the years, a diverse range of techniques, including remote sensing, has been employed for seepage detection. The remote sensing approach holds a great promise for this aim because it is a fast and cost-effective tool applicable to different operational scales for both direct and indirect seepage mapping. In the marine environment, this technology already provides a variety of sensing methods comprising laser florescence, synthetic aperture radar (SAR), and thermal infrared, to name a few (Leifer et al., 2012a). Terrestrial seepage detection, however, has relied heavily on spectral remote sensing collected by multi-, and hyperspectral instruments in the visible-near infrared (VNIR; 0.4–1.0 µm), the short-wave infrared (SWIR; 1.0–2.5 µm), and very occasionally in the longwave infrared (LWIR; 8–14 µm) wavelengths. This methodology has been employed to detect oil and gas seeps in a direct manner and the footprints of HC leakage in sedimentary units indirectly.

Whereas direct detection of HCs is a new research topic conducted mainly by airborne imaging spectroscopy, the indirect HC prospecting commenced with the launch of the first Landsat-MSS multispectral scanner (ERTS) in the 1970s (Simpson, 1978). The advent of the hazy anomaly over some productive/prospective fields was simultaneous to a renewed interest in microseepage concept (Donovan, 1974) and thereupon led to several research studies, including a NASA-Geosat joint project, to evaluate the use of satellite technology for HC prospecting (Abrams et al., 1985). Since then, sporadic case studies have appeared in the literature demonstrating the potentials and premises of this approach for oil and gas exploration (see van Der Meer (2002) and Yang (2000) and references therein) that in comparison to extensively studied offshore cases are almost negligible (Fingas and Brown, 2014; Leifer et al., 2012a).

Despite all the merits of spectral remote sensing, the technique does not yet hold a good place amongst unconventional exploration methods for oil and gas resources nor is acknowledged as a mature exploration tool by explorationists. In this article, we set out to discover the full potential of this state-of-the-art technology for seepage characterization and understand the reasons for which the approach is somehow underutilized by the petroleum industry. To fulfill this aim, we provide a systematic and critical overview of the subject based on well-described and reliable remote sensing case studies reported in the literature and thereupon attempt to address methodological shortcomings and inadequacies in data gathering, processing, and interpretations. In addition, we go through the seminal papers published on microseepage theory and interrelated geochemical and geophysical techniques from a remote sensing standpoint to enrich the discussion and highlight the unexplored capabilities of the technique in accomplishing the objectives of exploration sector. We also attempt to engage the attention of the community to useful case studies conducted over exhumed HC reservoirs as an analogy to depleted traps. In the case of direct seepage detection, we review the few present case studies and contrast their achievements to the findings of interrelated disciplines to underline a wide range of possibilities from spectral products. Lastly, we accentuate the uncertainties about the role of natural seepage in greenhouse gas emission and the possible ways to resolve it. This paper benefits from illustrative products generated over two case studies located in the Ventura Basin, California, USA and the Tucano Basin in Bahia state, Brazil known to host distinctive macro-, and microseepage systems, respectively.

2.2. Petroleum seepage

Surface manifestation of oil and gas can be divided into two broad categories namely macro- and microseepage (Fig. 1). M*acroseepage* is the surface expression of a leakage pathway, typically related to tectonic discontinuities, along which natural liquid or gaseous HCs is (has been) flowing from a subsurface source (Clarke and Cleverly, 1991; Link, 1952; Macgregor, 1993). M*icroseepage*, in contrast, refers to the slow, invisible, but pervasive migration of light alkanes ($C_1 – C_5$) and volatiles from accumulation to the surface (Etiope, 2015; Price, 1986; Schumacher, 1999). A microseepage is not related to faults but it can be enhanced by the presence of faults and large fractures (Richers et al., 1982). Here we emphasize that this division is not certain and in reality, there is a seepage continuum from minute microseeps to visible macroseeps (Etiope, 2015; Schumacher, 1999). Through this paper, '*seeps'* and '*seepage'* are used as generic terms to refer to both of the seepage systems. Furthermore, 'petroleum' and 'hydrocarbon' terms are interchangeably used to denote oil and gas, and any similar organic (natural/artificial) compounds, however, 'oil' is exclusively used to refer to 'crude oil'.

2.2.1 Macroseepage systems

Macroseeps have been classified according to their geological context, underlying tectonic regime, activity manner, migration pathway, leakage rate (intensity), occurrence environment, and associated alteration patterns (Abrams, 2005; Clarke and Cleverly, 1991; Link, 1952; Macgregor, 1993; Thrasher et al., 1996). However, for the purpose of this paper, we only distinguish them to be either *oil* or *gas* seepage, respectively dominated by liquid and gaseous phases. In terms of activity, flowing (active) oil seeps shall be differentiated from paleo-seeps (impregnations) wherein no evidence for their replenishment exists (Macgregor, 1993). Depending on their leaking state, oil seeps can incorporate liquid and solid petroleum, tar, and bitumen.

Some macroseeps have a diffused lateral flow in near-surface fractures thereupon trigger the formation of local alterations around the principle oil/gas seep (Fig. 1). To discriminate such alterations from those associated with a microseepage system (section 2-2), a new term called *miniseepage* has been introduced in the literature (Etiope, 2015). Here we adopt the same terminology to discern between the two seepage-related alterations, but since the final mineral assemblages are rather identical, we discuss them under the same section. Some studies have demonstrated that around 30% of the known macroseeps are associated with surficial anomalies of one type or another (Clarke and Cleverly, 1991), which based on the given definition, can indicate the overall proportion of miniseeps.

2.2.2 Microseepage systems

The long-term leakage of HCs in a microseepage system normally induces an array of diagenetic physio-chemical and mineralogical transformations in the *chimney* column above HC accumulations (Fig. 1). The activity and by-products of bacteria and other microbes are believed to change the pH-Eh of the overlying stratigraphic column and initiate a series of diagenetic changes including (i) biological (microbial/geobotanical) anomalies; (ii) mineralogical alterations; (iii) electrochemical changes and resistivity abnormalities; (iv) magnetic iron oxides and sulfides; and (v) radiation

anomalies (Etiope and Martinelli, 2002; Price, 1986; Saunders et al., 1999; Schumacher, 1996; Tedesco, 1995; Warren, 2012) (Fig. 1).

The mineralogical changes of the classic microseepage model, which is the focus of this review paper, comprises the following patterns (Fig. 1):

Carbonate precipitation: carbonate cements is quite common in microseepage systems. The diagenetic carbonate species detected so far is very diverse and ranges from calcite and ferroan calcite to dolomite, ankerite, siderite, rhodochrosite, and aragonite; though calcite by far is the dominant secondary carbonates (Al Shaieb et al., 1994; Saunders et al., 1999; Schumacher, 1996). The extent of carbonate cementation is not commonly known, albeit, in Cement Field, it was reported to cover an area of about 3.0 km^2 in total (Kirkland et al., 1995). Several studies have indicated that carbon of carbonates can be originated from HC oxidation (Donovan et al., 1974), however, it is not always the case and carbon can originate from other sources like meteoric or Interstitial water (Beitler et al., 2005). The ultimate species of carbonate in diagenetic facies is shown to be related to the concentrations of Fe, Mg, and Ca ion in the system (Al Shaieb et al., 1994).

Bleached red-beds: the discoloration, or bleaching, occurs due to the removal of ferric iron oxides (i.e. hematite) coating from red-beds. This phenomenon, which is chiefly controlled by the fabric of the host rock, is responsible for most of the visual color changes in the chimney column (Donovan, 1974; Schumacher, 1996). The iron discoloration can be due to *bleaching*, which involves chemical reduction and removal of iron from the system, or *enrichment*, which is the secondary precipitation of additional patchy iron oxide cement in the strata (Nielson et al., 2014) (see also Fig. 6e). The bleached ferric iron can ultimately take the following forms: (i) reduction into ferrous species; (ii) partial/total removal from the system; and (iii) re-precipitation as iron oxide concretions (Parry et al., 2009). Whereas some ferrous irons are likely to be carried away from a system, a significant proportion of it can be introduced into lower strata by descending meteoric water. Subsequent oxidation of these iron-rich horizons can bring about ironstones and ferricrete features (Nielson et al., 2014).

Clay formation: clay formation/transformation is mostly related to slightly acidic conditions in the chimney column. Kaolinite, which by far is the prevalent clay in microseepage-induced alterations, is believed to form after the alteration of feldspars or the conversion of then unstable illitic/smectitic clays (see Fig. 6d). This process can liberate potassium (K) from clays, and thus lead to low gamma-ray radiations above HC deposits (Saunders et al., 1993a). The conductive anomalies noted over some chimney column are attributed to the presence of dioctahedral and trioctahedral clays in the system (Warren, 2012).

Sulfide formation: pyrite constitutes the dominant sulfide mineral, though other sulfides like pyrrhotite, marcasite, galena, greigite, and native sulfur are locally abundant (Al Shaieb et al., 1994; Schumacher, 1996). The abundance of pyritic zone depends on several parameters including the abundance of sulfur (S) in the environment. The sulfur itself has been postulated to be originated either from H_2S associated with seeping HCs or from sulfates after bacterial metabolism (Lilburn and Zuhair, 1984). The native sulfur can be the result of hydrogen sulfide oxidation or the reaction of non-methane HCs with anhydrite (Worden et al., 1997). Sulfidization is a rather uncommon process in alteration facies.

The occurrence of pyrrhotite and greigite along with maghemite and magnetite are known to be responsible for anomalous magnetic susceptibility above many chimney columns (Schumacher,

1996). Due to the role of magnetotactic bacteria in forming such magnetic bodies, some have postulated that magnetic anomalies shall approximate the 'fossil' bacterial activities within a system (Foote, 2013; Saunders et al., 1999).

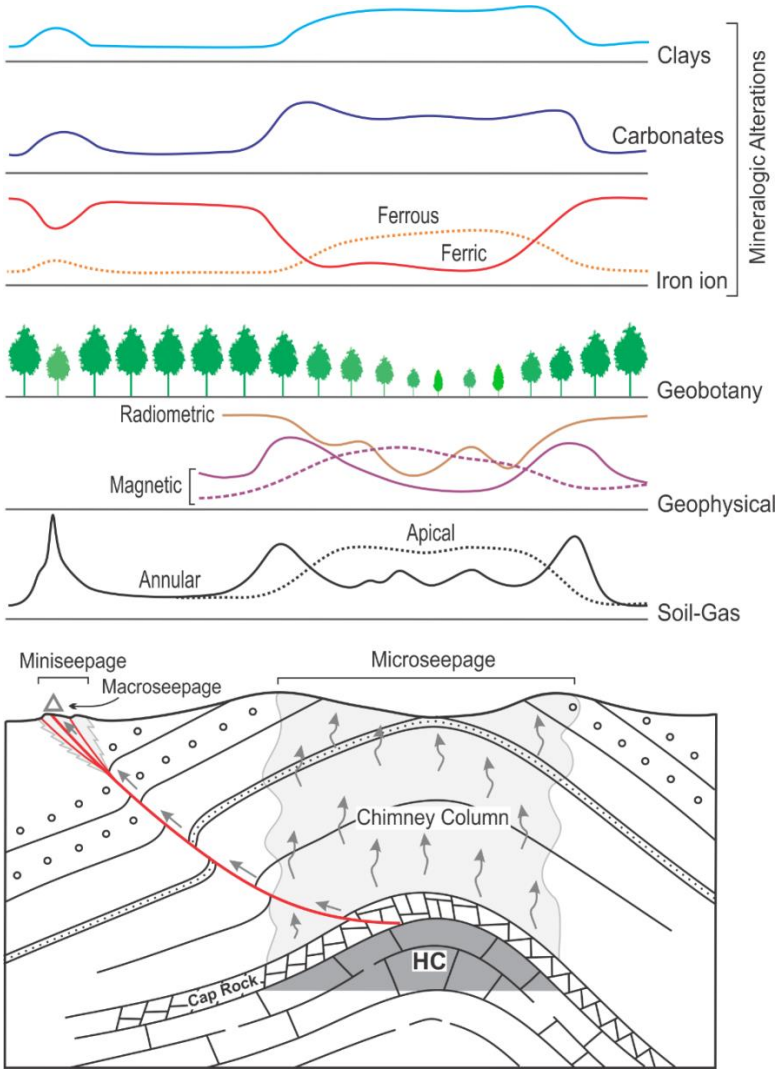

Fig. 1. Schematic representation of microseepage, miniseepage, and macroseepage concepts and associated HC-induced anomalies (adapted and modified partly after (Duchscherer, 1982; Eventov, 2000; Saunders et al., 1993b; Saunders et al., 1999; Schumacher, 1996; van der Werff, 2006)). The HC accumulation and induced alterations are shown by dark-gray and gray shades, respectively. The anomalous signatures encompass geochemical (soil-gas), geophysical, and remote sensing (geobotany and mineralogical alterations) sub-groups.

For decades, the microseepage theory has been controversial among explorationists. However, now there are several (empirical and theoretical) lines of evidence that support this postulation: (i) the C_1 to C_5 composition of the migrating gas that is similar to the gas phase of the pool; (ii) the isotopic content of the seeping HCs that matches that of the underlying reservoir; (iii) the dynamic character of the leakage in responding to reservoir depletion or recharge; (iv) the close conformity of the anomalies with the surface projection of HC reservoir; (v) the bloom of HC-degrading bacteria over accumulations; (vi) the direct evidence provided by well-cutting measurements; and (vii) the *gas chimneys* imaged by high resolution 3D seismic data (Connolly et al., 2013; Donovan, 1974; Duchscherer, 1980; Foote, 2013; Holysh and Toth, 1996; Horvitz, 1980; Horvitz, 1985; Jones and Drozd, 1983; Price, 1986; Tedesco, 1995; Thompson et al., 1994). The concept of chimney cube,

which highlights the chaotic behavior of seismic characters due to gas migration (Aminzadeh et al., 2001), not only is providing a new imaging indicator for HC leakage, but also is bridging the gap between seismic and geochemical surveys and likely is going to end the skepticism and debate on the concept.

The migration processes of gaseous HCs is not well understood, but some computer simulations have shown that pressure-driven continuous gas-phase flow through fractures or the transition from bubble regimes to continuous phase flow can be the plausible mechanism (Brown, 2000; Etiope and Martinelli, 2002). The proposed mechanism was demonstrated, for example, to successfully explain the rapid emerge of gaseous anomalies over gas storage fields.

2.2.3. Macro- versus microseepage

Macroseeps typically migrate laterally for a few to dozens of kilometers through major tectonic discontinuities (Link, 1952; Macgregor, 1993; Thrasher et al., 1996) and thus occur distal to accumulations, whereas microseeps migrate almost vertically and hence are proximal to the pool. Different aspects of macro-, and microseepage systems are compared in Table 1.

Table 1. A comparison between different aspects of micro- and macroseepage systems.

Parameter	Microseepage	Macroseepage
Detection approach	Analytical	Visual
Agent (HC content)	Light HCs (C_1 - C_5), volatiles (aromatics)	Heavy/light crude oil, volatiles, gaseous HCs
Migration mechanism	Continuous gas-phase flow in micro-fractures	Effusion through tectonic discontinuities
Abundance in petroliferous areas	More than 80%	Around 20%
Migration fashion	Mostly vertical	Mostly lateral
Alteration	Yes, extensive (chimney column)	Yes/No, limited (miniseep)
Spatial (areal) extent	Pervasive halo	Point targets
Associated trap	Structural, stratigraphic	Mainly structural
Flux ($mg/m^2/d$)	Tens	Hundreds to thousands
Exploration significance	Indirect	Direct
Types of activity	Active vs fossil (inactive)	Active vs passive Flowing vs impregnations
Targeting potential	Yes	No
Relationship to accumulation	Simple, proximal	Complex, distal

Historically, macroseeps have been the stimulus for early exploration drilling, and statistics reveal that almost all of the world's major oil fields at the beginning of 20th century have been associated with seeps. By 1950s, at least half of the proven reserves were discovered by drilling near macroseeps, whereas by the end of the last century, this proportion plummeted to around 20% (Clarke and Cleverly, 1991; Hunt, 1996; Link, 1952; Macgregor, 1993). On the other hand, more than 80% of the HC reservoirs are associated with microseeps of different manifestations (Schumacher, 2010), and unlike macroseeps that are general basinal scale prospecting indicators (Macgregor, 1993), microseeps have proved to be a targeting tool for subsurface HC accumulations. For example, with surface geochemical methods, 70-82% of subsequently commercial discoveries, and about 90% of the eventually dry holes have been reliably predicted, which are respectively equivalent to 11% false-negative and 18% false-positive proportions (Schumacher, 2010). Other statistics have revealed

that 76% of the prospects with soil gas anomaly above them have ended up as commercial discoveries (Schumacher, 2000). Further discussion on this topic is provided in section 6-2-8.

Even though many onshore macroseeps have already been accounted for and their attributes are reflected in compiled global databases (Clarke and Cleverly, 1991; Etiope, 2009), the distribution and varieties of onshore microseeps have remained largely unknown. The great number of documented seeps (>10,000 entities) indeed implies that there is a staggering potential for microseeps yet to be detected. Fortunately, the spectral remote sensing is quite capable of closing the gap in our knowledge of petroleum seepage systems. Typically, the higher flux of macroseeps (several to hundreds of kgd^{-1}) permits remote sensing techniques to 'directly' target them using the diagnostic spectral signatures of escaping HCs (solid/liquid/gas), whereas the existence of microseeps (with fluxes in the order of a few to tens of $mgm^{-2}d^{-1}$) are usually inferred 'indirectly' using the alteration mineralogy induced by microseepage systems.

In the literature, there is no account of the sizes of macroseepage as potential targets for remote sensing. However, our experience indicates that the majority of them constitute tiny targets of a few meters (See for instance Fig. 5e–5f). In contrast, microseepage-induced changes are proportional to accumulation extent and thus set larger targets.

2.3. Remote sensing of macroseepages

2.3.1. Spectroscopy of petroleum

A number of structural bonds in petroleum including CH, CH_2, CH_3, and C=C give rise to several fundamental absorption bands between 3000–9000 nm wavelengths (Cloutis, 1989; Coates, 2006; Lammoglia and Souza Filho, 2011) (Table 2). Within the VNIR–SWIR wavelengths, petroleum also retains a series of absorption bands due to overtones and combinations of noted stretching fundamentals. The most notable feature in the VNIR–SWIR range includes a triplet between 1700–1750 nm and a doublet between 2290–2360 nm (Fig. 2 and Fig. 5d). Occasionally, subtle features around 1120–1230 and 1350–1550 nm can be present in the spectra of some lighter oils (Fig. 2).

To promote our understanding of the absorbing bands involved in each feature, we measured a light oil sample (API=43.2) with an ASD FieldSpec-4 spectrometer between 0.35–2.5 μm (the details of the sample and measurement protocol is provided in Asadzadeh (2016a)). The resulting spectrum was then analyzed by means of the Modified Gaussian Model (MGM) deconvolution technique (Sunshine et al., 1990). Results indicated that around 20 individual absorption bands are resolvable (Fig. 2) with each of the noted features being decomposed into multiple overlapping bands. A case in point is the prominent features at 2300 and 1700 nm that are respectively comprised of seven and five overlapping bands of varying width and intensities (Fig. 2).

Based on the fundamental absorption features reported in Table 2 and the physical basis of spectroscopy, we attempted to calculate the position of resolved absorption bands by considering the overtones and combinations of the fundamentals bands along with the combination of overtones, the combination of overtones and fundamentals, and the overtone of combinations. The findings are reported in Table 3 and illustrated in Fig. 2. Seemingly, there is a close match between the observed bands and the results of mathematical calculations. Moreover, each individual band

usually corresponds to a collection of theoretical possibilities (Table 3), indicating the complexity of the spectra and the difficulty of such interpretations. In a highly complicated compound like petroleum, a sheer number of absorbing bands can develop in the SWIR range, albeit the majority of them are averaged out due to overlapping and only the more pronounced bands or the contributors with constructive bands are left out as absorption features in the spectrum (Clark et al., 2009; Cloutis, 1989). This process is responsible for an intense backdrop absorption and a low overall reflectance (<20%) of petroleum. Based on band assignment outcomes, we roughly divided the SWIR wavelength into three sub-ranges: (i) between 1050–1350 nm in which first/second overtones are dominant; (ii) between 1350–2100 in which first overtones and combinations are dominant; and (iii) beyond 2100 nm in which the combinations are the dominant absorption mechanism (Fig. 2).

Table 2. The main fundamental absorption bands of organic compounds and HCs (Modified after (Cloutis, 1989; Coates, 2006)).

Band	Fundamental Wavelength/wavenumber		Fundamental group/assignment
	μm	cm^{-1}	
a	2.81	3550	Hydroxyl O-H stretch
b	3.28–3.30	3030–3050	Alkene, aromatic (aryl) C-H stretch
c	3.38	2960	Asymmetric CH_3 (methyl) stretch
d	3.42	2925	Asymmetric CH_2 (methylene) stretch
e	3.48	2870	Symmetric CH_3 stretch
f	3.50	2855	Symmetric CH_2 stretch
g	5.78–5.88	1700–1730	Carbonyl-carboxyl C=O stretch
h	6.06–6.25	1600–1650	Aromatic, alkenyl carbon (C=C) stretch
i	6.82–6.90	1450–1465	Asymmetric CH_3, CH_2 bend
j	7.27	1375	Symmetric CH_3 bend
k	9.71	1030	Aromatic C-H bend

The pronounced absorption feature occurring in the VNIR was modeled by a single Gaussian band and ascribed to the pi (π) electronic bonding between carbon atoms (Fig. 2). Typically, the minimum of this broad absorption shifts towards longer wavelength as the average number of aromatic rings per molecule (i.e. the aromatization type and degree) and the complexity of HC molecules increases (Cloutis, 1989; Cloutis, 1990). In other words, as oil becomes heavier in terms of API gravity, the absorption tends to become broader while its minimum moves towards longer wavelengths (see Fig. 3 in (Asadzadeh and Souza Filho, 2016a)). As a result, the stretching wing of this broad absorption that reaches up to 1500 nm and beyond can potentially mask the subtle (and rather weaker) features occurring at 1200 and 1400 nm. We should emphasize that except the feature centered at 1400 nm that coincides with atmospheric water vapor, other features can be potentially used for remote sensing HC detection.

Due to similar processes, methane and other lightweight alkane molecules (C_1–C_{10+}) show distinctive absorption features in the SWIR as well as in the mid-infrared (MIR) and LWIR (e.g. between 7–8 μm) ranges. Reflectance spectroscopy of alkane series has been comprehensively covered in (Clark et al., 2009) (see also section 3-3).

Table 3. Calculated wavelength positions of HCs arising from overtones and combinations of the fundamental absorption bands listed in Table 2. The numbers "2", "3", and "+" symbols denote the 1st and 2nd overtones and combinations of relevant bands, respectively.

Overtone/combination band	Wavelength (μm)	Observed (μm)	Overtone/combination band	Wavelength (μm)	Observed (μm)
3e	1.16	1.16	2f	1.75	1.76
3f, 2(f + j)	1.17	1.19	2j + d	1.76	
2k + 2b	1.22	1.21	2g + 2k	1.81	1.81
2d + i	1.37	1.39	3g	1.93	1.94
2d + j	1.38		d + h	2.18	2.19
2a, 2b + k	1.40	1.41	c + h, f + g	2.19	
2d + k	1.45	1.45	e + h, f + h	2.23	2.23
a + c	1.68		c + i	2.27	2.28
2c, b + e	1.69	1.70	c + j, e + i	2.31	2.31
b + f, c + d	1.70		d + j, 2i + j	2.32	
2d, c + e	1.71		e + j	2.35	2.35
c + f, 2i + d	1.72		f + j	2.36	
d + f	1.73	1.73	2j + i	2.38	2.39
2e	1.74		b + k	2.46	2.47

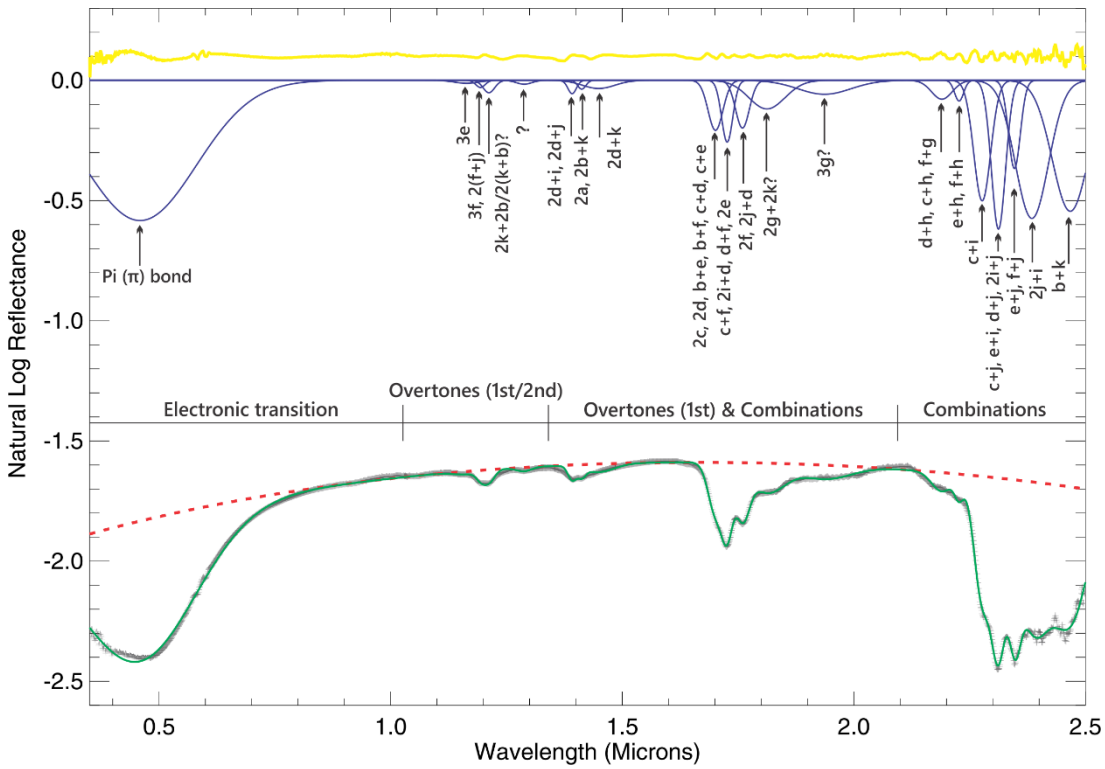

Fig. 2. Spectral deconvolution of a light HC spectrum using the MGM deconvolution technique. Plausible absorption mechanisms are provided for each of the resolved bands in the figure. The details of the band assignment are indicated in Table 3. The dashed red line represents the continuum modeled as a second-order polynomial. The solid yellow line illustrates the fitting error. The solid blue lines are the individual Gaussian bands. The original spectrum in gray is superimposed by the modeled spectrum (the combination of resolved bands) and shown in green.

2.3.2. An overview of the case studies

Despite he merit of spectral remote sensing in macroseepage detection, only a handful of case studies are available in the literature on the subject that is summarized in Table 4. The first study of this kind was facilitated by the Geosat Committee in 1998, wherein a Probe-1 hyperspectral scanner was flown over some well-known seeps in the Ventura basin, California. The study demonstrated that the seeping oil is spectrally mappable by a far-range system provided that the sensor has enough spatial and spectral resolution (Ellis et al., 2001). Although this study merely focused on the 2300 nm features, later investigations employed the feature centered at 1700 nm for HC detection using some simulated oil contaminations (Table 4). On the basis of that experiment, a 'hydrocarbon index' was proposed for HC targeting using the mere radiance dataset (Hörig et al., 2001; Kühn et al., 2004). In 2013, the oil seeps at the Ventura basin were surveyed once more using a bundle of hyperspectral imaging systems in the VNIR–SWIR–LWIR ranges. The conducted experiment with the SWIR subset indicated that a systematic remote sensing survey in a mature basin not only can map the extent of already known seeps accurately but also can lead to new seepage discoveries (Prelat et al., 2013). Correspondingly, in the case study reported in Fig. 5, the authors were unaware of the existence of Seep-4 beforehand and it was recognized just after spectral analysis (see Fig. 5f).

Table 4. Summary of the case studies directed towards the detection of petroleum macroseeps (and onshore oil contaminations) using spectral remote sensing techniques. The detection limit is defined as the smallest dimension of a target detectable spectrally. The acronyms used in the table are **BR**: Band ratioing, **RBD**: relative absorption band-depth, **DT**: decision tree, **SAM**: spectral angle mapper, **MF**: matched filtering, **MTMF**: mixture-tuned matched filtering, **GSD**: ground sampling distance, **HSI**: Hyperspectral imaging.

Location	target	Background	Remote sensing tools		Employed Features (nm)	Processing methodology	Detection limit (%)	Results	Reference
			Distal (GSD)	Proximal (Spectroscopy)					
Ventura, Santa Barbara, **USA**	Natural oil seeps	Monterey, Saugus, and Fernando Formations	Probe-1 (5 m)	Field spectroscopy	2300	Spectral unmixing	25	Excellent	(Ellis et al., 2001)
Spandau, **Germany**	Simulated oil contamination	Sand	HyMap (4 m)	In situ spectroscopy	1700	Color composite, BR	2.5	Promising	(Hörig et al., 2001; Kühn et al., 2004)
Ventura basin, **USA**	Natural oil seeps	Saugus and Fernando Formations	HSI (not specified) (3 m)	Field spectroscopy	2300	SAM + DT	Unknown	Excellent	(Prelat et al., 2013)
Barataria Bay, **USA**	Oil-impacted sites	coastline	AVIRIS (3.5–7.6 m)	Field spectroscopy	2300 & 1700	Feature fitting via MICA	16	Excellent	(Kokaly et al., 2013)
North Charleston, **USA**	Environmental pollution (diesel, crude oil, motor oil)	Estuarine area (dominated by clay minerals)	AVIRIS (20 m)	Field/Lab spectroscopy	1700 & 2300	Mixture simulation and spectral matching via SAM/MF ratio	10	Promising	(Smailbegovic et al., 2009)
Doberitzer Heide, **Germany**	Contaminations & plastics	Different soils	HyMap (5-6 m)	Lab simulation & in situ spectroscopy	2300, 1700 & 1200	BR	Unknown	Promising	(Winkelmann, 2005)
Casper, Wyoming, **USA**	Simulated oil-show	Sandy, clayey, and dolomitic soils	Simulated WV-3 (4–7.5 m)	In situ spectroscopy	1700	BR, RBD, MF	~35	Promising	Asadzadeh and de Souza Filho, 2016a)
			ProspecTIR (0.6m)		2300 & 1700	MTMF	5	Excellent	(Scafutto, 2015)

More recently, case histories over contaminated sites in Barataria Bay marshes affected by the Deepwater Horizon spill indicated the applicability of imaging spectrometer data to assess oiled sites onshore[2]. Spectral feature analysis incorporated both the HC features centered at 2300 and 1700 nm

2- Such case histories are noted here because of their direct relevance to oil-seep detection.

to compare the AVIRIS (Airborne Visible/InfraRed Imaging Spectrometer) data to the reference spectra collected in the field (Table 4). The maps derived from AVIRIS datasets with different dates and spatial resolutions were demonstrated to be ~90% accurate in predicting the oiled shorelines. The smallest oily target detected at 7.6 m Ground Sampling Distance (GSD) was shown to be 1.2 m, which is equivalent to 16% detection limit for the sensor (Kokaly et al., 2013).

In another environmentally oriented study, a suite of clean soil samples was impregnated by 10%, 30%, and 50% volume of four different HCs namely crude oil, motor oil, diesel, and kerosene and then measured spectrally. The compiled spectral library was then used to detect polluted sites in an estuarine area using AVIRIS data acquired at a GSD of 20 m (Table 4). The data successfully detected muddy and/or sandy estuarine sediments contaminated by around 10% HC contents. The HC features, however, were more apparent in the sandy samples rather than darker loamy mud (Smailbegovic et al., 2009). The study was incapable of distinguishing between the varieties of HC contaminations in the sediments.

So far, HC detection has been confined to airborne hyperspectral sensors, but some simulation experiments have demonstrated that the newly launched WorldView-3 (WV-3) satellite system can potentially be employed for direct HC detection using the bands coinciding with HC's 1700 nm feature (Table 4) (Asadzadeh and Souza Filho, 2016a). If further case studies with real datasets support this finding, the WV-3 satellite data then will be the first spaceborne instrument capable of detecting onshore oil directly and unambiguously.

2.3.3. Gas-plume sensing

Contrary to the marine environment in which gas seeps are easily noticeable, terrestrial seepage detection has been biased towards more visually evident oil seeps and, thereof, many tiny gas leakages have gone unnoticed. Generally, where the emission of such seeps occur as concentrated point sources, the escaping methane can form discernible gas-plume in the atmosphere and thus constitute detectable target. Current orbital gas remote sensing instruments like the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) and the Greenhouse Gas Observing Satellite (GOSAT) have the capability to measure methane and other trace gasses but only in continental scales (Watanabe et al., 2015). The delineation of point source plumes at local-scale, however, is a niche that is filled with high spatial (GSD<20 m) and spectral resolution hyperspectral imaging instruments.

The detection and mapping of methane plumes from natural and anthropogenic sources relies on its diagnostic absorption features in the SWIR (~2.3 μm) and/or the LWIR (~7.6 μm) wavelengths. In the SWIR, AVIRIS and AVIRIS-NG are the highly explored systems for this aim. For example, high-glint AVIRIS imagery was successful in detecting the concentrated sources of CH_4 in the Coal Oil Point (COP) seeps in the Santa Barbara Channel, California. In this test site, radiance data were processed using band ratioing and MODTRAN-based simulated radiance spectra to generate a 'methane index' over some known marine gas seeps (Bradley et al., 2011; Roberts et al., 2010) (Table 5). In the case of terrestrial methane mapping where emissions occur over non-uniform and heterogeneous background albedo, encouraging results were achieved by more advanced processing algorithms such as the Cluster-Tuned Matched Filter (CTMF) and the Iterative Maximum a Posteriori Differential Optical Absorption Spectroscopy (IMAP-DOAS) (Frankenberg et al., 2016; Thorpe et al., 2013) (Table

5). This approach was not only effective in detecting a plume but also was capable of quantifying the flux of CH_4 in the atmosphere (see section 6-1-1).

Some parallel studies using the LWIR wavelength has already given rise to encouraging results for methane detection. For instance, data from the airborne Hyperspectral Thermal Emission Spectrometer (HyTES) at ~2m GSD were successful to map several individual methane plumes over oil fields in the San Joaquin Valley, California (Table 5).

Table 5. Summary of the case studies directed towards gas-plume (methane) sensing over natural seepage systems.

Location	Target	Instrument (GSD)	Employed feature	Processing methodology	Results	Reference
Coal Oil Point, California, **USA**	Natural seep (offshore)	AVIRIS (7.5 m)	2.298 µm (SWIR)	Band ratioing Spectral residuals	Promising	(Bradley et al., 2011; Roberts et al., 2010)
Inglewood Oil Field, LA, **USA**	Natural seep (onshore)	AVIRIS (~3 m)		CTMF	Promising	(Thorpe et al., 2013)
San Juan Basin, **USA**	Natural seep (onshore)	AVIRIS-NG (1-3 m)		IMAP-DOAS	Promising	(Frankenberg et al., 2016)
San Joaquin Valley, Ca, **USA**	Natural seep? (onshore)	HyTES (~2 m)	7.66 µm (LWIR)	CMF	Promising	(Hulley et al., 2016)

Aside from the emerging airborne technology, ground-based sensing instruments have been utilized for gas seepage detection. For a comprehensive review of active ground-based systems, including laser and Lidar sensors, the reader is referred to Chapter 4 of Etiope's book (Etiope, 2015) and references therein. In the lab, reflectance spectroscopy has been further attempted to facilitate the detection of adsorbed gas (and probably aromatics/petroleum?) within clay particles. The trial with soil samples collected over three oil and gas fields indicated that meaningful spectral trends coexist over petroliferous areas (McCoy et al., 2001).

It is worth mentioning that X-band radar was indicated to be capable of detecting atmospheric seepage anomaly (ASA) above terrestrial accumulations. The scientific concept along with some case studies was discussed in a series of papers by T. Bailey (Bailey, 1996; Bailey and Grubb, 2006; Bailey and Skolnik, 1996).

2.4. Microseepage remote sensing

Since the majority of the mineralogical assemblages in microseepage systems (section 2-2) show diagnostic spectral features within VNIR–SWIR wavelengths (Hunt, 1977), they have been the focus of several remote sensing investigations. This section firstly attempts to overview the case studies available on the subject and then summarize the closely related surveys on exhumed HC reservoirs.

2.4.1. An overview of the case-studies

Near-surface diagenetic changes have been known for a long time (e.g. Davidson (2004) and references therein); however, the first systematic and comprehensive study on the topic was conducted only in the 1970s by Donovan (1974), which was simultaneous to the concept of 'hazy anomaly' from ERTS data. A term that was brought about to explain the peculiar subtle tonal areas

observed in enhanced imagery over some productive/prospective onshore fields[3] (Simpson, 1978). Thereafter, different generations of remote sensing data have been employed for this aim, each revealing further details of the diagenetic mineralogy within microseepage systems.

In this paper, we define three levels for remote sensing-derived anomalies including (i) tonal (hazy) anomaly, (ii) spectral anomaly, and (iii) mineral map (Table 6). The tonal anomaly is interpreted visually, thus its exact mineralogical attribute is not clear. The spectral anomaly is routinely derived from multispectral imagery following simple spectral analysis. In this level, altered zones are discriminated in a broad sense without attempting to discriminate mineral species (e.g. OH-bearing minerals derived from Landsat data). In latter, however, the majority of the minerals appearing in a microseepage system are either distinguished (e.g. Al-OH) or identified individually (e.g. kaolinite).

Key case studies demonstrating the application of remote sensing methods for microseepage delineation are summarized in Table 6. We devised a variety of attributes comprising field characteristics, geologic/mineralogic features (e.g. exposed stratigraphic units), remote sensing approach (i.e. distal airborne/spaceborne and proximal sensing tools), specifications of the yielded target, ground-truthing strategy along with complementary analytical techniques, and then tabulated the case studies on their basis. In this table, we distinguished active microseepage systems with ongoing leakage to the surface from fossil (inactive) systems wherein the HC migration has halted. In addition, for clarification purposes, case studies addressing miniseeps (See section 2-1 and Fig. 1) were segregated from microseepage systems. More importantly, the potential targets revealed by fieldwork were contrasted against mapping results achieved through spectral analysis in order to emphasize the weakest link(s) in the processing chain. We made a subjective judgment about the shape of the resulting anomaly and its correspondence with the surface projection of the accumulation and the overall outcome of every case study. Subsequently, we assigned the following scores to each of the attributes: very-high>high>medium>low>slim for the former and excellent>promising>average>poor for the latter (Table 6).

Due to the significance of ground-truthing in demonstrating the capability of remote sensing, in this table, we also reported the number of collected samples, the likely spectroscopy of the sample suite, the sampling scheme (traverse vs. selective), and the locality of sampling sites ‘on’ and ‘off’ the petroliferous (altered) zones considering the same geologic unit. Ultimately, because many case studies were supplemented by geochemical and/or geophysical techniques of one type or another, we accounted for them under a category named analytical approach (Table 6). To better reflect the trends in the compiled case histories (28 individual cases), we generalized the attributes of Table 6 to construct a series of bar charts illustrated in Figs. 3-4.

Apparently, the case studies are not uniformly distributed, for the majority of them are conducted in the United States or China and paramount petroliferous regions such as the Middle East are not adequately investigated (Fig. 3a). Although the cases have evenly covered oil and gas accumulations (Fig. 3b), they are biased heavily towards the structural traps and the important stratigraphic entities are not satisfactorily incorporated (Fig. 3c).

3- According to our experiment, a hazy anomaly arises likely due to the dissolution/re-deposition of authigenic minerals and subsequent disappearance of the texture and fabric of the original facies.

Table 6. Summary of the remote sensing case studies directed towards the detection and characterization of alterations induced by onshore mini-, and microseepage systems (the miniseep cases are marked in gray).

Location	Field characteristics			Geologic/mineralogic features			Remote sensing approach			Target specification				Fieldwork and sampling					Analytical approach	Results	Reference
	HC charge	Trap	Seepage activity & type	Outcrop lithology	Age	Exploration signatures	Distal tool	Proximal tool	Processing method	Anomaly type	Anomaly shape	Mapped target(s)	Correspondence	Locality: On	Locality: Off	Strategy: Selective	Strategy: Traverse	No			
Lisbon Valley; Utah, **USA**	Oil & gas	Structural (faulted trap)	Active	Wingate sandstone	Early Jurassic	• Bleaching of red beds • High clay content • High precipitation of limonite and carbonate • Anomalous weathering patterns	MSS TMS TM	Spectroscopy	BR	Spectral anomaly	Elongated patches at the hinge points of the fold	Rocks poor in limonite with high clays	Medium	x	x	x	-	15	- XRD - Petrography - Whole-rock geochemistry	Promising	(Segal and Merin, 1989; Segal et al., 1986)
							ASTER JERS	Spectroscopy	ML, PCA	Mineral map		Bleached facies	Medium	x	-	x	-	7 (altered)	- ICP-AES - Petrography	Promising	(Petrovic et al., 2008)
							HyMap	ASD spectroscopy	SAM, SFF	Mineral map		Kaolinite, calcite, hematite	Medium	x	x	x	-	37	- Petrography - ICP-MS/AES - Isotope Study: O, C	Promising	(Petrovic et al., 2012)
Navajo; Utah **USA**	Oil/gas	-	Ancient	Navajo Sandstone	Jurassic	• Iron Oxide beaching & goethite • Clays • Carbonate (calcite)	HyMap	ASD & FTIR spectroscopy	MF	Mineral map	Elongated patches of minerals	Goethite, hematite, carbonate	Medium	x	x	x	-	170	- Petrography - XRD - ICP-MS	Excellent	(Bowen et al., 2007)
Patrick Draw Wyoming, **USA**	Oil	Stratigraphic	Active	Calcareous/silty sandstone	Eocene	• Clays (illite, kaolinite/smectite) • Iron oxides (ferrihydrite, goethite)	Hyperion	ASD spectroscopy	SAM	Mineral map	Scattered patches	Bleaching?	Low	x	-	x	-	40	- ICP - XRD - Isotope Study: C	Average	(Khan and Jacobson, 2008)
Santa Barbara County, **USA**	Heavy oil	Structural	Active (Miniseep)	Saugus and Fernando Fm?	Pliocene	• Jarosite, siderite, calcite	AVIRIS	-	SAM, SFF	Mineral map	Scattered patches	?	Not verified	-	-	-	-	-	-	Poor	(Freeman, 2003)
Sulphur maountain, Ventura, Santa Barbara, **USA**	Oil	Structural	Active (Miniseep)	Fernando Fm.	Pliocene	• Peripheral alteration around macroseeps characterized by goethite and calcite	Probe-1 (HyMap)	-	SAM, CCSM	Mineral map	Small patches	Goethite	Slim	-	-	-	-	-	-	Average	(van Der Meer et al., 2002)
Dutton Basin, Wyoming, **USA**	Oil	Anticline	Ancient/ inactive (Miniseep)	Nugget/Chugwater Fm. (siltstone, Sandstone)	Triassic	• Fe hydroxides (goethite) • Calcite veins	AVIRIS	Spectrometer	BR	Mineral map	Pointy	Goethite	Medium	x	-	x	-	3	- XRD	Average	(Jengo and Vincent, 1999)
Paradfurdo; Matra Mountains, **Hungary**	Gas	?	Active? (Miniseep)	?	?	• Botanical anomaly • Mineralogical alterations	DAIS	Spectrometer	Unmixing Knowledge-based method	Spectral anomaly	Scattered patches	Clays?	Medium	x	-	x	-	56	- (Aerial photography)	Promising	(van der Werff, 2006)
Sheep Mountain, Bighorn basin, Wyoming, **USA**	Oil	Anticline	Active	Cleverly Fm. (sandstone, shale)	Triassic	• Absence of kaolinite in shale • Red bed bleaching	TM	Spectrometer	BR	Tonal anomaly	?	Bleaching	?	x	x	x	-	Several	- XRD - Ph measurement	Average	(Malhotra et al., 1989)
North Tucano Basin, **Brazil**	Gas	Anticline	Active	Marizal Fm. and Tona Sequence (Sandstone, siltstone)	Tertiary	• Bleaching and/or limonitic yellow soils • Development of clay minerals • Silicification & kaolinite depletion • Geobotanical anomalies	ASTER ETM+	ASD spectroscopy	BR, PCA SAM, MTMF, NN	Spectral anomaly	Scattered over the field	Clays, iron oxides	Medium	x	-	x	-	7	- Gasometry	Promising	(Lammoglia et al., 2008)
							TM	ASD spectroscopy	BR (2/3, 4/3)	Spectral anomaly	Integrated over gaseous anomaly	Iron oxides	High	x	-	-	x	25	- Gasometry - Radiometry - Petrography	Promising	(Almeida-Filho et al., 2002; Almeida-Filho et al., 1999)
Remanso do Fongo, **Brazil**	Gas	?	Active (Miniseep)	Alluvial sediments	Holocene	• Ferric iron and kaolinite • Geobotanical anomalies	ASTER TM	-	BR, PCA	Spectral anomaly	Integrated	Kaolinite iron oxides	Medium	-	-	-	-	-	- Radiometry - Magnetic survey	Average	(Curto et al., 2011; Souza Filho et al., 2008)
Table Rock, Wyoming, **USA**	Gas	Anticline	Active?	Green River & Wasatch Fm. (sandstone, mudstone, etc.)	Eocene	• Mixed-layer illite-smectite, kaolinite, chlorite • Hematite, pyrite, jarosite, gypsum • Carbonate cement	TMS	Airborne spectroradiometer	- Visual interpretation - BR	Tonal Anomaly & Mineral map	Integrated along the crest of anticline	?	High	x	-	-	x	10s	- XRD - Petrography - Ph measurement	Average	(Marrs and Paylor., 1987)

Table 6 (Continue).

Location	Field characteristics			Geologic/mineralogic features			Remote sensing approach			Target specification				Fieldwork and sampling					Analytical approach	Results	Reference
	HC charge	Trap	Seepage activity & type	Outcrop lithology	Age	Exploration signatures	Distal tool	Proximal tool	Processing method	Anomaly type	Anomaly shape	Mapped target(s)	Correspondence	Locality		Strategy		No			
														On	Off	Selective	Traverse				
Palm Valley, Amadeus basin, N.T; **Australia**	Gas	Anticline	Active	Hermannsburg sandstone	Devonian	• Iron oxide anomaly (magnetite high and hematite low) • Calcrete (dolomite, calcite) • Clays (montmorillonite and kaolinite) • Surface weathering/crust	TM TMS	IRIS spectrometer	BR	Tonal anomaly	Integrated	Probably iron oxides	Medium	x	?	-	x	60 spectra	- Gasometry - Isotope studies - Ph measurement	Average	(Simpson et al., 1991)
							GEOSCAN AMSS	IRIS spectrometer	SFF	Mineral Map	Integrated	Calcite, silica	Medium	x	-	x	-	several	-	Average	(Agar, 1999)
Los Chihuidos High, **Argentina**	Gas & oil	Exhumed reservoir		Huincul Fm. (fine to coarse-grained sandstone)	Upper Cretaceous	• Bleaching of red beds attributed to the migration of HCs	ETM+	-	Visual Interpretation	spectral anomaly	Pervasive halo	Bleached beds	High	x	x	-	x (7 profiles)	40	- Petrography - ICP-AES/MS - XRD - Microprobe - FTIR spectroscopy	Promising	(Rainoldi et al., 2014)
Baku Region, **Azerbaijan**	Oil	Structural	Active (miniseep)	Alluvial (Clay and sand)	Pliocene	• Clays (Montmorillonite, chlorite, illite, kaolinite) • Bleached/discolored red sandstones • Ferrous carbonate	ASTER InSAR	FieldSpec FR spectrometer	VMESMA	Mineral Map	Patchy above mud volcanos	Smectite, kaolinite	-	x	-	x	-	several	-	Promising	(Scholte, 2005; Scholte et al., 2003)
Ramhormoz, **Iran**	Oil & gas	Anticline	Active (miniseep)	Gachsaran & Mishan Fm. (Marl)	Miocene	• The formation of **gypsum** and S in marly limestone. • Removal of Fe oxides and calcite	ASTER WorldView-2	ASD spectroscopy (In the lab)	BR, PCA BRT	Spectral anomaly	Pointy	Bleaching?	Slim	x	-	x	-	88	- ICP-OES - XRD - Stable isotopes: C,O,S	Poor	(Salati et al., 2014a; Salati et al., 2014b)
Masjed Soleiman, **Iran**	Oil	Anticline	Active (miniseep)	Gachsaran formation (gypsum)	Miocene	• Decrease in gypsum/anhydrite (Gach-e-Turush) • Increase in calcite, dolomite, aragonite, native Sulphur (S) • Jarosite, Na-alunite, illite, montmorillonite	-	-	-	-	-	-	-	x	-	x	-	18	- ICP-AES - XRD - Isotope Study: C - Ph measurement	Promising	(Tangestani and Validabadi, 2014)
Kor Mor Field; Kurdistan, **Iraq**	Oil	Anticline	Active	Upper & lower Fars Fm. (sandstone, siltstone, marl)	Miocene	• Calcite, dolomite • Illite, chlorite • Iron oxides + jarosite	ETM+ ASTER	ASD spectroscopy (In the lab)	BR, PCA SFF, SAM	Mineral Map	Pervasive apical	Bleaching, calcite	Very high	x	-	x	-	>30	-	Promising	(Perry and Kruse, 2010)
Qiulitage Anticline, **China**	Oil	Thrust-and-fold	Active	Red beds with sandstone & gypsum	Pliocene	• Bleaching • Carbonates & gypsum • Clay & OH bearing minerals	ASTER	ASD spectroscopy (In the lab)	BR (2/1, 4/9)	Spectral anomaly	Pervasive apical	Bleached beds	High	x	x	x	-	10	- Petrography - Scanning Electron Microscope (SEM) - XRD	Excellent	(Shi et al., 2012)
Fula'erji, Songliao Basin, **China**	Heavy oil	Stratigraphic	Inactive	Alluvial flood-plain	Holocene	• Bleaching and clay anomaly?	ETM+	-	BR, PCA	Tonal anomaly	?	?	?	-	-	-	-	19	- Magnetic susceptibility - Delta carbonate	Poor	(Zhang et al., 2009)
Dushanzi Anticline, **China**	Oil	Structural	Active	Red beds interbedded with conglomerate	Pliocene	• Bleaching (hem-Fe^{3+}) to (para-Fe^{3+}) • Carbonates (calcite, siderite) • Chlorite	ASTER	ASD spectroscopy (In the lab)	BR (2/1, 4/6, 4/8)	Tonal anomaly	Pervasive apical	?	Very high	x	-	x	-	8	- XRD - XRF - Mössbaur spectroscopy - XANES	Excellent	(Fu et al., 2007; Zheng et al., 2010)
Karamay Field, **China**	Oil & gas	Structural	Active	Sandy breccia conglomeratic sandstone	Triassic	• Clays (Chlorite, montmorillonite) • Carbonate (siderite)	Hyperion	ASD spectroscopy (In the field)	MTMF, SAM	Mineral Map	Narrow linear patches	Siderite, clays	Slim	x	-	-	x?	?	-	Poor	(Qin et al., 2016)
Lake Albert Basin, **Uganda**	Oil & gas	Structural	Active	Alluvial Sediments	Pleistocene-Holocene	?	ASTER ETM QuickBird	-	?	Microseepage anomaly?	Pervasive apical	Altered units?	Very high	-	-	-	-	-	-	Excellent	(Frassy et al., 2015)
Ordos Basin, **China**	Oil	Structural?	Active	Yanchang Fm. (Sandstone, siltstone)	Upper Triassic	• Clays (Illite, kaolinite) • Carbonates (Siderite, calcite)	CASI/SASI	ASD spectroscopy (In the field)	SFF	Mineral map	Scattered along the field	Siderite, calcite, illite, kaolinite	Medium	x	-	x	-	12	- XRD - AAS	Average	(Chen et al., 2016)
East-Sichuan Fold Belt, **China**	gas	Structural	Active	?	Triassic?	• Clays (montmorillonite, illite, kaolinite, and chlorite) • Carbonates (calcite and dolomite) • Ferric & ferrous iron	TM	Hitachi VIS–NIR spectrometer	BR	Spectral anomaly	Scattered & patchy	Clays, bleaching facies	Low	x	x	-	x	23	- XRD - Chemical analysis	Poor	(Wang and Ding, 2000)

The acronyms used in the table are the following: **TM**: Landsat Thematic Mapper, **ETM+**: Landsat Enhanced Thematic Mapper plus, **TMS**: Thematic Mapper Simulator, **ASTER**: The Advanced Spaceborne Thermal Emission and Reflection Radiometer, **ASD**: Analytical Spectral Devices, **BR**: Band Ratioing, **PCA**: Principal Component Analysis, **ML**: Maximum Likelihood, **SAM**: Spectral Angle Mapper, **SFF**: Spectral Feature Fitting, **MF**: Matched Filtering, **MTMF**: Mixture Tuned Matched Filtering, **CCSM**: Cross-Correlogram Spectral Match, **NN**: Neural Network, **VMESMA**: Variable Multiple Endmember Spectral Mixture Analysis, **BRT**: Boosted Regression Tree, **XANES**: X-ray Absorption Near-edge Structure analysis, **XRD**: X-Ray Diffraction, **SEM:** Scanning Electron Microscope, **XRF**: X-Ray Fluorescence, **ICP**: Inductively Coupled Plasma, **AAS**: Atomic Absorption Spectroscopy.

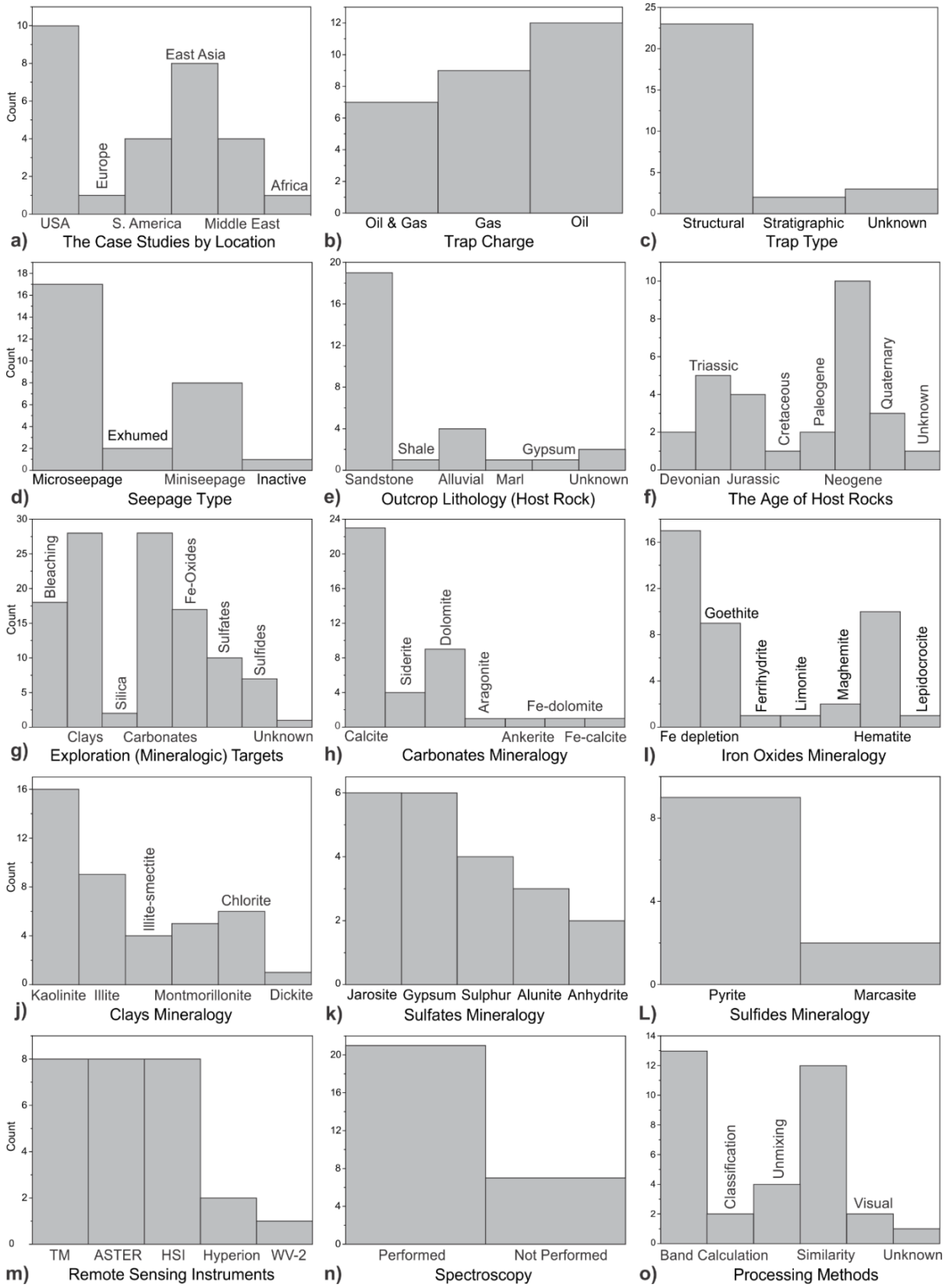

Fig 3. Bar chart representation of different attributes of the case studies summarized in Table 6. **a)** distribution of case studies around the world. **b)** charge of the reservoir/accumulation **c)** type of trap. **d)** seepage type (microseep vs miniseep) and activity. **e)** exposed lithologic units affected by microseepage system (here sandstone represents sandstone, siltstone, and conglomerate). **f)** generalized age of the host rocks (Quaternary encompasses Holocene and Pliocene sediments). **g)** described mineralogical signature of the alteration array (here the bleaching is distinguished from the excess of iron oxides). **h)** reported mineralogy of carbonates. **i)** detailed mineralogy of iron oxides. **j)** diversity of clay species. **k & l)** mineralogy of sulfates and sulfides. **m)** deployed remote sensing instruments (all the Landsat series are

reported as TM and airborne hyperspectral imaging systems as HSI). **n)** number of cases performing field/lab spectroscopy. **o)** processing methods employed in the analysis chain (band calculation indicate PCA and band ratio techniques). For more explanation about the processing methods, the reader is referred to (Asadzadeh and Souza Filho, 2016b). Fig.3g-3k were prepared by incorporating the data from Table 7.

In terms of seepage type, several case studies have focused on miniseeps (Fig. 3d), which according to the definition provided in section 2-1, constitute smaller point targets typically off the accumulations. The chiefly accounted host rocks over microseepage systems are sandstones and other relevant clastic sediments (Fig. 3e) that extend in time from Devonian to Holocene (Fig. 3f). This bias is likely because the footprints left by migrating HC in these units are vast and easily recognizable. Concerning the type of alteration mineralogy, iron oxides (i.e. bleaching and/or enrichment) sounds to be the dominant entity followed closely by clays and carbonates (Fig. 3g). Within each of these assemblages, a variety of minerals were observed, albeit some individuals came up to be more important. For example, while calcite is the dominant recorded carbonate (Fig. 3h) dolomite and siderite can be potentially important targets. In the literature, great emphasis has been given to kaolinite as the most prevalent clay alteration; however, our graph (Fig. 3j) clearly shows that other clay minerals such as illite, smectites, and chlorites are equally significant and indicative of diagenetic facies. Likewise, iron oxides and sulfates (Fig. 3i and 3k) assemblages reveal remarkably rich varieties of spectrally active minerals.

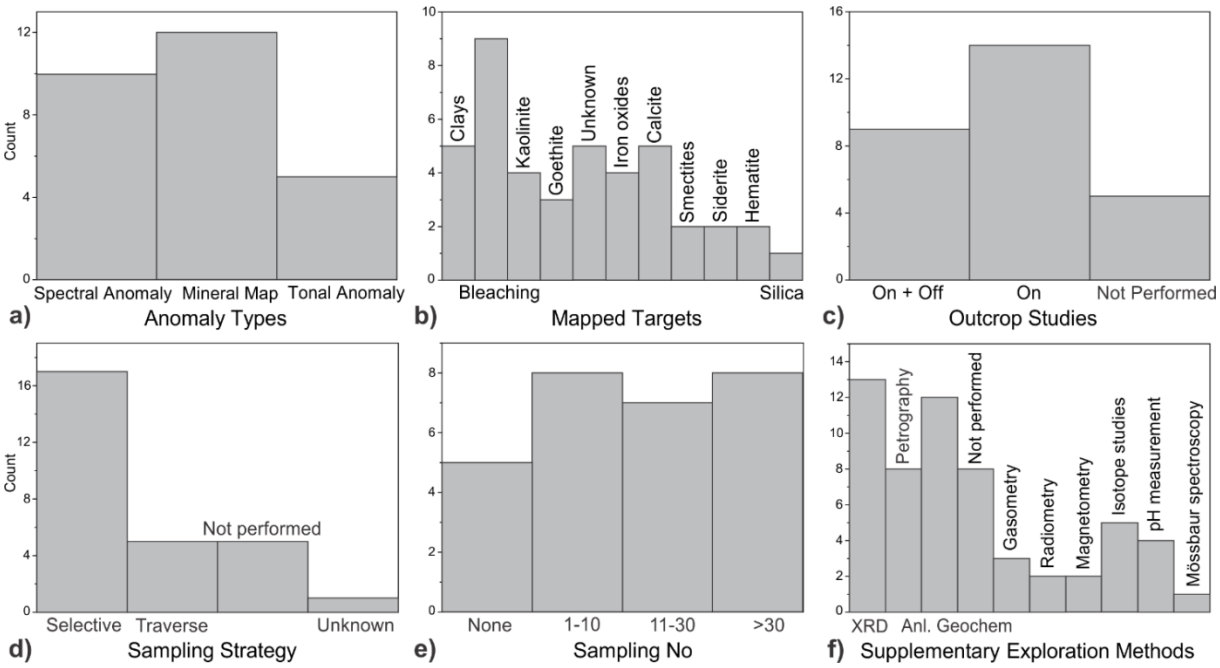

Fig 4. Bar chart representation of different attributes of the case studies summarized in Table 6. **a)** different types of anomalies yielded from remote sensing imageries. **b)** mapped targeted minerals. **c)** locality of the samples collected in outcrop studies defined to be 'on' or 'off' the affected zones. **d)** strategy in sample collection. **e)** number of samples collected in each case study. **f)** supplementary exploration methods. The analytical geochemistry (Anl. Geochem) column depicts XRF, ICP-MS/AES, and other analytical methods commonly used to analyze major/minor/trace elements.

In the outlined case studies, different generations of remote sensing instrument have been employed (Fig. 3m) to map the anomalous zones and field/lab spectroscopy has been regularly used to verify the results (Fig. 3n). Because the processing of spectral data has relied largely on simple techniques such as band ratios (Fig. 3o), many of the yielded anomalies reside in the category of spectral anomaly (Fig. 4a). As a result, the variation/pattern in the mineralogy of microseepage

systems is ill-understood (Fig. 4b). According to the adopted sampling strategies, a large proportion of the cases have collected <30 samples from localities 'on' the affected zones in a selective manner (Figs. 4c-4e). Ultimately, the studies have been supplemented by a variety of analytical techniques, including petrography and XRD analysis (Fig. 4f).

2.4.2. Exhumed HC reservoirs

There is a constellation of well-explored case studies on exhumed HC reservoirs/fairways in southwestern US basins (Table 7) within which the documented mineralogical suite (in microscopic, outcrop, and occasionally regional scales) remarkably resemble the microseepage-induced alterations discussed earlier (Eichhubl et al., 2004; Garden et al., 2001; Rainoldi et al., 2014). We speculate that such cases, which are characterized by intense alteration and chemical bleaching (e.g. Beitler (2005)), exemplify an endmember of the spectrum of alterations induced by HC migration/entrapment. Therefore, a deeper understanding of such changes could be an asset to the understanding of microseepage phenomenon and the characters it retains in different geologic settings.

Aside from exhumed HC reservoirs, the array of secondary mineralogy occurring within currently active oil and gas reservoirs and oil-sand deposits indeed indicate similar trends in diagenetic alteration (Cloutis et al., 1995; Matthews, 1986; Parry et al., 2009). A case in point is the Athabasca's oil-sands in which the bitumen is accompanied by clays (i.e. kaolinite, illite, montmorillonite, and chlorite), Fe-carbonates (siderite), and pyrite (Cloutis et al., 1995). This trend, however, does not imply that all the alteration/bleaching patterns observed in sedimentary environments arise from thermogenic HCs; because the circulation of other reducing agents such as biogenic methane, organic acids, CO_2, and H_2S can give rise to similar alteration arrays (Chan et al., 2000; Nielsen et al., 2009). To denote this analogy, we included two case histories on CO_2-induced transformations in Table 7. This underlines the significance of novel mineralogical indicators to distinguish the two phenomena apart. The subject is discussed in sections 6-2-2 and 6-2-6.

2.5. Demonstration datasets

Within this article, examples of remote sensing data over two study areas with distinctive macro-, and microseepage systems are presented as illustrative products. The first area located in the Ventura Basin (CA, USA) hosts a typical oil seepage system (e.g. (Ellis et al., 2001; Prelat et al., 2013; van Der Meer et al., 2002)) among outcrops of sandstone, conglomerate, mudstone, and shale units belonging to Pico, Sisquoc, and Monterey Formations (Fig. 5b). The data covering this test site were collected on August 26, 2014, during the HyspIRI Preparatory Campaign by AVIRIS instrument at an altitude of 20 km and a GSD of 15.9 m (Fig. 5a). The georectified reflectance dataset was acquired from the AVIRIS webpage and then was spectrally processed by a combination of matched filtering (MF) and logical operators designed to track the diagnostic absorption features of petroleum at 2300 and 1700 nm. The resulting anomaly map is presented in Fig. 5c. Selected anomalies namely Seep-1, Seep-4, Roof-1, and Plastic Paint were verified spectrally (Fig. 5d) and visually (Fig. 5e-5h) using respectively the spectral content of the imagery and high-resolution satellite imagery.

The second study area is located in Northern Tucano basin in Bahia, Brazil. The area is ruled by a semi-arid climate and consequently, the bedrock is fully exposed to satellite sensors (Fig. 6c). The

main rock units in the area are conglomerate, sandstone, shale, and silicified limestone that collectively belong to Tona Sequence and Marizal Formation (Fig. 6a). Possible HC accumulation in a structural trap in this area is accompanied by marginal soil-gas anomalies (Almeida-Filho et al., 2002; Lammoglia et al., 2008). The soil-gas data were employed to assess the remotely-sensed spectral anomalies. For this aim, ethane to pentane readings were firstly normalized by dividing them to their maximums and then summed up to represent the total HC content of gasses heavier than methane ($\sum C_2 - C_5$). The data were subsequently interpolated using the Empirical Bayesian Kriging technique available in ArcMap 10.3 software (www.esri.com) and illustrated in color-coded raster format (Fig. 6b). The datasets from Landsat-OLI (Fig. 6c) and ASTER multispectral instruments acquired respectively on 2016/10/06 and 2006/11/04 were used to map the intrinsic alterations within this microseepage system. In the spectral processing stage, we mapped the distribution of clays, ferrous iron (e.g. Fe-carbonates, chlorite, etc.), and ferric iron oxides/oxyhydroxides (e.g. goethite, hematite $\pm$ ferrihydrite) by employing a partial unmixing technique and proper endmembers derived manually from ASTER-SWIR and Landsat imageries. The resultant maps are shown in Fig. 6d–6f.

Table 7. Summary of case studies directed towards exhumed HC reservoirs and paleo-fluid fairways in permeable sedimentary facies. The summarized mineralogical contents of this table were incorporated into bar charts of Fig. 3.

Location	Target	Outcrop Lithology	Alteration signatures	Outcrop study	Mapping tool(s)	Reference
Valley of Fire; NV, **USA**	Exhumed reservoir	Aztec sandstone	- Clay minerals (kaolinite, illite/smectite, dickite) - Goethite & hematite banding (without pyrite) - Quartz overgrowth & carbonate cement - Sulfates (alunite & jarosite)	×	Color aerial photograph	(Eichhubl et al., 2004)
Moab anticline, Utah, **USA**	Exhumed reservoir	Entrada sandstone	- Iron oxides - Calcite cementation - Pyrite, ankerite, kaolinite	×	Color aerial photograph	(Garden et al., 2001)
Southeast Utah, **USA**	Paleo-fluid migration footprints	Navajo sandstone	- Iron cementation - Iron bleaching	×	Color aerial photograph	(Nielson et al., 2014)
Zion NP; Utah, **USA**	Remnants of reducing fluid	Navajo sandstone	- Iron bleaching - Iron enrichment	×	-	(Nielsen et al., 2009)
Southern Utah, **USA**	Exhumed reservoir	Navajo sandstone	- Secondary iron oxides - Late calcite cement (+ dolomite) - kaolinite & illite	×	Landsat ETM	(Beitler et al., 2005)
Kaibab, Utah, **USA**	Exhumed reservoir	Navajo sandstone	- Secondary hematite cement - Calcite - Illite and kaolinite	×	-	(Parry et al., 2004)
Elaterite Basin, Utah, **USA**	Exhumed reservoir	Permian White Rim sandstone	- Bleaching of red beds - Secondary diffused or concretionary iron - Pseudomorphs of pyrite - Calcite precipitation - Illite and kaolinite	×	GPS, geologic map	(Gorenc and Chan, 2015)
Colorado Plateau, **USA**	Exhumed reservoir	Glen Canyon sandstones	- Bleaching of iron oxides - Clays	×	Landsat ETM	(Beitler et al., 2003)
Onshore wells, **Denmark**	Unknown	Skagerrak Fm. (Arkosic red sandstone)	- Dolomite, anhydrite - Mixed-layer illite/smectite, kaolinite - Chlorite (Mg-rich in red and Fe-rich in whitish parts) - Pyrite	-	-	(Weibel, 1998)
O'Neill forebay CA, **USA**	Gas microseepage?	Pleistocene Corcoran clay	- Alunite, jarosite, and gypsum - pyrite and iron sulfates - Native sulfur and Acidic pH	-	-	(Prokopovich et al., 1971)
San Rafael Swell, UT, **USA**	Roll front of a CO_2 reservoir	Navajo sandstone	- (Oxyhdr) oxides (lepidocrocite/goethite/hematite) - Clays - Carbonate cementation (dolomite)	×	-	(Potter-McIntyre et al., 2013)
Green River, Utah, **USA**	Exhumed natural CO_2	Entrada sandstone	- Bleaching - Pore-filling calcite, dolomite, ferroan dolomite - Gypsum veins - Illite-smectite	×	-	(Wigley et al., 2012)

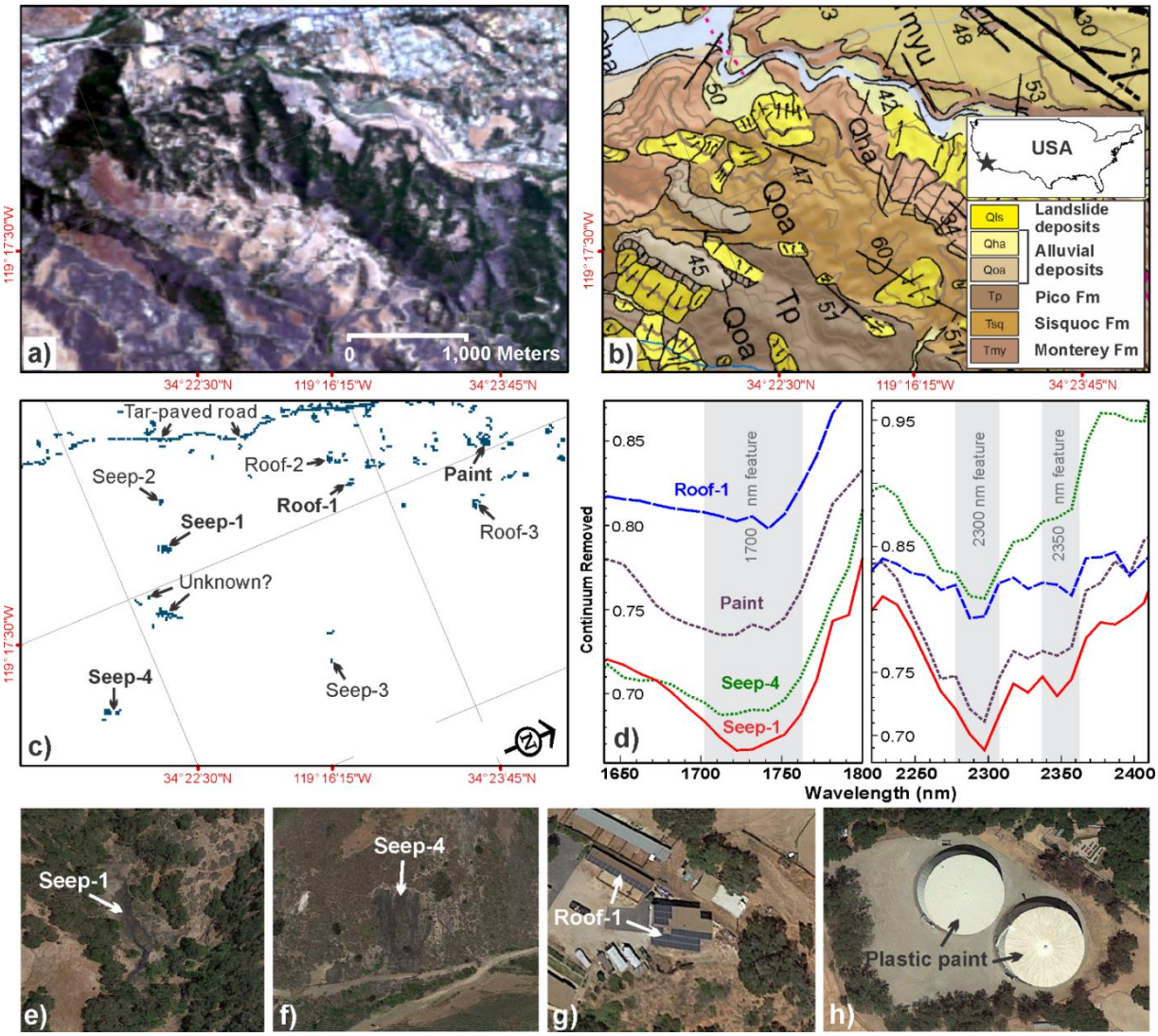

Fig 5. Macroseepage demonstration case study located in the Ventura Basin, California, USA. **a)** natural color composite of the AVIRIS data. **b)** geologic units of the area adapted from 1:100,000-scale geologic map of the Santa Barbara 30' × 60' quadrangle compiled by Gutierrez (2008). The location of the study area is shown in the inset map. **c)** anomaly map extracted from hyperspectral data following spectral analysis. Arrows indicate the location of the oil seeps and other spectrally similar petroleum-bearing compounds. **d)** continuum-removed reflectance spectra of four typical anomalies mapped in (c) achieved by averaging the relevant pixels in the hypercube data. The two major absorption features of petroleum centered at 1700 nm and 2300 nm are shown as gray columns in the left and right panels, respectively. **e-h)** high-resolution satellite imagery (GSD of ~1 m) of selected anomalies shown in (c) and (d), including oil seep-1 (e) located at coordinates 34°22'43.7"N and 119°17'40.5"W; oil seep-4 (f) located at 34°22'13.2"N and 119°16'59.5"W, roof-1 (g), and the plastic paint (?) of twin tanks (h). Source: Google Earth, imagery acquired in 1st May 1, 2015.

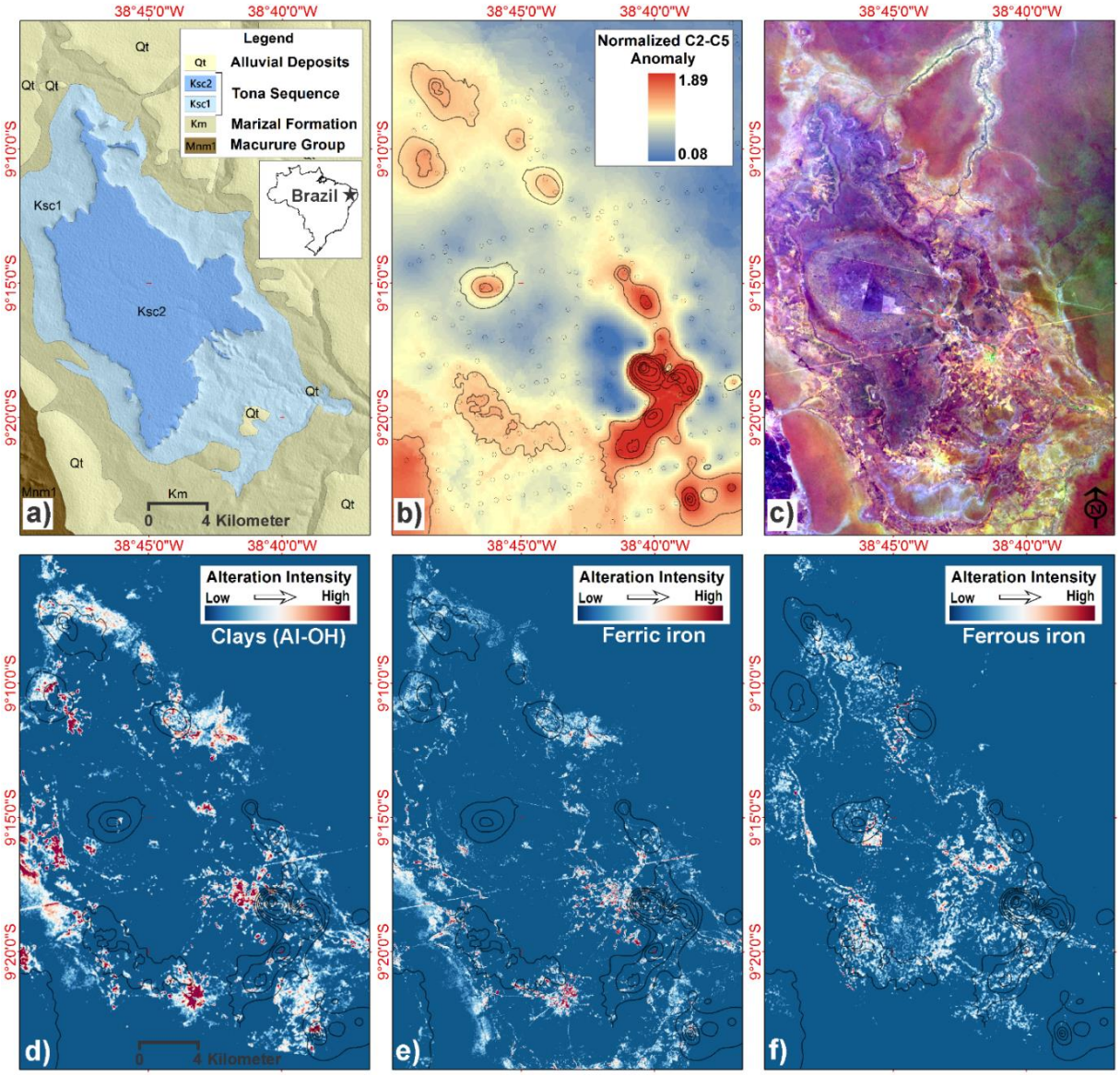

Fig 6. Microseepage demonstration case study located in the Northern Tucano basin, Bahia, Brazil. **a)**. generalized geologic map of the area. The location of the study area is shown in the inset map. **b)** Interpolated soil-gas anomaly map calculated by summing the normalized $C_2 - C_5$ readings. The sampling sites are shown by open circles. **c)** false color composite ($R = b7\ (2201), G = b5\ (864), B = b2\ (482\ nm)$) imagery of the Landsat-OLI data. **d)** relative abundance of clay alteration (including kaolinite, illite, and/or smectites) mapped using ASTER data. **e)** relative abundance of ferric iron minerals (e.g. goethite and hematite) extracted from Landsat data. **f)** relative abundance of ferrous iron minerals (e.g. Fe-carbonate, chlorite, and so forth) extracted from ASTER data. The overlaid contour lines in (d) to (f) were achieved by setting a threshold of 0.33 for the map shown in (b). The mineral abundances were calculated relative to image-derived endmembers.

2.6. Discussion

2.6.1. Direct sensing methods

The capability of remote sensing to detect and map oil and gas seeps can be used to: (i) screen frontier basins for any seepage manifestations, (ii) record the size, type, and possibly the likely replenishment of a leakage; (iii) update/map the seepage activity in mature/productive basins; (iv) assess the instant and long-term flux of the known seeps; and (v) compile global thematic databases for natural HC seeps.

Gas-plume sensing capability depends heavily on the seepage flux, wind speed, the specification of the deployed sensor (i.e. the spectral resolution and signal-to-noise ratio; SNR), and background cover, with spectrally uniform images being more advantageous relative to spectrally and thermally heterogeneous scenes (Frankenberg et al., 2016; Leifer et al., 2012b; Thorpe et al., 2014). Unlike the SWIR range, which is dependent on surface albedo, sensors in the LWIR range rely merely on the thermal emission and thermal contrast between ground and target gas. Hence, gas sensing in the LWIR could be more robust over a wider variety of land covers (Hulley et al., 2016). Indeed, simultaneous SWIR–LWIR data acquisition is required to investigate this notion. In both ranges, however, high spectral resolution data are required to distinguish the signatures of trace gas from interfering components (Hulley et al., 2016; Leifer et al., 2012b; Thorpe et al., 2016). Gas-plume sensing case studies thus far have been confined to methane detection that is highly significant for partitioning the sources of greenhouse gasses; albeit for oil and gas exploration, ethane (C_{2+}) constitutes a better exploration indicator (Jones and Drozd, 1983). Because methane emissions from geologic sources can incorporate 2–6% ethane on average (Etiope and Ciccioli, 2009), the recorded signal over natural plumes can be the overlap of methane and ethane signatures. The possibility of tracing ethane across geologic plumes is yet to be explored.

On the other hand, oil seepage detection is a function of seepage areal coverage, flowing rate, petroleum type, and geologic context, to name a few. Typically, petroleum sensing is bound to diagnostic absorption features centered at 1700 and 2300 nm (Fig. 5d), for the features centered at 1200 and 1400 nm are respectively uncommon and interfered by atmospheric water vapor. Within a seepage indicator, crude oil tends to mix physically with its background; thus, both of the absorption features are consistently modified by the mixed material(s). The 1700 nm feature is only overlapped by an absorption feature of sulfates (i.e. gypsum and alunite) centered at ~1750 nm (Fig. 7). However, in vegetated areas, the feature could be confused with non-photosynthetic vegetation. Furthermore, due to its proximity to water vapor band at 1900 nm, the long-wave side of the feature is prone to residual atmospheric contamination (Kokaly et al., 2013).

The feature at 2300 nm is particularly noticeable and persistent, albeit it is susceptible to be overlapped by clays and carbonates (Scafutto and Souza Filho, 2016) (Fig. 7). Although the feature centered at 1200 nm has been infrequently used for HC delineation (Clark et al., 2010), investigations have demonstrated its usefulness in differentiating false anomalies arising from plastic contaminations (Winkelmann, 2005). In the case study shown in Fig. 5, the limitations imposed by SNR of the data (likely due to poor atmospheric correction) hampered the efforts to accomplish this goal. However, marked differences in the 1700 nm feature of oil seeps and artificial compounds were noticed (left panel in Fig. 5d).

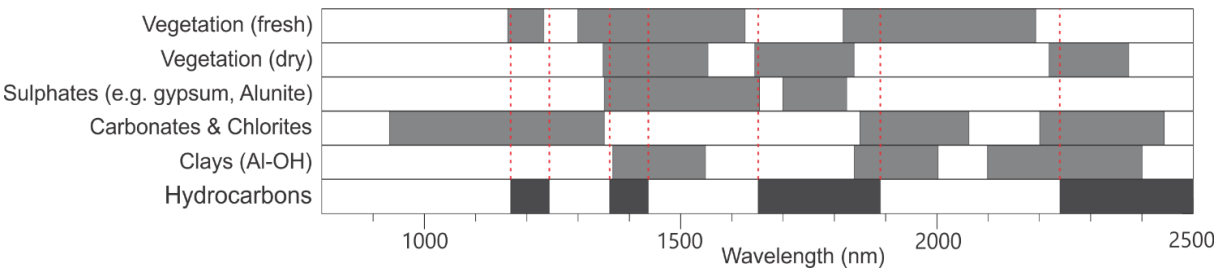

Fig 7. Spectral signature diagram of typical oil (obtained from Fig. 2) contrasted to absorption features of common sedimentary minerals as well as dry/fresh vegetation. The width of black bars indicate the relative widths of relevant absorption features.

To detect oil seeps effectively, a sensor should be equipped with proper spectral bands to resolve the features at 2300 and/or 1700 nm wavelengths (Fig. 5). For this reason, published case studies are confined to airborne systems (Table 4). Alternative spaceborne platform for this aim is due to be attained; albeit trials with WV-3 data has indicated the potentials of the sensor for this aim (Asadzadeh and Souza Filho, 2016a).

In terms of processing methodology, techniques like anomaly detection can be inadequate for HC mapping (Winkelmann, 2005). Instead, knowledge-based approaches (Asadzadeh and Souza Filho, 2016b) or hybrid methodologies that incorporate spectral-based decision-making system with feature tracking have been proven to yield superior results (Kokaly et al., 2013; Prelat et al., 2013) (see also the anomaly map in Fig. 5c).

The small size of oil seeps implies that high spatial resolution data (sub-decameter; Table 4) would be required to characterize them. The inclusion of alteration halos associated with a good portion of the seeps (i.e. miniseeps described in Table 6), however, can be used to facilitate their detection. A basic algorithm of this kind is developed by van der Werff (2006).

We should emphasize that petroleum detection includes but is not limited to the SWIR range. Crude oil is demonstrated to retain distinctive absorption features in the LWIR indeed (Lammoglia and Souza Filho, 2011). Nevertheless, due to the lack of a typical spectral library for crude oil diversity (and similar organic and petrochemical compounds) between 2.5–15 µm and also a limited number of hyperspectral sensors operating in this range, the potential applications of the LWIR for oil seepage detection has remained largely unknown.

2.6.1.1. The detection limit of HCs

The detection limit in remote sensing is defined as the smallest areal extent of a target (here a seep) within a pixel detectable spectrally. The detection limit for the case studies in Table 4 was reported to vary between 2.5 to 25% of a pixel. However, due to limited studies, this finding cannot be conclusive. In soil sciences, by comparison, the detection limit for the Total Petroleum Hydrocarbon (TPH) yielded from infrared spectroscopy varies from a few percent to a few hundredths of a percent (Chakraborty et al., 2015; Chakraborty et al., 2010; Correa Pabón and Souza Filho, 2016; Okparanma et al., 2014; Schwartz et al., 2013; Stallard et al., 1996). Correspondingly, the Total Bitumen Content (TBC) of oil-sand ores at around 1% wt. is routinely estimated via close-range spectroscopy (Lyder et al., 2010; Shaw and Kratochvil, 1990).

Some examinations have demonstrated that oil detection in the environment has a dependency upon physical and chemical properties of the soil matrix. For example, oil in siliciclastic or coarse-grained soils was easier to detect (at lower concentration levels) than in calcareous or fine-grained soils. The detection limit was also observed to be a function of API gravity, with heavier oils setting easier targets for spectral detection. The latter was likely due to the fact that heavier oils impregnated only the surface of the grains without being absorbed by porous media (Scafutto et al., 2016).

Overall, close-range investigations indicate that oil in the environment could be detected spectrally in very low quantities, albeit this finding cannot be directly linked to seepage reconnaissance. Apart from the fact that many studies were carried out by HCs other than crude oil,

the experimental settings incorporated instant physical mixing of oil with some predetermined soils in volumetric fashion, whereas seeps and their pollution counterparts appear as areally (or combined areally and volumetrically) mixed targets in the imagery. This could complicate the underlying principles of mixture analysis and affect the smallest extent of oiled surfaces detectable spectrally. Further research is required to appreciate the possibility of upscaling the lab findings to remote sensing imagery and figure out the practical detection limit of far-range systems.

Concerning the methane detection limit, in controlled release experiments, thermal sensing technology was shown to detect methane fluxes as small as 4–5 kg/h (Hulley et al., 2016; Tratt et al., 2014), whereas in the SWIR range, this level, depending on wind speed, was varied between 2–5 kg/h (Frankenberg et al., 2016; Thorpe et al., 2016). Based on field measurements, the leakage flux for high-level microseepage is >50 mg $m^{-2}d^{-1}$ that increases to an average of 2×10^4 gd^{-1} (~1 kg/h) for macroseeps (Etiope, 2015). It means that current technology can detect macroseeps that are emitting average to high levels of gaseous HCs, albeit it is not appropriate to detect microseepage-level fluxes. Above all, available technology can only afford to detect distinct point sources. For areal (diffused) sources, which are most common in petroliferous areas, a sensor with higher sensitivity (lower detection limit) would be required.

2.6.1.2. Petroleum quantification

In the seepage context, spectral techniques can be employed to quantify the dimension of a seep and possibly the amount of HCs in the environment. In close-range experiments, the TPH is assessed by a predictive model initially derived from a suite of training samples using the Partial Least Squares Regression (PLSR) technique (e.g. (Okparanma et al., 2014)). This approach, however, is not appealing to image data mostly because it is case dependent and requires auxiliary data. An alternative approach is to use the continuum band-depth of HC's diagnostic absorption features for abundance quantification in relative or absolute terms (Asadzadeh and Souza Filho, 2016b).

The few studies on the use of 1700 and 2300 nm features have indicated the merits of spectral parameters, especially the 'depth' and 'area' of the absorption, for HC quantification (Correa Pabón and Souza Filho, 2016; Scafutto et al., 2016). Nonetheless, the significance of each feature for this aim has been a matter of debate. For instance, in an analysis based on the correlation coefficients of different wavelet scales with TBC, it was demonstrated that the feature at 2300 nm appears correlated in several scales (from 1–7), whereas the 1700 nm feature shows correlation with only two (4 and 5) scales (Lyder et al., 2010). Based on this persistency, it was concluded that the 2300 nm feature is more successful in delineating the HC content of oil-sand samples. In the demonstration case study (Fig. 5a), the depth of these features was calculated after continuum removal, but due to the low spatial resolution of the imagery (~16 m), it was difficult to identify a meaningful trend and thus was discarded.

Even though the depth of the feature at 1700 nm was shown to be linearly correlated to specific levels of HCs (Correa Pabón and Souza Filho, 2016), in general, the correlation between HC content and the noted spectral parameters tends to exhibit nonlinear behaviors. Experiments with the incremental level of oil also showed that in a certain level, the soil sample becomes saturated, thereby breaking the correlative relationship. Identical to the detection limit, the saturation point was proved to be a function of oil type and intrinsic soil characters (Scafutto et al., 2016). By comparison, the feature at 1700 nm reaches to saturation at higher HC levels than the 2300 nm

feature. A related point to consider is that the saturation level was observed in volumetric experiments and it is not known as if it would be the case for the areal mixture.

In conclusion, such observations imply that potential (nonlinear) predictive models for HC quantification from spectral data would be valid between the 'detection limit' and 'saturation point'. Further research is required to better understand the behavior of noted spectral parameters against different levels of petroleum in a sample/pixel.

2.6.1.3. Petroleum characterization

Infrared spectroscopy is unlikely to be able to identify petroleum types (e.g. kerosene vs. diesel), particularly where it is mixed with soil particles. Instead, it was proved to be effective in characterizing oil types based on API gravity (a measure of how heavy liquid HC is, compared to water) and SARA (Saturate, Aromatic, Resin, and Asphalten) fractionation index (Aske et al., 2001; Lammoglia and Souza Filho, 2011). The approach incorporates a mathematical model experimentally constructed using a sort of correlation algorithm (e.g. PLSR) to predict the noted chemical properties. Although the methodologies have been successful in characterizing the API and SARA index of given oils in the lab, so far, only the API has been remotely assessed (mapped) in the marine environment using multispectral ASTER data (Lammoglia and Souza Filho, 2012) and similar onshore demonstration is yet to be provided. The latter, on the other hand, has remained a laboratory experiment and not attempted via image data. Although the spectroscopic-based API model was developed using a specific collection of crude oils (°API between 14–40), the underlying principle can be extrapolated to other compounds like fuel HCs.

2.6.2. Indirect RS methods

In oil and gas exploration, remote sensing could be utilized to screen a frontier basin for signs of a petroleum system, generate new exploration prospects, and evaluate the existence of microseepage anomaly over previously defined exploration leads (van Der Meer et al., 2002). Alike other surveying methods relying on microseepage theory, this approach neither can indicate the depth of an accumulation nor predict the economic success of wildcat drilling. In contrast to the soil-gas method that can directly verify the presence of HC, this technique should be applied to 'infer' the HC presence in a trap (Jones and Drozd, 1983; Schumacher, 1999). This technique shall be effective in dealing with bedrocks covered by sparse vegetation (i.e. <30%). Above this threshold, the use of mineralogical signature becomes limited and one should consider other realms of remote sensing including geobotany for microseepage detection. Obviously, spectral technique merely measures the reflected/emitted energy from topmost parts of the surface and is unable to penetrate into it.

A close inspection of the documented case studies in Table 6 reveals a large gap between the true potentials of remote sensing and the achieved results. Here, we highlight several of the shortcomings and subsequently give some guidelines for further case studies in the future.

2.6.2.1. The shortcomings of the case studies

Unfortunately, the number of well-documented publications on the subject, as summarized in this paper, is not considerable (<30 cases), while a large proportion of them was dedicated to miniseeps that are indications of macroseeps rather than microseepage systems. A typical remote

sensing study in local/regional scale should be augmented by detailed fieldwork and lab analysis. Nevertheless, a look at Table 6 indicates that the number of collected samples in each case study is inadequate and the sampling schemes are mostly disordered (Figs. 4d-4e), resulting in poor relationships between sampling sites and remote sensing anomalies. The benefits of detailed outcrop studies as a missing link in such investigations is best exemplified by the outlined studies in Table 7.

Because target minerals of a microseepage system are already common in sedimentary basins, it is important to draw the comparison in a relative term. In other words, a microseepage-affected zone should be compared to a 'reference area' inside the intact (unaffected) extent of the same lithologic unit. This effective strategy, which is commonly used to determine the background in geophysical and geochemical surveys (Abrams, 2005; Saunders et al., 1991), has been adopted by very few investigations (Fig. 4c). In addition, due to basic processing algorithms (Fig. 3o), the approach was only able to separate HC affected zones (anomaly) from their surroundings (background), without delineating individual minerals of the assemblage (Fig. 4a). To accomplish this objective and complement the results, spectroscopy (*in situ* or in the lab) could be used, but in the reviewed cases, this capacity has been rather underutilized. Spectroscopy not only can facilitate the identification of minerals, but can also be utilized to characterize slight compositional changes present in clays, carbonates, iron oxides/oxyhydroxides, and sulfates with reasonable accuracy.

A couple of studies have already shown that microseepage-induced alterations could be associated with spatial zoning. For example, some comprehensive fieldworks denoted that an extensively bleached area in red-beds surrounds an intense carbonate cementation zone along with pyrite patches occurring over productive limits of the reservoir (Al Shaieb et al., 1994). Despite the capability of image data to reveal possible spatial patterns of alterations, due to the scarcity of similarly detailed accounts, the zoning within microseepage systems has remained largely unknown. Furthermore, the spatial relationships between surface mineralogy relative to the 3D architecture of a reservoir are poorly understood. The only clue in this regard provided by measurements of magnetic susceptibility of exploration drill cuttings demonstrated that authigenic magnetic minerals (i.e. maghemite) are mostly distributed in depth range of 60–600 m above HC reservoirs (Foote, 1992; Foote, 1996; LeSchack and Van Alstine, 2002). Comparative studies between remote sensing anomalies and seismic data (e.g. (Shi et al., 2012)) or drill cuttings can be an asset for this aim.

All the outlined case studies investigate charged (productive/prospective) plays and none is focused on failure cases or unproven/dry prospects to have a fair unbiased estimate for false-positive and false-negative anomalies (see the next section). From a mineralogical point of view, it is entirely unclear if secondary anomalies are present where a trap retains no accumulation, or if the features are unique to already charged traps. Based on exhumed reservoir case studies, however, it seems that several (but not all) of the classic characteristics (including bleaching) are shared between fossil (paleo) and active microseepage systems. Additionally, the characters of underlying traps including underpressured vs. overpressured reservoirs (given that pressure is the main drive for HC migration), stratigraphic vs. structural traps, and the sealing are not yet properly addressed.

The effect of microseepage on sediment covers (e.g., transported floodplain alluvium and loess) is not extensively evaluated in the literature. Although some have reported that transported covers (i.e. glacial drifts) can obscure the microseepage-induced radiometric signals (Price, 1996; Saunders and Terry, 1985), there are numerous other studies that denote it is not the case (e.g. (LeSchack,

1997)). The surficial nature of radiometric survey (<25cm penetration depth) implies that secondary alterations can extend into a cover (provided a physically stable cover), and thus remote sensing approach should be able to detect the effects of microseepage in a similar manner. Some observations on the pace of mineralogical transformations can support such notion (see section 6-2-5).

So far, the outlined case studies deal with nearly mono-lithologic sandstone units (Fig. 3e). Although clastic rocks and oxidized continental facies (red-beds) form at the latest depositional cycle of a sedimentary basin, they cannot represent the whole variety of the bedrocks present above HC accumulations. Sandstones at best account for around 25% of the sedimentary rocks (Boggs, 2009) and most likely cover the sedimentary basins with the same proportion. Nonetheless, a ubiquitous proportion of HC traps are either overlaid by sediments other than sandstones, are multi-lithologic in essence, or are concealed by recent sediment covers. To have a complete picture of the variations, studies shall be diverted towards more diverse lithologies exposed over HC accumulations. The sandstones themselves retain large within-group variations that are neither considered in remote sensing studies nor in the classic microseepage model (see also section 6-2-4).

2.6.2.2. False-positive/negative anomalies

While there are statistics for positive and negative geochemical anomalies above exploration targets (section 2-3), such data have not been compiled for mineralogical-based remote sensing studies (see also section 6-2-8). In a sedimentary basin, any false-positive anomaly should be regarded as important as true anomalies, because a cross-comparison between these two groups can provide insights about how to distinguish them spectrally. As discussed by Brown (2000), present (and past) surface geochemical anomalies and associated alterations may result from uneconomic petroleum accumulations or migration pathways. Whereas remote sensing cannot distinguish between active microseepage systems arising from uneconomic and economic accumulations, it can be potentially used to exclude some other sources of false-positive anomalies. Generally, the following sources of false-positive anomalies in sedimentary basins should be recognized: (i) anomalies arising from processes unrelated to HC accumulations; (ii) anomalies occurring above depleted traps that once retained an active microseepage system; and (iii) anomalies intrinsic to sedimentary units.

In the first group, a number of already known geological processes give rise to roughly similar alteration assemblages. For example, the calcite cement in several soil profiles may be ‘caliche’, which is an amorphous soil salt originating from near-surface processes (Price, 1996). Pedogenic processes linked to climate conditions can initiate redox zones and precipitate maghemite, thereby yielding characters that may bear a resemblance to microseepage effects (Klusman, 2002; Schumacher, 1996). Similar alterations can also arise from shallow gasses of biogenic rather than thermogenic sources (Schumacher, 1999). Moreover, the circulation of abnormal amounts of H_2S, CO_2, and organic acids in permeable sediments may develop signatures similar to microseepage-induced alterations (Parry et al., 2004; Schumacher, 1996) (see also Table 7).

As petroleum system implies, HCs are preserved in a trap if the sum of leakage or destruction is less than the petroleum charge; otherwise, a trap would be devoid of HCs (Hunt, 1996; Magoon and Beaumont, 1999). Therefore, paleo-microseepage systems over depleted traps would possibly leave mineralogical footprints that are similar to active systems. The best analogy where this phenomenon

could be investigated is exhumed HC reservoirs cropping out throughout the Colorado Plateau (Table 7). This phenomenon substantiates that when the intensity of alteration is severe, the induced changes can survive through geologic time and emerge as false-positive anomalies in present-day surveys. Parallel outcrop studies also demonstrated that several episodes of fluid flow (meteoric and interstitial water and migrating HCs) could coincide to shape the final alteration facies (see Table 7). Finally, because the seeking diagenetic minerals (e.g. clays and carbonates) are intrinsically abundant in the background strata, occasionally, they could become a source of false-positive anomalies. Such anomalies are subject to inter-, and intra-unit variations.

All of the abovementioned instances create ambiguity in remote sensing data interpretations. Whereas part of the ambiguity is intrinsic to the approach and unavoidable, we believe the other part could be avoided by applying an efficient methodology, enhanced imaging tools, and above all, by promoting our understanding of the diagenetic changes. A case in point is the problem of alike background mineralogy that could be circumvented by setting a reference area off the affected zone for cross-comparison. The gained experience in mineral exploration implies that spectroscopic products including abundance, composition, and association of minerals derived from hyperspectral data may offer potentials for segregating real anomalies from false ones (see also section 6-2-6). In this regard, the authors believe that comparative studies between active microseepage systems and exhumed HC reservoirs could facilitate the development of mineralogical indicators to eliminate false-positive anomalies over fossil systems. Altogether, true anomalies typically are expected to cover vast areas and conform to regular spatial patterns (Fig. 6; see also section 6-2-7).

In the literature, there is no unbiased evaluation of remotely-sensed false-negative anomalies. To assess this parameter, the outcomes of this technique should be compared to geochemical data or evaluated by wildcat drilling indeed.

2.6.2.3. Sensor obstacles

Part of the reasons for the limited success of remote sensing in microseepage delineation is due to limitations imposed by multispectral sensors. In contrast to technological refinement in geochemical and geophysical analytical methods, the operational sensing capability of this approach has not witnessed dramatic advancements in the last decades. In Earth resource exploration, in particular, the only promotion in three decades has been the launch of ASTER multispectral instrument and more powerful hyperspectral technology has been pursued solely from airborne platforms. Newer satellite systems with better performances such as Sentinel-2 (2015) and WV-3 (2014), which could be potentially used for enhanced spectral mapping, were launched only recently. A leap that is due to revolutionize this discipline, however, is the arrival of spaceborne hyperspectral imaging systems like the EnMap satellite (http://www.enmap.org/). Because microseepage phenomena are bound to yield large targets, they would be fully resolved with moderate spatial resolution (i.e. 30 m) of such instruments. In the case study shown in Fig. 6, the 30 m resolution of ASTER/Landsat data were enough to resolve the targets, however, due to a limited number of bands, the discrimination of mineral species was not straightforward.

Thus far, the majority of published studies for microseepage characterization are limited to the VNIR–SWIR window and the literature is deficient in case studies using thermal infrared (LWIR) data. A suite of diagenetic changes in the chimney column such as silica enrichment (Almeida-Filho et al., 2002; Thompson et al., 1994) or feldspars depletion (alteration) are solely mappable via this dataset.

The LWIR wavelength or an integrated multi-wavelength spectral analysis can also promote the mapping of aforementioned alterations with greater accuracy.

2.6.2.4. The inadequacy of the microseepage model

In remote sensing data analysis, the microseepage model is used to define target minerals and subsequently to attribute the mapped minerals to microseepage effects. We believe the inadequacy of the present model is causing considerable uncertainties in the interpretation of resultant maps because it fails to fully delineate the quantity, quality, and diversity of mineralogical components present in a system. The existing theory, which evolved during the 1970s and 80s principally to explain anomalous gas levels over HC accumulations, was ultimately formulated in 1990s to account for alteration facies in a general term (Saunders et al., 1999; Schumacher, 1996; Thompson et al., 1994). Unfortunately, this model simply assumes global identical features for the phenomenon without taking into consideration the effects of local environmental variables (i.e. climate, hydrogeology, erosion, and soil-forming processes) and geologic setting (i.e. lithology and stratigraphy of the near-surface units) in which the microseepage occurs.

Microseepage and hydrothermal processes are somehow comparable phenomena. In a hydrothermal system, magma, as a source of energy, triggers fluid circulations and oxidation-reduction reactions in the host rock, leading to wall-rock alterations and ore deposition (Pirajno, 2009). In both systems, fractures provide the conduits for fluid transportation and subsequent mineral deposition, with macroscopic fractures playing a critical role in the overall mass transfer. Using this analogy, one can expect a similar trend in the diversity of alteration products within a chimney column, of which the collection described by the present model representing only a fraction. We postulate that upon providing a bigger picture of the mineral diversity, they could be categorized into a series of descriptive models, each tailored for specific geological settings. Such a multiplicity would then facilitate spectral processing and likely would contribute towards more efficient HC exploration. Below we attempt to outline some aspects of the mineralogy not reflected in the existing model.

The accepted notion about diagenetic clay minerals is based on kaolinite enrichment, but as we discussed earlier (e.g. Fig. 3j), a suite of indicative clays consisting of illite, smectites, and chlorites can be expected above HC accumulations. For instance, in Fig. 6f, the ferrous mineral map is expected to incorporate chlorite as well. Moreover, sulfates constitute a key mineralogical signature that has been overlooked altogether in the classic microseepage model. Due to a steady influx of O_2, the vadose zone above the water table is strongly oxidized, thus many of the reduced minerals, including pyrite and Fe-carbonates, become unstable and eventually weather (partially or fully) to other more stable forms. The oxidation of pyrite, in particular, can trigger the formation of several metastable pH-sensitive iron sulfates such as jarosite, copiapite, melanterite, and schwertmannite in the system. Subsequent oxidation of this array gives rise to a series of ferric iron oxyhydroxides in the diagenetic facies, including ferrihydrite, maghemite, lepidocrocite, goethite, and occasionally secondary hematite (Elwood Madden et al., 2004). Unlike pyrite, all the subsequent sulfates and oxides are spectrally active in the VNIR–SWIR windows, thus are potentially detectable remotely (Crowley et al., 2003). A case in point is the suite of Fe-sulfate and oxide minerals successfully mapped using AVIRIS data over the Dutton Basin Anticline (Staskowski et al., 2004). In a very arid environment, however, jarosite has reported to endure weathering and constitute a potential target over several microseepage-related facies (Bell et al., 2010; Elwood Madden et al., 2004; Everett et

al., 2002; Perry and Kruse, 2010). Ultimately, secondary hematite cement formed due to severe and prolonged oxidation of initial ferrous minerals have been observed in several oil and gas fields (Donovan et al., 1979-1982; Kirkland et al., 1995; Segal et al., 1986). The weathering, in a similar manner, can be responsible for the absence of magnetic bodies (e.g. maghemite) in near-surface (<60 m on average) section of the strata as recorded by well cuttings measurements (Foote, 1992; Foote, 1996).

Although many have emphasized the combined role of pH/Eh in diagenetic changes, the pH of the soils and sediments is not commonly reported (Fig. 4f) and its contribution in shaping the mineralogical facies is not fully addressed. This parameter is highly significant for it controls the mineralogy of clays and regulates the species of iron oxyhydroxides in the strata (Marrs and Paylor., 1987).

2.6.2.5. Microseepage within time

Microseepage anomaly, as scrutinized by soil-gas and microbial techniques, is a dynamic and rapid phenomenon that can appear/disappear within several months (Rice et al., 2002; Schumacher, 2000; Tedesco, 1999). By contrast, the pace of mineralogical changes over chimney column has not yet been realized, mainly because of the lack of systematic investigations and different transformation timescales. Some sporadic observations, however, have demonstrated that the pace could be very quick in geologic time. For instance, a study conducted over a gas-storage area revealed that anomalous magnetic susceptibility in soil appears in less than 24 yrs. (Saunders et al., 1991). Observations in sanitary landfills also have revealed that mineralogical transformation above a seeping methane source, which supposedly bears some resemblance to microseepage system, is rapid and happens in less than a decade or two (Ellwood and Burkart, 1996). In a quite exceptional case, more rapid transformation, in the order of several months, has been witnessed in a dam construction area (Prokopovich et al., 1971). Such rapid rate of changes conforms to our observations of soil bleaching by the gas flow in a controlled environment.

According to Saunders (1999), microseepage is a dynamic phenomenon and gradually moves from the center of the accumulation to the outer edges. When the gas anomaly merely shifts or becomes extinct altogether, the evidence of the leakage, including anomalous secondary minerals, are preserved and accumulated in the sediments as ‘fossil’ geochemical anomalies (Saunders et al., 1991). Due to the accumulative nature of such fossilized footprints within time, they typically constitute major targets for remote sensing, even though it can bring about false-positive anomalies over depleted traps. Subsequent weathering/erosion is likely to fade the footprints out, albeit profound authigenic changes, much like exhumed reservoirs (Table 7), would survive into the present time. As an example, the spectral anomaly yielded from satellite data over an oil field in Tennessee has been observed to diminish over 25 years of production (Perry, 2006). As stated earlier, novel mineralogical indicators should be quested to tell the active and fossil systems apart. A relevant point to consider is that the secondary cementation of the strata and the accumulative character of diagenetic changes are the likely reasons behind some disagreements between soil-gas and mineralogical anomalies (Klusman and Saeed, 1996). In the case study shown in Fig. 6, however, the soil-gas and alteration signatures maintain a close spatial relationship (e.g. Fig. 6d).

2.6.2.6. The necessity for quantitative mineral maps

In the literature on microseepage topic, it is uncommon to account quantitatively for mineralogical changes. A typical example is the 'bleaching' term regularly used to indicate the loss of ferric iron without attempting to quantify the original and final levels. Based on few reports available, anomalous terrains were recognized by an increase in clay content between two to five times; sulfide content between two to three times; and 45% more total carbonates (Marrs and Paylor., 1987; Schumacher, 1996; Segal and Merin, 1989). In absolute term, such anomalies were characterized by a reduction in iron content from 2.9% in the periphery to just 1.0% in the center of the alteration; a concentration of up to 2.0% wt. kaolinite; the occurrence of 1.6–5.7% sulfides; and 40–50% added carbonate cement (Donovan et al., 1975; Kirkland et al., 1995; Schumacher, 1996).

In order to increase the efficiency of remote sensing for microseepage delineation, we recommend adopting a quantitative approach for spectral processing using methodologies embedded in either knowledge-based or data-driven approaches (Asadzadeh and Souza Filho, 2016b). For example, it is possible to achieve a semi-quantitative map for hematite, kaolinite, and calcite by calculating the depth of the diagnostic absorption features centered at ~850, ~2200, and 2340 nm, respectively (e.g. Fig. 6d-6f). To eliminate the effect of background mineralogy and intra-unit variations, however, the interpretation shall be based upon the concept of 'reference area' introduced earlier.

The lack of quantitative maps has also hampered the efforts to establish relationships between alteration intensity/zoning and seeping activity over HC pools. To unravel such likely relationships, quantitative remote sensing results should be interpreted in the context of geochemical anomalies or wildcat drilling (see section 6-2-9).

Apart from abundance, spectral data can be used to map the variations in composition and crystallinity of selected minerals. A case in point is the capability of hyperspectral data to track the shift in the wavelength of absorption minimum at ~900 nm to discriminate between different iron oxide/oxyhydroxide species, namely maghemite, goethite, hematite, and ferrihydrite. This unexplored capability has important implications for exploration, because maghemite, for instance, is known to be responsible for the bulk of micromagnetic anomalies detected by geophysical surveys (Foote, 1996; Foote, 2013; Holysh and Toth, 1996; Saunders et al., 1991). Moreover, due to the sensitivity of Fe-bearing minerals to the pH of the environment (Swayze et al., 2000), a map like the one noted could be used to predict the pH over microseepage systems. Similar spectral products could be developed to differentiate carbonate species (dolomite, siderite, ankerite, calcite, and rhodochrosite) using the wavelength of the absorption feature between 2320–2350 nm (Gaffey, 1987). Possible variations in the physicochemistry of clays comprising illite-smectites, chlorites, and kaolinite (see (Asadzadeh and Souza Filho, 2016b) for the details) has not been deeply investigated; albeit in the case of chlorite, some well data analysis has shown dominant occurrence of Fe-rich chlorite in bleached zones (Weibel, 1998). More research is required to uncover the full potential of such spectral products for microseepage exploration and indeed false-positive elimination.

2.6.2.7. The shape of anomalies

A proper understanding of the shape of the microseepage-induced anomalies and their compliance with underlying accumulation is critical to oil and gas exploration (Jones and Drozd, 1983; Xuejing, 1992). Because the overall shape of anomalies largely depends upon the employed prospecting tool, here we first overview the known shapes yielded from other exploration methods and then discuss the expected and observed anomalies from image data.

Soil-gas anomalies can have various shapes that in the order of importance are: (i) *halo* (annular/ring) pattern, in which anomalous readings are aligned over the edges of the underlying accumulation (Figs. 8a and Fig. 6b); (ii) *apical* (bell-shaped), in which an area of high values is surrounded by a low background (Fig. 8b); (iii) *crescent* type, which is found over accumulations trapped against a fault; and (iv) *linear*, which is observed over surface traces of faults, likely due to the effusion of HCs along fault planes (Duchscherer, 1980; Duchscherer, 1982; Horvitz, 1980; Price, 1986; Price, 1996; Xuejing, 1992).

The noted patterns may be either continuous or discontinuous (patchy/broken), consisting of several unconnected patches (Fig. 8 and Fig. 6b) over an accumulation (Xuejing, 1992). Typically, due to offsets between the location of an anomaly and the underlying reservoir, geochemical anomaly is unlikely to closely resemble the shape of the underlying reservoir (Brown, 2000; Holysh and Toth, 1996; Jones and Drozd, 1983; Richers et al., 1986; Saunders et al., 1999; Thrasher et al., 1996).

Likewise, (micro)magnetic anomalies over HC accumulations can have four different types: (i) double hump anomaly, corresponding to more intensive values on the edges; (ii) ripples, which is composed of high frequency signals superimposed on higher amplitude regional trend; (iii) positive anomaly that is composed of increased magnetism over the entire field; and (iv) negative anomaly due to pyrite precipitation (Eventov, 2000). Radiometric anomalies, on the other hand, are reported to occur either as halo or saddle-shaped (apical) forms over reservoirs (Fig. 1). Whereas geophysical anomalies are not highly accurate in outlining accumulations (Saunders et al., 1993b; Saunders et al., 1999), microbial prospecting is known to yield apical anomaly and coincide exactly with the extent of the field (Price, 1986; Price, 1996).

Overall, the ‘halo anomaly’ is acknowledged to be the most common form among exploration techniques (see Fig. 6). Several reasons have been given to explain the occurrence of this pattern including: (i) higher density of fractures over the edges of a structure, (ii) lower bacterial activity over the edges (relative to apical zone), (iii) conformity to gas-water and oil-water contact, and (iv) caprock clogging (Eventov, 2000; Horvitz, 1980; Price, 1986; Saunders et al., 1999). Based on numerical modeling, Brown (2000) concluded that the halo anomaly and irregular distribution of apical anomalies are mainly due to fracture distribution. The locality of such anomalies is believed to overlap with the areas of maximum stress in a structure (Eventov, 2000).

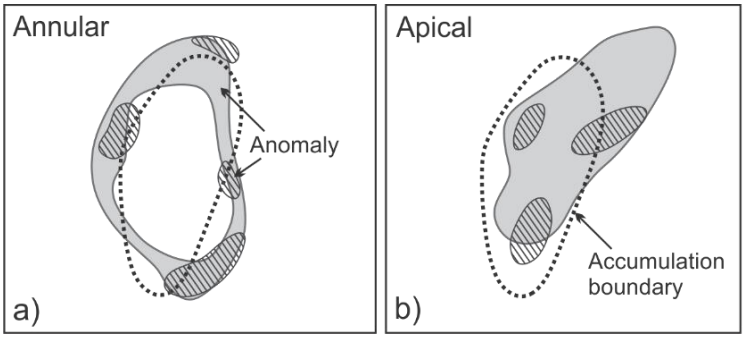

Fig. 8. Common surface geochemical anomalies over HC accumulations. **a)** Annular (halo). **b)** Apical. Solid and hatched fills indicate continuous and broken anomalies, respectively (Modified after (Xuejing, 1992)).

Based on field observations, a number of patterns for microseepage-induced mineralogy are already proposed as is illustrated in Fig. 1. However, in the absence of quantitative methodology, several of the illustrated forms have not been replicated via spectral data. Although some have pointed that the alteration anomalies over HC accumulations would exhibit a halo pattern (Klusman, 2002) (see also the case study in Fig. 6), in reality, the clogging of escaping routes (Saunders et al., 1999) along with the cumulative effect of alterations imply that the patterns should incorporate both apical and halo shapes. In other words, it could resemble an extended apical anomaly that depending on the mineralogy type will show positive or negative configuration. On the other hand, according to the descriptions provided in Table 6, it seems the 'disconnected apical' is the most likely pattern for mineralogical manifestations.

Saunders (1993a) has stated that "radiometric anomalies may be found only over portions of fields and thus it cannot be used to determine the production boundary accurately". In a similar way, it is rather unlikely that mineralogical indicators accurately conform to the shape of the underlying reservoir, albeit it is still a useful tool to locate microseepage systems arising from HC accumulations.

2.6.2.8. The overall efficiency of remote sensing approach

Owing to the fact that failure cases are absent in the reviewed case studies, an overall assessment of the efficiency of remote sensing in oil and gas exploration, much like geochemical method (e.g. (Schumacher, 2010)), is not feasible. Therefore in this section, we only overview the accounted success rates appeared in the literature.

In the early days of satellite data, it was reported that among the 57 hazy anomalies extracted from ERTS imagery in the Anadarko basin, 42 coincide with producing oil fields, corresponding to 73% coincidence (Short, 1977). In the era of Landsat, 75% of all mapped tonal anomalies (equivalent to 59 individual anomalies) were shown to be associated with economic HC plays (Feder, 1985). After drilling through 1177 geomorphic/tonal anomalies achieved from Landsat data in the Rocky Mountain, the average efficiency was assessed to be 54% (Land, 1996), whereas in search of stratigraphic traps by an integrated exploration approach, success rates were reported to vary between 29–53% (Saunders et al., 1999). In an objective assessment of unconventional exploration methods, the performance of Landsat data in predicting the outcome of wildcat drilling over conventionally generated prospects was evaluated to be better than 75%, which placed it among the best prospecting techniques evaluated (Calhoun, 1991).

In the era of 'spectral anomaly' extracted from ASTER data, 85% coincidence (18 out of 21) between spectral signatures and follow up fieldwork and geochemical sampling was recorded (Everett et al., 2002). In a recent study conducted in the Lake Albert basin in the East Africa Rift system, mineralogical anomalies due to HC migration were mapped using multisensory and multi-temporal satellite data (Table 6). Remarkably, spectral anomalies were detected in 18 out of 19 discovered (18 oil + 1 gas) fields, corresponding to 95% success rate (Frassy et al., 2015). This study, however, neither provided the characteristics of the anomalies nor the results of probable field verifications.

Conceptually, the performance of spectral techniques could be compared to radiometric and magnetic methods, as both aim to detect microseepage-induced mineralogical indicators in a direct way. As an example, after investigating the radiometric data collected over 706 oil and gas fields in the US, it was discovered that 72.7% of them are associated with typical K and U anomalies (Saunders et al., 1993a). Based on other accounts, the radiometric and magnetic techniques were reported to correctly predict a production hole, respectively, at 59–85% and 58–75% of the times (Land, 1996; Potter II et al., 1996). With improvements in analytical instrumentations and processing methodology, higher success rates were witnessed by the noted techniques. A case in point is the statistics yielded from micromagnetic surveys indicating that over 80% of producing petroleum fields are associated with anomalous magnetic susceptibilities (Foote, 1992; Saunders et al., 1991; Wolleben and Greenlee, 2002). By using modern sensing technologies and processing methods, we can expect similar incremental trends in the performance of spectral remote sensing.

Microseepage-based prospecting tools are reportedly successful in predicting dry holes. For instance, 29 out of 30 (96%) or 15 out of 19 (79%) negative geochemical anomalies drilled in the field were ended up to be dry holes (Davidson, 2004). The geophysical methods, on the other hand, have been 70–90% successful, on average, in predicting a dry hole (Foote, 1992; Potter II et al., 1996). A similar calculation is yet to be carried out for remote sensing studies.

2.6.2.9. Guidelines for future studies

The study of microseepage phenomenon shall incorporate multi-scale data from fieldwork, lab analysis, and regional surveys. Outcrop investigation in this sequence, which includes but is not limited to sampling, maintains the essential link between regional and microscopic studies. To our knowledge, sampling along profiles is the best approach to intersect the bulk of variations induced by HC microseepage. We recommend establishing at least one reference area in unaffected (intact) part of the target unit(s) for cross-comparison of the induced changes. Certainly, this approach would require *a priori* knowledge about the stratigraphy of the units provided by a detailed geologic map. The overall number of samples collected from on/off zones should be large enough (typically >30) to reflect the concurrent effects of intra-unit and induced variations within a given area. Instead of simply differentiating anomaly from the background, it is more appropriate to characterize and quantify the mineralogy of anomalous zones (e.g. Fig. 6) by employing a proper spectral technique (Asadzadeh and Souza Filho, 2016b). This approach would help determine the shape of anomaly and reveal the spatial relationships between the mapped anomaly and possible subsurface pool.

The best supplementary analytical methods to corroborate remote sensing studies are those that provide fresh insights into the particularity of minerals in a system. Hence, beside the indispensable reflectance (and emittance) spectroscopy, the investigation should be rather complemented by XRD,

optical microscopy, and Mössbaur spectroscopy. The latter has superiority in identifying the species of ferric/ferrous minerals in given samples. A deeper understanding of fluid-rock interactions, however, would be achieved when the studies are accompanied by stable isotope geochemistry.

Eventually, remote sensing data shall not be used in isolation. The integration of resultant maps into other non-seismic prospecting techniques not only can provide new insights into microseepage phenomenon, but can also add more value to the existing exploration data and improve the success of subsequent drilling (Prelat et al., 2013; Rice et al., 2016; Saunders et al., 1999).

2.7. Seepages and the environment

Although seeps have been a topic of concern to explorationists for a long time, until recently, there was little understanding of their role in carbon emission to the atmosphere. Based on recent assessments, macro-, and microseeps in total constitute the second most important sources of natural methane (and also ethane/propane) emissions to the air (Etiope and Ciccioli, 2009; Etiope and Klusman, 2010; Etiope et al., 2008). It has been estimated that between 21–36% of the geologic methane budget is emitted by seeps, of which microseepage-prone areas and macroseeps emit 10–25% and 11%, respectively (Etiope, 2015). The revised global budget via isotope data, however, suggested that methane emission from geologic sources is 60–110% greater than current estimates (Schwietzke et al., 2016). The estimation of macroseepage flux has been relatively straightforward as a large portion of them are already known and monitored in the field. However, the global emission of microseepage is still uncertain because the calculations are based on averaging field contributions from identifiable homogeneous areas. The current flux is based on a database of 563 measurements in dry soils (Etiope and Klusman, 2010).

The potential microseepage-prone areas in the globe are estimated to be in the order of 3.5–4.2 million Km^2 distributed within 937 petroliferous provinces or basins in 112 countries (Etiope and Klusman, 2010) that are approximately equivalent to 7% of the global dryland areas. Positive fluxes are typically a few to tens of mg $m^{-2}d^{-1}$ that can reach hundreds of mg $m^{-2}d^{-1}$ over widely tectonized zones (Etiope, 2015). Etiope and Klusman (2010) admitted that their calculation should be considered as a first spatial disaggregation of emission factors. They stated, "…*the uncertainties in global emission estimates are mainly due to a poor knowledge of the dryland area of invisible microseepage*". It is evident that all microseepage terrains occur within petroliferous provinces, but so far the actual microseepage areas have remained unknown (Etiope and Klusman, 2010). We postulate that there is a possibility to detect a large portion of microseeping areas using the capability of spectral remote sensing to map induced mineralogical signatures. Such maps would then facilitate evaluating the contribution of every sedimentary basin or petroleum field in natural methane inventory.

From another perspective, the oxidation of pyrite in a sulfide-rich microseepage system may produce acid rock drainage and affect the ecosystem and the quality of drinking water in nearby areas (Swayze et al., 2000). Screening of microseepage areas by spectral remote sensing can help geologists gain a better understanding of the relative distribution of secondary minerals formed after sulfide oxidation and predict potential acid generation areas.

Unlike the natural seepage of crude oil into the marine environment, onshore seepage is not considered an important source of oil pollution. However, owing to the fact that several aspects of

natural seepage are shared by anthropogenic oil spillage, any progress towards remote sensing seepage characterization can be an asset to environmental protection and vice versa.

2.8. Conclusion

Whereas oil production from offshore basins is rapidly growing, in terms of yet-to-find reserves, it has been estimated that around half of the world's total conventional oil would still come from onshore basins, of which about half is expected to be new discoveries (Schenk, 2012; 2000). The macro-, and microseepage systems associated with these unexplored (and explored) onshore accumulations have been proven to contribute substantially towards natural methane emission and global warming. Consequently, any attempt to develop methods for seepage characterization has dual implications: in one way, it can be employed as a reliable indicator for oil and gas exploration; and in the other, it can be used to disaggregate the geological emission factors in environmental assessment.

Spectral remote sensing is offering a unique opportunity to detect the full range of onshore seepage indications; typically, the VNIR–SWIR wavelengths have been used to map the alteration footprints of microseepage systems and the SWIR–LWIR wavelengths to detect the manifestation of oil and gas macroseeps. Despite the encouraging results outlined here, we believe the potentials of this state-of-the-art technology for seepage inspection is not yet fully exploited. In the case of oil seepage, this approach not only can detect the oil-shows but also has a great capacity to quantify and characterize their HC content provided that the sensor has high spatial and spectral resolutions. The emerging capability of this technique in mapping trace gas anomalies is very promising; however, it should be expanded to sense other lightweight HCs. Correspondingly, alongside methane sensing, which is particularly valuable for environmental issues, further experiments should be devised for mixed methane and ethane (C_{2+}) detection, as the latter is a very useful indicator for oil and gas exploration. Overall, we need to develop additional case studies over petroliferous terrains with different oil and gas seepage properties and flow rates to evaluate the robustness of the techniques in determining the outline, content, and quantity of leakages. Such studies shall benefit from sensors available in the SWIR and LWIR wavelength ranges.

In the case of microseepage systems, the existing ambiguities in anomaly interpretation are believed to arise from several contributory factors including data restrictions, incomplete study cases, a simplistic methodology for data analysis, and above all, an immature conceptual model. In order to increase the efficiency of remote sensing approach, studies should be enriched by advanced spectral products such as the abundance and physicochemistry of minerals and supplemented by a novel mineralogical indicator; an objective that is mainly achievable by hyperspectral remote sensing. Moreover, further studies should be directed towards a diverse range of host-rocks and geologic settings by considering the basic guidelines provided in this article aiming to give insights into the full range of secondary changes and evaluate their detectability via spectral techniques. Ultimately, a new composite microseepage model capable of accounting for the variety of secondary changes should be devised and employed in the future investigations.

Acknowledgement

We are thankful to Edward Cloutis (University of Winnipeg) and Celio Pasquini (University of Campinas) for their helpful discussion on petroleum spectroscopy. Rosa E. Pabón is acknowledged for providing the petroleum spectrum used for spectral deconvolution. We also appreciate the USGS and NASA for facilitating our access to the datasets used in the demonstration case studies. The authors would like to thank FAPESP for the research grant No. 2015/06663-7. C. R. Souza Filho thanks CNPq for research grant No. 2008-7/303563.

References

Abrams, M.A., 2005. Significance of hydrocarbon seepage relative to petroleum generation and entrapment. Marine and Petroleum Geology, 22(4): 457-477.

Abrams, M.J., Conel, J.E., Lang, H.R. and Paley, H.N., 1985. The Joint NASA/Geosat Test Case Project: final report. American Association of Petroleum Geologists.

Agar, B., 1999. Spectral and mineralogical characterisation of alteration associated with hydrocarbon seepage using Geoscan AMSS MKII data over Palm Valley Thirteenth International Conference and Workshops on Applied Geologic Remote Sensing. ERIM international, Vancouver, British Columbia, Canada, pp. 110-117.

Al Shaieb, Z., Cairns, J. and Puckette, J., 1994. Hydrocarbon-induced diagenetic aureoles: Indicators of deeper, leaky reservoirs. Association of Petroleum Geochemical Explorationists Bulletin, 10: 24–48.

Almeida-Filho, R., Miranda, F.P., Galvão, L.S. and Freitas, C.C., 2002. Terrain characteristics of a tonal anomaly remotely detected in an area of hydrocarbon microseepage, Tucano Basin, north-eastern Brazil. International Journal of Remote Sensing, 23(18): 3893-3898.

Almeida-Filho, R., Miranda, F.P. and Yamakawa, T., 1999. Remote detection of a tonal anomaly in an area of hydrocarbon microseepage, Tucano basin, north-eastern Brazil. International Journal of Remote Sensing, 20(13): 2683-2688.

Aminzadeh, F., Berge, T., B., de Groot, P. and Valenti, G., 2001. Using Gas Chimneys as an Exploration Tool. World Oil, 222(5): 50-56.

Asadzadeh, S. and Souza Filho, C.R., 2016a. Investigating the capability of WorldView-3 superspectral data for direct hydrocarbon detection. Remote Sensing of Environment, 173: 162-173.

Asadzadeh, S. and Souza Filho, C.R., 2016b. A review on spectral processing methods for geological remote sensing. International Journal of Applied Earth Observation and Geoinformation, 47: 69-90.

Aske, N., Kallevik, H. and Sjöblom, J., 2001. Determination of Saturate, Aromatic, Resin, and Asphaltenic (SARA) Components in Crude Oils by Means of Infrared and Near-Infrared Spectroscopy. Energy & Fuels, 15(5): 1304-1312.

Bailey, T.C., 1996. Application of X-band radar to sense hydrocarbon seepage. Oil and Gas Journal, 94(50): 72-75.

Bailey, T.C. and Grubb, J.M., 2006. Drilling confirms radar-mapped atmospheric seepage anomalies. Oil and Gas Journal, 104(39): 32-37.

Bailey, T.C. and Skolnik, M.I., 1996. Science behind sensing hydrocarbon seepage using X-band radar. Oil and Gas Journal, 94(49): 85-88.

Beitler, B., Chan, M.A. and Parry, W.T., 2003. Bleaching of Jurassic Navajo sandstone on Colorado Plateau Laramide highs: Evidence of exhumed hydrocarbon supergiants? Geology, 31(12): 1041-1044.

Beitler, B., Parry, W.T. and Chan, M.A., 2005. Fingerprints of Fluid Flow: Chemical Diagenetic History of the Jurassic Navajo Sandstone, Southern Utah, U.S.A. Journal of Sedimentary Research, 75(4): 547-561.

Bell, J.H., Bowen, B.B. and Martini, B.A., 2010. Imaging spectroscopy of jarosite cement in the Jurassic Navajo Sandstone. Remote Sensing of Environment, 114(10): 2259-2270.

Boggs, S.J., 2009. Petrology of sedimentary rocks. Cambridge University Press, NY, USA, 600 pp.

Bowen, B.B., Martini, B.A., Chan, M.A. and Parry, W.T., 2007. Reflectance spectroscopic mapping of diagenetic heterogeneities and fluid-flow pathways in the Jurassic Navajo Sandstone. AAPG Bulletin, 91(2): 173-190.

Bradley, E.S., Leifer, I., Roberts, D.A., Dennison, P.E. and Washburn, L., 2011. Detection of marine methane emissions with AVIRIS band ratios. Geophysical Research Letters, 38(10): n/a-n/a.

Brown, A., 2000. Evaluation of Possible Gas Microseepage Mechanisms. AAPG Bulletin, 84(11): 1775-1789.

Calhoun, G.G., 1991. How 12 Geochemical Methods Fared in GERT Project in Permian. Oil and Gas Journal, 89(19): 62-68.

Chakraborty, S., Weindorf, D.C., Li, B., Ali Aldabaa, A.A., Ghosh, R.K., Paul, S. and Nasim Ali, M., 2015. Development of a hybrid proximal sensing method for rapid identification of petroleum contaminated soils. Science of The Total Environment, 514: 399-408.

Chakraborty, S., Weindorf, D.C., Morgan, C.L.S., Ge, Y., Galbraith, J.M., Li, B. and Kahlon, C.S., 2010. Rapid Identification of Oil-Contaminated Soils Using Visible Near-Infrared Diffuse Reflectance Spectroscopy. Journal of Environmental Quality, 39(4): 1378-1387.

Chan, M.A., Parry, W.T. and Bowman, J.R., 2000. Diagenetic Hematite and Manganese Oxides and Fault-Related Fluid Flow in Jurassic Sandstones, Southeastern Utah. AAPG Bulletin, 84(9): 1281-1310.

Chen, S., Zhao, Y., Zhao, L., Liu, Y. and Zhou, C., 2016. Hydrocarbon micro-seepage detection by altered minerals mapping from airborne hyper-spectral data in Xifeng Oilfield, China. Journal of Earth Science: 1-10.

Clark, R.N., Curchin, J.M., Hoefen, T.M. and Swayze, G.A., 2009. Reflectance spectroscopy of organic compounds: 1. Alkanes. Journal of Geophysical Research: Planets, 114(E3): 19.

Clark, R.N., Swayze, G.A., Leifer, I., Livo, K.E., Kokaly, R., Hoefen, T., Lundeen, S., Eastwood, M., Green, R.O., Pearson, N., Sarture, C., McCubbin, I., Roberts, D., Bradley, E., Steele, D., Ryan, T., Dominguez, R. and Team, a.t.A.b.V.I.I.S.A., 2010. A method for quantitative mapping of thick oil spills using imaging spectroscopy, U.S. Geological Survey Open-File Report Number 2010–1167.

Clarke, R.H. and Cleverly, R.W., 1991. Petroleum seepage and post-accumulation migration. Geological Society, London, Special Publications, 59(1): 265-271.

Cloutis, E.A., 1989. Spectral reflectance properties of hydrocarbons: Remote-sensing implications. Science, 245(4914): 165-168.

Cloutis, E.A., 1990. Identification, detection and characterization of individual tar sand phases using diffuse reflectance spectroscopy (0.35-2.6 µm). AOSTRA Journal of Research, 6: 17-27.

Cloutis, E.A., Gaffey, M.J. and Moslow, T.F., 1995. Characterization of minerals in oil sands by reflectance spectroscopy. 74(6): 874-879.

Coates, J., 2006. Interpretation of Infrared Spectra, A Practical Approach, Encyclopedia of Analytical Chemistry. John Wiley & Sons, Ltd, pp. 1-23.

Connolly, D., Aminzadeh, F., Brouwer, F. and Nielsen, S., 2013. Detection of Subsurface Hydrocarbon Seepage in Seismic Data: Implications for Charge, Seal, Overpressure, and Gas-hydrate Assessment. In: F. Aminzadeh, T.B. Berge and D.L. Connolly (Editors), Hydrocarbon Seepage: From source to surface. SEG & AAPG, Tulsa, USA, pp. 199-220.

Correa Pabón, R.E. and Souza Filho, C.R., 2016. Spectroscopic characterization of red latosols contaminated by petroleum-hydrocarbon and empirical model to estimate pollutant content and type. Remote Sensing of Environment, 175: 323-336.

Crowley, J.K., Williams, D.E., Hammarstrom, J.M., Piatak, N., Chou, I.-M. and Mars, J.C., 2003. Spectral reflectance properties (0.4–2.5 µm) of secondary Fe-oxide, Fe-hydroxide, and Fe-sulphate-hydrate minerals associated with sulphide-bearing mine wastes. Geochemistry: Exploration, Environment, Analysis, 3(3): 219-228.

Curto, J.B., Pires, A.C.B., Moreira, A.S. and Crósta, Á.P., 2011. Strategies for detecting hydrocarbon microseepages using airborne geophysics integrated with remote sensing data, 12th International Congress of the Brazilian Geophysical Society Brazilian Geophysical Society, Rio de Janeiro, Brazil, pp. 1-6.

Davidson, M.J., 2004. Evolution of scientific surface geochemical exploration. Oil and Gas Journal, 102(23): 35-38.

Donovan, T.J., 1974. Petroleum Microseepage at Cement, Oklahoma; Evidence and Mechanism. AAPG Bulletin, 58(3): 429-446.

Donovan, T.J., Friedman, I. and Gleason, J.D., 1974. Recognition of Petroleum-bearing Traps by Unusual Isotopic Compositions of Carbonate-cemented Surface Rocks. Geology, 2(7): 351-354.

Donovan, T.J., Noble, R.L., friedman, I. and Gleason, J.D., 1975. A possible petroleum related geochemical anomaly in surface rocks, USGS, Open-File Report No 75-47

Donovan, T.J., Roberts, A.A. and Dalziel, M.C., 1979-1982. Epigenetic zoning in surface and near-surface rocks resulting from seepage-induced redox gradients, Velma oil field, Oklahoma: A synopsis. The Shale Shaker Digest XX, XXX-XXXII: 175-181.

Duchscherer, J.W., 1980. Geochemical methods of prospecting for hydrocarbons. Oil and Gas Journal, DEC: 194-208.

Duchscherer, J.W., 1982. Geochemical exploration for hydrocarbons, no new tricks - but an old dog. Oil and Gas Journal, July: 163-176.

Eichhubl, P., Taylor, W.L., Pollard, D.D. and Aydin, A., 2004. Paleo-fluid flow and deformation in the Aztec Sandstone at the Valley of Fire, Nevada—Evidence for the coupling of hydrogeologic, diagenetic, and tectonic processes. Geological Society of America Bulletin, 116(9-10): 1120-1136.

Ellis, J.M., Davis, H.H. and Zamudio, J.A., 2001. Exploring for onshore oil seepas with hyperspectral imaging. Oil and Gas Journal, 99(37): 49-58.

Ellwood, B.B. and Burkart, B., 1996. Test of hydrocarbon-induced magnetic patterns in soils: the sanitary landfill as laboratory. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expresseion. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 91-98.

Elwood Madden, M.E., Bodnar, R.J. and Rimstidt, J.D., 2004. Jarosite as an indicator of water-limited chemical weathering on Mars. Nature, 431(7010): 821-823.

Etiope, G., 2009. GLOGOS, A new global onshore gas-oil seeps dataset, Search and discovery, Article #70071. AAPG online journal.

Etiope, G., 2015. Natural Gas Seepage: The Earth's Hydrocarbon Degassing. Springer International Publishing, Switzerland.

Etiope, G. and Ciccioli, P., 2009. Earth's Degassing: A Missing Ethane and Propane Source. Science, 323(5913): 478-478.

Etiope, G. and Klusman, R.W., 2010. Microseepage in drylands: Flux and implications in the global atmospheric source/sink budget of methane. Global and Planetary Change, 72(4): 265-274.

Etiope, G., Lassey, K.R., Klusman, R.W. and Boschi, E., 2008. Reappraisal of the fossil methane budget and related emission from geologic sources. Geophysical Research Letters, 35(9): L09307.

Etiope, G. and Martinelli, G., 2002. Migration of carrier and trace gases in the geosphere: an overview. Physics of the Earth and Planetary Interiors, 129(3–4): 185-204.

Eventov, L., 2000. The nature and interpretation of geophysical and geochemical anomalies over oil and gas fields. The Leading Edge, 19(5): 488-490.

Everett, J.R., Staskowski, R.J. and Jengo, C., 2002. Remote sensing and GIS enable future exploration success. World Oil, 223(11): 59-63.

Feder, A.M., 1985. Contemporary Remote Sensing for Hydrocarbon Exploration and Development-with Selected Case Histories. Oil and Gas Journal.

Fingas, M. and Brown, C.E., 2014. Oil Spill Remote Sensing, Handbook of Oil Spill Science and Technology. John Wiley & Sons, Inc, pp. 311-356.

Foote, R.S., 1992. Use of Magnetic Field Aids Oil Search. Oil and Gas Journal, 90(18): 137-142.

Foote, R.S., 1996. Relationship of near-surface magnetic anomalies to oil- and gas producing areas. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expresseion. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 111-126.

Foote, R.S., 2013. Value of MSRI measurement of Selma exploration drill cuttings. Oil and Gas Journal, 111(10): 66-73.

Frankenberg, C., Thorpe, A.K., Thompson, D.R., Hulley, G., Kort, E.A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C., Conley, S., Bue, B.D., Aubrey, A.D., Hook, S. and Green, R.O., 2016. Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region. Proceedings of the National Academy of Sciences, 113(35): 9734-9739.

Frassy, F., Maianti, P., Marchesi, A., Nodari, F.R., Dalla Via, G., De Paulis, R., Biffi, P.G. and Gianinetto, M., 2015. Satellite remote sensing for hydrocarbon exploration in new venture areas, 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 2884-2887.

Freeman, H., 2003. Evaluation of the use of hyperspectral imagery for identification of microseeps near Santa Barbara, California, West Virginia University, 22 pp.

Fu, B., Zheng, G., Ninomiya, Y., Wang, C. and Sun, G., 2007. Mapping hydrocarbon-induced mineralogical alteration in the northern Tian Shan using ASTER multispectral data. Terra Nova, 19(4): 225-231.

Gaffey, S.J., 1987. Spectral reflectance of carbonate minerals in the visible and near infrared (0.35–2.55 um): Anhydrous carbonate minerals. Journal of Geophysical Research: Solid Earth, 92(B2): 1429-1440.

Garden, I.R., Guscott, S.C., Burley, S.D., Foxford, K.A., Walsh, J.J. and Marshall, J., 2001. An exhumed palaeo-hydrocarbon migration fairway in a faulted carrier system, Entrada Sandstone of SE Utah, USA. Geofluids, 1(3): 195-213.

Gorenc, M.A. and Chan, M.A., 2015. Hydrocarbon-induced diagenetic alteration of the Permian White Rim Sandstone, Elaterite Basin, southeast Utah. AAPG Bulletin, 99(5): 807-829.

Gutierrez, C.I., Tan, S.S. and Clahan, K.B., 2008. Geologic map of the east half Santa Barbara 30' × 60' Quadrangle, California. California Geological Survey.

Holysh, S. and Toth, J., 1996. Flow of formation waters: Likely cause for poor definition of soil gas anomalies over oil fields in east-central Alberta. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expression. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 255-277.

Hörig, B., Kühn, F., Oschütz, F. and Lehmann, F., 2001. HyMap hyperspectral remote sensing to detect hydrocarbons. International Journal of Remote Sensing, 22(8): 1413-1422.

Horvitz, L., 1980. Near-surface evidence of hydrocarbon movement from depth, SG 10: Problems of Petroleum Migration. AAPG, pp. 241-269.

Horvitz, L., 1985. Geochemical Exploration for Petroleum. Science, 229(4716): 821-827.

Hulley, G.C., Duren, R.M., Hopkins, F.M., Hook, S.J., Vance, N., Guillevic, P., Johnson, W.R., Eng, B.T., Mihaly, J.M., Jovanovic, V.M., Chazanoff, S.L., Staniszewski, Z.K., Kuai, L., Worden, J., Frankenberg, C., Rivera, G., Aubrey, A.D., Miller, C.E., Malakar, N.K., Sánchez Tomás, J.M. and Holmes, K.T., 2016. High spatial resolution imaging of methane and other trace gases with the airborne Hyperspectral Thermal Emission Spectrometer (HyTES). Atmos. Meas. Tech., 9(5): 2393-2408.

Hunt, G.R., 1977. Spectral signatures of particulate minerals in the visible and near infrared. Geophysics, 42(3): 501-513.

Hunt, J.M., 1996. Petroleum geochemistry and geology. W. H. Freeman, New York.

Jengo, C.M. and Vincent, R.K., 1999. Hyperspectral imaging of ancient hydrocarbon seeps, Wind River basin, Wyoming, Thirteenth International Conference and Workshops on Applied Geologic Remote Sensing. ERIM international, Vancouver, British Columbia, Canada, pp. 118-125.

Jones, V.T. and Drozd, R.J., 1983. Predictions of oil or gas potential by near-surface geochemistry. AAPG Bulletin, 67(6): 932-952.

Khan, S.D. and Jacobson, S., 2008. Remote sensing and geochemistry for detecting hydrocarbon microseepages. Geological Society of America Bulletin, 120(1-2): 96-105.

Kirkland, D.W., Denison, R.E. and Rooney, M.A., 1995. Diagenetic alteration of Permian strata at oil fields of south central Oklahoma, USA. Marine and Petroleum Geology, 12(6): 629-644.

Klusman, R.M., 2002. Interpretation and display of surface geochemical data. In: D. Schumacher and L.A. LeSchack (Editors), Surface exploration case histories: Application of geochemistry, magnetics, and remote sensing. AAPG and SEG, Tulsa, OK, U.S.A., pp. 1-24.

Klusman, R.M. and Saeed, M.A., 1996. Comparison of light hydrocarbon microseepage mechanisms. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expresseion. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 157-168.

Kokaly, R.F., Couvillion, B.R., Holloway, J.M., Roberts, D.A., Ustin, S.L., Peterson, S.H., Khanna, S. and Piazza, S.C., 2013. Spectroscopic remote sensing of the distribution and persistence of oil from the Deepwater Horizon spill in Barataria Bay marshes. Remote Sensing of Environment, 129(0): 210-230.

Kühn, F., Oppermann, K. and Hörig, B., 2004. Hydrocarbon Index – an algorithm for hyperspectral detection of hydrocarbons. International Journal of Remote Sensing, 25(12): 2467-2473.

Lammoglia, T. and Souza Filho, C.R., 2011. Spectroscopic characterization of oils yielded from Brazilian offshore basins: Potential applications of remote sensing. Remote Sensing of Environment, 115(10): 2525-2535.

Lammoglia, T. and Souza Filho, C.R., 2012. Mapping and characterization of the API gravity of offshore hydrocarbon seepages using multispectral ASTER data. Remote Sensing of Environment, 123(0): 381-389.

Lammoglia, T., Souza Filho, C.R. and Filho, R.A., 2008. Characterization of hydrocarbon microseepages in the Tucano basin, (Brazil) through hyperspectral classification and neural network Analysis of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, pp. 1195-1200.

Land, J.P., 1996. Nonseismic methods can provide many views of a drillsite. Oil and Gas Journal, 94(9): 69-73.

Leifer, I., Lehr, W.J., Simecek-Beatty, D., Bradley, E., Clark, R., Dennison, P., Hu, Y., Matheson, S., Jones, C.E., Holt, B., Reif, M., Roberts, D.A., Svejkovsky, J., Swayze, G. and Wozencraft, J., 2012a. State of the art satellite and airborne marine oil spill remote sensing: Application to the BP Deepwater Horizon oil spill. Remote Sensing of Environment, 124(0): 185-209.

Leifer, I., Tratt, D.M., Realmuto, V.J., Gerilowski, K. and Burrows, J.P., 2012b. Remote sensing atmospheric trace gases with infrared imaging spectroscopy. Eos, Transactions American Geophysical Union, 93(50): 525-525.

LeSchack, L.A., 1997. Results of magnetic HGI and radiometric surveys in W. Canada. Oil and Gas Journal, 95(20): 84-89.

LeSchack, L.A. and Van Alstine, D.R., 2002. High-resolution ground-magnetic (HRGM) and radiometric surveys for hydrocarbon exploration: Six case histories in Western Canada. In: D. Schumacher and L.A. LeSchack (Editors), Surface exploration case histories: Application of geochemistry, magnetics, and remote sensing. AAPG and SEG, Tulsa, OK, U.S.A., pp. 67-156.

Lilburn, R.A. and Zuhair, A.-S., 1984. Geochemistry and isotopic composition of Hydrocarbon-Induced Diagenetic Aureole (HIDA), Cement Field Oklahoma. The Shale Shaker Digest XI, XXXIII-XXXV: 46-72.

Link, W.K., 1952. Significance of oil and gas seeps in world oil exploration. AAPG Bulletin, 36(8): 1505-1540.

Lyder, D., Feng, J., Rivard, B., Gallie, A. and Cloutis, E., 2010. Remote bitumen content estimation of Athabasca oil sand from hyperspectral infrared reflectance spectra using Gaussian singlets and derivative of Gaussian wavelets. Fuel, 89(3): 760-767.

Macgregor, D.S., 1993. Relationships between seepage, tectonics and subsurface petroleum reserves. Marine and Petroleum Geology, 10(6): 606-619.

Magoon, L.B. and Beaumont, E.A., 1999. Petroleum Systems. In: E.A. Beaumont and N.H. Foster (Editors), Treatise of Petroleum Geology / Handbook of Petroleum Geology: Exploring for Oil and Gas Traps. AAPG, pp. 3-1 - 3-34.

Malhotra, R.V., Birnie, R.W. and Johnson, B.G., 1989. Detection of surficial changes associated with hydrocarbon seepage Sheep Mountain anticline, Bighorn basin, Wyoming, Seventh Thematic Conference on Remote Sensing for Exploration Geology. ERIM international, Calgarly, Alberta, Canada, pp. 1097-1110.

Marrs, R.W. and Paylor., E.D., 1987. Investigation of a surface spectral anomaly at Table Rock gas field, Wyoming. GEOPHYSICS, 52(7): 841-857.

Matthews, M.D., 1986. Effects of hydrocarbon leakage on earth surface materials. In: M.J. Davidson (Editor), Unconventional methods in prospecting for petroleum and natural gas IV. outhern Methodist University Press, Institute for the Study of Earth and Man, Dallas, pp. 27-44.

McCoy, R.M., Blake, J.G. and Andrews, K.L., 2001. Detecting hydrocarbon microseepage using hydrocarbon absorption bands of reflectance spectra of surface soils. Oil and Gas Journal, 99(22): 40-45.

Nielsen, G.B., Chan, M.A. and Petersen, E.U., 2009. Diagenetic coloration facies and alteration history of the jurassic Navajo sandstone, Zion national park and vicinity, Southwestern Utah, CD-ROM.

Nielson, G.B., Chan, M.A. and Bowen, B., 2014. Iron-rich horizons in the Jurassic Navajo Sandstone, southwestern Utah: progressive cementation and permeability inversion. In: J.S. MacLean, R.F. Biek and J.E. Huntoon (Editors), Geology of Utah's Far South. Utah Geological Association Publication, pp. 215-236.

Okparanma, R.N., Coulon, F. and Mouazen, A.M., 2014. Analysis of petroleum-contaminated soils by diffuse reflectance spectroscopy and sequential ultrasonic solvent extraction–gas chromatography. Environmental Pollution, 184: 298-305.

Parry, W.T., Chan, M.A. and Beitler, B., 2004. Chemical bleaching indicates episodes of fluid flow in deformation bands in sandstone. AAPG Bulletin, 88(2): 175-191.

Parry, W.T., Chan, M.A. and Nash, B.P., 2009. Diagenetic characteristics of the Jurassic Navajo Sandstone in the Covenant oil field, central Utah thrust belt. AAPG Bulletin, 93(8): 1039-1061.

Perry, A.F., 2006. Unconventional exploration technologies: Take another look. World Oil, 227(9): 45-52.

Perry, S.L. and Kruse, F.A., 2010. Evidence of hydrocarbon seepage using multispectral satellite imagery, Kurdistan, Iraq, AAPG International Convention and Exhibition, AAPG Search and Discovery Article #90108, Calgary, Alberta, Canada, pp. 7.

Petrovic, A., Khan, S.D. and Chafetz, H.S., 2008. Remote detection and geochemical studies for finding hydrocarbon-induced alterations in Lisbon Valley, Utah. Marine and Petroleum Geology, 25(8): 696-705.

Petrovic, A., Khan, S.D. and Thurmond, A.K., 2012. Integrated hyperspectral remote sensing, geochemical and isotopic studies for understanding hydrocarbon-induced rock alterations. Marine and Petroleum Geology, 35(1): 292-308.

Pirajno, F., 2009. Hydrothermal Processes and Mineral Systems. Springer Science.

Potter-McIntyre, S., Allen, J., Lee, S.Y., Han, W.S., Chan, M. and McPherson, B., 2013. Iron precipitation in a natural CO2 reservoir: Jurassic Navajo Sandstone in the northern San Rafael Swell, UT, USA. Geofluids, 13(1): 82-92.

Potter II, R.W., Harrington, P.A., Silliman, A.H. and Viellena, J.H., 1996. Significance of Geochemical Anomalies in Hydrocarbon Exploration: One Company's Experience. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expression. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 431-440.

Prelat, A., Gunaratne, S., Huebner, L., Freeman, C., Cook, A. and Soriano, C., 2013. Airborne hyperspectral detection of natural offshore and onshore hydrocarbon seeps. In: F. Aminzadeh, T.B. Berge and D.L. Connolly (Editors), Hydrocarbon Seepage: From Source to Surface. American Association of Petroleum Geologists, Tulsa, OK, U.S.A., pp. 171-182.

Price, L.C., 1986. A critical review and proposed working model of surface geochemical exploration. In: M.J. Davidson (Editor), Unconventional methods in exploration for petroleum and natural gas. Southern Methodist University Press, Dallas, pp. 245–304.

Price, L.C., 1996. Research-derived insights into surface geochemical hydrocarbon exploration. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expression. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 285-307.

Prokopovich, N.P., Cole, R.C. and Nishi, C.K., 1971. Alteration of sediments by natural gases in western Merced County, California. AAPG Bulletin, 55(6): 826-832.

Qin, Q., Zhang, Z., Chen, L., Wang, N. and Zhang, C., 2016. Oil and gas reservoir exploration based on hyperspectral remote sensing and super-low-frequency electromagnetic detection. Journal of Applied Remote Sensing, 10(1): 016017-016017.

Rainoldi, A.L., Franchini, M., Beaufort, D., Patrier, P., Giusiano, A., Impiccini, A. and Pons, J., 2014. Large-Scale Bleaching of Red Beds Related To Upward Migration of Hydrocarbons: Los Chihuidos High, Neuquén Basin, Argentina. Journal of Sedimentary Research, 84(5): 373-393.

Rice, G.K., Belt Jr, J.Q. and Berg, G.E., 2002. Soil-gas Hydrocarbon Pattern Changes During a West Texas Waterflood. In: D. Schumacher and L.A. LeSchack (Editors), Surface exploration case histories: Application of geochemistry, magnetics, and remote sensing. AAPG and SEG, Tulsa, OK, U.S.A., pp. 157-174.

Rice, G.K., King, G. and Henson, J., 2016. Integration of geology, seismic, and geochemical data — Theory and practice in Cheeseburger Field, Stonewall County, Texas, USA. Interpretation, 4(2): T215-T225.

Richers, D.M., Jones, V.T., Matthews, M.D., Maciolek, J., Pirkle, R.J. and Sidle, W.C., 1986. The 1983 Landsat soil-gas geochemical survey of Patrick Draw area, Sweetwater County, Wyoming. AAPG Bulletin, 70(7): 869-887.

Richers, D.M., Reed, R.J., Horstman, K.C., Michels, G.D., Baker, R.N., Lundell, L. and Marrs, R.W., 1982. Landsat and soil-gas geochemical study of Patrick Draw oil field, Sweetwater County, Wyoming. AAPG Bulletin, 66(7): 903-922.

Roberts, D.A., Bradley, E.S., Cheung, R., Leifer, I., Dennison, P.E. and Margolis, J.S., 2010. Mapping methane emissions from a marine geological seep source using imaging spectrometry. Remote Sensing of Environment, 114(3): 592-606.

Salati, S., van Ruitenbeek, F., van der Meer, F. and Naimi, B., 2014a. Detection of alteration induced by onshore gas seeps from ASTER and WorldView-2 data. Remote Sensing, 6(4): 3188-3209.

Salati, S., van Ruitenbeek, F.J.A., de Smeth, J.B. and van der Meer, F.D., 2014b. Spectral and geochemical characterization of onshore hydrocarbon seep-induced alteration in the Dezful embayment, southwest Iran. AAPG Bulletin, 98(9): 1837-1857.

Saunders, D.F., Burson, K.R., Branch, J.F. and Thompson, C.K., 1993a. New method of aerial and surface radiometric prospecting for oil, gas. Oil and Gas Journal, 91(38): 104-109.

Saunders, D.F., Burson, K.R., Branch, J.F. and Thompson, C.K., 1993b. Relation of thorium-normalized surface and aerial radiometric data to subsurface petroleum accumulations. Geophysics, 58(10): 1417-1427.

Saunders, D.F., Burson, K.R. and Thompson, C.K., 1991. Observed relation of soil magnetic susceptibility and soil gas hydrocarbon analyses to subsurface petroleum accumulations. AAPG Bulletin, 75(3): 389-408.

Saunders, D.F., Burson, K.R. and Thompson, C.K., 1999. Model for hydrocarbon microseepage and related near-surface alterations. AAPG Bulletin, 83(1): 170-185.

Saunders, D.F. and Terry, S.A., 1985. Onshore exploration using new geochemistry and geomorphology. Oil and Gas Journal, September, 16(38): 126-130.

Scafutto, R.D.P.M., 2015. Sensoriamento remoto hiperespectral aplicado à identificação remota de solos e rochas impregnados com hidrocarbonetos, University of Campinas, Campinas, Brazil, 258 pp.

Scafutto, R.D.P.M. and Souza Filho, C.R., 2016. Quantitative characterization of crude oils and fuels in mineral substrates using reflectance spectroscopy: Implications for remote sensing. International Journal of Applied Earth Observation and Geoinformation, 50: 221-242.

Scafutto, R.D.P.M., Souza Filho, C.R. and Rivard, B., 2016. Characterization of mineral substrates impregnated with crude oils using proximal infrared hyperspectral imaging. Remote Sensing of Environment, 179: 116-130.

Schenk, C.J., 2012. An estimate of undiscovered conventional oil and gas resources of the world, 2012.

Scholte, K.H., 2005. Hyperspectral remote sensing and mud volcanism in azerbaijan (PhD thesis), Delft University of Technology, Netherlands, 147 pp.

Scholte, K.H., Hommels, A., Munoz-Sabater, J., Hanssen, R.F., Van der Meer, F.D., Kroonenberg, S.B., Aliyeva, E., Huseynov, D. and Guliev, I., 2003. Preliminary ASTER and InSAR imagery combination for mud volcano dynamics, Azerbaijan, Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International, pp. 1573-1575.

Schumacher, D., 1996. Hydrocarbon-induced alteration of soils and sediments. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expression. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 71-89.

Schumacher, D., 1999. Surface Geochemical Exploration for Petroleum. In: E.A. Beaumont and N.H. Foster (Editors), Treatise of Petroleum Geology / Handbook of Petroleum Geology: Exploring for Oil and Gas Traps. AAPG, pp. 18-1 - 18-27.

Schumacher, D., 2010. Integrating hydrocarbon microseepage data with seismic data doubles exploration success, 34th annual conference and exhibition. Proceedings, Indonesian petroleum association, Indonesia, pp. 11.

Schumacher, D.D., 2000. Surface geochemical exploration for oil and gas: New life for an old technology. The Leading Edge, 19(3): 258-261.

Schwartz, G., Ben-Dor, E. and Eshel, G., 2013. Quantitative Assessment of Hydrocarbon Contamination in Soil Using Reflectance Spectroscopy: A “Multipath” Approach. Applied Spectroscopy, 67(11): 1323-1331.

Schwietzke, S., Sherwood, O.A., Bruhwiler, L.M.P., Miller, J.B., Etiope, G., Dlugokencky, E.J., Michel, S.E., Arling, V.A., Vaughn, B.H., White, J.W.C. and Tans, P.P., 2016. Upward revision of global fossil fuel methane emissions based on isotope database. Nature, 538(7623): 88-91.

Segal, D.B. and Merin, I.S., 1989. Successful use of Landsat Thematic Mapper data for mapping hydrocarbon microseepage-induced mineralogic alteration, Lisbon Valley, Utah. Photogrammetric Engineering and Remote Sensing, 55(8): 1137-1145.

Segal, D.B., Ruth, M.D. and Merin, I.S., 1986. Remote detection of anomalous mineralogy associated with hydrocarbon production, Lisbon Valley, Utah. The Moutain Geologist, 23(2): 51-62.

Shaw, R.C. and Kratochvil, B., 1990. Near-infrared diffuse reflectance analysis of Athabasca oil sand. Analytical Chemistry, 62(2): 167-174.

Shi, P., Fu, B., Ninomiya, Y., Sun, J. and Li, Y., 2012. Multispectral remote sensing mapping for hydrocarbon seepage-induced lithologic anomalies in the Kuqa foreland basin, south Tian Shan. Journal of Asian Earth Sciences, 46(0): 70-77.

Short, N.M., 1977. Exploration for fossil and nuclear fuels from orbital altitudes. In: W.L. Smith (Editor), Remote-sensing applications for mineral exploration. Dowden, Hutchinson, and Ross, Pennsylvania.

Simpson, C.J., 1978. LANDSAT: developing techniques and applications in mineral and petroleum exploration BMR Journal of Australian Geology & Geophysics, 3: 181-191.

Simpson, C.J., Wilford, J.R., Macias, L.F. and Korsch, R.J., 1991. Remote sensing of petroleum gas seepage alteration Palm Valley Gas Field, Amadeus Basin, central Australia., Proceedings of the Eighth Thematic Conference on Geologic Remote Sensing. ERIM, Denver, Colorado, USA, pp. 1463-1472.

Smailbegovic, A., Gray, K., Johnson, K., Murphy, V. and Holbrook, R., 2009. Spectroscopic and hyperspectral evaluation of possible hydrocarbon occurrences in estuarine sediments, North Charleston, South Carolina, USA, 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp. 1-5.

Souza Filho, C.R., Augusto, V., Oliveira, W.J. and Lammoglia, T., 2008. Detecção de exsudações de hidrocarbonetos por geobotânica e sensoriamento remoto multi-temporal: estudo de caso no Remanso do Fogo (MG). Revista Brasileira de Geociências, 38(2): 228-243.

Stallard, B.R., Garcia, M.J. and Kaushik, S., 1996. Near-IR Reflectance Spectroscopy for the Determination of Motor Oil Contamination in Sandy Loam. Applied Spectroscopy, 50(3): 334-338.

Staskowski, R.J., Jengo, C. and Everett, J.R., 2004. Utility of ASTER for detecting hydrocarbon [Abstract], American Association of Petroleum Geologists Annual Convention, Dallas, Texas, pp. 6.

Sunshine, J.M., Pieters, C.M. and Pratt, S.F., 1990. Deconvolution of mineral absorption bands: an improved approach. Journal of Geophysical Research, 93: 6955-6966.

Swayze, G.A., Smith, K.S., Clark, R.N., Sutley, S.J., Pearson, R.M., Vance, J.S., Hageman, P.L., Briggs, P.H., Meier, A.L., Singleton, M.J. and Roth, S., 2000. Using Imaging Spectroscopy To Map Acidic Mine Waste. Environmental Science & Technology, 34(1): 47-54.

Tangestani, M.H. and Validabadi, K., 2014. Mineralogy and geochemistry of alteration induced by hydrocarbon seepage in an evaporite formation; a case study from the Zagros Fold Belt, SW Iran. Applied Geochemistry, 41: 189-195.

Tedesco, S.A., 1995. Surface geochemistry in petroleum exploration. Chapman and Hall, Inc, New York.

Tedesco, S.A., 1999. Anomaly shifts indicate rapid surface seep rates. Oil and Gas Journal, 97(13): 69-72.

Thompson, C.K., Saunders, D.F. and Burson, K.R., 1994. Model advanced for hydrocarbon microseepage, related alterations. Oil and Gas Journal, 92(46): 95-99.

Thorpe, A.K., Frankenberg, C., Aubrey, A.D., Roberts, D.A., Nottrott, A.A., Rahn, T.A., Sauer, J.A., Dubey, M.K., Costigan, K.R., Arata, C., Steffke, A.M., Hills, S., Haselwimmer, C., Charlesworth, D., Funk, C.C., Green, R.O., Lundeen, S.R., Boardman, J.W., Eastwood, M.L., Sarture, C.M., Nolte, S.H., McCubbin, I.B., Thompson, D.R. and McFadden, J.P., 2016. Mapping methane concentrations from a controlled release experiment using the next generation airborne visible/infrared imaging spectrometer (AVIRIS-NG). Remote Sensing of Environment, 179: 104-115.

Thorpe, A.K., Frankenberg, C. and Roberts, D.A., 2014. Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: application to AVIRIS. Atmos. Meas. Tech., 7(2): 491-506.

Thorpe, A.K., Roberts, D.A., Bradley, E.S., Funk, C.C., Dennison, P.E. and Leifer, I., 2013. High resolution mapping of methane emissions from marine and terrestrial sources using a Cluster-Tuned Matched Filter technique and imaging spectrometry. Remote Sensing of Environment, 134: 305-318.

Thrasher, J., Fleet, A.J., Hay, S.J., Hovland, M. and Düppenbecker, S., 1996. Understanding Geology as the key to using seepage in exploration: The spectrum of seepage styles. In: D. Schumacher and M.A. Abrams (Editors), Hydrocarbon migration and its near-surface expression. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 223-241.

Tratt, D.M., Buckland, K.N., Hall, J.L., Johnson, P.D., Keim, E.R., Leifer, I., Westberg, K. and Young, S.J., 2014. Airborne visualization and quantification of discrete methane sources in the environment. Remote Sensing of Environment, 154: 74-88.

USGS, 2000. U.S. Geological Survey World Petroleum Assessment 2000: Description and Results. 60.

van Der Meer, F., van Dijk, P., van Der Werff, H. and Yang, H., 2002. Remote sensing and petroleum seepage: a review and case study. Terra Nova, 14(1): 1-17.

van der Werff, H., 2006. Knowledge based remote sensing of complex objects recognition of spectral and spatial patterns resulting from natural hydrocarbon, Enschede, The Netherlands, 138 pp.

van der Werff, H.M.A., Bakker, W.H., van der Meer, F.D. and Siderius, W., 2006. Combining spectral signals and spatial patterns using multiple Hough transforms: An application for detection of natural gas seepages. Computers & Geosciences, 32(9): 1334-1343.

Wang, Y. and Ding, X., 2000. Hydrocarbon Alteration Characteristics of Soils and Mechanism for Detection by Remote Sensing in East Sichuan Area, China. Natural Resources Research, 9(4): 295-305.

Warren, R.K., 2012. Near-surface resistivity for hydrocarbon detection. The Leading Edge, 31(10): 1182-1191.

Watanabe, H., Hayashi, K., Saeki, T., Maksyutov, S., Nasuno, I., Shimono, Y., Hirose, Y., Takaichi, K., Kanekon, S., Ajiro, M., Matsumoto, Y. and Yokota, T., 2015. Global mapping of greenhouse gases retrieved from GOSAT Level 2 products by using a kriging method. International Journal of Remote Sensing, 36(6): 1509-1528.

Weibel, R., 1998. Diagenesis in oxidising and locally reducing conditions — an example from the Triassic Skagerrak Formation, Denmark. Sedimentary Geology, 121(3–4): 259-276.

Wigley, M., Kampman, N., Dubacq, B. and Bickle, M., 2012. Fluid-mineral reactions and trace metal mobilization in an exhumed natural CO2 reservoir, Green River, Utah. Geology, 40(6): 555-558.

Winkelmann, K.H., 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, Brandenburg University of Technology, Brandenburg, Germany, 236 pp.

Wolleben, J.A. and Greenlee, D.W., 2002. Successful application of micromagnetic data to focus hydrocarbon exploration. In: D. Schumacher and L.A. LeSchack (Editors), Surface exploration case histories: Application of geochemistry, magnetics, and remote sensing. AAPG and SEG, Tulsa, OK, U.S.A., pp. 175-191.

Worden, R.H., Smalley, P.C. and Fallick, A.E., 1997. Sulfur cycle in buried evaporites. Geology, 25(7): 643-646.

Xuejing, X., 1992. Local and regional surface geochemical exploration for oil and gas. Journal of Geochemical Exploration, 43(1): 25-42.

Yang, H., Meer, F.V.D., Zhang, J. and Kroonenberg, S.B., 2000. Direct detection of onshore hydrocarbon microseepages by remote sensing techniques. Remote Sensing Reviews, 18(1): 1-18.

Yergin, D., 1992. The Prize: The Epic Quest for Oil, Money, and Power. Simon & Schuster.

Zhang, G., Zou, L., Shen, X., Lu, S., Li, C. and Chen, H., 2009. Remote sensing detection of heavy oil through spectral enhancement techniques in the western slope zone of Songliao Basin, China. AAPG Bulletin, 93(1): 31-49.

Zheng, G., Fu, B., Takahashi, Y., Kuno, A., Matsuo, M. and Zhang, J., 2010. Chemical speciation of redox sensitive elements during hydrocarbon leaching in the Junggar Basin, Northwest China. Journal of Asian Earth Sciences, 39(6): 713-723.

Chapter 3

A review on spectral processing methods for geological remote sensing[4]

In this chapter, many of the fundamental and advanced spectral processing methods available to geologic remote sensing are reviewed. A novel categorization scheme is proposed that groups the techniques into knowledge-based and data-driven approaches, according to the type and availability of reference data. The two categories are compared and their characteristics and geologic outcomes are contrasted. Using an oil-sand sample scanned through the sisuCHEMA hyperspectral imaging system as a case study, the effectiveness of selected processing techniques from each category is demonstrated. The techniques used to bridge between the spectral data and other geoscience products are then discussed. Subsequently, the hybridization of the two approaches is shown to yield some of the most robust processing techniques available to multi- and hyperspectral remote sensing. Ultimately, current and future challenges that spectral analysis are expected to overcome and some potential trends are highlighted.

3.1. Introduction

Spectroscopy is the measurement of light as a function of wavelength reflected or emitted from a material (Clark, 1999; Hapke, 1993). The resultant spectrum conveys information about the state of the target that in geology is usually, but not necessarily composed of rocks and minerals. Pioneering work of John Hunt and Ronald Lyon in the early 70s paved the way for the interpretation of such spectra using quantum mechanics concepts (e.g. (Hunt and Salisbury, 1971; Lyon and Burns, 1963)). Their work established a link between observed variation in reflectance/emittance spectra with chemical and physical properties of minerals and demonstrated their potential use in remote sensing (Hunt, 1977, 1979). Minerals, rocks, and other terrestrial compounds like hydrocarbons exhibit diagnostic absorption features in either the visible-near infrared (VNIR) (0.4-1.0 µm), shortwave infrared (SWIR) (1.0-2.5 µm), mid infrared (MIR) (3-5 µm), and/or longwave infrared (LWIR) (8-14 µm) wavelength ranges due to electronic and vibrational processes, as well as overtones and combinations of the fundamental (Clark, 1999; Gaffey et al., 1993; Hapke, 1993; Hook et al., 1999; Hunt and Salisbury, 1971, 1974).

Historically, remotely sensed multispectral imaging (MSI) has been used to produce colorful photographs for visual interpretation of lithologic units and geologic structures (Goetz and Rowan, 1981; Gregory and Moore, 1975). Meanwhile, its multispectral content has been processed by simple techniques, like band arithmetic, to discriminate broad alteration patterns (Goetz and Rowan, 1981; Rowan et al., 1974; Sabins, 1999). Early experiments with airborne imaging spectrometer (AIS) prototypes revealed its potential for remote mineral detection, which subsequently led to the development of NASA's Airborne Visible-InfraRed Imaging Spectrometer (AVIRIS) hyperspectral imaging (HSI) sensor (Goetz et al., 1985; Vane and Goetz, 1991). HSI has matured to such extent that advanced systems of this kind are currently orbiting Earth and Mars (e.g. Hyperion and OMEGA) (Bell, 2008; Pearlman et al., 2003). This technology has also evolved as a tool for field spectroscopy (Goetz, 2009; Thompson et al., 1999), drill core and chips logging (Mason and Huntington, 2012; Roache et

[4] A version of this chapter is published in the International Journal of Applied Earth Observation and Geoinformation.

al., 2011; Tappert et al., 2011), wall-rock imaging (Kruse et al., 2012; Kurz et al., 2012; Murphy and Monteiro, 2013; Ragona et al., 2006), and sensor-based mineral sorting (Goetz et al., 2009). Overall, proximal and distal sensing technologies in the VNIR-SWIR have been matured and readily available (Goetz et al., 1985), whereas the LWIR hyperspectral data are only now becoming routinely available (Hook et al., 2013; Mason and Huntington, 2012; Vaughan et al., 2003).

HSI with hundreds of contiguous spectral bands has resulted in plethora of near laboratory-quality spectra for every pixel (Clark and Swayze, 1996; Goetz, 2009; Goetz et al., 1985), thus creating its own breed of spectral analysis methods (e.g. (Adams et al., 1986; Vane and Goetz, 1991)). Spectral processing (also known as spectral mapping, or spectral analysis) refers to "the extraction of quantitative and/or qualitative information from remotely sensed reflectance (or emittance) spectra based on the albedo-, and wavelength-dependent properties of the material" (Mustard and Sunshine, 1999). It encompasses most of the techniques proposed for detection, classification, discrimination, identification, characterization, and quantification of materials in a given hyper-, or multispectral scene (Chang, 2003, 2007; Schott, 2006).

There are numerous review papers devoted to the topic of spectral analysis and geologic remote sensing in the last two decades. In a tutorial paper on spectral unmixing by Keshava & Mustard (Keshava and Mustard, 2002), linear versus nonlinear mixing is clarified and algorithms for linear unmixing are discussed. Recent advances in this subject including geometrical, statistical, and sparse regression-based approaches, along with unmixing challenges are highlighted in (Bioucas-Dias et al., 2012; Plaza et al., 2011). There are also papers concentrated on very specific themes like subpixel detection algorithms (Chang, 2003), nonlinear unmixing (Heylen et al., 2014), image classification (Lu and Weng, 2007; Richards, 2005), support vector machine (Mountrakis et al., 2011), or the evolution of HSI technology (Goetz, 2009; Schaepman et al., 2009; Vane and Goetz, 1991).

On the other hand, a wealth of review papers is dedicated to the application of remotely sensed imagery for natural resource assessment (Agar and Coulter, 2007; Bedell et al., 2009; Gregory and Moore, 1975; Rajesh, 2004; Sabins, 1999). Van der Meer et al. (van der Meer et al., 2012) provided a balanced review of multispectral and hyperspectral remote sensing data, the common products, and their applications to different geologic areas, with a brief discussion on historic and current processing techniques used for both data types. More in-depth evaluation of analytical techniques for extraction of compositional mineralogical information from hyperspectral remote sensing data was provided by Cloutis, some two decades ago (Cloutis, 1996).

While these review papers are seminal and have made science impacts, they either focus on the application of remote sensing in geology, or take stock in specific algorithmic research areas, or are not wide-ranging and up-to-date. None of them provide a categorization strategy for the vast spectral processing methodologies, nor study them in a comparative manner.

In this paper, many of the known and off-the-shelf spectral analysis methods currently available for geologic remote sensing are reviewed. According to the availability and usage of reference data, a categorization scheme is proposed that groups the techniques into knowledge-based and data-driven approaches. The two categories are compared and their outcomes in terms of geologic information are discussed. The methods used to bridge the spectral data and mineralogical, lithological and geochemical datasets are considered. Subsequently, current and potentially new hybridization

concepts are discussed, and future challenges that spectral processing methods are expected to overcome are highlighted.

3.2. Test dataset

Throughout this paper, a hyperspectral datacube of an oil-sand sample is processed and used to illustrate the effectiveness of selected processing techniques discussed within the text. The sample was taken from an exhumed hydrocarbon reservoir located in the eastern edge of the Paraná basin, some 170 km to the NW of São Paulo city, Brazil. The area consists of bitumen accumulations in early Triassic sandstones (de Araújo et al., 2006). XRD analysis shows that the sample is dominated by quartz and montmorillonite, plus titanomagnetite, brushite, and orthoclase as minor phases. Montmorillonite is present as inter-layers and small spots in the sandy matrix, probably as a result of alteration due to hydrocarbon migration (Fig. 1a). The sample was scanned using the sisuCHEMA-SWIR hyperspectral imaging instrument (Roache et al., 2011). Using a 31mm lens, a spatial resolution of 390 µm in length and 680 µm in width was achieved, and later resampled to equal-sized pixels using 0.57 multiplying factor. The 256 spectral bands between 928-2524 nm were transformed into reflectance using built-in instrument routines.

We retained 240 spectral bands and omitted 16 noisier bands at both ends of the spectra (13 between 928-1004 nm, and 3 between 2512-2524 nm). The tray background and at least two pixels at the margins of the sample were masked out. To cancel the illumination variation between the scan lines, we transformed the image into the frequency domain using the fast Fourier transform. In this domain, the sub-horizontal frequencies related to noise and striping was left off from the data, and the remaining part was transformed back into image domain (de Souza Filho et al., 1996). Finally, the spectra of each pixel underwent spectral smoothing using the Sav-Gol filter of 2nd order polynomial (section 3-1-1). To check out the validity of the results, we collected 9 representative spectra using an ASD FieldSpec spectrometer (Goetz, 2009), and a Spectralon panel as a reference to convert the measurements into reflectance. These curves are compared with their relevant image-derived spectra in Fig. 1b. Note the correspondence in overall spectral shape between the two series and specific features at 1900, 2200, 2300, and 2350 nm. The image spectra at around 1650 nm, however, have a higher albedo and are noisier between 1500 to 1800 nm ranges. In general, the image spectra show greater spectral contrast than ASD spectra. The bitumen and montmorillonite are the only spectrally (SWIR) active compounds of the sample and the image spectra are dominated by their diagnostic absorption features (Fig. 1b).

3.3. Categorization and description of the algorithms

There is neither a standardized, universally accepted methodology for the spectral processing of remotely sensed data, nor a comprehensive framework to categorize the existing methods. In the literature, the methods are grouped according to (i) their date of emergence (conventional, or traditional, vs. new, or advanced (Landgrebe, 2003; van der Meer and De Jong, 2002)); (ii) presumed randomness (parametric vs. non-parametric (Keshava et al., 2000; Tso and Mather, 2009)); (iii) type of data they are applied to (multispectral vs. hyperspectral (Richards and Jia, 2006; Schott, 2006)); (iv) the way pixels are treated (hard, or per-pixel vs. soft, or sub-pixel classifier (De Jong and van der Meer, 2005; Lu and Weng, 2007; Schowengerdt, 2007)); (v) the need for training data (supervised vs.

unsupervised (Richards and Jia, 2006; Tso and Mather, 2009)); (vi) and data representation fashion (geometric vs. statistics, or statistical vs. non-statistical (Keshava et al., 2000; Landgrebe, 2003)).

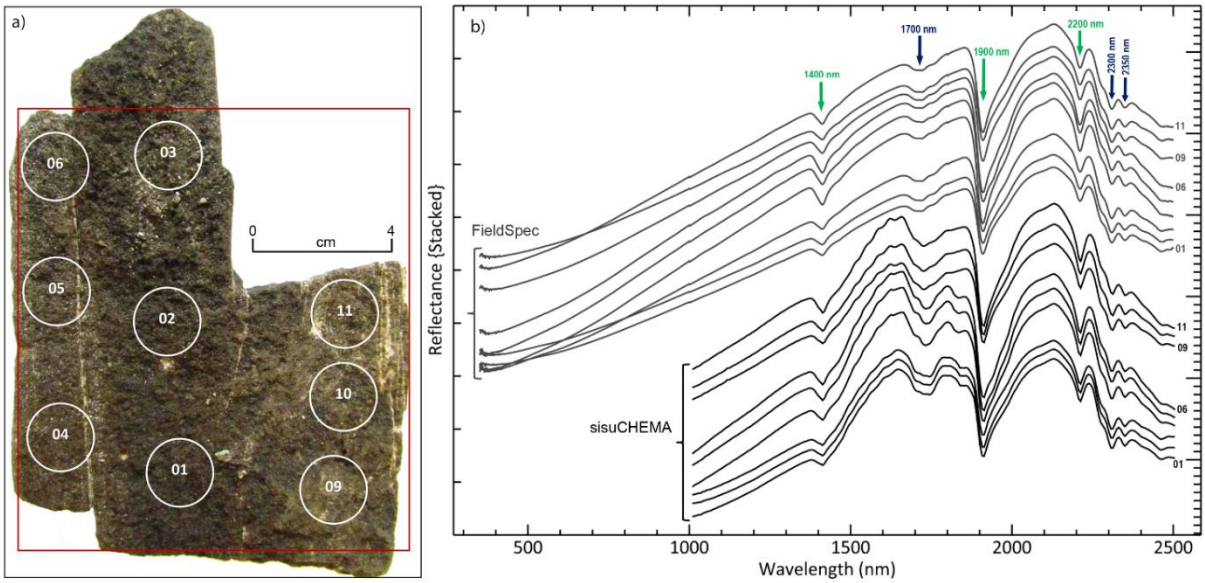

Fig. 1. a) Color photograph of the bituminous sandstone sample, which was scanned to produce the hyperspectral datacube using a sisuCHEMA-SWIR imaging system. The red box (≈ 12 x 11 cm) illustrates the subset used during the processing. Circles are parts measured by the ASD spectrometer. Montmorillonite is evident as white inter-layers and pockets enclosed by dark-colored bitumen. b) Comparison of the representative reflectance spectra collected using the ASD FieldSpec-4 spectrometer (upper stack) and extracted from sisuCHEMA imagery (lower stack). The latter spectra are obtained by averaging the pixels inside each circle. The numbers match the circles in (a) and the spectra are stacked. Absorption bands related to bitumen are indicated by blue arrows (@ ≈1700, 2300, and 2350 nm), and those related to montmorillonite by green (@ ≈1400, 1900, and 2200 nm).

Mustard & Sunshine (Mustard and Sunshine, 1999) proposed three basic categories for spectral processing, including: (i) simple methods of spectral analysis for the definition of broad-scale units, (ii) feature mapping and the absorption band modeling, and (iii) full spectral mapping for material quantification. Schott (Schott, 2006) divided the multitude of spectral analysis algorithms into three perspectives including the geometric, stochastic and spectral feature.

The basis in which a spectral processing technique requires a priori reference data, or not, is used here to establish a categorization scheme. In the case of no reference data, the method is usually able to make direct use of spectral patterns available in a pixel (or measured spectra). In contrast, there are those techniques that try to describe the spectral content of a pixel according to some predefined representative facts known as reference data, or endmembers. This initial difference gives rise to two distinct categories for spectral processing methods: the knowledge-based approach, and the data-driven approach. This division is followed here to review, describe and compare the majority of the spectral processing methods. To make the manuscript more concise, we will avoid providing details on the mathematical formulation of each method, and the reader should refer to the cited work for specifics.

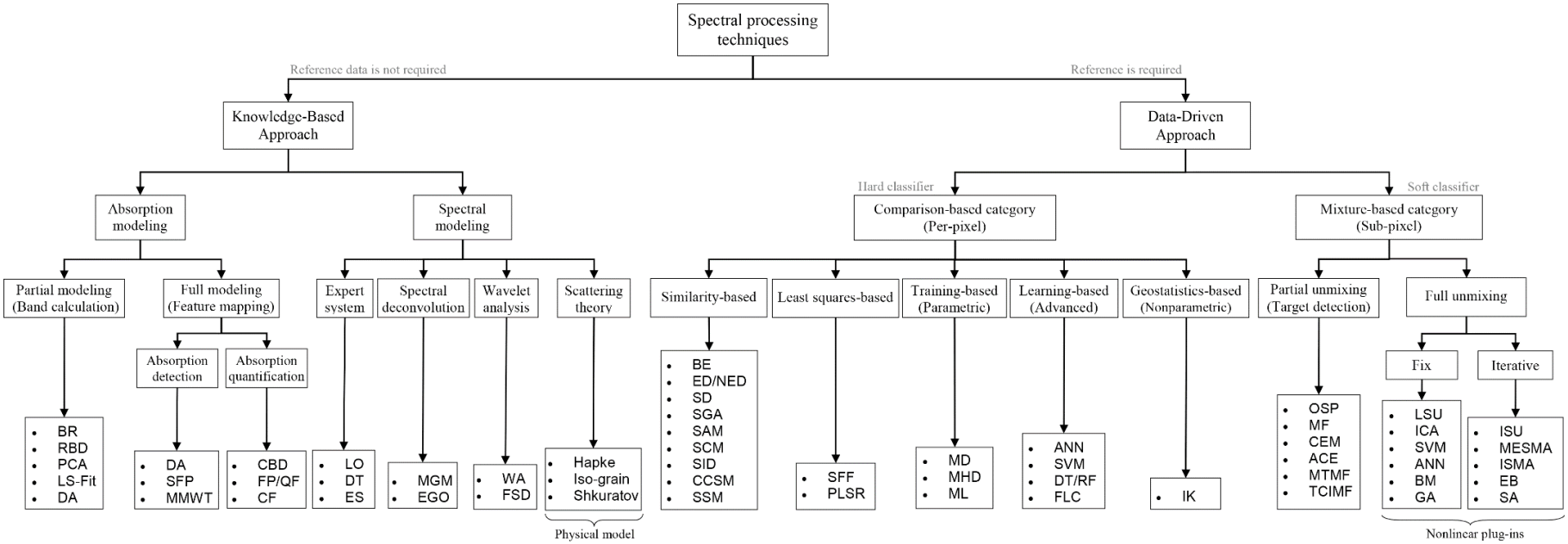

Fig. 2. Taxonomic tree describing the spectral processing methods as proposed and discussed in this work. The "nonlinear plug-ins" indicates that these methods can be plugged into nonlinear algorithms for abundance corrections. The acronyms used in the tree and throughout the text are: BR: Band ratio, RBD: Relative absorption band-depth, PCA: Principal component analysis, LS-Fit: Least-square fitting, DA: Derivative analysis, SFP: Spectral fingerprints, MMWT: Maximum modulus wavelet transform, CBD: Continuum band-depth, FP: Fitted polynomial, QF: quadratic fitting, CF: Curve fitting, LO: Logical operator, DT: Decision tree, ES: Expert systems, MGM: Modified Gaussian model, EGO: exponential Gaussian optimization, WA: Wavelet analysis, FSD: Fourier self-deconvolution, Hapke: Hapke (bidirectional) scattering theory, Iso-grain: Iso-grain scattering theory, Shkuratov: Shkuratov scattering theory, BE: Binary encoding, ED: Euclidean distance, NED: Normalized Euclidean distance, SD: Spectral distance, SGA: spectral gradient angle, SAM: Spectral angle mapper, SCM: Spectral correlation mapper, SID: Spectral information divergence, CCSM: Cross-correlogram spectral match, SSM: Spectral similarity mapper, SFF: Spectral feature fitting, PLSR: Partial least square regression, MD: Minimum distance, MHD: Mahalanobis distance ML: Maximum likelihood, ANN: Artificial neural network, SVM: Support vector machines, DT: Decision tree, RF: random forests, FLC: fuzzy logic classifier, IK: indicator kriging, OSP: Orthogonal subspace projection, MF: Matched filtering, CEM: Constrained energy minimization, ACE: Adaptive coherence estimator, MTMF: Mixture tuned matched filtering, TCIMF: Target-constrained interference-minimized filter, LSU: Linear spectral unmixing, ICA: Independent component analysis, SVM: Support vector machines, ANN: Artificial neural network, BM: Bayesian model, GA: Genetic algorithm, ISU: Iterative spectral unmixing, MESMA: Multiple endmember spectral mixture analysis, ISMA: Iterative spectral mixture analysis, EB: Endmember bundles, SA: Simulated annealing.

3.3.1. Knowledge-based approach

The knowledge-based approach incorporates the user knowledge about the spectral behavior of a target to extract meaningful information from individual spectrum without (at least no direct) reliance on reference data. The building block of knowledge-based approach is the distinct characteristics of the absorption features (i.e. position, depth, asymmetry, and width) in different materials (Clark, 1999; Hunt, 1979; Mustard and Sunshine, 1999; van der Meer, 2004). Generally, each spectrum consists of three basic components: (i) a continuum (also called “hull” or “base line”); (ii) absorption bands; and (iii) residual or noise (Maddams, 1980; Pontual et al., 2008b). Virtually all the knowledge-based methods strive to give an estimate of the quantity or quality of one or more of these components in an interactive or automated way. In this work, we have divided the diversified knowledge-based techniques into two broad categories named “absorption modeling” and “spectral modeling” (Fig. 2). In the former, a limited portion of the spectrum covering a typical absorption is considered for the analysis, whereas in the latter all the absorptions and components of the spectrum are incorporated. The absorption modeling includes band calculation (partial modeling), and feature mapping (full modeling). The spectral modeling on the other hand encompasses several groups including an expert system, spectral deconvolution, wavelet analysis, and scattering theory (Fig. 2). The following section provides a description of the methods available to each of these groups.

3.3.1.1. Band calculation

Band arithmetic is the simplest and most common image processing method. It provides an estimate of the shape or gradient of the absorption feature using basic math operations. The band ratio (BR) uses the difference in reflectance between an absorption band and one of its shoulders (Goetz and Rowan, 1981; Rowan et al., 1974) (Fig. 3b-3f). While it is more resistant to many scene variations, including the topography, the outcome is often ambiguous (Agar and Coulter, 2007). To overcome this limitation, the average of the channels from both feature shoulders was proposed and coined relative absorption band-depth (RBD) (Crowley et al., 1989). The RBD is typically used for the detection of compounds with strong absorption bands (e.g. Al-OH), and can provide a semi-quantitative measure of mineral abundance and/or the “crystallinity” (Clark et al., 1993; Cudahy et al., 2008) (Fig. 3g-3i). Principal component analysis (PCA) makes use of spectral gradients, but in a statistical fashion. It entails a linear projection of the selected bands into a new orthogonal space. The features of interest are then located in a PC band according to the eigenvector values (Crosta and McMoore, 1989) (Fig. 3k, 3l). While BR and RBD are still in use with both MSI and HSI datasets, PCA has been mostly confined to multispectral imagery, perhaps because it relies merely upon empirically chosen input bands, or because of the difficulties in equating PCs to specific features in the imagery (Crosta et al., 2003; Crosta and McMoore, 1989). PCA used together with contrast stretching is comprised in a technique coined the “decorrelation stretch” and has been used to enhance image color and highlight specific targets in MSI data (e.g. silica in TIMS data) (Mustard and Sunshine, 1999). The trend in a feature can also be modeled with least-squares fitting (LS-Fit) and then subtracted from the original spectrum to help predict anomalous regions associated with specific absorbing bands (Green and Craig, 1984) (Fig. 3m, 3n).

In the case study shown in Fig.3, the BR highlights both targets, but the results rely on the selected feature and its shoulder. For bitumen, the results of each feature are different (Fig. 3c-3e)

with the 2300 nm absorption feature providing the best results. The image scores are sharper for the RBD and PCA method, but again they are prone to variation between absorbing bands. On the other hand, the LS-fit shows no superiority over other band calculation methods.

The derivative of a spectrum (first, second, or higher order) involves the calculation of reflectance variation relative to the wavelength (Tsai and Philpot, 1998). It is commonly calculated using a finite approximation method; hence, it bears a resemblance to band calculation. The derivative is a parameter that is more sensitive to the shape rather than the magnitude of the spectra (Zhang et al., 2004). Lower order derivatives seem to be more sensitive to the spectral inflections, whereas higher orders are relatively insensitive to illumination variations (Demetriades-Shah et al., 1990; Tsai and Philpot, 1998). In geological remote sensing, derivative analysis (DA) is exploited for deriving parameters like band position and bandwidth from absorption features in both direct, and indirect manners (Huguenin and Jones, 1986). For example, DA is used directly to separate ore from gangue, or estimate ore content (Murphy and Monteiro, 2013). It is indirectly applied to eliminate background signal, enhance the spectral contrast, or derive other target parameters (Demetriades-Shah et al., 1990; Huguenin and Jones, 1986; Zhang et al., 2004) (see section 5 and 3-1-2-2).

The DA is notoriously sensitive to noise, hence some sort of preprocessing for noise suppression is always required (Cloutis, 1996; Tsai and Philpot, 1998). The most popularly used spectral filters for smoothing spectral data include moving average (median and mean), Savitzky-Golay, Kawata–Minami, cubic spline, geostatistical filter, and wavelet-based thresholding (Mitchley et al., 2009; Oskouie and Busch, 2008; Schmidt and Skidmore, 2004; Tsai and Philpot, 1998). Among them, Sav-Gol is most commonly used (Fig. 1b), because it can provide simultaneous data smoothing and differentiation (Tsai and Philpot, 1998). While the spectral smoothing should be exerted with caution to avoid any loss of information (Cloutis, 1996), overall, it is functional, and has been shown to improve the accuracy of many processing algorithms (Monteiro et al., 2009).

3.3.1.2. Feature mapping

Feature mapping routines aim to fully quantify an absorption band using characteristics like wavelength position, depth, width, and asymmetry. Typically, they demand the absorption be defined or detected, and the continuum to be removed first.

3.3.1.2.1. Continuum removal

To isolate the absorption bands, the continuum should be initially removed. The continuum is the background absorption due to a different process with overall concave shape onto which other absorption bands are superimposed (Clark and Roush, 1984). Though its physical meaning is not thoroughly understood, the continuum is thought to be the manifestation of non-selective multiple scattering due to the matrix effect, Fresnel reflectance, and the presence of spectrally inactive minerals (Clark and Roush, 1984; Mustard and Sunshine, 1999; Thompson et al., 1999). The continuum may also be influenced by the physical (particle size, texture, roughness, etc.) and the chemical properties of the surface, along with the illumination condition (Clark, 1999; Clark and Roush, 1984; Hapke, 1993; Mustard and Sunshine, 1999; Roy et al., 2009). As a result, the prediction of an appropriate function for the continuum is not a straightforward task.

In the empirical way of continuum removal (CR), a convex hull is fitted over the top of the spectrum using straight (tangent) line segments, linking the reflectance maxima and bridging over all absorption features. Next, the original spectrum is divided by this continuum to produce a continuum-removed or hull quotient spectra (Clark and Roush, 1984; Kruse et al., 1993b). Nevertheless, there are several disadvantages around this method: (i) it can be problematic at the endpoints and where the absorption bands are subtle, (ii) it may suppress the broad absorption bands associated with electronic processes in the VNIR range, and (iii) the results may not be similar and comparable (Pontual et al., 2008b; Roy et al., 2009). The modified variant of this routine calculates a linear “local hull” to reach more appropriate results (Clark et al., 2003).

The original version of the modified Gaussian model technique uses a straight line as a continuum in logarithmic reflectance ordinate (section 3-1-4). This representation of the continuum has been adapted for HSI and is shown to outperform the empirical approach (Combe et al., 2006), while remaining immune against grain size effects (Sunshine and Pieters, 1993). To give more flexibility to the continuum, the second order polynomial in wavelength space, the Gaussian low-pass filter, and the low-frequency cubic spline are also suggested (Berman et al., 1999; Clenet et al., 2013; Roy et al., 2009). The Gaussian filter divides each pixel spectrum by its trend curve to reach a normalized reflectance without the need to model the entire spectrum (Roy et al., 2009). On the other hand, the spline approach implements simultaneously the fitting and the mineral identification (Berman et al., 1999). Due to the important role that continuum plays in spectral analysis, its modeling is the subject of active research.

3.3.1.2.2. Absorption detection

Traditionally, absorption detection has been implemented manually, but now there are a number of algorithms to automate this process. A common routine is to search directly for the local spectral minima using the continuum-removed spectra (Clénet et al., 2011; Kruse et al., 1993b); however, since an absorption corresponds to an inflection in the spectrum, DA can be a choice for its detection. The absorption occurs where the fifth derivative of a spectrum equals zero, the fourth derivative has a positive sign, and the second derivative is negative (Brown, 2006; Huguenin and Jones, 1986). In the scale-space representation of the spectrum, inflection points seem to remain stable at different scales, therefore methods like spectral fingerprints (SFP) can robustly recover them (Piech and Piech, 1990; Piech and R., 1987). The SFP applies a convolution with a Gaussian kernel (with incremental variance) to achieve the scale-space representation, and the first-order derivative to identify the points themselves. The maximum modulus wavelet transform (MMWT) is the generalized form of SFP that instead makes use of second-order derivative in wavelet domain (Hsu, 2003). Other detection methodologies worth mentioning are local boundary hunting and unimodal segmentation (van der Meer, 1994; Zhouyu et al., 2007). The continuum removal and absorption detections are prerequisites to absorption quantification.

3.3.1.2.3. Absorption quantification

There are correlations between spectral feature characteristics (wavelength position, shape, and asymmetry) of absorption bands and the mineralogic content of a target. The wavelength is related to the chemistry of a mineral, whereas the intensity (depth) of the feature is proportional to the abundance of the compound (Clark and Roush, 1984; Duke, 1994; Hunt, 1979). Typically, the

abundance of a material is quantified by calculating the depth of its diagnostic absorption feature relative to the continuum background (Clark and Roush, 1984; Cudahy et al., 2008; Sunshine and Pieters, 1993). However, there are several drawbacks to the continuum band-depth (CBD) technique for abundance estimation: (i) the depth of an absorption is more or less proportional to particle size and amount of opaque materials (Clark, 1999; Gaffey et al., 1993), (ii) the parameter may become saturated for certain minerals (Pompilio et al., 2009; Pontual et al., 2008a), (iii) it may behave nonlinearly in relation to areal/weight percentage due to intimate mixing (Dalton et al., 2004; Shipman and Adams, 1987; Thompson et al., 1999), and (iv) it is likely for the absorption bands to overlap each other (Cudahy et al., 2008). Even so, the CBD is still the most accepted spectroscopic-based method for abundance quantification (Fig. 3o, 3p) (e.g. (Haest et al., 2012)). Recently there have been attempts to boost this criterion. For example, a regression model named vegetation corrected continuum depth (VCCD) is designed to compensate for the obscuring effect of vegetation on the 2.2 µm band depth of the Al-OH absorbing species (Rodger and Cudahy, 2009).

The “asymmetry” of an absorption is defined by the wavelength difference between a minimum and its two shoulders (van der Meer, 2004), or by the difference in the area of the two halves, whereas the “width” is typically measured as the full-width at half-maximum (FWHM) (Clénet et al., 2011; Kruse et al., 1993b) (Fig. 3q, 3r). To track the shifts in wavelength position that are associated with compositional variation, the linear approximation is proposed and is shown to give a very rough estimate of the parameter (van der Meer, 2004), while fitted polynomial (FP) of higher orders is shown to achieve more accurate results. For instance, a fitted 4th order polynomial is used to model and map the level of Tschermak substitution in white micas (Cudahy et al., 2008). For broader absorptions like iron oxides, however, a 2nd order polynomial has been utilized (Cudahy and Ramanaidou, 1997) (Fig. 3s). Whereas the polynomial is commonly fitted to entire absorption, within a method called quadratic fitting (QF), three spectral bands are used to achieve the estimation (Rodger et al., 2012). The derivative of the fitted polynomial is then used to derive the wavelength information.

A more general form of FP is called curve fitting (CF), within which a curve of a specific type is fitted to the absorption (normally after CR), to facilitate the extraction of noted information. For example, absorption bands are modeled using the amplitude (α), a central frequency ($\nu 0$), and full-width at half-maximum (δ) of Gaussian, Lorentzian, or Voight functions. In order to achieve a greater number of measurement points however, a cubic spline fit can be used to interpolate the hyperspectral data (Brown, 2006) (Fig. 6b).

The CBD image in Fig.3, which is calculated after the continuum removal, has sharp boundaries for both compounds and portrays the relative content of the targets well (Fig. 3o). The asymmetry of the 2200 nm feature yields averages results. Over the white, montmorillonitic patches exclusively, the feature is almost symmetrical, but where the clay is mixed with bitumen, it is more asymmetrical (Fig. 3q). The minimum of the same feature calculated by the FP method varies between 2206 -2211 nm, with the higher wavelengths being related to isolated patches of pure montmorillonite. The total area of the 1700 nm absorption (Fig. 3r) is correlated to the abundance of bitumen, however, it is better describing the limits of the high bituminous parts of the rock (Fig. 3g).

3.3.1.3. Expert systems

The objective of expert systems (ES) is to automate the process of mineral identification by gathering together the spectroscopic knowledge and feature mapping procedures (Clark et al., 2003; Kruse et al., 1993b). In the case of the logical operator (LO), several BRs or RBDs are combined to attain a type of binary hard classifier. Although the LO was developed for the analysis of both multi-, and hyperspectral datasets (Brown, 2010; Mars and Rowan, 2006), transferable thresholds appear to be elusive. A similar solution may come from the knowledge-based decision tree (DT) partitioning techniques (Fig. 3t) (Friedl and Brodley, 1997; Tso and Mather, 2009), but again it suffers from the same problem. Other systems try to mimic human experts by extracting and measuring the noted spectral parameters, and then devising (hard) rules for mineral identification (Cudahy et al., 2008; Kruse and Lefkoff, 1993; Kruse et al., 1993b) (see also section 5).

3.3.1.4. Spectral deconvolution

Unlike the CF, spectral deconvolution techniques strive to deconvolve the whole spectrum into the three noted components (Maddams, 1980). The Modified Gaussian model (MGM) is a (parametric) spectral deconvolution method for modeling the electronic transition bands of reflectance spectra (Sunshine et al., 1990). MGM assumes that bond length, and the distribution of absorbing energies are random variables (Huguenin and Jones, 1986; Sunshine et al., 1990), hence their absorbing bands can be described by a Gaussian distribution (Huguenin and Jones, 1986; Sunshine et al., 1990). The MGM states that for a given absorption, there is a distribution in energy (x) with a standard deviation (δ), mean (μ), and amplitude (s). This is given by equation (1):

$$m(x) = s \times e^{\frac{-(x^{-1}-\mu^{-1})^2}{2\delta^2}} \qquad (1)$$

where m (x) is the modified Gaussian expressed as a function of energy. To establish an additive linear system, the reflectance data are first converted to natural log reflectance. The initial conditions for each band (position, width, strength) and the continuum are provided manually, and then a nonlinear least-squares algorithm is used to determine the MGM solution to the spectra in an iterative way (Sunshine et al., 1990). According to this model, absorption bands are essentially symmetric and any apparent asymmetry is caused by hidden overlapping bands (Brown, 2006; Sunshine and Pieters, 1993).

MGM deconvolution has been successfully applied in the lab to extract modal abundances and compositional information from charge transfer absorptions bands in the VNIR, in both linear and intimate mixing scenarios, as well as overtone and combination of OH absorption bands in the SWIR range (Mustard, 1992; Sunshine and Pieters, 1993). Recently, the original MGM has been modified to automatically handle large amounts of hyperspectral datasets (Clénet et al., 2011), and then is implemented to characterize the modal and chemical composition of *a priori* unknown mafic mineralogy on Earth and Mars (Clenet et al., 2013).

A recent variant of the MGM is called exponential Gaussian optimization (EGO), which is designed to account for the non-Gaussian behavior of the absorption features, and alike the MGM, it decomposes a spectrum into several EGO models superimposed on a continuum. The technique is shown to be able to model band asymmetry and flattening due to saturation effect and nested bands (Pompilio et al., 2010; Pompilio et al., 2009).

3.3.1.5. Wavelet analysis

MGM deconvolution and all the other processing methods are implemented in signal (spectral) domain, but there are methods specific to the wavelet or frequency domain as well. Wavelet analysis (WA) has been attractive to hyperspectral data processing, because the signal is varying in both amplitudes (feature depth), and scale (feature width) (Bruce and Jiang, 2001). The WA decomposes a spectrum into a series of shifted and scaled versions of the mother wavelet function as either continuous wavelet transforms (CWT), or discrete wavelet transforms (DWT) (Bruce et al., 2001). The CWT aims to deconvolve the spectrum into linearly additive wavelets, enabling the isolation of spectral features from their continuum over a broad spectral region. In such a representation, narrow absorption features in the original spectrum are captured by the low-scale wavelet component, while the continuum is associated with the higher scale components (Rivard et al., 2008). The lower components are chiefly used to map chemical variations associated with given minerals (Rivard et al., 2008) (Fig. 6).

In the frequency domain, Fourier self-deconvolution (FSD) is used to narrow the width of absorbance bands, without affecting the corresponding position, or its total area (Kauppinen et al., 1981). In this method, the spectrum is Fourier transformed to the frequency domain, multiplied by an exponential function, and then is transformed back to the spectral domain. The result is a mathematically enhanced spectrum with more distinct absorptions in the overlapping wavelengths (Griffiths and de Haseth, 2007). Clearly, FSD is an enhancement technique; nonetheless, it has been rarely used with reflectance spectra.

3.3.1.6. Scattering theory

Scattering theory utilizes a radiative transfer equation to describe the scattering behavior of light from particulate media (Hapke, 1993). An approximate analytic solution to this equation is provided by what is called scattering theories. The most popular of them are Hapke (Hapke, 1981), iso-grain (a derivative of the first) (Hiroi and Pieters, 1992), and Shkuratov scattering theories (Shkuratov et al., 1999). These models are able to give an accurate estimate of abundance and grain size in the case of the powdered surface. Both theories have proven to be effective in laboratory and field testing on deriving abundance as well as grain size information to within 5-10% accuracy (Mustard and Pieters, 1989; Poulet and Erard, 2004; Shipman and Adams, 1987). Even so, they are notoriously complex and require extensive empirical data to perform, which makes them notably unpopular (e.g. (Bioucas-Dias et al., 2012; Cloutis, 1996; Keshava and Mustard, 2002)). An alternative strategy has been the simpler, but physics-inspired nonlinear models (section 3-2-7).

3.3.2. Data-driven approach

Data-driven methods illustrate an alternative spectral analysis approach in which only the hyperspectral data themselves and some additional reference data (spectra) are required. Based on the algorithm involved, reference data are commonly called training classes, or endmember sets (e.g. (Chang, 2007; Richards and Jia, 2006)), each comprised of a single or multiple spectra (Boardman, 1989; Winter and Winter, 2000) (Fig. 4). The endmembers may be imported to the image (e.g. from a spectral library), or derived from it. The latter has distinct advantages, and therefore is mostly preferred and practiced (Chang, 2013).

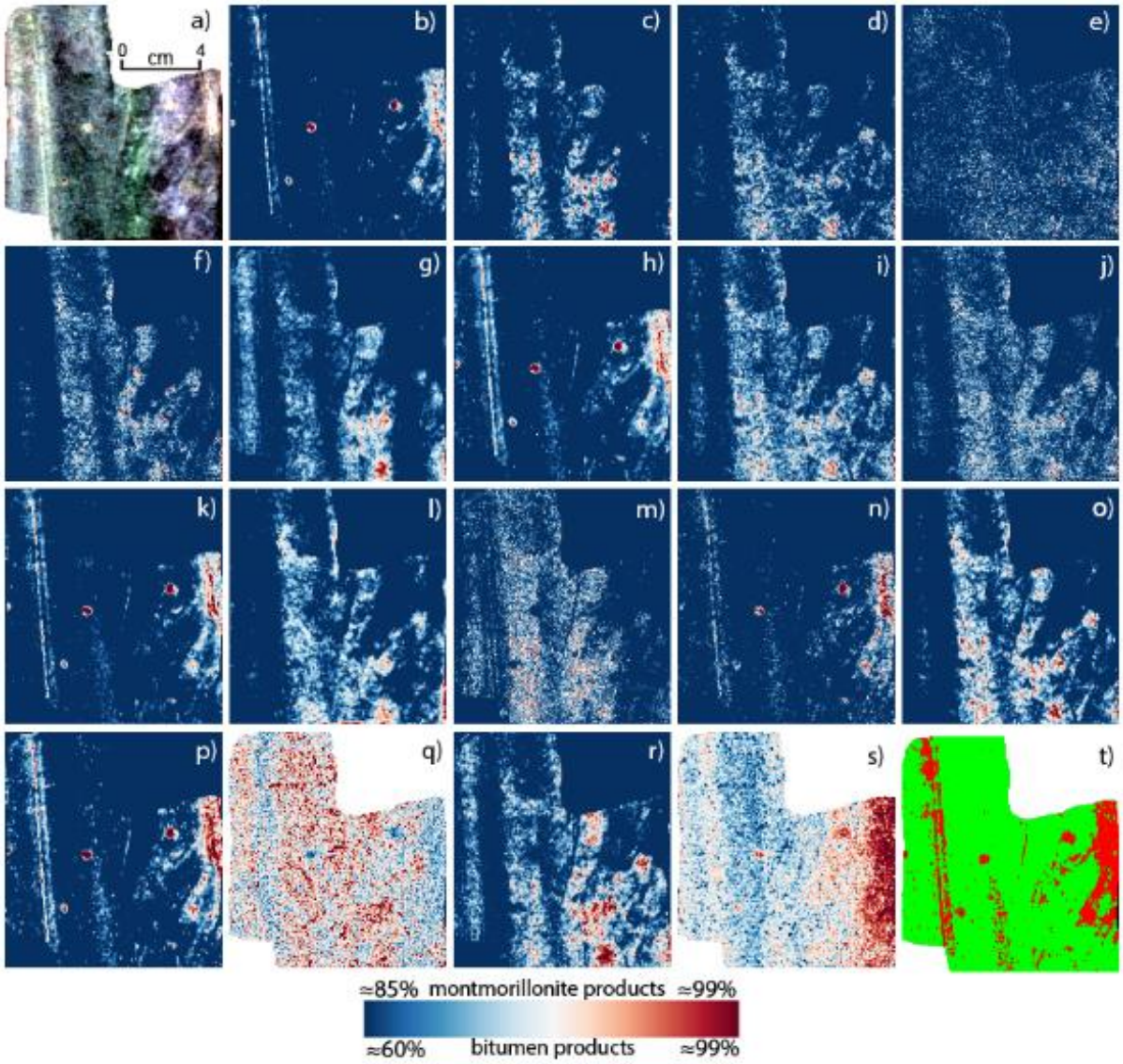

Fig. 2. Examples of mineralogic products of the oil-sand sample extracted from the sisuCHEMA-SWIR hypercube dataset using the knowledge-based approach. **a)** false color composite of the cube (RGB=bands at R_{1174}, R_{1801}, and R_{2425}). **b)** BR of the left shoulder of the 2200 nm absorption using R_{2170}/ R_{2207}. **c)** BR of the right shoulder of the 1700nm absorption using $\sum R_{1783-1801} / \sum R_{1720-1739}$. **d)** BR of the left shoulder of the 2300 nm absorption using R_{2282}/ R_{2307}. **e)** BR of lethe ft shoulder of the 2350 nm absorption using R_{2331}/ R_{2350}. **f)** BR of the right shoulder of the 2350 nm absorption using R_{2375}/ R_{2350}. **g)** RBD of the 1700 nm absorption using $\sum R_{1657-1664} + \sum R_{1789-1795} / \sum R_{1720-1739}$. **h)** RBD of the 2200 nm absorption using R_{2170} + R_{2251}/R_{2207} + R_{2213}. **i)** RBD of the 2300 nm absorption using R_{2282}+R_{2331}/R_{2307}+R_{2313}. **j)** RBD of the 2350 nm absorption using R_{2331}+R_{2375}/R_{2350}. **k)** Inverse of PC2 of the 2200 nm absorption using PCA of bands between R_{2145} –R_{2244}. **l)** PC1 of the 2300 nm absorption using PCA of bands between R_{2282} – R_{2331}. **m)** LS-Fit of the 1700nm absorption used with bands between R_{1776} – R_{1795}, predicting R_{1726}. **n)** LS-Fit of the 2200nm absorption used with bands between R_{2170} – R_{2188}, predicting R_{2213}. **o)** Relative abundance of bitumen calculated by CBD of the 2300nm absorption. **p)** Relative abundance of montmorillonite calculated by CBD of the 2200nm absorption. **q)** Relative asymmetry of the 2200nm absorption calculated using the area to the left-, and right-side of the absorption minimum between 2151-2238 nm (blue: almost symmetrical, red: asymmetrical). **r)** Total area of the 1700 nm absorption calculated using the 2nd order FP between 1657-1789 nm. **s)** Montmorillonite composition calculated using the 4th order FP between 2151-2244 nm (blue: 2206 nm, red: 2211 nm). **t)** DT-based classification using a combination of knowledge-based techniques and interactive thresholding (green: bitumen, red: montmorillonite). The thresholds vary for each product, but for montmorillonite-related products is between 85%-99.4%, and for bitumen between 60-99.4%. The absorption features are defined in Fig. 1b.

We have divided the sheer number of data-driven processing methods into two broad categories named “per-pixel” and “sub-pixel” (Chang, 2003; Lu and Weng, 2007; Tso and Mather, 2009) (Fig. 2). The per-pixel category, which is also called a hard classifier, compares each reference spectra to

unknown pixels one by one, based on criteria like similarity metric, image statistics, or least-square estimation. By contrast, in the sub-pixel category, or soft classifier, multiple and variable labels at each pixel is permissible (Keshava and Mustard, 2002; Schowengerdt, 2007). The mixture-based group may be further divided into two main sub-groups, known as partial and full unmixing. The following section provides a description of the methods available to each group.

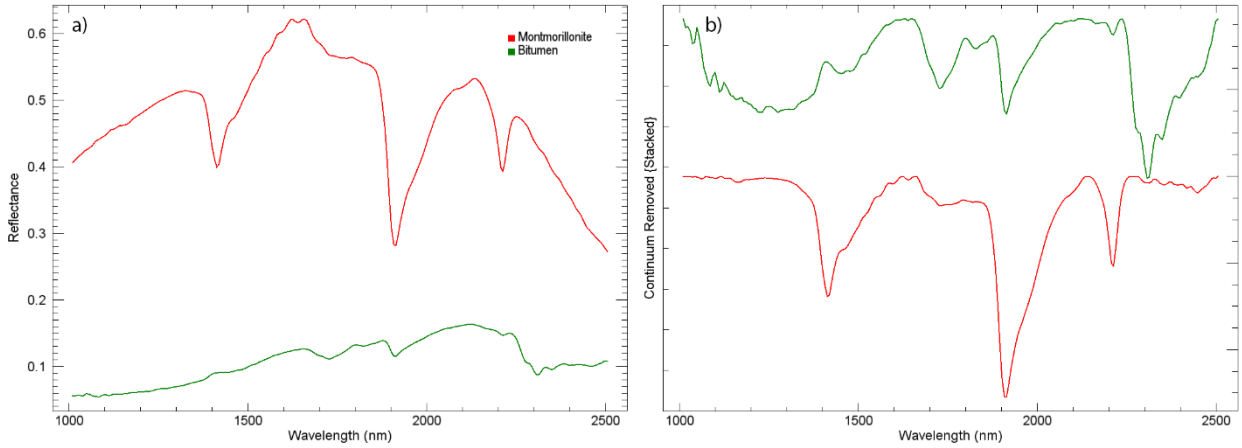

Fig. 4. The two endmembers used with the data-driven approach. Montmorillonite was automatically extracted from the imagery (using the sequential maximum angle convex cone (SMACC) tool embedded in the ENVI (Environment for visualizing images) software from Exelis Visual Information Solutions, Boulder, Colorado), whereas bitumen was defined manually based on the knowledge-based results, and then nonlinearly tuned to reduce the widespread mixing effect of montmorillonite. a) Normal representation, b) The continuum removed and stacked of (a).

3.3.2.1. Similarity-based group

The spectral similarity (or matching) techniques strive to find a measure of mathematical or physical similarity between a known reference spectrum, ***x***, and an unknown test (target) spectrum, ***y*** (van der Meer, 2006a). The binary encoding (BE) technique encodes the test and reference spectra into 0 and 1, based on the mean of the spectrum, and then uses an exclusive OR function to measure their similarity (Mazer et al., 1988). The reference and test spectrum can also be compared however, based on the "angle" or the "distance" between them in n-dimensional space. The spectral angle mapper (SAM) technique assesses the spectral angle between ***x*** and ***y*** by applying a dot product multiplication between them (Kruse et al., 1993a) (Fig. 5a). The Euclidean distance (ED), on the other hand, measures the distance of ***x*** and ***y*** in n-dimensional space (Richards and Jia, 2006) (Fig. 5b). It has been shown that SAM is essentially the ED when the spectral angle is small (Du et al., 2004; van der Meer, 2006a). The normalized Euclidean distance (NED) works in the same manner as the ED, but it normalizes the vectors first, causing the values to range between 0 and 1 (Keshava, 2004; Robila and Gershman, 2005) (Fig. 5c). Spectral distance (SD) is another measure in this family that is very similar to NED, except that the calculation is carried out in natural logarithm reflectance, and the ***x*** and ***y*** are continuum removed beforehand (Combe et al., 2005).

To enhance the precision of SAM, two new variants called RAF-SAM (Wang et al., 2009) and BAO-SAM (Keshava, 2004) were proposed. The former represses the impact of an additive factor on the spectral angle value in the feature space, while the latter attempts to iteratively increase the angular separability between ***x*** and ***y*** by selecting the optimum bands. Another variation is the Spectral

Gradient Angle (SGA), which calculates gradient or slope changes for **x** and **y**. Some experiments, however, have shown no superiority of SGA over SAM (Robila and Gershman, 2005).

SAM is regarded as a variant of the more general Pearson correlation coefficient, and based on that, the spectral correlation mapper (SCM) method was introduced (Carvalho Junior and Menezes, 2000). The SCM has distinct advantages in providing a direct measure of the similarity between the shapes of two spectra and has the ability to detect false positive results. The major difference between SAM and SCM is that SCM standardizes the data, centralizing the cloud in the mean of ***x*** and ***y;*** therefore the results inevitably range between -1 to 1 (Carvalho Junior and Menezes, 2000) (Fig. 5d). Cross-correlogram spectral matching (CCSM) is another similarity measure based on correlation (van der Meer and Bakker, 1997). Here, a cross correlogram is constructed by calculating the cross correlation coefficient between ***x*** and ***y*** at different match positions, ***m***, by shifting the ***x*** spectrum. The cross correlogram for a perfectly matching reference and test spectrum is a parabola around the central matching number (*m*=0) with a peak correlation of 1. Deviations from this shape indicate a different test spectrum (van der Meer and Bakker, 1997).

The other similarity technique is spectral information divergence (SID) that calculates the distance between the probability distributions produced by the spectral signatures of the two spectra (***x*** and ***y***) using the means of Kullback-Leibler information measure (Du et al., 2004) (Fig. 5e). In this measure, spectral variations among the spectral bands can be captured more effectively in a stochastic manner [32].

There are other similarity algorithms in which two measures are combined to generate a hybrid method inheriting the benefits of both sides. The spectral similarity mapper (SSM) calculates two numbers for each pixel; the first is the Euclidean distance between the ***x*** and ***y*** and the second is a correlation value, which respectively gives a measure of brightness difference and similarity in shape between ***x*** and ***y*** (Granahan and Sweet, 2001; www.exelisvis.com/ProductServices/ENVI.aspx):

$$SSM\ (\boldsymbol{x},\boldsymbol{y}) = \sqrt{d_e^2 + \hat{r}^2} \quad , \hat{r} = 1 - r^2 \quad (2)$$

where $\boldsymbol{d}_e$ is the ED and $\boldsymbol{r}^2$ is the correlation coefficient between the target and reference spectrum. The spectral similarity value will range between 0 and $\sqrt{2}$ (Granahan and Sweet, 2001) (Fig. 5f). In a similar way, the SID and SAM are combined by trigonometric functions into mixed measures (Du et al., 2004) (Fig. 5g).

In the case study, shown in Fig.5, the performance of similarity metrics considering most methods is very close. The ED has lower performance in delineating the boundary of montmorillonite and bitumen (Fig. 5b), whereas the SID ×sin (SAM) shows better results in describing the boundaries of the target, specifically for bitumen (Fig. 5g).

The significance of similarity measures to search out the spectral libraries or to analyze multi-, and hyperspectral images have led to several comparative studies. Van der Meer (van der Meer, 2006a) compared the performance of deterministic-empirical measures (SAM, ED, and CCSM) relative to the stochastic measure (SID), and concluded that: (i) SID outperforms other techniques, (ii) CCSM better exploits the overall shape of the spectrum, (iii) SAM and ED give nearly similar results, and (iv) CCSM is more sensitive to noise. The major problem associated with similarity measures, however, is their inability to deal with mixed spectra as well as subjective thresholding.

3.3.2.2. Least squares-based group

Least squares regression techniques aim to model dependent variables by the means of an independent variable (Esbensen, 2006). Spectral feature fitting (SFF) is a feature-based methodology that uses linear least square regression to work out a fit between a continuum removed reference (***x***) and test (***y***) spectra (Clark et al., 1990). The fit (matching) between absorption features comprised in ***y*** and ***x*** is provided by the total root mean square (RMS) error of the regression and the coefficient of determination (R^2). This method is able to use single or multiple features over a spectrum, and accepts user-defined constraints (Clark et al., 1990; Clark et al., 2003; Xu et al., 2011) (Fig. 5h). In practice, SFF uses the user knowledge of the features and CR procedure to do the regression; hence, it can be considered a hybrid method (section 5) as well.

Partial least square regression (PLSR) is another technique of this family that is now gaining popularity in spectral analysis. The PLSR, which inherits features from principal component analysis and multiple regression, establishes a linear regression model to concentrate information contained in the spectra in a few latent variables that are optimized to produce the best correlation with the desired property of interest (Esbensen, 2006). PLSR is mostly used to relate spectral data to other non-spectral variables. For example, it is utilized to build a predictive mineral model from VNIR-SWIR spectra, or to compare LWIR spectra with X-ray diffraction (XRD) results, as well as thin section studies (Cudahy et al., 2001; Goetz et al., 2009; Hecker et al., 2012).

3.3.2.3. Training-based group

Traditional training-based classifiers aim to cluster the imagery by comparing the test spectrum with the training classes using a statistical criterion (Landgrebe, 2003; Tso and Mather, 2009). The minimum distance (MD) classifier takes into account the Euclidean distances, whereas the maximum likelihood (ML) classifier calculates a probability distance using the mean and covariance matrices of the clusters. The Mahalanobis distance (MHD) classifier, on the other hand, is direction-sensitive, but assumes an equal covariance for all classes (Landgrebe, 2003; Richards and Jia, 2006; Schowengerdt, 2007; Tso and Mather, 2009). While these algorithms are helpful for land-cover classification, generally they are found to be inefficient at the practice of lithology discrimination on both multi- and hyperspectral datasets (Agar and Coulter, 2007; Mustard and Sunshine, 1999; Sabine, 1999).

3.3.2.4. Learning-based group

Different studies show the considerable advantages of artificial neural networks (ANN) over conventional training-based methods (Licciardi and Del Frate, 2011; Mas and Flores, 2007; Richards and Jia, 2006). ANN has the ability to learn the relationship between a set of example patterns, generalize and combine the results, and then apply it to new input patterns (Mas and Flores, 2007; Yang, 1999). ANN is free of distribution assumptions, is capable to generalize even in noisy environments, and works rapidly once it is trained (Foody et al., 1997; Licciardi and Del Frate, 2011). A variety of ANN, including back-propagation neural network (BPN) and self-organizing maps (SOMs) are shown to be good supervised (hard) classifiers for HSI datasets (Mas and Flores, 2007; Villmann et al., 2003; Yang, 1999) (Fig. 5i). On the other hand, it is possible to train the ANN with endmembers or with mixed pixels and derive abundance maps in both linear and nonlinear mixing scenarios (Foody et al., 1997; Licciardi and Del Frate, 2011). In linear scenarios, multilayer perceptron (MLP) models are used for feature reduction, as well as abundance estimation (Licciardi and Del Frate,

2011). In nonlinear scenarios, ANN has been used, for example, to estimate the SiO_2 content of igneous rocks (Ninomiya, 1995). A thorough review of ANN for the analysis of remotely sensed data is provided by Mas & Flores (Mas and Flores, 2007).

Support Vector Machines (SVM) draws on statistical learning theory for pattern recognition (Vapnik, 1998). This non-parametric approach is based on constructing a separating hyperplane (or decision boundary) within an n-dimensional feature space using the properties of training samples. The so-called structural risk minimization is used to iteratively optimize the margins between the hyperplane and the closest training samples, known as support vectors (Vapnik, 1998). The classifier only requires this small subset at class boundary for classification, even with the case of high dimensional datasets (Melgani and Bruzzone, 2004; Tso and Mather, 2009). In geologic remote sensing, the SVM is mostly utilized in lithology classification and occasionally in mineral detection and ore discrimination (Cracknell and Reading, 2014; Gilmore et al., 2008; Monteiro et al., 2009; Waske et al., 2009) (Fig. 5j). The state-of-the-art and diverse applications of SVM in data mining are reviewed by Mountrakis et al. and many others (Camps-Valls and Bruzzone, 2009; Lu and Weng, 2007; Mountrakis et al., 2011; Plaza et al., 2009). Typically the prime issue around SVM and ANN is reported to be parameter assignment (Lu and Weng, 2007).

Another non-parametric technique to be noted is the decision tree (DT), which hierarchically subdivides the dataset based on a set of tests defined at each of its branches. While univariate DT is shown to outperform the ML classifier or yield comparable results (Friedl and Brodley, 1997; Pal and Mather, 2003), its recent variant is known as Random Forests (RF) – logic-based learner, has been shown to be a superior choice for lithology classification when compared to SVM and ANN (Cracknell and Reading, 2014).

Where applied to the case study, SVM displays better performance relative to ANN (Fig. 5i, 5j); however, they cannot outperform the results produced by the DT based on user-provided thresholds (Fig. 3t).

Fuzzy logic classifier (FLC) accommodates multiple class membership for each pixel considering fuzzy rules (Bardossy and Samaniego, 2002; Lu and Weng, 2007; Wang, 1990). The hindrance of the method however, is to find the correct rules and to select the relevant features. The sub-pixel niche that FLC belongs to has been mostly filled by unmixing methodologies (section 3-2-7).

3.3.2.5. Geostatistics-based group

Indicator kriging (IK) has been proposed as an efficient geostatistical technique for image classification and extraction of absorption features for mineral mapping. The IK is a non-parametric method in which variables are transformed into (0, 1) depending on the presence or absence of a feature of interest, or whether a threshold is exceeded or not. It directly benefits from spectral information in a supervised manner and has the capability of dealing with spatial information (van der Meer, 1994, 1996, 2006b). Although areas smaller than a pixel can be estimated, the probability distributions of ordinary kriging was integrated with Bayesian statistics to yield a hard classifier (van der Meer, 2006b). The IK was as well used by Chiang (Jie-Lun et al., 2014) to estimate the class probabilities in feature space instead of image space. Individual pixels were then assigned to classes using the maximum class probability. It has been shown that this linear hard classifier can outperform nonlinear SVM method (Jie-Lun et al., 2014).

3.3.2.6. Partial unmixing group

In many applications, it is not essential to fully decipher the content of a pixel. Instead the aim is to isolate spectral features of interest from the background (Ahlberg and Renhorn, 2004). In this case, the problem is reduced to the detection of spectral signatures that match the known target (Camps-Valls et al., 2012; Chang, 2003; Manolakis et al., 2003; Mustard and Sunshine, 1999; Schott, 2006). Target detection algorithms vary from matched filtering (MF) (Boardman et al., 1995), constrained energy minimization (CEM) (Chang et al., 2000), orthogonal subspace projection (OSP) (Harsanyi and Chein, 1994), and adaptive coherence estimator (ACE)(Kraut et al., 2005), to target-constrained interference-minimized filter (TCIMF)(Ren and Chang, 2000), and mixture tuned matched filtering (MTMF) (Boardman and Kruse, 2011) (Fig. 2). An exhaustive list of target detectors is provided by (Chang, 2003) and (Manolakis et al., 2003). They are compared theoretically and practically in (Manolakis and Shaw, 2002).

In the OSP detector, the subspace of the background basis functions is removed from the analyzed pixel, leaving only the part related to the known endmember (Harsanyi and Chein, 1994) (Fig. 5k). In MF, the response of the target signature is maximized and the response of the background subspace is minimized by a likelihood ratio, thus matching the signature (Boardman et al., 1995) (Fig. 5l). The CEM utilizes a finite impulse response filter to pass through the target signature, while minimizing its output energy resulting from the composite background (Chang et al., 2000) (Fig. 5m). Mathematically though, the MF is a mean-centered version of the CEM (Chang, 2003). The TCIMF can be viewed as the extension of CEM, where the filter not only detects the desired target and eliminates the background, but also is constrained to exclude the response of non-targets (Ren and Chang, 2000) (Fig. 5n). The ACE detector is based on the generalized likelihood ratio and thus is invariant to relative scaling of the test data (Kraut et al., 2005) (Fig. 5o). In the MTMF, beside the MF, an “infeasibility” image is also calculated for each target signature and then the predominant material and its abundance is determined using the combined criteria of high MF and low Infeasibility scores (Boardman and Kruse, 2011).

In essence, these detectors carry out a partial unmixing and their output is a single score (abundance of the target) per pixel, which bear some resemblance to similarity measures. These methods are not yet comparatively studied for geologic applications. However, based on the current case study, it seems they yield acceptable results for the smaller bright target (Fig. 5k, 5m, 5n), with the exception of the MF and the ACE. These methods are only able to detect the strongest signals (Fig. 5i, 5o); but for the pervasive darker signal, they yield poor results. For example, MF and ACE failed to detect the signal altogether (Fig. 5i′, 5o′), while the CEM has confusion in discriminating between the targets (Fig. 5m′). The performance of these detectors is found to broadly improve using clustering or feature extraction techniques (e.g. (Funk et al., 2001)).

3.3.2.7. Full unmixing group

Full unmixing attempts to linearly (or nonlinearly) decompose the pixel spectrum into a collection of deterministic constituent spectra (endmembers) and to estimate their corresponding abundances (Adams et al., 1986; Smith et al., 1990). A linear mixture model (LMM) is valid where the mixing is from a checkerboard mixture of macroscopic scale and the incident light interacts with just one material, whereas multiple scattering between the light and materials of the scene results in nonlinear mixing (Hapke, 1993; Keshava and Mustard, 2002). The simplicity of the LMM has given

rise to many algorithmical developments in this era as are reflected in the relevant review papers (Bioucas-Dias et al., 2012; Keshava and Mustard, 2002). Spectral unmixing typically consists of three major steps: (i) searching for the number and the best endmember set to represent the entire scene variation; (ii) finding the best endmember subset that firmly accounts for the spectral variation in a pixel; and (iii) estimating the accurate abundances of each endmember within the pixel.

The first step is called endmember extraction (EE), and the last two are optimization and inversion steps which are performed simultaneously. Because the type and the number of endmembers has a profound effect on unmixing results, EE has been the focus of many studies leading to numerous algorithmic developments (e.g. N-FINDER (Winter and Winter, 2000)). A fair review of the advancements on this field is provided in (Chang, 2013) and the performance of the algorithms are compared in (Camps-Valls et al., 2012; Plaza et al., 2004; Winter and Winter, 2000). Despite this intense diversification, there is still no guarantee for the extraction of desired geological endmembers in a given scene (Rivard et al., 2009). One solution has been the partitioning of the input dataset (García-Haro et al., 2005; Zare and Gader, 2010) or the inclusion of spatial preprocessing (Zortea and Plaza, 2009), while the other inevitable key has been the supervised sample/spectral collection (Rivard et al., 2009). For an image scene, apart from the type, the number of endmembers are largely unknown *a priori*. Traditionally, minimum noise fraction (MNF)(Green et al., 1988) has been used to estimate the inherent dimensionality (ID) of the data, and recently a rich variety of concepts including the virtual dimensionality (VD) are offered to fulfil this requirement (Chein and Qian, 2004).

When endmembers are identified, the problem reduces to model inversion. In linear spectral unmixing (LSU), the unconstrained or constrained least-squared inversion, singular value decomposition etc. are used to solve the inversion problem (Boardman, 1989) (Fig. 5p). In addition, there are techniques like independent component analysis (ICA) (Comon, 1994; Nascimento and Bioucas Dias, 2005), SVM (Camps-Valls and Bruzzone, 2009), ANN (Licciardi and Del Frate, 2011), Bayesian model (BM) (Dobigeon et al., 2008), and genetic algorithm (GA)(Farzam et al., 2008) which are specifically adapted for linear unmixing process. Such unmixing has been performed on wavelet-transformed spectra as well (Mitchley et al., 2009).

Spectral unmixing may end up giving unrealistic results, because the selected endmembers might not account for the spectral variability present in a scene/pixel (Manolakis et al., 2003). The so-called iterative unmixing algorithms, including multiple endmember spectral mixture analysis (MESMA) (Roberts et al., 1998), iterative spectral unmixing (ISU) (van der Meer, 1999), iterative spectral mixture analysis (ISMA) (Rogge et al., 2006), multiple-endmember linear spectral unmixing model (MELSUM) (Combe et al., 2008), and endmember bundles (EB) (Bateson et al., 2000) have been developed to account for pixel-scale variability of endmember types and numbers (step ii). A recent algorithm of this kind is simulated annealing (SA) (Debba et al., 2006), which involves normalizing a random combination of initial endmember vectors and calculating the Euclidian distance between them and the target vector in an iterated way (Debba et al., 2006; Penn, 2002). The endmember variability topic is fully reviewed in (Somers et al., 2011). Classically, the optimization criteria in iterative (and fixed) algorithms have been the RMS error minimization (Roberts et al., 1998), decline in the rate of RMS (Rogge et al., 2006), anisotropy of RMS (van der Meer, 1999), or X^2 residual (Combe et al., 2008).

Although the iterative unmixing techniques attempt to give a reliable estimate of the contributing materials to a pixel's spectrum, the accuracy of the estimated abundances under linear assumption may not be assured (Keshava and Mustard, 2002; Mustard and Sunshine, 1999). The common solution has been the incorporation of nonlinear unmixing models and methods (Dobigeon et al., 2014; Keshava and Mustard, 2002). These family of algorithms benefit from nonlinear functions like neural networks, kernel methods, or machine learning approaches in their architecture (Camps-Valls and Bruzzone, 2009; Licciardi and Del Frate, 2011). However, such algorithms rely heavily on simplified assumptions, and largely overlook the physics of intimate mixing (section 3-1-6), and are rather complicated and difficult to implement. The details of important nonlinear unmixing methods, which are now rising in popularity, are given in (Bioucas-Dias et al., 2012; Camps-Valls and Bruzzone, 2009; Dobigeon et al., 2014; Heylen et al., 2014).

The inaccurate estimate of the abundance quantity could be circumvented by plugging a nonlinear inversion method (like those mentioned earlier (Camps-Valls et al., 2012)), or by including nonlinear regression into the end of the linear unmixing chain (Fig. 2). The bias in the estimation of the abundances is known to be induced by "camouflage" between mineral classes, and hence camouflage (CF) correction is proposed (Kuosmanen and Laitinen, 2008). The CF correction involves removing the bias from estimated abundances by a case dependent nonlinear polynomial function, helping to reach to a mean absolute residual error of around 1% for the case of mineral powders (Kuosmanen and Laitinen, 2008).

The reflectance spectra in a pixel can be imported fully and directly into most of the data-driven methods. Nevertheless, sometimes it is beneficial, or crucial, to select or extract specific bands. A "Feature selection" function aims to reduce the data dimensionality, improve the processing, and maximize the output reliability. Procedures like BandMax (www.exelisvis.com/ProductServices/ENVI.aspx), information-theory-based optimal bans sets (Shen and Bassett, 2002), genetic algorithm and SVM (Li et al., 2011), neural networks (Licciardi and Del Frate, 2011), Fuzzy ROC curves (Mitchley et al., 2009), spectral screening (Robila, 2005), and stable zone unmixing (Somers et al., 2010) are specifically designed to serve this need. In contrast, "feature extraction" aims to create a feature subset, by transforming the data into an uncorrelated new space with lower dimensionality and improved signal-to-noise ratio (SNR). Principal component analysis (PCA (Jolliffe, 1986), minimum noise fraction (MNF) (Green et al., 1988), and independent component analysis (ICA) (Comon, 1994) are the widely used feature extraction techniques.

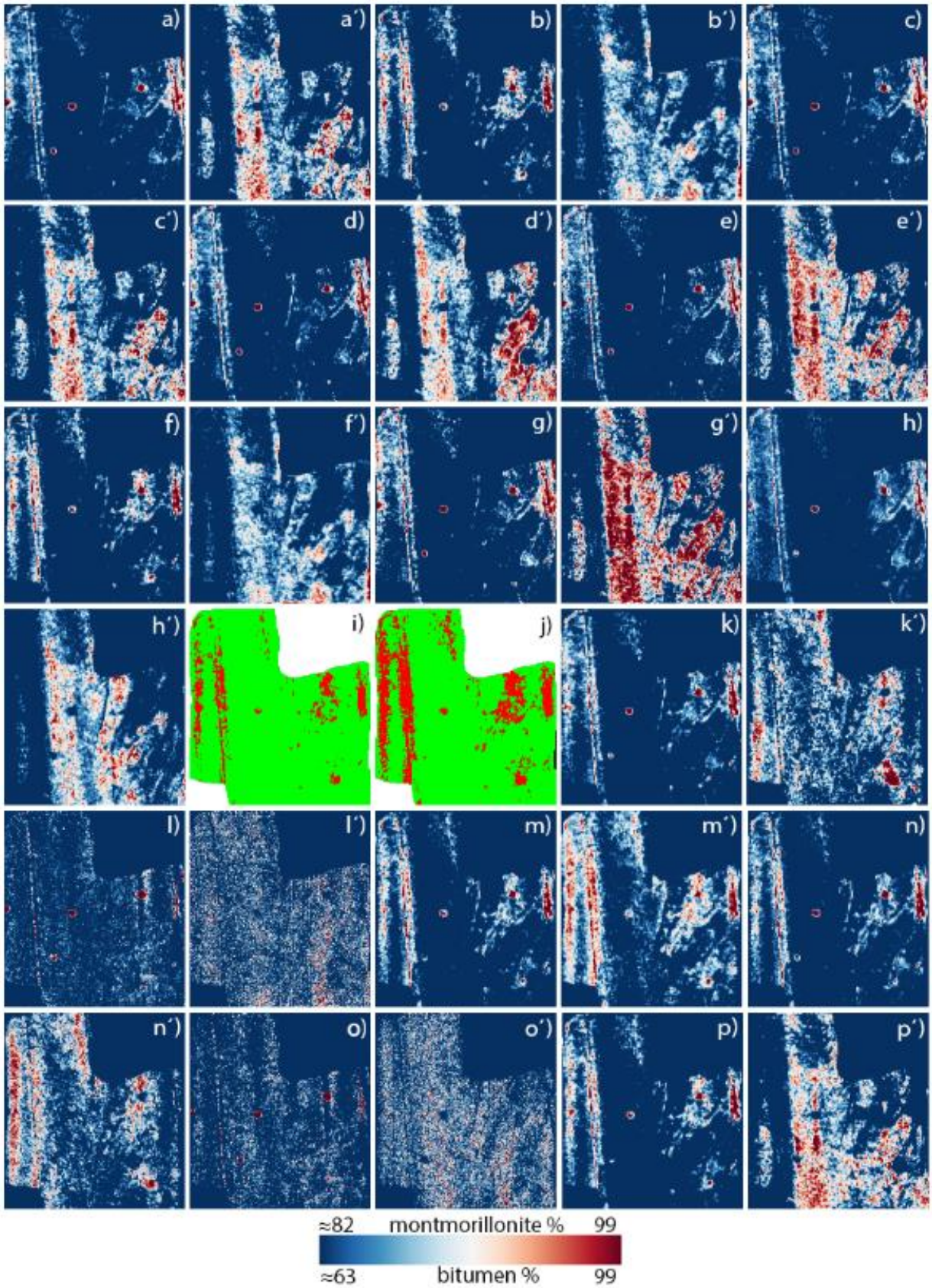

Fig. 5. Examples of mineralogic products of the oil-sand sample extracted from the sisuCHEMA-SWIR hypercube dataset using different techniques comprised of the data-driven approach and the twin endmembers shown in Fig. 4. a) SAM, b) ED, c) NED, d) SCM, e) SID, f) SSM, g) SID x sin (SAM), h) SFF in scale/RMS mode, i) ANN-based classification using hyperbolic activation (green: bitumen, red: montmorillonite), j) SVM-based classification using 2nd order polynomial kernel (green: bitumen, red: montmorillonite), k) OSP, l) MF, m)

CEM, n) TCIMF, o) ACE, p) LSU with sum to 1.0 constraint. The full spectra of the endmembers are used with all methods. In all cases the left figure belongs to montmorillonite and right to bitumen. The output scores or abundance images are all color coded. The target detectors are used without background estimation. The score or abundance thresholds used with color representation is typically between 82-99% for montmorillonite and between 63-99% for bitumen.

3.4. Comparative study of the approaches and their products

The knowledge-based and data-driven approaches derive from different disciplines, with dissimilar assumption, procedures, and algorithmic architecture. Nonetheless, both aim to decompose a multi-, or hyperspectral signal into meaningful, quantitative or qualitative geologic information. Accordingly, it is possible to compare the similarities as well as the differences among them (table 1).

The knowledge-based approach is physically-based and is derived from the concepts of spectroscopy. The techniques in this category can be used in a processing chain without the need for reference data. In contrast, the data-driven approach is mathematically-based and relies directly on reference data for information extraction. It treats each pixel as a ***n***-dimensional vector (with "***n***" being the number of spectral bands) in feature space, and attempts to model the "whole scene" variation by a set of endmembers (Landgrebe, 2003). Since ***n*** is inter-correlated, the dataset should undergo a feature extraction process prior to the processing. On the other hand, the knowledge-based approach strives to describe the variation observed in a "single spectrum" through absorption band modeling in spectral space, and as a preprocessing step, it only possibly demands the continuum to be removed. While the search for the endmembers (their numbers and types) in the data-driven approach is automated, the detection of absorptions (their numbers and positions) in the knowledge-based approach is largely manual and based on user knowledge (table 1); though recently, a number of automated algorithms have been proposed to serve this need (e.g. (Zhouyu et al., 2007)).

Due to their structure, knowledge-based methods are sensitive to data type, meaning that their input has to be in reflectance/emittance unit (except for PCA), whereas data-driven methods can be conducted on both radiance and reflectance/emittance data (except for SFF). Because the knowledge-based methods merely rely upon spectral space, it is feasible to compile regional to continental-scale mineral maps on their basis (e.g. (Cudahy et al., 2008)). The only critical requirements are an accurate atmospheric correction and seamless mosaicking (e.g. (Gao et al., 2009; Granahan and Sweet, 2001)). On the other hand, because data-driven methods are scene dependent (for either endmember selection or spectral mapping), the challenges posed for large-scale applications are more difficult to transcend. The reliance of the former techniques upon spectral space, however, make them increasingly vulnerable to noise. Hence, a spectral smoothing step may be required to be incorporated into the process (section 3-1-1). In comparison, data-driven techniques are less sensitive to noise, but where needed, they can make use of spatial, spectral, or frequency-domain filters (Monteiro et al., 2009; Schott, 2006; Schowengerdt, 2007).

Algorithms belonging to full unmixing and least square groups benefit from an embedded error metric. Training-, and learning-based classifiers use ground truth data for overall accuracy assessment (Tso and Mather, 2009). In contrast, knowledge-based methods completely lack such metrics (except for MGM and FP). Occasionally, however, the RMS error or R^2 of the regression

performed for validation or calibration purposes can be used as an indirect error metric (e.g. (Swayze et al., 2014)).

Table 1. A comparison between different aspects of the knowledge-based and data-driven approaches.

Parameter	Approach	
	Knowledge-based	**Data-driven**
Domain	Spectral space (spectral/absorption modeling)	Feature space (scene modeling)
Background theory	Spectroscopy (physically-based)	Statistical, geometrical, etc. (mathematically-based)
Reference data	Not required	Required
Input data	Reflectance/emittance Rarely radiance	Reflectance/emittance Radiance
Applicability	MSI (except for FP and MGM), HSI	HSI, MSI (except for SFF)
Spectral band range	Partial (2-10 bands) or full	Full or partial
Feature selection criteria	Essential/manual	Optional/ interactive, or automated
Treatment with pixels	Soft and hard	Hard and soft
Information unit	Absorption	Endmember/class
Required preparation	Absorption detection (assignment)	Endmember extraction/class definition
Preprocessing/normalization	CR and DA (over the spectrum)	MNF, PCA, ICA (over the scene)
Products	Thematic map (occurrence), abundance image, composition, and crystallinity maps	Thematic map, abundance image
Main capability	Discrimination, classification, identification, quantification	Detection, classification, quantification
Abundance metric	Depth (area?) of absorption	Fraction of endmember
Abundance quantity	Relative	Absolute
Algorithm type	Mainly linear (except for physical models)	Linear/nonlinear
Sensitivity to noise	Very sensitive	Sensitive
Noise suppression	Spectral filters	Spatial/spectral filters
Error metric	Not available (except for MGM and FP)	RMS error/ confusion matrix
Processing time	Short	Long
Pros and cons	- Simple, easy to execute and available - Diverse products - Transferable between scales - Less robust (esp. with MSI data)	- More mature and diverse - More robust - Complicated and time consuming - Unavailable

The data-driven approach properly implements both soft and hard classifications and has distinctive algorithms for each task (Fig. 2). The major product of this approach is either a “thematic map” or an “abundance image” (Schott, 2006; Schowengerdt, 2007). The thematic map in geologic remote sensing includes a classified lithology/mineralogy map obtained chiefly from the statistics-based group (Fig. 5). The abundance image, which represents the areal fraction of an endmember in a pixel, is obtained from mixture-based category (Keshava and Mustard, 2002). In contrast, the knowledge-based approach is in essence a soft classifier, because it pinpoints very specific spectral region(s) for identification and/or quantification (Fig. 3), hence raising the possibility for multiple mineral mapping using a single spectrum (e.g. (Cudahy et al., 2008)). Nevertheless, where the absorption bands are overlapping, the knowledge-based algorithms are inevitably switched to a hard classifier (the exception is MGM). Generally, the conversion of knowledge-based methods into a hard classifier is challenging, because every case needs its own threshold, which is not always available, nor universal (Mars and Rowan, 2006). The abundance image in this approach is achieved by

calculating the absorption band depth (Haest et al., 2012). From the perspective of outcomes, the knowledge-based approach is able to produce both abundance image and classified maps, but the latter is by no means comparable to the products of the data-driven approach. Although the data-driven approach is able to account for endmember variability between pixels, individual endmembers are still treated as spectrally rigid quantities. Nevertheless, the majority of the geologic materials (rocks and minerals) are chemically and therefore spectrally variable. The knowledge-based approach accounts for this variability and provides two parametric maps known as “composition” and “crystallinity” that respectively are indicators of chemical variation and crystal order of the minerals (Clark et al., 1993; Clénet et al., 2011; Cudahy et al., 2008).

The similarity-based methods (Fig. 2) are used to search for spectra of interest in an image data or within a spectral library. In contrast, knowledge-based methods (especially absorption modelers) are employed to search for a specific feature(s) in an image or in a suite of spectra. Their typical outcomes are discriminated alteration index or detected mineral index. Where the whole spectrum is involved, this approach (e.g. by an expert system) enables mineral (material) identification as well (Clark et al., 2003; Kruse et al., 1993b); a process that is not at all straightforward for data-driven methods.

The majority of the algorithm noted so far may have emerged as VNIR-SWIR data processing tools, but they can handle LWIR data as well. A case in point is the application of BR (Feng et al., 2006), least square (Feng et al., 2006), CR (Cudahy et al., 2009), FP (Cudahy et al., 2009), PLSR (Hecker, 2012), ANN (Ninomiya, 1995), LSU (Ramsey and Christensen, 1998), MF (Funk et al., 2001), MTMF (Kruse, 2015), MESMA (Funk et al., 2001), and WA (Feng et al., 2011) routines to radiance or emittance thermal datasets.

In general terms, the data-driven approach is “deductive”, since it looks on the spectra as a whole to find out the contents of every pixel, while the knowledge-based approach is “inductive” by resolving the contents of an individual pixel to understand the whole. Overall, the knowledge-based approach has the advantages of being simple, straightforward, easily attributable to mineralogy/geochemistry needs, transferable between different scales and cases, and more importantly, available to all. Its major drawback is that it cannot handle overlapping/mixing absorption features, and is not robust enough, specifically with MSI data (Fig. 3). In comparison, the data-driven approach is more mature and robust, but is typically complicated; it demands more computing power, has fewer outcomes (Fig. 5), and is out of access for many users. An initial stage is needed to set the reference data, or train the algorithm, which makes the data-driven techniques more time consuming (table 1). The best solution for geological application however, may come from the hybridization of these approaches.

3.5. Hybrid methods

Given the strength and limitations of the individual spectral processing algorithms, it would be favorable to combine (crossbreed) multiple perspectives to yield advanced algorithms. In the very simple form, methods like PCA of ratios, derivative ratio, or ratio classification have been suggested (Fraser, 1991; Philpot, 1991; Rud et al., 2006). In addition, there are many examples of mathematical hybridization within similarity-based methods including SSM (Eq. 2), MF/SAM ratio, and SID-SAM techniques (Du et al., 2004; Granahan and Sweet, 2001;

www.exelisvis.com/ProductServices/ENVI.aspx). Likewise, the joint use of CR procedure to these similarity measures, in order to enhance their performance, can be deemed as a hybridization action. A case in point is the attachment of CR to SAM, CCSM, ED, and SID routines (Bue et al., 2010; Kruse et al., 1993a; van der Meer, 2000). The last uses a weighted combination of continuum intact (CI) and CR for the spectral measurement. The continuum-removed spectra are also used with MESMA approach and are shown to improve its classification performance (Youngentob et al., 2011). The SFF technique discussed in section 3-2-2 is in reality a hybrid method that combines user knowledge of the feature(s) and CR prior to the least-square fitting (Clark et al., 1990).

CR is employed to level out or normalize the hyperspectral signal for cross-comparison (Clark and Roush, 1984); however owing to its quotient nature, its incorporation into unmixing procedure is believed to be problematic (Rivard et al., 2008). To scale up the endmembers during unmixing, some prefer to add in a shadow component (Keshava and Mustard, 2002), but this only cancels out the linear effect of illumination and cannot account for the continuum which in essence, is nonlinear (section 3-1-2-1). So far, only an unmixing-like routine called canonical variates analysis (CVA) has incorporated the CR procedure into its structure. The CVA simultaneously estimates both the continuum and the mineral abundances, and is reported to estimate the abundance of a mineral with 15% average accuracy (Berman et al., 1999). Other spectral normalization procedures prior to unmixing are the division of a (multispectral) spectra to its mean (called mean normalization) (Berman et al., 2004), standardization using the mean and standard deviation (García-Haro et al., 2005), and derivative unmixing (Zhang et al., 2004). Derivative spectral unmixing (DSU) is shown to quantitatively estimate the fraction of an endmember, in spite of having only a general knowledge of the spectral shapes of the remaining endmembers (Zhang et al., 2004). The derivative is also beneficial for estimation of abundances, or classification of spectrally similar, and mathematically correlated endmembers, as is the case with SA algorithm (Debba et al., 2006; Monteiro et al., 2009).

The fact that some spectral variation in the deterministic LMM is intrinsically stochastic has helped the stochastic mixing model (SMM) to emerge as another line of hybridization. The SMM combines the stochastic property of endmembers with LMM to capture the variations that otherwise could not be described by standard models (Eismann and Stein, 2007). Likewise, MTMF and kernel methods could be regarded as hybrids of deterministic and stochastic perspectives (Boardman and Kruse, 2011; Camps-Valls and Bruzzone, 2009).

An interesting hybridization between LMM and similarity measures has given rise to the optimized cross correlation mixture (OCCM) analysis (Coulter, 2006). Its basic philosophy is to match the entire shape of each pixel spectra to a linearly synthesized mixture of endmember spectra using the SCM method. The method tries to keep the maximum cross-correlation close to 1, and simultaneously optimizes the endmember weights in an iterative way (Coulter, 2006). The main difference between LSU and OCCM is that the former attempts to minimize the error of inversion, while the latter tries to maximize the “goodness of fit” through fraction optimization (Coulter, 2006; Keshava and Mustard, 2002). Therefore, it is more capable of tackling the issue of correlated endmembers. However, up to now, the performance of this technique has not been tested against iterative unmixing algorithms. The idea behind OCCM may be extended to construct other “similarity-based unmixing algorithms” or other optimization criteria.

The most recent and promising line of hybridization comes from WA. The wavelets can be added to boost the performance of other algorithms (Fig 6). For instance, instead of directly mapping the spectral feature, Bruce et al. incorporated wavelet coefficients' scalar energies as features into an automated statistical classification system (ML) for spectral mapping (Bruce et al., 2001). A similar method for comparison of spectral angles known as "Wavanglet" was also proposed, which defines a more effective way of measuring the spectral angles between the reference and test spectrum in wavelet domain (Schmidt et al., 2007). Furthermore, the linear additive character of the wavelet domain is used to facilitate the linear unmixing and abundance estimation (Rivard et al., 2008), which is exemplified in Fig. 6. Here the wavelet is combined with other methodologies to yield results that are more accurate. For example, the composition or abundance of montmorillonite (Fig. 6b, 6f) is better quantified relative to analogous maps in Fig 3s, 5p. The CWT is employed to minimize the combined influence of variable mineral grain size, illumination, and surface roughness (continuum) on the spectra, and increase the SAM classification accuracy of drill cores (Feng et al., 2011). Eventually, the derivative are computed more efficiently in the transformed wavelet domain (Bruce and Jiang, 2001).

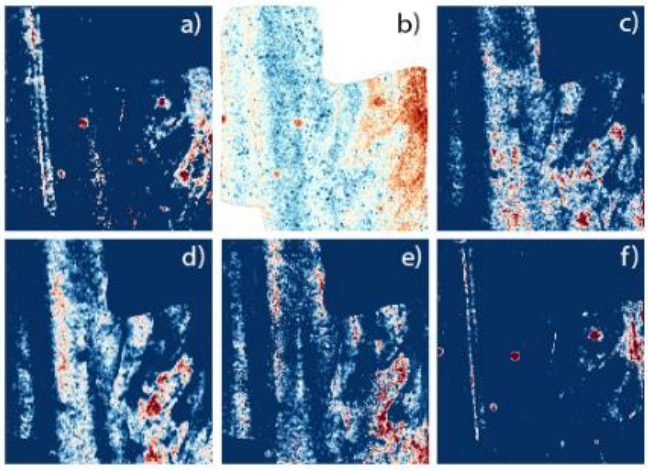

Fig. 6. Examples of spectral products using the wavelet-transformed hypercube data of the oil-sand sample. Three out of ten components (scales) most relevant to absorption features in the wavelet domain is retained, summed, and then used during the processing. **a)** The ratio between R_{2213} and R_{1914} is used as a measure of montmorillonite crystallinity masked by its abundance image (Blue: poor-ordered, red: well-ordered). **b)** Montmorillonite composition calculated using the Gaussian curve fitting between 2151-2244 nm after spline interpolation (blue: 2206 nm, red: 2212 nm). This product is equivalent to Fig. 3s. **c)** Total area of the 2300 nm absorption calculated using the 2nd order FP between 2282-2331 nm. **d)** NED similarity-based measure of bitumen endmember. **e)** and **f)** LSU of bitumen and montmorillonite calculated using the image-extracted endmembers. The contrast stretch is the same as used in Fig. 3 and Fig. 5.

While the above-mentioned hybrid methods are sensible, we do believe that the most advanced and rigorous processing methodologies are achieved when two completely different perspectives (i.e. physically- and mathematically-based approach) inter-breed. On this basis, we have conceived two major hybrid species: (i) those that incorporate the spectroscopy knowledge and mixing models to build enriched spectral libraries; and (ii) those that supplement *a priori* geological knowledge with mixture theory, and vice versa.

The core of the first species is a highly enriched spectral library and a decision making mechanism to compare with the image (test) spectra. One of the first of this kind was developed by Kruse et al. (Kruse and Lefkoff, 1993; Kruse et al., 1993b) and Kruse (Kruse, 2008). It benefits from a lab or image extracted spectral library and a set of spectroscopic-based rules (as described in section 3-1-3) to

implement the identification. The most sophisticated and successful form of mineral identification system based on reflectance spectra is indeed the Tetracorder package developed by the US Geological Survey (Clark et al., 2003). In this system, the spectral library is enriched by binary and ternary mineral mixtures (in both linear (Clark et al., 2003), and nonlinear (Dalton et al., 2004) scenarios), vegetation, etc.; and then grouped based upon spectroscopic similarities. Two metrics, the goodness of fit (R^2), and the band depth (D), are calculated on the basis of continuum removed image and library spectra (section 3-2-2), and are then used within an intelligent expert system decision-making framework to identify and map the geologic materials from AVIRIS hyperspectral data. Its superior performance is demonstrated in (Clark et al., 2003; Dalton et al., 2004; Swayze et al., 2014). A modified version of the early Tetracorder with graphical user interface (GUI) designation known as material identification and characterization algorithm (MICA) is described in (Kokaly, 2012).

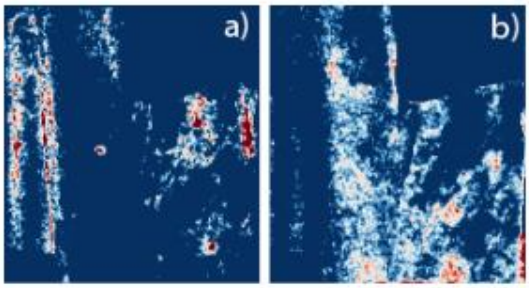

Fig. 7. Examples of linearly simulated spectral library and similarity-based measures for abundance estimation. **a)** Abundance image of the montmorillonite. The correlation between this mineral map and that shown in Fig. 5p is as high as 0.987. **b)** Abundance image of the bitumen. The contrast stretch applied here is similar to Fig. 5.

To rigorously define the mixture amounts in a field-adjusted HyMap data, Roy et al. (Roy et al., 2009) have developed a simulated spectral library consisting of three and four rock components to represent mantle and crustal sequences, respectively. The increment used is 0.1 for every pair of spectra. Both the simulated spectral library and image pixels are normalized for continuum by dividing them by their Gaussian low-pass trend. Finally, every pixel is compared to the library by the SD measure to find out its lithologic content and quantity (Roy et al., 2009). A linearly simulated mixture between bitumen and montmorillonite is compared to image spectra by a similarity measure in Fig. 7.

The state-of-the-art methodology belonging to the first species is the hypersensitive mineral identification method (HMIM), developed by ERSDAC, for the analysis of multispectral ASTER satellite data (Sanga and Tachikawa, 2006). The HMIM comprises a very sophisticated spectral library yielded by simulating various mixtures of 13 minerals for every 10% abundance using the iso-grain model cited above. The model takes into consideration not only the abundances, but also the refractive and reflectance indices, grain size, and the scattering coefficients (ERSDAC, 2006). To avoid complication, the simulation was carried out in two different batches namely acidic and phyllic-propylitic. The image pixel was then compared to the library by similarity measures and the content of the top five simulators were averaged and reported as minerals' abundance in the relevant pixel. This package combines the knowledge of alteration mineralogy with nonlinear scattering theories, and uses similarity metrics to search for the best answers to each pixel (ERSDAC, 2006; Sanga and Tachikawa, 2006).

In many geological applications (e.g. mineral exploration and rock type classification), the type of target minerals and their associations are usually predictable (Sabine, 1999; Sillitoe and Thompson, 2006; Thompson et al., 1999). This *a priori* knowledge can be involved within the spectral processing chain. In other words, unmixing can be performed in the geologic context. Despite the possible theoretical framework, a hybrid method of the second species has not yet been developed. However, the usefulness of such *a priori* knowledge has been sparsely explored for mapping hydrothermal systems (e.g. by methodologies like MESMA (Bedini et al., 2008), HMIM (Sanga and Tachikawa, 2006), OCCM (Coulter, 2006)) and lithologic variations (e.g. (Gilmore et al., 2008; Roy et al., 2009)).

Other hybrid methods worthy noting are those unmixing procedures with roots in physical models, and the ability to fully unmix the spectra of a pixel, or pixels of a scene by combined linear and/or nonlinear models (e.g. (Close, 2011)).

3.6. Discussion

Geologists find spectroscopy appealing because it is fast, cost-effective, non-destructive, and more importantly, multiscale. It is a method capable to provide data and information from proximal to distal sensing. Traditionally, spectral data have been used to detect surficial alteration minerals, but now this versatile tool is used to quantify a diverse range of chemical and physical parameters related to a wealth of Earth Sciences disciplines, as summarized in table 2. These parameters are extracted directly, or inferred indirectly from the four basic products of spectral analysis (table 1) in a qualitative or semi-quantitative way. The spectral solution has been widely acknowledged by the mining industry and to a lesser extent the energy sector, and the interest in spectral technology for geological applications is steadily growing (van der Meer et al., 2012).

Table 2. The common geoscience parameters estimated and quantified by spectral data. The spectral solution is based on the basic spectral products reported in table 1. The "spectral range" indicates the appropriate range(s) for information extraction.

Parameter	Spectral solution	Spectral range	References
Material detection (Minerals, rocks, and other compounds)	Thematic map	SWIR-LWIR-VNIR	(Kozak et al., 2004; Sgavetti et al., 2006; Thompson et al., 1999)
Spectral assaying (elements and other compounds)	Abundance image Composition map	VNIR-SWIR-LWIR	(Dai et al., 2013; Lyder et al., 2010; Murphy and Monteiro, 2013)
Whole rock geochemistry	Abundance image	LWIR-VNIR-SWIR	(Walter and Salisbury, 1989)
Temperature, pressure, and metamorphic grade	Thematic map (mineral occurrence) Composition map Crystallinity map	SWIR-LWIR	(Duke, 1994)
Weathering	Crystallinity map Composition map Thematic map (mineral occurrence)	VNIR-SWIR	(Cudahy et al., 2008)
Fluid pathway	Composition map	SWIR-LWIR	(Cudahy et al., 2008; Herrmann et al., 2001)
Ph/Eh	Thematic map (mineral occurrence)	VNIR	(Swayze et al., 2000)

Spectral mapping techniques available to geologic remote sensing largely aim to retrieve information about mineralogic, lithologic, and to a lesser extent, chemical content of a target (table 1

& 2). The mineralogic pattern (2D/3D) is a key factor in understanding the geological processes in general, and mineral systems in particular (e.g. (Holliday and Cooke, 2007; Sillitoe and Thompson, 2006; van der Meer et al., 2012; Wyborn et al., 1994)). Owing to its fundamental role in resource exploration and recalling its close relationships to spectroscopic concepts, mineralogy has been the focus of many studies. Nevertheless, there are quite a few accounts of the absolute accuracy and precision of the processes used for abundance estimation. The physical models are reported to estimate the abundance to within 5-10% accuracy (Mustard and Pieters, 1989; Poulet and Erard, 2004; Shipman and Adams, 1987), whereas for hybrid methods like HMIM and CVA this figure is within the 10-15% range (Berman et al., 1999; ERSDAC, 2006). Although values as precise as 2% are reported for CBD (Kruger et al., 1998), there have been cases of fractional error by as much as 30% absolute or more (Keshava and Mustard, 2002; Kuosmanen and Laitinen, 2008). From the spectroscopic viewpoint, minerals forming less than 1% in abundance in a mixture have been detected spectrally (Pontual et al., 2008a); though, in general, minerals that encompass less than 5% in a rock are usually difficult to identify. As a basic rule, spectroscopic detection limits for bright and dark minerals are considered to be 10 and 20%, respectively (Thompson et al., 1999). Accordingly, the accuracy of estimation and the detection limit is dependent not only on the algorithm in use, but also on the type of target material and its spectral contrast. The accuracy of abundance estimation is as well affected by the sensing approach (proximal vs. distal), the sensor technology, and the imaging scale.

At present, the detection limit of the current sensing approach for individual minerals is largely unknown and only a few algorithms yield an estimation of the abundance accuracy. We believe that in the absence of real and independent ground truth data, the described abundances in the literature are at most a relative quantity, as with the case study shown here. To obtain absolute quantities for the abundance, many have used a training or correction stage, respectively, at the beginning or at the end of the processing chain, using ancillary data (e.g. (Kuosmanen and Laitinen, 2008; Lyder et al., 2010)).

Regarding the compositional variations present in several mineral species (including white mica, chlorite, alunite, amphiboles, epidote, montmorillonite, feldspar, etc. (Cudahy et al., 2009; Duke, 1994; Hecker, 2012; Herrmann et al., 2001; Mustard, 1992; Pontual et al., 2008b; Roache et al., 2011; Swayze et al., 2014; Thompson et al., 1999)), there are quite a few effective and practical methodologies to quantitatively and robustly map them with hyper-, or multispectral data. Such maps have been shown to have great potential for revealing the physicochemistry of minerals not only in hydrothermal systems, but also in metamorphosed, metasomatized, and sedimentary environments (Cudahy et al., 2008; Duke, 1994; Herrmann et al., 2001; Kurz et al., 2012; Laukamp et al., 2011; Sgavetti et al., 2006; Thompson et al., 1999) (table 2). Moreover, despite the partial sensitivity of the spectral technique to mineral crystallinity (Thompson et al., 1999), current solutions for mapping this parameter remotely appear to be very embryonic (Clark et al., 1993; Cudahy et al., 2008).

Since the VNIR reflectance spectra of rare earth elements (REE) are dominantly relevant to isolated ions rather than ligands (Adams, 1965; Hunt, 1977), spectroscopy affords a unique opportunity to directly detect elements. In the scarce literature on the subject, band depth has been the common method used for REE characterization (Dai et al., 2013; Huntington et al., 2012). Other major/minor elements (including the transition metals like Fe, Mn, Cu, etc.) are merely indirectly

assayed. For this purpose, different regression techniques (section 3-2-2) are used to construct a mathematical model by correlating a spectral parameter (i.e. depth, ratio of depths, derivatives, abundance, wavelength, and area of absorption) with other reliable, but costly, geoscience products acquired from independent analytical tools (Cloutis, 1996; Cudahy et al., 2009; Cudahy et al., 2001; Haest et al., 2012; Herrmann et al., 2001; Murphy and Monteiro, 2013; Mustard, 1992; Pontual et al., 2008b; Post and Noble, 1993; Swayze et al., 2014; Tappert et al., 2011; Thompson et al., 1999; Walter and Salisbury, 1989). The degree of substitution of the elements like Al, Si, Mg, Fe, K, etc. in the structure of selected minerals may also be estimated using the compositional maps derived from knowledge-based techniques.

Spectral analysis can effectively deal with mono-mineralic rocks (e.g. limestone) (Combe et al., 2006; Kozak et al., 2004; Kurz et al., 2012), but given the multi-mineralic nature of the many rock types, it is challenging to characterize them using current techniques (Rivard et al., 2009; Sgavetti et al., 2006). For instance, the statistically based classification algorithms (e.g. IK, RF, SVM, and ANN) cannot compete with the outcomes of visual interpretation by an analyst (Cracknell and Reading, 2014; Jie-Lun et al., 2014; Mas and Flores, 2007). This is a gap that may be partly bridged by more advanced methodologies, but a better and effective solution is provided by incorporating the LWIR data (e.g. (Cracknell and Reading, 2014; Feng et al., 2011; Roy et al., 2009)),

While VNIR-SWIR data are valuable to study the alteration mineralogy, LWIR data are able to provide information on the composition of the rocks and rock forming minerals, as they have their fundamental vibrational bands (e.g. Si-O stretching in silica tetrahedral) in this region (Gaffey et al., 1993; Hook et al., 1999; Hunt and Ashley, 1979; Walter and Salisbury, 1989). The LWIR sensing can augment the noted spectral outcomes (table 2) and, at the same time, holds promise to bring new spectral products into existence. Examples are the deduction of mineral orientation (Tappert et al., 2013), rock hardness/crushability (Huntington et al., 2010), and modal mineralogy of the rocks (Hamilton and Christensen, 2000). LWIR data can facilitate the detection of compounds like gaseous hydrocarbons as well (Johnson et al., 2014; Thorpe et al., 2013).

In the self-similar geological environment, spectral mixing is an established fact that should be acknowledged by all the processing methods at all scales. The self-similarity principle, however, does not imply that every mappable phenomenon is shared among the scales. For example, while some products are already shared between proximal and distal sensing techniques (e.g. white mica composition); ore detection is possibly going to remain exclusive to proximal methods. Presently, there are challenges regarding the transferability of a product between scales (e.g., the chemical variation of chlorite) which need to be tackled in the future.

The performance of spectral processing methods is seriously affected by the spectral resolution and intrinsic SNR of the sensor, along with the quality of the atmospheric compensation (Clark and Swayze, 1996; Green et al., 1998). Currently, there exist rigorous correction algorithms based on radiative transfer codes (e.g. (Gao et al., 2009)), but their residual error hampers the detection and identification of materials by hybrid methods like Tetracorder. To attain absolute reflectance, a further refinement step is taken by incorporating ground measurements (Clark et al., 2003). In the case of MSI, however, ground-based or cross-sensor calibration has been the only reliable method for accurate radiance to reflectance conversion (Cudahy et al., 2008; ERSDAC, 2006; Mars and Rowan, 2010).

3.7. Conclusions

Despite important algorithmic developments in recent years, there is still no universal and optimal recipe for remote identification, classification, and quantification of geologic materials. In the past decade, linear spectral unmixing has received a great attention from algorithm developers, but so far, the outcomes have not been utterly convincing. Recently, unmixing has been augmented by incorporating the contextual (spatial) information, or by bringing nonlinear methods into the scene (e.g. (De Jong and van der Meer, 2005; Heylen et al., 2014; Plaza et al., 2009)). The nonlinear unmixing methods are proposed to give more accurate estimates of abundances, whereas the spatial-spectral unmixing is exploited to incorporate the pictorial character of the image. Indeed, we anticipate other types of hybridizations to take form. As discussed in this paper, there are many opportunities and promises in hybridization between spectral and feature domains.

In geological remote sensing, the most complicated mixtures happen between spectrally similar minerals (from the same absorbing species), which unfortunately are associated with each other in real geological environments. In such cases, distinctive absorption features are very near or overlapping and the spectra (endmembers) are highly correlated. This correlation hinders any attempts to identify or discriminate the minerals by routine ways. While FP is promising for those solid solutions that manifest themselves as linear wavelength shifts in the absorption minima, there are few remedies for more complicated overlaps. We anticipate the solution may come from hybrid methods like “similarity-based unmixing”, or “simulation-tuned similarity measurement” algorithms. The successful Tetracorder package is a good example of the latter solution, although the decision-making system of Tetracorder and its successor are still governed by hard rules.

In addition, unlike physical models, current unmixing algorithms are not sensitive to the type of spectra (minerals) they are unmixing. Such knowledge, already available in the spectra of each pixel, can supplement the unmixing procedure to pinpoint the camouflaging spectra, or predict the proper and probable combinations of endmembers present in each pixel, leading subsequently to better abundance estimation. This is what we call “unmixing in the geologic context”, a discipline-oriented hybrid model of the second species. Generally, the delivery of a set of abundance images has been regarded as the final step in the remote sensing processing chain. We do believe this chain is complete only when a sensible interpretation is given to these final maps in the context of the geologic system under survey.

Unlike the unmixing techniques that decompose a spectrum into its constituent endmembers, the MGM experiment has shown that a spectrum can also be decomposed into a continuum, and a collection of absorbing bands as physically meaningful quantities. The few data-driven techniques that incorporated the continuum-removed spectra have shown to give higher overall performance, yet none has adapted to account for the continuum components or decompose the spectrum into absorption bands. The continuum modeling itself is not yet satisfactorily matured, but there are clues that point towards the potentials of wavelet analysis for this aim.

Studies show that within the VNIR-SWIR range, we are at most measuring 50% of the minerals present in a system. The other half needs to be dealt with using LWIR sensing technology (Hook et al., 1999; Huntington et al., 2010). Up to now, LWIR (and partly VNIR) data have been processed and interpreted in isolation; however there are clues that underline the significance of simultaneous

"multiple wavelength processing" (e.g. (Chen et al., 2007; Huntington et al., 2010; Kruse, 2015)). The geologic remote sensing community is only beginning to understand and explore the potentials of this spectral range and the merits of integrated processing.

Regarding the rock type classification (in both close-, and far-ranges), we have not gone far from traditional "image classification", which is hardly comparable to the outcomes of visual techniques, and a system similar to Tetracorder for automated rock type identification is still absent. Given the richness of contextual information embedded in rocks, ores, veinlets, and alteration facies as texture or zoning, it is conceivable to tap into this information using spatial-spectral hybrid techniques. Such a system may inherit its character from hybrid techniques specifically adopted to analyze the spatial-spatial pattern, as well as multiple wavelength spectral ranges.

As discussed in this paper, the use of ancillary data, which are essential for accurate spectral quantification, makes the processing techniques case specific and nontransferable. As concluded in (van der Meer et al., 2012), this 'hampers automating processing chains and standardized (qualitative or quantitative) products'.

Thus far, the processing routines have been confined to sensor frames (scene/strip), which are not a match for orderly quadrangles used for standard geoscience maps. To have similar standard end products for the earth surface from orbital sensors, the spectral processing techniques are required to be applied to seamlessly mosaicked reflectance data. It means that the next generation of satellite sensors must be equipped with proper spectral bands to compensate for atmospheric effects.

Acknowledgments

We are indebted to Professor Celio Pasquini for kindly sharing unlimited time of the sisuCHEMA instrument based on the Chemistry Department at UNICAMP. We are thankful to Thomas Cudahy (Commonwealth Scientific and Industrial Research Organization; CSIRO, Australia) and Benoit Rivard (University of Alberta, Canada) for processing our dataset, respectively, by FP and wavelet techniques. We acknowledge the assistance of Samuel Murphy for IDL (interactive data language) programming and R. Perobelli for XRD analysis. We also thank the anonymous reviewers for thorough comments and helpful suggestions. The authors thank CAPES and CNPq for the research grants.

References

Adams, J.B., Smith, M.O., Johnson, P.E., 1986. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site. Journal of Geophysical Research 91, 8098-8112.

Adams, J.W., 1965. The visible region absorption spectra of rare-earth minerals. The American Mineralogist 50, 356-366.

Agar, B., Coulter, D.W., 2007. Remote sensing for mineral exploration – A decade perspective 1997-2007, in: Milkereit, B. (Ed.), Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration, pp. 109-136.

Ahlberg, J., Renhorn, I., 2004. Multi- and hyperspectral target and anomaly detection. Technical Report FOI-R-1526-SE, Swedish Defence Research Agency,, p. 41.

Bardossy, A., Samaniego, L., 2002. Fuzzy rule-based classification of remotely sensed imagery. Geoscience and Remote Sensing, IEEE Transactions on 40, 362-374.

Bateson, C.A., Asner, G.P., Wessman, C.A., 2000. Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis. Geoscience and Remote Sensing, IEEE Transactions on 38, 1083-1094.

Bedell, R., Crosta, A.P., Grunsky, E., 2009. Remote sensing and spectral geology. Reviews in Economic Geology 16.

Bedini, E., van der Meer, F., van Ruitenbeek, F., 2008. Use of HyMap imaging spectrometer data to map mineralogy in the Rodalquilar caldera, southeast Spain. International Journal of Remote Sensing 30, 327-348.

Bell, J., 2008. The Martian surface: Composition, mineralogy, and physical properties. Cambridge University Press, New York.
Berman, M., Bischof, L., Huntington, J., 1999. Algorithms and software for the automated identification of minerals using field spectra or hyperspectral imagery, Proceedings of 13th International Conference on Applied Geologic Remote Sensing, Vancouver, British Columbia, p. 222−232.
Berman, M., Kiiveri, H., Lagerstrom, R., Ernst, A., Dunne, R., Huntington, J.F., 2004. ICE: a statistical approach to identifying endmembers in hyperspectral images. Geoscience and Remote Sensing, IEEE Transactions on 42, 2085-2095.
Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., Chanussot, J., 2012. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5, 354−379.
Boardman, J.W., 1989. Inversion of imaging spectrometry data using singular value decomposition, Proceedings, IGARSS'89, 12th Canadian Symposium on Remote Sensing, pp. 2069-2072.
Boardman, J.W., Kruse, F.A., 2011. Analysis of Imaging Spectrometer Data Using N-Dimensional Geometry and a Mixture-Tuned Matched Filtering Approach. Geoscience and Remote Sensing, IEEE Transactions on 49, 4138-4152.
Boardman, J.W., Kruse, F.A., Green, R.O., 1995. Mapping target signatures via partial unmixing of AVIRIS data, Summaries, Fifth JPL Airborne Earth Science Workshop. JPL Publication 95-1, pp. 23-26.
Brown, A.J., 2006. Spectral curve fitting for automatic hyperspectral data analysis. Geoscience and Remote Sensing, IEEE Transactions on 44, 1601-1608.
Brown, M.A., 2010. Image processing manual for hydrothermal alteration mapping using HyMap data. Geological Society of Iran, Tehran, unpublished internal report,, p. 32.
Bruce, L.M., Jiang, L., 2001. Wavelets for computationally efficient hyperspectral derivative analysis. Geoscience and Remote Sensing, IEEE Transactions on 39, 1540-1546.
Bruce, L.M., Morgan, C., Larsen, S., 2001. Automated detection of subpixel hyperspectral targets with continuous and discrete wavelet transforms. Geoscience and Remote Sensing, IEEE Transactions on 39, 2217-2226.
Bue, B.D., Merényi, E., Csathó, B., 2010. Automated labeling of materials in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 48, 4059–4070.
Camps-Valls, G., Bruzzone, L., 2009. Kernel methods for remote sensing data analysis. John Wiley & Sons, Singapore.
Camps-Valls, G., Tuia, D., Gómez-Chova, L., Jiménez, S., Malo, J., 2012. Remote sensing image processing. Morgan & Claypool publishers.
Carvalho Junior, O.A., Menezes, P.R., 2000. Spectral Correlation Mapper (SCM): an improvement on the Spectral Angle Mapper (SAM), Proceedings of the tenth JPL Airborne Earth Science Workshop. JPL Publication 00-18, pp. 65-74.
Chang, C.-I., Liu, J.-M., Chieu, B.-C., Ren, H., Wang, C.-M., Lo, C.-S., Chung, P.-C., Yang, C.-W., Ma, D.-J., 2000. Generalized constrained energy minimization approach to subpixel target detection for multispectral imagery. Optical Engineering 39, 1275-1281.
Chang, C.I., 2003. Hyperspectral imaging: techniques for spectral detection and classification. Kluwer Academic Publishers, New York.
Chang, C.I., 2007. Hyperspectral data exploitation: theory and application. John Wiley & Sons.
Chang, C.I., 2013. Hyperspectral data processing: Algorithm design and analysis. John Wiley & Sons, Hoboken, NJ.
Chein, I.C., Qian, D., 2004. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. Geoscience and Remote Sensing, IEEE Transactions on 42, 608-619.
Chen, X., Warner, T.A., Campagna, D.J., 2007. Integrating visible, near-infrared and short-wave infrared hyperspectral and multispectral thermal imagery for geological mapping at Cuprite, Nevada. Remote Sensing of Environment 110, 344-356.
Clark, R.N., 1999. Spectroscopy of rocks and minerals and principles of spectroscopy, in: Rencz, A.N. (Ed.), Remote sensing for the earth sciences, 3rd ed. John Wiley & Sons.
Clark, R.N., Gallagher, A.J., Swayze, G.A., 1990. Material absorption band depth mapping of imaging spectrometer data using the complete band shape least-squares algorithm simultaneously fit to multiple spectral features from multiple materials, Proceedings of the third Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) workshop. JPL Publication 90-54, p. 176−186.
Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. Journal of Geophysical Research: Solid Earth 89, 6329-6340.
Clark, R.N., Swayze, G.A., 1996. Evolution in imaging spectroscopy analysis and sensor signal-to-noise: An examination of how far we have come, in: Green, R.O. (Ed.), Summaries of the Sixth Annual JPL Airborne Earth Science Workshop. JPL Publication, p. 5.
Clark, R.N., Swayze, G.A., Eric Livo, E., Kokaly, R.F., Sutley, S.J., Dalton, J.B., McDougal, R.R., Gent, C.A., 2003. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. Journal of Geophysical Research 108.
Clark, R.N., Swayze, G.A., Gallagher, A., 1993. Mapping minerals with imaging spectroscopy. U.S. Geological Survey, Office of Mineral Resources, 141–150.
Clenet, H., Pinet, P., Ceuleneer, G., Daydou, Y., Heuripeau, F., Rosemberg, C., Bibring, J.-P., Bellucci, G., Altieri, F., Gondet, B., 2013. A systematic mapping procedure based on the Modified Gaussian Model to characterize magmatic units from olivine/pyroxenes mixtures: Application to the Syrtis Major volcanic shield on Mars. Journal of Geophysical Research: Planets 118, 1632-1655.

Clénet, H., Pinet, P., Daydou, Y., Heuripeau, F., Rosemberg, C., Baratoux, D., Chevrel, S., 2011. A new systematic approach using the Modified Gaussian Model: Insight for the characterization of chemical composition of olivines, pyroxenes and olivine–pyroxene mixtures. Icarus 213, 404-422.
Close, R., 2011. Endmember and Proportion Estimation Using Physics-Based Macroscopic and Microscopic Mixture Models. University of Florida, Gainesville.
Cloutis, E.A., 1996. Review Article Hyperspectral geological remote sensing: evaluation of analytical techniques. International Journal of Remote Sensing 17, 2215-2242.
Combe, J.-P., Launeau, P., Carrère, V., Despan, D., Méléder, V., Barillé, L., Sotin, C., 2005. Mapping microphytobenthos biomass by non-linear inversion of visible-infrared hyperspectral images. Remote Sensing of Environment 98, 371-387.
Combe, J.P., Launeau, P., Pinet, P., Despan, D., Harris, E., Ceuleneer, G., Sotin, C., 2006. Mapping of an ophiolite complex by high-resolution visible-infrared spectrometry. Geochemistry Geophysics Geosystems 7, 1-11.
Combe, J.P., Le Mouélic, S., Sotin, C., Gendrin, A., Mustard, J.F., Le Deit, L., Launeau, P., Bibring, J.P., Gondet, B., Langevin, Y., Pinet, P., 2008. Analysis of OMEGA/Mars Express data hyperspectral data using a Multiple-Endmember Linear Spectral Unmixing Model (MELSUM): Methodology and first results. Planetary and Space Science 56, 951-975.
Comon, P., 1994. Independent component analysis: A new concept? Signal Processing 36, 287–314.
Coulter, D.W., 2006. Remote sensing analysis of alteration mineralogy associated with natural acid drainage in the Grizzly Peak Caldera, Sawatch Range, Colorado. Colorado School of Mines, Golden, Colorado.
Cracknell, M.J., Reading, A.M., 2014. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Computers & Geosciences 63, 22-33.
Crosta, A.P., De Souza Filho, C.R., Azevedo, F., Brodie, C., 2003. Targeting key alteration minerals in epithermal deposits in Patagonia, Argentina, using ASTER imagery and principal component analysis. International Journal of Remote Sensing 24, 4233–4240.
Crosta, A.P., McMoore, J., 1989. Enhancement of Landsat Thematic Mapper imagery for residual soil mapping in SW Minas Gerais State Brazil: a prospecting case history in greenstone belt terrain, Seventh Thematic Conference on Remote Sensing for Exploration Geology. ERIM, Calgary, Alberta, Canada, pp. 1173–1187.
Crowley, J.K., Brickey, D.W., Rowan, L.C., 1989. Airborne imaging spectrometer data of the Ruby mountains, Montana: mineral discrimination using relative absorption band-depth images. Remote Sensing of Environment 29, 121-134.
Cudahy, T.J., Hewson, R., Caccetta, M., Roache, A., Whitbourn, L., Connor, P., Coward, D., Mason, P., Yang, K., Huntington, J., Quigley, M., 2009. Drill core logging of Plagioclase feldspar composition and other minerals associated with Archean gold mineralization at Kambalda, Western Australia, using a bidirectional thermal infrared reflectance system. Reviews in Economic Geology 16, 223–235.
Cudahy, T.J., Jones, M., Thomas, M., Laukamp, C., Caccetta, M., Hewson, R., Rodger, A., Verrall, M., 2008. Next generation mineral mapping: Queensland Airborne HyMap and satellite ASTER Surveys. CSIRO Exploration & Mining Report P2007 / 364, 152.
Cudahy, T.J., Ramanaidou, E.R., 1997. Measurement of the hematite:goethite ratio using field visible and near-infrared reflectance spectrometry in channel iron deposits, Western Australia. Australian Journal of Earth Sciences 44, 411-420.
Cudahy, T.J., Wilson, J., Hewson, R., Linton, P., Harris, P., Sears, M., Okada, K., Hackwell, J.A., 2001. Mapping variations in plagioclase feldspar mineralogy using airborne hyperspectral VNIR-SWIR-TIR imaging data, Proceedings of Geoscience and Remote Sensing Conference, IEEE International, Sydney, Australia, p. 4.
Dai, J., Wang, D., Wang, R., Chen, Z., 2013. Quantitative estimation of concentrations of dissolved rare earth elements using reflectance spectroscopy. APPRES 7, 073513-073513.
Dalton, J.B., Bove, D.J., Mladinich, C.S., Rockwell, B.W., 2004. Identification of spectrally similar materials using the USGS Tetracorder algorithm: the calcite–epidote–chlorite problem. Remote Sensing of Environment 89, 455-466.
de Araújo, C.C., Yamamoto, J.K., Rostirolla, S.P., 2006. Tar sandstones in the Paraná Basin: a study of the occurrences in the Anhembi Structural High (in Portuguese). Petrobras Bulletin of Geosciences 14, 47-70.
De Jong, S.M., van der Meer, F.D., 2005. Remote sensing image analysis including the spatial domain. Kluwer Academic Publishers.
de Souza Filho, C.R., Drury, S.A., Dennis, A.M., Carlton, R.W.T., Rothery, D.A., 1996. Restoration of corrupted optical Fuyo-1 (JERS-1) data using frequency domain techniques. Photogrammetric Engineering & Remote Sensing 62, 1037-1047.
Debba, P., Carranza, E.J.M., van der Meer, F.D., Stein, A., 2006. Abundance Estimation of Spectrally Similar Minerals by Using Derivative Spectra in Simulated Annealing. Geoscience and Remote Sensing, IEEE Transactions on 44, 3649-3658.
Demetriades-Shah, T.H., Steven, M.D., Clark, J.A., 1990. High resolution derivative spectra in remote sensing. Remote Sensing of Environment 33, 55-64.
Dobigeon, N., Tourneret, J.Y., Chein, I.C., 2008. Semi-Supervised Linear Spectral Unmixing Using a Hierarchical Bayesian Model for Hyperspectral Imagery. Signal Processing, IEEE Transactions on 56, 2684-2695.
Dobigeon, N., Tourneret, J.Y., Richard, C., Bermudez, J.C.M., McLaughlin, S., Hero, A.O., 2014. Nonlinear Unmixing of Hyperspectral Images: Models and Algorithms. Signal Processing Magazine, IEEE 31, 82-94.
Du, Y., Chang, C.I., Ren, H., Chang, C.C., Jensen, J.O., D'Amico, F.M., 2004. New hyperspectral discrimination measure for spectral characterization. Optical Engineering 43, 1777–1786.
Duke, E.F., 1994. Near infrared spectra of muscovite, Tschermak substitution, and metamorphic reaction progress: Implications for remote sensing. Geology 22, 621-624.

Eismann, M.T., Stein, D.W.J., 2007. Stochastic mixture modeling, in: Chang, C.I. (Ed.), Hyperspectral data exploitation; Theory and applications. John Wiley & Sons, Hoboken, NJ, pp. 107-148.
ERSDAC, 2006. Preliminary copper exploration studies by remote sensing data analysis. National Iranian Copper Industries Company (NICICO), Tehran, unpublished internal report,, p. 328.
Esbensen, K.H., 2006. Multivariate Data Analysis in Practice: An Introduction to Multivariate Data Analysis and Experimental Design, 5th ed. Camo, Oslo.
Farzam, M., Beheshti, S., Raahemifar, K., 2008. Calculation of abundance factors in hyperspectral imaging using genetic algorithms, Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE), Niagara Falls, Ontario, pp. 837-842.
Feng, J., Rivard, B., Gallie, A., Sanchez-Azofeifa, A., 2011. Rock type classification of drill core using continuous wavelet analysis applied to thermal infrared reflectance spectra. International Journal of Remote Sensing 32, 4489–4510.
Feng, J., Rivard, B., Gallie, E.A., Sanchez, A., 2006. Quantifying total sulfide content of cores and cut-rock surfaces using thermal infrared reflectance. Geophysics 71, M1-M9.
Foody, G.M., Lucas, R.M., Curran, P.J., Honzak, M., 1997. Non-linear mixture modelling without end-members using an artificial neural network. International Journal of Remote Sensing 18, 937-953.
Fraser, S.J., 1991. Discrimination and identification of ferric oxides using satellite Thematic Mapper data: A Newman case study. International Journal of Remote Sensing 12, 635-614.
Friedl, M.A., Brodley, C.E., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 61, 399-409.
Funk, C.C., Theiler, J., Roberts, D.A., Borel, C.C., 2001. Clustering to improve matched filter detection of weak gas plumes in hyperspectral thermal imagery. Geoscience and Remote Sensing, IEEE Transactions on 39, 1410-1420.
Gaffey, S.J., McFadden, L.A., Nash, D., Pieters, C.M., 1993. Ultraviolet, visible, and near-infrared reflectance spectroscopy: laboratory spectra of geologic materials, in: Pieters, C.M., Englert, P.A.J. (Eds.), Remote geochemical analysis: elemental and mineralogical composition. Cambridge University Press, NY, USA, pp. 43-77.
Gao, B.-C., Montes, M.J., Davis, C.O., Goetz, A.F.H., 2009. Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. Remote Sensing of Environment 113, Supplement 1, S17-S24.
García-Haro, F.J., Sommer, S., Kemper, T., 2005. A new tool for variable multiple endmember spectral mixture analysis (VMESMA). International Journal of Remote Sensing 26, 2135-2162.
Gilmore, M.S., Bornstein, B., Merrill, M.D., Castaño, R., Greenwood, J.P., 2008. Generation and performance of automated jarosite mineral detectors for visible/near-infrared spectrometers at Mars. Icarus 195, 169-183.
Goetz, A.F.H., 2009. Three decades of hyperspectral remote sensing of the Earth: A personal view. Remote Sensing of Environment 113, Supplement 1, S5-S16.
Goetz, A.F.H., Curtiss, B., Shiley, D.A., 2009. Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy. Minerals Engineering 22, 490-499.
Goetz, A.F.H., Rowan, L.C., 1981. Geologic remote sensing. Science 211, 781–791.
Goetz, A.F.H., Vane, G., Solomon, J.E., Rock, B.N., 1985. Imaging spectrometry for earth remote sensing. Science 228, 1147−1153.
Granahan, J.C., Sweet, J.N., 2001. An evaluation of atmospheric correction techniques using the spectral similarity scale, Geoscience and Remote Sensing Symposium, IGARSS '01. IEEE 2001 International, pp. 2022-2024.
Green, A.A., Berman, M., Switzer, P., Craig, M.D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. Geoscience and Remote Sensing, IEEE Transactions on 26, 65-74.
Green, A.A., Craig, M.D., 1984. Integrated analysis of image data for mineral exploration, Proceedings of the International Symposium on Remote Sensing of Environment, Third Thematic Conference, Remote Sensing for Exploration Geology, Colorado Springs, Colorado, pp. 131-137.
Green, R.O., Eastwood, M.L., Sarture, C.M., Chrien, T.G., Aronsson, M., Chippendale, B.J., Faust, J.A., Pavri, B.E., Chovit, C.J., Solis, M., Olah, M.R., Williams, O., 1998. Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). Remote Sensing of Environment 65, 227-248.
Gregory, A.F., Moore, H.D., 1975. The role of remote sensing in mineral exploration with special reference to ERTS-1. Canadian Mining and Metallurgical Bulletin 68, 67–72.
Griffiths, P., de Haseth, J.A., 2007. Fourier transform infrared spectrometry, Second ed. John Wiley & Sons, Hoboken, New Jersey.
Haest, M., Cudahy, T., Laukamp, C., Gregory, S., 2012. Quantitative mineralogy from infrared spectroscopic data. I. Validation of mineral abundance and composition scripts at the Rocklea Channel iron deposit in Western Australia. Economic Geology 107, 209-228.
Hamilton, V.E., Christensen, P.R., 2000. Determining the modal mineralogy of mafic and ultramafic igneous rocks using thermal emission spectroscopy. Journal of Geophysical Research: Planets 105, 9717-9733.
Hapke, B., 1981. Bidirectional reflectance spectroscopy: 1. Theory. Journal of Geophysical Research: Solid Earth 86, 3039-3054.
Hapke, B., 1993. Theory of reflectance and emittance spectroscopy. Cambridge University Press, Cambridge, UK.
Harsanyi, J.C., Chein, I.C., 1994. Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach. Geoscience and Remote Sensing, IEEE Transactions on 32, 779-785.
Hecker, C., Dilles, J.H., van der Meijde, M., van der Meer, F.D., 2012. Thermal infrared spectroscopy and partial least squares regression to determine mineral modes of granitoid rocks. Geochemistry, Geophysics, Geosystems 13, Q03021.

Hecker, C.A., 2012. Mapping feldspars from above : A thermal infrared and partial least squares based approach, University of Twente Faculty of Geo-Information and Earth Observation (ITC), Enschede, p. 165.
Herrmann, W., Blake, M., Doyle, M., Huston, D., Kamprad, J., Merry, N., Pontual, S., 2001. Short Wavelength Infrared (SWIR) Spectral Analysis of Hydrothermal Alteration Zones Associated with Base Metal Sulfide Deposits at Rosebery and Western Tharsis, Tasmania, and Highway-Reward, Queensland. Economic Geology 96, 939-955.
Heylen, R., Parente, M., Gader, P., 2014. A Review of Nonlinear Hyperspectral Unmixing Methods. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of PP, 1-26.
Hiroi, T., Pieters, C.M., 1992. Effects of grain size and shape in modeling reflectance spectra of mineral mixtures, Proceedings of Lunar and Planetary Science. Lunar and Planetary Institute, Houston, TX, pp. 313-325.
Holliday, J.R., Cooke, D.R., 2007. Advances in geological models and exploration methods for copper ± gold porphyry deposits, in: Milkereit, B. (Ed.), Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration, Toronto, Canada, pp. 791-809.
Hook, S.J., Abbot, E.A., Grove, C., Kahle, A.B., Fallucani, F.D., 1999. Multispectral thermal infrared data in geological studies, in: Rencz, A.N. (Ed.), Remote sensing for the earth sciences, 3rd ed. John Wiley & Sons.
Hook, S.J., Johnson, W.R., Abrams, M.J., 2013. NASA's Hyperspectral Thermal Emission Spectrometer (HyTES), in: Kuenzer, C., Dech, S. (Eds.), Thermal infrared remote sensing: sensors, methods, applications. Springer, pp. 93-115.
Hsu, p., 2003. Spectral feature extraction of hyperspectral images using wavelet transform, Department of survey engineering. National Cheng Kung University, Taiwan.
Huguenin, R.L., Jones, J.L., 1986. Intelligent information extraction from reflectance spectra: Absorption band positions. Journal of Geophysical Research: Solid Earth 91, 9585-9598.
Hunt, G.R., 1977. Spectral signatures of particulate minerals in the visible and near infrared. Geophysics 42, 501-513.
Hunt, G.R., 1979. Near-infrared (1.3-2.4 pm) spectra of alteration minerals- Potential for use in remote sensing. Geophysics 44, 1974–1986.
Hunt, G.R., Ashley, R.P., 1979. Spectra of altered rocks in the visible and near infrared. Economic Geology 74, 1613-1629.
Hunt, G.R., Salisbury, J.W., 1971. Visible and near infrared spectra of minerals and rocks: I. Silicates. Modern Geology 1, 283–300.
Hunt, G.R., Salisbury, J.W., 1974. Midinfrared spectral behaviour of igneous rocks. U.S. Air Force Cambridge Research Laboratories Technical Report AFCRL-TR-74-0625, Cambridge, MA,.
Huntington, J., Hancock, L., Smith, B., Morin Ka, S., 2012. Rare earth reflectance spectroscopy basics and Australian NVCL applications, Symposium 32.4 at the 34th International Geological Congress, Brisbane, Australia.
Huntington, J.F., Whitbourn, L., Mason, P., Schodlok, M., Berman, M., 2010. Hylogging - Voluminous industrial-scale reflectance spectroscopy of the earth's subsurface, Proceedings of ASD and IEEE GRS; Art, Science and Applications of Reflectance Spectroscopy Symposium, Boulder, Colorado, p. 14.
Jie-Lun, C., Jun-Jih, L., Chiang, W., Ke-Sheng, C., 2014. A feature-space indicator kriging approach for remote sensing image classification. Geoscience and Remote Sensing, IEEE Transactions on 52, 4046-4055.
Johnson, W.R., Hulley, G., Hook, S.J., 2014. Remote gas plume sensing and imaging with NASA's Hyperspectral Thermal Emission Spectrometer (HyTES), pp. 91010V-91010V-91017.
Jolliffe, I.T., 1986. Principal component analysis. Springer-Verlag, New York.
Kauppinen, J.K., Moffatt, D.J., Mantsch, H.H., Cameron, D.G., 1981. Fourier Self-Deconvolution: A Method for Resolving Intrinsically Overlapped Bands. Appl. Spectrosc. 35, 271-276.
Keshava, N., 2004. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. Geoscience and Remote Sensing, IEEE Transactions on 42, 1552-1565.
Keshava, N., Kerekes, J.P., Manolakis, D.G., Shaw, G.A., 2000. An algorithm taxonomy for hyperspectral unmixing, Proceeding of SPIE 4049, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, Orlando, FL, pp. 42-63.
Keshava, N., Mustard, J.F., 2002. Spectral unmixing. Signal Processing Magazine, IEEE 19, 44-57.
Kokaly, R.F., 2012. Spectroscopic remote sensing for material identification, vegetation characterization, and mapping, in: Shen, S.S. (Ed.), Proceeding, SPIE 8390, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII. SPIE, Maryland, USA, pp. 839014-839014-839012.
Kozak, P.K., Duke, E.F., Roselle, G.T., 2004. Mineral distribution in contact-metamorphosed siliceous dolomite at Ubehebe Peak, California, based on airborne imaging spectrometer data. American Mineralogist 89, 701-713.
Kraut, S., Scharf, L.L., Butler, R.W., 2005. The adaptive coherence estimator: a uniformly most-powerful-invariant adaptive detection statistic. IEEE Transactions on Signal Processing 53, 427-438.
Kruger, G., Erzinger, J., Kaufmann, H., 1998. Laboratory and airborne reflectance spectroscopic analyses of lignite overburden dumps. Journal of Geochemical Exploration 64, 47–65.
Kruse, F.A., 2008. Expert system analysis of hyperspectral data, Proceedings, SPIE Defense and Security, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV, Conference DS43, Orlando, Florida, p. 12.
Kruse, F.A., 2015. Comparative analysis of Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), and Hyperspectral Thermal Emission Spectrometer (HyTES) longwave infrared (LWIR) hyperspectral data for geologic mapping, pp. 94721F-94721F-94713.
Kruse, F.A., Bedell, R.L., Taranik, J.V., Peppin, W.A., Weatherbee, O., Calvin, W.M., 2012. Mapping alteration minerals at prospect, outcrop and drill core scales using imaging spectrometry. International Journal of Remote Sensing 33, 1780-1798.
Kruse, F.A., Lefkoff, A.B., 1993. Knowledge-based geologic mapping with imaging spectrometers. Remote Sensing Reviews 8, 3–28.

Kruse, F.A., Lefkoff, A.B., Boardman, J.B., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993a. The Spectral Image Processing System (SIPS) - Interactive Visualization and Analysis of Imaging Spectrometer Data. Remote Sensing of Environment 44, 145–163.
Kruse, F.A., Lefkoff, A.B., Dietz, J.B., 1993b. Expert system-based mineral mapping in northern death valley, California/Nevada, using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). Remote Sensing of Environment 44, 309-336.
Kuosmanen, V.V., Laitinen, L.J., 2008. Quantitative Mineral Assessment of Apatite, Calcite/Dolomite, and Phlogopite Powder Mixtures by Using VSWIR Reflectance. Geoscience and Remote Sensing, IEEE Transactions on 46, 1774-1782.
Kurz, T.H., Dewit, J., Buckley, S.J., Thurmond, J.B., Hunt, D.W., Swennen, R., 2012. Hyperspectral image analysis of different carbonate lithologies (limestone, karst and hydrothermal dolomites): the Pozalagua Quarry case study (Cantabria, North-west Spain). Sedimentology 59, 623-645.
Landgrebe, D.A., 2003. Signal theory methods in multispectral remote sensing. John Wiley & Sons, Hoboken, NJ.
Laukamp, C., Cudahy, T., Thomas, M., Jones, M., Cleverley, J.S., Oliver, N.H.S., 2011. Hydrothermal mineral alteration patterns in the Mount Isa Inlier revealed by airborne hyperspectral data. Australian Journal of Earth Sciences 58, 917-936.
Li, S., Wu, H., Wan, D., Zhu, J., 2011. An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine. Knowledge-Based Systems 24, 40–48.
Licciardi, G.A., Del Frate, F., 2011. Pixel Unmixing in Hyperspectral Data by Means of Neural Networks. Geoscience and Remote Sensing, IEEE Transactions on 49, 4163-4172.
Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28, 823-870.
Lyder, D., Feng, J., Rivard, B., Gallie, A., Cloutis, E., 2010. Remote bitumen content estimation of Athabasca oil sand from hyperspectral infrared reflectance spectra using Gaussian singlets and derivative of Gaussian wavelets. Fuel 89, 760-767.
Lyon, R.J.P., Burns, E.A., 1963. Analysis of rocks and minerals by reflected infrared radiation. Economic Geology 58, 274-284.
Maddams, W.F., 1980. The Scope and Limitations of Curve Fitting. Appl. Spectrosc. 34, 245-267.
Manolakis, D., Marden, D., Shaw, G.A., 2003. Hyperspectral image processing for automatic target detection applications. Lincoln Laboratory Journal 14, 79–116.
Manolakis, D., Shaw, G., 2002. Detection algorithms for hyperspectral imaging applications. Signal Processing Magazine, IEEE 19, 29-43.
Mars, J.C., Rowan, L.C., 2006. Regional mapping of phyllic- and argillic-altered rocks in the Zagros magmatic arc, Iran, using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data and logical operator algorithms. Geosphere 2, 161-186.
Mars, J.C., Rowan, L.C., 2010. Spectral assessment of new ASTER SWIR surface reflectance data products for spectroscopic mapping of rocks and minerals. Remote Sensing of Environment 114, 2011-2025.
Mas, J.F., Flores, J.J., 2007. The application of artificial neural networks to the analysis of remotely sensed data. International Journal of Remote Sensing 29, 617-663.
Mason, P., Huntington, J.F., 2012. HyLogger 3 components and preprocessing: An overview. Northern Territory Geological Survey, Technical Note 2012-002,.
Mazer, A.S., Martin, M., Lee, M., Solomon, J.E., 1988. Image processing software for imaging spectrometry analysis. Remote Sensing of Environment 24, 201-210.
Melgani, F., Bruzzone, L., 2004. Classification of hyperspectral remote sensing images with support vector machines. Geoscience and Remote Sensing, IEEE Transactions on 42, 1778-1790.
Mitchley, M., Sears, M., Damelin, S., 2009. Target detection in hyperspectral mineral data using wavelet analysis, Geoscience and Remote Sensing Symposium, IEEE International, IGARSS, pp. IV-881-IV-884.
Monteiro, S.T., Murphy, R.J., Ramos, F., Nieto, J., 2009. Applying boosting for hyperspectral classification of ore-bearing rocks, Machine Learning for Signal Processing, MLSP 2009. IEEE International Workshop on, Grenoble, France, pp. 1-6.
Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66, 247–259.
Murphy, R.J., Monteiro, S.T., 2013. Mapping the distribution of ferric iron minerals on a vertical mine face using derivative analysis of hyperspectral imagery (430–970 nm). ISPRS Journal of Photogrammetry and Remote Sensing 75, 29-39.
Mustard, J.F., 1992. Chemical analysis of actinolite from reflectance spectra. American Mineralogist 77, 345-358.
Mustard, J.F., Pieters, C.M., 1989. Photometric phase functions of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra. Journal of Geophysical Research: Solid Earth 94, 13619-13634.
Mustard, J.F., Sunshine, J.M., 1999. Spectral analysis for earth science investigation: Investigations using remote sensing data, in: Rencz, A.N. (Ed.), Remote sensing for the earth sciences, 3rd ed. John Wiley & Sons.
Nascimento, J.M.P., Bioucas Dias, J.M., 2005. Does independent component analysis play a role in unmixing hyperspectral data? Geoscience and Remote Sensing, IEEE Transactions on 43, 175-187.
Ninomiya, Y., 1995. Quantitative estimation of SiO2 content in igneous rocks using thermal infrared spectra with a neural network approach. Geoscience and Remote Sensing, IEEE Transactions on 33, 684-691.
Oskouie, M.M., Busch, W., 2008. A Geostatistically Based Preprocessing Algorithm for Hyperspectral Data Analysis. GIScience & Remote Sensing 45, 356-368.
Pal, M., Mather, P.M., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment 86, 554-565.

Pearlman, J.S., Barry, P.S., Segal, C.C., Shepanski, J., Beiso, D., Carman, S.L., 2003. Hyperion, a space-based imaging spectrometer. Geoscience and Remote Sensing, IEEE Transactions on 41, 1160-1173.
Penn, B.S., 2002. Using simulated annealing to obtain optimal linear end-member mixtures of hyperspectral data. Computers & Geosciences 28, 809-817.
Philpot, W.D., 1991. The derivative ratio algorithm: avoiding atmospheric effects in remote sensing. Geoscience and Remote Sensing, IEEE Transactions on 29, 350-357.
Piech, M.A., Piech, K.R., 1990. Fingerprints and fractal terrain. Mathematical Geology 22, 457–485.
Piech, M.A., R., P.K., 1987. Symbolic representation of hyperspectral data. Applied Optics 26, 4018-4026.
Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J.C., Trianni, G., 2009. Recent advances in techniques for hyperspectral image processing. Remote Sensing of Environment 113, Supplement 1, S110-S122.
Plaza, A., Du, Q., Bioucas-Dias, J.M., Jia, X., Kruse, F.A., 2011. Foreword to the Special Issue on Spectral Unmixing of Remotely Sensed Data. Geoscience and Remote Sensing, IEEE Transactions on 49, 4103-4110.
Plaza, A., Martínez, P., Pérez, R., Plaza, J., 2004. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing 42, 650-663.
Pompilio, L., Pedrazzi, G., Cloutis, E.A., Craig, M.A., Roush, T.L., 2010. Exponential Gaussian approach for spectral modelling: The EGO algorithm II. Band asymmetry. Icarus 208, 811-823.
Pompilio, L., Pedrazzi, G., Sgavetti, M., Cloutis, E.A., Craig, M.A., Roush, T.L., 2009. Exponential Gaussian approach for spectral modeling: The EGO algorithm I. Band saturation. Icarus 201, 781-794.
Pontual, S., Merry, N., Gamson, P., 2008a. G-MEX—spectral analysis guides for mineral exploration—Practical applications handbook. Ausspec International Pty. Ltd, Kew, Victoria 3101, Australia,, p. 86.
Pontual, S., Merry, N., Gamson, P., 2008b. G-MEX—spectral analysis guides for mineral exploration—Spectral interpretation field manual. Ausspec International Pty. Ltd, Kew, Victoria 3101, Australia,, p. 189.
Post, J.L., Noble, P.N., 1993. The near-infrared combination band frequencies of dioctahedral smectites, micas, and illites. Clays and Clay Minerals 41, 639-644.
Poulet, F., Erard, S., 2004. Nonlinear spectral mixing: Quantitative analysis of laboratory mineral mixtures. Journal of Geophysical Research: Planets 109, E02009.
Ragona, D., Minster, B., Rockwell, T., Jussila, J., 2006. Field imaging spectroscopy: A new methodology to assist the description, interpretation, and archiving of paleoseismological information from faulted exposures. Journal of Geophysical Research: Solid Earth 111, B10309.
Rajesh, H.M., 2004. Application of remote sensing and GIS in mineral resource mapping—an overview. Journal of Mineralogical and Petrological Sciences 99, 83–103.
Ramsey, M.S., Christensen, P.R., 1998. Mineral abundance determination: Quantitative deconvolution of thermal emission spectra. Journal of Geophysical Research: Solid Earth 103, 577-596.
Ren, H., Chang, C.-I., 2000. Target-constrained interference-minimized approach to subpixel target detection for hyperspectral images. Optical Engineering 39, 3138-3145.
Richards, J.A., 2005. Analysis of remotely sensed data: the formative decades and the future. Geoscience and Remote Sensing, IEEE Transactions on 43, 422-432.
Richards, J.A., Jia, X., 2006. Remote sensing digital image analysis, 4th ed. Springer Pub.
Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sensing of Environment 112, 2850-2862.
Rivard, B., Zhang, J., Feng, J., Sanchez-Azofeifa, G.A., 2009. Remote predictive lithologic mapping in the Abitibi Greenstone Belt, Canada, using airborne hyperspectral imagery. Canadian Journal of Remote Sensing 35, S95–S105.
Roache, T.J., Walshe, J.L., Huntington, J.F., Quigley, M.A., Yang, K., Bil, B.W., Blake, K.L., Hyvärinen, T., 2011. Epidote–clinozoisite as a hyperspectral tool in exploration for Archean gold. Australian Journal of Earth Sciences 58, 813-822.
Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R.O., 1998. Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. Remote Sensing of Environment 65, 267-279.
Robila, S.A., 2005. Using spectral distances for speedup in hyperspectral image processing. International Journal of Remote Sensing 26, 5629-5650.
Robila, S.A., Gershman, A., 2005. Spectral matching accuracy in processing hyperspectral data, Signals, Circuits and Systems, International Symposium on, ISSCS, pp. 163-166 Vol. 161.
Rodger, A., Cudahy, T.J., 2009. Vegetation corrected continuum depths at 2.20 µm: An approach for hyperspectral sensors. Remote Sensing of Environment 113, 2243-2257.
Rodger, A., Laukamp, C., Haest, M., Cudahy, T., 2012. A simple quadratic method of absorption feature wavelength estimation in continuum removed spectra. Remote Sensing of Environment 118, 273-283.
Rogge, D.M., Rivard, B., Zhang, J., Jilu, F., 2006. Iterative Spectral Unmixing for Optimizing Per-Pixel Endmember Sets. Geoscience and Remote Sensing, IEEE Transactions on 44, 3725-3736.
Rowan, L.C., Wetlaufer, P.H., Goetz, A.F.H., Billingsley, F.C., Stewart, J.H., 1974. Discrimination of rock types and detection of hydrothermally altered areas in south-central Nevada by use of computer-enhanced ERTS images. Geological Survey Professional Paper 883, United States Geological Survey, U.S. Government Printing Office, Washington D.C,.
Roy, R., Launeau, P., Carrère, V., Pinet, P., Ceuleneer, G., Clénet, H., Daydou, Y., Girardeau, J., Amri, I., 2009. Geological mapping strategy using visible near-infrared–shortwave infrared hyperspectral remote sensing: Application to the Oman ophiolite (Sumail Massif). Geochemistry, Geophysics, Geosystems 10, Q02004.

Rud, R., Shoshany, M., Alchanatis, V., Cohen, Y., 2006. Application of spectral features' ratios for improving classification in partially calibrated hyperspectral imagery: a case study of separating Mediterranean vegetation species. Journal of Real-Time Image Processing 1, 143–152.
Sabine, C., 1999. Remote sensing strategies for mineral exploraion, in: Rencz, A.N. (Ed.), Remote sensing for the earth sciences, 3rd ed. John Wiley & Sons, pp. 375-447.
Sabins, F.F., 1999. Remote sensing for mineral exploration. Ore Geology Reviews 14, 157-183.
Sanga, T., Tachikawa, T., 2006. Highly accurate mineral mapping using ASTER data and Iso-grain model, 5th ASTER Workshop, Tokyo.
Schaepman, M.E., Ustin, S.L., Plaza, A.J., Painter, T.H., Verrelst, J., Liang, S., 2009. Earth system science related imaging spectroscopy—An assessment. Remote Sensing of Environment 113, Supplement 1, S123-S137.
Schmidt, F., Douté, S., Schmitt, B., 2007. WAVANGLET: An efficient supervised classifier for hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing 45, 1374–1385.
Schmidt, K.S., Skidmore, A.K., 2004. Smoothing vegetation spectra with wavelets. International Journal of Remote Sensing 25, 1167-1184.
Schott, J.R., 2006. Remote sensing: The image chain approach, 2nd edition ed. Oxford University Press, New York.
Schowengerdt, R.A., 2007. Remote sensing model and methods for image processing, 3rd ed. Elsevier Inc.
Sgavetti, M., Pompilio, L., Meli, S., 2006. Reflectance spectroscopy (0.3–2.5 μm) at various scales for bulk-rock identification. Geosphere 2, 142-160.
Shen, S.S., Bassett, E.M., 2002. Information-theory-based band selection and utility evaluation for reflective spectral systems, Proceedings of SPIE conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII. SPIE, Orlando, FL, pp. 18-29.
Shipman, H., Adams, J.B., 1987. Detectability of minerals on desert alluvial fans using reflectance spectra. Journal of Geophysical Research: Solid Earth 92, 10391-10402.
Shkuratov, Y., Starukhina, L., Hoffmann, H., Arnold, G., 1999. A Model of Spectral Albedo of Particulate Surfaces: Implications for Optical Properties of the Moon. Icarus 137, 235-246.
Sillitoe, R.H., Thompson, J.F.H., 2006. Changes in mineral exploration practice: Consequences for discovery. economic Geology special publication 12, 193-219.
Smith, M.O., Ustin, S.L., Adams, J.B., Gillespie, A.R., 1990. Vegetation in deserts: I. a regional measure of abundance for multispectral images. Remote Sensing of Environment 31, 1-26.
Somers, B., Asner, G.P., Tits, L., Coppin, P., 2011. Endmember variability in Spectral Mixture Analysis: A review. Remote Sensing of Environment 115, 1603-1616.
Somers, B., Delalieux, S., Verstraeten, W.W., van Aardt, J.A.N., Albrigo, G.L., Coppin, P., 2010. An automated waveband selection technique for optimized hyperspectral mixture analysis. International Journal of Remote Sensing 31, 5549-5568.
Sunshine, J.M., Pieters, C.M., 1993. Estimating modal abundances from the spectra of natural and laboratory pyroxene mixtures using the modified Gaussian model. Journal of Geophysical Research: Planets 98, 9075-9087.
Sunshine, J.M., Pieters, C.M., Pratt, S.F., 1990. Deconvolution of mineral absorption bands: an improved approach. Journal of Geophysical Research 93, 6955-6966.
Swayze, G.A., Clark, R.N., Goetz, A.F.H., Livo, K.E., Breit, G.N., Kruse, F.A., Sutley, S.J., Snee, L.W., Lowers, H.A., Post, J.L., Stoffregen, R.E., Ashley, R.P., 2014. Mapping Advanced Argillic Alteration at Cuprite, Nevada, Using Imaging Spectroscopy. Economic Geology 109, 1179-1221.
Swayze, G.A., Smith, K.S., Clark, R.N., Sutley, S.J., Pearson, R.M., Vance, J.S., Hageman, P.L., Briggs, P.H., Meier, A.L., Singleton, M.J., Roth, S., 2000. Using Imaging Spectroscopy To Map Acidic Mine Waste. Environmental Science & Technology 34, 47-54.
Tappert, M., Rivard, B., Giles, D., Tappert, R., Mauger, A., 2011. Automated drill core logging using visible and near-infrared reflectance spectroscopy: A case study from the Olympic Dam IOCG deposit, South Australia. Economic Geology 106, 289-296.
Tappert, M.C., Rivard, B., Tappert, R., Feng, J., 2013. Using reflectance spectroscopy to estimate the orientation of quartz crystals in rocks. The Canadian Mineralogist 51, 405-413.
Thompson, A.J.B., Hauff, P.L., Robitaille, A.J., 1999. Alteration mapping in exploration: Application of short-wave infrared (SWIR) spectroscopy. Society of Economic Geologists Newsletter 39, 16-27.
Thorpe, A.K., Roberts, D.A., Bradley, E.S., Funk, C.C., Dennison, P.E., Leifer, I., 2013. High resolution mapping of methane emissions from marine and terrestrial sources using a Cluster-Tuned Matched Filter technique and imaging spectrometry. Remote Sensing of Environment 134, 305-318.
Tsai, F., Philpot, W., 1998. Derivative Analysis of Hyperspectral Data. Remote Sensing of Environment 66, 41-51.
Tso, B., Mather, P.M., 2009. Classification methods for remotely sensed data, Second ed. Taylor & Francis Group, LCC, New York.
van der Meer, F., 1994. Extraction of mineral absorption features from high-spectralresolution data using non-parametric geostatistical techniques. International Journal of Remote Sensing 15, 2193-2214.
van der Meer, F., 1996. Classification of remotely-sensed imagery using an indicator kriging approach: application to the problem of calcite-dolomite mineral mapping. International Journal of Remote Sensing 17, 1233-1249.
van der Meer, F., 1999. Iterative spectral unmixing (ISU). International Journal of Remote Sensing 20, 3431-3436.
van der Meer, F., 2000. Spectral curve shape matching with a continuum removed CCSM algorithm. International Journal of Remote Sensing 21, 3179-3185.

van der Meer, F., 2004. Analysis of spectral absorption features in hyperspectral imagery. International Journal of Applied Earth Observation and Geoinformation 5, 55-68.
van der Meer, F., 2006a. The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. International Journal of Applied Earth Observation and Geoinformation 8, 3-17.
van der Meer, F., 2006b. Indicator kriging applied to absorption band analysis in hyperspectral imagery: A case study from the Rodalquilar epithermal gold mining area, SE Spain. International Journal of Applied Earth Observation and Geoinformation 8, 61-72.
van der Meer, F.D., Bakker, W., 1997. Cross Correlogram Spectral Matching (CCSM): application to surface mineralogical mapping using AVIRIS data from Cuprite, Nevada. Remote Sensing of Environment 61, 371–382.
van der Meer, F.D., De Jong, S.M., 2002. Imaging spectrometry: basic principles and prospective application. Kluwer Academic Publishers.
van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., Hecker, C.A., Bakker, W.H., Noomen, M.F., van der Meijde, M., Carranza, E.J.M., Smeth, J.B.d., Woldai, T., 2012. Multi- and hyperspectral geologic remote sensing: A review. International Journal of Applied Earth Observation and Geoinformation 14, 112-128.
Vane, G., Goetz, A.F.H., 1991. Terrestrial imaging spectrometry: current status, future trends. Remote Sensing of Environment 44, 117−126.
Vapnik, V.N., 1998. Statistical learning theory. John Wiley & Sons, New York.
Vaughan, R.G., Calvin, W.M., Taranik, J.V., 2003. SEBASS hyperspectral thermal infrared data: surface emissivity measurement and mineral mapping. Remote Sensing of Environment 85, 48-63.
Villmann, T., Merényi, E., Hammer, B., 2003. Neural maps in remote sensing image analysis. Neural Networks 16, 389-403.
Walter, L.S., Salisbury, J.W., 1989. Spectral characterization of igneous rocks in the 8- to 12-µm region. Journal of Geophysical Research: Solid Earth 94, 9203-9213.
Wang, F., 1990. Fuzzy supervised classification of remote sensing images. Geoscience and Remote Sensing, IEEE Transactions on 28, 194-201.
Wang, Y., Guo, L., Liang, N., 2009. Improving the classification precision of spectral angle mapper algorithm, Processing of Remote Sensing and GIS Data Processing and Other Applications. SPIE, Yichang, China, p. 6.
Waske, B., Benediktsson, J.A., Árnason, K., Sveinsson, J.R., 2009. Mapping of hyperspectral AVIRIS data using machine-learning algorithms. Canadian Journal of Remote Sensing 35, S106-S116.
Winter, E.M., Winter, M.E., 2000. Autonomous hyperspectral endmember determination methods, in: Fujisada, a.L. (Ed.), Proceedings PIE 3870, Sensors, Systems, and Next-Generation Satellites III, Florence, Italy, pp. 150–158.
www.exelisvis.com/ProductServices/ENVI.aspx.
Wyborn, L.A., Heinrich, C.A., Jaques, A.L., 1994. Australian Proterozoic Mineral Systems: Essential ingredients and mappable criteria, AUSIMM Annual Conference, Darwin, AU, pp. 109-115.
Xu, N., Hu, Y.X., Lei, B., Hong, Y.T., Dang, F.X., 2011. Mineral information extraction for hyperspectral image based on modified spectral feature fitting algorithm. Spectroscopy and Spectral Analysis 31, 1639–1643.
Yang, H., 1999. A back-propagation neural network for mineralogical mapping from AVIRIS data. International Journal of Remote Sensing 20, 97-110.
Youngentob, K.N., Roberts, D.A., Held, A.A., Dennison, P.E., Jia, X., Lindenmayer, D.B., 2011. Mapping two Eucalyptus subgenera using multiple endmember spectral mixture analysis and continuum-removed imaging spectrometry data. Remote Sensing of Environment 115, 1115-1128.
Zare, A., Gader, P., 2010. PCE: Piece-wise convex endmember detection. IEEE Transactions on Geoscience and Remote Sensing 48, 2620-2632.
Zhang, J., Rivard, B., Sanchez-Azofeifa, A., 2004. Derivative spectral unmixing of hyperspectral data applied to mixtures of lichen and rock. Geoscience and Remote Sensing, IEEE Transactions on 42, 1934-1940.
Zhouyu, F., Robles-Kelly, A., Caelli, T., Tan, R.T., 2007. On automatic absorption detection for imaging spectroscopy: a comparative study. Geoscience and Remote Sensing, IEEE Transactions on 45, 3827-3844.
Zortea, M., Plaza, A., 2009. Spatial Preprocessing for Endmember Extraction. Geoscience and Remote Sensing, IEEE Transactions on 47, 2679-2693.

Chapter 4

Iterative curve fitting: A robust technique to estimate the wavelength position and depth of absorption features from spectral data[5]

In this chapter, we introduce a robust method coined iterative curve fitting to estimate the wavelength position and depth from spectral absorption features. The technique iteratively fits a curve to a continuum-removed spectrum and subsets the bands based on the minimum of the previous fit until fulfilling a specified threshold for residual error. The minimum of the latest iteration and its substituted reflectance value are then retrieved as the feature wavelength and depth. Two variants of the technique named iterative Gaussian fitting and iterative polynomial fitting (IPF) are presented. The superiority of these algorithms over current methods is demonstrated using four different absorption features between 400 and 2500 nm collected from an array of sandstones in the laboratory. The methods can achieve rmse values of ±1.0 nm for the wavelength and 1% for the feature depth. The estimated wavelength position in a hyperspectral sensor with less than 10 nm sampling interval is demonstrated to be in error by at most ±3 nm at 95% confidence level. Experiments with varying signal-to-noise ratios (SNRs) indicate the robustness of the technique against noise. The IPF is able to estimate the wavelength of narrow features with an rmse of ±2.7 nm at an SNR of 150:1 and broad features with an rmse of ±4.2 nm at an SNR of 400:1. The method, which is embedded in a package named Automated Absorption-based Mineral Spectral Analyzer (AMISA), enables the simultaneous calculation of width, area, and asymmetry of spectral data acquired from imaging and nonimaging sensors.

4.1. Introduction

The analysis of spectral data collected in the visible-near infrared (VNIR), short-wave infrared (SWIR), or longwave infrared (LWIR) by imaging or non-imaging remote sensing systems can be performed using either data-driven or knowledge-based approach [1]. In the former, the data is compared against a set of endmembers and corresponding abundance images are obtained [2]. In the latter, every spectrum is assumed to be composed of a series of absorption bands superimposed on a continuum background [3]. The bands are routinely characterized by their wavelength position, depth, full-width at half maximum depth (FWHM), area, and asymmetry [1, 4-6]. By incorporating these collective spectral parameters with user-knowledge, semi-quantitative to quantitative information is obtained about the identity, abundance, and chemistry of a terrestrial or planetary remote sensing target [4, 6-8].

In the modified Gaussian model (MGM), all the components of a spectrum are simultaneously modeled by a series of Gaussian bands and a continuum line in the natural log reflectance space [9]. Normally, initial conditions (i.e. the center, FWHM, strength, and continuum) are provided by the user, and then the routine seeks an optimum nonlinear least-square solution for the provided collection. This method interactively decomposes all the features of a given spectrum into symmetric Gaussian bands, each of which characterized by a center, width, and strength parameter [9]. Albeit the merits of this technique, the initial estimation of the Gaussians, particularly for hyperspectral imaging (HSI) data, can be troublesome. Moreover, many applications do not demand for the spectral decomposition nor need to analyze the entire spectrum. Instead, a partial solution for a

[5] A version of this chapter is published in the IEEE Transaction on Geoscience and Remote Sensing journal.

limited portion of the spectrum covering a typical absorption feature is deemed appropriate for information extraction. In such situations, the absorption(s) of interest are directly described by the noted spectral parameters.

An accurate determination of the wavelength position is essential and prerequisite for absorption quantification, as it will significantly impact the accuracy of the depth and subsequently width and asymmetry of a feature [10]. Above all, this parameter is frequently used in geologic remote sensing for: (i) target detection and material identification [11, 12], (ii) discrimination among different species within a mineral group (e.g. carbonates) [13], (iii) delineation of subtle variations in the chemistry of a mineral due to cation substitution as an indicator of the physicochemistry of the environment (e.g. white mica compositional variation) [14], (iv) estimation of the relative quantity of a mineral in a solid solution (e.g. hematite/goethite ratio) [15], (v) determination of spectral mixing between different compounds, and (vi) occasionally, indication of grain-size variations (e.g. in iron oxides) [16]. The depth of an absorption feature, which is determined at the minimum wavelength, is widely used for abundance quantification [3, 17]. The whereabouts of the absorption feature itself is commonly revealed by the user; however, now several algorithms are proposed to automate this process [1].

The wavelength of absorption maximum may be directly determined from a spectrum, but this approach can be easily affected by noise. Accordingly, a preliminary spectral smoothing step has been embedded to this method for noise suppression [18]. Within this procedure, it is common to calculate the spectral derivatives and then apply them for wavelength determination [19]. The accuracy of such methods, however, is constrained by the instrument's sampling regime; in this way, an increase in the sampling interval gives rise to a decrease in the accuracy of wavelength position. Hence, the method is only appropriate for very high resolution (and noise-free) spectral data [20].

In the case of HSI data, a linear approximation method based on trigonometry was proposed by [5]. Nevertheless, it reduces the curvature of a feature into simply interpolated lines, thus causing a low accuracy (i.e. +/-15 nm) in wavelength estimation [5].

The fitted polynomials (FP) of higher order on the other hand, seem to be more suitable for this aim. In this approach, a polynomial of certain order (from second up to fifth) is fitted to the entire range of an absorption feature using a least-square technique, and then, based on the resultant equation, the wavelength and depth of the absorption feature are estimated [15, 17, 20-22]. Despite the widespread use, the accuracy of the technique is barely evaluated against independent datasets. Moreover, there is no accepted basis for choosing a polynomial of certain order for specific applications. In a study to determine the wavelength position of ferric absorption features centered at 900 nm, it was argued that a fifth order polynomial yielded better results compared to direct minimum determination. The average accuracy of this method was reported to be within +/- 8 nm at 4.6 nm sampling intervals [20].

In practical applications (e.g., mineral compositional mapping), the error in wavelength position estimation is required to be very low (i.e. few nanometers) in order for the real variations to stand out. As the spectrum of ferric absorption feature in Fig. 1 illustrates, second to forth order polynomials fitted to the entire range of the absorption between 776-1074 nm cannot successfully model the spectrum and cause errors in the estimation of wavelength position, depth, or both. To address this issue, three alternatives have been considered: (i) dramatic increase in the order of the

polynomial [20]. This solution can indeed improve the accuracy, though in practice, for example with HSI data, there are limitations upon the polynomial order (i.e. N-1 with N bands) increment. In addition, it is possible that a higher order polynomial overfit a noisy absorption band; (ii) restriction of the number of fitted bands to only three, including the absolute minimum and its neighbors. Such a procedure, named quadratic fitting to the minimum (QFM), was shown to be efficient for narrow absorptions present in SWIR region of HSI dataset [23]. However, it may not be appropriate for broader absorption features or datasets with higher spectral resolution (Fig. 1 and 6a); (iii) independent estimation of wavelength position and depth using two different FPs [17]. This solution may alleviate the results, but it cannot resolve the inherent uncertainties around FPs of fixed order (Fig. 1).

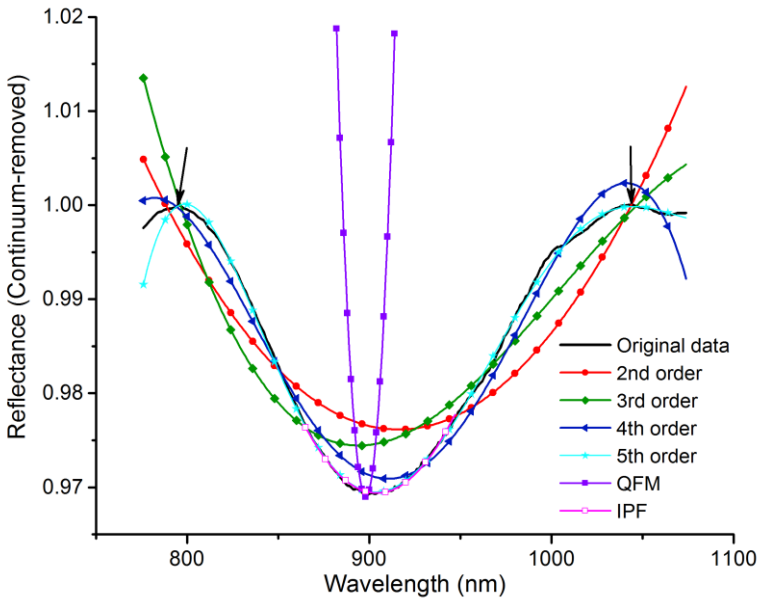

Fig.1. Spectral plots of a continuum-removed ferric absorption feature between 776–1074 nm, and fitted polynomials of second to fifth orders, quadratic fitting to the minimum (QFM), and iterative polynomial fitting (IPF) used to model the spectrum. The wavelength of the absorption minimum (determined directly after spectral smoothing) occurs at ≈ 903 nm with a feature depth of 0.03 relative the local continuum. The errors in estimating the wavelength position and feature depth from noted techniques relative to the reported values are, respectively, as following: second order (+13 nm, 20%), third order (-8 nm, 16.5%), fourth order (7 nm, 5%), fifth order (0.5 nm, 0.4%), QFM (-5 nm, -1.3%) and IPF (1.1 nm, 0%). Arrows indicate the shoulders of the feature.

Here, an alternative technique coined Iterative Curve Fitting (ICF) is introduced for the estimation of wavelength position and depth of spectral absorption features. The performance of the algorithm relative to current methods are assessed using four different absorption features collected from an array of sandstones in the laboratory. The algorithm is then tested against varying signal-to-noise ratios and different sampling intervals from ProspecTIR, HyMap, and the upcoming EnMap hyperspectral systems.

4.2. Iterative Curve Fitting (ICF)

The simplified flowchart of the ICF technique is outlined in Fig. 2. Initially, it removes the continuum from a spectrum, as this component tends to displace the absorption minimum [17, 19, 24]. We developed a simple methodology that iteratively calculates and removes the 'local' continuum over a given absorption feature. Let's consider the spectrum over an absorption feature

as a one-dimensional vector $\boldsymbol{s}\ (\lambda)$, confined between λ_1 and λ_2 wavelengths (normally provided by the user) with *K* discrete spectral bands. The steps for continuum removal are as follows:

(i) The maximum reflectance values, $\boldsymbol{s}\ (\lambda_{max})$ on either side of the median band (b_m) are determined.

(ii) A 'straight line' is fitted to the maximums (dashed line in Fig. 3) and the continuum is removed by dividing the spectrum to it.

(iii) In the resultant spectrum, the local maximums are determined again (step 1) and their relevant positions are projected over $\boldsymbol{s}\ (\lambda)$. A new line (dash-dot line in Fig. 3) is fitted to $\boldsymbol{s}\ (\lambda)$ and a continuum-removed spectrum is re-calculated (step ii).

(iv) Step (iii) is repeated until at least one spectral band on each side of b_m yields a value equal to 1.0 and there is no band with a value greater than 1.0.

The resultant continuum line (solid line in Fig. 3) should be tangent to the curve at both feature shoulders and the continuum-removed spectrum ($\boldsymbol{S}_{CR}(\lambda)$) should comprise values lower than or equal to 1.0.

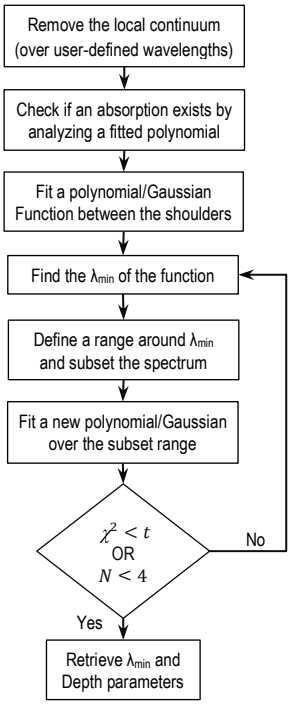

Fig. 2. Simplified flowchart of the proposed ICF technique for the estimation of feature wavelength position and depth. λ_{min} denotes minimum wavelength, t, threshold, N, band number, and χ^2, residual error according to Eq. 7.

Following the continuum removal, a preliminary curve is fitted to the spectrum (i.e. $\boldsymbol{S}_{CR}(\lambda)$). The algorithm assumes that the absorption of $\boldsymbol{S}_{CR}$ is located between the two shoulders, thus the fitting is confined between them (the arrows in Fig. 1). As a result, the number of available bands is reduced to N ($N \leq K$). We developed two variants of ICF named iterative Gaussian fitting (IGF) and iterative polynomial fitting (IPF), which respectively benefit from a Gaussian and a polynomial function for the fitting. A Gaussian function is defined as:

$$\boldsymbol{G}(\lambda) = \alpha . e^{-\frac{(\lambda-\mu)^2}{2\sigma^2}} \quad (1)$$

where μ is the wavelength of the center, α is the height of curve in the center, and σ is the width of the function.

A polynomial function is defined as:

$$\boldsymbol{P}(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_2\lambda^2 + a_1\lambda + a_0 \quad (2)$$

where $a_0, \cdots, a_n$ are constants and n is the order of the polynomial.

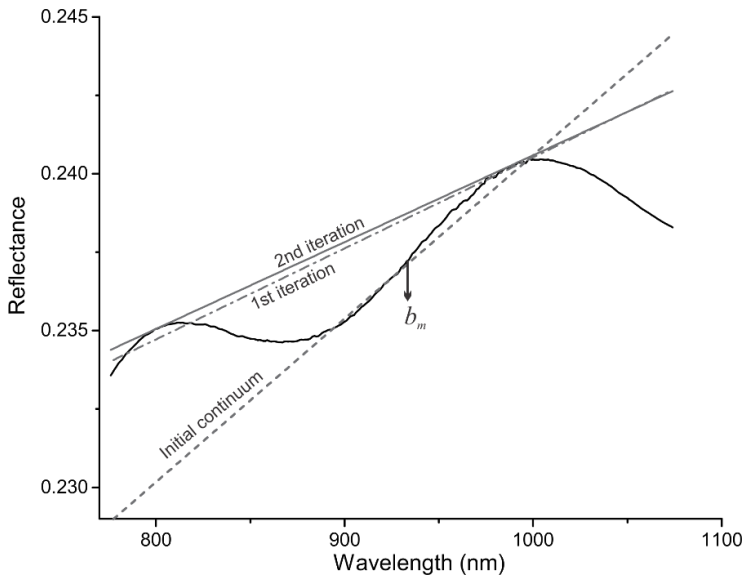

Fig. 3. Calculation of local continuum by the 'iterative continuum removal' algorithm. For a ferric (Feox) absorption feature, the continuum line tangent to the spectrum is depicted after two iterations.

The minimum of the fitted Gaussian function occurs at μ (Eq. 1), whereas that of the polynomial is determined from the roots of the explicit first derivative. The minimum for a quadratic function can be defined as:

$$\lambda_{\min} = \frac{-a_1}{2a_2} \quad (3)$$

For higher order polynomials, however, the minimum is determined from (the real part of the) roots of the first explicit derivative derived from the Lagrange method [25]. The minimum materializes where the following conditions are fulfilled:

$$\boldsymbol{P}^I(\lambda) = 0\,, \boldsymbol{P}^{II}(\lambda) > 0 \quad (4)$$

This initial minimum is considered as a center (Fig. 4), and a range is defined around it using the following formulation:

$$range = (0.5 - r) \times N \quad (5)$$

where N is the number of available bands and r is a constant ($0 < r < 0.5$), which is experimentally determined to be 0.25 for Gaussian and 0.2 for polynomial functions (see also section 4.8). The range is then used to subset the spectrum:

$$\boldsymbol{S}_{sub}(\lambda) = \boldsymbol{S}_{CR}[\lambda_{min} - range : \lambda_{min} + range] \quad (6)$$

A new curve is fitted to the subset spectrum (Fig. 4) and the residual error, χ^2 (chi squared) is calculated as:

$$\chi^2 = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{S}_{sub}(\lambda_i) - \boldsymbol{F}(\lambda_i))^2} \quad (7)$$

where $\boldsymbol{S}_{sub}(\lambda)$ is the subset spectrum, $\boldsymbol{F}(\lambda)$ is the fitted curve (either $\boldsymbol{G}(\lambda)$ or $\boldsymbol{P}(\lambda)$), and N the number of available bands at current iteration.

If the residual error is below a certain threshold, the iteration is ended; otherwise, it uses the newly fitted function and repeats the previous steps (i.e. subsetting and re-fitting; Fig. 2). These steps for both Gaussian and polynomial functions are illustrated in Fig. 4.

The threshold (t) values for χ^2 are dependent on the dataset under study and is experimentally determined (see section 4.8). With quasi noise-free lab datasets, values between 1×10^{-4} to 3×10^{-4} give rise to optimum results. In practice, however, besides checking the χ^2 threshold, when N reaches 4 for Gaussian and 3 for polynomial functions, the iteration is terminated.

When the iteration number (n) exceeds 2, an index is calculated from the λ_{min} values of the last three iterations in the following way:

$$in = \frac{\lambda_{\min(n-2)} + \lambda_{\min(n)}}{2\lambda_{\min(n-1)}} \quad (8)$$

The value of this index normally varies between 0.6 and 1.4 in converging conditions. If the index acquire a value outside this range, the algorithm is reinitiated with a different (normally lower) value for r. The algorithm does this modification for four times and if the desired conditions are not met, the loop is exited.

Once the fitting meet the required conditions, the minimum of the function from the latest iteration is reported as the wavelength position of the absorption feature (λ_{min}; Fig. 4). The depth of the absorption (D) for a Gaussian is proportional to α (Eq. 1) and for a polynomial, it is achieved by replacing λ_{min} to the equation and subtracting from 1.0:

$$D = 1.0 - \boldsymbol{P}(\lambda_{min}) \quad (9)$$

Before initiating the iteration process, the method checks if an absorption exists in the given range by analyzing the explicit first and second derivatives of a higher (commonly fifth or sixth) order fitted polynomial (Fig. 2). The algorithm is executed if at least one minimum exists between the two local maximums of the fitted polynomial (see also section 4.4).

While a fixed second order polynomial is used within the IPF routine, a more robust solution is to let the order vary according to the available number of bands, the level of noise, and the residual error. Correspondingly, the IPF can start with higher order and decline gradually to second degree polynomial, as the minimum allowed. We coined the former IPF-fixed, and the latter, IPF-float routines.

The polynomial and Gaussian fittings are performed using linear and non-linear least-square techniques [26]. All the routines were developed using IDL 8.3 from EXELIS visual information solution (http://www.exelisvis.com).

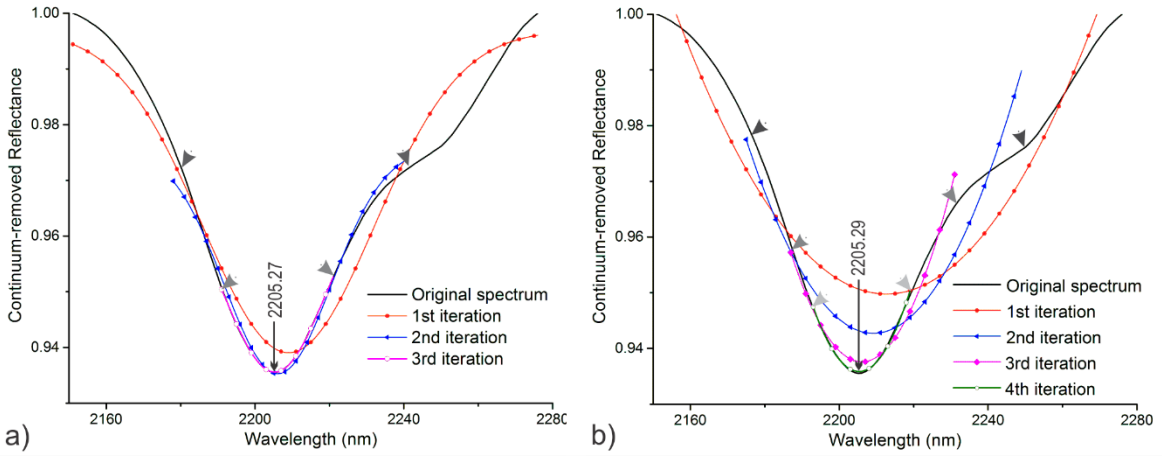

Fig. 4. Gaussian and polynomial fitted curves at different iterations to a continuum-removed clay (AlOH) absorption feature between 2151–2276 nm. **a)**. The IGF meets the threshold after 3 iterations and estimates the wavelength position (λ_{min}) at 2205.27 nm. **b)** The IPF demands 4 iterations for the same feature and estimates the wavelength position at 2205.29 nm. Arrow pairs indicate the subset of the spectrum used for fitting at each iteration. The threshold (t) used for both techniques was 1×10^{-4}.

4.3. Materials

We used a suite of 326 samples collected from Qom sedimentary basin in central Iran for experiments. The samples were selectively acquired from sandstone beds of the upper red bed formation. In the lab, the dried samples were measured using an ASD's FieldSpec-4 spectrometer (ASD Inc., Boulder, CO) under artificial illumination from a contact probe. This instrument covers the spectral range between 350–2500 nm with 1 nm sampling intervals. For each measurement, fifty individual scans were averaged to reduce the contribution of instrumental noise. The measured spectra were converted to absolute reflectance using a Spectralon white reference panel and, subsequently, were corrected for splice error between the three sub-systems. Four different spectral regions were selected from the collected spectra to evaluate the performance of the ICF methodology. The regions encompass wavelength ranges between 776–1074, 1845–2085, 2151–2276, and 2308–2387 nm, respectively related to ferric iron oxides and hydroxides, hydroxyl ion/adsorbed water, clays, and carbonates absorptions (Fig 5). The absorptions, which will be henceforth referred as ‘Feox’, ‘OH’, ‘AlOH’, and ‘Carbonate’, are virtually present in all samples; their approximate maximum depths occur at 900, 1900, 2210, and 2345 nm, respectively (Fig 5).

The collected spectra show slight noise, especially in the VNIR, which is probably due to the occurrence of magnetite or other dark minerals in sandstone matrix. While the noise was retained during the assessments of ICF algorithm, for direct determination of wavelength position, it was removed using a Savitzky-Golay convolution filter with a 17 nm smoothing window and second order polynomial.

This dataset was used to assess the performance of ICF relative to other fitting techniques (Fig. 1) and to compare the variants of ICF against each other. Additionally, it was used to analyze the

sensitivity and accuracy of the algorithm with varying noise, sampling intervals, spectral interpolation, and constant (e.g. r and t) levels.

An additive white Gaussian noise was systematically added to the dataset at different signal-to-noise ratios (SNRs). SNRs of 100, 150, 200, 250, 300, 400, and 500:1 were generated for this aim. The original dataset was also convolved to the spectral response of three different hyperspectral remote sensing instruments including (i) the ProspecTIR airborne system, with 357 spectral bands covering the wavelength ranges between 400–2450 nm (SpecTIR LLC, Reno, USA); (ii) the HyMap airborne system, with 126 spectral bands between 450–2500 nm [27]; and (iii) the upcoming Environmental Mapping and Analysis program (EnMap) satellite system, with 244 spectral bands spanning between 420–2450 nm [28].

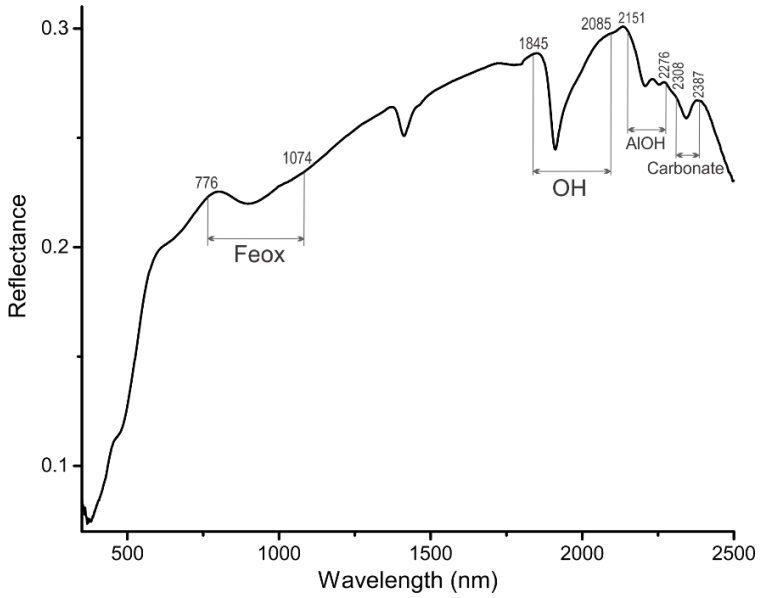

Fig. 5. A typical reflectance spectrum of sandstone with distinctive absorption features including ferric iron oxides/hydroxides (Feox), hydroxyl ion (OH), clays (AlOH), and carbonate. The features delimited within specific wavelength ranges were used to evaluate the performance of the ICF method.

4.4. Results and discussion

4.4.1. Iterative local continuum removal

We calculated the 'average' iterations required for removing the local continuum from OH, Carbonate, Feox, and AlOH features to be respectively equal to 1.03, 1.16, 1.22, and 1.3. This average may soar to 2.34 for more complex features like those yielded by goethite (not discussed in this paper). The method demonstrated to be highly successful with laboratory spectra, though, it encountered difficulties when it was applied to noise-added datasets (see section 4.7).

4.4.2. The consistency of the ICF technique

The wavelength position and depth estimated by IPF, IGF, QFM, and FP techniques were compared to similar parameters directly determined from the features following spectral smoothing. In the case of FP, a polynomial of second to sixth order was fitted to the features and the parameters were derived from the roots of the explicit first derivative. Except for the second order, all the others

were fitted to the entire range. The former was limited between the shoulders to avoid ill-fitting. The wavelength difference between each of these techniques and the direct method was calculated and used for accuracy assessment. In order to portray an intelligent depth parameter, the absolute difference was normalized against the depth of the direct method and expressed in percentage. The results are illustrated in Fig. 6.

Evidently, the performance of FPs of different orders are not identical and depends on the absorption under study. The polynomials of second to fourth order are shown to be inaccurate in at least three cases (Fig. 6a to c) with the estimated wavelength to be in error by as much as 29 nm (e.g. OH absorption; Fig. 6b). In the case of Feox and Carbonate (Fig. 6a and d), the wavelength position was underestimated, whereas for OH and AlOH features the same parameter was overestimated. An increase in the polynomial order has increased the estimation accuracy, but even the fifth or sixth order polynomial has not guaranteed the accurate simultaneous estimation of both parameters (e.g. Fig. 6a). The only feature successfully modeled by a FP (of sixth order) belongs to Carbonate (Fig. 6d).

In contrast, the IPF and IGF outperformed all of these techniques in terms of accuracy, precision and consistency. The root-mean-square error (RMSE) in wavelength estimation from both variant was better than 1.3 nm for all features (i.e. Feox: 1.3, OH: 0.45, Carbonate: 0.65, and AlOH: 0.33 nm) and their means were very close to zero, which signify their higher precision. For broad Feox absorption, however, our visual inspection revealed the results from IPF/IGF to be more viable than the direct method. Therefore, the RMSE of estimation by these techniques could be expected to be better than 1 nm. Correspondingly, both IPF and IGF methods yielded superior results in depth estimation (Fig. 6), with an average error not exceeding 1.5%.

As expected, the QFM technique achieved relatively similar results for the narrow absorptions (i.e. OH, AlOH, and Carbonate features; Fig. 6b to d), whereas for the broader Feox feature, it underestimated the wavelength minimum. This is evident by the RMSE of −20 nm in the determination of wavelength positions (Fig. 6a).

4.4.3. IPF versus IGF

Due to similarities in performance, the IPF and IGF techniques can retrieve very close results (Fig. 4 and 6). We used the same data of Fig. 6 to evaluate comparatively the performance of these methods. The results are summarized in Table 1.

There are strong correlations between the outcomes of IPF and IGF for all cases. The average difference in feature position and feature depth between them lies below 0.5 nm and ≈ 1%, respectively (Table 1). The IGF demands lower iteration to converge and its RMSE and χ^2 error are slightly lower for all features, which is probably due to the ability of the Gaussian function to better model the curvature of an absorption feature (Fig. 4). However, in their current architecture, the IPF is 4 to 17 (average 8.5) times faster than IGF, using identical computer hardware (Table 1). Indeed, this capability makes the IPF more suitable for massive data processing characteristic of remote sensing. Owing to this fact, only the IPF was evaluated in the remaining part of paper; however, we expect the conclusion to be equally valid for IGF as well. In different situations, we employed both IPF-fixed and IPF-float variants of the technique. Whereas they yielded similar results, the latter proved to be more straightforward when the feature was asymmetric.

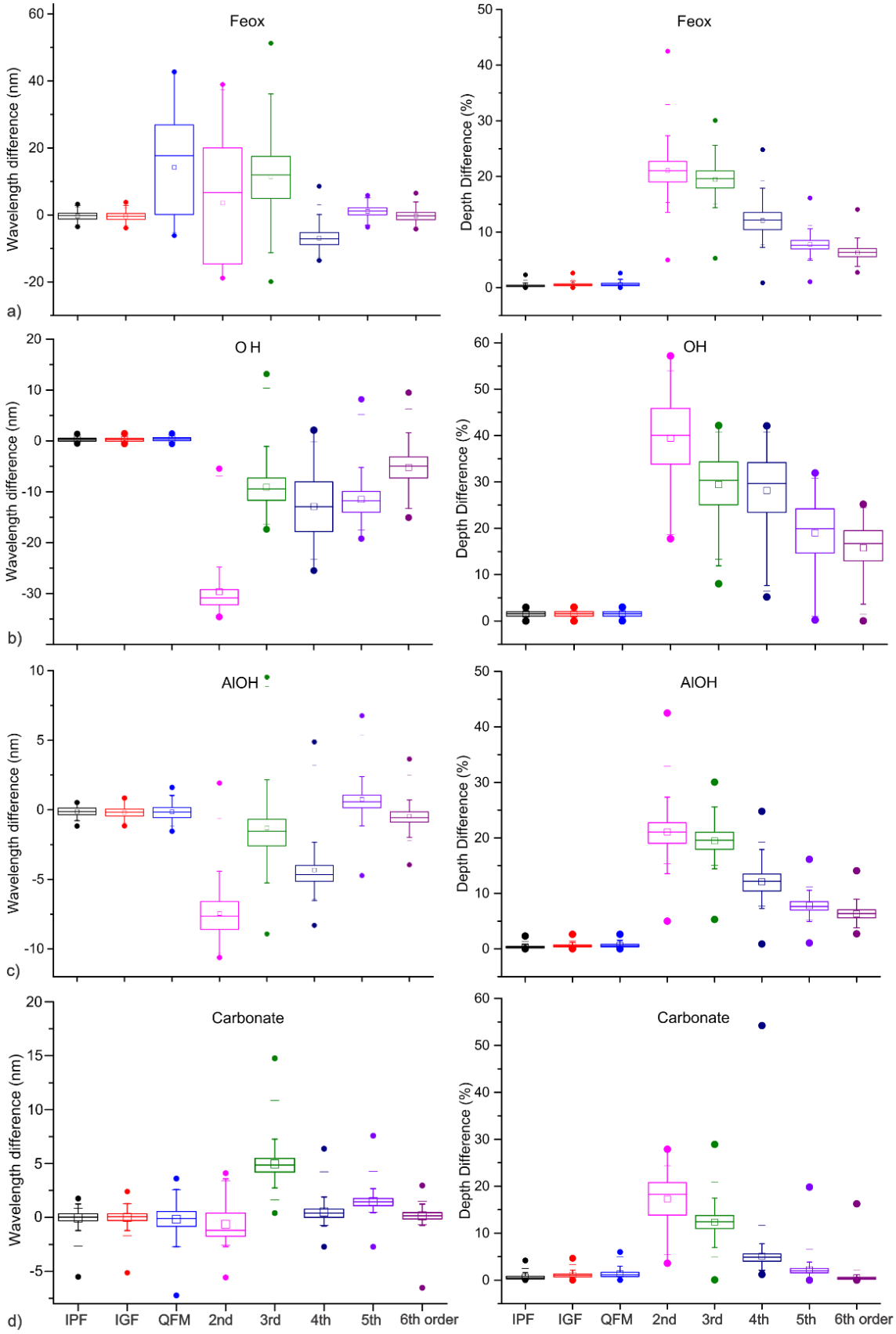

Fig. 6. Variability of the wavelength position and depth of selected absorption features **(a)** Feox, **(b)** OH, **(c)** AlOH, and **(d)** Carbonate. The left panel represents the wavelength difference (in nm) and the right panel the depth difference (in percent). The calculations are

performed relative to 'direct' method. The percentage is calculated using $\frac{|D-D_{direct}|}{D_{direct}} \times 100$, where D is the depth yielded from tested technique. The abbreviations corresponds to the legend in Fig. 1. The graphical elements of the box plots indicate the following statistical attributes: Box boundaries the 25th and 75th percentiles; the smaller square and the line bisecting the box the median and mean of the data, respectively; whiskers the 5th and 95th percentiles; smaller bar the 1st and 99th percentiles; and filled circles the maximum and minimum of the data series. In all settings, t was fixed to 2×10^{-4} and r was 0.25 and 0.2 for IGF and IPF, respectively.

Table 1. Comparison of the results obtained from IGF and IPF techniques over the selected spectral regions (Δ = Average difference).

Spectral region	Wavelength position		Depth		Relative processing time (IGF/IPF)
	R^2	Δ (nm)	R^2	Δ (%)	
Feox	0.992	0.46	0.999	1.05	4.13
OH	0.999	0.09	1.000	0.07	17.08
AlOH	0.996	0.06	0.999	0.28	6.44
Carbonate	0.952	0.19	0.999	0.97	5.91

4.4.4. The effect of spectral shape

As noted in section 2, a built-in procedure verifies if an absorption exists in a given spectrum in advance. As such, the algorithm is capable to tackle absorption features with highly diverse shapes. For instance, the IPF was successful in estimating the wavelength position of a suite of Carbonate features with varying levels of depth, width, and asymmetry (Fig. 7a). However, the ultimate fitting and the accuracy of the results are affected by the threshold (t) used for residual error (section 4.8; Eq. 7). A lower threshold may give rise to a better fit, but the wavelength determined on its basis is not necessarily a good representative of the absorption feature (e.g. feature # 3 in Fig.7a).

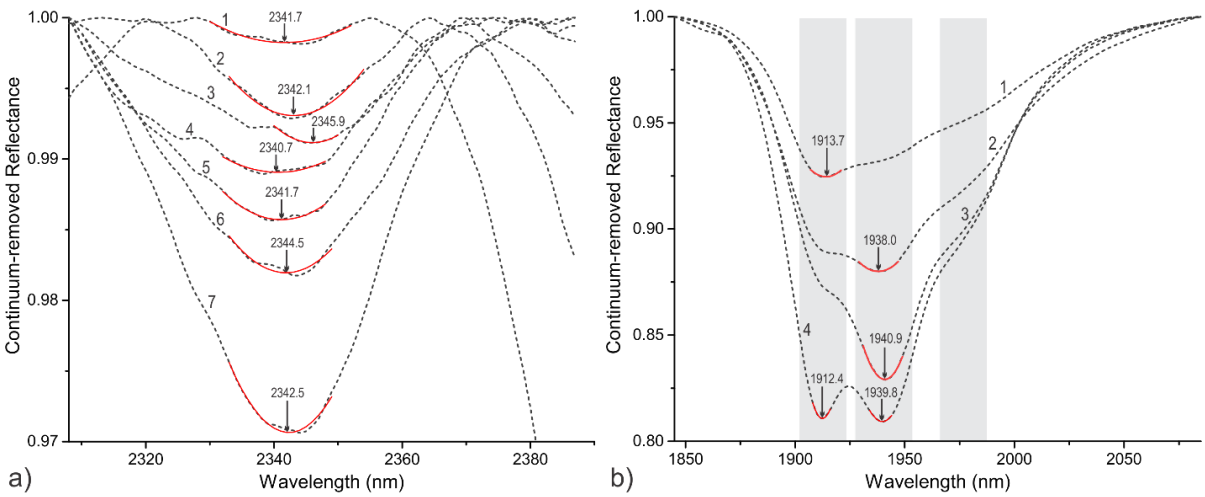

Fig. 7. The performance of IPF in estimating the wavelength position of various spectral features. **a)** Carbonate absorption features with varying depth, asymmetry, and width. **b)** OH features with multiple (triple) overlapping absorption bands defined by gray columns in the background (due to OH and adsorbed water of gypsum). The spectra are illustrated by dashed lines and the fitted polynomials by solid red. The arrows indicate the corresponding estimated minimum(s).

The core of the IPF/IGF algorithms can handle a single feature at a time. However, the algorithm also comprises complementary procedures that examine a spectrum for multiple absorption features. Firstly, during the analysis of an absorption feature (Fig. 2), if multiple absorptions are detected following the analysis of the derivatives of the fitted polynomial, the range is divided into multiple sub-ranges, each consisting of a single absorption. Secondly, if the constants of the final fit are not consistent with a minimum ($a_2 < 0$ for quadratic and $\alpha > 0$ for Gaussian; Eq. 1 and 4), then

the returned wavelength is regarded as a local maximum, the range is broken into two sub-ranges, and the algorithm is rerun for each of them separately (Fig. 7b).

Correspondingly, in the case of overlapping absorption features, the algorithms can resolve a feature if it is accompanied by a local minimum. For example, in the spectra shown in Fig. 7b, IPF can only resolve a 'doublet' for feature # 4, though in all cases at least three bands are overlapping. Accordingly, the performance of ICF techniques would not be identical to the deconvolution methods discussed earlier [9, 29].

4.4.5. Consistency in different sampling intervals

Because a hyperspectral sensor flown in an aircraft or satellite cannot sense the OH absorption between 1845–2085, we excluded this feature from our resampled dataset. The remaining three spectral regions were firstly resampled to the spectral resolution of ProspecTIR, HyMap and EnMap sensors, and then processed by the IPF algorithm for the estimation of wavelength position and depth. We interactively tested different initial parameters to find the ideal settings. The optimum parameters are shown on the caption of Figs. 8 to 10.

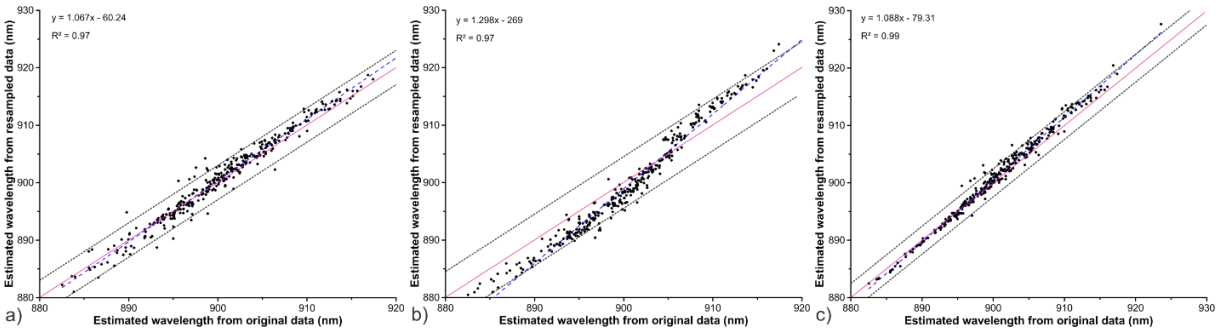

Fig. 8. The performance of IPF in estimating the wavelength position of Feox feature at different sampling intervals, considering the spectral resolution of **(a)** ProspecTIR, **(b)** HyMap, and **(c)** EnMap sensors. The dotted lines are, respectively, $\pm$ 3, $\pm$ 4.5, and $\pm$ 2.5 nm from the 1-to-1 line (solid magenta), as determined by WE95. The dashed (blue) lines represent the line best fitted to the data. The equation and R^2 of the fitting are reported on each plot. The settings (t, r) were equal to (4×10^{-4}, 0.1), (5×10^{-4}, 0.12), and (2×10^{-4}, 0.12), respectively.

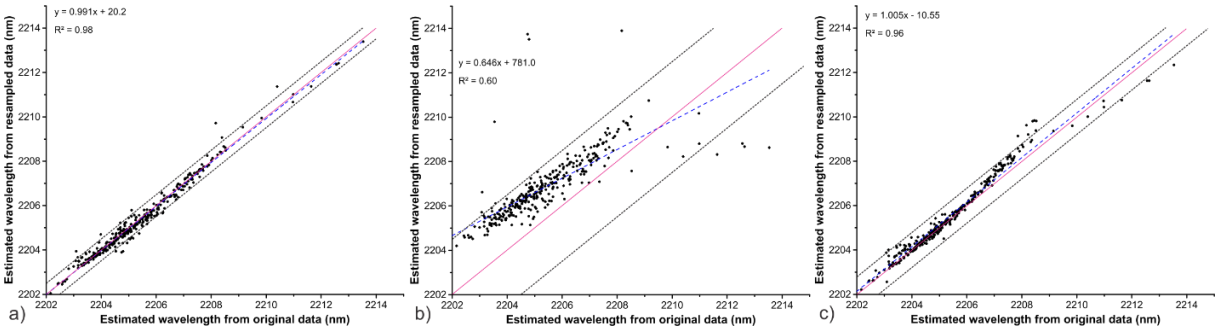

Fig. 9. The performance of IPF in estimating the wavelength position of AlOH feature at different sampling intervals, considering the spectral resolution of **(a)** ProspecTIR, **(b)** HyMap, and **(c)** EnMap sensors. The dotted lines are, respectively, $\pm$ 0.5, $\pm$ 2.4, and $\pm$ 0.8 nm from the 1-to-1 line, as determined by WE95. The settings (t, r) were equal to (2×10^{-4}, 0.1), (1×10^{-3}, 0.1), and (2×10^{-4}, 0.15), respectively.

To evaluate the results in terms of accuracy, a new parameter called wavelength position error at 95% confidence level (WE95) was defined. The WE95 means that the error in wavelength position is represented within the stated accuracy 95% of the times. A similar parameter named DE90 was

defined for the depth at 90% confidence level and expressed by a percentage relative to reference data (Table 2).

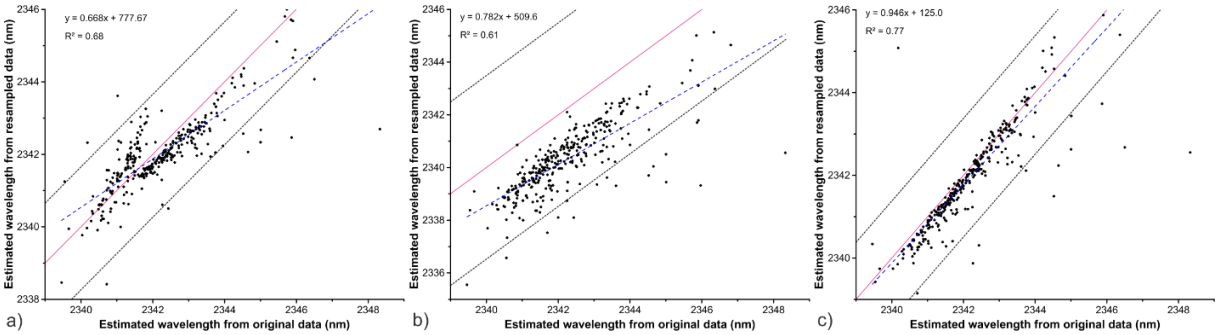

Fig. 10. The performance of IPF in estimating the wavelength position of Carbonate feature at different sampling intervals, considering the spectral resolution of **(a)** ProspecTIR, **(b)** HyMap, and **(c)** EnMap sensors. The dotted lines are, respectively, $\pm$ 1.7, $\pm$ 3.5, and $\pm$ 1.5 nm from the 1-to-1 line, as determined by WE95. The settings (t, r) were equal to (1×10^{-4}, 0.12), (1×10^{-3}, 0.3), and (2×10^{-4}, 0.2), respectively.

The results of the resampled Feox feature revealed the WE95 to be equal to 3, 4.5, and 2.5 nm for ProspecTIR, HyMap and EnMap, correspondingly (Fig. 8). The same parameter for AlOH was 0.5, 2.4, and 0.8 nm (Fig. 9), and for Carbonate 1.7, 3.5, and 1.5 nm, respectively (Fig. 10). All calculations were performed relative to the original higher resolution dataset reported in Fig. 6.

Table 2. The performance of IPF in estimating the depth of resampled Feox, AlOH, and Carbonate features at DE90 (%).

Sensor/Feature	DE90 (%)		
	Feox	AlOH	Carbonate
ProspecTIR	8.70	3.09	4.16
HyMap	6.11	14.10	16.46
EnMap	5.82	3.34	8.92

Evidently, the sampling regime of a sensor has pronounced effect on the accuracy of the IPF algorithm, because the highest WE95 values (e.g. 4.5 and 3.5 nm) are related to the wider sampling interval (17 nm) of the HyMap system. However, such relationship is not straightforward. For example, an increase in sampling interval from 7 to 17 nm decreases the accuracy of Feox estimation by 2 nm (Fig. 8), while in the case of AlOH, it gives rise to a decrease of only 1.6 nm (Fig. 9). Nevertheless, the sampling interval alone is not determining the accuracy of IPF technique, because with a constant sampling interval given by the ProspecTIR system (6 nm), different accuracy levels (between 0.5–3 nm) are achieved for different features. We believe the shape of an absorption is playing a role in the accuracy of wavelength determination. In this context, higher WE95 values for HyMap maybe as well related to the depth or asymmetry of the absorption features under study. Overall, the accuracy presented by the IPF algorithm is slightly better than or equal to the QFM method reported in [23].

A similar trend is observed in the depth estimation as well (Table 2); as long as the sampling interval is below 10 nm (e.g. ProspecTIR sensor), the accuracy at DE90 is below 9% (relative to the original data) and an increase of sampling interval diminishes the accuracy to around 17%. Nonetheless, the reduction rate as well seems to be dependent upon the type and characteristics of the absorption feature under study.

4.4.6. The effect of interpolation

The processing was repeated for HyMap-resampled data after applying a linear two point interpolation procedure as the optimum approach [29]. The results indicated that the WE95 of Feox could be improved by around 0.5 nm, whereas the WE95 of Carbonate declined by 2.2 nm. It seems that it is not appropriate to incorporate interpolation with the IPF algorithm, and when included, any possible improvement should be validated.

4.4.7. The robustness of IPF against noise

The IPF has a stable performance against added noise up to 100:1 SNR level. However, the RMSE, which was calculated using the difference between noise-free and noise-added signals, steadily increases with decreasing SNR. The RMSE of wavelength estimation is below $\pm$ 2.7 nm at SNR of 150:1 for narrow AlOH and Carbonate features in the SWIR region. Indeed, at 500:1 level, the technique is able to estimate the feature wavelengths to an RMSE of $\pm$ 1.6 nm (Fig. 11a). Similar to sampling interval, the effect of noise varies among features, with the Feox being the mostly affected feature by noise. The RMSE of wavelength estimation for this feature at SNR of 400:1 is equal to $\pm$ 4.2 nm. Such uncertainty is likely inherent to any broad absorption features (e.g. in the case of Feox, the FWHM is equivalent to 160 nm) and will probably arise in similar conditions.

The depth parameter estimated by IPF suffers from a serious error. For instance, at SNR of 150:1, the error expressed by DE90 is 57%, 19%, and 74%, respectively, for Feox, AlOH, and Carbonate features (Fig. 11b). The high error for the Carbonate feature may be connected to its comparatively shallow feature; however, in general, this error is introduced by inaccurate estimation of the continuum line. The noise tends to elevate the touching points of the continuum, thus giving rise to depth overestimation. Further improvement in the continuum removal methodology (section 2) is required to tackle this problem.

At each level of this experiment, a specific threshold for residual error (t) was interactively determined and applied to the algorithm (section 4.8).

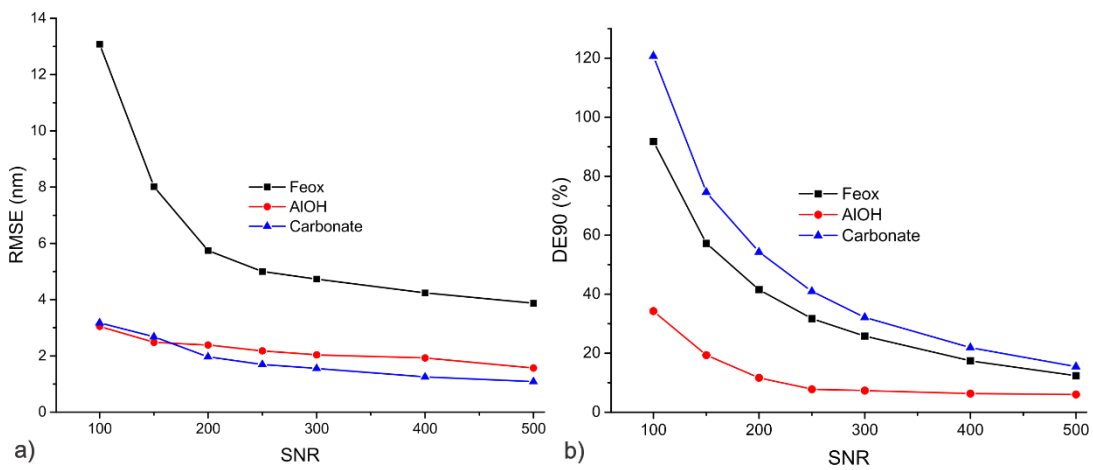

Fig. 11. The effect of SNR on wavelength position and depth determination of Feox, AlOH, and Carbonate features using the IPF (float) algorithm. **a)** Variation of RMSE of wavelength position against SNR. **b)** Variation of DE90 (%) against SNR. The RMSE was calculated from the difference between estimated values at specific noise level relative to original data.

4.4.8. Setting the constants

The two variants of ICF are sensitive to the range (r) and threshold (t) constant values (Eq. 5 and 7). These constants can influence the number of iterations, as well as the accuracy of estimations. A low value for r increases the likelihood of finding the best λ_{min}, but it will demand more iterations. In contrast, a high value for it will reduce the processing time, but likely, the algorithm will not find the optimum wavelength position or miss it altogether.

By using the parameters directly determined from the AlOH feature as a reference, we repeatedly changed the r values between $0.05 \leq r \leq 0.45$ and recorded the average number of iterations (from 326 spectra), as well as the average error between estimated and direct wavelength position using the IPF algorithm. The t value was set to 5×10^{-4} and remained constant in all attempts. We deliberately increased the threshold to let the effects of r to be easily seen. The plot of iteration and error against values of r is illustrated in Fig. 12. In the case of the AlOH feature, the optimum value for r using the IPF algorithm is 0.25. However, values between $0.1 \leq r \leq 0.3$ seems to be balanced. When we repeated the same assessment using the Feox feature and the IGF algorithm, relatively similar results ($0.15 \leq r \leq 0.35$) were obtained. For this reason, we set r for laboratory spectra equal to 0.2 and 0.25, respectively, for the IPF and IGF methods. The values used for resampled datasets (Figs. 8 to 10) were kept in the same range; though, we let them vary in each case to achieve the best results.

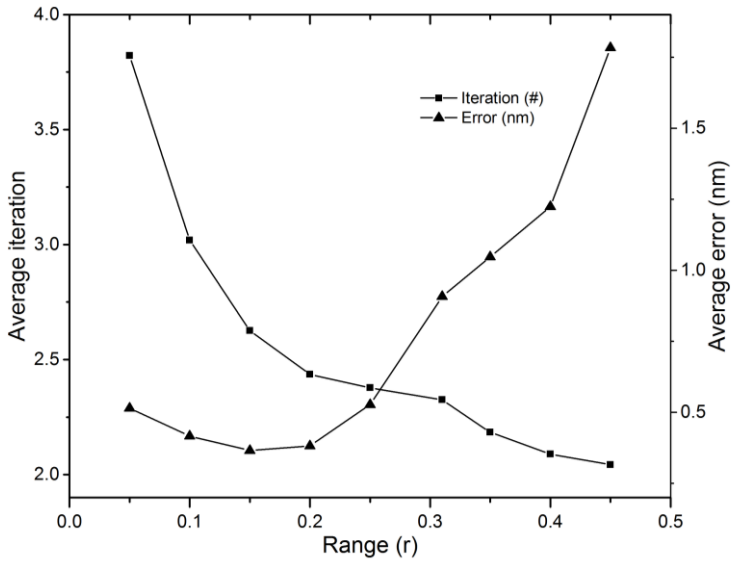

Fig. 12. Diagram of the average iteration and average error at varying range (r) levels. This plot was calculated using the AlOH feature and the IPF algorithm with a threshold set to 5×10^{-4}. The results are the average of all the 326 spectral features.

The threshold (t) that is used with residual error (Eq. 7) to leverage the accuracy and terminate the iteration has dependency on noise level, as well as the sampling interval of the dataset. An analysis of t against different SNR levels using the Feox feature and the IPF algorithm is shown in Fig. 13. Obviously, they are inversely related, because with a decrease in SNR, a higher threshold value is needed to maintain the accuracy. While this diagram can provide some hints on how to set the threshold, in practice, different values should be evaluated. For instance, our experiment with the Feox feature resampled to the ProspecTIR resolution indicated that a lower t does not necessarily give rise to greater accuracy (Fig. 8a). Ideally, an automated criterion could be used to estimate the

best initial t value directly from the dataset – a feature we shall develop in future versions of the code.

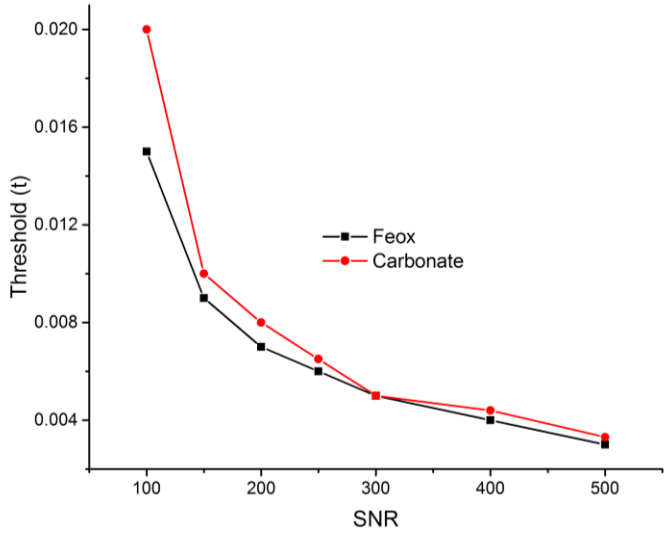

Fig. 13. The relationship between threshold (t) and SNR level using the Feox and Carbonate features and the IPF algorithm. The r constant was set to 0.2 in all experiments.

4.4.8. Other spectral parameters

The described techniques are embedded in a package preliminarily named Automated Absorption-based Mineral Spectral Analyzer (AMISA), in which beside wavelength position and depth, enables the calculation of FWHM, area, and asymmetry of absorption features. In the case of FWHM, its whereabouts is estimated by half the depth (D/2) value on both sides of λ_{min}; next, a line is fitted to each of these points and their immediate neighbors. The exact wavelengths are determined using the equations of these fitted lines, and then the absolute difference between the retrieved pairs is reported as FWHM. This strategy will ensure the accurate determination of the FWHM in varying sampling intervals and noise levels. The area of the absorption (A) is calculated by the trapezoidal summation technique. The asymmetry (S) is estimated using either area or width of the absorption (Eq. 10a and 10b), or the residual error of mirroring one-half of the absorption over the other half:

$$S_a = \frac{2A_{left}}{A_{all}} - 1 \quad (10a)$$

$$S_w = \frac{2W_{left}}{FWHM} - 1 \quad (10a)$$

where A_{left} is the area of the absorption from left shoulder to λ_{min}, A_{all} is the total area of an absorption, and W_{left} is the width of the absorption to the left of λ_{min} at half maximum depth [10]. The values of S_a and S_w range between -1.0 to 1.0, and in practice are strongly ($R^2 > 0.9$) correlated. In the case of mirroring, when an absorption is symmetric, the residual error is around zero and it escalates with increasing asymmetry.

The AMISA embodies means to analyze a given spectral data for the existence of an absorption and is able to return multiple absorption features. It is capable to process both spectral data from a spectrometer and HSI imagery and provide the spectral parameters as spreadsheet or image format.

4.5. Conclusion

A new methodology for the retrieval of wavelength position and depth from spectral data was proposed. Two variant of the technique including iterative polynomial and iterative Gaussian fitting (IPF and IGF) were introduced and tested. Instead of fitting the curve to the entire feature, these techniques attempt to minimize the fitting error merely around the desirable portion of the spectrum containing the minimum reflectance. The experiments proved that with high resolution data from a spectrometer, they can achieve accuracies better than $\pm$ 1.0 nm for the wavelength position and 1% for feature depth.

Although the IGF proved superior in modeling an absorption feature and gave rise to lower residual error, the IPF showed advantages in terms of processing speed. With IPF, the initial fitting helps reveal the approximate location of a minimum and the later iterations optimize the results. This algorithm unifies all the available polynomial fitting techniques into a single unit capable of handling different absorption features present in VNIR–SWIR and probably LWIR range. In fact, the QFM method discussed in [23] can be deemed as a special case of IPF on which the number of bands has reached down to three.

This method could be adapted to process hyperspectral imagery as well. Several parameters including the sensor's sampling interval, noise level, and the intrinsic character (shape) of the absorption feature, however, will contribute towards the accuracy of wavelength position and depth estimations. Our experiment indicated that when the sampling interval is better than 10 nm, the IPF could estimate the wavelength position of absorption features to within $\pm$ 3 nm at WE95. To maintain this accuracy level for narrow features, the imaging system is required to have SNR better than 150:1. For broad features, however, the SNR should be equivalent to 400:1 or higher. The experiment with Feox feature showed that at such level, the wavelength position could be estimated to an RMSE of $\pm$ 4.2 nm. The IPF is able to provide a proper estimation of depth from HSI dataset as well. The estimation accuracy was demonstrated to be better than 16.5% at DE90, but this could be profoundly affected by noise. To sustain this level of accuracy, a noise-resilient continuum removal methodology would be demanded.

The described ICF technique is embedded in a package called AMISA that as well allows the simultaneous calculation of width, area, and asymmetry. Further work is planned to equip this package with expert system rules to enable the identification of mineral species and predict possible mixtures using the quintuple spectral parameters.

Acknowledgement

We are thankful to Luis Guanter for the provision of EnMap spectral response functions. The exploration directorate of the National Iranian Oil Company is acknowledged for supporting the fieldworks. S.A. thanks CNPq and FAPESP (Proc. No. 2015/06663-7) for the research grants.

References

[1] S. Asadzadeh and C. R. de Souza Filho, "A review on spectral processing methods for geological remote sensing," *International Journal of Applied Earth Observation and Geoinformation,* vol. 47, pp. 69-90, 5// 2016.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing," *Signal Processing Magazine, IEEE,* vol. 19, pp. 44-57, 2002.

[3] R. N. Clark and T. L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," *Journal of Geophysical Research: Solid Earth,* vol. 89, pp. 6329-6340, 1984.

[4] B. H. N. Horgan, E. A. Cloutis, P. Mann, and J. F. Bell Iii, "Near-infrared spectra of ferrous mineral mixtures and methods for their identification in planetary surface spectra," *Icarus,* vol. 234, pp. 132-154, 2014.

[5] F. van der Meer, "Analysis of spectral absorption features in hyperspectral imagery," *International Journal of Applied Earth Observation and Geoinformation,* vol. 5, pp. 55-68, 2004.

[6] F. A. Kruse, A. B. Lefkoff, and J. B. Dietz, "Expert system-based mineral mapping in northern death valley, California/Nevada, using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sensing of Environment,* vol. 44, pp. 309-336, 1993.

[7] T. J. Cudahy, M. Jones, M. Thomas, C. Laukamp, M. Caccetta, R. Hewson*, et al.*, "Next generation mineral mapping: Queensland Airborne HyMap and satellite ASTER Surveys," *CSIRO Exploration & Mining Report P2007 / 364,* p. 152, 2008.

[8] F. A. Kruse, "Expert system analysis of hyperspectral data," in *Proc. SPIE 6966, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV*, Orlando, FL, 2008, pp. 69660Q-69660Q-12.

[9] J. M. Sunshine, C. M. Pieters, and S. F. Pratt, "Deconvolution of mineral absorption bands: an improved approach," *Journal of Geophysical Research,* vol. 93, pp. 6955-6966, 1990.

[10] K. Okada and A. Iwashita, "Hyper-multispectral image analysis based on waveform characteristics of spectral curve," *Advances in Space Research,* vol. 12, pp. 433-442, 1992/07/01 1992.

[11] R. N. Clark, G. A. Swayze, E. Eric Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton*, et al.*, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *Journal of Geophysical Research,* vol. 108, 2003.

[12] M. Berman, L. Bischof, and J. Huntington, "Algorithms and software for the automated identification of minerals using field spectra or hyperspectral imagery," in *Proceedings of 13th International Conference on Applied Geologic Remote Sensing*, Vancouver, British Columbia, 1999, p. 222–232.

[13] S. J. Gaffey, "Spectral reflectance of carbonate minerals in the visible and near infrared (0.35–2.55 um): Anhydrous carbonate minerals," *Journal of Geophysical Research: Solid Earth,* vol. 92, pp. 1429-1440, 1987.

[14] F. J. A. van Ruitenbeek, T. Cudahy, M. Hale, and F. D. van der Meer, "Tracing fluid pathways in fossil hydrothermal systems with near-infrared spectroscopy," *Geology,* vol. 33, pp. 597-600, July 1, 2005 2005.

[15] T. J. Cudahy and E. R. Ramanaidou, "Measurement of the hematite:goethite ratio using field visible and near-infrared reflectance spectrometry in channel iron deposits, Western Australia," *Australian Journal of Earth Sciences,* vol. 44, pp. 411-420, 1997.

[16] R. V. Morris, H. V. Lauer, C. A. Lawson, E. K. Gibson, G. A. Nace, and C. Stewart, "Spectral and other physicochemical properties of submicron powders of hematite (α-Fe2O3), maghemite (γ-Fe2O3), magnetite (Fe3O4), goethite (α-FeOOH), and lepidocrocite (γ-FeOOH)," *Journal of Geophysical Research: Solid Earth,* vol. 90, pp. 3126-3144, 1985.

[17] M. Haest, T. Cudahy, C. Laukamp, and S. Gregory, "Quantitative mineralogy from infrared spectroscopic data. I. Validation of mineral abundance and composition scripts at the Rocklea Channel iron deposit in Western Australia," *Economic Geology,* vol. 107, pp. 209-228, 2012.

[18] F. Tsai and W. Philpot, "Derivative Analysis of Hyperspectral Data," *Remote Sensing of Environment,* vol. 66, pp. 41-51, 1998.

[19] R. L. Huguenin and J. L. Jones, "Intelligent information extraction from reflectance spectra: Absorption band positions," *Journal of Geophysical Research: Solid Earth,* vol. 91, pp. 9585-9598, 1986.

[20] R. J. Murphy, S. Schneider, and S. T. Monteiro, "Consistency of Measurements of Wavelength Position From Hyperspectral Imagery: Use of the Ferric Iron Crystal Field Absorption at 900 nm as an Indicator of Mineralogy," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 52, pp. 2843-2857, 2014.

[21] E. Ramanaidou, M. Wells, D. Belton, M. Verrall, and C. Ryan, "Mineralogical and Microchemical Methods for the Characterization of High-Grade Banded Iron Formation-Derived Iron Ore," in *SEG Reviews*. vol. 15, S. Hagemann, C. Rosière, J. Gutzmer, and N. J. Beukes, Eds., ed: Society of Economic Geologists, 2008, pp. 129-156.

[22] T. E. Townsend, "Discrimination of iron alteration minerals in visible and near-infrared reflectance data," *Journal of Geophysical Research: Solid Earth,* vol. 92, pp. 1441-1454, 1987.

[23] A. Rodger, C. Laukamp, M. Haest, and T. Cudahy, "A simple quadratic method of absorption feature wavelength estimation in continuum removed spectra," *Remote Sensing of Environment,* vol. 118, pp. 273-283, 2012.

[24] J. K. Crowley, D. E. Williams, J. M. Hammarstrom, N. Piatak, I.-M. Chou, and J. C. Mars, "Spectral reflectance properties (0.4–2.5 μm) of secondary Fe-oxide, Fe-hydroxide, and Fe-sulphate-hydrate minerals associated with sulphide-bearing mine wastes," *Geochemistry: Exploration, Environment, Analysis,* vol. 3, pp. 219-228, 2003.

[25] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd Edition ed.: Cambridge University Press, 2007.

[26] C. B. Markwardt, "Non-linear least squares fitting in IDL with MPFIT," presented at the Astronomical Data Analysis Software and Systems XVIII, ASP Conference Series, Quebec, Canada, 2009.

[27] T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields, "The HyMap airborne hyperspectral sensor: The System, calibration and performance," presented at the First EARSeL Workshop on Imaging Spectroscopy, Zurich, 1998.

[28] H. Kaufmann, K. Segl, S. Chabrillat, S. Hofer, T. Stuffier, A. Mueller*, et al.*, "EnMAP A Hyperspectral Sensor for Environmental Mapping and Analysis," in *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, 2006, pp. 1617-1619.

[29] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 44, pp. 1601-1608, 2006.

Chapter 5

A new insight into the microseepage-induced alterations in redbeds through detailed spectroscopic analysis[6] [7]

According to microseepage theory, hydrocarbon reservoirs leak to the surface and trigger an array of diagenetic physicochemical and mineralogical changes in overlying sediments. The induced alterations have been widely used for microseepage targeting, albeit, the ambiguities associated with such indirect techniques have hindered its usefulness in facies determination. This ambiguity is partly due to gaps in our understanding of the diversity of changes in microseepage environments. In this article, an integrated spectroscopic approach was adopted to investigate the details of diagenetic mineralogical changes induced by active microseepage systems over Alborz oilfield in the Qom district, Iran. For this aim, a large suite of samples collected from petroleum-affected zones was spectrally characterized within 350–2500nm wavelengths and then contrasted with unaffected samples from a control area using statistical approaches. Spectral analysis indicated that smectites (montmorillonite, nontronite), secondary illite/smectite (Al-rich), and Fe-rich chlorites constitute the clay alteration in affected zones. The high iron content of the redbeds coupled with the arid climates prevented the sediments from complete bleaching. The chemical weathering of the relict irons contributed towards the formation of secondary goethite and ferrihydrite. Consequently, besides bleaching, modifications in the mineralogy of iron oxyhydroxides is proposed as an indicator for microseepage characterization. It was also observed that pyrite could be weathered to gypsum at the expense of carbonate dissolution. The resulting sulphates (gypsum ± jarosite) along with secondary carbonates constitute an additional indication of diagenetic facies. The geochemistry of the redbeds as revealed by infrared spectroscopy was verified using X-Ray Florescence (XRF) analysis. Subsequently, based on the silica contents of the host-rocks, the microseepage-induced alteration assemblages were classified into two broad categories namely felsic (SiO_2>60%) and mafic (45<SiO_2<60%) lithofacies wherein the latter is represented by the collection delineated in this study. This research has implications for the characterization of microseepage systems using spectral remote sensing.

5.1. Introduction

According to microseepage theory, the caprocks above hydrocarbon (HC) traps are not perfectly efficient and thus light gaseous HCs can migrate to the surface over time. The long-term leakage of HCs coupled with bacterial activities feeding on them triggers an array of physicochemical and mineralogical changes in the stratigraphic column above petroleum accumulations (Saunders et al. 1999; Schumacher 1996; Thompson et al. 1994). The bacterial and microbial oxidation of HCs and their by-products is believed to change the pH-Eh of the environment, thereby altering the mineralogy of the HC-affected soils and sediments. These changes broadly comprise clay alteration, carbonate and sulfide precipitations, and bleached facies (Al Shaieb et al. 1994; Donovan 1974). The array of changes in petroleum microseepage is considered to be highly significant for oil and gas exploration because it provides a context in which powerful surface exploration technologies could

[6] This chapter is under preparation to be submitted to the American Association of Petroleum Geologists (AAPG) Bulletin.

[7] A version of this chapter was presented as oral talk in the 18th Brazilian Remote sensing Congress (XVIII SBSR), Santos, Brazil.

be developed to help locate subsurface HC accumulations (Schumacher 2012). Moreover, any development in the methodology for microseepage characterization could benefit the environmental sciences, because it could be used to outline areas prone to methane emission to the atmosphere (Asadzadeh and Souza Filho 2017).

The majority of the microseepage-induced alterations retain diagnostic absorption features within the Visible-Near Infrared (VNIR) or the Shortwave Infrared (SWIR) wavelengths (Hunt 1977). Correspondingly, over the years, close-, and far-range spectroscopic techniques have been employed to characterize the mineralogic indicators associated with microseepage systems (Asadzadeh and Souza Filho 2017). Segal et al. (1986) used spectroscopy to detect diagenetic changes in Wingate Sandstone cropping out at Lisbon Valley, Utah. The unaffected bedrocks in this area were observed to contain large amounts of illite and mixed-layer clays and iron oxides. In contrast, the bleached facies was characterized by the removal of hematite and a large amount of kaolinite that was then postulated to be derived from the dissolution of mixed layer clays (i.e. illite). It was indicated that the high proportion of kaolinite in the area could approximate the limits of subsurface HC pools (Segal and Merin 1989). Other studies, however, have noticed a reverse trend in clay mineralogy. For instance, the transition from hematite towards more goethite compositions was noticed to be associated with the disappearance of the kaolinite doublet feature in favor of more illitic clays (Bowen et al. 2007; Petrovic et al. 2008).

Other studies have indicated a widespread carbonate cementation diagenetically formed over microseepage-affected areas (Petrovic et al. 2008; Schumacher 1996; Segal et al. 1986; Shi et al. 2012). The species of carbonates is reported to be highly diverse, varying from calcite, ferroan calcite, and dolomite to ferroan dolomite, ankerite, and siderite (Schumacher 1996). The dominant carbonate mineralogy was indicated to correspond to relative concentrations of Ca, Fe, Mg, and Mn ions in the environment; however, calcite (Ca) constitutes by far the dominant carbonate mineralogy in diagenetic facies (Al Shaieb et al. 1994). Whereas carbonates typically occur as pore-filling cement, there have been cases in which gypsum was entirely replaced by secondary calcite (Donovan 1974). In contrast to other analytic techniques, infrared spectroscopy alone is capable of differentiating between various types of carbonates minerals (Gaffey 1987), hence a useful tool for microseepage research.

Bleaching of ferric iron oxides is a common phenomenon observed in microseepage affected zones (Al Shaieb et al. 1994; Asadzadeh and Souza Filho 2017; Bowen et al. 2007; Donovan 1974; Fu et al. 2007; Schumacher 1996). This process involves the chemical dissolution of iron oxide coating (primary hematite) from redbeds through the reduction of ferric (Fe^{3+}) iron into ferrous (Fe^{2+}). The released iron is either removed from the system or is embedded into ferrous-bearing minerals comprising sulfides and ferroan carbonates. The bleaching accounts for most of the visual color changes in microseepage-prone areas, though is not exclusively correlative. Because spectroscopy is highly sensitive to iron oxide content (Balsam et al. 2014), it provides the best tool to delineate any changes in the mineralogic contents of iron oxides in altered facies.

Even though spectroscopy has been incorporated in many microseepage remote sensing surveys, the technique has been utilized only to verify anomaly maps derived from multi-, and hyperspectral remote sensing through analysis of a handful of samples collected from affected zones (see Asadzadeh and Souza Filho (2017)). Thus far, very few studies have systematically investigated the

diagenetic changes in soils and sediments using reflectance spectroscopy and, consequently, the potential of this powerful tool for microseepage characterization has remained rather underexplored. Spectroscopy is a fast, non-destructive and cost-effective analytic tool. It is not only capable of identifying diagenetic mineralogy in microseepage systems, but also offers great potential for characterizing the physicochemistry and composition of minerals with reasonable accuracy. Spectroscopic outcomes can be readily integrated with remote sensing image data to collectively characterize the distribution, proportion, and composition of diagenetic minerals at outcrop to reservoir/basin scales and thereby facilitate the correlation of remotely sensed anomalies to microseepage phenomenon.

In this article, we adopted an integrated spectroscopic approach to investigate the details of the mineralogical variations induced by an active microseepage system over the Alborz oilfield located in the Qom district, Iran. For this aim, a large suite of samples systematically collected from affected zones was spectrally characterized and then were contrasted with unaffected samples from a control area using a statistical approach.

5.2. Geologic setting of the study area

The Qom study area is located near the city of Qom, some 100 km to the south of Tehran, Iran, and encompasses an area of about 1400 km^2. The area coincides with the western edges of a back-arc basin known as the Central Basin of the Iranian Plateau. It includes a proven reservoir named ‘Alborz’, which occurs in an anticlinal trap with axial length and width of ~50 and 6–8 km, respectively (Fig. 1). The Alborz oilfield, which was only productive for a short duration, has potential reserves that are estimated at 20–100 MMBLO (million barrels of crude oil). The reservoir consists of limestones of the Qom Formation sealed by thick evaporitic sequences of the late Miocene and buried beneath >2500 m of Oligocene sediments of the Upper-Red Formation (URF) (Berberian and King 1981; Morley et al. 2009). The URF, which hosts the Alborz microseepage system, contains a thick sequence of interbedded sandstone, siltstone, conglomerate, and mudstone (collectively called sandstone henceforth), with local layers of marl, shale, and gypsum (Amini 2001). The thickness of the units varies from millimetric laminated strata to beds of several meters. The URF sandstone is largely a lithic arenite as it is predominantly (>50%) composed of lithic fragments (Fig. 2a). Over the reservoir, the URF beds are commonly friable and thus have been very prone to erosion (Fig. 2b).

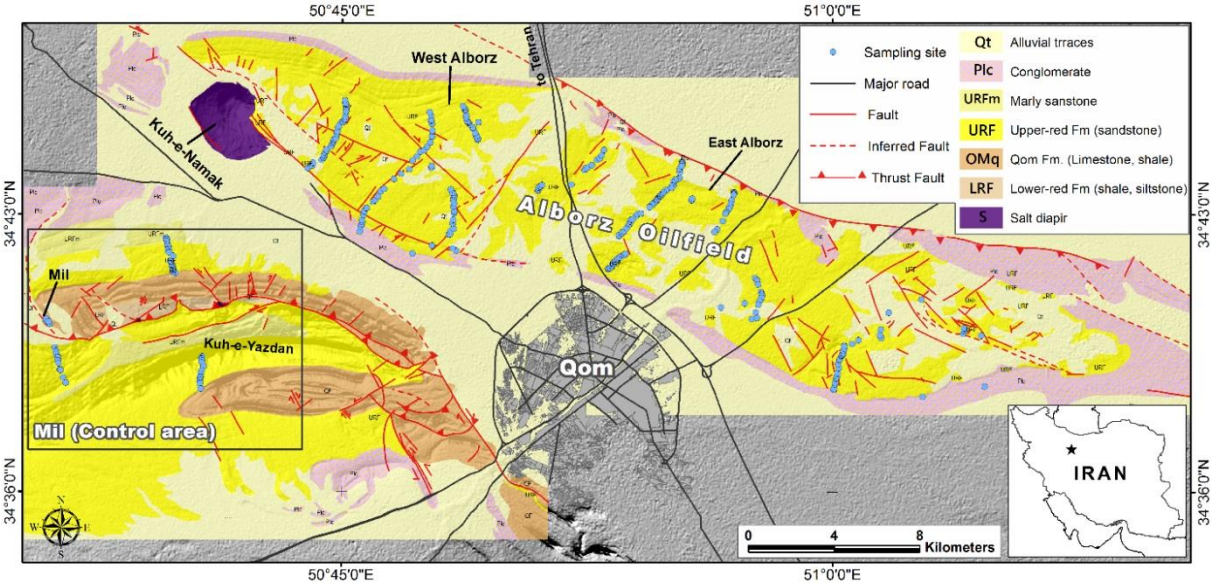

Fig. 1. Generalized geologic map of the study area superimposed on a shaded relief digital elevation model. The sampling sites are shown by cyan circles along 14 profiles crossing the Alborz oilfield and the Mil (control) area.

Fig. 2. Photographs showing sandstone beds of the Upper-Red Formation (URF) in the Qom area. **a)** The unaffected sandstones over the Mil area off-field are hard and reddish brown. **b)** The URF beds overlying the Alborz oilfield are commonly yellowish gray, friable, and prone to cavities.

5.3. Materials and methods

Fieldwork and sampling were carried out along carefully planned profiles perpendicular to stratigraphic variations (Fig. 1). To draw a fair comparison, we covered off-field areas with unaltered bedrocks in parallel with altered zones associated with the Alborz oilfield (Fig. 1). The off-field control area (also referred to as Mil area) was used to define the mineralogic background and cross-validate the diagenetic changes observed over HC-affected zones. Overall, we collected about 340 samples (280 from Alborz + 60 from Mil) along 14 profiles, resulting in an average of one sample per ∼120 m profile length (Fig. 1). The samples covered a variety of lithofacies including fresh cuts from sandstone beds, siltstones, shales, conglomerates, and gypsum as well as the soil covers. The majority of the samples were collected from exposed areas, but vertical faces were also sampled.

In the Lab, the entire sample suite was spectrally measured using a FieldSpec-4 spectrometer with 1 nm sampling intervals between the 350–2500 nm range. The measurements were performed under artificial illumination of a contact probe; and then the acquired data were converted to reflectance using a Spectralon panel. For each record, 50 individual scans were averaged to suppress the contribution of instrumental noise. Consequently, >2400 spectra (2100 from Alborz + 320 from Mil) were collected and integrated into two separate spectral libraries for Alborz and Mil areas. The residual splice drift of the spectrometer was corrected by extrapolating the position of the channels immediately before and after the SWIR1 detector, using a pair of channels on either side of the array. The offsets were then applied to the VNIR and SWIR2 channels one-by-one to yield uniform spectral data over the entire range. The data were ultimately converted to absolute reflectance using the response of the Spectralon reference (Clark et al. 1990).

All the noted corrections and subsequent processing were performed using the AMISA (Automated Absorption-based Mineral Spectral Analyzer) package developed in the Interactive Data Language (IDL, version 8.3) program. The package calculates the wavelength position of a given absorption feature (called ‘wavelength position’ henceforth) along with its depth, width, area, and asymmetry in a fully automated fashion (Fig. 3) (Asadzadeh and Souza Filho 2016). For a user-specified spectral range, the algorithm fits a local continuum and then divides the data by it to yield a continuum-removed spectrum. When there is no absorption based on some predefined criteria, it returns no values for the spectral parameters. In case more than one absorption feature is present in the specified range, the algorithm returns multiple values for every spectral parameter (Asadzadeh and Souza Filho 2016).

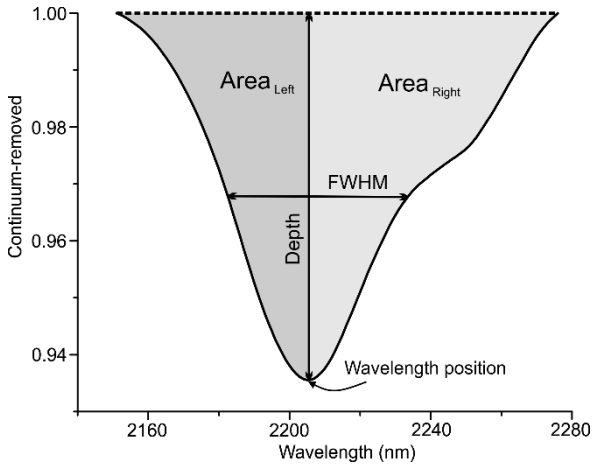

Fig. 3. Spectral parameters retrieved from each absorption feature following a local continuum removal using the AMISA code: wavelength position of the absorption feature, depth of the absorption, the full-width at half maximum depth (FWHM) or simply the width, and areas to the left and right side of the minimum (used to calculate the absorption asymmetry). For details, see Asadzadeh and Souza Filho (2016).

Visual inspection of the compiled library indicated that nine diagnostic absorption features are present in the VNIR–SWIR window, as illustrated in Fig. 4 and listed in Table 1. A spectral range was carefully defined for each absorption feature (see Table 1) and then used as inputs into an iterative polynomial fitting (IPF) algorithm to retrieve the desired spectral parameters. The threshold to end the iterations was varied from 1.5×10^{-4} to 2.5×10^{-4} (Table 1). For the ferrous iron feature, however, a fixed two-iteration polynomial fitting (PF) was employed to estimate the wavelength position to within an accuracy of ± 18 nm at 95% confidence level. To make the interpretation easier, the spectra of some dark samples (e.g. rich in manganese oxides) and a few malachite-bearing specimens were omitted from the collection. The outcomes of AMISA was a spreadsheet containing the noted spectral parameters for every individual measurement. The data provided by this procedure were then used to create histograms and scatter plots to address the inter-, and intra-relationships between different spectral parameters of the diagnostic absorption features. An example of a histogram generated in this way is illustrated in Fig. 6.

Table 1. The major absorption features detected in the spectral data and mechanism of absorption. The fitted range defines the spectral range along which a local continuum is calculated and removal. RMSE denotes the thresholds used to end the iterative polynomial fitting (IPF). The last column represents the accuracy of wavelength position estimation at 95% confidence level.

Feature	Absorption mechanism	Fitted Range (nm)	Methodology	Threshold (RMSE)	Accuracy (nm)
Ferric iron (Fe^{+3})		600–800	IPF	1.7×10^{-4}	±1
Feox (Ferric oxides)	Crystal Field transition	776–1074	IPF	2.5×10^{-4}	±1.5
Ferrous iron (Fe^{+2})		1000–1370	PF	-	± 18
OH/ H_2O	OH stretch (1st overtone) H_2O stretch	1364–1560	IPF	2.0×10^{-4}	±1
H_2O	H_2O stretch + bend	1845–2060	IPF	2.0×10^{-4}	±1
Al–OH	OH stretch + Al–OH bend	2151–2276	IPF	2.0×10^{-4}	±1
Fe–OH1 (Chlorite)	OH stretch + Fe–OH bend	2231–2274	IPF	1.5×10^{-4}	±1
Fe–OH2 (Nontronite)	OH stretch + Fe–OH bend	2276–2306	IPF	1.5×10^{-4}	±1
CO_3 (Carbonates)	C–O stretch (2nd overtone)	2308–2387	IPF	2.0×10^{-4}	±1

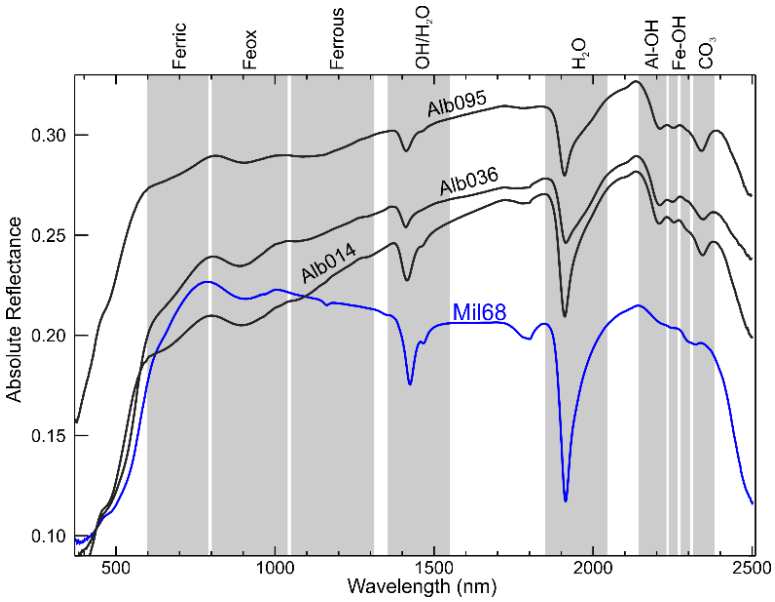

Fig. 4. Representative examples of spectra collected over the Alborz oilfield and characteristics of the absorption features in the VNIR-SWIR window. Gray bars indicate the spectral ranges analyzed by the AMISA (see Table 1 for the details). The names given to each feature (shown on top of the columns) are adopted to present the results. A spectrum from the Mil library (solid blue) is included to highlight the difference in the pattern and albedo of the spectra.

To verify the geochemistry of the redbeds as revealed by spectroscopy, we performed X-Ray Florescence (XRF) analysis to a subset of 36 samples (28 for Alborz + 8 for Mil) aiming to determine the major-oxides compositions of the sandstones. The results are listed in Table 2.

5.4. Results and discussion

Visually speaking, the originally dark reddish gray sandstones (10R, 4/1 based on Munsell rock color chart) display varying color changes that range from light gray (5Y, 7/1), due to the bleaching of iron oxide (hematite) coatings, to pale yellow (5Y, 8/2), related to the neomineralization of iron oxyhydroxides or changes in the abundance of ferric minerals (Fig. 5). The changes in color are controlled by the iron oxide content, and the fabric of the original units. Beds with higher iron concentrations marked by dark reddish color (Fig. 5a) are transformed into moderate yellow facies (Fig. 5c), whereas beds with lower iron concentrations (Fig. 5b) are altered to light gray beds (Fig. 5d). The color transformation is also associated with changes in the morphology of the rocks. The hard and uniform strata (Fig. 2a) becomes friable and porous, hence very prone to erosion (Fig. 2b). This geomorphological change is presumably due to the dissolutin of sadstone cements.

Table 2. Major-oxide compositions (weight %) from XRF analysis for unaltered (Mil) and altered (Alborz) facies in the URF sandstones.

Sample	SiO_2	Al_2O_3	Fe_2O_3	TiO_2	MnO	MgO	CaO	Na_2O	K_2O	P_2O_5	LOI	SO_3	Total
Mil-25	49.43	17.07	6.41	0.720	0.115	1.93	6.92	7.14	0.77	0.113	8.74	-	99.4
Mil-31	48.05	16.53	6.40	0.674	0.108	2.78	7.20	7.14	0.76	0.114	9.79	-	99.5
Mil-55	49.59	18.36	6.34	0.687	0.108	2.83	5.50	8.45	0.43	0.114	7.88	-	100.3
Mil-61*	36.79	8.72	3.54	0.347	0.053	1.62	23.43	1.11	1.52	0.093	22.29	-	99.5
Mil-64	47.25	17.33	7.57	0.732	0.137	3.68	4.77	8.19	0.67	0.132	9.93	-	100.4
Mil-73	49.26	16.50	7.49	0.816	0.120	2.66	6.26	6.34	0.73	0.148	9.45	-	99.8
Mil-84	48.70	18.14	7.99	0.674	0.142	3.17	7.25	5.50	0.34	0.133	7.64	-	99.7
Mil-92	50.09	18.27	6.40	0.571	0.112	2.77	4.97	9.15	0.38	0.071	7.06	-	99.8
Alb-16	54.28	12.33	3.74	0.437	0.080	1.18	11.15	3.22	1.47	0.099	11.30	-	99.3
Alb-25	51.80	11.33	4.18	0.451	0.102	1.80	13.22	2.27	1.48	0.082	13.18	-	99.9
Alb-29	53.49	12.32	4.10	0.535	0.088	1.44	11.83	3.20	1.40	0.107	10.97	-	99.5
Alb-32	51.20	11.83	3.79	0.503	0.096	1.23	13.54	3.04	1.29	0.116	12.68	-	99.3
Alb-38	50.98	10.71	3.91	0.500	0.116	1.18	14.13	3.49	1.06	0.066	13.64	-	99.8
Alb-61	55.76	10.25	3.02	0.358	0.106	1.10	12.67	3.06	1.24	0.043	12.35	-	100
Alb-73	55.81	11.84	4.30	0.499	0.081	1.36	10.43	3.37	1.31	0.098	10.13	-	99.2
Alb-80	52.31	10.75	3.82	0.469	0.115	1.05	13.47	3.76	1.03	0.048	13.07	-	99.9
Alb-85	48.17	9.93	3.67	0.461	0.108	1.20	16.27	3.17	1.02	0.063	14.66	1.16	99.9
Alb-88	54.22	11.23	3.75	0.428	0.097	1.68	12.33	3.42	1.19	0.048	11.52	-	99.9
Alb-110	50.01	9.33	3.43	0.517	0.153	1.38	16.34	1.97	1.09	0.099	15.17	-	99.5
Alb-130	48.53	15.26	7.07	0.624	0.136	2.35	9.22	6.98	0.50	0.091	9.11	-	99.9
Alb-131	51.29	10.71	4.62	0.557	0.116	1.14	13.64	3.35	1.07	0.083	13.31	-	99.9
Alb-150	53.80	10.70	4.46	0.544	0.090	1.44	12.49	2.54	1.37	0.059	12.37	-	99.9
Alb-156	50.06	14.51	5.62	0.593	0.103	2.09	9.51	5.17	1.01	0.112	10.73	-	99.5
Alb-162	51.45	11.89	4.51	0.542	0.103	1.27	12.60	3.35	1.41	0.105	12.29	-	99.5
Alb-178	34.39	7.56	2.90	0.335	0.059	1.64	17.54	5.73	1.04	0.040	16.07	12.44	99.7
Alb-186	45.94	9.73	3.41	0.449	0.129	1.10	17.50	2.95	0.98	0.053	16.10	0.12	98.5
Alb-194	55.78	11.31	3.27	0.366	0.076	1.11	11.78	3.80	1.20	0.046	11.13	-	99.9
Alb-198	52.51	10.78	4.75	0.479	0.104	1.62	12.54	3.87	0.96	0.059	12.23	-	99.9
Alb-201	40.90	9.41	6.67	0.897	0.154	0.98	17.84	2.92	0.85	0.049	12.95	5.85	99.5
Alb-217	61.99	10.96	4.69	0.677	0.096	1.17	7.91	3.12	1.68	0.097	7.18	-	99.6
Alb-231	54.13	10.98	3.77	0.462	0.093	1.26	12.17	3.49	1.21	0.065	12.24	-	99.9
Alb-238	50.37	10.74	3.99	0.498	0.116	1.26	14.54	3.18	1.08	0.061	14.04	-	99.9
Alb-243	48.86	9.99	6.35	0.927	0.192	1.32	14.80	3.00	1.05	0.052	13.36	-	99.9
Alb-253	51.68	10.76	6.75	1.038	0.115	1.63	12.81	2.51	1.26	0.114	10.56	-	99.2
Alb-257	46.66	10.17	3.06	0.336	0.083	0.97	15.47	3.76	0.93	0.032	12.00	6.4	99.9
Alb-282	48.69	10.02	4.11	0.564	0.151	1.66	16.08	2.61	0.94	0.056	14.99	-	99.9
Mean: Mil	48.91	17.46	6.94	0.70	0.12	2.83	6.12	7.42	0.58	0.12	8.64	-	-
Mean: Alb	50.41	10.90	4.32	0.53	0.11	1.39	13.70	3.36	1.16	0.07	12.81	5.19	-

* This sample was collected from the alteration halo around a macroseep, hence was considered along with Alborz samples.

5.4.1. Diagenetic mineralogy within the VNIR range

The original redbeds in the Mil area are characterized by a broad absorption feature centered at ~880 nm, corresponding to hematite coating (see Fig. 5a and 6), whereas the specimens collected from petroliferous zones are mainly associated with a shift towards longer wavelengths (between ~890–930 nm) that is interpreted to be representative of goethite (Sherman et al. 1982) (Fig. 6a). The goethite-rich facies, which is exclusive to petroliferous areas, is presumably linked to the transformation of hematite to goethite in the environment (Schwertmann 1971). In the reducing condition imposed by microseeps, hematite is known to dissolve by transformation into ferrous iron. In vadose zones, in which both oxygen and percolating meteoric water are present in high quantities, goethite (FeOOH) can precipitate from ferrous iron through oxidation and hydrolysis (reaction 1) or by the conversion of other metastable oxyhydroxides (Cornell and Schwertmann 2003).

$$Fe^{2+} + ¼\ O_2 + 1.5\ H_2O \rightarrow\ FeOOH\ + 2H^+ \quad (1)$$

The pervasive and evenly distributed goethite over the petroleum deposit indicates that the reducing environment near the surface has been overprinted by oxidation promoted by chemical weathering. The dominance of goethite could also be indicative of a high concentration of carbon-rich organic compounds in the environment (Schwertmann 1971).

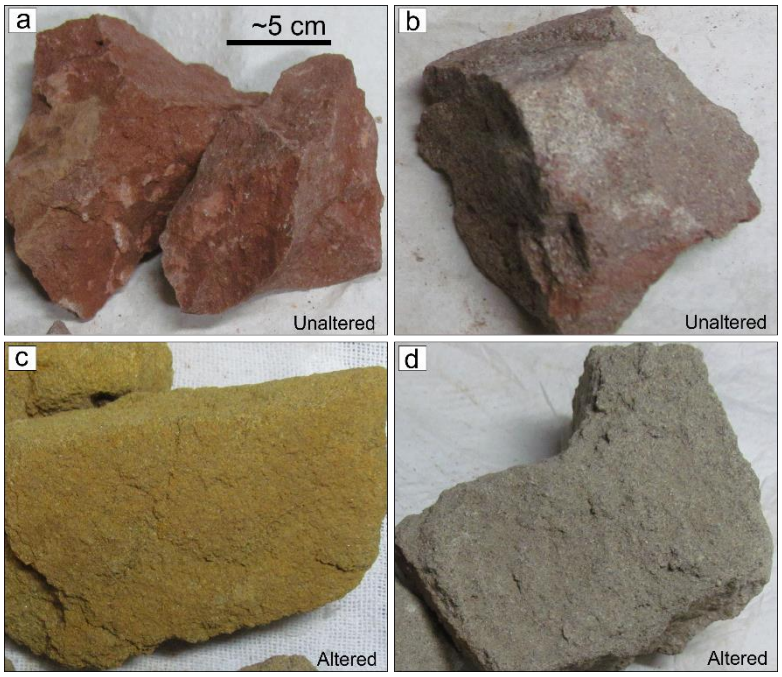

Fig. 5. Photographs exhibiting the variations in color of the URF sandstones in hand specimens. **a-b)** unaltered, and **c-d)** altered bedrocks. The changes in color are controlled by the iron oxide content (and the fabric) of original units. Beds with higher iron concentrations, marked by dark, reddish color (a), are transformed into a yellowish rock (c). Beds with lower iron concentrations (b) are altered to light gray facies (d). The scale in (a) applies to all photographs.

An alternation between oxidation-reduction phases is reported to favor the formation of maghemite in soils (see Ellwood and Burkart (1996) and the references therein). Maghemite and ferrihydrite display distinctive absorption features respectively at ~955 and ~915 nm (Sherman et al. 1982). The wavelength positions exhibited in Fig. 6a is confined to 925 nm, hence are not evidential of maghemite presence. However, when ferrihydrite is present, most likely it is physically mixed with goethite or hematite (Cornell and Schwertmann 2003) and thereby, is not easily distinguished by reflectance spectroscopy. We reckon that the formulation of ferrihydrite proportion in intimate mixture with goethite requires some further spectral simulation experiments, which is out of the scope of this work.

Goethite is also characterized by a distinctive absorption feature centered at ~668 nm (Sherman et al. 1982). The wavelength positions observed within the 600–800 nm window, however, extend across a wider range, thereby implying that the features are not solely related to goethite (Fig. 7a). Apart from goethite, four additional minerals are known to have diagnostic absorption features between 600–800 nm wavelengths: nontronite (at ~650 nm), hematite (subtle feature below 660 nm), jarosite (at ~665 nm), and Fe-chlorite/chamosite (at ~720 nm) (Bishop et al. 2008; Hunt and Ashley 1979). Correspondingly, the features identified below 660 nm could be related to nontronite and/or hematite and those beyond 690 nm to chlorite. Since nontronite constitutes a subtle

component of the samples (interpreted from the SWIR data discussed in section 4.2), it is unlikely that its iron feature contributes much in this range; therefore, the wavelength positions should be assigned to hematite (or a mixture of hematite + goethite). The jarosite feature overlaps with goethite in the 900 nm range and, thus, is not distinguishable in the histogram. Seemingly, there are intermediate phases that arise from a physical mixture of these minerals, particularly between goethite and chlorite. Goethite by far is the most dominant mineral in this range, for the mean of the histogram (μ = 675 nm) closely corresponds to its wavelength position (Fig. 7a). It is worth mentioning that due to the presence of carbonates, lepidocrocite (the polymorph of goethite) was considered to be unlikely to form in the system (Cornell and Schwertmann 2003) and, consequently, was dismissed from the spectral processing.

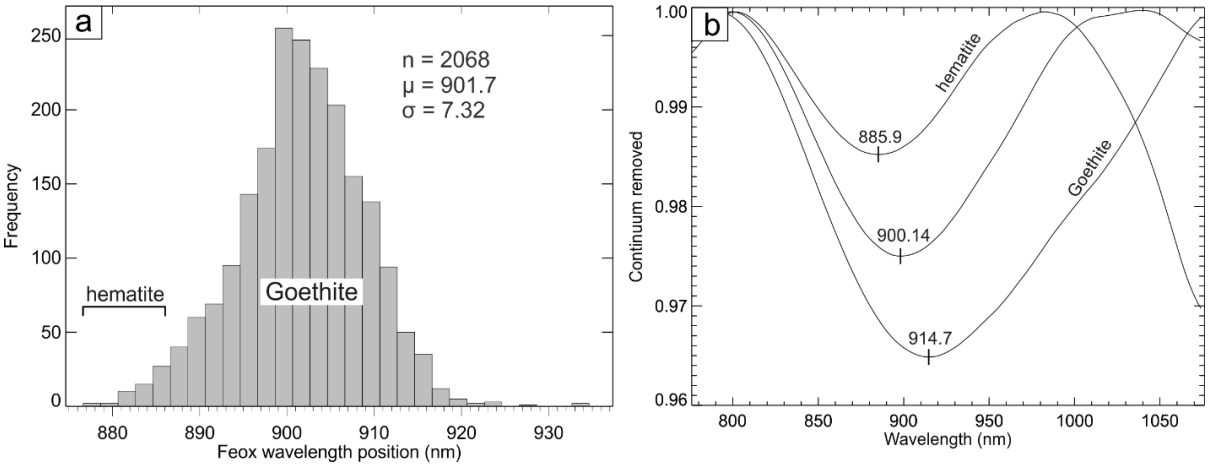

Fig. 6. a) Histogram of the wavelength position of Feox (ferric oxides) absorption feature extracted from the Alborz spectral library. n, μ, and σ represent the number of contributed features, the data mean, and standard deviation, respectively. **b)** Continuum-removed spectra typifying short (hematite-rich), medium, and long (goethite-rich) endmembers of the Feox absorption feature wavelength position. The numbers denote the calculated wavelengths positions. A Sav-Gol filter was used to smooth the spectra.

Aluminum can substitute both hematite and goethite and shift their wavelength position by up to 30 nm (from 860 to ~875 nm for hematite and from 895 to ~930 nm for goethite) (e.g. Buckingham and Sommer (1983)). On this basis, we deduced that primary hematite in the redbeds is Al-substituted by as much as 20 mole percentage aluminum. Regarding the secondary goethite, it is not straightforward to estimate its substituted Al percentage; however, the restricted range of goethite wavelength position (confined between 900–925 nm ranges) implies that in the case of substitution, the amount should be very low. In fact, if we consider the wavelengths between 900–910 nm (μ + 1σ; see Fig 6a) to be representative of goethite in this area, then the Al-substitution results to be <5 mole percent (Buckingham and Sommer 1983). In other words, the secondary goethite is ordered and well-crystalline. Further investigation is required to develop Al-substitution in goethite (and hematite) into an indicator for the discrimination of secondary iron oxides and indeed for microseepage characterization.

In a microseepage system similar to Alborz, wherein iron is in excess (Fe_2O_3 > 6% in unaltered beds vs. < 4% in altered units; see Table 2), siderite or ankerite are likely to form as ferroan carbonate (Fe content: siderite > ankerite > ferroan dolomite). Such Fe-carbonates should be revealed by diagnostic ferrous features centered, respectively, at ~1270 and ~1050 nm (Gaffey 1987). However, the analysis of the Ferrous spectral range (see Fig. 4) demonstrated that the wavelength positions occur between 1050–1230 nm (with the mean at around 1140 nm; Fig. 7b) that is characteristic of chlorite

rather than Fe-carbonate. Most likely, any possible ferroan carbonates in the system have been chemically weathered into iron oxyhydroxides (i.e. goethite) by the following reaction:

$$FeCO_3 + ½\ H_2O + ¼\ O_2 \rightarrow FeOOH + CO_2 \quad (2)$$

Siderite Goethite

The subtle feature was interpreted to be due to Ferrous-rich chlorite, because the derived wavelengths are confined between 1050–1230 nm, thereby closely correspond to chlorite that has absorptions centered at the 1050–1160 nm range. It is worth mentioning that adsorbed water exhibits a sharp absorption near 1135 nm due to fundamental O–H stretching and bending (Hunt 1977) that potentially could overlap with ferrous features. However, the meaningful correlation between ferrous and Fe–OH feature (see Fig. 10d) indicates that the observed absorptions in this range originate mainly from chlorite.

5.4.2. Diagenetic mineralogy within the SWIR range

The occurrence of chlorite is also corroborated by additional Fe–OH absorption feature at ~2253 nm (Bishop et al. 2008). As is illustrated in Fig. 7f, the majority of the microseepage-affected specimens retain an absorption that stretches between 2245–2260 nm wavelengths. There are several lines of evidence that support the interpretation of this feature in favor of chlorite rather than palygorskite, which happens to have an absorption feature at similar position: (i) palygorskite's OH/water absorption occurs at 1421 nm but, in the histogram (Fig. 7c), very few spectra correspond to this wavelength; (ii) the H_2O absorption for palygorskite occurs at around 1918 nm, whereas in the histogram displayed in Fig. 7d absorption features within such wavelengths are almost absent (Pontual et al. 2008).

The presence of clay minerals was confirmed by considering four characteristic absorption features namely Al–OH, Fe–OH, OH, and H_2O within the SWIR window (Fig. 7c-g). The rounded singlet absorption at around 2205 nm (the Al–OH histogram in Fig. 7e) was interpreted to be related to Al-smectites (montmorillonite) that is known to hold a diagnostic absorption feature at the same wavelengths. Montmorillonite is differentiated from other white micas through the absence of an additional absorption at ~2345 nm and the OH and H_2O absorptions centered at 1904 and ~1412 nm, respectively (Bishop et al. 2008; Hunt 1977). The subtle feature at ~2290 nm arising from Fe–OH bond was assigned to nontronite, which is a Fe endmember of the smectite series (Bishop et al. 2008) (Fig. 7g). Likewise, nontronite retains deep water absorptions at 1908 and ~1415 nm, which is characteristic of smectites (Fig. 7c-d). Remarkably, the spectroscopic analysis indicates that no kaolinite exists in neither the HC-affected nor unaltered rocks of the Qom region. The situation in which kaolinite is absent in either environment is barely reported in the literature, but presumably is not uncommon (Fu et al. 2007; Perry and Kruse 2010).

The other mineral that substantially contributes to aforementioned features is gypsum. The gypsum has multiple features in the SWIR range, including a triplet at around 1490 nm, a singlet at 1750 nm, a broad water band at 1945 nm, and a doublet centered at ~2215 nm (Bishop et al. 2014) (Fig. 7c-e). The latter was observed to overlap with the Al–OH feature at a longer wavelength, albeit, based on the wavelength position of Al–OH against H_2O, we interpreted the features beyond ~2210 nm to be mostly due to gypsum (see Fig. 7c-e and Fig. 14 for the details).

The few samples containing jarosite were identified by a characteristic absorption feature at ~2265 nm (Hunt and Ashley 1979) (Fig. 7f). Jarosite was only present in some iron-rich localized veins among the URF beds and was missing from the background lithology of the area. It seems the high concentration of calcite has maintained the pH neutral/alkaline and thus prevented the development of jarosite in the environment. In extreme cases, however, jarosite was shown to occur as a prevalent alteration product in microseepage environment (Perry and Kruse 2010).

Our spectral analysis also revealed that calcite (with a feature at ~2340 nm) is the dominant carbonate mineral present in the samples. However, trace amounts of ankerite (with absorptions between 2332–2338 nm) was also recorded in the samples (Fig. 7h). Even though calcite dramatically increases in the samples collected above the reservoir, it has not contributed much in the re-cementation of the sandstones, for the carbonate-rich facies are commonly friable and highly eroded (e.g. Fig. 2b). In contrast, where gypsum is present, it substantially takes part in the cementation of the units, hence forms the local highlands and cliffs. In such circumstances, gypsum shows a negative correlation with calcite abundances. Presumably, gypsum has precipitated at the expense of calcite dissolution. We speculate that the weathering of probable pyrite in nonacidic conditions imposed by the presence of calcite (along with oxygen and water) could give rise to more neutral minerals such as gypsum and ferric hydroxides, according to the following reaction (Bethke 2008; Ritsema and Groenenberg 1993):

$$FeS_2 + 2CaCO_3 + 11/2\ H_2O + 15/4\ O_2 \rightarrow Fe(OH)_3 + 2CaSO_4 . 2H_2O + 2CO_2 \quad (3)$$

Pyrite Calcite ic hydroxide Gypsum

Considering this reaction, even though pyrite is oxidized, due to the buffering effect of calcite, the pH will not drop below 6 (Bethke 2008). However, where the oxidizing situation is pervasive and there is no carbonate, the oxidation of pyrite can trigger the precipitation of metastable, pH-sensitive iron sulphates, including jarosite, copiapite, melanterite, and schwertmannite. Subsequent oxidation of this array could produce a range of ferric oxyhydroxides, comprising ferrihydrite, goethite, and eventually hematite (Asadzadeh and Souza Filho 2017; Elwood Madden et al. 2004).

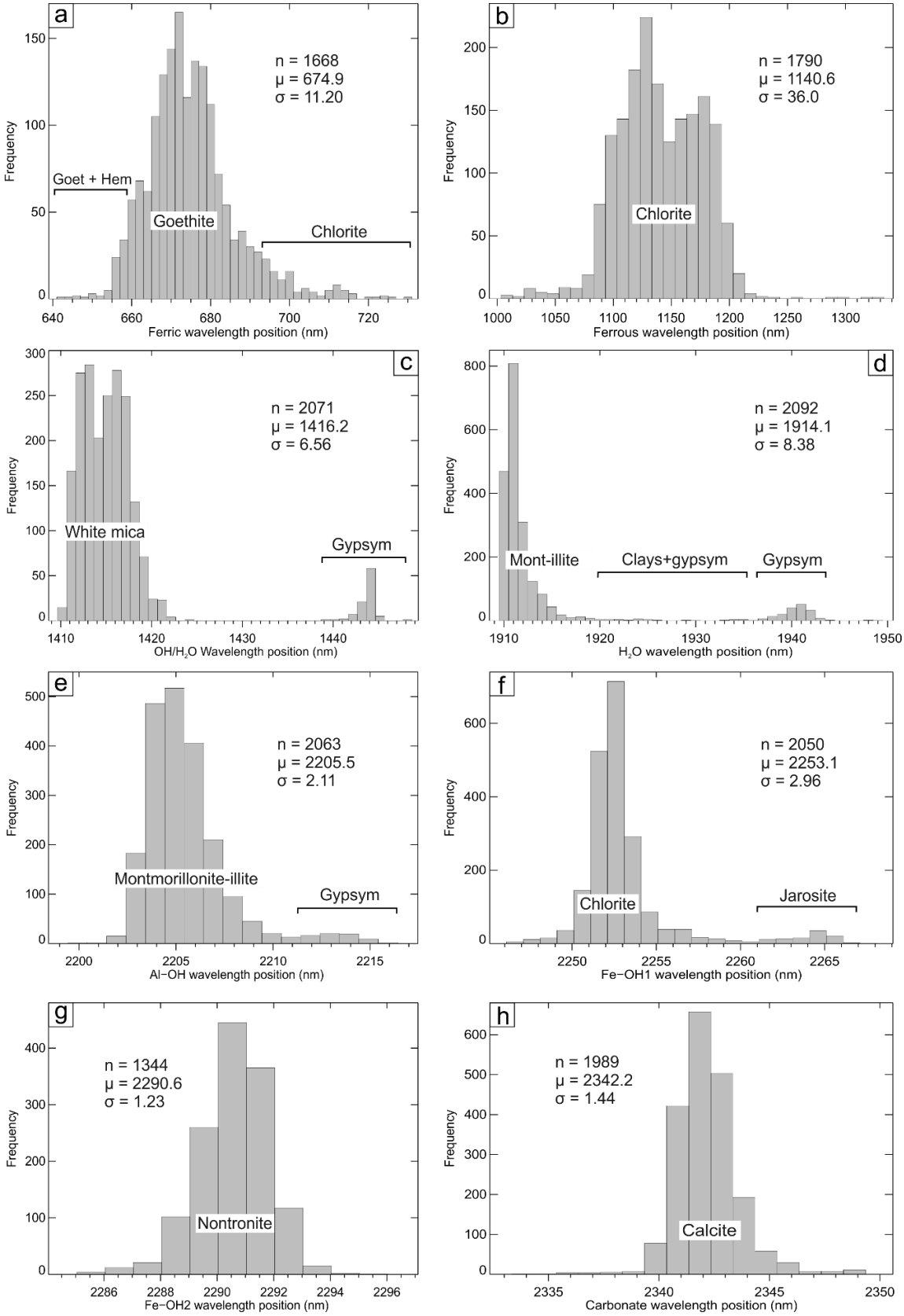

Fig. 7. a) Histograms of the wavelength position of different absorption features extracted from the Alborz spectral data overlaid with interpreted mineralogy. **a)** Ferric iron feature. **b)** Ferrous iron feature. **c)** OH/H_2O features. In this plot, the feature of jarosite at ~1473 nm was excluded for clarity. **d)** H_2O feature. **e)** Al–OH feature. **f)** Fe–OH(1) feature. **g)** Fe–OH(2) feature. **h)** Carbonate feature. n, μ, and σ represent the number of contributed features, the mean of the data, and the standard deviation, respectively. The acronyms are Goet: goethite, Hem: hematite, Mont: montmorillonite. The 'white mica' comprises montmorillonite and mixed-layer illite/smectite clays.

Reaction (3) can potentially explain (i) the unusual abundance of gypsum as a secondary cement in the altered facies, (ii) the complete absence of pyrite in outcrop studies, (iii) the remarkable excess of ferric irons, and (iv) the scarcity of calcite in gypsum-dominated samples. The presence of gypsum and ferric hydroxides indicate the extensive weathering of the host-rocks in the area and, indeed, an increase in the pH level of the soils and sediments (Elwood Madden et al. 2004). The presence of gypsum as a diagenetic mineral in seepage system is also reported by Shi et al. (2012) and Salati et al. (2014).

Even though gypsum was detected only in a limited number of samples through its diagnostic absorption feature (Fig. 7c-e), the plot of the H_2O wavelength position against its asymmetry denotes that the mineral is pervasive as a minor component. This is deduced from the fact that the shift in the water feature towards longer wavelength makes the feature more asymmetric, attesting the subtle presence of gypsum (Fig. 8).

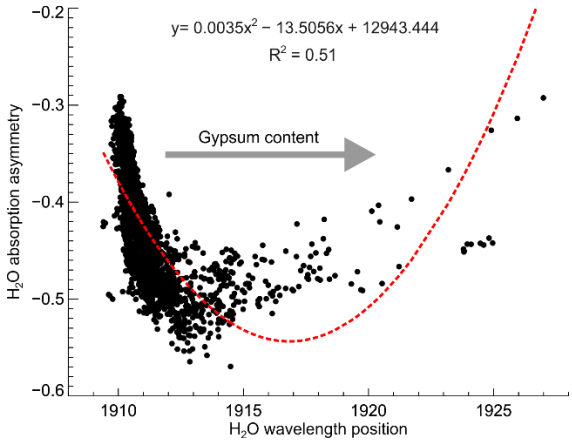

Fig. 8. Scatterplot exhibiting the variation of the wavelength position of H_2O at around 1900 nm against its asymmetry. The introduction of gypsum causes the feature to shift toward longer wavelengths and become more asymmetric.

5.4.3. The paradox of iron oxides abundances

The plot of Feox wavelength against its depth and asymmetry show that the transformation of iron oxides from hematite to goethite is associated with an unusual increase in the depth of the absorption feature (Fig. 9a) along with a decrease in the feature asymmetry (Fig. 9b). The decrease in the asymmetry (corresponding to an increase in the area of the right wing of the feature; see Fig. 3) could be linked to either overlapping ferrous feature (i.e. chlorite) or higher albedo of goethite beyond 1100 nm. The increase in the depth of the absorption is proportional to a rise in the amount of ferric iron. On the other hand, the formation of secondary goethite has been associated with iron oxide enrichment. This unusual tendency, which is in contrast to the established pattern in microseepage model (i.e. bleaching and iron oxide depletion), could presumably be due to (i) the introduction of iron from upper strata into the underlying layers by percolating meteoric water and then precipitation as meta-stable iron oxides; (ii) the oxidation of detrital magnetite or other secondary ferrous minerals noted above (eq. 2-3) into ferric iron during multiple oxidation-reduction phases; (iii) modification in the spectral behavior of the samples (e.g. the continuum slope or total reflectance) during the transformation from hematite to goethite; (iv) a decline in the amount of Al substitution in goethite lattice compared to Al-rich hematite, hence deepening the absorption feature; (v) a decrease in the particle size of iron oxides.

The first hypothesis is incoherent, because the bulk-rock geochemistry revealed by XRF analysis indicated that, on average, around 2.5% of the total Fe_2O_3 of the redbeds is depleted (Fe_2O_3 drops to <4% in affected facies; see Table 2 for more details). Moreover, the concentration of TiO_2 decreases together with Fe_2O_3, (Table 2) presumably indicating that Ti is released from titanomagnetite ($Fe_{3-x}Ti_xO_4$) grains. Concerning the second and third assumptions, our calculations (considering the maximum reflectance at ~1350 ± 70 nm) showed that the albedo of Alborz spectra is, on average, 30% higher than the Mil spectra (compare, for example, the Mil68 spectrum to the remaining spectra in Fig. 4). The increase in the albedo of the altered rocks could partly be related to the accumulation of clays; however, we believe the depletion of opaque minerals, including magnetite, is similarly important. Regarding the forth assumption, as discussed earlier, the amount of Al-substitution in secondary goethite in the HC-affected area is much lower than hematite (<5 mole %). Because Al-substitution can modify both the depth and wavelength position of goethite (Scheinost et al. 1999), a well-crystalline goethite gives rise to a deeper absorption and is likely to bring about a decrease in the asymmetry of the feature, as documented in Fig. 9a-b. The fifth and more plausible explanation, however, is a decrease in grain-size of iron mineralogy that is shown to be associated with a dramatic increase in the depth of the 900 nm absorption feature (e.g. Ramanaidou et al. (2008)). Overall, we believe a combination of these factors (grain-size effect alongside changes in the Al-substitution and albedo) should be responsible for the unusual trend documented in this study. To convert the ferric absorption depth into a direct measure of iron oxides abundances, one should develop some sort of corrections for the grain-size effects.

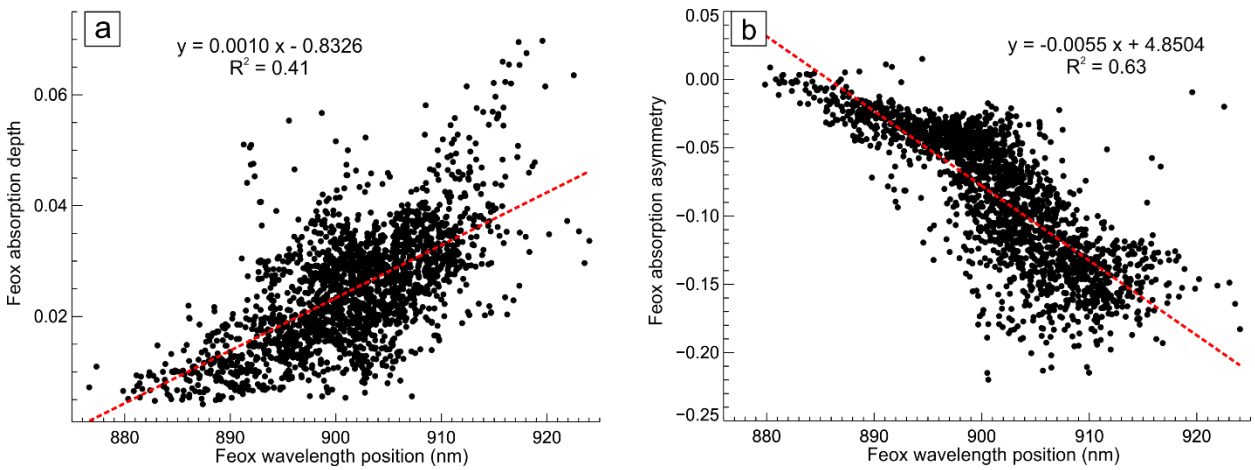

Fig. 9. Scatterplots exhibiting the variation of the wavelength position of the Feox feature against its depth and asymmetry. The step at 900 nm in (b) could be indicative of a change in the mineralogy of the sample from a hematite + goethite mixture into a dominant goethite phase.

5.4.4. The co-occurrence of diagenetic minerals

We investigated the co-occurrence of diagenetic minerals by applying statistical analysis to the wavelength position and depth of their diagnostic absorption features (Fig. 10). The weak negative correlation between Al–OH and goethite (Feox) wavelength positions shown in Fig. 10a provides an additional indication for a subtle Al-substitution in goethite, corroborating the abovementioned deduction. This trend could also suggest that goethite mainly coexists with Al-rich species of white micas characterized by shorter absorption wavelengths. The depth of the Al–OH feature has no meaningful correlation (zero correlation; not shown here) with Feox wavelength position. This means that the clay alteration/transformation is not interrelated with specific goethite concentration (recall

that the wavelength position of Feox feature defines the proportion of goethite to hematite in a binary system (Ramanaidou et al. 2008)), and therefore, it forms in a wide range of conditions alongside goethite.

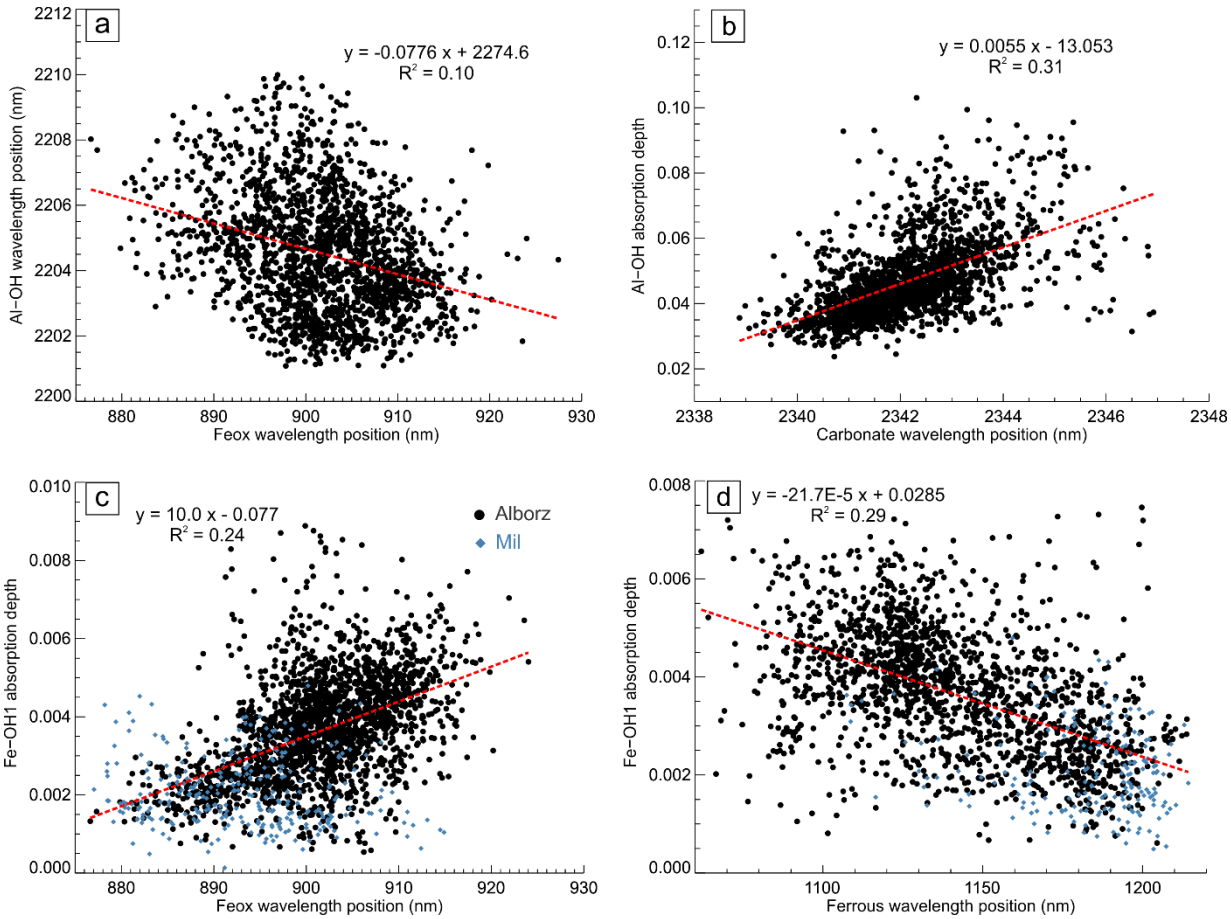

Fig. 10. Scatterplots exhibiting the relationships between different spectral parameters of diagnostic absorption features. **a)** the wavelength position of Feox (ferric iron) against the wavelength of Al–OH absorption. **b)** The wavelength position of carbonate feature against Al–OH absorption depth at 2200 nm. **c)** the wavelength position of Feox against the depth of Fe–OH (chlorite) absorption feature depth at 2245 nm. **d)** the wavelength position of ferrous feature against the depth of Fe–OH absorption feature depth. The legend in (c) applies to (d) as well.

One of the remarkable relationships observed among the spectral parameters was between the wavelength position of carbonate and Al–OH absorption depth, on which an increase in clay content pushes the feature towards longer wavelengths (Fig. 10b). We interpreted this positive relationship to be due to a second Al–OH feature beyond 2340 nm (Hunt 1977) overlapping with the carbonate feature in the same region. The development of this feature indicates that the secondarily formed clays in microseepage system tend to be mixed-layer illite/smectite. The composition of Al–OH clays is further discussed in section 4.6.

There is a positive correlation between the Feox wavelength position and the Fe–OH1 (chlorite) absorption depth (Fig. 10c), indicating that the transformation of hematite into goethite is associated with chlorite formation. This could imply that both goethite and chlorite originate from the same source, which is the reduced ferric iron produced in the microseepage system. Furthermore, the co-occurrence of goethite and chlorite indicates that the secondary chlorite in the environment is relatively iron-rich. As a reference, we have also included chlorites detected over the Mil area in Fig.

10c. Obviously, the noted relationship is typical of the microseepage-affected zones and the samples from the control area do not show a similar trend. Ultimately, we observed a meaningful negative correlation between the chlorite absorption depth and the ferrous feature (Fig. 10d). This correlation further supports the interpretation of the ferrous feature in favor of chlorite rather than Fe-carbonates. Moreover, it indicates that as the Fe–OH(1) absorption depth increases, the ferrous feature shifts towards shorter wavelengths, possibly implying that the higher abundance of chlorite is interrelated with Fe-rich species manifested by shorter-wavelength ferrous features. The few chlorites identified in the Mil samples are scattered at the long-wavelength end and do not exhibit a similar trend. The composition of chlorite is further discussed in section 4.6.

To illustrate the similarity and interrelationships between the delineated mineralogy, we conducted cluster analysis by using the depth of the corresponding absorption features of a subset of spectra wherein all the nine features were already present. The result is shown in Fig. 11. On this basis, three different groups could be defined including (i) carbonates (CO_3), Feox (ferric iron), Al–OH, (ii) chlorite (Fe–OH1), nontronite (Fe–OH2), ferrous iron, goethite, and (iii) OH and H_2O as two distant clusters. The dendrogram displays a cluster of carbonates, Feox, and aluminous clays together, indicating that they originate from the same process comprised in the microseepage-induced alteration. The second cluster encompasses iron-rich minerals consisting of chlorite (chlorite + ferrous ± ferric feature) and nontronite.

Fig. 11. Dendrogram showing clustering of the depth of the absorption features over the microseepage affected zones. The height of the graph represents the Euclidian distance between the absorption features. A subset of the samples wherein all the feature depths are present was used to draw this graph. See Fig. 7 and Table 1 for details about the labels.

5.4.5. The uniformity of the alterations

To investigate how distinctive the features are, we draw a statistical comparison between the samples from microseepage-affected zones and those of the control area (see Fig. 1 for the location) using the student's T-test. To simplify the comparison, the test was restricted to the wavelength position and depth of the diagnostic absorption features only. The results are reported in Table 3 and illustrated in Figs. 12-13. The same results are displayed in map format in Figs. 17 and 18 (see section 4.7).

Table 3. Comparison of wavelength positon and absorption depth over microseepage-affected and control areas using the statistical T-test (assuming equal variances). The checkmark (✓) shows a meaningful difference between the pair, whereas the cross (X) indicates a negligible difference.See Table 1 for more details about feature labels.

Feature	Ferric	Feox	Ferrous	OH	H_2O	Al–OH	Fe–OH1	Fe–OH2	Carbonate
Wavelength positon	✓	✓	✓	✓	X	✓	✓	✓	✓
Absorption depth	X	✓	✓	✓	✓	✓	✓	✓	✓

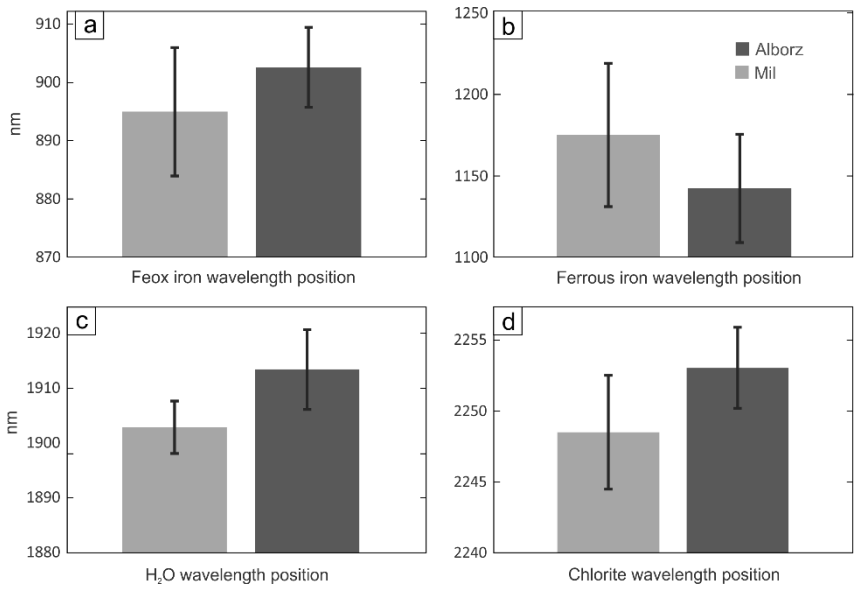

Fig. 12. Bar-chart representation of the differences in the wavelength position of particular absorption features from microseepage-affected (Alborz) and control (Mil) areas. The bar and whisker represent the mean and one standard deviation, respectively.

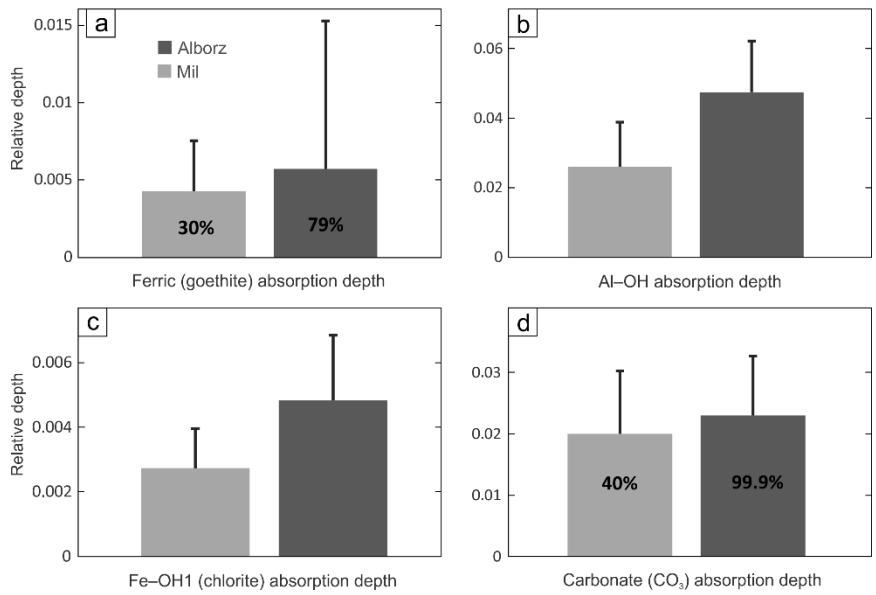

Fig. 13. Bar-chart representation of the differences in the absorption depth of particular spectral features from microseepage-affected (Alborz) and control (Mil) areas. The bar and whisker represent the mean and one standard deviation, respectively. The percentages printed on the bars represent the proportion of that specific feature in the sample collection.

The analysis revealed that except the depth of the ferric iron feature (occurring between 600–800 nm) and water, other spectral parameters in rocks over the oilfield show meaningful differences from their counterparts in the control area. More specifically, there was a difference of more than 8 nm, on average, in the wavelength position of the Feox feature (Fig. 12a), >30 nm in the wavelength of the ferrous iron (Fig. 12b), >10 nm in the wavelength of H_2O feature (Fig. 12c), and >5 nm in the wavelength of chlorite (Fig. 12d) between the two datasets. Regarding differences in the abundance of minerals as characterized by the absorption depth, we observed that although the pair of Ferric (goethite) features show no meaningful differences (Table 3), the distribution of the mineral within the control area is less than half of the petroliferous zones (Fig. 13a). In contrast, the Al–OH clays and chlorites absorptions were shown to be, respectively, >80% and >75% more pronounced in the

affected area (Fig. 13b-c). The increase in the proportion of clays in the altered sandstones may indicate that the secondary illite-smectites formed after feldspars/plagioclase alterations (see next section). Albeit the difference for the calcite declines to 15% only, the carbonates are present in only 40% of the samples of the control area, while it is present in virtually all the specimens collected over the oilfield (Fig. 13d).

As regards the relative proportion of the minerals in the microseepage environment, the Feox absorption is by far the dominant feature amongst the collected spectral data. Based on the relative abundances derived from spectroscopic analysis, the diagenetic minerals could be sorted as follows: goethite > smectite-illite ≥ chlorite ≥ calcite > nontronite > gypsum ≫ hematite ≫ jarosite (see the number of spectra "n" in Figs. 6 and 7).

5.4.6. Composition and crystallinity of clays

To explore the chemical composition and crystal structure of clay minerals, we plotted the wavelength position of the Al–OH feature against the crystallinity of white micas calculated by ratioing the absorption depth of the Al–OH to its nearby water feature at 1900 nm (Fig. 14). This plot could facilitate the discrimination of smectites from mixed-layer illite/smectite, because the higher the crystallinity of the white micas, the deeper the Al–OH feature against H_2O and, hence, the greater the calculated index (Pontual et al. 2008). In this plot, smectites could be characterized by wavelength ranges between 2200–~2210 nm and a crystallinity index <0.3, whereas the mixed-layer illite/smectite clays are outlined by values >0.3. Gypsum's feature coincides with the same crystallinity index as smectite; however, it occupies wavelengths longward the ~2210 nm. On this basis, the majority of clays from the Mil area are smectites. A large part, however, has near zero values and hence lay outside the defined clusters (Fig. 14; see also Fig. 17).

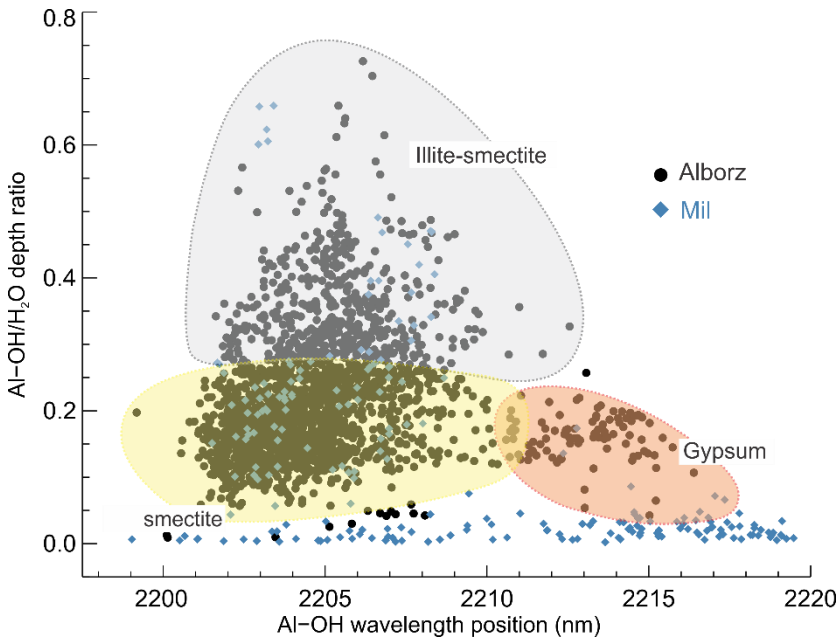

Fig. 14. Scatterplot showing the crystallinity of white mica calculated by the ratio of the Al–OH absorption depth to the water depth at 1900 nm and plotted as a function of Al–OH wavelength position. Three major categories are recognized: smectite, mixed-layer illite/smectite, and gypsum that happens to have an absorption at the same wavelength range.

As noted earlier, both the compositions and abundances of white micas and chlorites vary between unaltered to altered facies. To better understand these changes, we plotted the concentration of MgO and K_2O, respectively, against the Fe–OH and Al–OH spectral features (Fig. 15). The three endmembers of chlorites, namely Mg-, Mg/Fe-, and Fe-rich species, display their Fe–OH minimum at ~2245, ~2253, and ~2261 nm, respectively (Bishop et al. 2008). In the plot of Fig. 15a, the samples from the Mil area occur at around ~2247 nm and comprise the highest MgO content, whereas the HC-affected specimens tend to occur at wavelengths between 2250–2260 nm and show lower MgO concentration, thereby corresponding to Fe-rich chlorites. Remarkably, the abundance of chlorite (exhibited by the size of the symbols) shows a similar tendency and increases towards longer wavelenths (Fig. 15a; see also Fig. 18). Chlorite has been reported to occur in other microseepage environments (e.g. Fu et al. (2007)), however, thus far there was no account of its compositional variations in the system. In the literature, the transition from Mg-, to Fe-rich chlorite was only notified, to best of our knowledge, in bleached sandstones from onshore wells (Weibel 1998).

The white micas show a positive correlation with K_2O content and increases from unaltered to altered rocks (Fig. 15b). The extra potassium, which at least partly replaces the structural water and fills the interlayer sites, increases the crystallinity of the white micas (as discussed above) and brings about K-rich dioctahedral clays, including illite or mixed-layer illite/smectite. These secondary clay minerals are characterized by wavelength positions between 2207–2202, which in comparison to the Mil samples occur at shorter wavelengths and thus correspond to Al-rich white micas (Duke 1994) (Fig. 15b; see also Fig. 17).

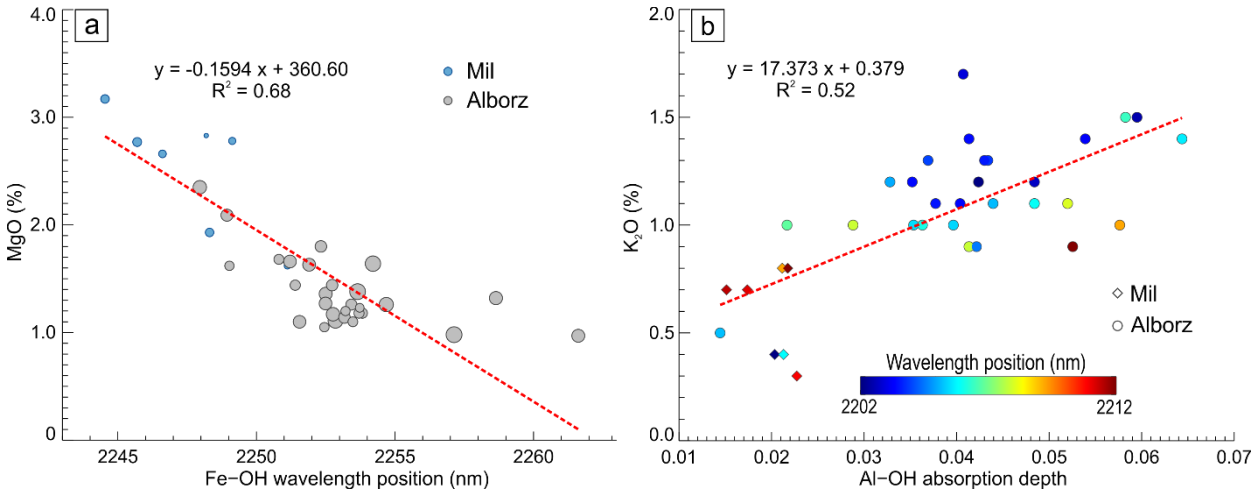

Fig. 15. a) Correlation of Fe–OH wavelength position against MgO (wt. %). The size of symbols represents the relative abundance of chlorite derived from the Fe–OH(1) absorption depth. **b)** Correlation of Al–OH absorption depth against K_2O (wt. %). The color-coded symbols exhibit the wavelength position of Al–OH feature. This parameter itself is negatively correlated (R^2 = 0.42) to K_2O content.

The plot of Al–OH absorption depth against Na_2O displays a nonlinear trendline (Fig. 16), indicating that the secondary clays are formed at the expense of sodium depletion from the system. The threefold decrease in the Na concentration possibly corresponds to either the alteration of feldspatic frameworks grains (i.e. albite) or the transfromation of dioctahedral smectites into illite. Since the total amounts of clays increases in the diagenetic facies (Fig. 13b), then it is more likely that the new clays are formed after fedspar alterations. Detailed petrographic studies are required to understand the processess involved in the formation and alteration of clay minerals in the altered lithofacies.

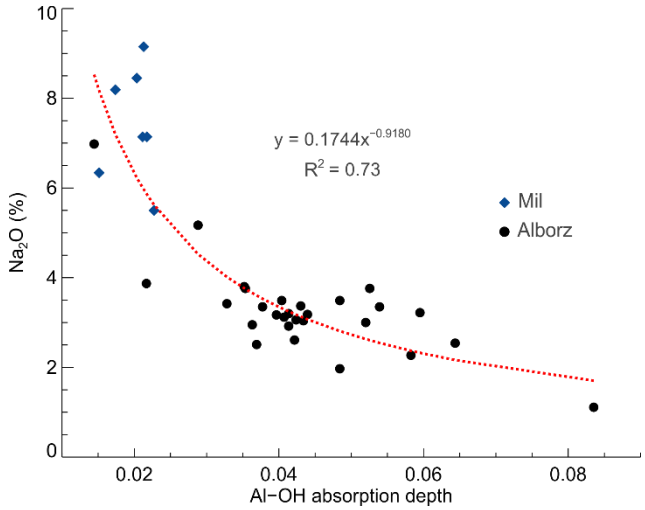

Fig. 16. Correlation of Al–OH absorption depth against Na_2O (wt. %), indicating that clays are formed at the expense of Na depletion of the system.

5.4.7. Spatial distribution of the diagenetic mineralogy

To show the spatial distribution of the minerals delineated in this study, we averaged the spectral parameters of every sample and assigned it to its geographic coordinates to yield point maps. The results are exhibited in Figs. 17 and 18. These maps clearly show the anomalous trends in the compositions and abundances of the diagenetic minerals over the Alborz microseepage system. Examples are the higher abundances of Al-rich and well-crystalline white micas (Fig. 17a-c) and the frequent occurrences of calcite (Fig. 17d). Chlorites on the other hand, tend to be Fe-rich and more abundant (Fig. 18a-b), whereas ferric iron oxides transform to goethite-rich facies over the reservoir (Fig. 18c). The only exception is an unbleached patch in the center of the anticline that is outlined by the dashed line (see Fig. 18c). Nontronite, which is also characteristic of microseepage environment (Fig. 18d), specifically co-occurs with goethite and white micas (compare Fig. 18d with 18c and 17c).

5.4.8. The correspondence with the standard microseepage model

The microseepage-induced mineralogy described in this study, which is summarized in Fig. 19, is consistent with low-temperature moderate alteration regimes characterized by neutral to alkaline conditions and high levels of Fe/Mg typical of mafic lithofacies. This mineralogic assemblage is fundamentally different from the array commonly noted in the literature (e.g. Asadzadeh and Souza Filho (2017)). Remarkably, if we consider the total bleaching and kaolinite occurrences as one end of the spectrum, the other end, which is described here, includes partial bleaching and the lack of kaolinite altogether (Fig. 19).

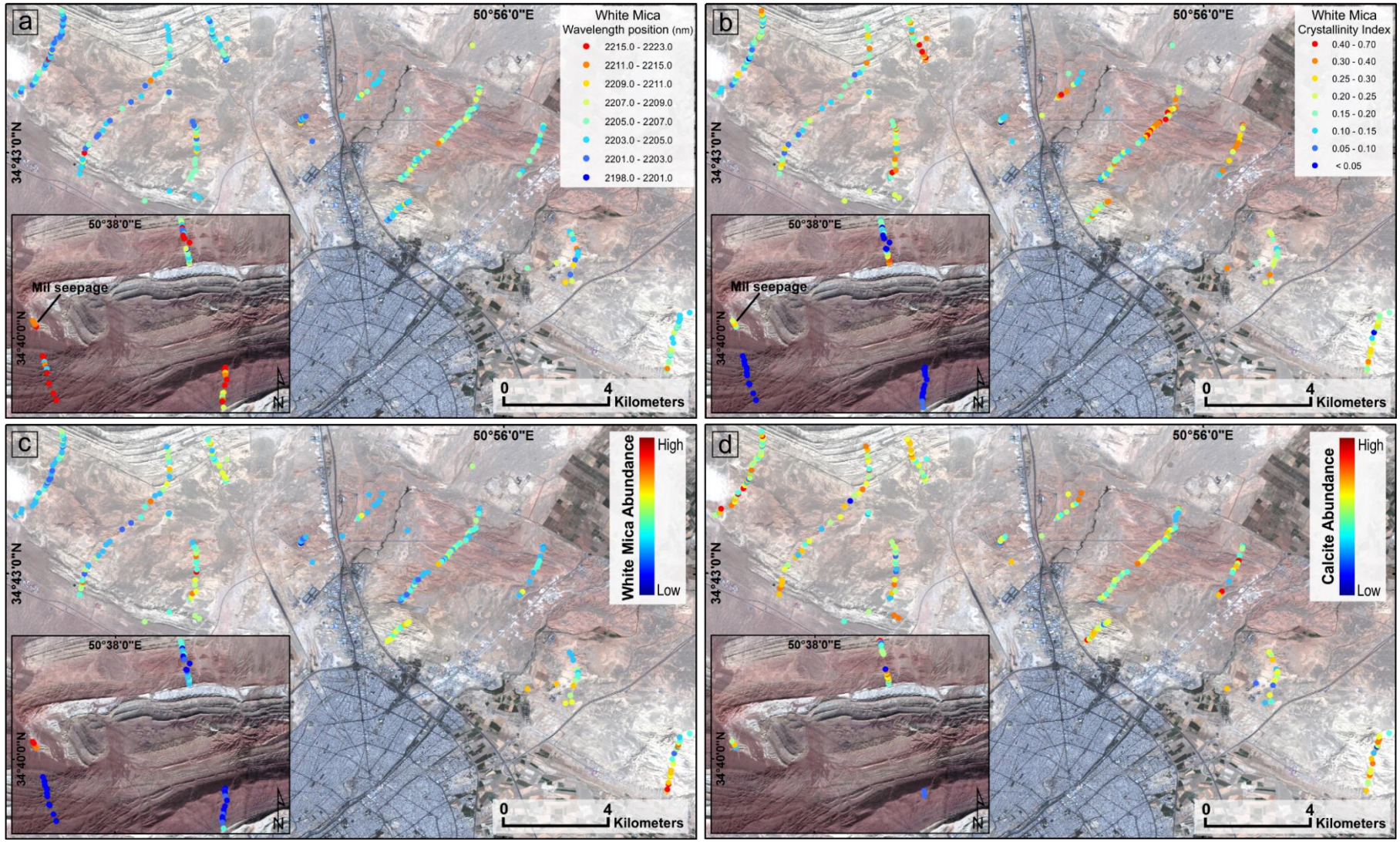

Fig. 17. Spatial distribution of diagenetic mineralogy interpreted from spectroscopic data overlaid on naturel color composite of Sentinel-2 imagery. a) wavelength position of white micas. b) crystallinity of white micas. c) white micas abundances corresponding to the Al–OH absorption depth. d) calcite abundance. Note that a subset of the area in Fig. 1 is exhibited here and the control area to the SW of the area is shown as an inset map. The open circles in (d) indicate the absence of calcite in the samples. The samples marked as 'Mil seepage' in the inset map were collected from the alteration halo of Mil oil-seepage.

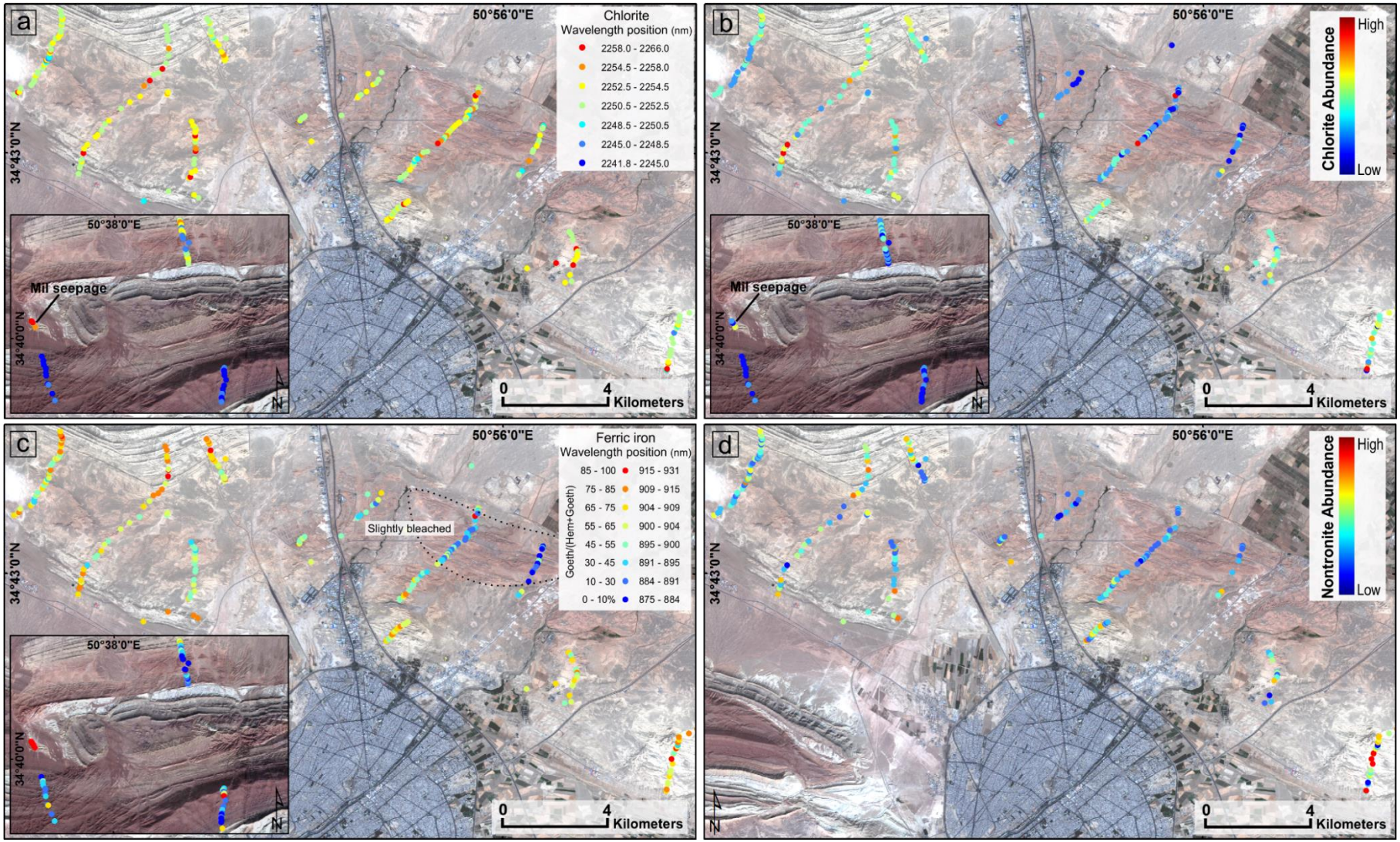

Fig. 18. Spatial distribution of diagenetic mineralogy interpreted from spectroscopic data overlaid on naturel color composite of Sentinel-2 imagery. **a)** wavelength position of chlorite. **b)** chlorite abundances corresponding to the Fe–OH(1) absorption depth. **c)** wavelength position of ferric iron and the goethite to hematite ratio (Goeth/Goeth+Hem) calculated by the formulation provided by Ramanaidou et al. (2008). **d)** nontronite abundances corresponding to the Fe–OH(2) absorption depth. The control area is omitted in (d) to better show the pattern over the reservoir.

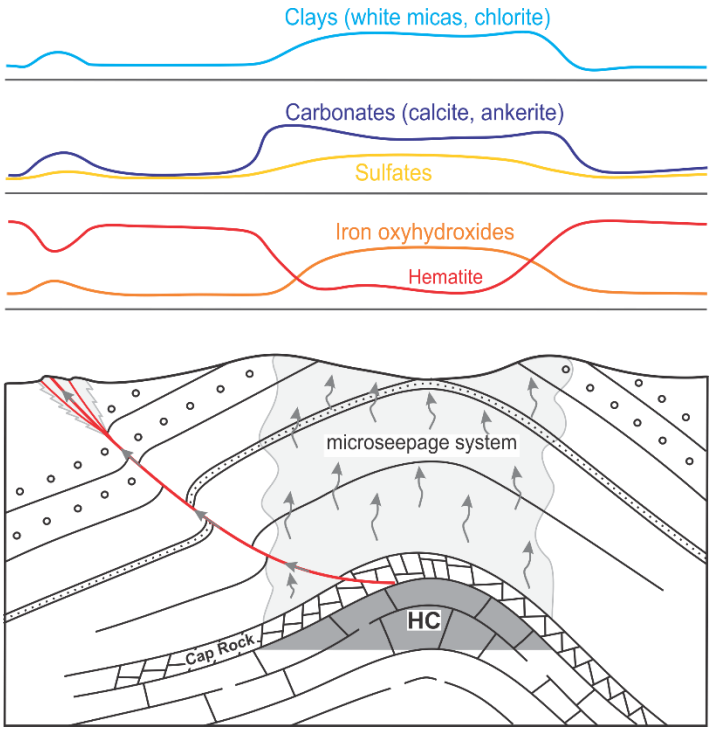

Fig. 19. An alternative model proposed to account for microseepage-induced alterations in mafic lithofacies. See Table 4 for the details of the mineralogy.

We speculate that the geochemistry of host-rock, as well as the weathering overprint induced by climate conditions, play a key role in shaping the ultimate alteration facies and mineralogic outcomes in a microseepage environment. Based on the silica content of the host-rock lithology, we tentatively classified the diverse range of diagenetic mineralogy induced by microseepage systems into two broad categories, namely felsic (SiO_2 > 60%) and mafic (45% < SiO_2 < 60%) assemblages. The results are summarized in Table 4.

Table 4. A tentative classification of the diagenetic alteration mineralogy commonly observed in microseepage environments.

Alteration mineralogy	Host-rock lithology	
	Felsic	**Mafic**
Clays	Kaolinite Illite (?)	Montmorillonite Illite (Al-rich) Mixed-layer illite/smectite Nontronite Chlorite (Fe-rich)
Ferric iron	Intense hematite bleaching	Moderate bleaching Goethite Ferrihydrite Maghemite Secondary hematite
Ferrous iron	Pyrite Siderite	Fe-chlorite Ankerite Pyrite (?)
Sulphates	(?)	Gypsum Jarosite, etc.
Carbonates	Calcite Dolomite Ferroan dolomite Ankerite	Calcite Siderite (?) Ankerite (?)

5.5. Conclusion

Infrared spectroscopy coupled with a multi-band spectral processing approach provided an unparalleled wealth of information about the abundances, compositions, and crystallinity of the microseepage-induced alterations over the Alborz oilfield. The study demonstrated that a rich variety of clay minerals, including smectites (montmorillonite, nontronite), mixed-layer illite/smectite, and chlorites are present in the HC-affected zones. The white micas composition in the altered facies proved to be Al-rich with a higher degree of crystallinity (structural order), whereas the chlorite composition showed a change to Fe-rich in the microseepage environment. Though the loss of ferric iron (bleaching) is widely noted to be indicative of petroleum microseepage, it was demonstrated that in the iron-rich host-rocks, which is characteristic of the mafic lithofacies in the Qom region, ferric iron might not be completely depleted, thus the chemical weathering of the relict ferrous irons could bring about secondary oxyhydroxides, including goethite and ferrihydrite. The variation in the composition (and crystallinity) of this assemblage (marked by a shift in the absorption feature centered at ∼900 nm) is considered to constitute an additional indicator for microseepage systems. Despite some earlier studies that denoted the replacement of gypsum by calcite, it was observed that possible pyrite in the bedrocks could weather into gypsum at the expense of calcite dissolution. The diagenetic sulphate minerals (gypsum and, in parts jarosite) could be considered as additional targets for remote sensing microseepage detection. Regarding the carbonates, our spectroscopic analysis revealed that calcite constitutes the dominant carbonate mineralogy in the sandstone beds. In the absence of ferroan carbonates and sulfides, Fe-rich chlorite appears to be the only ferrous iron mineral within the Alborz microseepage system.

Overall, the authors support that the ultimate mineral assemblages in a microseepage system are controlled by the characteristics of the petroleum reservoir, as well as the geochemistry of the host-rocks, and the chemical weathering overprint and in turn, the local climate conditions. The incorporation of the characterized mineralogic indicators could give rise to a more robust composite model for delineation of microseepage-induced alterations in soils and sediments.

Acknowledgement

The exploration directorate of the National Iranian Oil Company is acknowledged for supporting the fieldwork and gaining permission to publish the data. The authors would like to thank FAPESP for the research grant No. 2015/06663-7. C.R.S.F. thanks CNPq for the research grant No. 2008-7/303563.

References

Al Shaieb, Z., Cairns, J., & Puckette, J. (1994). Hydrocarbon-induced diagenetic aureoles: Indicators of deeper, leaky reservoirs. *Association of Petroleum Geochemical Explorationists Bulletin, 10*, 24–48

Amini, A. (2001). Red colouring of the upper red formation in central part of its basin, Central zone, Iran. *Journal of Sciences, Islamic Republic of Iran, 12*, 145-156

Asadzadeh, S., & Souza Filho, C.R. (2016). Iterative Curve Fitting: A Robust Technique to Estimate the Wavelength Position and Depth of Absorption Features From Spectral Data. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 5964-5974

Asadzadeh, S., & Souza Filho, C.R. (2017). Spectral remote sensing for onshore seepage characterization: A critical overview. *Earth-Science Reviews, 168*, 48-72

Balsam, W., Ji, J., Renock, D., Deaton, B.C., & Williams, E. (2014). Determining hematite content from NUV/Vis/NIR spectra: Limits of detection. *American Mineralogist, 99*, 2280-2291

Berberian, M., & King, C.C.P. (1981). Toward a paleogeography and tectonic evolution of Iran. *Canadian Journal of Earth Sciences, 18*, 210-265

Bethke, C.M. (2008). *Geochemical and biogeochemical reaction modeling*. Edinburgh, UK: Cambridge University Press

Bishop, J.L., Lane, M.D., Dyar, M.D., & Brown, A.J. (2008). Reflectance and emission spectroscopy study of four groups of phyllosilicates: smectites, kaolinite-serpentines, chlorites and micas. *Clay Minerals, 43*, 35-54

Bishop, J.L., Lane, M.D., Dyar, M.D., King, S.J., Brown, A.J., & Swayze, G.A. (2014). Spectral properties of Ca-sulfates: Gypsum, bassanite, and anhydrite. *American Mineralogist, 99*, 2105-2115

Bowen, B.B., Martini, B.A., Chan, M.A., & Parry, W.T. (2007). Reflectance spectroscopic mapping of diagenetic heterogeneities and fluid-flow pathways in the Jurassic Navajo Sandstone. *AAPG Bulletin, 91*, 173-190

Buckingham, W.F., & Sommer, S.E. (1983). Mineralogical characterization of rock surfaces formed by hydrothermal alteration and weathering; application to remote sensing. *Economic Geology, 78*, 664-674

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., & Vergo, N. (1990). High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research: Solid Earth, 95*, 12653-12680

Cornell, R.M., & Schwertmann, U. (2003). *The Iron Oxides: Structure, Properties, Reactions, Occurences and Uses*. (2nd ed.). Wiley-VCH Verlag GmbH & Co. KGaA

Donovan, T.J. (1974). Petroleum Microseepage at Cement, Oklahoma; Evidence and Mechanism. *AAPG Bulletin, 58*, 429-446

Duke, E.F. (1994). Near infrared spectra of muscovite, Tschermak substitution, and metamorphic reaction progress: Implications for remote sensing. *Geology, 22*, 621-624

Ellwood, B.B., & Burkart, B. (1996). Test of hydrocarbon-induced magnetic patterns in soils: the sanitary landfill as laboratory. In D. Schumacher, & M.A. Abrams (Eds.), *Hydrocarbon migration and its near-surface expresseion* (pp. 91-98). Tulsa, OK, U.S.A.: AAPG Memoir 66

Elwood Madden, M.E., Bodnar, R.J., & Rimstidt, J.D. (2004). Jarosite as an indicator of water-limited chemical weathering on Mars. *Nature, 431*, 821-823

Fu, B., Zheng, G., Ninomiya, Y., Wang, C., & Sun, G. (2007). Mapping hydrocarbon-induced mineralogical alteration in the northern Tian Shan using ASTER multispectral data. *Terra Nova, 19*, 225-231

Gaffey, S.J. (1987). Spectral reflectance of carbonate minerals in the visible and near infrared (0.35–2.55 um): Anhydrous carbonate minerals. *Journal of Geophysical Research: Solid Earth, 92*, 1429-1440

Hunt, G.R. (1977). Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics, 42*, 501-513

Hunt, G.R., & Ashley, R.P. (1979). Spectra of altered rocks in the visible and near infrared. *Economic Geology, 74*, 1613-1629

Morley, C.K., Kongwung, B., Julapour, A.A., Abdolghafourian, M., Hajian, M., Waples, D., Warren, J., Otterdoom, H., Srisuriyon, K., & Kazemi, H. (2009). Structural development of a major late Cenozoic basin and transpressional belt in central Iran: The Central Basin in the Qom-Saveh area. *Geosphere, 5*, 325-362

Perry, S.L., & Kruse, F.A. (2010). Evidence of hydrocarbon seepage using multispectral satellite imagery, Kurdistan, Iraq. In, *AAPG International Convention and Exhibition, AAPG Search and Discovery Article #90108* (p. 7). Calgary, Alberta, Canada

Petrovic, A., Khan, S.D., & Chafetz, H.S. (2008). Remote detection and geochemical studies for finding hydrocarbon-induced alterations in Lisbon Valley, Utah. *Marine and Petroleum Geology, 25*, 696-705

Pontual, S., Merry, N., & Gamson, P. (2008). *G-Mex—spectral analysis guides for mineral exploration: spectral interpretation field manual*. (third ed.). Kew, Victoria, Australia: Ausspec International Pty. Ltd.

Ramanaidou, B., Wells, M., Belton, D., verrall, M., & Ryan, C. (2008). Mineralogical and Microchemical Methods for the Characterization of High-Grade Banded Iron Formation-Derived Iron Ore. In S. Hagemann, C.A. Rosière, J. Gutzmer, & N.J. Beukes (Eds.), *Banded Iron Formation-Related High-Grade Iron Ore*: Society of Economic Geologists

Ritsema, C.J., & Groenenberg, J.E. (1993). Pyrite Oxidation, Carbonate Weathering, and Gypsum Formation in a Drained Potential Acid Sulfate Soil. *Soil Science Society of America Journal, 57*, 968-976

Salati, S., van Ruitenbeek, F.J.A., de Smeth, J.B., & van der Meer, F.D. (2014). Spectral and geochemical characterization of onshore hydrocarbon seep-induced alteration in the Dezful embayment, southwest Iran. *AAPG Bulletin, 98*, 1837-1857

Saunders, D.F., Burson, K.R., & Thompson, C.K. (1999). Model for hydrocarbon microseepage and related near-surface alterations. *AAPG Bulletin, 83*, 170-185

Scheinost, A.C., Schulze, D.G., & Schwertmann, U. (1999). Diffuse reflectance spectra of Al substituted goethite; a ligand field approach. *Clays and Clay Minerals, 47*, 156-164

Schumacher, D. (1996). Hydrocarbon-induced alteration of soils and sediments. In D. Schumacher, & M.A. Abrams (Eds.), *Hydrocarbon migration and its near-surface expression* (pp. 71-89). Tulsa, OK, U.S.A.: AAPG Memoir 66

Schumacher, D. (2012). Pre-drill prediction of hydrocarbon charge: Microseepage-based prediction of charge and post-survey drilling results. In, *CWLS GeoConvention 2012* (p. 9). Calgary, AB, Canada: AAPG

Schwertmann, U. (1971). Transformation of Hematite to Goethite in Soils. *Nature, 232*, 624

Segal, D.B., & Merin, I.S. (1989). Successful use of Landsat Thematic Mapper data for mapping hydrocarbon microseepage-induced mineralogic alteration, Lisbon Valley, Utah. *Photogrammetric Engineering and Remote Sensing, 55*, 1137-1145

Segal, D.B., Ruth, M.D., & Merin, I.S. (1986). Remote detection of anomalous mineralogy associated with hydrocarbon production, Lisbon Valley, Utah. *The Moutain Geologist, 23*, 51-62

Sherman, D.M., Burns, R.G., & Burns, V.M. (1982). Spectral characteristics of the iron oxides with application to the Martian bright region mineralogy. *Journal of Geophysical Research: Solid Earth, 87*, 10169-10180

Shi, P., Fu, B., Ninomiya, Y., Sun, J., & Li, Y. (2012). Multispectral remote sensing mapping for hydrocarbon seepage-induced lithologic anomalies in the Kuqa foreland basin, south Tian Shan. *Journal of Asian Earth Sciences, 46*, 70-77

Thompson, C.K., Saunders, D.F., & Burson, K.R. (1994). Model advanced for hydrocarbon microseepage, related alterations. *Oil and Gas Journal, 92*, 95-99

Weibel, R. (1998). Diagenesis in oxidising and locally reducing conditions — an example from the Triassic Skagerrak Formation, Denmark. *Sedimentary Geology, 121*, 259-276.

Chapter 6

Characterization of microseepage-induced diagenetic changes in the Upper-Red Formation, Qom district, Iran[8 9]

This work explores the surficial signatures of seeping hydrocarbons in parts of the Saveh-Qom sedimentary basin, Iran. The area includes the effectively sealed Alborz oilfield buried under Oligocene sediments of the Upper Red Formation (URF). The aim of the study was to prove the presence of leaking HCs over the reservoir, understand the chemical and physical processes responsible for diagenetic changes in the strata, and translate the diagenetic alterations into exploration indicators for remote sensing microseepage exploration worldwide. To achieve such objectives, the work integrated evidence from (i) petrographic and geochemical studies in the laboratory, (ii) outcrop investigations in the field, and (iii) broad-scale anomaly mapping using ASTER and Sentinel-2 orbital multispectral datasets. To build up a clear picture of the changes, the mineralogy, geochemistry, and petrography of the microseepage-affected lithofacies were contrasted with its unaffected counterpart in a nearby control area. The study indicated that even in the case of near perfect evaporitic capping, the reservoir still leaks HC to the surface. The induced physio-chemical and mineralogical changes were shown to be a function of the geochemistry of the lithofacies hosting the microseepage and the local environmental variables (i.e. climate and hydrogeology among others). The principally mafic URF lithofacies in this area coupled with an alternation between oxidizing and reducing regime have given rise to a novel array of mineralogical and geochemical indicators for microseepage systems. The major oxides, specifically the ratio of $CaO+K_2O/Na_2O+MgO+CaO+K_2O$, were substantiated to be a reliable indicator of microseepage-induced alteration. Regarding the iron oxides, we indicated that iron oxyhydroxides (i.e. goethite) constitute an additional indicator for microseepage remote sensing. The clay mineralogy was revealed to be dominantly smectites, mixed-layer illite/smectite, and chlorite with no traces of kaolinite. Based on isotopic studies, a combination of three sources of carbon including oxidized hydrocarbons was recognized to participate in the carbonate cements of the strata. Satellite remote sensing was successful in mapping the reservoir-scale footprints of the seeping petroleum. It indicated that the extent of microseepage-induced alteration, and thereby the petroleum reservoir, is much larger than previously thought. The anomalous zones over the reservoir were shown to conform to an annular pattern. This study demonstrated that an integrated exploration approach could facilitate the characterization of active microseepage systems and reduce the exploration and development risks in frontier and mature basins.

6.1. Introduction

In spite of traditional principles, seals above hydrocarbon (HC) reservoirs are not perfectly efficient and, as a result, HC accumulations leak to the surface over time. When the surface manifestation of oil and gas is clearly visible to the naked eye, it is termed macroseepage. A microseepage, on the other hand, constitutes minute traces of invisible light HCs in soils and sediments that is solely detectable by analytical methods and careful geochemical sampling (Horvitz, 1985; Tedesco, 1995). Macroseeps typically show on the surface due to major tectonic discontinuities cutting the reservoir, and penetrating the surface (Link, 1952; Macgregor, 1993),

[8] This chapter is under preparation for submission to the AAPG Bulletin.
[9] A part of this chapter was presented as oral talk in the 48° Congresso Brasileiro de Geologia, Porto Alegre, Brazil.

whereas microseeps are the results of imperfect reservoir sealing wherein HCs migrate in an almost vertical fashion (Price, 1986).

Microseepage-induced settings are characterized by a diverse range of anomalies including (1) anomalous gas concentration; (2) microbial and geobotanical anomalies; (3) mineralogical changes such as clay alterations/transformation, and the formation of carbonates, sulfides (e.g. pyrite), and elemental sulfur; (4) bleached facies; (5) electrochemical changes; (6) magmatic iron oxides and sulfides; and (7) radiation anomalies (Donovan, 1974; Price, 1986; Saunders et al., 1999; Schumacher, 1996; Tedesco, 1995) (Fig. 1). Bacteria and other microbes are believed to play a key role in the oxidation of seeping HCs. Their activity and by-products can change the pH-Eh of the overlying stratigraphic column and initiate the specified array of diagenetic physio-chemical and mineralogical transformations in the environment.

The detection of seepage systems is significant for oil and gas exploration, because it provides conclusive evidence for the formation of petroleum systems in a given sedimentary basin and, more importantly, because it supplies a powerful targeting tool for HC prospecting (Schumacher, 2012). Owing to this significance, a diverse range of unconventional exploration methodologies has emerged aiming to detect the footprints of HC migration in either direct or indirect way. The motivation for using remote sensing techniques in petroleum exploration has been the occurrence of several diagenetic minerals (Fig. 1) in the microseepage systems overlying HC accumulations that are spectrally active within the visible-near infrared (VNIR) and shortwave infrared (SWIR) wavelengths (Hunt, 1977).

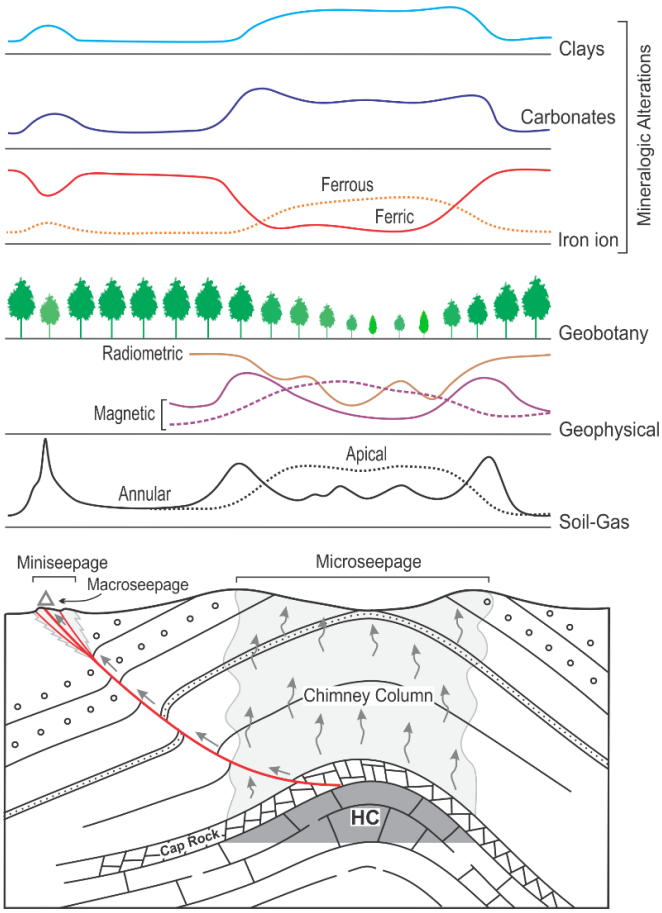

Fig1. Schematic representation of microseepage-induced anomalies overlying HC accumulations (after Asadzadeh and Souza Filho (2017)).

Despite the merits of remote sensing techniques in microseepage detection, decent case studies on the subject are rare in the literature. Only a few cases have used a comprehensive approach to scrutinize microseepage-induced anomalies from microscopic to regional scales. Moreover, the analytical techniques are seldom used in conjunction with remote sensing data for better understanding of the changes induced by microseepage at a multitude of scales.

6.2. Objectives

To close the gap in remote sensing microseepage detection and our understanding of the microseepage phenomenon, we adopted a multi-scale and multidisciplinary approach to study the well-exposed bedrocks overlying Alborz oilfield in the Qom region, Iran. The aim was to prove the presence of leaking HCs over the reservoir, understand the chemical and physical processes responsible for diagenetic changes in the strata, and then translate the diagenetic alterations into exploration indicators for remote sensing microseepage exploration worldwide. To achieve these objectives, the work integrated evidence from (i) petrographic and geochemical studies in the laboratory, (ii) outcrop investigations in the field, and (iii) broad-scale anomaly mapping using ASTER and recently launched Sentinel-2 multispectral datasets. To build up a clear picture of the changes, the mineralogy, geochemistry, and petrography of the microseepage-affected lithofacies were contrasted with its unaffected counterpart in a control area.

6.3. Study area and sampling

6.3.1. Geological background

The study area is located near the city of Qom, about 100 km to the south of Tehran. The approximate coordinates of the site are between 50°35'15" to 51°11'0" east, and 34°33'27" to 34°47'48" north covering an area of >1400 km^2 (Fig. 2). The climate of the region is arid to semi-arid with the temperature fluctuating between -15–45°C and <150 mm of annual precipitation on average, which slightly increases towards southwest in the mountainous parts. The dominant physiographic features of the area include the NW–SE running Kuh-e-Yazdan ridges, the prominent Kuh-e-Namak (salt diapir) topographic feature to the west, and low-relief and hilly land surfaces (laying at altitudes between 850–1000 m) corresponding to Alborz anticline. The Qom area is covered by sparse vegetation and thus have well-exposed bedrocks both on the ground and within the satellite imagery.

Geologically, this area is a part of the Central Basin of Iran Plateau referred to here as the Saveh-Qom basin. This wedge-shaped basin is bounded by Eocene pyroclastics to the north and tertiary volcanics of the Urumieh-Dokhtar Magmatic Arc (UDMA) to the south and southwest whilst to the southeast it connects to the Central Basin (Amini, 2001; Morley et al., 2009) (Fig. 2). The UDMA, which stretches hundreds of kilometers from northwest to the southeast of the country, is due to the subduction of Neo-Tethys oceanic crust beneath the continental crust of Central Iran. The Saveh-Qom basin is believed to form in the extensional regime within the back-arc of the subduction zone with the volcanic arc to the north corresponding to back-arc rifting (Berberian and King, 1981). Eocene volcanic rocks underlie the Oligocene to Miocene sequence in the Saveh-Qom basin (Berberian and King, 1981; Mostofi and Gansser, 1957). The main stratigraphic units cropping out in the basin comprises the Lower Red Formation (LRF; Oligocene), the Qom Formation (Oligocene-early

Miocene), and the Upper Red Formation (URF; Miocene) (Morley et al. (2009) and references therein) (Figs. 3-4). The lithology of the LRF, which retains limited outcrops to the southwest of the area (Fig. 3), includes shales, siltstones, marls, sandstones, and conglomerates. At the base of the LRF, there is a sequence of the halite-dominated evaporitic unit which in parts has been attributed to intrusion into upper units as salt pillows and diapirs (Morley et al., 2009).

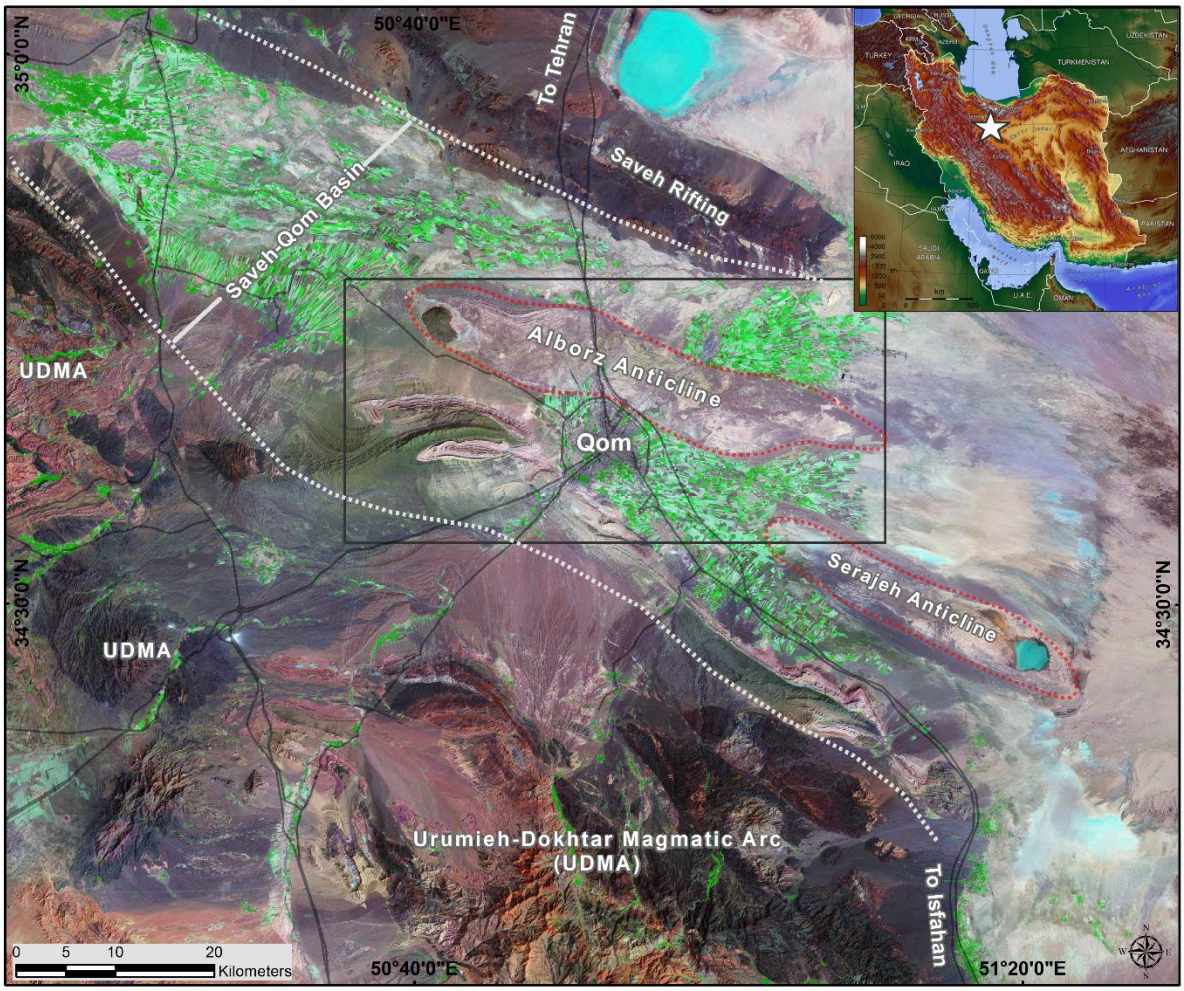

Fig. 2. The location of the study area (black quadrangle), and the outline of Alborz and Serajeh anticlines (red dashed line) overlaid on the mosaic of Landsat satellite imagery (RGB = 741). The boundary of the Saveh-Qom Basin is marked by a white dashed line. The inset map displays the locality of the study area in Iran.

The Qom Formation is a 1200 m thick carbonate-dominated unit that is composed of carbonates, marls, sandy limestones, and bituminous shales, with occasional anhydrite layers. The formation is stratigraphically divided into six different members starting from “a” at the bottom of the column to “f” at the top. It is believed that “e” and “f” members of this sequence (also known as Qom limestones) form the reservoir rock, with the evaporitic units at the base of the URF (composed of gypsum and salt) constituting the caprock (Gansser, 1957) (Fig. 4). The thickness of the caprock was indicated to reach up to 200 m in the drilled wells over the Alborz reservoir (Mostofi and Gansser, 1957). Qom Formation has outcrops to the west and southwest of the area, but towards the northeast, it is buried beneath 3000 m or more of the younger detrital-dominated deposits from the URF (Fig. 3).

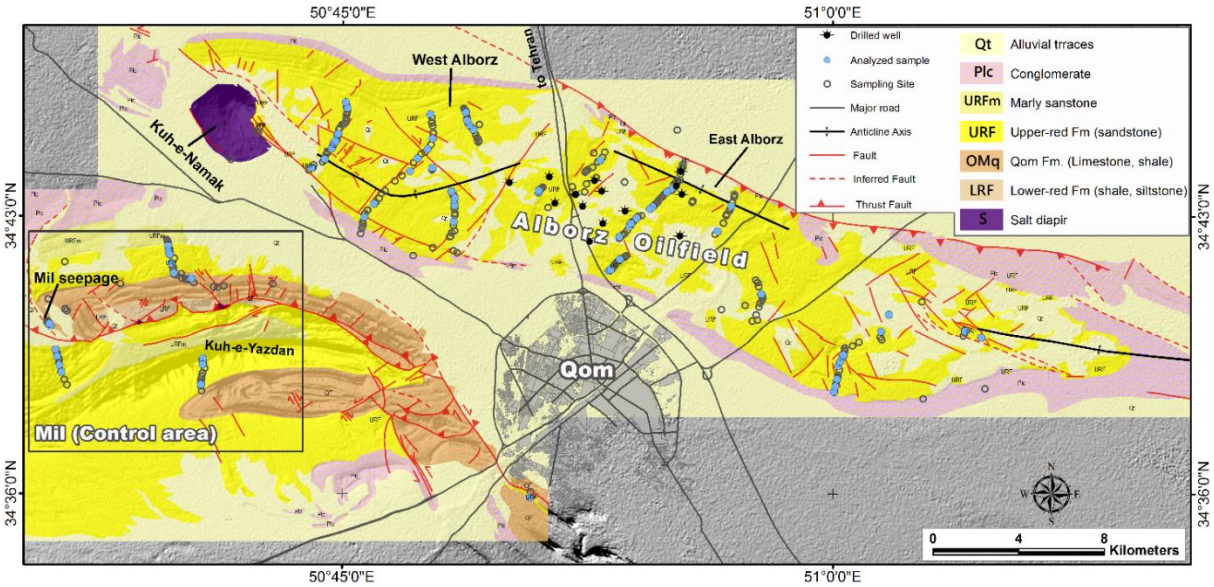

Fig. 3. Generalized geologic map of the Qom area superimposed on shaded digital elevation model compiled from 1:100,000, Qom, sheet 6159 and 1:250,000, Qom, sheet 046 geologic maps published by Geological Survey of Iran (www.gsi.ir).

The URF consists of a thick sequence of interbedded reddish sandstones, marly siltstone, conglomerate, and mudstone with local layers of marl, shale, and gypsum (Amini, 2001). The sandstone beds of the URF vary in thickness from millimetric laminated strata to beds of several meters. Most of the URF section was deposited during the transtensional phase of the basin in the early-middle Miocene wherein the clastic components were eroded away from nearby highlands (Amini, 2001; Reuter et al., 2007). The clastic contents dominated by volcanic detritus were demonstrated to be driven from two major sources located at either margin of the basin using palocurrent indicators. The volcanics originating from southern end were indicated to be aphanitic lithics and more basic in composition, whereas those driven from the northern margin were dominated by more acidic lithics (Amini, 1997).

The main character of the URF is its red color that points to an oxidizing continental environment during/after the deposition. The URF, which constitutes the target lithology of this study, is composed of at least three members: (i) lower member consisting of dark reddish gypsiferous sandstone, shale, siltstone and green marl; (ii) middle member consisting of intercalations of cavernous sandstone, conglomerate and shale; and (iii) upper member consisting of siltstone, bright yellow gypsiferous marl interbedded by sandstone. The Upper Red and Qom Formations are considered to be approximately time equivalent, respectively to the Agajari and Asmari Formations in the Zagros oil belt (Morley et al., 2009).

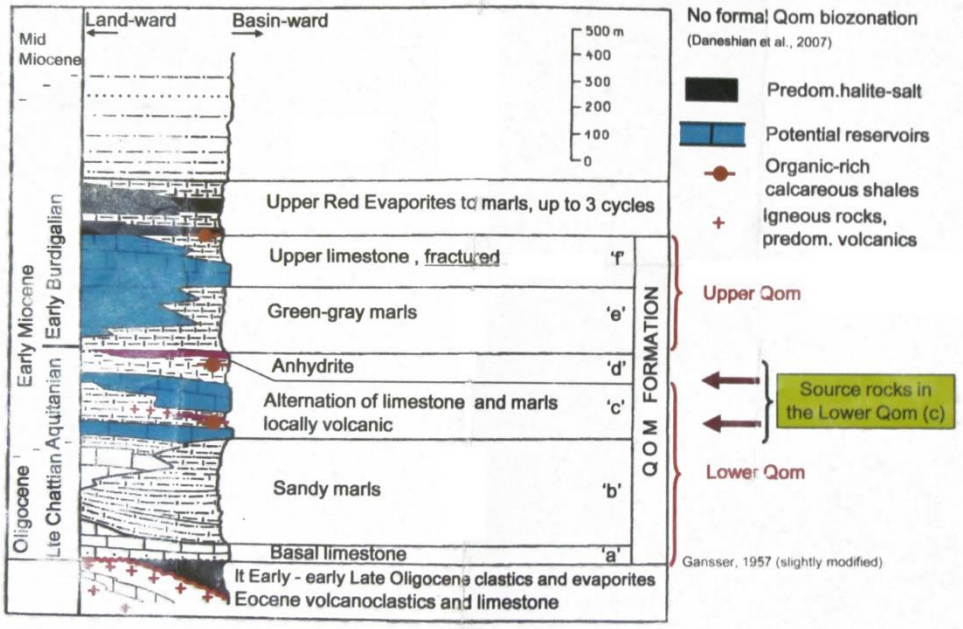

Fig. 4. Simplified stratigraphic column for the Saveh-Qom Basin sedimentary rocks. Adapted from Gansser (1957).

The transpressional deformation phase occurring during the Miocene has given rise to the formation of 'Alborz' and 'Serajeh' anticlines (Morley et al., 2013)(Fig. 2). The Alborz anticline, which is the focus of this study, is an asymmetric, overturned structure with over 50 km long and ~12 km wide from syncline to syncline (Mostofi and Gansser, 1957) (Fig. 1). The exploration drilling in the 1950s indicated that oil and gas accumulations are present, respectively, in the Alborz and Serajeh structures. Unfortunately, the first well breaching Alborz reservoir in 1956 blew out, and depleted more than 7 MMBLO (million barrels of crude oil) from the reservoir (Mostofi and Gansser, 1957). The oilfield that was productive for only a short duration, has potential reserves (oil in place) that are estimated at 217–274 MMSTB (million stock tank barrel of oil; NIOC internal report). The specific gravity (API°) of the oil from this field was estimated to be around 37 with the total sulfur content of 0.85% (Mostofi and Gansser, 1957). The URF is subject to HC migration from the reservoir to the surface. Thus far, however, the characteristics of a possible microseepage system in the area and its interrelationships with underlying HC accumulations has remained unknown. The only known macroseepage in the area called 'Mil' is located 25 km west of the Qom city (Fig. 3). This seep issues from gypsiferous units surrounded by very steep southwest plunge marly members of the Qom Formation.

6.3.2. Sampling

Fieldwork was carried out along 14 profiles defined to cross the distinct stratigraphic variations of the area using color composite satellite imageries. Surface outcrop samples were collected along these profiles aiming to represent all of the commonly observed diagenetic changes within the URF. The studies covered two main sites 'on' and 'off' the oilfield; the 'on-field' site corresponds to outcrops overlying the Alborz oilfield, whereas the 'off-field' site, which devised to serve as a control area, relates to Mil and Kuh-e-Yazdan areas (Fig. 2). The control area that is consisting of unaffected outcrops of the URF were used to define the petrographic, geochemical, and mineralogical background of the host-rock and cross-validate the diagenetic changes observed over microseepage-affected zones. Around 20% of the samples came from the control area and the remaining were

gathered from the microseepage-affected zones (respectively called 'Mil' and 'Alborz' area henceforth). In total, some 360 representative samples were collected along ~43 km line profiles resulting in an average of one sample per ~120 m length. The samples covered a variety of targets consisting of rock chips from weathered and fresh lithofacies (mainly sandstones/siltstones with occasional conglomerates), soil covers, and oil seeps among others. The majority of the samples were collected from the rocks exposed to orbital sensors; however, vertical faces hidden from overhead sensors were also sampled. We catalogued the variations in the color of the hand samples using a Munsell color chart and photographed typical geologic phenomena in the field. The locations of the sampling sites and field observations were recorded using a handheld GPS.

The entire sample suite was analyzed in the laboratory using reflectance spectroscopy. The details of the spectral analysis are provided in a companion article on the area (Chapter 5). A selection of the suite (n = 73), chosen by considering the visual characteristics of the samples, its locality along the profiles, and the spectroscopic outcomes, is considered here for detailed geochemical and petrographical studies as is summarized in Table 1. The

Table 1. Summary of the analytic techniques employed to study the samples from Qom area.

Analysis method	Samples
Infrared spectroscopy	360
Petrography	25
X-Ray Diffraction	41
X-Ray Florescence	36
ICP-MS	22
Stable isotopes (C, O)	15

6.4. Material and methods

6.4.1. Petrography

Representative thin sections (n = 25; 17 from Alborz + 8 from Mil) of intact and altered URF sandstones were examined petrographically using an optical microscope. The soft and friable samples from altered facies were first stabilized by impregnation with epoxy resin before undergoing cutting and polishing processes. The grains, cements, and coatings of altered facies were qualitatively studied against unaltered samples from Mil area. The percent mineralogy was estimated by point counting (500 points) at a spacing of 0.1 mm. Because spectroscopic analysis indicated the dominance of calcite in the samples (Chapter 5), we declined to stain the thin section for carbonate identification, and assumed all the carbonates to be indicative of calcite.

6.4.2. Geochemical analysis

We analyzed 25 variably altered specimens from Alborz area and further 10 specimens from Mil area with X-ray diffraction (XRD). The specimens were grinded using agate mortar and pestle and then were analyzed directly to detect the main mineralogic phases. For the aim of clay identification, the clay particles (<2 μm fraction) of a subset of 13 specimens (8 + 5) were extracted by peptizing in water and two rounds of centrifuging to make dual oriented smears on glass blades for each specimen. The blades were analyzed following heating and glycolation for 24h for the identification

of mixed-layer clay minerals. In addition, 6 specimens from the Alborz suite were analyzed by quantitative XRD technique.

We analyzed 36 representative specimens (28 from Alborz + 8 from Mil) with X-ray fluorescence (XRF). The pulverized specimens were converted into glass disks and analyzed for eleven major oxides including SiO_2, Al_2O_3, Na_2O, K_2O, CaO, Fe_2O_3, MgO, TiO_2, P_2O_5, MnO, and SO_3. The results are reported in Table 2 of Chapter 5. Moreover, 22 pulverized specimens (18 + 4) were analyzed by ICP-MS (inductively coupled plasma-mass spectrometry) to determine 42 minor/trace elements. The REE data were normalized against C1 Condrite standard (McDonough and Sun, 1995) before illustrating in diagrams.

Carbon (^{13}C) and oxygen (^{18}O) isotopes of 15 variably altered specimens were measured from calcite embedded in carbonate-rich sandstones aiming to infer the origins and generations of cements and fluids and determine their relationship with hydrocarbon microseepage. The samples were pulverized and dissolved with 100% phosphoric acid to liberate CO_2 (Rainoldi et al., 2014; Simpson et al., 1991). Then, the collected gas was measured by an isotope ratio mass spectrometer. The results are reported relative to the VPDB standard.

6.4.3. Reflectance spectroscopy

As noted earlier, all the collected samples in Qom area were analyzed using reflectance spectroscopy. Here, we utilized a subset of the collected spectral data to corroborate the XRD analysis, interrelate the mineralogic interpretations, and verify the results yielded from remote sensing studies. The details of the data collection and processing approach are provided in a companion article in Chapter 5.

6.4.4. Satellite remote sensing

For reservoir-scale alteration mapping, we deployed two orbital datasets obtained from ASTER and Sentinel-2 satellite sensors. ASTER has 3, 6, and 5 spectral bands in the VNIR, SWIR, and TIR wavelengths with 15, 30, and 90 m spatial resolution, respectively (Abrams and Hook, 2000). Sentinel-2, on the other hand, retains 9 multispectral bands between 450–2450 nm of which 7 are located within the VNIR range. Hence, it provides a powerful tool to map transitional elements, particularly iron oxides and hydroxides, at 10–20 m spatial resolution thus complementing the SWIR bands of the ASTER imagery. An ASTER Level-1B scene acquired on June 2, 2004 and two adjacent tiles of Sentinel-2 data (level-1C) acquired on October 5, 2016 were selected to cover the study area. Essential radiometric and atmospheric corrections were applied to both datasets. The radiance at-sensor SWIR data were converted to apparent surface reflectance using ATCOR atmospheric correction software (ReSe Applications LLC, CH). The Sentinel-2 data were also converted to apparent reflectance (Level-2A) using Sen2Cor toolbox provided by SANP platform. The results were then mosaicked to yield seamless imagery over the area.

Based on field observations and spectroscopic analysis, a suite of microseepage-induced mineralogic targets was specified and then mapped using a combination of feature tracking (e.g. band ratioing) and match filtering techniques applied to either of satellite datasets.

6.5. Results and discussion

6.5.1. Field observations

The most noticeable phenomenon observed in the field was a significant difference in the color of the bedrocks between unaffected and microseepage-affected zones (Fig. 5). The color of unaffected sandstone and siltstone units (collectively called sandstone henceforth) in Mil area and elsewhere was varied from dark reddish gray (10R, 4/1) to pale red (7.5R, 6/3) and from reddish brown (2.5YR, 5/3) to weak red (10R, 5/4) (e.g. Fig. 5a). Over the reservoir where the sediments were affected by microseepage, a wide range of colors was observed varying from pale red (2.5YR, 7/2) and reddish yellow (7.5YR, 7/6) to yellow (2.5Y, 7/6) and pale yellow (5Y, 8/2) (e.g. Fig. 5b). In extreme circumstances, light gray (5Y, 7/1) and white (2.5Y, 8/1) lithofacies were also observed (Fig. 5c).

The catalogued variations in the hue of the sediments correspond to changes in their iron oxide contents and compositions as well as crystal size and isomorphic substitution (Nielson et al., 2014). In the yellowish sandstone beds (Fig. 5b-c), which happens to be the most indicative feature of the affected zones, the original iron oxide (hematite) coating has transformed into goethite (see Chapter 5 for a comprehensive overview of the findings), whereas in the intensely bleached facies, iron oxide grain coats has been completely dissolved by chemical processes and removed from the system. The noted variability was especially evident in the sandstone lithofacies. In in conglomerates, however, due to the breakdown of the rock unit in affected zones, such changes were not easy to notice and record (Compare, for instance, Fig. 5d with 5e). The shaly inter-layers, on the other hand, demonstrated a reverse trend. In the unaffected zones, these inter-layers tend to appear as light to medium gray units, whereas in the affect zones, they retain more yellowish color likely because of secondary iron enrichment (Fig. 5f-g) integrated from nearby units during iron mobilization.

The delineated color changes were selective and controlled by stratigraphy and fabric of the facies among which the more permeable strata (i.e. coarse-grained sandstones) were undergoing the most dramatic changes (Fig. 5c and 6a). Our fieldwork in Mil area indicated that the variability in the colors of the units is also a function of primary hematite proportion/infiltration, for the hematitic pigment was not uniformly distributed among the lithofacies in the first place (e.g. Fig. 5 in Chapter 5). A similar trend is for example, reported in Oklahoma oilfields wherein fine-grained sandstones were selectively bleached with other lithofacies like claystones and siltstones remaining largely unaffected (Ferguson, 1976-1979).

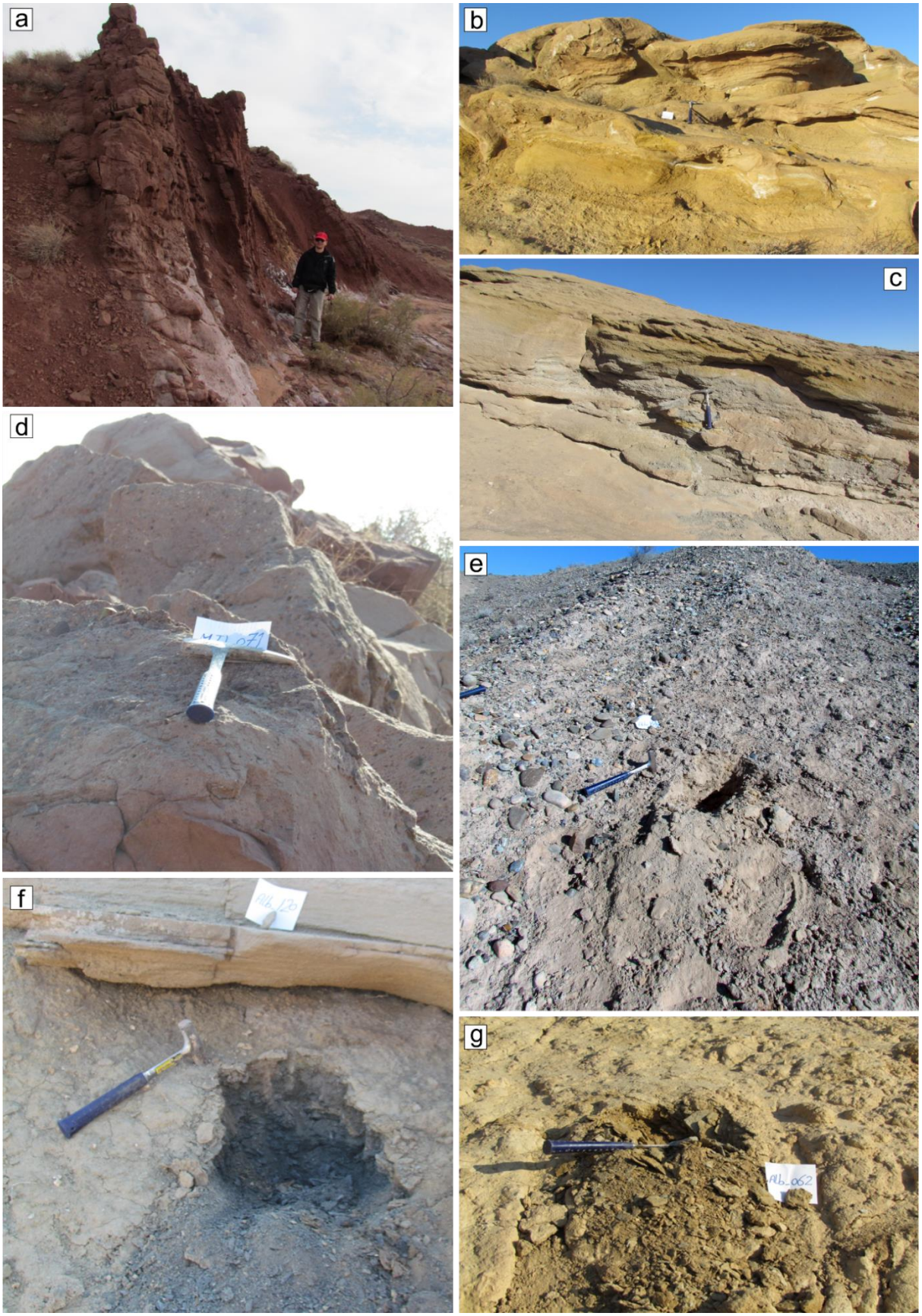

Fig. 5. Photographs contrasting the microseepage-affected beds of the Upper Red Formation with unaffected counterparts in the Qom region. **a)** unaffected sandstone beds of the URF at Mil-32 sampling site. **b-c)** affected sandstone beds of the URF, respectively at Alb-217 and Alb-166 sampling sites. **d-e)** unaffected and affected conglomerate beds of the URF, respectively at Mil-71 and Alb-144 sampling sites. **f-g)** unaffected and affected shale beds of the URF, respectively at Alb-120 and Alb-62 sampling sites.

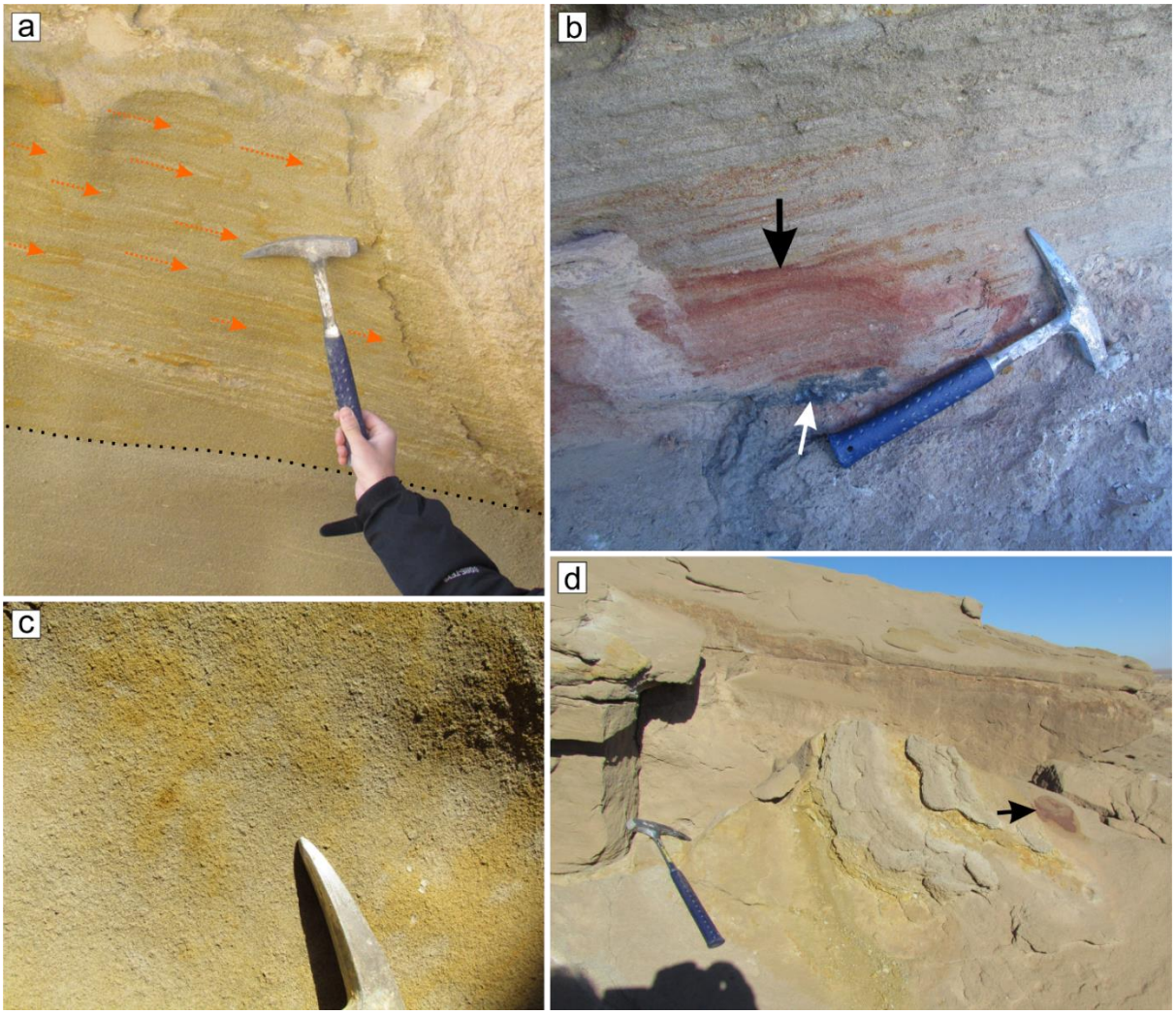

Fig. 6. Photographs illustrating the evidence for dissolution and movement of iron within the sandstone beds due to microseepage phenomenon. **a)** the miniature roll-fronts of ferric iron oxides formed due to the oxidation of soluble/mobile ferrous iron; photographed at Alb-160 sampling site. **b)** remnant of the initial hematite coating (banded pattern; black arrow) in bleached strata at the Alb-158 sampling site. The white arrow indicates manganese oxides. **c)** diffuse pattern in the distribution of secondary iron oxides (goethite) at the Alb-74 sampling site. **d)** concretionary pattern in the distribution of secondary iron oxides (shown by the black arrow) at Alb-165 sampling site.

Several patterns related to secondary iron oxide distributions were observed in the URF lithofacies including diffuse, banded, spotty, and irregular concretionary among others (Fig. 6; see also Fig. 3 in Nielson et al. (2014)). The patterns are known to form due to chemical weathering (re-oxidation) of the reduced iron (Fe^{2+}) by meteoric water in the vadose zone (Nielson et al., 2014). Goethite, as the most common secondary iron oxide, shows a diffuse pattern discernible by evenly distributed yellowish color (Fig. 5b and 6c). The dominance of goethite coats not only indicates that the environment has been in excess of oxygen but also imply that there has been an alternation between oxidizing and reducing regime in the environment. Alternation of redox potential is known to arise in Mediterranean soils wherein interstitial waters change from reducing during cold seasons to oxidizing during hot, dry summers (Ellwood and Burkart (1996) and references therein). It seems that the long and hot summers in the Qom region have imposed alternating oxidizing conditions on the soils and sediments within the microseepage system. The risen redox potential then has triggered the precipitation of mobilized irons (Fe^{+2}) as diffuse oxyhydroxides overprinting the reducing background (e.g. Fig. 5b). The same process could also be responsible for the oxidation of possible ferrous minerals such as pyrite and ankerite into metastable iron oxides (i.e. ferrihydrite and

goethite). The alternation in redox potential could also explain well-ordered goethite with subtle Al-substitution characterized by shorter wavelength position at ~900 nm absorption feature (see Chapter 5 for the details).

It should be emphasized that while the role of groundwater flow in displacement or obliteration of geochemical anomalies is still disputed (Schumacher, 2000), there is no doubts that underground water flow is responsible for mass transfer (deposition/depletion) into or out of the system in proportion to flow gradient and the distribution of discharge and recharge zones (Holysh and Toth, 1996).

The noted processes have likely participated in shaping the physiography of the area. Given the sandstone beds, to the south and southwest it tends to be hard and resistant to erosion, hence constitute the rugged topography (i.e. Kuh-e-Yazdan in Fig. 1). Over the oilfield, however, it tends to be friable and porous; consequently the progressive erosion has formed a flat-lying topography in large parts of the area.

Other secondary features in the affected lithofacies worth mentioning are localized occurrances of colliform manganese oxides growing in joints and alongside the secondary iron oxides (Fig. 6b), secondary gypsum cements associated with bleached facies, pore-filling silicic veins, and sparse copper mineralizations (mainly malachite) in permeable strata. Presumably, the gypsum cement is formed by the oxidation of sulfides (i.e. pyrite) at the expense of carbonate dissolution, as indicated in Chapter 5.

Over the reservoir, there was no structural conduit or any macroseepage indication which is characteristics of the lightly deformed basins associated with the back-arc setting (Macgregor, 1993). The only macroseepage in the area (Mil oil seep) is located farther southwest and issues from a narrow gypsiferous layer bounded by marly units (Figs. 2 and 7a-b). The seepage appears at the endpoint of a reservoir penetrating low-angle thrust fault. The fault seems to direct the petroleum towards gypsiferous layer before manifesting to the surface, for oil was observed to be trapped in between the gypsum cleavage (Fig. 7d). Around the Mil seepage, parts of the gypsiferous strata were altered to what is known as Gach-i-Turush in the literature (Clarke and Cleverly, 1991; Thomas, 1952) that is the association of oxidized petroleum, calcite, grey powdery gypsum, and jarosite among others (Fig. 7c). The other local alteration observed in the marly unit (Fig. 7a) bear a striking resemblance to microseepage-induced diagenetic changes. Here, the diffused near-surface flow of gaseous phases of the seepage has altered the originally red marls into yellow units characteristics of hematite to goethite transformation (Fig. 7a). Even though the involved physiochemical processes are identical, we adopt the 'miniseepage' term to describe the phenomenon and distinguish it from the alterations induced by near-vertical HC migration over the accumulation (Asadzadeh and Souza Filho, 2017). The details of mineralogical (and geochemical) changes induced in marly and gypsiferous inter-layers are discussed in Section 5-3. The relevant alteration maps are illustrated in section 5-4.

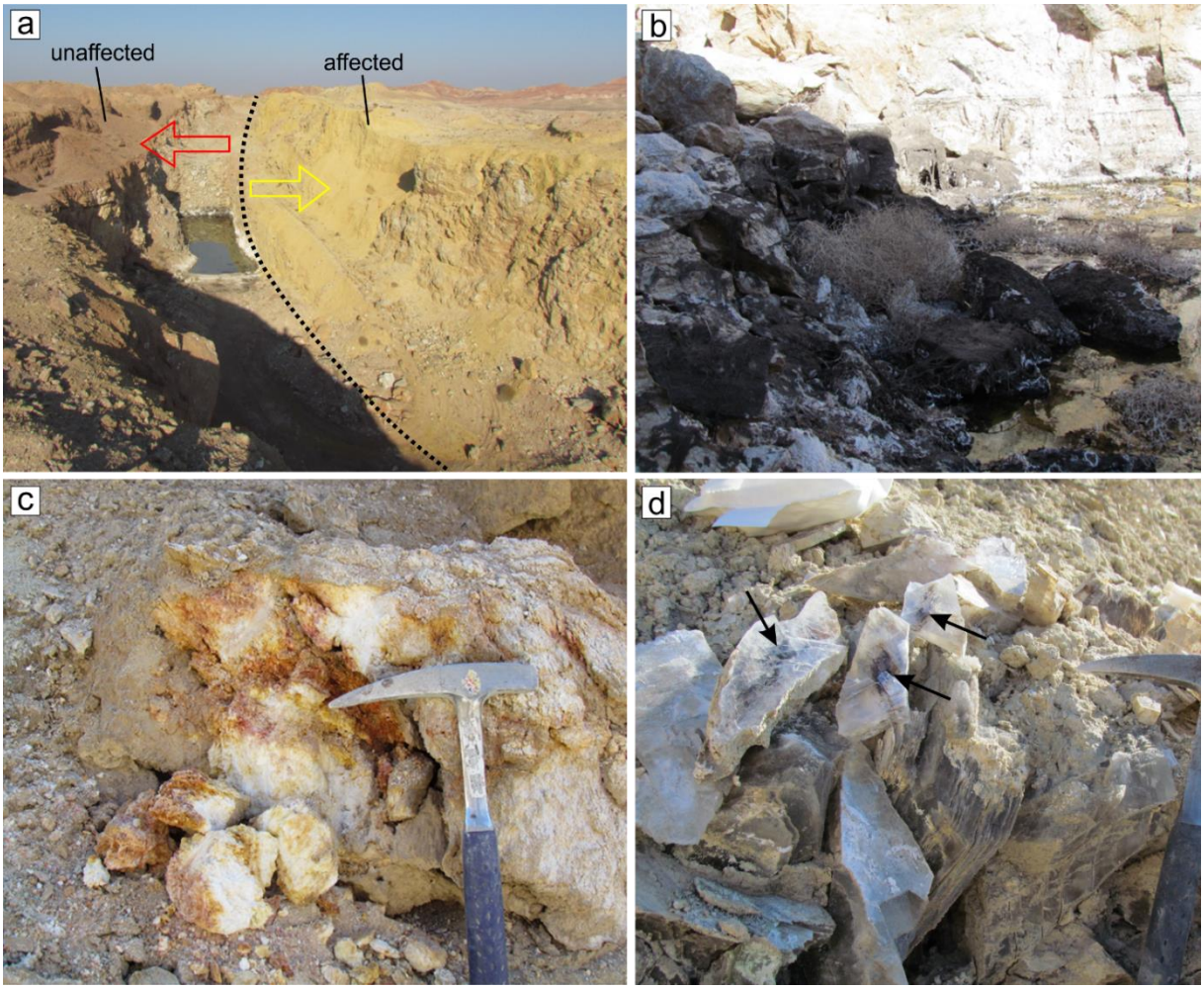

Fig. 7. Photographs illustrating the oil-shows and associated alterations around Mil seepage. **a)** alteration of the marly unit by focused seeping petroleum. The red marl to the left side of the quarry (developed to exploit the gypsum inter-layer; red arrow) is transformed into yellowish hue indicating hematite to goethite transformation. **b)** seeping oil impregnating the rocks at the bottom of the quarry. **c)** local alteration of the gypsiferous unit (Gach-i-Turush) around the seepage. **d)** leaked oil trapped in between the flakes of gypsum crystals.

6.5.2. Petrography

The URF sandstones are texturally immature, poorly rounded, medium to coarse-grained, poorly-sorted, lithic arenite sandstone. The detrital mineralogy was predominantly composed of lithic fragments (>50%), feldspars, quartz, opaque minerals (>10%), and to a lesser extent micas and pyroxenes (Fig. 8a-b). The poorly rounded grains are indicative of the short distance between provenance and the deposition site (Fig. 8b). The unaffected samples were characterized by a pervasive hematite coating (Fig. 8a). The occurrence of iron oxide coatings between the grain contacts implies that oxidizing condition was prevalent in the early interstitial fluids. In the affected zones, the iron coatings are (partially or totally) bleached, but in parts is concentrated as patches of iron oxides. In this facies, calcite was observed as pore-filling cement as well as lithic clasts. The clasts (primary calcite), likely derived from Qom Formation, was distinguished by its shape and by iron oxide coating around the grains (Fig. 8d). The pore-filling cements presumably were precipitated due to microseepage-related alteration. In comparison, the dominant calcite phase in the Mil samples was lithic clasts and pore-filling calcite was rarely observed. The pervasive calcite precipitations in the former correspond to an increased fluid-rock interaction following the dissolution of early hematitic cement and porosity enhancement (Bowen et al., 2007).

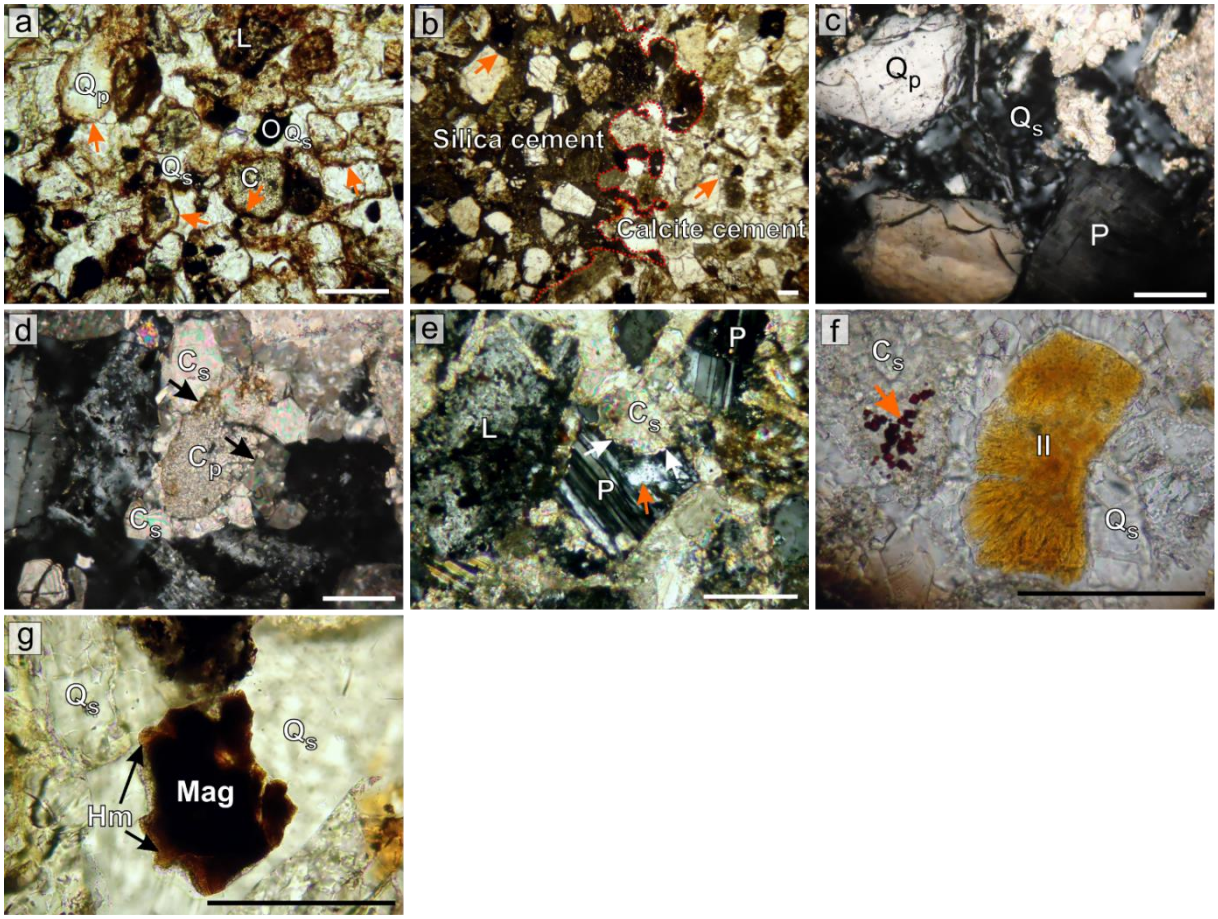

Fig. 8. Photomicrographs of the diagenetic facies and induced alteration in the URF sandstones. **a)** framework grains of the sandstones composing of lithics (L), quarz (Q), plagioclase (P), calcite (C), and opaque (O) minerals. The unaffected sandstones are associated with pervasive hematite grain-coating (arrows). **b)** secondary silica and calcite cements in affected facies. The cementation is associated with partial or complete removal of iron oxide coatings (arrows). **c)** silicification (secondary quartz; Q_s) growing around the primary quartz grain (Q_p). **d)** secondary calcite cements (C_s) growing around the primary calcite clast (C_p) that is partly coated by iron oxides. **e)** alteration of feldspar grains to clays (orange arrow) and dissolution and replacement by calcite cement. **f)** authigenic illite (Il) in the context of silica cements. The red dots (arrow) within secondary calcite are likely magnezioferite crystals. **g)** transformation of magnetite (Mag) into hematite (Hm) at the rims of the grain. Scale bars = 0.1 mm.

Secondary clays were also observed to be abundant in the affected samples. Large grains of secondary illite were detected in between the secondary calcite cements (Fig. 8c). The morphology and locality of the grains pinpoint towards its authigenic origin. Some red grains formed in the vicinity of illite grains (Fig. 8c) were interpreted to be magnezioferite crystals. Silicification (secondary quartz) was also observed to frequently occur in the affected samples, mostly overgrowing around the primary quartz grains (Fig. 8f). The lack of coating around the authigenic quartz grains implies that silicification is a late stage alteration that takes place after hematite coating removal, thus likely corresponds to microseepage effects.

6.5.3. Geochemistry

6.5.3.1. Whole-rock analysis

The results of major oxides analysis indicated significant differences between the specimens from microseepage-affected and unaffected lithofacies within the control area. The weight percent of SiO_2, K_2O, and LOI (loss on ignition) in the affected rocks were enriched, whereas Fe_2O_3, MgO, Na_2O and Al_2O_3 were substantially depleted. The concentration of Fe_2O_3 in the sample suite varies by over 5 wt. % between altered and unaltered lithofacies; whereas the Fe_2O_3 in the unaffected specimens reaches to 8 wt. %, in the affected zones it declines below 3.0 wt.%. The most significant difference, however, belongs to CaO in which its percentage is almost doubled in the altered facies. The average of major oxide composition in either of the environments is illustrated in Fig. 9. SO_3 was determined to be present in just a handful of specimens (n = 5) of the Alborz area, thereby excluded from the charts.

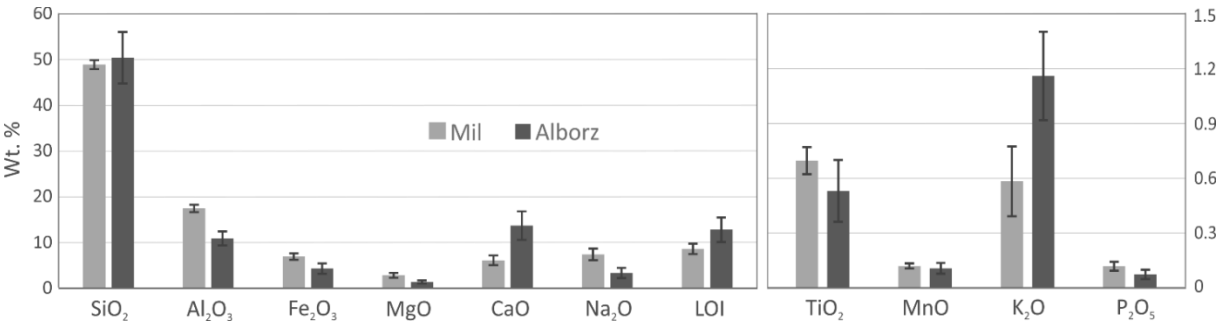

Fig. 9. Bar-chart representation of the major oxide composition of the Alborz and Mil sample suite, respectively corresponding to microseepage-affected and unaffected lithofacies. The bar and whisker represent the mean ± standard deviation, respectively. Note that the two charts have different scales.

Statistical analysis (student's T-test) was deployed to compare the average percentage of corresponding major oxides. The result is summarized in Table 2. The analysis revealed that except SiO_2 and MnO, all the other oxides show meaningful differences from unaffected to microseepage-affected facies. The variations were interpreted to arise from the interaction of microseepage-induced solutions with the host-rocks.

The enrichment of CaO and K_2O corresponds to high concentrations of calcite and clays (illite), respectively. In sandstones, sodium and potassium typically constitute alkali feldspars and muscovite, but are also present in illite and smectites. Calcium, on the other hands, is embedded in calcitic plagioclase, calcite cements, and, to a minor extent, smectitic clays (Boggs, 2009). The loss of Na_2O and Al_2O_3 together with the enrichment of K_2O and CaO is indicative of illite formation at the expense of feldspars alteration as demonstrated in Chapter 5. The alteration is also evident by a significant negative correlation between Al_2O_3 and CaO percentage (R^2 = 0.83), indicating the removal of calcitic plagioclase and simultaneous clay and calcite precipitations. The fact that potassium increases over the reservoir is in contrast to conventional microseepage model in which the acidic solutions associated with bacterial activity stimulate the loss of potassium of the system (Saunders et al., 1999). In addition, the positive correlation of Al_2O_3 with Fe_2O_3 (R^2 = 0.57) might suggest that a portion of the aluminum was depleted from hematite structure corresponding to isomorphic substitution.

Magnesium is contained in chlorite, smectite, and if present, dolomite cements (Boggs, 2009). In Alborz microseepage, MgO was demonstrated to release mainly from Mg-chlorite (see Chapter 5 for details). In the absence of Mg-carbonates to incorporate the released magnesium, it seems the element has been removed from the system resulting in MgO loss in the XRF results. In contrast, the released calcium is incorporated into the structure of calcitic cements. Moreover, the rise in LOI was assumed to correspond to an increase in the CO_2 contents of the bleached lithofacies (because there is a good agreement between CaO and LOI constituents; R^2 = 0.85), and probably also structural water and organic matters of the bedrocks.

Even though silica (SiO_2) shows statistically insignificant differences between the two facies, it is associated with some enrichment in the altered rocks, which is consistent with the silicification witnessed in the petrographic studies and field observations (Sections 5-2 and 5-1, respectively). In contrast to Zheng et al. (2010) who reported the enrichment of P_2O_5 in the petroleum-affected zones (presumed to link to bacterial activity), we observed a reverse trend in the concentration of P_2O_5 over the HC-affected zones, likely because of the alteration of detrital minerals such as apatite. Finally, the loss of Fe_2O_3 is closely related to a drop in TiO_2 percentage (R^2 = 0.69). This could imply that parts of the Ti loss are due to chemical weathering of detrital titanomagnetite, which was observed to occur frequently in sandstone beds of the area.

Table 2. Comparison of the major oxide composition of the affected and unaffected lithofacies using the statistical T-test (assuming equal variances). The checkmark (✓) indicates that the observed difference between the pair is significant, whereas the cross (**X**) indicates insignificant differences.

SiO_2	TiO_2	Al_2O_3	Fe_2O_3	MnO	MgO	CaO	Na_2O	K_2O	P_2O_5	LOI
X	✓	✓	✓	X	✓	✓	✓	✓	✓	✓

In the literature, several indices are proposed to describe lithogeochemical alterations including Ishikawa Index, Silicification Index ($SiO_2/SiO_2 + Al_2O_3$), Sericite Index (K_2O/K_2O+Na_2O), and Chemical Index of Alteration to name but a few (Pirajno, 2009). We found the current alteration indices insufficient to delineate the alterations induced by microseepage system, for all of them were proposed to characterize hydrothermal alteration systems. For this aim, we developed a new index named Microseepage-Induced Alteration Index (MIAI):

$$\text{MIAI} = \frac{\text{CaO}+K_2\text{O}}{Na_2\text{O}+K_2\text{O}+\text{CaO}+\text{MgO}}$$

It constitutes the gained elements divided by the sum of gained and lost elements in the microseepage environment. By applying this index to the dataset, we noticed that in the unaffected zones, the MIAI is generally <0.45, whereas in the microseepage-affected zones, it varies between 0.5 < MIAI < 0.9, but mostly beyond 0.7. The scatterplots for MIAI versus Silicification and Sericite indices are illustrated in Fig. 10. In these plots, the altered lithofacies also retains relatively higher indices of Silicification and Sericite corroborating the formation of clays and silica in the affected facies as interpreted above. The two specimens (Alb-130 and Alb-156) with moderate indices correspond to slightly bleached (altered) facies, whereas the specimen collected over Mil miniseepage (see Section 5-1 and Fig. 7a) retained the highest MIAI and Sericite Index (Fig. 10).

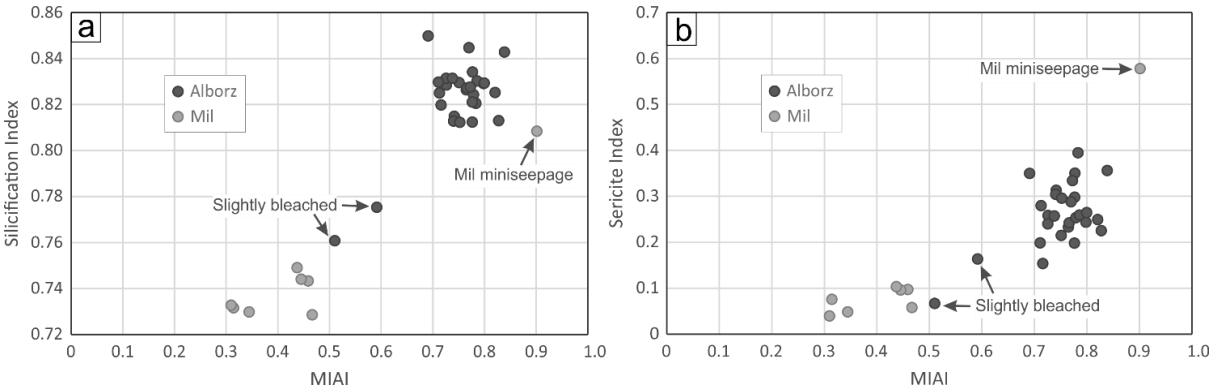

Fig. 10. Scatterplots exhibiting the variation of Microseepage-induced Alteration Index (MIAI) against a) silicification index and b) sericite index. 'Alborz' and 'Mil' correspond to microseepage-affected and unaffected lithofacies, respectively. The indices are defined along the text.

Trace element data indicated a negative anomaly in V, Co, and Sc over HC accumulations compared to Mil area. The elements though show a strong correlation with iron oxide (Fe_2O_3) contents of the samples (R^2 = 0.87, 0.83, and 0.94, respectively). It might imply that the elements were adsorbed by hematite (Beitler et al., 2005), thereby the dissolution of hematite coats stimulated the depletion of these elements as well.

Heavy metals such as U, Pb, and Th were slightly enriched over petroleum deposit (respectively 0.4, 6.4, and 1.7 ppm, on average), whereas Mn, Cr, Ni, Cu, Zn, Mo, Ba, and Sr showed no clear trends. The fact that minor/trace elements show little or no differences between the two facies together with the moderate iron bleaching and abundance of clays indicate that the lithofacies have not been thoroughly flushed with fluids. Such condition might arise from the scarcity of meteoric/ interstitial water and indeed from the alternation of the redox potential hypothesized earlier.

6.5.3.2. REE geochemistry

The patterns of REEs in both Alborz and Mil area show enrichment of LREE (<Eu) relative to HREE, but the two series fail to conform to each other (Fig. 11). This inconstant pattern implies that the provenance of igneous rock debris has not been the same throughout the sedimentary basin, as the REEs are resistant to fractionation during weathering or diagenesis processes (Rollinson, 1993). The REE patterns (together with Th/Sc ratio (<1.0), Th/U ratio (>3.0), and Eu concentration (≈1.0) described in Boggs (2009)) point toward two sediment provenances, which is consistent with the findings of Amini (1997). The contribution of acidic debris from the northern margin seems to be responsible for the enrichment of LREE relative to HREE in Alborz samples. This phenomenon is known to occur due to the abundance of hornblende or accessory minerals such as zircon in the debris (Rollinson, 1993). The great abundance of Zr over the oilfield (114 vs. 68 ppm, which is 67% higher on average) presumably could account for the variations observed in the REEs patterns (Fig. 11). In this context, the two lithofacies (Alborz vs. Mil) are equivalent, hence comparable with each other for the sake of this study.

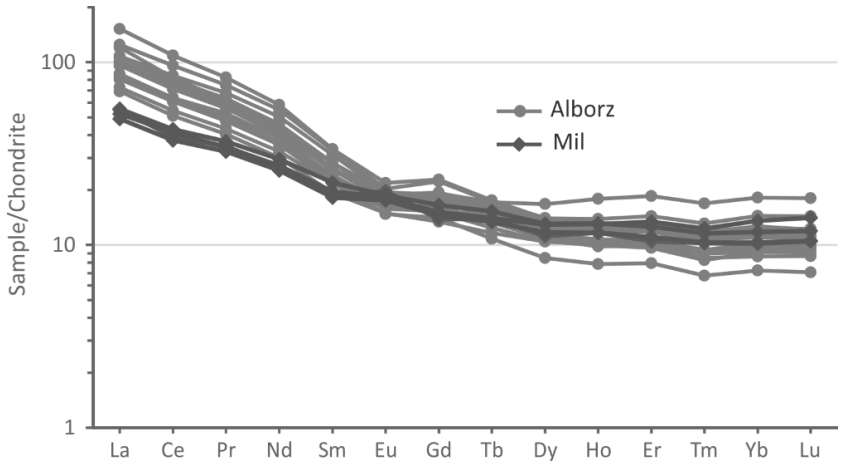

Fig. 11. Chondrite-normalized REE diagrams of the microseepage-affected (Alborz) and unaffected (Mil) lithofacies.

6.5.3.3. X-ray diffraction

X-ray diffraction patterns revealed the mineralogy of the diagenetic alteration to be composed of illite, montmorillonite, palygorskite, calcite, and chlorite (clinochlore) with occasional indications of gypsum (Fig. 12). Although kaolinite was reported to occur in a few samples, the dominant clay mineralogy was illite and smectite (montmorillonite) in which were absent in the Mil sample suite. Calcite was sporadically detected in Mil samples, but it was the dominant carbonate species in the microseepage-affected rocks. Palygorskite was detected in only one sample from the affected zone, albeit was discovered to be abundant in Mil sample suite (Fig. 12).

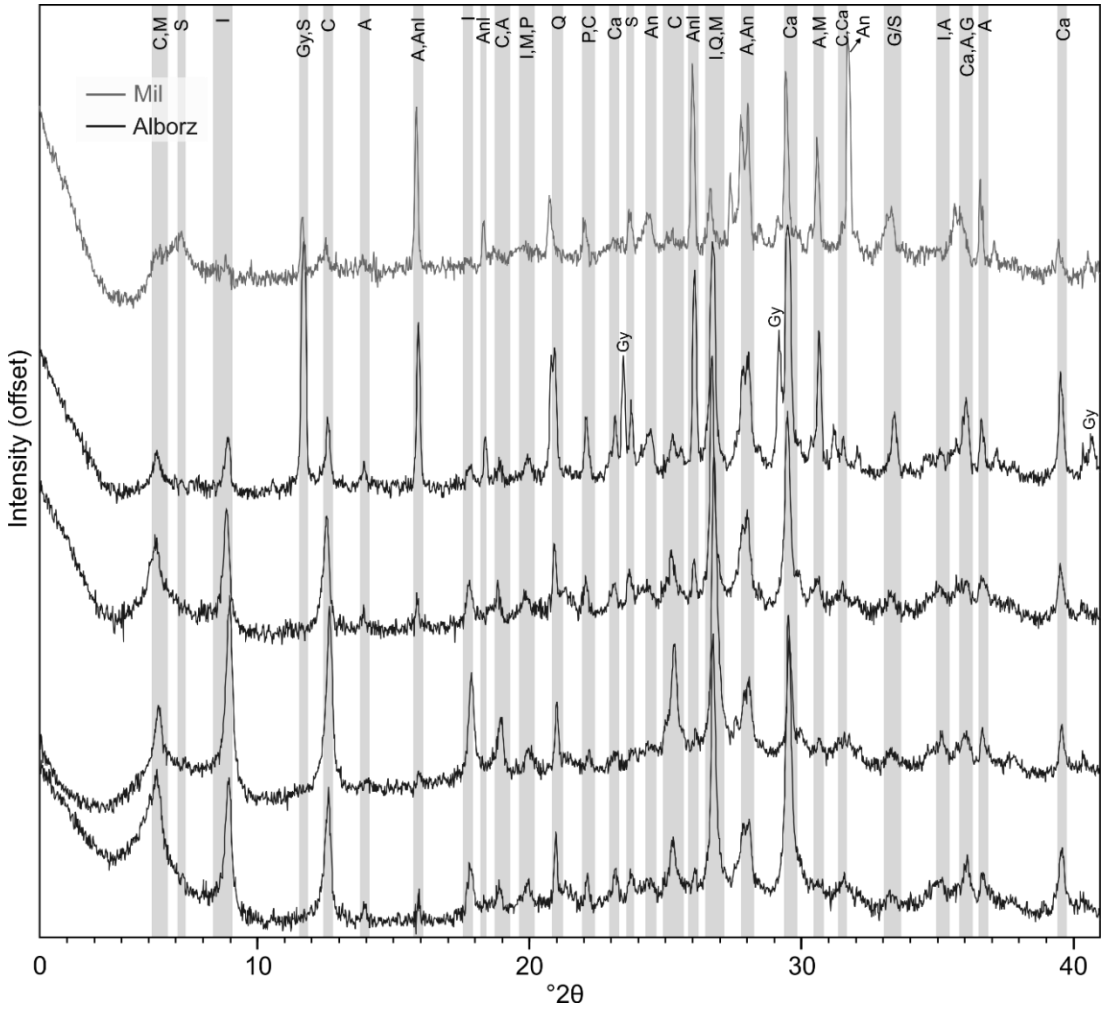

Fig. 12. Examples of X-ray diffraction patterns from affected and unaffected lithofacies along with mineralogical interpretations. The labeled features: chlorite (C), montmorillonite (M), illite (I), palygorskite (P), sepiolite (S), gypsum (Gy), goethite (G), calcite (Ca), quartz (Q), albite (A), anorthite (An), analcime (Anl).

In contrast to XRD, reflectance spectroscopy is highly sensitive to very low levels (<1%) of diagenetic minerals including Fe-chlorite, palygorskite, and nontronite. Accordingly, simultaneous use of the two techniques provided an accurate picture of the changes induced in the microseepage environment.

Table 3. Summary of the mineralogy derived from quantitative XRD analysis. Q = quartz, C = calcite, Pl = plagioclase, M = montmorillonite, Gy = gypsum, Mc = mica, Z = zeolite, Py = pyroxenes, H = hematite, G = goethite, Ch = chlorite.

Mineral	Q	Pl	Py	Mc	Z	M	Gy	C	H/G	Ch
Percentage	7–26	13–26.6	0.8–1.1	1–12	1.4–4.1	0–1.2	0–60	6–30	0–0.6	1–3.5

6.5.3.4. Reflectance spectroscopy

Reflectance spectroscopy indicated the mineralogy of the samples is composed of montmorillonite, illite (and mixed-layer illite/smectite), chlorite, nontronite, hematite-goethite, calcite, with occasional appearance of jarosite and gypsum. The Mil samples tend to contain hematite and traces of palygorskite, chlorite, and sometimes sepiolite (Fig. 13), whereas in the affected zones the same features alter to goethite and smectite-illite. The spectroscopic analysis is discussed further in Chapter 5.

In contrast to XRD, reflectance spectroscopy is highly sensitive to very low levels (<1%) of selected diagenetic minerals including Fe-chlorite, palygorskite, and nontronite. Accordingly, simultaneous use of the two techniques provided an accurate picture of the changes induced in microseepage environment.

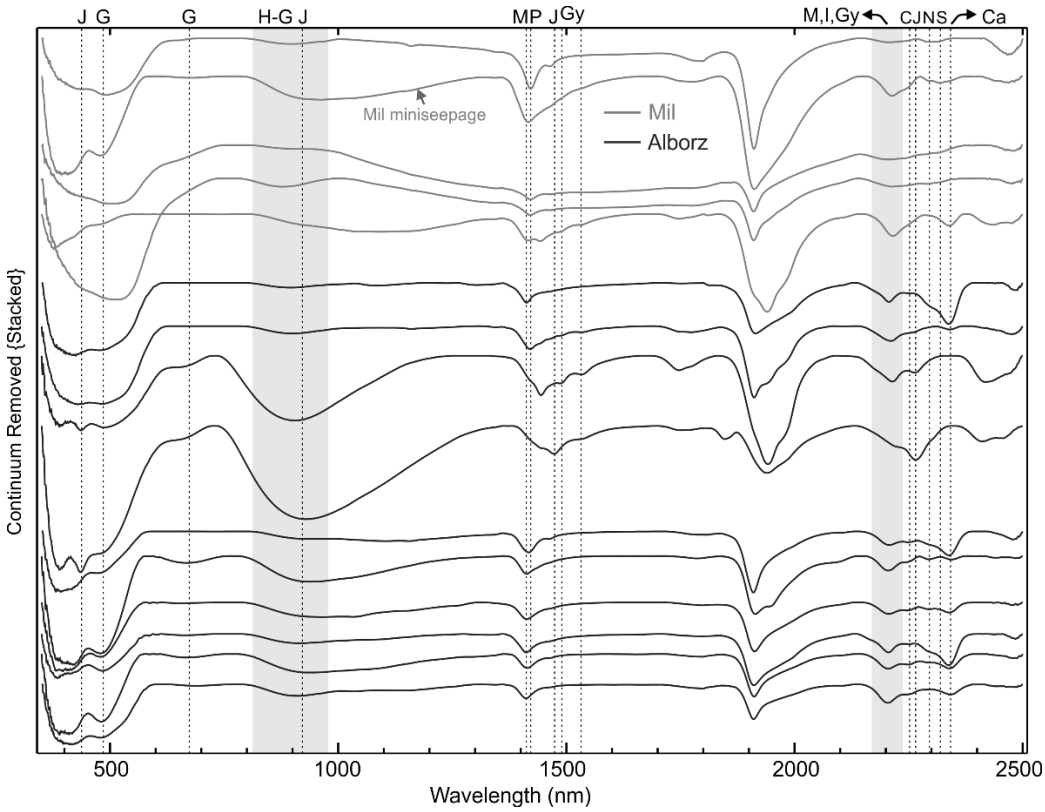

Fig. 13. Representative spectra collected from Alborz and Mil area and corresponding mineralogical interpretations. The labeled absorption features: jarosite (j), goethite (G), hematite (H), montmorillonite (M), palygorskite (P), gypsum (Gy), illite (I), chlorite (C), nontronite (N), sepiolite (S), and calcite (Ca).

6.5.3.5. Stable isotopes

The $\delta^{18}O$ and $\delta^{1}C$ of the calcite vary from -1.23 to -4.67‰ and from -5.77 to -8.36‰, respectively. The isotopic data represent calcite from three different origins and generations: (i) a cluster (encircled in Fig. 14) characterized by $\delta^{13}C$ ranging from -2.1 to -1‰ and $\delta^{18}O$ between -6 to -7‰ corresponding to the bleached marly sandstones located to the north of the West-Alborz. (ii) the samples scattered at one end of the trendline characterized by $\delta^{13}C$ of ∼-2.0 and $\delta^{18}O$ ranging between -8 to -8.5‰ corresponding to unaffected lithofacies. The $\delta^{18}O$ and $\delta^{13}C$ of these samples are slightly lighter when compared to the first group. (iii) the third generation of calcite cement typified by continual depletion of ^{13}C from -2 to -5‰ and enrichment of $\delta^{18}O$ corresponding to highly affected (bleached) facies. The multiple origins of the calcite cements was also corroborated by the lack of meaningful trend between either $\delta^{18}O$ or $\delta^{13}C$ against CaO and carbonate absorption depth.

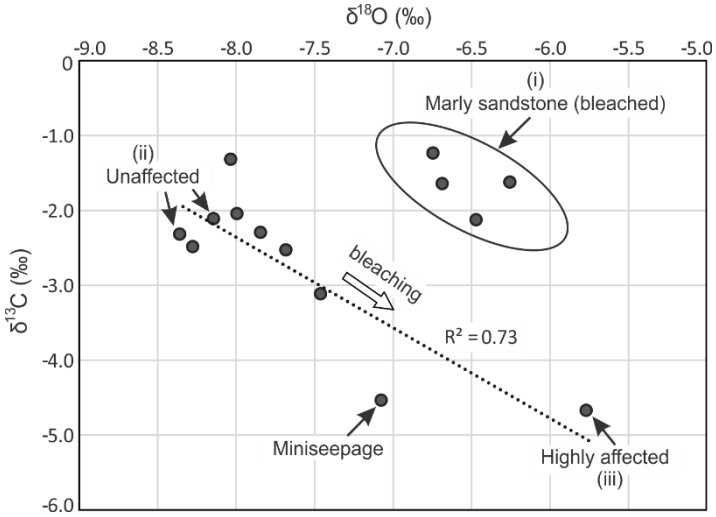

Fig. 14. Cross-plot of $\delta^{18}O_{VPDB}$ against $\delta^{13}C_{VPDB}$ from the carbonate (calcite-dominant) cements/clasts of the URF sandstones. The numbers correspond to the groups described in the text.

The carbonate transported to the system (detrital calcite) has definitely incorporated into the measured isotopic composition; however, as petrographic studies confirmed, these clasts account for a mere 10–20% of the carbonate contents of the samples signifying that the isotopic composition of the bedrocks is determined in large parts by pore-filling calcite cements. The $\delta^{13}C$ and $\delta^{18}O$ composition of the Qom limestone as retreived from well cuttings tend to be ~1 and ~-5‰, respectively (Warren et al., 2014).

On this basis, the unaffected lithofacies (ii) with the severely depleted $\delta^{18}O$, was attributed to a freshwater source, though some contribution from detrital calcite is plausible. On the other hand, the isotopic composition of the altered units in the group (i) was infered to be the results of the interaction between a CO_2-charged brine with meteoric water in shallow groundwater. The carbon of the brine could have been inherited from the dissolution of early marine carbonates (i.e. Qom limestone), albeit it might incorporate carbon from the oxidation of seeping hydrocarbons as well. Presumably, the same fluid is responsible for the pervasive gypsum cementation in the corresponding strata (as described in Chapter 5). The effect of the oxidizing hydrocarbons is more evident in the highly altered specimens (Fig. 14). Here, the meaningful trend (R^2= 0.73) between the two endmembers is further evidence for a mixture of two fluid sources, including the one charged with oxidized seeping hydrocarbons.

Compared to other case studies in the literature (e.g. Donovan et al. (1974)), the $\delta^{13}C$ is not profoundly depleted. This tendency was inferred to reflect the effect of the meteoric water (and in turn the surface evaporation of the meteoric water and Rayleigh distillation) in shaping the final alteration facies (Chan et al., 2000; Donovan et al., 1974). In the altered facies, there exist a positive correlation between $\delta^{18}O$ and iron oxides wavelength position centered at ~900 nm (R^2= 0.24). It could indicate that the enrichment of $\delta^{18}O$ is associated with goethite formation. In other words, the secondary goethite has inherited its oxygen from distilled meteoric water.

6.5.4. Remote sensing alteration mapping

6.5.4.1. Microseepage remote sensing

Visually, the color variability documented in the field is as well evident in the satellite imagery wherein the originally dark red facies in the control area is transformed into reddish to whitish hues over petroleum deposit (Fig. 15a). Moreover, the stratigraphic sequences of the URF, which is clearly visible in the unaffected areas, fades away over the reservoir, presumably because of accelerated erosion associated with authigenic cements dissolution, and because of elimination of the texture/fabric of the original bedrocks[10] (Compare C–D to A–B in Fig. 15a). The lithofacies within the boundary of the structure, however, is not equally affected by petroleum microseepage. This is clearly indicated on the map of hematite shown in Fig. 15b. Here, a large patch of red-beds to the north of East Alborz appears to be hematite-rich and unaffected by diagenetic processes. Nevertheless, a large extent of the URF is affected by microseepage-induced alterations. The regional-scale alteration is characterized by the partial/complete removal of hematite coatings and the formation of secondary iron oxyhydroxides comprising goethite and possibly ferrihydrite/maghemite (Fig. 15c). Goethite shows its greatest abundance along the crest of the anticline, but towards the edges, its proportion shrinks and transitions to bleached facies. The bleached facies then coincides with an excess of clays and secondary carbonates (Fig. 15c-e) that collectively could be indicative of extensive microseepage-induced alterations.

Given the clay alteration map as the indication of diagenetic changes (Fig. 15d), three microseepage-prone areas could be recognized over the anticline including a large zone to the south of East Alborz of which is comprised of several broken anomalies, and two distinct zones on either side of the West Alborz. The anomalies represent an asymmetric microseepage system relative to the crest of the anticline. Whereas the phenomenon could correspond to an annular anomaly over the accumulations (Asadzadeh and Souza Filho, 2017), the exploration drilling (Wells No. 1-4) indicated that at least under a portion of the unaltered terrains no accumulations exists (Fig. 15a). We speculate that this asymmetric pattern mirrors the shape of the reservoir beneath conforming to a thrusted anticline. In other words, the petroleum accumulations tend to occur towards the south of the axis of the anticline; a fact that is corroborated by drilling (Well No. 5 that blew out) and is consistent with the microseepage pattern. The absence of alterations on the northern margin of the East Alborz reservoir along which the thrust faulting occurs is not yet understood. The only changes induced on this end are some narrow and elongated goethitic anomalies that presumably correspond to surface manifestation of thrust faults (Fig. 15c).

Thus far, however, the extent to which the microseepage-induced anomalies coincide with the boundary of the underlying reservoir is not known (at least to the authors). Observations over Serajeh gas field (not discussed here) indicated that microseepage anomaly occurs as haloes around (and over) the surface projection of HC reservoir. Given the annular anomaly pattern, the potential reservoir of Alborz could overlap with altered zones and extend beneath the unaffected red-beds. Based on alteration mapping outcomes, three prospective zones were defined over Alborz reservoir of which only Prospect-1 is tested by drilling (Fig. 15d). Two other microseepage anomalies alongside a large portion of the primary target are yet to be tested by drilling.

[10] This phenomenon is believed to justify the 'hazy anomalies' reported over productive oilfields in the early days of satellite remote sensing.

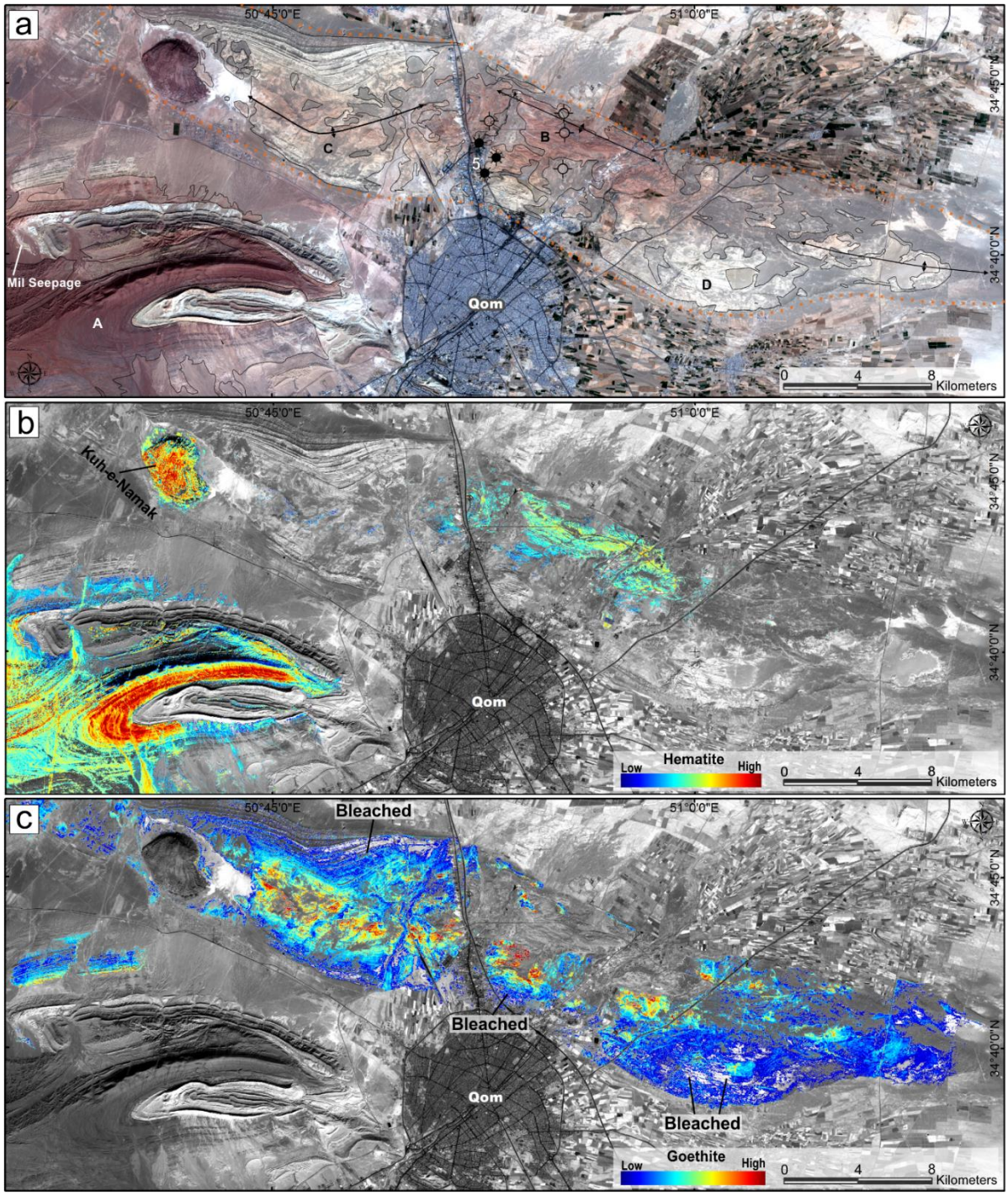

Fig. 15. Microseepage alteration maps derived from Sentinel-2 and ASTER satellite data in Qom study area. **a)** natural color composite imagery. **b)** the abundance and distribution of hematite corresponding to unaffected lithofacies. **c)** the abundance and distribution of goethite/ferrihydrite corresponding to affected lithofacies.

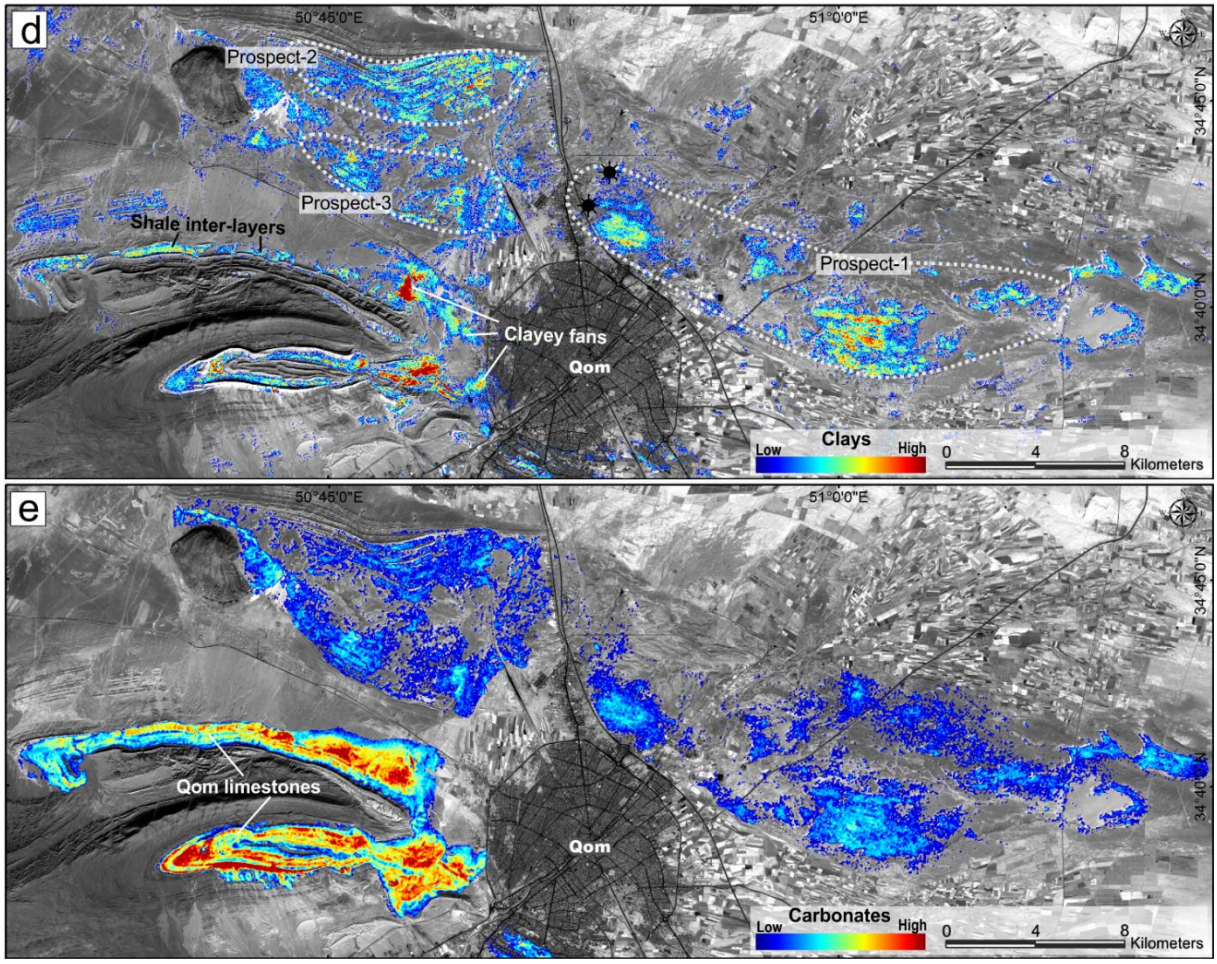

Fig. 15 (Continue). d) the abundance and distribution of clay minerals (smectites and mixed-layer illite/smectite). **e)** the abundance and distribution of carbonates (calcite). The dashed white lines indicate the outline of prospects over the structure.

The concern arose as to whether the recorded anomalies are induced by microseepage systems. Besides the evidence presented above, there are several lines of evidence that support the interpretation is favor of microseepage-induced alterations rather than an intrinsic property of stratigraphic variations: (i) the diagenetic changes were observed only over the Alborz reservoir. In Mil area (and another control site some 60 km to the north of this area not reported here), the delineated indicators were largely absent. (ii) the changes fail to show any spatio-temporal relationships with the stratigraphy of the area; for instance, the same lithofacies altered to the south of East Alborz is unaffected in northern edge. Even along the same beds from east to west, a similar pattern could not be followed (Fig. 15a). (iii) there are unaffected inter-layers between the intensely bleached facies. (iv) the scale in which iron is mobilized (and then removed or re-precipitated) is very large; such a regional-scale iron oxides reduction could only arise from a reservoir-wide microseepage system.

Overall, the resulting mineral maps show a good agreement with field observations and spectroscopic studies. However, due to the limitations imposed by multispectral data, it was not possible to differentiate between species of clay minerals or map the compositional variations associated with, for example, white micas.

6.5.4.2. Miniseepage remote sensing

Remote sensing data was successful in mapping the local alterations induced by Mil miniseepage (Fig. 16). The marly unit surrounding the gypsiferous inter-layer tends to retain red color that is indicative of hematite coating (Fig. 16b). Albeit, due to miniseepage effects, its eastern contacts with gypsum layer has transformed into goethite coats (Fig. 16c). On the other hand, the ASTER data collected before the exploitation of the gypsiferous layer provided an opportunity to target sour gypsum (Gach-i-Turush) through the area (Fig. 16a). We presumed that the miniseepage affected zones should retain very strong water absorptions at 2200 nm (Hunt, 1977) and cover wide areas, for the alteration is known to yield powdery gypsum. Accordingly, we decided to vist the localities wherein gypsum showed abnormal intensities (Fig. 16a). By this way, we verified the occurrence of Gach-i-Turush is several localities. Moreover, a new series of paleo-seeps were detected along the outcrops of the gypsiferous unit in the vicinity of altered gypsum. The locations are indicated by stars in Fig. 16a.

These maps show that in comparison to microseeps, miniseepage alterations around oil and gas indications constitute smaller targets. Accordingly, their detection would not be straightforward using satellite remote sensing technology.

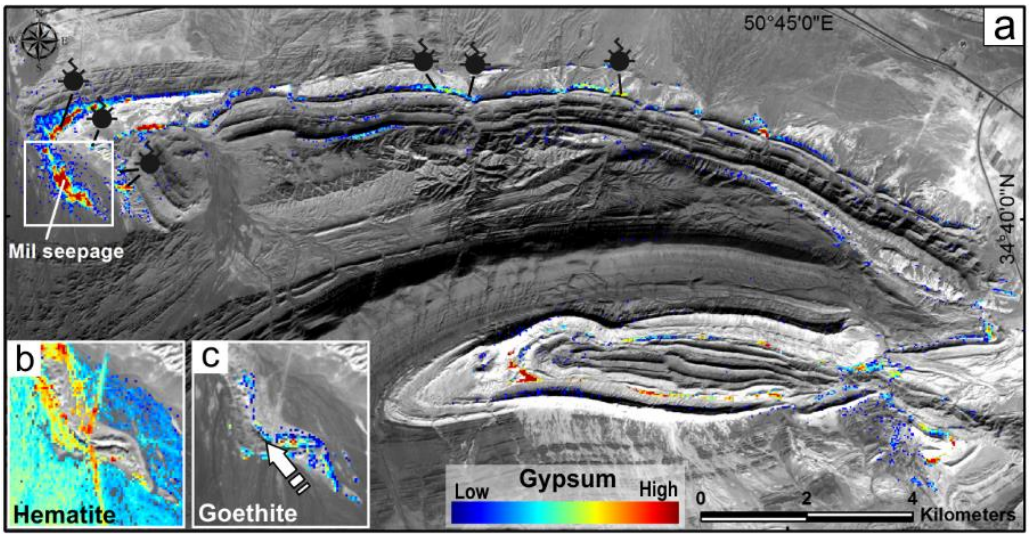

Fig. 16. Mineral maps derived from satellite data over Mil area. **a)** abundance and distribution of gypsiferous inter-layer. The circles pinpoint the location of paleo-seeps verified on the ground. **b)** abundance of hematite and **c)** goethite surrounding the Mil seepage (miniseepage). The maps in b-c correspond to the rectangle exhibited in (a). The arrow in (c) indicates the position of the photograph shown in Fig. 7a.

6.6. Conclusion

This study indicated that even in the case of near perfect evaporitic capping, the reservoir still leaks HC to the surface. The induced physio-chemical and mineralogical changes were shown to be a function of the geochemistry of the lithofacies hosting the microseepage, the local environmental variables (i.e. climate and hydrogeology among others) and likely the characteristics of the reservoir (i.e. pressure, leakage rate, etc.) and the chemistry of the accumulation (i.e. H_2S proportion). The mafic lithofacies in this area (i.e. URF) coupled with an alternation between oxidizing and reducing regime had given rise to a new array of mineralogical and geochemical indicators for microseepage

systems. The major oxides proportions were substantiated to be a reliable indicator of microseepage-induced alteration. A new index based on the major oxides was proposed for geochemical evaluation of microseepage alterations. Regarding the iron oxides, we indicated that the composition of iron oxides/oxyhydroxides along with its proportion (bleaching) could constitute additional indicator for microseepage remote sensing. The clay mineralogy on the other hand, was revealed to be dominantly smectites, mixed-layer illite/smectite, and chlorite with no traces of kaolinite. Nevertheless, the collection could not be mapped remotely due to the limitation imposed by multispectral data.

The extent of microseepage-induced alteration within this area was indicated to be much larger than what was known before. The distribution of anomalous zones was shown to conform to the annular pattern. Further drilling is required to verify the occurance of economic petroleum accumulations in the prospects.

The findings of this study would be helpful in distinguishing charged traps from dry plays and likely will increase the efficiency of remote sensing techniques for oil and gas exploration through eliminating false-positive and false-negative anomalies. The global coverage of the datasets would facilitate the mapping of active microseepage systems all around the world in a fast and cost-effective way.

Acknowledgement

We are greatly thankful to Ricardo Perobelli and Rômulo Angélica for instructions on XRD analysis, Luiza Mendes for help with clay separations, Aline Belila for assistance in optical microscopy, Joao Paulo Pitombeira for tips on REE data processing, Lena V. Soares (USP) for performing stable isotope analysis, and Javad Bagheri for his accompany during fieldworks. The exploration directorate of the NIOC (National Iranian Oil Company) is acknowledged for supporting the fieldworks and sample collections, and permission to publish the data. We appreciate the logistics provided by Hezareh Sevom Company in the field. The authors thank FAPESP for the research grants No. 2015/06663-7.

References

Abrams, M., Hook, S., 2000. ASTER User Handbook. Jet Propulsion Laboratory, Earth Remote Sensing Data Analysis Center, Pasadena, CA.

Al Shaieb, Z., Cairns, J., Puckette, J., 1994. Hydrocarbon-induced diagenetic aureoles: Indicators of deeper, leaky reservoirs. Association of Petroleum Geochemical Explorationists Bulletin 10, 24–48.

Amini, A., 1997. Provenance and depositional environment of the Upper Red Formation, Central Zone, Iran. University of Manchester, Manchester, UK, p. 320.

Amini, A., 2001. Red colouring of the upper red formation in central part of its basin, Central zone, Iran. Journal of Sciences, Islamic Republic of Iran 12, 145-156.

Asadzadeh, S., Souza Filho, C.R., 2017. Spectral remote sensing for onshore seepage characterization: A critical overview. Earth-Science Reviews 168, 48-72.

Beitler, B., Parry, W.T., Chan, M.A., 2005. Fingerprints of Fluid Flow: Chemical Diagenetic History of the Jurassic Navajo Sandstone, Southern Utah, U.S.A. Journal of Sedimentary Research 75, 547-561.

Berberian, M., King, C.C.P., 1981. Toward a paleogeography and tectonic evolution of Iran. Canadian Journal of Earth Sciences 18, 210-265.

Boggs, S.J., 2009. Petrology of sedimentary rocks, Second ed. Cambridge University Press, NY, USA.

Bowen, B.B., Martini, B.A., Chan, M.A., Parry, W.T., 2007. Reflectance spectroscopic mapping of diagenetic heterogeneities and fluid-flow pathways in the Jurassic Navajo Sandstone. AAPG Bulletin 91, 173-190.

Chan, M.A., Parry, W.T., Bowman, J.R., 2000. Diagenetic Hematite and Manganese Oxides and Fault-Related Fluid Flow in Jurassic Sandstones, Southeastern Utah. AAPG Bulletin 84, 1281-1310.

Clarke, R.H., Cleverly, R.W., 1991. Petroleum seepage and post-accumulation migration. Geological Society, London, Special Publications 59, 265-271.
Donovan, T.J., 1974. Petroleum Microseepage at Cement, Oklahoma; Evidence and Mechanism. AAPG Bulletin 58, 429-446.
Donovan, T.J., Friedman, I., Gleason, J.D., 1974. Recognition of Petroleum-bearing Traps by Unusual Isotopic Compositions of Carbonate-cemented Surface Rocks. Geology 2, 351-354.
Eichhubl, P., Taylor, W.L., Pollard, D.D., Aydin, A., 2004. Paleo-fluid flow and deformation in the Aztec Sandstone at the Valley of Fire, Nevada—Evidence for the coupling of hydrogeologic, diagenetic, and tectonic processes. Geological Society of America Bulletin 116, 1120-1136.
El Bishlawy, S., Sehim, A., El Sabbagh, M., Habo, M., 2001. Soil gas survey detects microseepage through thick evaporitic sequence in Egypt. Oil and Gas Journal 99.
Ellwood, B.B., Burkart, B., 1996. Test of hydrocarbon-induced magnetic patterns in soils: the sanitary landfill as laboratory, in: Schumacher, D., Abrams, M.A. (Eds.), Hydrocarbon migration and its near-surface expresseion. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 91-98.
Ferguson, J.D., 1976-1979. The subsurface alteration and mineralization of Permian red beds overlying several oil fields in southern Oklahoma. The Shale Shaker Digest IX XXVII-XXIX, 17-33.
Gansser, A., 1957. Die geologische erforschung der Qum gegend, Iran. Bulletin der Verengung Schweizerisches Petroleum-Geologen und -Ingenieur 23, 1-16.
Garden, I.R., Guscott, S.C., Burley, S.D., Foxford, K.A., Walsh, J.J., Marshall, J., 2001. An exhumed palaeo-hydrocarbon migration fairway in a faulted carrier system, Entrada Sandstone of SE Utah, USA. Geofluids 1, 195-213.
Holysh, S., Toth, J., 1996. Flow of formation waters: Likely cause for poor definition of soil gas anomalies over oil fields in east-central Alberta, in: Schumacher, D., Abrams, M.A. (Eds.), Hydrocarbon migration and its near-surface expression. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 255-277.
Horvitz, L., 1985. Geochemical Exploration for Petroleum. Science 229, 821-827.
Hunt, G.R., 1977. Spectral signatures of particulate minerals in the visible and near infrared. Geophysics 42, 501-513.
Kirkland, D.W., Denison, R.E., Rooney, M.A., 1995. Diagenetic alteration of Permian strata at oil fields of south central Oklahoma, USA. Marine and Petroleum Geology 12, 629-644.
Lilburn, R.A., Zuhair, A.-S., 1984. Geochemistry and isotopic composition of Hydrocarbon-Induced Diagenetic Aureole (HIDA), Cement Field Oklahoma. The Shale Shaker Digest XI XXXIII-XXXV, 46-72.
Link, W.K., 1952. Significance of oil and gas seeps in world oil exploration. AAPG Bulletin 36, 1505-1540.
Macgregor, D.S., 1993. Relationships between seepage, tectonics and subsurface petroleum reserves. Marine and Petroleum Geology 10, 606-619.
McDonough, W.F., Sun, S.s., 1995. The composition of the Earth. Chemical Geology 120, 223-253.
Ming-Kuo, L., Bethke, C.M., 1994. Groundwater flow, late cementation, and petroleum accumulation in the Permian Lyons Sandstone, Denver Basin. AAPG Bulletin 78, 217-237.
Morley, C.K., Kongwung, B., Julapour, A.A., Abdolghafourian, M., Hajian, M., Waples, D., Warren, J., Otterdoom, H., Srisuriyon, K., Kazemi, H., 2009. Structural development of a major late Cenozoic basin and transpressional belt in central Iran: The Central Basin in the Qom-Saveh area. Geosphere 5, 325-362.
Morley, C.K., Waples, D.W., Boonyasaknanon, P., Julapour, A., Loviruchsutee, P., 2013. The origin of separate oil and gas accumulations in adjacent anticlines in Central Iran. Marine and Petroleum Geology 44, 96-111.
Mostofi, M., Gansser, A., 1957. The story behind the 5 Alborz. Oil and Gas Journal 55, 78-84.
Nielson, G.B., Chan, M.A., Bowen, B., 2014. Iron-rich horizons in the Jurassic Navajo Sandstone, southwestern Utah: progressive cementation and permeability inversion, in: MacLean, J.S., Biek, R.F., Huntoon, J.E. (Eds.), Geology of Utah's Far South. Utah Geological Association Publication, pp. 215-236.
Parry, W.T., Chan, M.A., Nash, B.P., 2009. Diagenetic characteristics of the Jurassic Navajo Sandstone in the Covenant oil field, central Utah thrust belt. AAPG Bulletin 93, 1039-1061.
Petrovic, A., Khan, S.D., Thurmond, A.K., 2012. Integrated hyperspectral remote sensing, geochemical and isotopic studies for understanding hydrocarbon-induced rock alterations. Marine and Petroleum Geology 35, 292-308.
Pirajno, F., 2009. Hydrothermal Processes and Mineral Systems. Springer Science.
Price, L.C., 1986. A critical review and proposed working model of surface geochemical exploration, in: Davidson, M.J. (Ed.), Unconventional methods in exploration for petroleum and natural gas. Southern Methodist University Press, Dallas, pp. 245–304.
Rainoldi, A.L., Franchini, M., Beaufort, D., Patrier, P., Giusiano, A., Impiccini, A., Pons, J., 2014. Large-Scale Bleaching of Red Beds Related To Upward Migration of Hydrocarbons: Los Chihuidos High, Neuquén Basin, Argentina. Journal of Sedimentary Research 84, 373-393.
Reuter, M., Piller, W.E., Harzhauser, M., Mandic, O., Berning, B., Rogl, F., Kroh, A., Aubry, M.-P., Wielandt-Schuster, U., Hamedani, A., 2007. The Oligo-Miocene Qom Formation (Iran): Evidence for an early Burdigalian restriction of the Tethyan Seaway and closure of its Iranian gateways. International Journal of Earth Sciences 98, 627-650.
Rollinson, H.R., 1993. Using geochemical data: evaluation, presentation, interpretation, First ed. Longman Scientific & Technical, UK.
Salati, S., van Ruitenbeek, F.J.A., de Smeth, J.B., van der Meer, F.D., 2014. Spectral and geochemical characterization of onshore hydrocarbon seep-induced alteration in the Dezful embayment, southwest Iran. AAPG Bulletin 98, 1837-1857.
Saunders, D.F., Burson, K.R., Thompson, C.K., 1999. Model for hydrocarbon microseepage and related near-surface alterations. AAPG Bulletin 83, 170-185.

Schumacher, D., 1996. Hydrocarbon-induced alteration of soils and sediments, in: Schumacher, D., Abrams, M.A. (Eds.), Hydrocarbon migration and its near-surface expression. AAPG Memoir 66, Tulsa, OK, U.S.A., pp. 71-89.
Schumacher, D., 2012. Pre-drill prediction of hydrocarbon charge: Microseepage-based prediction of charge and post-survey drilling results, CWLS GeoConvention 2012. AAPG, Calgary, AB, Canada, p. 9.
Schumacher, D.D., 2000. Surface geochemical exploration for oil and gas: New life for an old technology. The Leading Edge 19, 258-261.
Simpson, C.J., Wilford, J.R., Macias, L.F., Korsch, R.J., 1991. Remote sensing of petroleum gas seepage alteration Palm Valley Gas Field, Amadeus Basin, central Australia., Proceedings of the Eighth Thematic Conference on Geologic Remote Sensing. ERIM, Denver, Colorado, USA, pp. 1463-1472.
Tedesco, S.A., 1995. Surface geochemistry in petroleum exploration. Chapman and Hall, Inc, New York.
Thomas, A.N., 1952. "Gach-i-turush" and associated phenomena in southwest Persia, V11 Convegno Nazionale Del Metano E Del Petrolio Taormina, section 1, preprint.
Warren, J., Morley, C.K., Charoentitirat, T., Cartwright, I., Ampaiwan, P., Khositchaisri, P., Mirzaloo, M., Yingyuen, J., 2014. Structural and fluid evolution of Saraburi Group sedimentary carbonates, central Thailand: A tectonically driven fluid system. Marine and Petroleum Geology 55, 100-121.
Worden, R.H., Smalley, P.C., Fallick, A.E., 1997. Sulfur cycle in buried evaporites. Geology 25, 643-646.
Zheng, G., Fu, B., Takahashi, Y., Kuno, A., Matsuo, M., Zhang, J., 2010. Chemical speciation of redox sensitive elements during hydrocarbon leaching in the Junggar Basin, Northwest China. Journal of Asian Earth Sciences 39, 713-723.

Chapter 7

Investigating the capability of WorldView-3 superspectral data for direct hydrocarbon detection[11]

The recently launched WorldView-3 (WV-3) satellite is a high spatial resolution instrument with eight multispectral bands in the visible and near-infrared and an additional eight bands in the short-wave infrared (SWIR). Three of the SWIR bands, including bands 9, 12, and 16 (centered at 1210, 1730, and 2330 nm) respectively overlap with diagnostic absorption features of hydrocarbons (HCs) at 1200, 1700, and 2300 nm. This chapter (paper) aims to investigate the capability of this superspectral instrument for direct HC detection. For this purpose, we have conducted several simulation experiments using multiple datasets comprising (i) spectral libraries of different HCs measured in the laboratory, (ii) close-range hyperspectral imagery of a well-known tar-sand sample acquired with a sisuCHEMA imaging system, and (iii) far-range ProSpecTIR hyperspectral imagery collected over twelve simulated HC-shows. These datasets were convolved to the spectral resolution of WV-3 and analyzed using a variety of spectral processing techniques. The absorption features of HCs manifest themselves in all cases albeit with varying intensity. The effect of a series of parameters on the detectability of the HCs was also scrutinized; these included background geology, spectral mixing, HC type, endmember set, spatial resolution, noise level, and topography. We demonstrate that the HC absorption feature in WV-3's band 12, accompanied by shoulders sustained at bands 11 and 13 (centered at 1660 and 2165 nm), is resilient enough and persist under various conditions. Potential applications of these finding include hydrocarbon exploration in frontier basins and environmental monitoring.

7.1. Introduction

Hydrocarbon (HC) is an organic compound found in petroleum (crude oil), bitumen, tar, and kerogen. The detection and mapping of HC is important, because it provides a compelling direct evidence for the generation of thermogenic HC in basins and plays (Macgregor 1993), and simultaneously, because it is currently one of the large contributors to anthropogenic pollution (Brown 2009; Jha et al. 2008). In offshore basins, a diverse range of remote sensing technologies are used to map the extent of oil spills and/or oil seeps, including synthetic aperture radar (SAR), laser fluorescence, thermal infrared (based on apparent thermal inertia), light detection and ranging (Lidar), and hyperspectral imaging (Fingas and Brown 2014; Leifer et al. 2012). SAR is by far the most extensively used spaceborne technology for this aim. However, it is unable to identify oil type and to estimate the thickness of oil spills (Brekke and Solberg 2005).

A number of structural bonding in HCs, including C–H, CH_2, and CH_3 give rise to stretching fundamentals between 3000-7000 nm regions. The overtones and combinations of these fundamentals fall within the near-infrared (NIR) and short-wave infrared (SWIR) wavelengths. Pronounced HC features in the NIR-SWIR window include a triplet between 1700-1750 nm due to simultaneous overtone/combination, and a doublet between 2290-2360 nm due to combination of C–H stretching. Occasionally, a singlet subtle absorption feature between 1150-1230 nm, related to the second overtone of the C-H stretching, is also present (Clark et al. 2009; Cloutis 1989; Kallevik et al. 2000; Lammoglia and Souza Filho 2011).

[11] This chapter is published in the Remote Sensing of Environment journal.

Several case studies have demonstrated the merit of hyperspectral data for direct HC detection using the aforementioned absorption features. For instance, oils occurring naturally among onshore seeps in a sedimentary basin, or the oil-contaminated sands in an urban area have been detected by airborne hyperspectral imageries (Ellis et al. 2001; Hörig et al. 2001; Kühn et al. 2004; Prelat et al. 2013). Similarly, the distribution of oil from the Deepwater Horizon spill in coastal marshes was successfully mapped with Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) hyperspectral data (Kokaly et al. 2013). In offshore basins, imaging spectroscopy was demonstrated to have the groundbreaking potential to quantify oil thickness (for slicks >0.1 mm thick) and the oil to water ratio (Clark et al. 2010; Leifer et al. 2012). Moreover, infrared spectroscopy has proved to be an effective tool in identifying oil types with varying API gravity (Lammoglia and Souza Filho 2011).

Despite the encouraging results from optical remote sensing for direct[12] oil detection and characterization, such studies have been confined to airborne systems, which are generally of limited access, costly, and with small areal coverage. In contrast, current broadband multispectral orbital instruments with global coverage (e.g. the Landsat series or Advanced Spaceborne Thermal Emission and Reflection Radiometer; ASTER), are seldom used for this aim. The main reason is probably the lack of spectral bands required to resolve the C-H absorption features (Fig. 1). For instance, AVIRIS data degraded to 30 m spatial resolution of Landsat ETM were 74% accurate in detecting oiled marshes (against 93% of AVIRIS at 7.6 m resolution); however, the same data resampled to the spatial and spectral resolution of Landsat ETM (six spectral bands; Fig. 1) failed to detect any oiled targets (Kokaly et al. 2013). The ASTER instrument is likely to resolve the 2300 nm absorption feature when not mixed with other compounds (Fig. 1). However, the ASTER SWIR detectors are no longer functioning and data acquired since April 2008 are not useable. Regarding oil characterization, ASTER instrument has been merely used to estimate the API gravity of oil slicks emanating from offshore subsurface reservoirs (Lammoglia and Souza Filho 2012). Other indirect assessment methods like the time series analysis of vegetation greenness due to oil spills extracted from Landsat ETM imageries (Mishra et al. 2012) was believed to be too complicated to reflect the real impacts of oil on an ecosystem (Kokaly et al. 2013).

The successful launch of WorldView-3 (WV-3) satellite by Digital Globe in August 13/2014 has opened up new opportunities to map mineral assemblages, vegetation cover and man-made materials with greater spatial and spectral resolution (Kruse and Perry 2013). The coincidence of specific WV-3 bands with HC features centered at 1200, 1700, and 2300 nm motivated us to investigate whether the sensor has the capability to directly and unambiguously detect HCs.

In the absence of real WV-3 data acquired over a known HC outcrop, we have simulated WV-3 data using datasets from i) spectral libraries of different HCs measured in the laboratory, ii) a close-range hyperspectral imagery (sisuCHEMA), and iii) a far-range (airborne) ProSpecTIR hyperspectral imagery collected over twelve man-made HC-shows. Each dataset was convolved to the spectral response of the WV-3 instrument, and then analyzed using several spectral processing techniques and a suite of parameters including background geology, spectral mixing, HC type, endmember set, spatial resolution, noise level, and topography.

[12] Here we emphasize the “direct” detection to differentiate it from “indirect” sensing, which is used for oil and gas exploration through mapping of alteration minerals above HC accumulations; for details see (Van Der Meer et al. 2002).

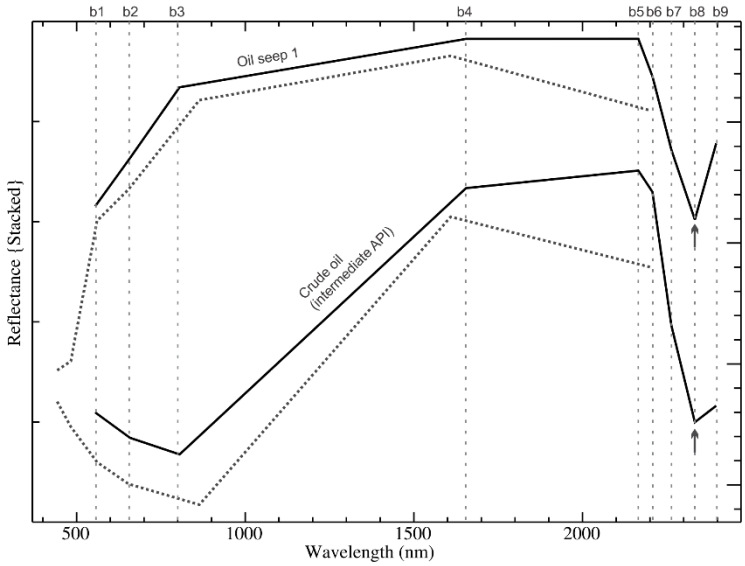

Fig. 1. Reflectance spectra of two oil-bearing samples resampled to the response functions of ASTER (solid line) and Landsat-8 (dashed line). Note that only ASTER can resolve the HC feature at 2300 nm (shown by the arrow). Vertical dotted lines define ASTER band positions (numbered b1 to b9). Samples are described in section 3-1.

7.2. Specifications of the WorldView-3 instrument

WV-3 is a sun-synchronous, high spatial resolution commercial satellite with moderately large spectral bands. In the visible and near infrared, like its predecessor WorldView-2, it provides four standard color bands and four additional bands named coastal, yellow, red edge, and near-IR2. The major advancement of WV-3, however, is the addition of eight SWIR bands in the 1195-2365 nm window (Fig. 2). Furthermore, twelve additional CAVIS (Clouds, Aerosols, Vapors, Ice, and Snow) bands guarantee proper estimates of aerosol and water vapor for atmospheric compensation of the data. The platform operates at an altitude of 617 km. The spatial resolutions at nadir are 0.31 m (panchromatic), 1.24 m (VNIR), 3.7 m (SWIR), and 30 m (CAVIS). The SWIR dataset, however, will only be released commercially at 7.5 m resolution in 14-bits dynamic range. The swath width of the instrument is 13.1 km, with the descending node at 10:30 A.M.(DigitalGlobe 2014).

A comparison between the spectral bands of WV-3 with that of ASTER and OLI/Landsat-8 reveals the full strengths of this new instrument (Fig.2). In the VNIR range, WV-3 surpasses both of them with several additional bands (Fig. 2), which potentially enables a more accurate mapping of iron oxides. WV-3 also introduced four SWIR bands in 1200-1750 nm region relative to a single band in both ASTER and Landsat-8 instruments (Fig. 2). WV-3's SWIR bands 13 to 16 (spanning between 2160-2330 nm; Table 1) are approximately identical to ASTER's SWIR bands 5 to 8 (spanning between 2165-2330 nm). However, the equivalent to ASTER's SWIR band 9 (centered at 2395 nm) is missing in WV-3. Consequently, while most of ASTER's capabilities are expected to be inherited by WV-3, it is unlikely that WV-3 could map mineral groups like Mg-OH or carbonates as accurate as ASTER. In general, though, WV-3 should significantly outperform ASTER because of additional SWIR bands and higher spatial resolution (Fig.2).

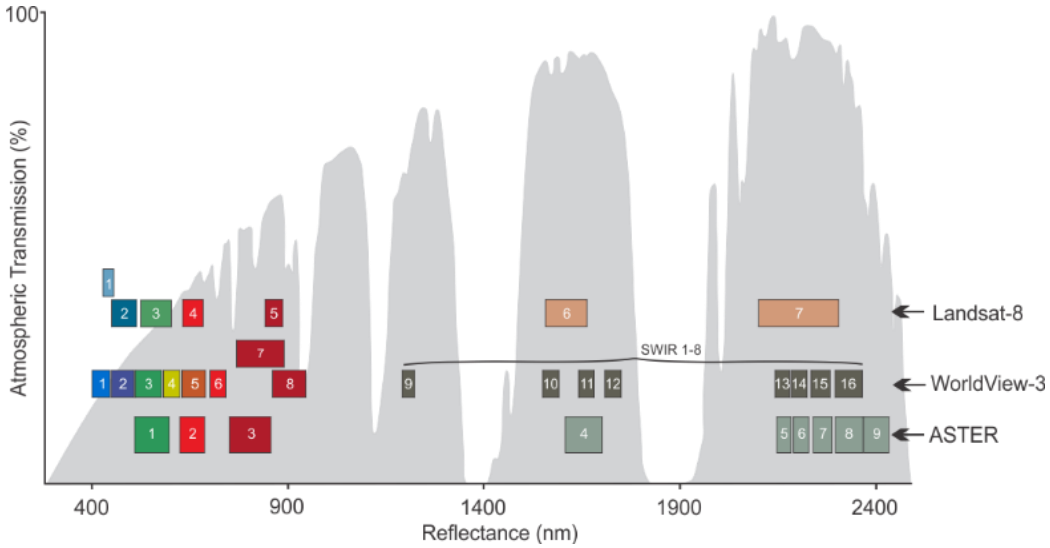

Fig. 2. Comparison of WV-3 spectral bands with ASTER and OLI/Landsat-8 multispectral instruments. Bars represent the full width at half-maximum (FWHM) coverage of each numbered band. The atmospheric transmission that is adapted from USGS (http://landsat.usgs.gov/landsat8.php) is plotted in the background in grey.

Table 1. Characteristics and wavelength ranges of WV-3 spectral bands (source: Digital Globe).

Sub-system	Band number	Band name	Wavelength range (nm)	Band centers (nm)	Sub-system	Band number	Band name	Wavelength range (nm)	Band centers (nm)
VNIR	1	Coastal	400 - 450	425	SWIR	9	SWIR-1	1195 - 1225	1210
	2	Blue	450 - 510	480		10	SWIR-2	1550 - 1590	1570
	3	Green	510 - 580	545		11	SWIR-3	1640 - 1680	1660
	4	Yellow	585 - 625	605		12	SWIR-4	1710 - 1750	1730
	5	Red	630 - 690	660		13	SWIR-5	2145 - 2185	2165
	6	Red edge	705 - 745	725		14	SWIR-6	2185 - 2225	2205
	7	Near-IR1	770 - 895	825		15	SWIR-7	2235 - 2285	2260
	8	Near-IR2	860 - 1040	910		16	SWIR-8	2295 - 2365	2330

7.3. Materials and methods

7.3.1. Spectral library

To investigate the spectral behavior of HC-bearing materials at WV-3 spectral resolution, four sample suites were collected. Table 2 lists the four suites, their location, and any accompanying minerals. The crude oil suite extracted from Brazilian basins is comprised of three samples with different API gravities: heavy (API<22), intermediate (22<API<30), and light (API>30) oils. The tar-sand suite (a heavy hydrocarbon with API<10 and varying clay content) was extracted from an exhumed hydrocarbon reservoir host in early Triassic sandstones of the Parana basin. The sampling site was near the town of Anhembi, some 170 km to the NW of Sao Paulo city in Brazil. The oil seep suite was collected from an active onshore macroseep occurring in gypsiferous strata. The seep is associated with low-angle thrust faults penetrating the Alborz structural trap (Miocene) in Qom area,

central Iran. The tar-like, oxidized solid oil was taken from the remnant of a wildcat oil gusher over the Alborz anticline in Qom area (Table 2).

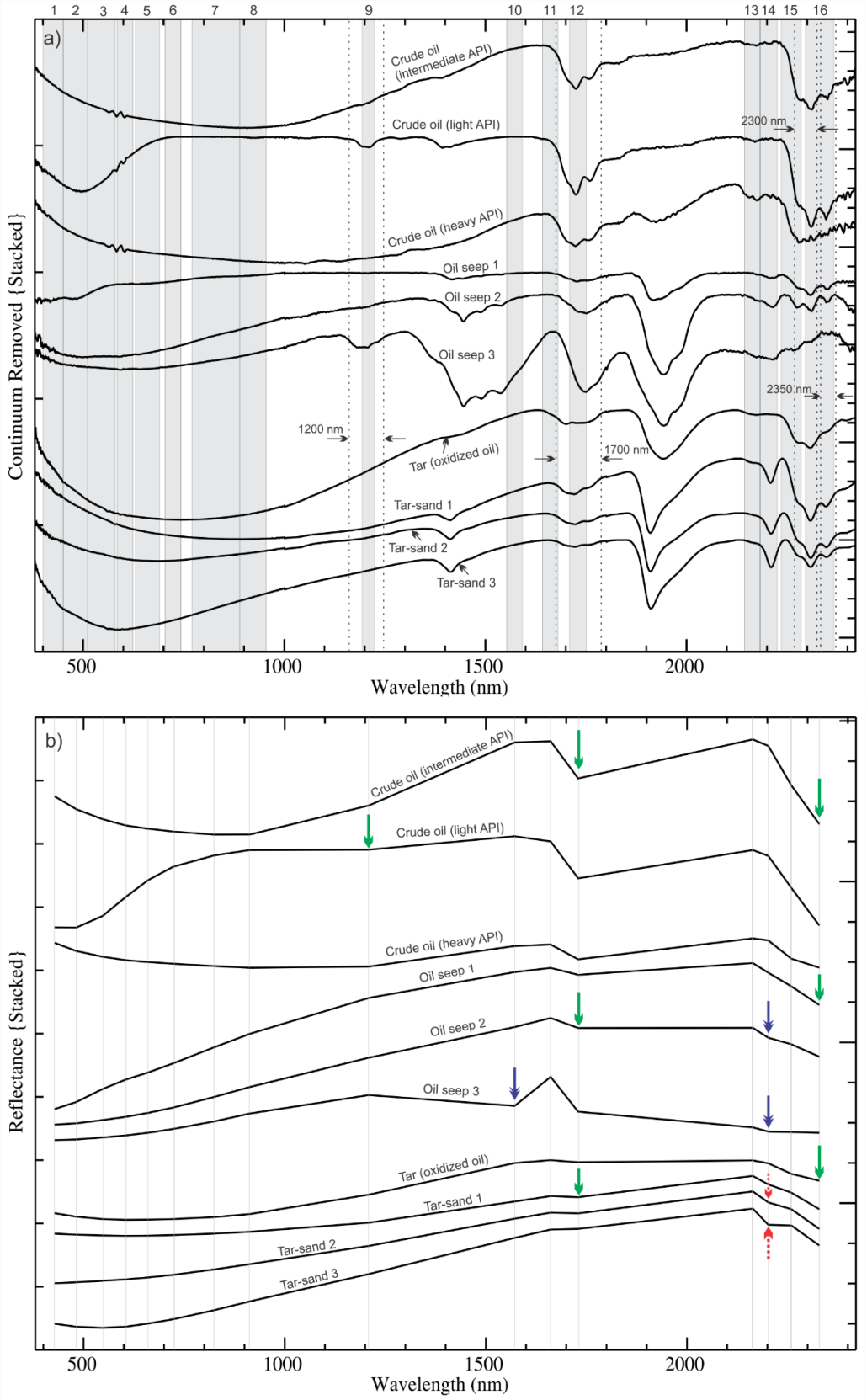

Fig. 3. Reflectance spectra of HC-bearing compounds acquired from a variety of samples including crude oils, tar-sands, tar, and oil-seeps of Iran and Brazil. **a)** Continuum removed spectra measured in the laboratory using a FieldSpec-4 spectrometer. The wavelength range of

the WV-3 bands is represented by grey bars, and band numbers are printed on top of each bar. The wavelength region of the main HC absorption features is shown by parallel dashed lines and also are enclosed by horizontal arrows. **b)** The same spectra shown in (a) convolved to WV-3 spectral response; vertical solid lines indicate the central wavelength of each band. Vertical arrows denote the absorption features resolved by the sensor; those related to HC are shown in green, gypsum in blue (double arrow), and montmorillonite in dashed red. Note that spectra given in (b) are not continuum removed.

The solid samples were collected from surface outcrops and kept in fully sealed, thick plastic bags. The liquid samples were collected in specially sealed jars that are standard in the petroleum industry. The samples were collected during 2011 to 2013 and since then have been stored at ambient temperature in the laboratory. The four suites were measured in the laboratory using an ASD's FieldSpec-4 spectrometer (covering the spectral range between 350-2500 nm at 1 nm intervals) under artificial illumination from a contact probe. For each measurement, 50 scans were used to reduce the contribution of instrument noise. The measured spectra were converted to absolute reflectance using a Spectralon panel. All the measurements were corrected for splice error and then convolved to WV-3 spectral resolution using spectral response functions of the sensor provided by Digital Globe. The full resolution and resampled spectra are shown in Fig. 3.

Table 2. Summary of the samples used to study in the laboratory.

Sample suite	Number of samples	Accompanying material(s)	Sampling site
Crude oil	3	-	Brazilian basins
Tar-sand	3	Montmorillonite, quartz	Anhembi area, Brazil
Oil seep	3	Gypsum	Qom area, Iran
Tar	1	-	Alborz anticline, Iran

7.3.2. SisuCHEMA hyperspectral dataset

In the close-range laboratory experiment, we used hypercube data of the tar-sand sample. The sample is dominated by bitumen, but montmorillonite is also present as inter-layers and small spots in a quartz-dominant matrix (Fig. 4). This particular tar-sand sample was chosen for the study because of its unique mineralogy (a binary mixture; as quartz is spectrally featureless in the SWIR region), and because of its compositional and spectral similarity to oxidized petroleum. More importantly, once proved to be detectable by WV-3's SWIR bands, tar-sands could be potentially targeted by this sensor in exploration surveys.

The spectral image was collected using a sisuCHEMA^{TM} hyperspectral imaging instrument (Specim Ltd., Finland). The sisuCHEMA system consists of a moving sample tray, an illumination lamp, and a high-resolution spectral camera that measures radiance in the SWIR (Table 3). The image cube acquired by this system is converted to reflectance using a white Spectralon reference panel and built-in routines of the instrument. The tar-sand sample shown in Fig. 4 was scanned by the system and an image cube with 256 spectral bands between 930-2500 nm and average spatial resolution of 150 µm was obtained. In order to produce symmetric pixels, the image was resampled to 300 µm using 0.53 and 0.43 multipliers for X and Y directions, and a pixel aggregate function. In order to suppress noise, statistical and Sav-Gol smoothing filters were applied to the imagery. The final cube was then resampled to WV-3's spectral response, yielding eight spectral bands in the SWIR region (Fig. 7b). In the processing stage, we worked out a series of simple calculations over the main absorption features of bitumen by applying band ratioing (BR) and relative-absorption band depth (RBD) techniques (Crowley et al. 1989) to the resampled and original datasets.

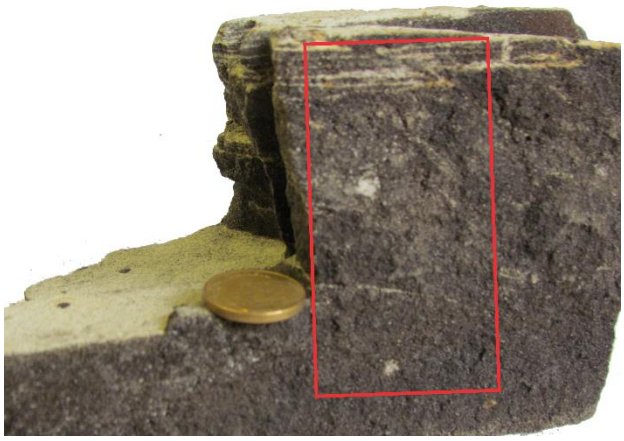

Fig. 4. Photograph of the tar-sand sample used in the close-range study. The area inside the red box was scanned by the sisuCHEMA imaging system.

Table 3. Technical specification of the sisuCHEMA imaging system

Spectral range (nm)	Spectral bands	Spectral resolution (nm)	Pixel size (μm)
930-2500	256	10	30-300

7.3.3. ProSpecTIR dataset

For the far-range experiment, we employed the dataset acquired over simulated (man-made) HC-shows in the Casper testing center, located in Wyoming, USA. In this campaign, twelve boxes (2.44 × 2.44 × 0.15 m) were filled with three types of mineral substrates and four types of liquid HCs. The mineral substrates were sandy, dolomitic and clayey soils with varying grain sizes (Fig. 5). The liquid HCs used were diesel, gasoline, light oil (API = 29), and heavy oil (API = 41) (Table 4). Each box was impregnated with 36 liters of HCs (Fig. 5d), and 30 minutes later, they were imaged by the ProSpecTIR-VS hyperspectral instrument. The spatial arrangement of the boxes and a description of their content are provided in Fig. 6 and Table 4.

The ProSpecTIR™ airborne instrument has two major sub-systems operating in the VNIR and SWIR. They cover the wavelengths ranges between 400-970-2450 nm with 357 spectral bands at respectively 2.9 and 8.5 spectral resolutions. The signal-to-noise ratio (SNR) of the instrument is higher than 500:1. In this study, the data were acquired with a ground sampling distance (GSD) of 0.6 m. The data were atmospherically compensated and then spectrally resampled to WV-3 spectral resolution, yielding sixteen spectral bands with 0.6 m spatial resolution (Fig. 6).

Primarily, we aimed to use the airborne hyperspectral dataset to understand the detectability of different HCs in intimate mixtures with distinct soils using three breeds of processing algorithms, notably BR and RBD, spectral correlation mapper (SCM), and match filtering (MF) (Boardman et al. 1995; Carvalho Junior and Menezes 2000). Next, to appreciate the effect of GSD on the detectability of HCs, we resampled the dataset to lower resolutions (i.e. 1.2, 2.4, 3.7, 4.2, 6.0, and 7.5 m; in which the 3.7 m is the standard WV-3 resolution and 7.5 m is the commercial release of SWIR sub-system) using a pixel aggregate function.

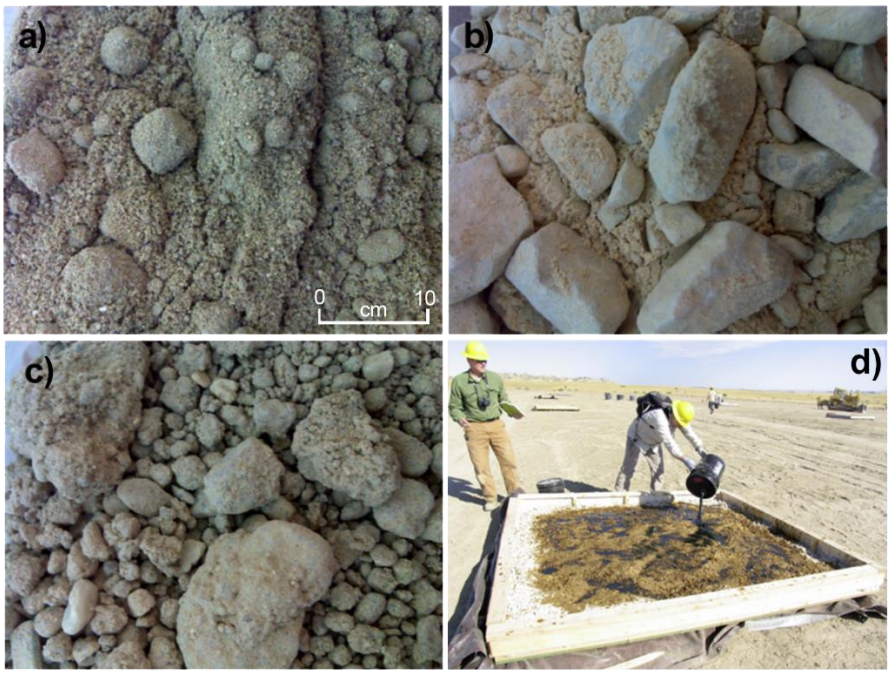

Fig. 5. Photograph of the geologic substrates used to fill each box in the airborne simulation experiment. **a)** Sandy soil. **b)** Dolomitic soil. **c)** Clayey soil. **d)** Illustration of the impregnation of the substrates with HCs (here light crude oil). The scale for (b) and (c) is similar to (a).

Table 4. A simple description of the mineralogic and HC content of boxes used with airborne simulation experiment. The arrangement of the boxes is shown in Fig. 6.

Box No	Substrate	HC type
1	Sand	Crude oil (°API 29)
2	Sand	Crude oil (°API 41)
3	Sand	Diesel
4	Sand	Gasoline
5	Dolomite	Crude oil (°API 29)
6	Dolomite	Crude oil (°API 41)
7	Dolomite	Diesel
8	Dolomite	Gasoline
9	Clay	Crude oil (°API 29)
10	Clay	Crude oil (°API 41)
11	Clay	Diesel
12	Clay	Gasoline

Primarily, we aimed to use the airborne hyperspectral dataset to understand the detectability of different HCs in intimate mixtures with distinct soils using three breeds of processing algorithms, notably BR and RBD, spectral correlation mapper (SCM), and match filtering (MF) (Boardman et al. 1995; Carvalho Junior and Menezes 2000). Next, to appreciate the effect of GSD on the detectability of HCs, we resampled the dataset to lower resolutions (i.e. 1.2, 2.4, 3.7, 4.2, 6.0, and 7.5 m; in which the 3.7 m is the standard WV-3 resolution and 7.5 m is the commercial release of SWIR sub-system) using a pixel aggregate function.

Fig. 6. Natural color composite of the ProSpecTIR data resampled to WV-3 resolutions showing the arrangement of twelve boxes containing different substrates impregnated with distinctive HCs. The arrow indicates the flight direction.

Furthermore, to test the robustness of the HC features against noise, scaled random Gaussian noise was systematically added to the cube using the following formula:

$$\rho_N(\lambda) = \rho(\lambda)\left(1 + \frac{R(\lambda)}{SNR}\right) \quad (1)$$

where $\rho_N(\lambda)$ is the noise-added pixel spectrum, $\rho(\lambda)$ is the original WV-3 spectrum, R(λ) is a random number from Gaussian distribution (generated using the computer clock), and SNR is the desired SNR level (Rodger et al. 2012). Thirteen SNR levels (30, 40, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, and 500:1) were used in the computational experiment.

Since direct HC detection has been hindered by brightness confusion with low albedo pixels resulting from shade or moisture (Ellis et al. 2001; van der Werff 2006), we employed a rugged terrain (with 800 m elevation difference) to simulate the effect of topography on the HC signatures considering sun illumination at mid- latitude winter (20 of December) and summer (20 of June) as follows:

$$Modeled\ image = Normal\ image\left(\frac{Cos\ i}{Cos\theta}\right) \quad (2)$$

where i is the local incidence angle, and θ is the solar zenith angle.

$$Cos\ i = Cos\beta Cos\theta + Sin\beta Sin\theta Cos(\gamma - \varphi) \quad (3)$$

where β is the terrain slope, γ is the local terrain aspect, and φ is the solar azimuth angle (Feng et al. 2003).

7.4. Results and discussion

7.4.1. FieldSpec spectral dataset

The spectral plots of Fig. 3 shows that WV-3's bands 9, 12, and 16 approximately overlap with the absorption features of HCs at 1200, 1700, and 2300 nm. Among them, band 12 is the most interesting for HC detection because it forms a relatively deep and stable absorption feature associated with a sharp left-side shoulder (band 11), as well as another broader right-side shoulder (band 13) (Fig. 3b). Band 9 is not associated with a proper "shoulder" in its vicinity, thus it is unlikely that this single band could uniquely resolve the 1200 nm absorption feature linked to HCs. However, in the wavelength region between bands 9 and 10 (920-1640 nm) some heavier HCs (e.g. bitumen and oil seeps; Fig.3a) display a uniform positive slope. The resampled spectral library also revealed a negative slope between bands 15 and 16 and occasionally between bands 13 to 16. These patterns, however, are easily affected and even reversed due to the spectral mixing of HCs with mineral substrates. For example, the presence of gypsum in oil-seep 3 or increasing amount of clays (montmorillonite) in tar-sand 1 to 3 (Fig. 3b) affect the absorptions features centered at bands 12 and 16, or their corresponding shoulders at bands 11, 13, and partly band 15. In general, the 2300 nm features result in more intense absorptions, however, it is potentially obscured by the absorption features of clays or carbonates and is not fully resolved by the WV-3 sensor. The 1700 nm absorption feature, on the other hand, albeit less pronounced, is not overlapped by absorption features of common minerals (Cloutis 1989; Kühn et al. 2004). However, there are reports of false positive anomalies due to the interfering effect of dried vegetation with absorption features at the same wavelength region (Elvidge 1990; Kokaly et al. 2013).

7.4.2. SisuCHEMA hyperspectral dataset

We calculated a series of BRs including b10/b9, b11/b12, b13/b12 and b15/b16, as well as the RBD index using (b11+b13)/b12 in order to test the performance of WV-3 bands for detecting bitumen in the tar-sand and map the extend of bitumen in the rock. The results are illustrated in Fig. 7. We used the original hypercube data and calculated an RBD index to cross-validate the image maps produced at WV-3 spectral resolution (Fig. 7h). The index ($R_{1657} + R_{1664} + R_{1801} + R_{1808}$ / $R_{1720} + R_{1726} + R_{1733} + R_{1739}$) was yielded by incorporating reflectance bands (R) positioned at i) the shoulders and ii) the absorption minima of the 1700 nm feature as numerator and denominator, respectively.

The band ratios were in good agreement with the bitumen content of the sample (evident as dark patches in Fig. 7b), with the exception of band ratio b11/b12. These maps have successfully discriminated montmorillonite (white patches in Fig. 7a and b) from bitumen and defined the relative extend of bituminous sands well. The correlation coefficient between the RBD index obtained from the original hypercube data and those obtained from the resampled data are reported in Table 5. The best correlation (0.93) is achieved when a combination of three bands (b11, b12, b13) are used, whereas poor results are yielded when using two of these bands only (b11/b12). The b10/b9 and b15/b16 ratios, which benefit from the spectral slope of HCs, produced moderately successful results. It seems, however, that the poor performance of the b11/b12 ratio is likely due to either nonlinear mixing effect of montmorillonite (section 4.1), the steep slope of bitumen towards shorter wavelength, or the inherent noise of the hyperspectral imaging system. This is because removal of the continuum from the spectra rose the correlation coefficient to 0.96. Thus, the relative

importance in the performance of the WV-3 bands for detecting HC features are as follows: b12 >b16 >b9. This simple band calculation using a sisuCHEMA hyperspectral dataset shows that the spectral response of WV-3 can preserve the HC absorption bands, particularly around 1700 nm, and more importantly, the feature can be used for HC mapping. It is worth noting that based on the results achieved from hyperspectral studies, the performance of HC absorption features for characterizing petroleum could be prioritized as: 2300 nm >>1700 nm >>1200 nm (Clark et al. 2010; Ellis et al. 2001; Hörig et al. 2001; Kokaly et al. 2013).

The depth of the 1700 nm absorption feature of tar-sands is shown to be correlated with the volumetric bitumen content of a sample (Donkor et al. 1995; Shaw and Kratochvil 1990). Here, the continuum-removed band-depth of this feature at 1726 nm was calculated using the original hyperspectral data and compared to similar parameter computed from the resampled WV-3 data at band 12. The high correlation coefficient for the two images (> 0.95%) implies that WV-3 is probably able to also estimate the HC fraction of a pixel.

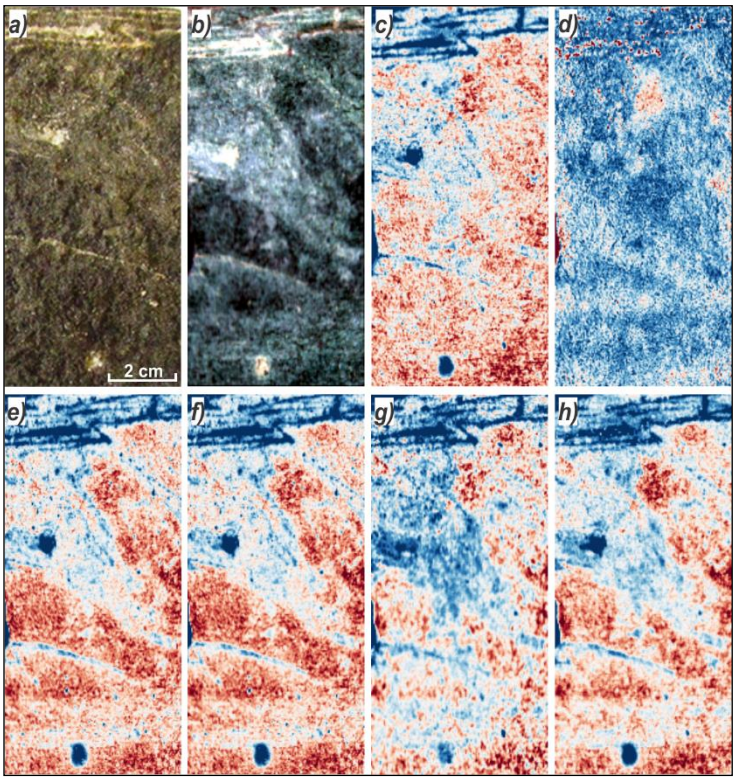

Fig. 7. a) Photograph (true color) of the tar-sand sample scanned by the sisuCHEMA imaging system. **b)** False color composite image (RGB=b9, b12, b16) of the tar-sand sample after resampling to WV-3's SWIR resolution. **c)** b10/b9 ratio image used to highlight the positive slope in HC spectra between 900-1600 nm. **d)** b11/b12 ratio image. **e)** b13/b12 ratio image. **f)** RBD of the 1700 nm feature using (b11+b13)/b12. **g)** b15/b16 ratio image. **h)** RBD ($R_{1657} + R_{1664} + R_{1801} + R_{1808} / R_{1720} + R_{1726} + R_{1733} + R_{1739}$) of the original hypercube data used to map the 1700 nm feature. The stretching thresholds for all color-coded figures (c) to (h) are between 3.5% for the lower values (shown in blue) and 99.5% for the higher values (shown in red). The scale in (a) applies to all images.

Table 5. The correlation coefficient calculated between the RBD image from the original hypercube data and BRs and RBD computed from the resampled WV-3 dataset.

Correlation coefficient		WV-3 bands				
		b10/b9	b11/b12	b13/b12	(b11+b13)/b12	b15/b16
SisuCHEMA bands	$(R_{1657} + R_{1664} + R_{1801} + R_{1808}) / (R_{1720} + R_{1726} + R_{1733} + R_{1739})$	0.83	0.22	0.91	0.93	0.88

7.4.3. ProSpecTIR dataset

We processed the resampled WV-3 resolution dataset using three breeds of image processing techniques namely BR/RBD, SCM, and MF. The criterion to evaluate the results was the number of oil-bearing pixels highlighted in each box after applying a user-defined threshold. This is reported as hit pixels per box or pixel # (Fig. 8). To avoid any bias in thresholding, we used predefined fractions (e.g. >99.55%), but case-to-case changes were also made to eliminate commission errors as much as possible (especially with the null margins of the image). Ultimately, only those pixels that occurred exclusively inside the limits of the boxes or in their immediate vicinity were counted (Fig. 8). The limit of each box was initially determined by processing the original ProSpecTIR data.

7.4.3.1. Assessing the performance of the WV-3 bands

Here, we repeated the same image processing techniques explored in section 4.2., but also analyzed the performance of each index for spectral mixing. The results are displayed in Table 6 as pixel # scores. The RBD index detected the presence of oils in all of the boxes (Fig. 8a) and also highlighted the highest number of pixels. The b13/b12 ratio yielded the poorest results (Fig. 8b), because three of the boxes were overlooked altogether using this ratio. The b11/b12 ratio provided the second best performance; however, one of the targets was missed by this ratio (box 9; Table 6).

Table 6. The number of oil-bearing pixels detected in each box using BR and RBD methods applied to the WV-3 dataset (GSD=0.6 m). Number 1 to 12 denote the boxes described in Table 4.

BR/RBD	(score; pixel #) per box												Total	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Pixels**	**Boxes**
b11/b12	9	23	17	20	18	32	25	31	0	14	17	10	216	11
b13/b12	0	9	9	10	0	12	17	14	0	9	6	4	90	9
b10/b9	10	15	14	14	8	19	14	17	4	4	8	2	129	12
b15/b16	8	9	9	8	2	10	13	12	9	10	9	8	107	12
(b11+b13)/2×b12	9	20	18	22	13	34	27	36	4	19	16	13	231	12

The background mineralogy and oil-type content of the boxes (Fig. 9) indicate that the b13/b12 ratio is very sensitive to HC type. The presence of clays in the background could also affect both b13/b12 and b10/b9 ratios. Three of the ratios (b13/b12, b10/b9, and b15/b16) showed average performance levels for HC detection. Heavier crude oil (°API 29; oil-29) proved to be the most difficult target to detect by spectral band ratios, whereas the detection of lighter oil (°API 41; oil-41) was relatively straightforward (Fig. 9b). Dolomitic soils stood out as the best background for HC detection examined here even though the b15/b16 ratio showed poorer results (Fig. 9a), which is likely due to the overlap of HCs and carbonate absorption bands around 2300nm. Clayey soil, in contrast, appears to be a problematic background for HC detection in the same way discussed in

section 4.2. In general, the combination of b11, b12, and b13 seems to work well for differentiating various HCs in varied background geology.

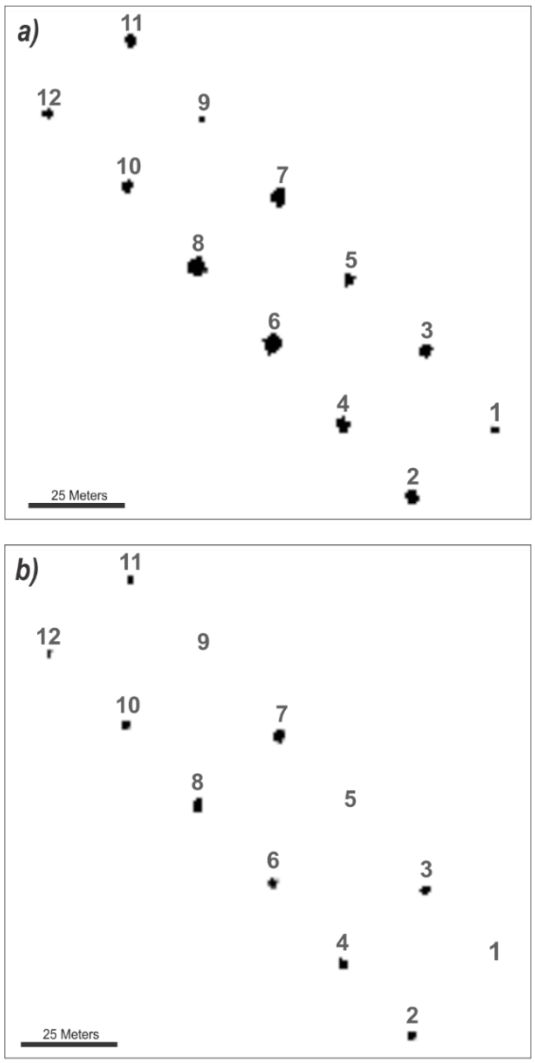

Fig. 8. Examples of the thresholded HC-bearing pixels extracted from simulated WV-3 data shown in Fig 6. **a)** The RBD index detected all the targets (box 1 to 12) and highlighted 231 pixels in total. **b)** The b13/b12 ratio detected nine targets (neglecting box 1, 5, and 9) and highlighted 90 pixels in total.

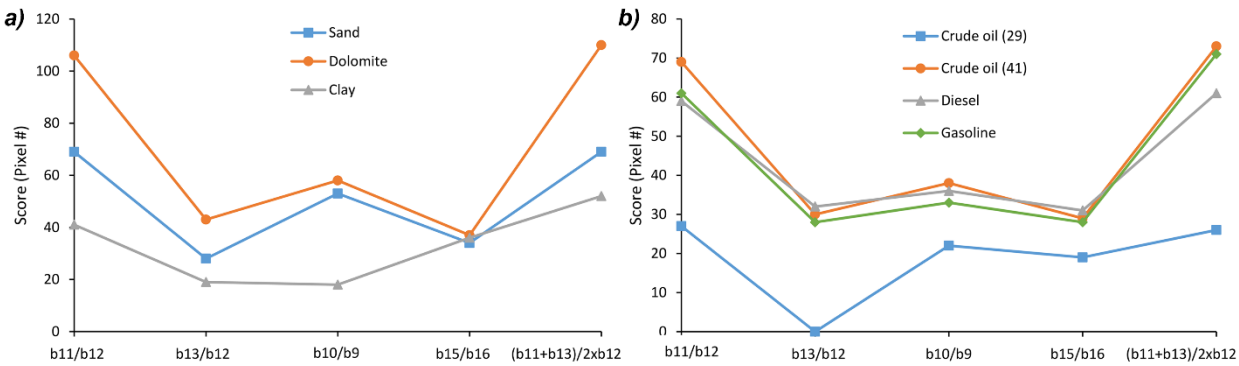

Fig. 9. The performance of different band ratios and RBD index in detecting HCs using simulated WV-3 data. **a)** HC detectability in different background soils. **b)** HC detectability considering the type of hydrocarbon.

7.4.3.2. The effect of endmember, background soil and band entry

By selecting a pixel in the center of each box (Fig. 6), twelve image-derived endmembers were defined. They were then grouped and averaged, based on background mineralogy of the soil and HC types, bringing about seven different endmembers. Additionally, all the twelve imaged-derived spectra and the spectra shown in Fig. 3b were averaged to prepare two more generic endmembers (Fig. 10). The immediate aim was to examine the sensitivity of HC detection against endmember setting. This is important because in real cases there is limited or no knowledge about the spectral characters of the target. The performance of SCM and MF processing techniques relative to each endmember is shown in Fig. 11.

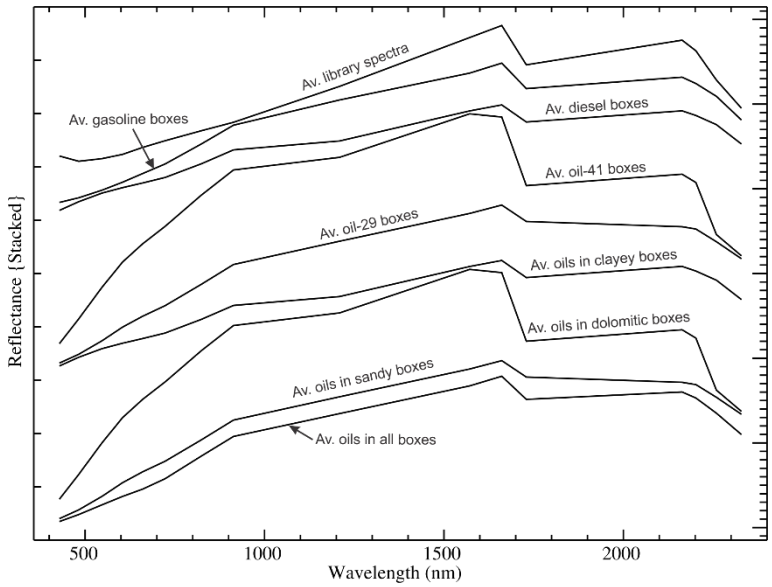

Fig. 10. Endmember suite used with SCM and MF processing techniques. The endmembers were extracted from the center of each box and averaged spectrally based on oil type or background mineralogy as defined in table 4. The library spectra are averaged from the resampled library shown in Fig. 3b. Av= average.

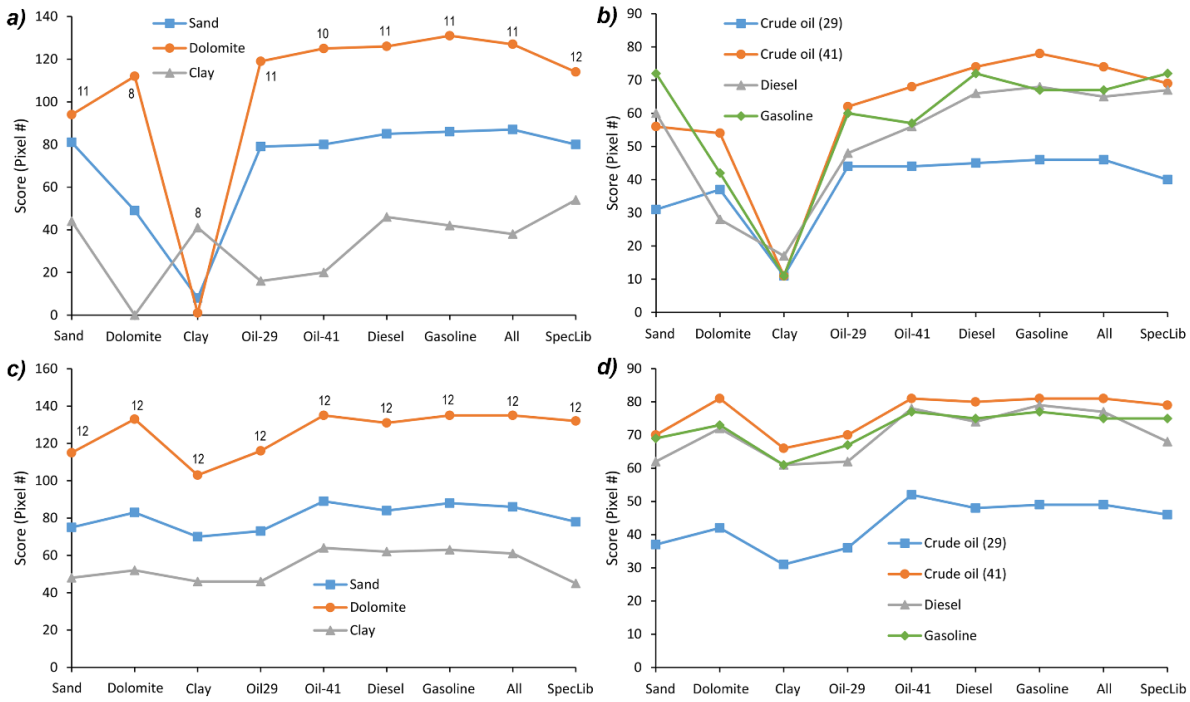

Fig. 11. Performance of the SCM and MF processing techniques in detecting HCs using different endmember sets and simulated WV-3 SWIR bands. **a)** and **b)** The effect of endmembers considering the type of soil substrate and HCs using the SCM method. **c)** and **d)** The effect of endmembers considering the type of soil substrate and HCs using the MF method. The endmembers abbreviated in the x-axis are illustrated in Fig. 10. The digits above the graphs in (a) and (c) define the number of target boxes detected using that specific endmember.

SCM and MF techniques perform poorly when an endmember is selected from soils impregnated by different HCs (e.g. the average of four HC-bearing pixels in the clayey background), as shown by the clay in the x-axis (Fig. 11 a and c). This means that the average of different HCs from similar background does not constitute a good endmember for HC detection, even when the endmember and the target are alike (e.g. clay graph versus clay endmember in Fig. 11a and c). On the other hand, endmembers selected from similar HCs in different background soils yield better results. This is evident for the higher scores achieved with endmembers such as oil-41 or gasoline (Fig. 11a to d). Furthermore, oil types are not correlated with their relevant endmembers. For example, in Fig. 11d, the maximum score of crude oil-29 is associated with oil-41 endmember and the highest score of diesel is correspondent to gasoline endmember**.** Based on these results, it seems unlikely that WV-3 could distinguish between the oil types, especially where they are mixed with other materials, which is consistent with the results achieved using a classic hydrocarbon index from hyperspectral data (Kühn et al. 2004).

The reliability of detecting HCs varies depending on soils type. Dolomitic soils impregnated with the four HCs gave rise to the highest scores, making oil sensing straight-forward in this background (Fig. 11c). Clayey soils mixed with HCs, however, showed the lowest scores (Fig. 11c and a). Sandy soils have an intermediate level of detectability (Fig. 11a and c). The low clayey soil scores are likely due to absorption of the liquid HCs and/or nonlinear spectral mixture with HCs suppressing the band 12 feature. Three of the lightest HCs (gasoline, diesel, and oil-41) are approximately equally responsive to the endmembers and show similar detectability (Fig. 11b and d). Oil-29, as the heaviest oil in the set, clearly singles out as a difficult target, implying the critical role of API gravity in the detectability of oils by remote sensing instruments. In general, where both the HC targets and the endmembers show deep 1700 nm absorption feature (Fig. 10 and 3a), they create easier targets or better endmembers for spectral mapping.

The SCM method calculates the correlation of an endmember with each pixel spectra and yields a score between -1 to + 1 (Carvalho Junior and Menezes 2000**)**. The results achieved with this method (Fig. 11a and b) illustrate that the SCM is very sensitive to the type of endmember or target environment. This instability in the spectral signature, which is likely due to the physical/spectral mixing of liquid petroleum with background soils, makes similarity measures inappropriate for HC detection. In contrast, the MF technique provided higher overall scores and detected all 12 targets in various scenarios (Fig. 11c). Moreover, the MF was more stable against variations in endmember setting, which was a function of oil type, API gravity, and spectral mixing (Fig. 11c). Furthermore, the good performance of generic endmembers (constructed from averaging the image-derived or library-derived spectra; Fig. 10) is promising, because in practice, with little knowledge about the oil type or its occurrence environment, they could be effectively applied to partial unmixing techniques to detect HC-prone areas.

We repeated the same MF processing scheme described above, but employed the sixteen WV-3 (VNIR+SWIR) bands instead of eight SWIR bands. The best results were achieved when only the SWIR bands were used for spectral analysis (Fig. 12), which is likely because most relevant absorption

features of HCs are concentrated in the SWIR part of the spectrum. We have implemented the other processing steps in this work based on the findings described in this section.

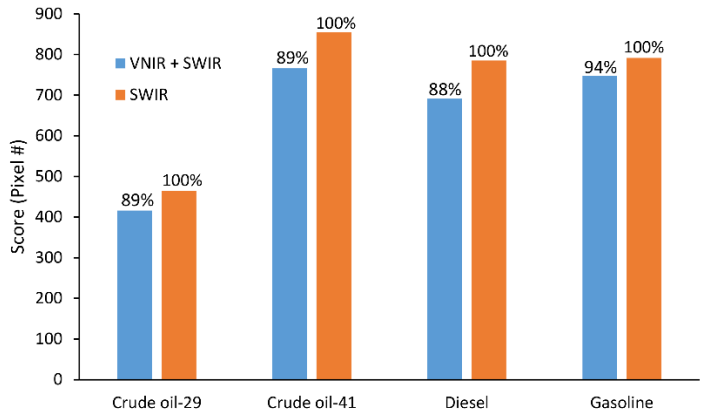

Fig. 12. The assessment of HC detecting using MF technique with different endmembers applied to simulated WV-3 with sixteen spectral bands (VNIR+SWIR) versus the SWIR only bands. The GSD is equal to 0.6 m and the scores are cumulative of all settings. The performance of VNIR+SWIR bands are calculated relative to SWIR bands and are printed on top of each bar.

7.4.3.3. The effect of GSD

To acknowledge the effect of spatial resolution on HC detectability, we resampled the original dataset to higher GSDs and then repeated the processing using both the MF and RBD techniques (Fig. 13). As GSD increases the average scores and the total number of detected oil targets decrease but at different rates for the two techniques. At 3.7 m, the RBD detected seven boxes (Fig. 13a), with the dolomitic soil holding the strongest signal. At the same GSD, oil-41 was the most easily detectable oil, while oil-29 was not traceable (Fig. 13c). At this resolution, the areal ratio of target to pixel is 0.43, meaning that around half of the pixel is covered by oil (Fig. 13c). At 7.5 m resolution, although the oil signal from dolomitic soil still remained, the HC signal from other soils vanished altogether. At this resolution, three out of four targets with dolomitic substrate (including only lighter oils) were detected, wherein each target encompassed merely 11% of a pixel.

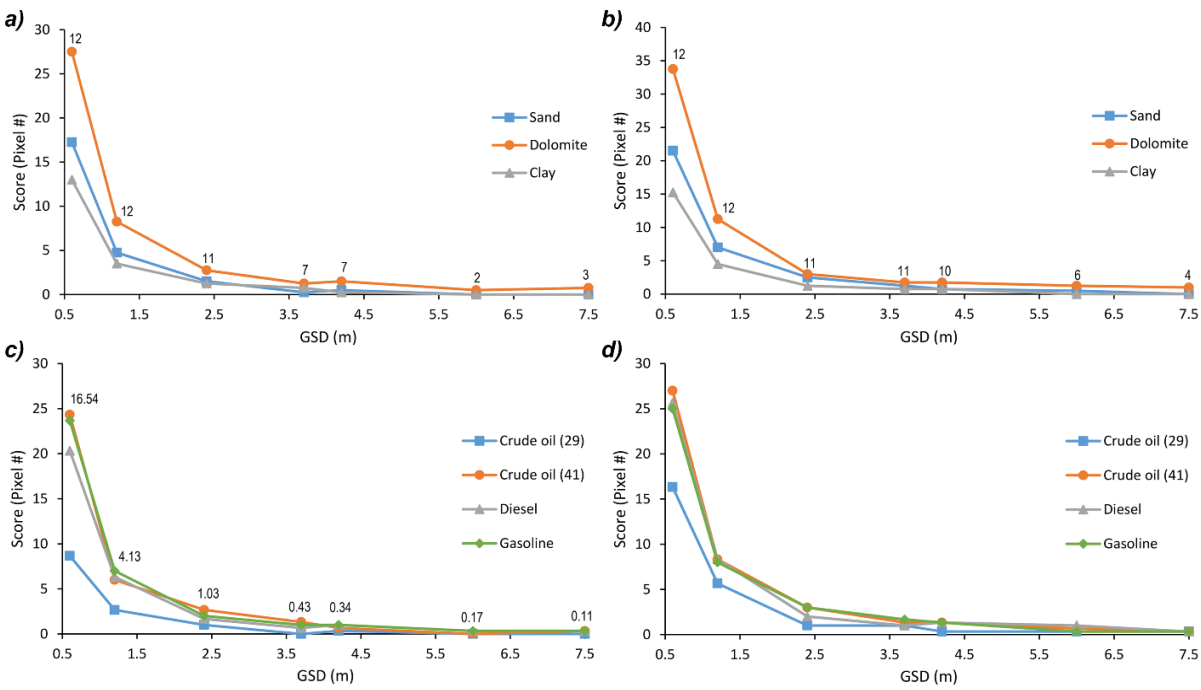

Fig. 13. The effect of GSD on HC detectability using simulated WV-3 data. **a)** The effect of different soil substrates using the RBD (b11+b13/b12) processing method**. b)** The effect of different soil substrates using the MF processing method. **c)** The effect of different HC types using the RBD method. **d)** The effect of different HC types using the MF method. The MF technique was used with SWIR bands and the endmember were yielded from averaging of all-boxes. The numbers above each GSD in (a) and (b) indicate the total number of boxes detected. The numbers in (c) define the target-to-pixel areal ratio.

MF is a partial unmixing technique and can resolve spectral signatures at sub-pixel level more efficiently. Despite this fact, the responses of HCs at different substrates were not similar. At the GSD of 3.7 m, the spectral response of dolomitic soils impregnated with oil was the strongest. Eleven targets out of twelve were successfully detected (Fig. 13b); with the gasoline achieving the highest score (Fig. 13d). When the GSD reached up to 7.5 m, the weaker response shown by clayey and sandy soils at lower resolutions faded away completely but that of dolomite persisted for all the four types of HCs (Fig. 13d).

The detection limit for oil in dolomitic, sandy, and clayey soils was respectively 11%, 17%, and 34% (Fig. 13c and d). This means that the HC spectral response is relatively resilient, and wherever minor quantities of oil is present in a pixel, it is detectable spectrally. Nonetheless, to have at least one pixel selected per oil type on average, a GSD of about 4.0 m is required (Fig. 13c), implying that the HC target needs to cover 37% of a pixel (or take part with the same ratio in the outgoing signal).

Very few studies in the literature have addressed the detectability of petroleum targets via spectral methods. In coastal zones, for instance, AVIRIS data with spatial resolutions of 7.6 m have been used to map oil-damaged marshes extending 1.2 m or less inland, which is equivalent to 16% sub-pixel detection limit. The same dataset resampled to 30 m GSD was still able to detect sub-pixel occurrences of 4 m oiled targets, that is equal to 13% detection limit (Kokaly et al. 2013). In the case of onshore oil seepage, our field experience indicates that a large portion of seeps provide targets with only a few meters wide. Such targets have been successfully detected, for example, using a hyperspectral dataset with 3 m spatial resolution (Prelat et al. 2013). In conclusion, the ultimate detection limit of WV-3's SWIR bands (at 7.5 m commercial release) for targeting HCs will only be established when real datasets are evaluated at varying environmental conditions.

7.4.3.4. The effect of SNR

We analyzed the sensitivity of HC spectral response against different noise levels using the MF method. Fig 14a and b shows the results using the 0.6 m, and Fig 14 c and d using the 3.7 m resolution datasets.

We have used the original dataset (before adding noise) with relevant GSD as a reference and then calculated the minimum SNR required to detect the HCs by considering a 95% confidence level. The results are shown by arrows above each diagram in Fig. 14. The minimum level of SNR depends upon the HC background and oil type. For example, while this level was 100:1 for dolomitic soil, for clayey substrate this minimum increased to 200:1 (Fig. 14a). When oil type was accounted for the level was relatively constant and was less than or equal to 150:1 (Fig. 14b).

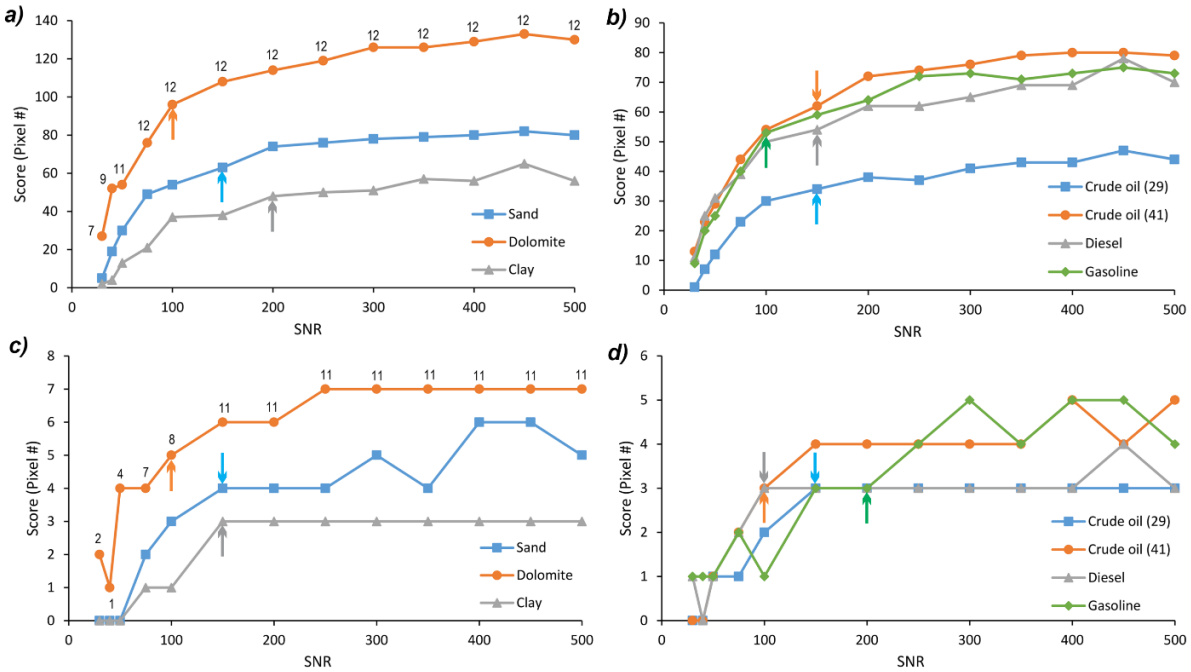

Fig. 14. The effect of SNR on HC detectability by applying the MF technique to simulated WV-3 data. **a)** Variation in HC detectability considering different soil substrates as a function of SNR at 0.6 m GSD. **b)** Variation in HC detectability considering different HCs as a function of SNR at 0.6 m GSD. **c)** Variation in HC detectability considering different soil substrates as a function of SNR at 3.7 m GSD. **d)** Variation in HC detectability considering different HCs as a function of SNR at 3.7 m GSD. The processing was performed using SWIR bands and an endmember yielded from the average of all-boxes. The numbers above each SNR in (a) and (c) indicates the total number of boxes detected. The arrows show the minimum SNR required to detect HC responses at 95% confidence level.

The same calculation applied for data resampled to 3.7 m resolution yielded variable performance inconsistent with those verified at higher resolution. There are two possibilities for such observations: (i) fewer pixels (commonly one pixel in each target) make them highly vulnerable to random noise, (ii) potential instability and error in thresholding.

In general, with appropriate spatial resolution, a SNR of 150:1 should be adequate for HC detection. At this level, the positive detection of each target type is highly likely (Fig. 14a and c).

7.4.3.5. The effect of topography

Our experiment with simulated topography during two different seasons of the year indicates that the detectability of HCs is influenced by shadow (Fig. 15), with the wintertime showing the largest impact upon the signal. Shadowing in the wintertime reduced the total number of pixels by 18% (relative to original WV-3 dataset). For the summer time, this rate declined to merely 10%. Consequently, as long as a proper acquisition time is used from the instrument, the HC targets can be detected without serious disturbance by the shadow.

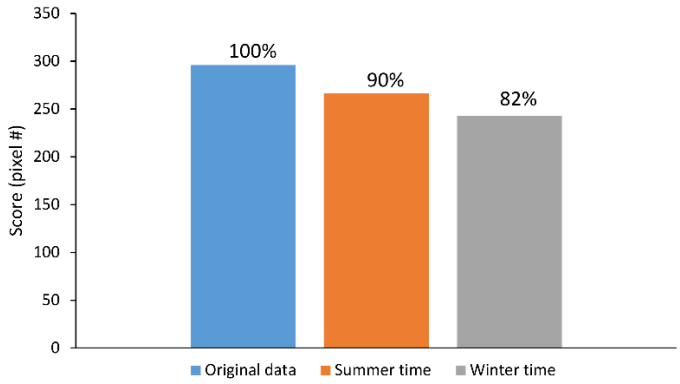

Fig. 15. The effect of rugged terrain and shadowing on the detectability of HCs in a simulated WV-3 dataset. The mid-latitude summer topography correspond to June the 20th, and that of winter season to December, the 20th. All experiments were conducted using SWIR bands at 0.6 m GSD.

7.5. Conclusion

Unlike other multispectral sensors, the WV-3 instrument is equipped with suitable bands to detect HCs. Several of the SWIR bands are located in spectral regions where HCs have diagnostic absorption features, including bands 9, 12, and 16, which are respectively centered at around 1200, 1700, and 2300 nm. Our simulation experiment with close-range and far-range datasets demonstrated the relevance of these bands for HC detection. Band 12 was found to be superior because it is accompanied by two "shoulders" given by bands 11 and 13. Therefore, a simple index such as b11+13/b12 could unambiguously highlight the presence of HC in a pixel.

The HC spectral response however, is not immune to mixing. For example, petroleum in brighter soil substrates with coarser grain size seems to form easier target for the sensor relative to fine-grained darker clays. Besides, our finding indicated that the detectability of oils is dependent on API gravity and lighter oils with higher API are relatively easy to identify.

Since HCs are virtually featureless in the VNIR part of the electromagnetic spectrum, only WV-3's SWIR bands are adequate for HC detection. Our work clearly shows that the use of SWIR bands alone (preferentially acquired at high solar elevation angle) provides superior results in detecting HCs using partial unmixing techniques. Furthermore, despite the lack of details on SNR of the SWIR sub-system, the results indicate that as long as the SNR level of the sensor is higher than 150:1, the detection of petroleum in various environments can be achieved to within a 95% confidence level.

Regarding sub-pixel target detection, our trial with a linear pixel aggregate function revealed that oil fractions of just 0.1 are potentially detectable, however, the oil fraction must typically exceed 0.33 for confident detection. This has significance for the detection of subtle oil leakage from natural reservoir or anthropogenic installations.

Further research with real datasets are required to fully understand the spectral behavior of petroleum as viewed by WV-3 and assess the effect of spectrally similar materials such as non-photosynthetic vegetation or plastics (Clark et al. 2009; Elvidge 1990) on the signal recorded by this sensor. It is also essential to address the accuracy of sub-pixel detection and fraction estimation for petroleum targets using the WV-3's SWIR bands at 7.5 m resolution.

We believe the results achieved here are evident in general. However, offshore scenarios will demand additional research to fully understand the behavior of oil slicks in aquatic environment as seen by WV-3 sensor.

The demonstrated capability of WV-3 has implications in environmental protection and exploration activities in the oil and gas industry. The global coverage of this sensor will facilitate the monitoring of pipelines, petroleum installations (wells, tanks, etc.), and oil terminals for any oil leakage. It will also help identify contaminants and assess the oil-impacted sites or oil spills onshore and offshore. In the energy sector, it will allow prospectors to screen frontier basins for oil seeps and will likely promote the detection and mapping of tar-sands and oil-shales.

Acknowledgement

We are greatly thankful to Rebbeca Scafutto for the provision of the ProSpecTIR dataset, Samuel W. Murphy for help with English, and Alessandro Batazelli for his guidance on collecting the tar-sand samples. The authors acknowledge CNPq for the research grants.

References

Boardman, J.W., Kruse, F.A., & Green, R.O. (1995). Mapping target signatures via partial unmixing of AVIRIS data. In, *Summaries, Fifth JPL Airborne Earth Science Workshop* (pp. 23-26): JPL Publication 95-1

Brekke, C., & Solberg, A.H.S. (2005). Oil spill detection by satellite remote sensing. *Remote Sensing of Environment, 95*, 1-13

Brown, R.A. (2009). Hydrocarbon pollution detection and quantification using hyperspectral remote sensing and geophysics. In, *Institute for Geo-Information Science and Earth Observation (ITC)* (p. 83). Enschede, the Netherlands

Carvalho Junior, O.A., & Menezes, P.R. (2000). Spectral Correlation Mapper (SCM): an improvement on the Spectral Angle Mapper (SAM). In, *Proceedings of the tenth JPL Airborne Earth Science Workshop* (pp. 65-74): JPL Publication 00-18

Clark, R.N., Curchin, J.M., Hoefen, T.M., & Swayze, G.A. (2009). Reflectance spectroscopy of organic compounds: 1. Alkanes. *Journal of Geophysical Research: Planets, 114*, 19

Clark, R.N., Swayze, G.A., Leifer, I., Livo, K.E., Kokaly, R., Hoefen, T., Lundeen, S., Eastwood, M., Green, R.O., Pearson, N., Sarture, C., McCubbin, I., Roberts, D., Bradley, E., Steele, D., Ryan, T., Dominguez, R., & Team, a.t.A.b.V.I.I.S.A. (2010). A method for quantitative mapping of thick oil spills using imaging spectroscopy. In (p. 51): U.S. Geological Survey Open-File Report Number 2010–1167

Cloutis, E.A. (1989). Spectral reflectance properties of hydrocarbons: Remote-sensing implications. *Science, 245*, 165-168

Crowley, J.K., Brickey, D.W., & Rowan, L.C. (1989). Airborne imaging spectrometer data of the Ruby mountains, Montana: mineral discrimination using relative absorption band-depth images. *Remote Sensing of Environment, 29*, 121-134

DigitalGlobe (2014). WorldView-3 datasheet (URL: https://www.digitalglobe.com/sites/default/files/DG_WorldView3_DS_forWeb_0.pdf). In

Donkor, K.K., Kratochvil, B., & Thompson, G.R. (1995). Analysis of Athabasca oil sand by near-infrared-diffuse reflectance spectroscopy. *Analyst, 120*, 2713-2717

Ellis, J.M., Davis, H.H., & Zamudio, J.A. (2001). Exploring for onshore oil seepas with hyperspectral imaging. *Oil and Gas Journal, 99*, 49-58

Elvidge, C.D. (1990). Visible and near infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing, 11*, 1775-1795

Feng, J., Rivard, B., & Sánchez-Azofeifa, A. (2003). The topographic normalization of hyperspectral data: implications for the selection of spectral end members and lithologic mapping. *Remote Sensing of Environment, 85*, 221-231

Fingas, M., & Brown, C.E. (2014). Oil Spill Remote Sensing. *Handbook of Oil Spill Science and Technology* (pp. 311-356): John Wiley & Sons, Inc

Hörig, B., Kühn, F., Oschütz, F., & Lehmann, F. (2001). HyMap hyperspectral remote sensing to detect hydrocarbons. *International Journal of Remote Sensing, 22*, 1413-1422

Jha, M., Levy, J., & Gao, Y. (2008). Advances in Remote Sensing for Oil Spill Disaster Management: State-of-the-Art Sensors Technology for Oil Spill Surveillance. *Sensors, 8*, 236

Kallevik, H., Kvalheim, O.M., & Sjöblom, J. (2000). Quantitative Determination of Asphaltenes and Resins in Solution by Means of Near-Infrared Spectroscopy. Correlations to Emulsion Stability. *Journal of Colloid and Interface Science, 225*, 494-504

Kokaly, R.F., Couvillion, B.R., Holloway, J.M., Roberts, D.A., Ustin, S.L., Peterson, S.H., Khanna, S., & Piazza, S.C. (2013). Spectroscopic remote sensing of the distribution and persistence of oil from the Deepwater Horizon spill in Barataria Bay marshes. *Remote Sensing of Environment, 129*, 210-230

Kruse, F., & Perry, S. (2013). Mineral Mapping Using Simulated Worldview-3 Short-Wave-Infrared Imagery. *Remote Sensing, 5*, 2688-2703

Kühn, F., Oppermann, K., & Hörig, B. (2004). Hydrocarbon Index – an algorithm for hyperspectral detection of hydrocarbons. *International Journal of Remote Sensing, 25*, 2467-2473

Lammoglia, T., & Souza Filho, C.R.d. (2011). Spectroscopic characterization of oils yielded from Brazilian offshore basins: Potential applications of remote sensing. *Remote Sensing of Environment, 115*, 2525-2535

Lammoglia, T., & Souza Filho, C.R.d. (2012). Mapping and characterization of the API gravity of offshore hydrocarbon seepages using multispectral ASTER data. *Remote Sensing of Environment, 123*, 381-389

Leifer, I., Lehr, W.J., Simecek-Beatty, D., Bradley, E., Clark, R., Dennison, P., Hu, Y., Matheson, S., Jones, C.E., Holt, B., Reif, M., Roberts, D.A., Svejkovsky, J., Swayze, G., & Wozencraft, J. (2012). State of the art satellite and airborne marine oil spill remote sensing: Application to the BP Deepwater Horizon oil spill. *Remote Sensing of Environment, 124*, 185-209

Macgregor, D.S. (1993). Relationships between seepage, tectonics and subsurface petroleum reserves. *Marine and Petroleum Geology, 10*, 606-619

Mishra, D.R., Cho, H.J., Ghosh, S., Fox, A., Downs, C., Merani, P.B.T., Kirui, P., Jackson, N., & Mishra, S. (2012). Post-spill state of the marsh: Remote estimation of the ecological impact of the Gulf of Mexico oil spill on Louisiana Salt Marshes. *Remote Sensing of Environment, 118*, 176-185

Prelat, A., Gunaratne, S., Huebner, L., Freeman, C., Cook, A., & Soriano, C. (2013). Airborne hyperspectral detection of natural offshore and onshore hydrocarbon seeps. In F. Aminzadeh, T.B. Berge, & D.L. Connolly (Eds.), *Hydrocarbon Seepage: From Source to Surface* (pp. 171-182). Tulsa, OK, U.S.A.: American Association of Petroleum Geologists

Rodger, A., Laukamp, C., Haest, M., & Cudahy, T. (2012). A simple quadratic method of absorption feature wavelength estimation in continuum removed spectra. *Remote Sensing of Environment, 118*, 273-283

Shaw, R.C., & Kratochvil, B. (1990). Near-infrared diffuse reflectance analysis of Athabasca oil sand. *Analytical Chemistry, 62*, 167-174

Van Der Meer, F., Van Dijk, P., Van Der Werff, H., & Yang, H. (2002). Remote sensing and petroleum seepage: a review and case study. *Terra Nova, 14*, 1-17

van der Werff, H. (2006). Knowledge based remote sensing of complex objects recognition of spectral and spatial patterns resulting from natural hydrocarbon. In, *International Institute for Geo-Information Science and Earth Observation (ITC)* (p. 138). Enschede, The Netherlands

Chapter 8

Multi-scale mapping of oil-sands in Anhembi (Brazil) using imaging spectroscopy[13]

In this work oil-sand outcrops in the Anhembi area in the Paraná Basin, Brazil, were studied using multi-scale imaging spectroscopy. The study incorporated VNIR-SWIR (400-2500nm) multispectral data from the WorldView-3 satellite instrument and hyperspectral data collected on the ground and from an airplane using an AisaFENIX imaging system. This was followed by outcrop studies and sampling in the field and close-range spectroscopy of the samples in the laboratory using both imaging (i.e. sisuCHEMA) and nonimaging (i.e. FieldSpec-4) instruments. The bitumen content of the deposit, which proved to be as high as ∼12 wt. %, was revealed by close-range imaging to be unevenly distributed, even in the scale of a hand sample. The mean bitumen content was estimated to vary from 6 % in hand samples to ∼5 % in bitumen-rich vertical walls, to 4 % in the aerial oil-sand outcrops. The continual decrease in the mean bitumen contents was inferred to be linked to the averaging effect of growing pixel sizes of the imagery. The bitumen impregnation was shown to be controlled mainly by the permo-porosity of the facies, with sand sheet facies corresponding to highest bitumen contents. The induced alteration mineralogy was restricted to clay formations and iron oxides bleaching. The clays, identified as montmorillonite, were ubiquitous and mixed intimately with bitumen at all studied scales. These collective observations were interpreted to be indicative of a single-phase flow of very viscous (API° ∼5) HCs to the reservoir. The study showed that bitumen determination on the ground requires a system with high signal-to-noise ratio and good illumination conditions. Spectral denoising was also shown to be crucial and a prerequisite for spectroscopic analysis. WV-3 data showed capable of resolving HC's feature at 1700 nm wavelength by its SWIR band-4 over targets encompassing ∼25–30% of the resampled SWIR pixel (7.5 m); albeit the data was not successful in determining the total bitumen content of the outcrops. This study demonstrated that a multi-scale spectroscopic approach could provide a complete picture of the variations in the content and composition of bitumen and associated alteration minerals. The HC signature, especially the one centered at 2300 nm, is consistent and comparable among scales, and upon employing proper calibration data and analysis approach is capable of estimating the bitumen content of oil-sands at all imaging scales.

8.1. Introduction

Oil-sands are a complex mixture of bitumen (extra heavy hydrocarbons, with API gravity <10°), clays, water, and consolidated sands consisting of quartz, feldspars, and other accessory minerals (Cloutis et al. 1995). Oil-sand represents a type of unconventional oil deposit that is exploitable by surface mining (Speta et al. 2015). Natural bituminous deposits are reported in many countries, but giant reserves are found in the province of Alberta, Canada. In Brazil, the most important oil-sand accumulations occur in the eastern margin of the Paraná Basin to the southeast of the country. So far, 19 occurrences of oil-sands are reported to exist on this basin, of which the most significant and likely mineable deposit occurs in the Anhembi area (Araújo et al. 2005; Thomaz Filho et al. 2008).

Hydrocarbons (HC) display several diagnostic absorption features between the visible-near infrared (VNIR) to short-wave infrared (SWIR) wavelengths due to overtones and combinations of the stretching fundamentals. Pronounced features in the SWIR window include a triplet between

[13] This chapter is submitted to the Remote Sensing of Environment Journal.

1700–1750 nm and a doublet between 2290–2360 nm. HCs also show a broad but prominent absorption feature in the VNIR due to the pi (π) bonding between carbon (C=C) atoms (Asadzadeh and Souza Filho 2017; Cloutis 1989; Lammoglia and Souza Filho 2011). Owing to the presence of these diagnostic absorption features, reflectance spectroscopy has been extensively used as a rapid and non-invasive tool for characterization of oil-bearing compounds. In soil sciences, this technique is employed to track and quantify HC contamination of the soil (Correa Pabón and Souza Filho 2016; Schwartz et al. 2013). In the oil-sand industry, in particular, it has been widely used for on-line estimation of the total bitumen content (TBC) during ore processing (Rivard et al. 2010). Close-range imaging spectroscopy, on the other hand, was shown to be a valuable tool in the core logging and the determination of TBC of the core samples (Speta et al. 2015). From the air, hyperspectral imaging is deployed to detect and quantify oil contaminations or naturally occurring seeps in the environment (Ellis et al. 2001; Kokaly et al. 2013). Thus far, however, the capability of orbital remote sensing data for direct petroleum detection has remained unexplored, mainly because of the lack of proper spectral bands in the operational instruments to resolve HC's absorption features. Some initial spectral simulations demonstrated the potential of WorldView-3 (WV-3) satellite data for this aim (Asadzadeh and Souza Filho 2016a), albeit this possibility has not yet been tested using real datasets.

In geologic remote sensing, the gap between proximal measurements in the lab and distal imaging spectroscopy collected overhead is closed by means of ground-based hyperspectral instruments. This emerging technology is demonstrated to be invaluable to study the petrographic and mineralogic aspects of the near-vertical rock faces at submillimeter to centimeter resolution. For instance, it has proved useful to characterize sedimentary stratigraphy (Greenberger et al. 2015; Okyay et al. 2016), carbonate lithofacies in quarries (Kurz et al. 2012), ferric iron minerals in vertical mine faces (Murphy and Monteiro 2013), and outcrops of organic-rich oil shales (Greenberger et al. 2016). Nonetheless, the usefulness of these high spatial/spectral resolution data in the oil-sand mining industry is unaddressed in the literature.

This work aims to study oil-sand outcrops of the Anhembi area in Paraná basin, Brazil, using multi-scale imaging spectroscopy. The motivation to explore this site was the well-exposed outcrops of the oil-sands, the easy accessibility of the site for field studies, and the availability of WV-3 imagery over the entire area provided by DigitalGlobe. The area was further surveyed by ground-based and airborne imaging systems using an AisaFENIX hyperspectral instrument (Specim Ltd., Finland). It was then followed by outcrop studies and sampling in the field and close-range spectroscopy in the laboratory using both the imaging (i.e. sisuCHEMA) and nonimaging (ASD FieldSpec) instruments. The research in this area was twofold: (i) to detect and characterize the bituminous sandstones using spectral technology and evaluate the consistency of HC's signatures at multiple scales ranging from microscopic to deposit scales for bitumen content estimation, and (ii) to delineate the diagenetic alteration mineralogy associated with bitumen emplacement in sandstone beds and draw an analogy between this phenomenon and microseepage-induced alterations.

8.2. Geological context

The Anhembi study area is located in the eastern edge of the Paraná basin, near the town of Anhembi, some 170 km to the NW of Sao Paulo city in southeastern Brazil (Fig. 1). The approximate coordinates of the site are between 48°01'10" to 48°01'30" west, and 22°45'35" to 22°45'45" south, covering an area of <2 km^2. The climate of the region is mild, with an average annual temperature of 20°C and 1200 mm precipitation.

The sandstone beds of the Pirambóia Formation cover almost the entire study area. The Pirambóia Formation comprises an unconformity-bounded aeolian succession of fluvial and aeolian sandstones of the lower Triassic. The sandstones, which vary in thickness from 20–200 m, are medium to fine-grained and show cross-stratification in cross section (Fig. 2b). The depositional environment of this formation is indicated to be fluvial with channels of floodplain facies. At its base, it comprises thick aeolian sand sheet facies, which are overlain by aeolian dune and interdune strata (Dias and Scherer 2008). Overall, four eolian facies are defined in the Pirambóia Formation comprising (i) foreset beds, (ii) bottomset beds, (iii) wet and (iv) dry interdune facies. The first two facies are characterized by high to intermediate permo-porosity, whereas the interdune facies are characterized by low permo-porosity and abundant argillaceous minerals (Martins et al. (2015) and references therein).

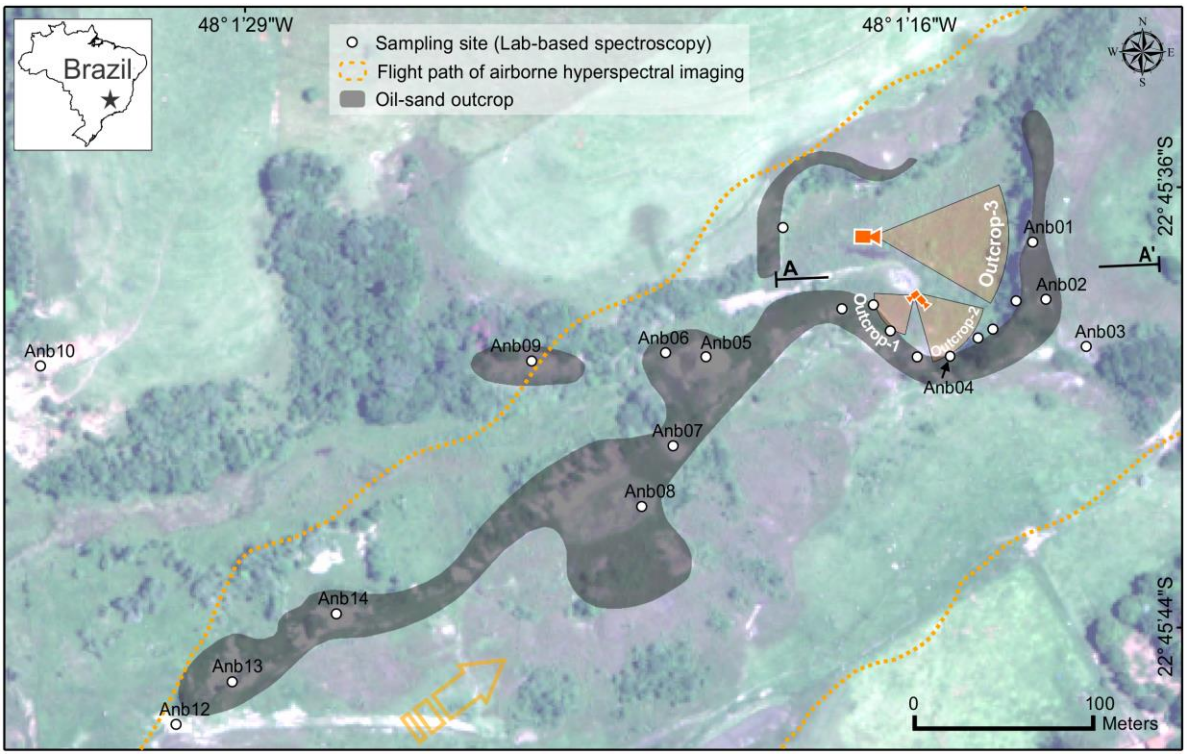

Fig. 1. The worldview-3 natural color composite imagery of the Anhembi area located in southeastern Brazil (inset map). The circle sectors (arcs) and the orange dashed polygon respectively represent the wall faces scanned by ground-based hyperspectral imaging and the flight path of the airborne system. The gray-shaded polygon outlines the oil-sand outcrops. The white circles indicate the sampling sites with the labeled ones corresponding to anomalies derived from either WV-3 or airborne data and verified on the field. Note that Pirambóia Formation covers the entire area. A schematic geologic section along A–A' is shown in Fig. 3.

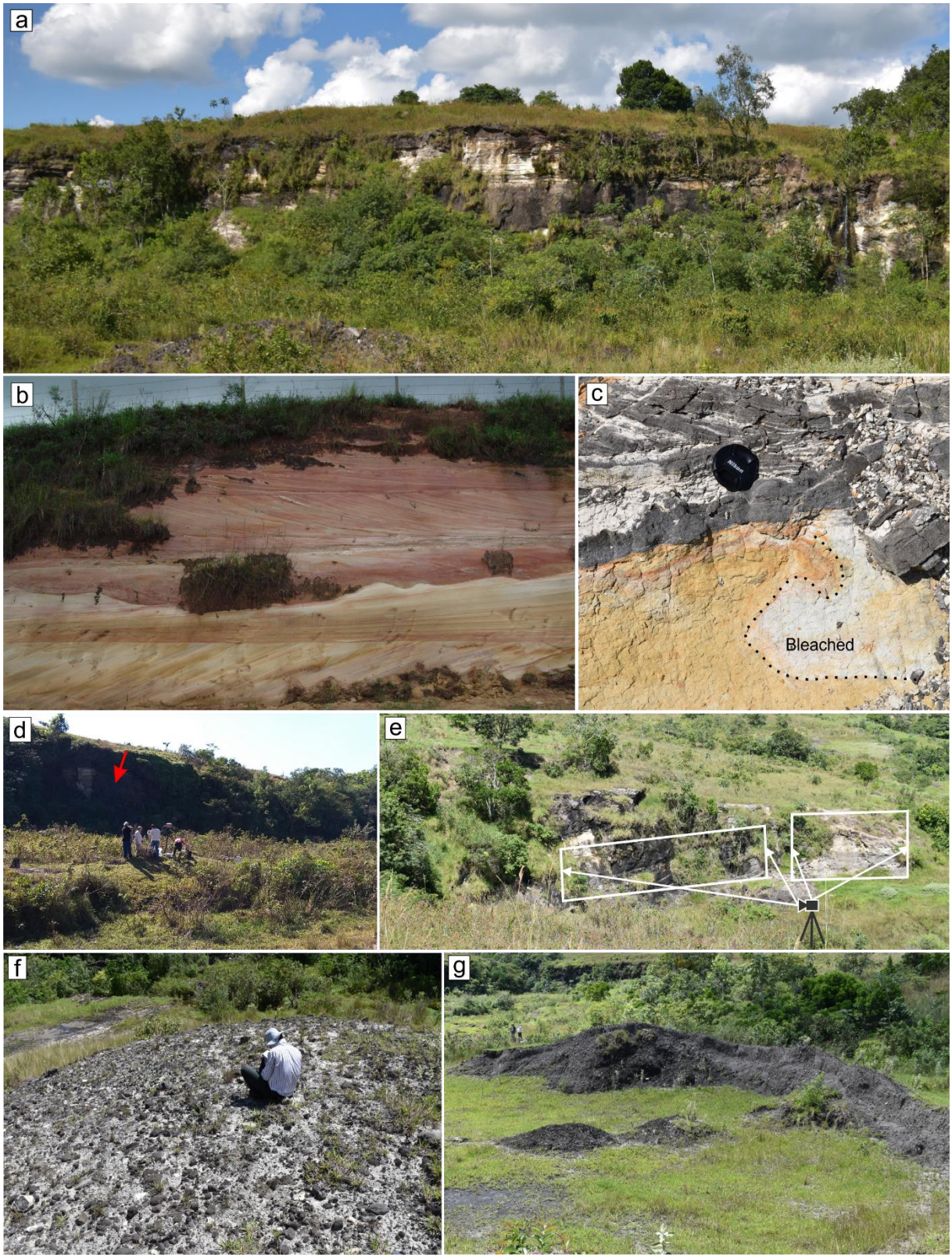

Fig 2. Photographs illustrating the Pirambóia sandstones and oil-sand outcrops in the Anhembi area. **a)** examples of oil-sand walls cropping out in the study area. **b)** cross-stratification in the eolian oil-free sandstones of the Pirambóia Formation in nearby areas. **c)** bleaching of the iron oxide coats that is confined to bituminous zones. **d)** a tripod-mounted AisaFENIX hyperspectral camera in operation scanning Outcrop-3. The arrow exemplifies a wet oil-sand wall in the shade. **e)** the outlines of Outcrop-1 (right) and Outcrop-2 (left) scanned by the imaging system. **f-g)** outcrops of oil-sands detected by airborne hyperspectral (AisaFENIX) and spaceborne multispectral (WorldView-3) data corresponding to Anb05 and Anb08 sampling sites, respectively. The locality of vertical walls (d-e) and the sampling sites (f-g) are illistrated in Fig. 1.

The bitumen accumulations in the Paraná basin occur mainly along fault zones or around structural highs on erosion-preserved areas (Fig. 2a) (Araújo et al. 2006). Petroleum in this area was derived from the Permian Irati shales (specifically the organic-rich Assistência Member), thermally matured due to Serra Geral magmatism, and then migrated to the lower Pirambóia Formation. Presumably, oil migrated along the contacts of diabase dykes with sedimentary units and then impregnated the eolian sandstones (Fig. 3). The interdune argillaceous beds and fluvial facies along with the intruded dykes then acted as local barriers against, the vertical and horizontal HC migrations, respectively (Araújo et al. 2005; Thomaz Filho et al. 2008). The restriction of accumulations in the basinward (west) side of the dykes (see Fig. 3) is hypothesized to be due to hydrodynamic flow from the inner parts of the basin to its periphery (Araújo et al. 2006; Araújo et al. 2005). Due to the role of diabase dykes in the maturation and migration of oil, some scholars regarded this Permian petroleum system as an atypical system (Araujo et al. 2000).

As noted earlier, the largest known accumulation in this basin is located in the Anhembi area. This deposit, which extends along the edges of a NE–SW trending valley (Figs. 1 and 2a), is estimated to contain about 5.7 million barrels of oil in-place based on average bitumen content of 5.5 wt. %. The oil in this deposit is immature, very heavy, and highly viscose (API ~5°), with 2–3 wt. % sulfur content (Thomaz Filho et al. 2008). Studies suggest that petroleum at this accumulation has been heavily degraded microbially, with the dry interdune and bottomset facies showing the highest and the lowest degree of biodegradation, respectively (Martins et al. 2015). During the 1980s, there have been some mining activities at this site (see for example Fig. 2g). The excavated oil-sands have been used mainly for road pavements of the nearby areas.

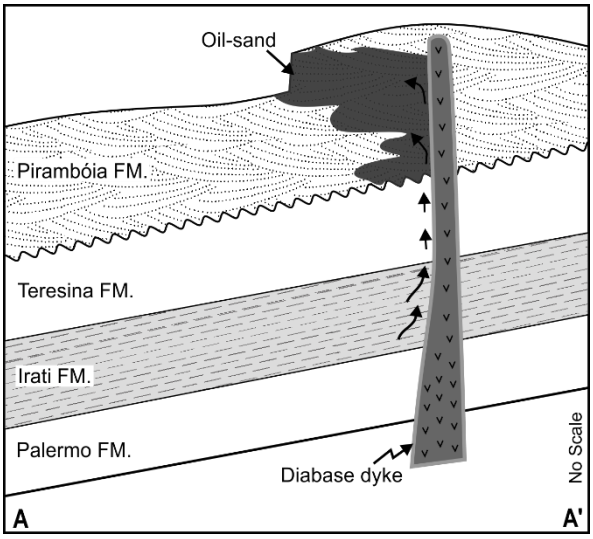

Fig. 3. Schematic geologic section representing the mechanism of oil migration along the diabase dykes to eolian Pirambóia sandstones. The section corresponds to A–A' in Fig. 1 (Adapted from Thomaz Filho et al. (2008)).

8.3. Data collection and processing

8.3.1. Fieldwork and sampling

Fieldwork in this area was carried out at three different levels. Initially, we tried to visit the reported outcrops of the oil-sands in the vicinity of Anhembi, particularly those to the east of Botucatu city. While several occurrences of oil-sands were reported to exist around this deposit (Araújo et al. 2006; Araújo et al. 2005), oil-sands were observed only in one locality[14] called Bofete to the south of Bofete town (23°11'36.5"S, 48°13'59.5"W). Notwithstanding, due to voluminous data collected over the Anhembi test site and the sheer size of accumulations, we restricted the scope of our studies to this deposit.

At deposit-scale, we collected several representative samples (n=15) from oil-impregnated outcrops (mostly vertical walls) and the associated alterations (Fig. 1). To obtain the best results in the lab, fresh cuts of the oil-sands were collected during sampling. In the field, the relationships between the sandstone facies and their bitumen content, as well as the induced alterations, were investigated. Moreover, the anomalies derived from WV-3 satellite data were visited and verified in the field and sampled (n=10) for follow-up spectral study in the lab (Fig. 1). The same sample suite was also used to verify the result of airborne hyperspectral imagery.

We also studied some well-exposed outcrops of the Pirambóia Formation devoid of bitumen from neighboring areas (particularly in road cuts) aiming to achieve a better understanding of the secondary mineralogic changes induced by bitumen emplacement and collect representative samples (n=11) for comparative spectroscopic studies. The coordinates of the sampling sites were recorded using a handheld GPS receiver to within an accuracy of ±3 m.

To facilitate the referral, henceforth, the oil-impregnated sandstones of the Pirambóia Formation are referred to as ‘oil-sands’, whereas the beds devoid of bitumen are called ‘Pirambóia sandstones’.

8.3.2. Infrared spectroscopy

In the lab, both sets of samples (oil-sands and Pirambóia sandstones) were air-dried (if recognized to be wet) and then measured using a FieldSpec-4 spectrometer between 350–2500 nm wavelength range at 1 nm sampling intervals. A contact probe with artificial illumination was used to perform the measurements and a Spectralon reference panel was used to convert the results to reflectance. To reduce the effects of instrumental noise, fifty consecutive scans were incorporated into each recorded spectrum. Subsequently, the results were integrated into two separate spectral libraries namely oil-sands and Pirambóia sandstone suite. The procedure explained in Chapter 5 was used to correct the residual splice drift of the spectrometer between VNIR and SWIR2 channels. For comparison with ground-based data, the datasets were converted to absolute reflectance by multiplying the results to the response of the Spectralon white reference. Besides measuring the reflectance spectra of the sample suite, infrared spectroscopy also provided benchmark for checking the accuracy of other imaging instruments in retrieving quality reflectance spectra.

[14] Two known occurrences at Guareí (23°22'11.97"S, 48°15'32.58"W) and Mina farm (22°41'7.80"S, 48°18'3.61"W) were not visited during this study.

We used the Absorption-based Mineral Spectral Analyst (AMISA) to analyze the spectral datasets (Asadzadeh and Souza Filho 2016b). Three absorption features including ferric iron oxides (Feox), H_2O, and Al–OH, respectively centered at ~900, ~1900, and ~2200 nm (Fig. 4), were considered to compare the oil-sands with Pirambóia sandstones. More specifically, we used the wavelength position, depth, and asymmetry of the features for cross-comparison. Beside these absorptions, which observed to be common in both datasets, some infrequent features occurring at ~2340 and ~1065 nm were also inspected. For processing of all absorption features, a fixed threshold of $2x10^{-4}$ was applied within the AMISA.

In order to verify the mineralogy of the samples as revealed by spectroscopy, and understand the composition of framework grains of the sandstones, we analyzed three oil-sand samples with X-ray diffraction (XRD).

8.3.3. Close-range hyperspectral imaging

Alongside the point measurement, some of the oil-rich samples were also selected for imaging using a sisuCHEMA™ hyperspectral imaging instrument (Specim Ltd., Finland), aiming to examine the distribution and spatial relationships of clays and bitumen in oil-sands. The sisuCHEMA system consists of a moving tray to put the sample into, an illumination source, and a high-resolution hyperspectral camera measuring radiance within the SWIR wavelength range (Table 1). The system can automatically measure a large sample by repeating the scanning along parallel scan lines. The acquired image cube is then converted to reflectance using an internal Spectralon reference by instrument's built-in routnines.

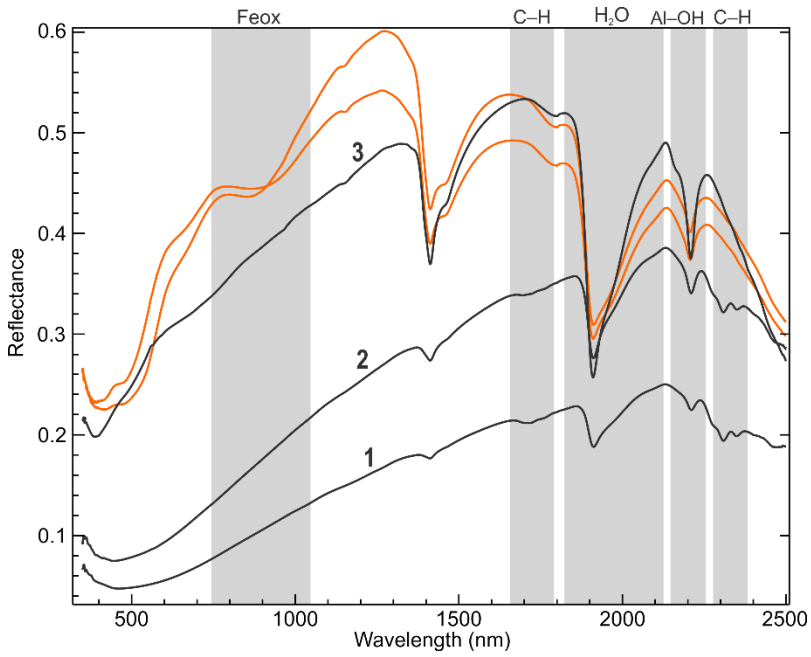

Fig. 4. Representative spectra of the Pirambóia sandstones (orange) and oil-sands (gray). Spectra No.1–3 exemplify a bitumen-rich sample, bitumen in (areal) mixture with montmorillonite, and a bitumen-free sample dominated by montmorillonite, respectively. The gray bars exhibit the analyzed wavelength ranges comprising iron oxides (Feox), hydrocarbons (C–H), water (H_2O), and clays (Al–OH).

For quality scanning, the oil-sand sample was firstly sliced appropriately to form even surfaces. Next, the sample (shown in Fig. 15a) was scanned by using a 31 mm lens to yield a hypercube data with 256 spectral bands between 930–2500 nm and spatial resolutions of 390 and 680 μm,

respectively along the length (X) and width (Y) of the sample. To obtain equal-sized pixels, however, the image was resampled to 680 µm using 0.57 multiplier factor in the Y direction and a pixel aggregate function. Sixteen bands at either side of the cube (consisting of thirteen bands between 928–1004 nm and three between 2512–2524 nm) were determined to be very noisy and thus excluded. The tray in the background and at least two pixels in its vicinity was also nullified. To relieve the illumination variations among scan lines, the imagery was transformed into the frequency domain using the Fast Fourier transform (FFT). There, the frequencies related to striping and illumination variations were discarded, and the remaining image was transformed back into the spatial domain (see also section 3-4). Ultimately, to suppress the residual noise, a spectral smoothing filter (Sav-Gol filter with a second-order polynomial convolution) was applied to the image on a pixel-by-pixel basis. The spectral content of the resulting imagery was validated against some representative spectra derived from the same surface using a contact probe connected to a spectrometer. The processing of the denoised images is described together with the ground-based data in the next section.

8.3.4. Ground-based hyperspectral imaging

On June 17, 2016, the vertical walls of the oil-sand deposit were scanned on the ground using an AisaFENIX hyperspectral system (Specim Ltd., Finland). This instrument collects data over the entire 380–2500 nm range in a single continuous image with up to 620 spectral bands at 3.5 (VNIR) to 12 (SWIR) spectral resolution (Table 1). The signal-to-noise ratio (SNR) of the instrument varies between 600–1000:1 and it works as a pushbroom line scanner with 384 pixels per line.

Table 1. Specification of the hyperspectral imaging systems used in this study.

Imaging system	Spectral range (nm)	Spectral bands	Spectral resolution (nm)	Spatial resolution
AisaFENIX	380–2500	360–620*	3.5 (VNIR) 12 (SWIR)	384 pixels**
sisuCHEMA	930–2500	256	10	30–300 µm

* it depends on the spectral binning option.
** it depends on the altitude of the sensor.

The sensor was used in the ‘field mapping’ mode, which includes a rotating stage attached to a tripod, and configured to collect data by along-track rotation of the camera at a determined velocity (Fig. 2d). The image data were collected in three binning options namely 2x2, 4x2, and 8x2. Whereas the 2x2 data series (with 624 spectral bands) were discovered to be very noisy, the 8x2 series were found to retain low spectral resolution to resolve the features of interest (with 363 spectral bands). Thereby, we opted for the 4x2 data series with intermediate properties (corresponding to 450 spectral bands and high SNR). In total, three different faces named Outcrop1–3 were imaged by the instrument from two scan positions (Figs. 1 and 2e, and Table 2). The distance of the sensor from the walls was varied from around 10 to 80 m (Table 2). This configuration yielded to a ground sampling distance (GSD) of 1.5 to 8 cm. Both Outcrop-1 and 2 were illuminated directly by sunlight (Fig. 2e), but Outcrop-3 sat in dark shades (see arrow in Fig. 2d as an example). The shadow coupled with dripping wet surfaces, gave rise to very low SNR over this outcrop, and thus the data was eliminated from further processing.

Table 2. Characteristics of the imaged outcrops and the solar position during the time of ground-based imaging spectroscopy (17th June 2016). See Fig. 1 for the location of the outcrops.

Outcrop	Outcrop azimuth	Local time	Solar elevation	Solar azimuth	Distance (m)	GSD (cm)	Illumination condition	Physical condition
1	~135°	10:30	37.2°	30.4°	~10	~1.5	Direct sunlight	dry
2	~75°	11:12	41.3°	19.1°	~30	~3.0	Direct sunlight	dry
3*	~5°	12:58	42.5°	346.2°	~80	~8.0	Indirect sunlight	wet

* This outcrop was found to be very wet and deep in shadow, hence eliminated from further processing.

For atmospheric compensation, we used a single calibration panel in the field of view (FOV) of the imaging system. The panel consisted of a 30 x 50 cm wooden board plastered by a thick layer of highly reflective barium sulfate ($BaSO_4$).

Dark current was recorded in the field by blocking the light entrance of the system while averaging the data over about 10 s. This data was then subtracted from the images on a line-by-line basis for dark current removal (Fig. 6). The spectral data were calibrated to at-sensor radiance by using manufacturer's calibration factors. The spectral smile of the system is reported by the manufacturer to be below 0.2 pixel, thereby was omitted from correction.

Within the preprocessing stage, the atmospheric effect was first corrected through dividing each pixel of the image to the average radiance of the reference panel, which was achieved by incorporating the pixels occurring on the center of the panel (Fig. 6). Next, the resulting image was multiplied by the absolute reflectance of the panel acquired by the following steps (Fig. 5): (i) the reflectance of the barium sulfate plaster was recorded under artificial light using a pistol connected to a FieldSpec-3 spectrometer; (ii) an average of ten measurements collected from different parts of the panel was calculated and then multiplied to the reflectance of the Spectralon reflectance provided by its manufacturer (Fig. 5); (iii) the resulting absolute reflectance was resampled to the response of the AisaFENIX dataset.

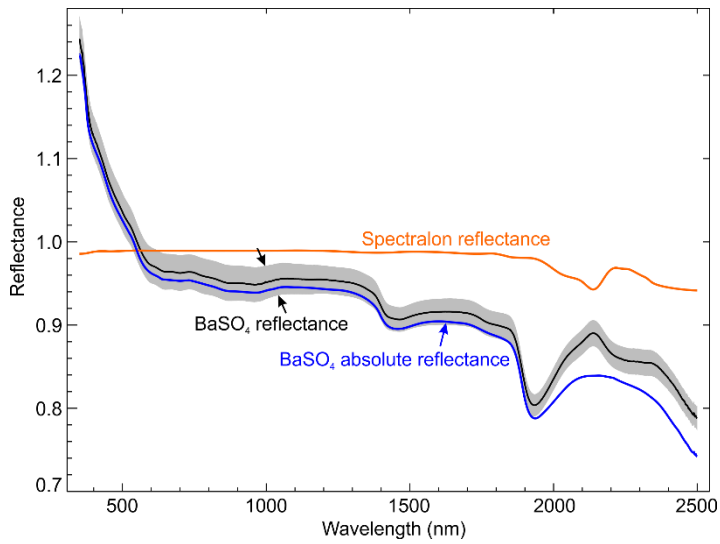

Fig. 5. The conversion of relative reflectance of barium sulphate ($BaSO_4$) panel (gray) into absolute reflectance (blue) using the Spectralon reflectance (red). The black line and the gray error bar illustrate, respectively, the mean ± one standard deviation of the reference panel.

The following bands were considered as bad-bands and discarded from the data: (i) water vapor bands between 1338–1455 and 1812–1989 nm (22 and 33 channels, respectively) and (ii) noisy bands between 380–411 and 2445–2500 nm (11 and 12 channels, respectively). The remaining 372 spectral bands (n) were retained for further processing (Fig. 6). The stripping noise, which arises from radiometric calibration differences between charge coupled device (CCD) detector arrays (Kurz et al. 2013), was removed in the frequency domain following the Fast Fourier Transform (Fig. 6). This process was indicated to be more effective when it is applied to derived products with enhanced noise components (Watson 1993). Three transformations were evaluated for noise enhancement: independent component analysis (ICA), principal component analysis (PCA), and minimum noise fraction (MNF) (see Asadzadeh and Souza Filho (2016c) and references therein). Whereas the MNF transform was found to scatter the stripping noise throughout the components, thereby making the filtering complicated, the ICA revealed to be very sensitive to slight changes in the contents of the components (associated with noise removal) during the inverse transformation. PCA, on the other hand, showed superior performances, hence opted for this aim.

Following forward PCA transform, a subset of *m* components ($PC_{1\text{-}m}$, where $m \ll n$) determined to be spatially coherent (based on visual inspection and eigenvalues) was segregated from the remaining components dominated by stationary noise[15]. The segregated array was treated in three different ways: (i) the higher components free from systematic/random noise were preserved intact, (ii) the components recognized to retain systematic horizontal or vertical noise (e.g. stripping) were transformed into the frequency domain for filtering, and (iii) the remaining coherent images associated with random noise were treated by statistical filters. In the frequency domain, the frequencies related to striping were discarded and then an inverse FFT was used to transform the image back to the spatial domain. Because PCs have distinct noise properties, they were treated independently in the frequency domain. An inverse PC was then used to transform the incorporated *m* denoised bands into image domain. Finally, the image was spectrally smoothed by a Savitzky–Golay (savgol) filter, configured with a second-order polynomial convolution (Fig. 6).

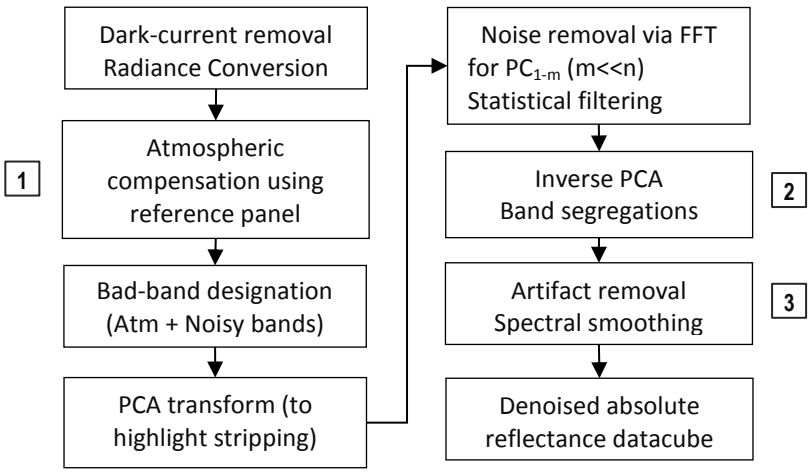

Fig. 6. Workflow diagram showing the preprocessing steps applied to ground-based hyperspectral data. The acronyms are Atm: atmospheric, PCA: Principal component analysis, and FFT: Fast Fourier Transform.

[15] In the case of ground-based data, 'm' was indicated to vary between 20–25 components.

The resulting images were spectrally processed to yield the TBC (total bitumen content) image along with the Al–OH wavelength position and depth images, respectively indicative of Al-smectites composition and abundances. For this aim, a second-order polynomial was fitted to the absolute minimum of the Al–OH feature and its immediate neighboring bands between 2151–2244 nm wavelengths following continuum removal (Rodger et al. 2012). The coefficients of the polynomial were then used to calculate the minimum of the feature and its corresponding depth relative to a local continuum. The fitting was performed to the imagery before spectral smoothing to avoid altering the position of the absorption feature. The TBC of the samples and walls were estimated by employing the two-bands normalization method provided by Rivard et al. (2010), which proved to estimate the TBC with a dispersion of ∼1.5% in controlled illumination conditions. Albeit, instead of $R_{2220nm,}$ we used R_{2230nm} as the shoulder of the Al–OH absorption feature. The distribution of oil-sands was first mapped by employing the spectral angle mapper (SAM) (Kruse et al. 1993) technique and manual endmember selection through averaging dozens of image spectra from corresponding imageries. The input image was confined to the spectral bands between 1630–1784 and 2231–2396 nm wavelengths. The result was then thresholded to yield a mask to outline only the oil-sands and then applied to the TBC image. The average of R_{2220nm} and R_{2230nm} were thresholded (<0.1) and used to mask very dark shadowy pixels.

8.3.5. Airborne hyperspectral imaging

The outcrops of the oil-sands in Anhembi area were also imaged by the AisaFENIX hyperspectral scanner on board a twin-engine Seneca II (Piper) airplane at an altitude of 915 m on May 2, 2016 (See Fig. 1 for the areal coverage). The hyperspectral data were acquired in a clear sky condition with a nominal GSD of ∼0.35 m in the 8x2 binning mode within which 363 co-registered spectral bands are contained between 380–2500 nm wavelengths (see also section 3-4). A single strip best covering the area (Fig. 1) acquired at 14:14 PM local time was geometrically corrected (orthorectified) using sensor's built-in navigation (GPS and IMU) system to within an accuracy of ±10 cm. The radiance data were then converted to apparent surface reflectance using the ATCOR atmospheric correction software (ReSe Applications LLC, CH). During this process, the data were resampled to 1.0 m spatial resolution using pixel aggregate function.

The denoising approach described in the previous section was deployed to suppress noise components in the imagery. A combination of C–H and Al–OH absorption features were used to isolate the outcrops of the oil-sands within the area. Then, the continuum-removed depth of the C–H and Al–OH features, respectively centered at 2300 and 2209 nm were calculated aimed at mapping the bitumen and clay contents of the oil-sands.

8.3.6. Spaceborne multispectral imaging

Following the demonstration of WV-3's capability to detect HCs using simulated data, under our request, DigitalGlobe (Westminster, Colorado, USA) acquired VNIR+SWIR imagery over the Anhembi test site on 12th February, 2016. WV-3 is a high spatial resolution satellite operational at an altitude of 617 km with equatorial crossings at 10:30 AM. The platform comprises sixteen multispectral bands in the VNIR–SWIR wavelengths. The SWIR sub-system, which is the focus of this study, contains eight spectral bands in the 1195–2365 nm window acquired at 3.7 m spatial resolution in 14-bits dynamic range. The dataset, however, is resampled to 7.5 m resolution before releasing for

commercial use. The sensor benefits from embedded CAVIS bands to estimate aerosol and water vapor within the scene and yield atmospherically corrected at-surface reflectance data (Asadzadeh and Souza Filho 2016a).

In this study, WV-3 data were provided to us as at-surface reflectance data at 7.5 m pixel size. A combination of feature tracking (applied to continuum-removed SWIR data) and match filtering technique (Asadzadeh and Souza Filho 2016c) was used to map the distribution of oil-sands in the area.

8.4. Results and discussion

8.4.1. Field observations

In the field, we observed that where bitumen is present, the sandstone has undergone intense bleaching (due to the depletion of hematitic coatings) and the amount of clay minerals has gone up (Fig. 2a). However, the bleaching is confined mainly to the bituminous zones; within their immediate margins, ferric iron oxides reappear (Fig. 2c). Moreover, while many of the white patches were observed to be due to the excess of clay alterations, not all the white walls were necessarily argillaceous, as the bleaching of aeolian sandstone yield alike white appearance. It was witnessed that the degree of HC impregnation is a function of sandstone facies. The sand sheet facies with higher permo-porosity present greater impregnation by HCs, whereas the interdune facies is somehow devoid of HCs. Albeit, this is a rule of thumb, and there are local variations in the bitumen contents of the sandstones, as indicated by ground-based imaging spectroscopy (see section 4-5).

8.4.2. Infrared spectroscopy

The composition of oil-sands as revealed by infrared spectroscopy was fairly simple: the SWIR wavelengths was dominated by characteristics C–H, Al–OH, OH, and H_2O signatures, respectively at around 1700/2300, 2200, 1400, and 1900 nm (Figs. 4, 7, and 8b) arising from embedded bitumen and clay minerals (Fig. 7). The TBC of the oil-sand samples, estimated by the two-band normalization method described in Section 3-4, was observed to vary between 6 to 11.5 wt. %, and showed proportionality to the depth of the distinctive absorption features at ~1700 and 2300 nm. Whereas the feature at 2300 nm is very distinct and more pronounced in the continuum-removed spectra (compare Fig. 7a with 7b), the feature at 1700 nm is remarkably more correlated to the TBC of the samples (R^2 0.975 vs. 0.954), with a lower standard deviation of errors (0.19 vs. 0.25; see Table 3). This strong correlation between the TBC of the samples and the continuum-removed depth, as indicated by others (e.g. (Correa Pabón and Souza Filho 2016; Scafutto and Souza Filho (2016); Shaw and Kratochvil 1990)), opens up alternative opportunities for the estimation of TBC content using hyperspectral imagery. The TBC was also indicated to be inversely proportional to the clay content of the samples. Accordingly, as the clay content of the samples increases, the overall albedo of the sample along with the intensity of the Al–OH, OH, and H_2O tends to increase, while the distinctive HC features fading away (see Fig. 4).

Table 3. The performance of the continuum-removed HC absorption features in duplicating the TBC of the oil-sands derived from the two-bands normalization model. D_x represents the depth of the feature after local continuum removal.

Model	True TBC (wt. %)		
	Correlation	R^2	σ (%)
TBC = 48.469$D_{1700\ nm}$ + 6.9265	0.987	0.975	0.19
TBC = 16.287$D_{2300\ nm}$ + 6.5851	0.976	0.954	0.26

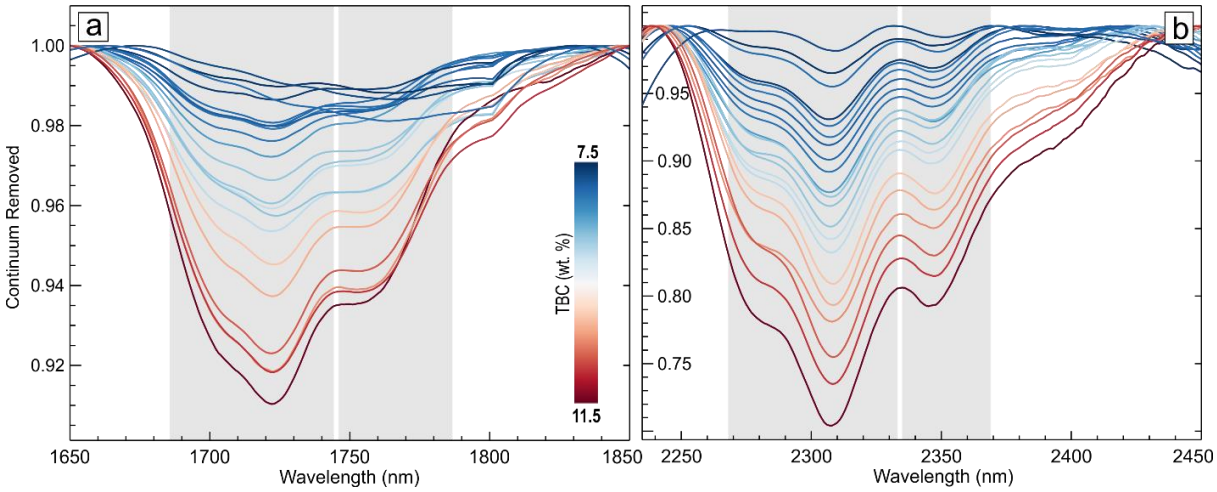

Fig. 7. Examples of the continuum-removed HC absorption features with variable bitumen contents (TBC wt. %) in the Anhembi deposit. **a)** absorption features centered at ~1700 nm. **b)** absorption features centered at ~2300 nm. Note that the feature at 2300 nm (b) is almost 2x deeper relative to the local continuum. The gray columns in the background represent the individual absorbing bands. Both features were indicated to be closely correlated with the TBC of the sands. The color table applies to both graphs.

The clay mineralogy was identified to be montmorillonite, which is marked by the distinctive Al–OH absorption feature at ~2208 nm and the lack of extra features beyond 2300 nm (Hunt 1977). Some unpublished studies have reported illite and mixed-layer illite/smectite in the oil-sands of this area using scanning electron microscopy. Such mineral assemblages, in contrast to montmorillonite, show additional absorption features between 2342–2350 nm due to Al–OH bonds (Bishop et al. 2008; Hunt 1977). In the spectral collection, however, only 4 out of 200 spectra displayed this extra feature (Fig. 8b), and thus attributable to mixed-layer illite/smectite. The majority of the samples are featureless within this range, corroborating the dominance of montmorillonitic clay. This mineral is present in almost all of the samples collected from the area, including the impregnated sands, implying that the clays are in an intimate mixture with bitumen (this subject is addressed in details using sisuCHEMA-derived imagery). The other subordinate mineral detected within the SWIR range was calcite (in Anb09; see Fig. 1), determined by a diagnostic absorption feature at 2340 nm (Fig. 8b).

Within the VNIR range, the Feox feature characteristic of the Pirambóia sandstones disappears (e.g. Spectra 1–3 in Fig. 4), presumably because of the bleaching of the hematite coatings of the facies. Wherever iron oxides are present (normally within the clay-rich zones), however, the minimum of the feature shifts towards longer wavelengths (e.g. >910 nm) that is indicative of goethite (Fig. 8a) formed after the re-oxidation of reduced ferric coatings. Ferrous iron, on the other hand, was rarely seen in the samples and the only ferrous mineral identified thus far belongs to Fe-chlorite characterized by triple iron absorptions at ~750, ~930, and ~1100 nm (Bishop et al. 2008) (Fig. 8a). In the bituminous samples, the ferrous feature is prone to be obscured by the broad absorption feature of HCs arising from π (C=C) bonding between 340–1400 nm (Fig. 4).

The XRD results corroborated the presence of montmorillonite (and quartz) as the dominant constituents of the samples. It also revealed that K-feldspars (orthoclase, sanidine),

titanomagnetite, brushite, and kaolinite constitute the minor mineralogic phases within the oil-sand samples (Fig. 9). The mineralogy of the framework grains also clarified why the spectra are dominated by the montmorillonite features, because the other mineralogic components (except a subtle feature due to kaolinite; Fig. 9) are spectrally neutral in the VNIR–SWIR range.

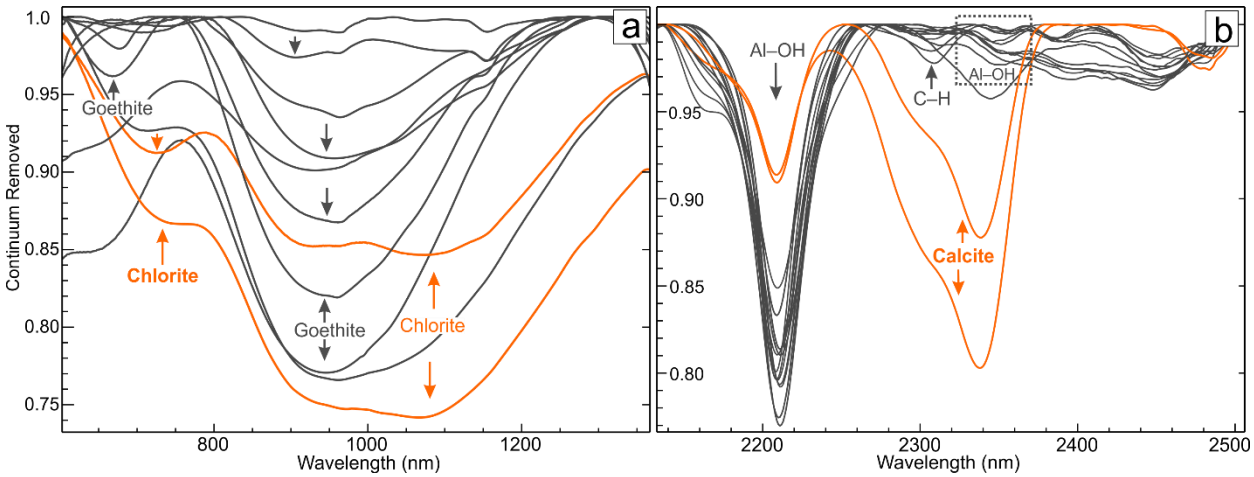

Fig. 8. Examples of oil-sands spectra and the interpreted mineralogy. **a)** features in the VNIR comprising crystall-field band absorptions of ferrous (chlorite) and ferric (goethite) minerals. **b)** absorption features in the SWIR comprising Al–OH, hydrocarbons (C–H), and calcite. The subtle features centered at ~2347 nm (enclosed by a dashed square) arise from the second Al–OH absorption of the mixed-layer illite/smectite. The two spectral collections do not necessarily come from the same samples.

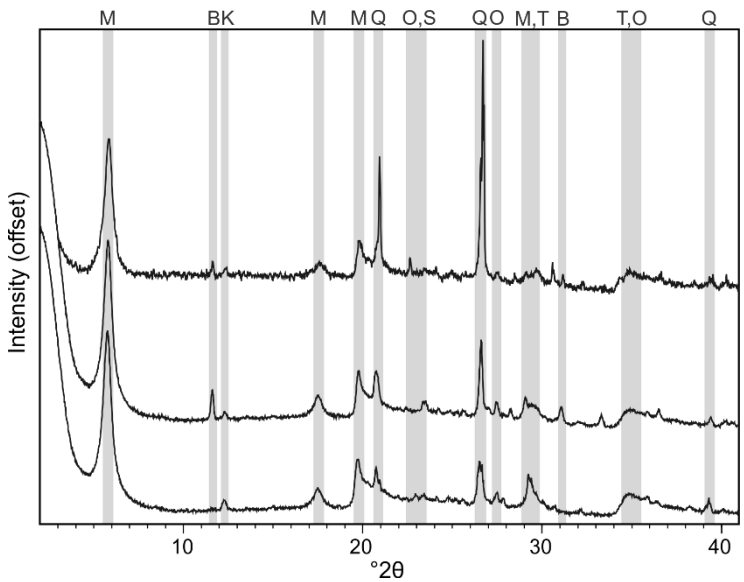

Fig. 9. X-ray diffraction patterns revealing the mineralogy of framework grains of the oil-sands. Shaded bars represent the major mineral peaks labeled as: M = montmorillonite, B = brushite, Q = quartz, O = orthoclase, S = sanidine, T = titanomagnetite, K = kaolinite.

In comparison to the samples from Pirambóia sandstones, the oil-sands are distinguished by the absence of a ferric feature (due to bleaching) and an increase in the proportion of montmorillonitic clays, presumably caused by the alteration of the host siliciclastic rocks (i.e. K-feldspars; Fig. 9). To quantify the clay content, we calculated the depth of the Al–OH feature (Fig. 8b) from both datasets and showed them as abundance histograms (Fig. 10a-b). If we put the mixed clays aside and consider the clay-dominant facies only (Fig. 10b), then the absorption feature, and thus the proportion of clays, is at least 25% higher in oil-sand samples compared to original sandstones (Fig. 10a), meaning that the emplacement of bitumen has been associated with clay formation. The

composition of the Al-smectites, however, showed no significant differences between the two collections (Fig. 10c-d). The Al–OH wavelength position in the Pirambóia sandstones spans between 2206–2211 nm that is similar to wavelength ranges of the oil-sand samples (2207–2212 nm), implying that the presence of oil has not affected nor modified the composition of existing, or newly formed, Al-smectites.

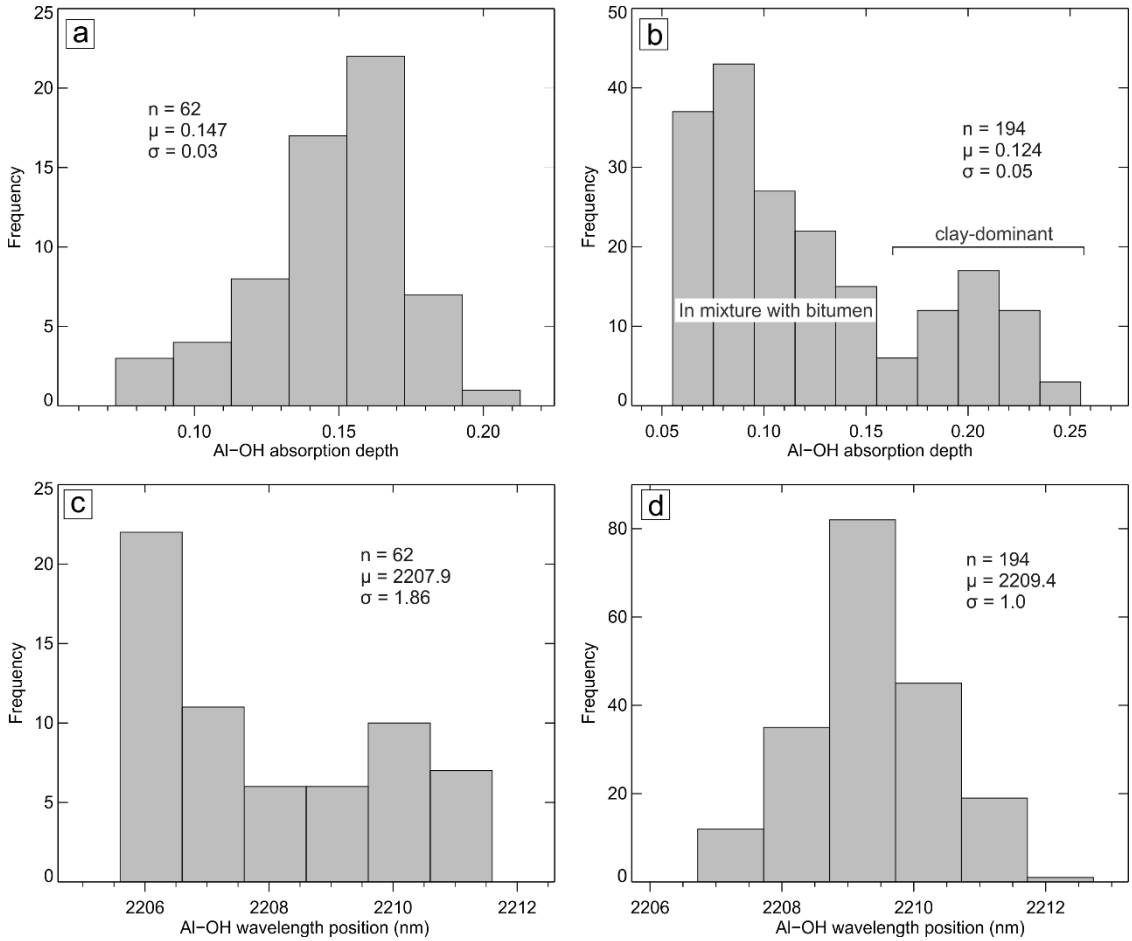

Fig. 10. Histograms of the wavelength position and depth of Al–OH absorption features. **a)** Al–OH absorption depth of Pirambóia sandstones. **b)** Al–OH absorption depth of oil-sand samples. **c)** Al–OH wavelength position of Pirambóia sandstones. **d)** Al–OH wavelength position of oil-sand samples.

Two basic similarities between some microseepage systems and the alterations induced by oil-sands were noticed: (i) the strata is devoid of kaolinite and, instead, smectites dominate the clay alterations; (ii) the bleaching of hematite coating from sandstone beds and occasional occurrences of goethite. Albeit, the secondary goethite displays its minimum absorption at wavelengths >940 nm indicating higher levels of Al-substitution in the goethite lattice and the excess of aluminum in the system. Overall, the oil-sands in the Anhembi area display very limited alteration diversity compared to a microseepage system (Fe-chlorite and calcite were limited to a few samples only) or counterpart reserves in for example, Alberta, Canada (Asadzadeh and Souza Filho 2017; Cloutis et al. 1995). We speculate that it is related to the characteristics of bitumen emplacement in the area. Likely, the oil has migrated as a very viscous and single-phase liquid, thereby inducing limited mineralogic changes in the environment. This assumption is supported by the confined iron bleaching (indicative of no gas-phase involvement) and the crucial role of permo-porosity of the

facies in oil emplacement (indicative of very viscous oils). In practice, the reappearance of iron oxide coatings marks the front of oil migrations (e.g. Fig. 2c).

8.4.3. Quality control of the image spectra

Before going to the details of the results achieved by imaging spectroscopy, we first provide a quality assessment of the spectral data provided by imaging systems in the lab and on the ground.

8.4.3.1. SisuCHEMA image spectra

Regarding the sisuCHEMA datacube, we collected representative spectra over the oil-rich and clay-rich parts of the sample using the contact probe of an ASD spectrometer and compared them to the average of corresponding pixels in the imagery. The results are shown in Fig. 11. Obviously, there is a correspondence in the overall spectral shape of the two datasets, especially over the absorption features centered at ~1900, 2200, 2300, and 2350 nm. Closer inspections, however, revealed three important differences between the two series: (i) the reflectance of the image spectra at around 1650 nm is unusually high (Fig. 11; encircled); (ii) the image data is noisier between 1500–1800 nm; (iii) the shape of the HC feature at 1700 nm is modified in the image spectra. The higher reflectance coupled with the altered absorption feature could give rise to a bias in the estimation of bitumen content using the feature derived from this type of hyperspectral imagery.

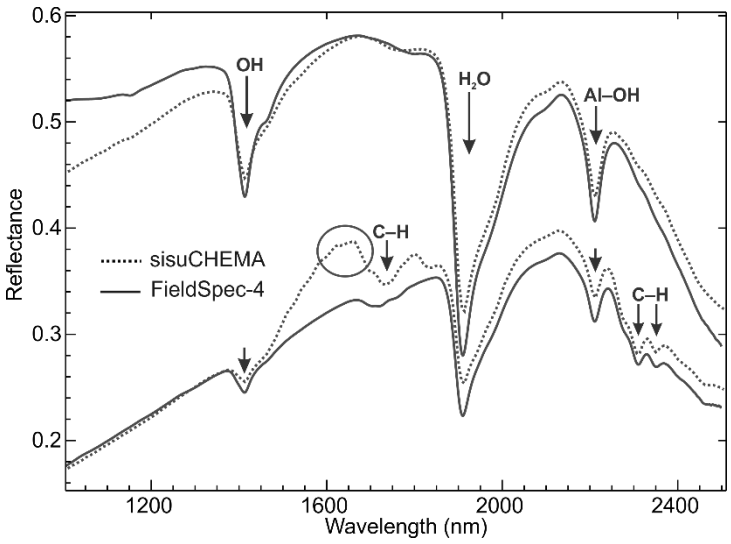

Fig. 11. Comparison of the spectra derived from of a FieldSpec-4 point-spectrometer and a sisuCHEMA hyperspectral scanner. The spectra were collected from the same sample by averaging the pixels corresponding to the aperture of a contact probe. The circle highlights a change in the reflectance (continuum) of the spectrum derived from sisuCHEMA imagery.

8.4.3.2. Ground-based image spectra

We calculated the mean and standard deviation of the pixels occurring in the center of the reference panel (n≈450 pixels) within the imagery of Outcrop-1 and then divided the mean of each band to the standard deviation (μ/σ) to yield a graph of signal quality (SNR) against wavelength (Fig. 12). This measure was used to assess the impact of preprocessing steps (marked by number 1–3 in the workflow of Fig. 6) in improving the quality of the signal of the ground-based hyperspectral data. The results denote that the steps devised in the workflow continually improve the quality of

the signal. In particular, the SNR is greatly enhanced following the atmospheric compensation, indicating the significance of this step for ground-based data. The de-noising steps, on the other hand, were highly effective beyond 2300 nm wavelengths, where HC's major absorption features occur. Considering, for example, the wavelengths at 600 and 2300 nm, the improvement in the SNR was >45% (Fig. 12). The de-noising was demonstrated not only to improve the SNR of the imagery, but also was crucial in retrieving the overall shape of the HC's absorption features in the SWIR range (Fig. 13).

We repeated the same processes for the image of the other walls. The results indicate that with the current sensor technology, direct solar illumination is critical for spectral imaging of dark targets, including the oil-sands. The data collected over Outcrop-3 were eliminated because of the low SNR linked to illumination conditions and surface wetness. Furthermore, over Outcrop-2, the poor illumination not only weakened the HC feature at 2300 nm, but also was responsible for the disappearance of HC feature centered at 1700 nm (see Fig. 14). The albedo of the image spectra located in the shadow was >2x lower compared to the well-illuminated oil-sands nearby. Although the weak absorption features could be enhanced by continuum removal, the obtained spectrum does not show a good match with its well-illuminated counterpart (Fig. 14). This is particularly problematic during the estimation of TBC over poorly-illuminated facies, as discussed in Section 4-5. In the circumstance of poor illumination, the atmospheric effect presumably dominates the recorded signal, thereby hampering the recovery of the HC's signal during atmospheric compensation. This is already evident by the downward trend of the spectra on either side of the water absorption window (note the horizontal arrows in Fig. 14).

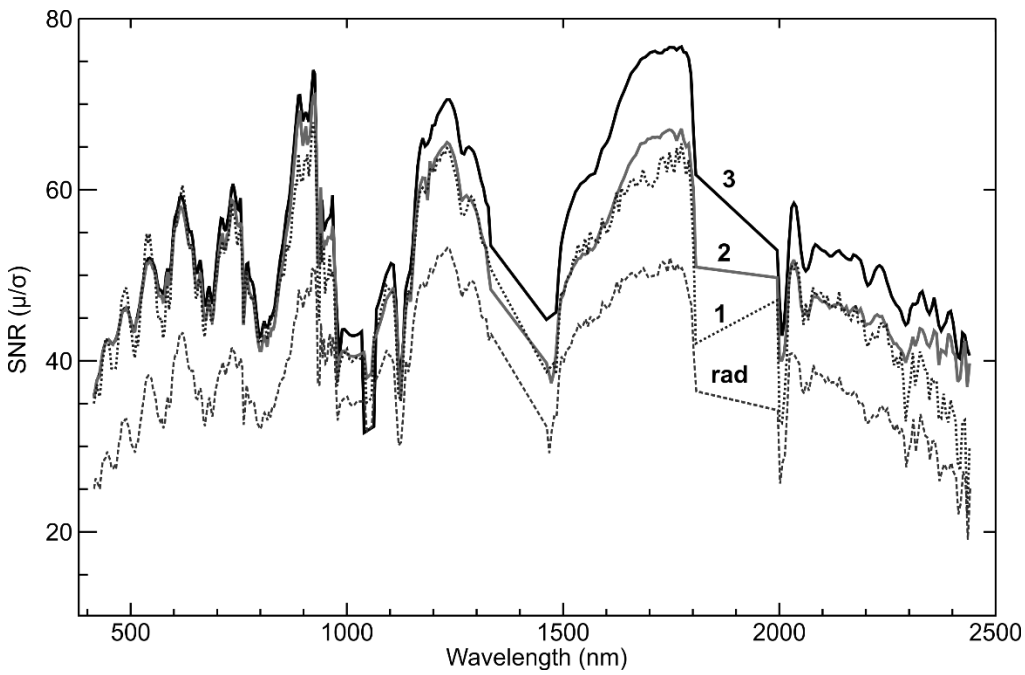

Fig. 12. Improvement in the quality (SNR) of the image following the preprocessing steps shown in Fig. 6. The SNR was calculated by ratioing the mean (μ) of the pixels over the reference panel in Outcrop-1 to the standard deviation (σ). The numbers correspond to the steps marked in the workflow. "rad" stands for 'radiance data'.

It worth mentioning that the idea of averaging several successive images over the same outcrop proved to further improve the image quality and increase the SNR in the same way as the

implemented steps of the workflow did. However, since the spatial overlap of the sequential images in this site was not considerable, we discarded this step from our preprocessing chain.

Because the spatial resolution of the ground-based data is within 1.5–3 cm range, a pixel occupies the same area as a contact probe does, so the image spectra shall be essentially comparable to laboratory spectra. Accorddingly, representative spectra for bitumen and montmorillonite from well-illuminated parts of the scene were extracted. The results are illustrated in Fig. 15. A higher albedo in the ground-based spectra compared to laboratory data is noted. The albedo measured at ∼1300 nm (± 50) was, on average, 60–70% higher in the ground-based spectra. While this difference could be due to the illumination condition of the imagery, we conjecture that it could also arise from the single white reference panel deployed for atmospheric compensation. Most likely, by using multiple reference pannels (dark, white, and gray) and the empirical line correction approach, the issue of higher albedo could be alleviated. Another reason for this difference could be the sharp contrast between the very dark oil-sand targets with the surrounding high-albedo clays that make the adjustment of the sensor complicated during data collection.

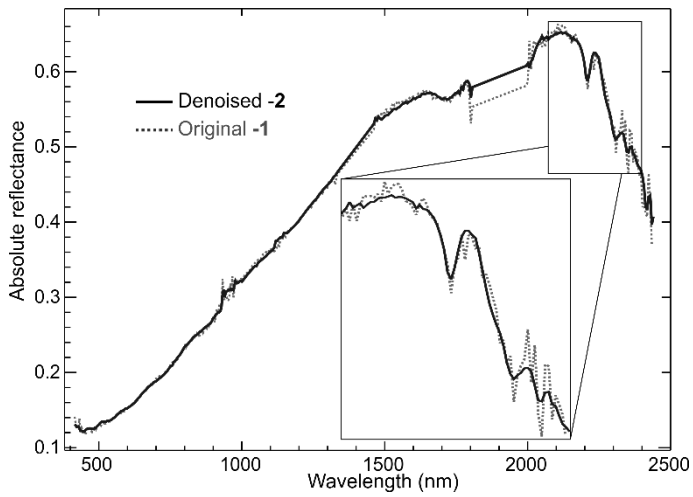

Fig. 13. Improvement in the quality of the signal and the pattern of absorption features of the image spectra after denoising. The numbers correspond to the steps defined in the workflow of Fig. 6.

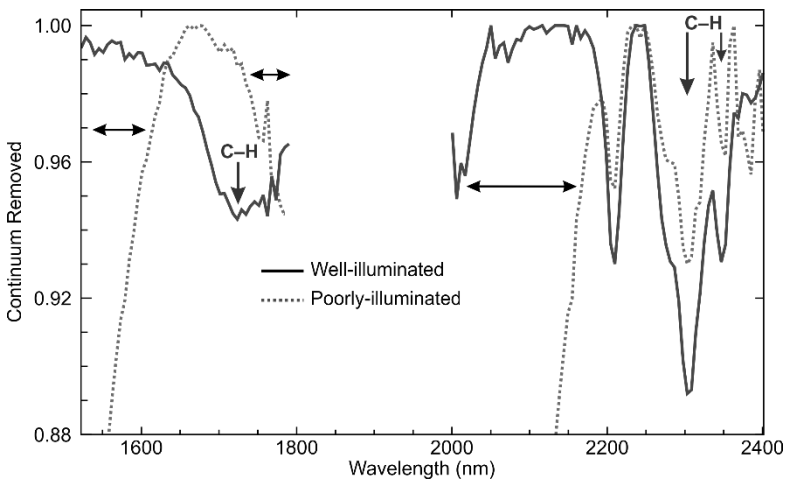

Fig. 14. The effect of illumination on bitumen absorption features. The pixels lying in the shadow retain weaker absorption features at 2300 nm and lose the feature at 1700 nm altogether. Each spectrum is the mean of 20 adjacent pixels from Outcrop-2 imagery shown in Fig. 19a. The horizontal arrows indicate the dominance of atmospheric water absorption in the poorly-illuminated signal.

Furthermore, it was observed that the long-wave side of the 1700 nm feature is noisier in the ground-based data (Fig. 15; encircled). This problem, which was also noticed in airborne data (see Kokaly et al. (2013)), is due to the impact of residual atmospheric contamination. This fact indeed underpins the need for accurate atmospheric correction using the empirical line method.

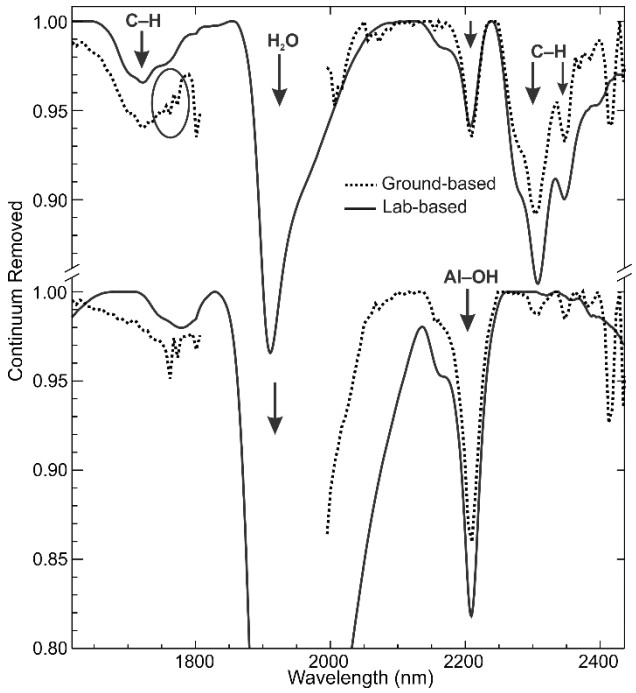

Fig. 15. Comparison of the continuum-removed spectra derived from the ground-based imaging system and a spectrometer in the lab. The clay-rich (Al–OH) and oil-rich (C–H) spectra were yielded by averaging four pixels over homogeneous and well-illuminated portions of the image shown in Fig. 19a. The circle indicates the residual atmospheric contamination. The laboratory spectra were converted to absolute reflectance for this aim.

8.4.4. Close-range spectral imaging

The bitumen content of the sample calculated by the two-band normalization method revealed that the oil is not homogenously distributed within the host sandstone (Figs. 16b and 17b). While the TBC of the sample reaches up to 10.8 wt. %, it conforms to the normal distribution with the means and standard deviations equal to 6.0 and 1.4 wt. %, respectively (Fig.17a). The lowest bitumen levels correspond to the clay-rich interlayers and vice versa. Mathematically speaking, the correlation coefficient of bitumen and clay contents (derived from Al–OH absorption depth; not shown here) was calculated to be -0.48. Although not visible to the naked eye (e.g. Fig. 16a), montmorillonitic clay (manifested by Al–OH absorption feature) is remarkably ubiquitous and occurs virtually in all pixel spectra. The depth of the feature in the bituminous pixels was at least one third that of the clay-rich pixels.

The image of Al–OH wavelength position derived from sisuCHEMA data also show a meaningful trend. The bituminous pixels tend to show their minimum Al–OH wavelengths between 2206–2208 nm, whereas this parameter for clayey interlayers is 2212 nm (Fig. 16c). It is not yet certain, however, if the changes in the wavelength position of Al–OH features are linked to compositional variation in smectitic clays or it arises from the mixing effect of HCs, especially from a subtle asymmetric feature at 2200 nm (see Fig. 2 in Asadzadeh and Souza Filho (2017)).

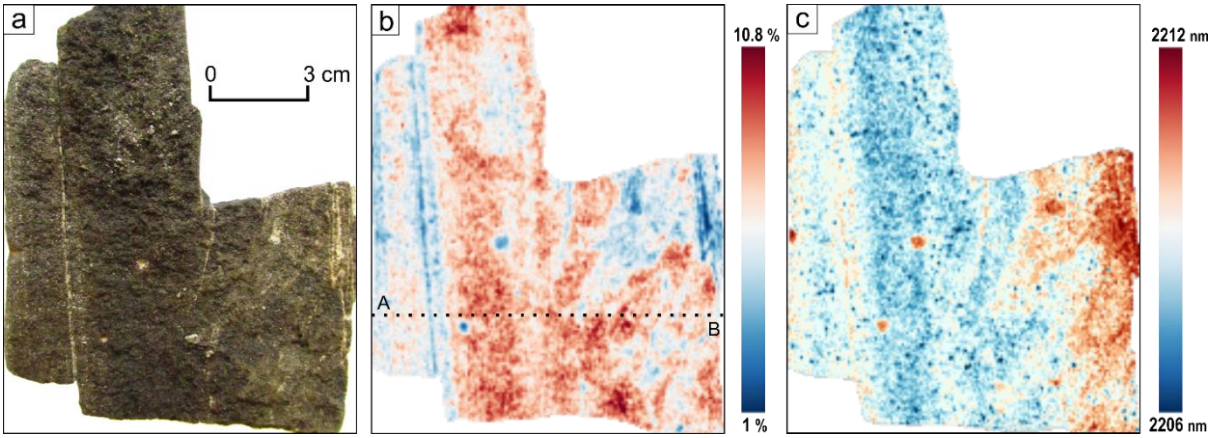

Fig. 16. Spectral products derived from close-range imaging spectroscopy. **a)** photograph of the sample scanned by a sisuCHEMA imaging system. **b)** Total Bitumen Content (TBC; wt. %) image calculated by employing the two-bands normalization method and stretched by 1% from the low end. A TBC profile along A–B is shown in Fig. 17b. **c)** montmorillonite compositional map corresponding to Al–OH wavelength position derived from polynomial fitting to 2151–2244 nm wavelengths. Figs (a) and (c) are modified after Asadzadeh and Souza Filho (2016c).

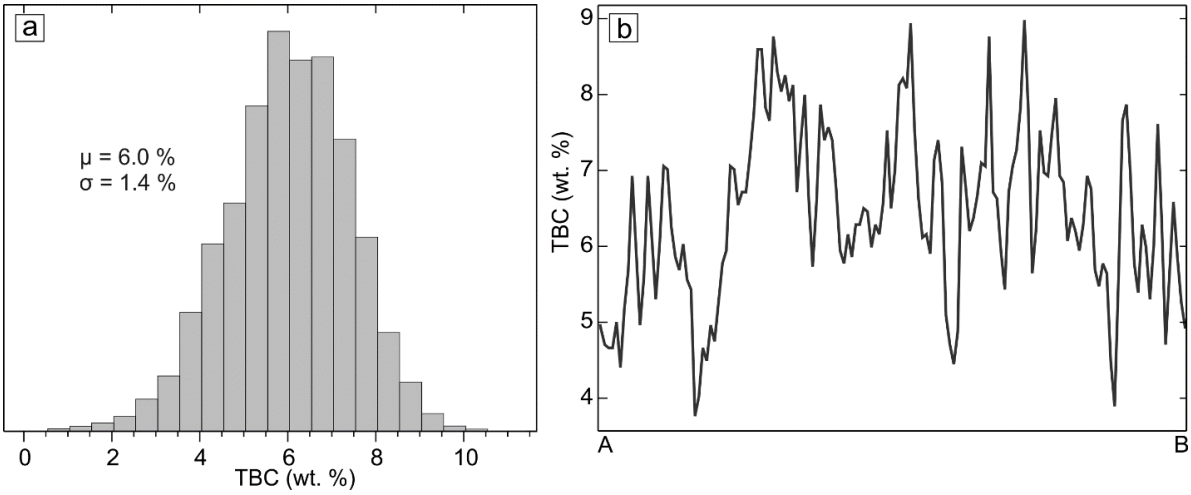

Fig. 17. a) Histogram of the TBC of the oil-sand sample exhibited in Fig. 16a. **b)** TBC profile corresponding to line A–B in Fig. 16b.

8.4.5. Ground-based spectral imaging

The TBC derived from the two-band normalization method revealed to overestimate the bitumen in both outcrops by 1–2.9%. The overestimation was especially noticeable over shadowy or poorly illuminated faces (e.g. A/B in Figs. 18a and 19a). The consequence was a bimodal distribution in the histogram of the TBC images, particularly over Outcrop-2. To resolve this issue, we calibrated the continuum-removed depth (CRD) of the HC feature at 2303 nm against the TBC image by using carefully selected pixels (n = 61) from well-illuminated parts of the image. The established relationships between the two parameters (R^2 = 0.93 and σ = 0.77 wt. %) is indicated in Eq. 1:

$$TBC = 54.188 \times CRD_{2303} + 0.1784 \qquad (1)$$

As is illustrated in Fig. 19b, by reducing the variations in the albedo/illumination of the imagery, continuum removal can improve the estimation of TBC of the facies. We avoided using the feature at 1700 nm, because as discussed in Section 4-3-2, it is prone to be missed in the poorly-illumination portion of the image. By deploying Eq. 1, the TBC was mapped seamlessly throughout

the imageries (Fig. 18b and 19c). The only exception was a patch in Fig. 19c, on which due to low SNR, the TBC was still overestimated. In this product, we also discovered some contamination from the non-photosynthetic vegetation (see, for example, the arrows in Fig. 19a), which is likely due to the similarity of the absorption features of the two compounds, particularly around 1700 nm wavelengths (Fig. 20). In comparison, the feature of non-photosynthetic vegetation, however, is wider at 1700 nm wavelengths and shows its minimum at longer wavelengths. Moreover, the slope of the continuum line between 1650–2100 nm is negative in non-photosynthetic vegetation and it exhibits an additional absorption feature centered at 2100 nm, which is absent in bitumen (Fig. 20). To omit dry vegetation, we first mapped its extent within the image by applying the SAM technique to a spectral subset between 1600–2210 nm wavelengths using an image-derived endmember (see Fig. 19a). Then, the result was thresholded and transformed into a mask to exclude the non-photosynthetic vegetation. The refined result is shown in Fig. 19c.

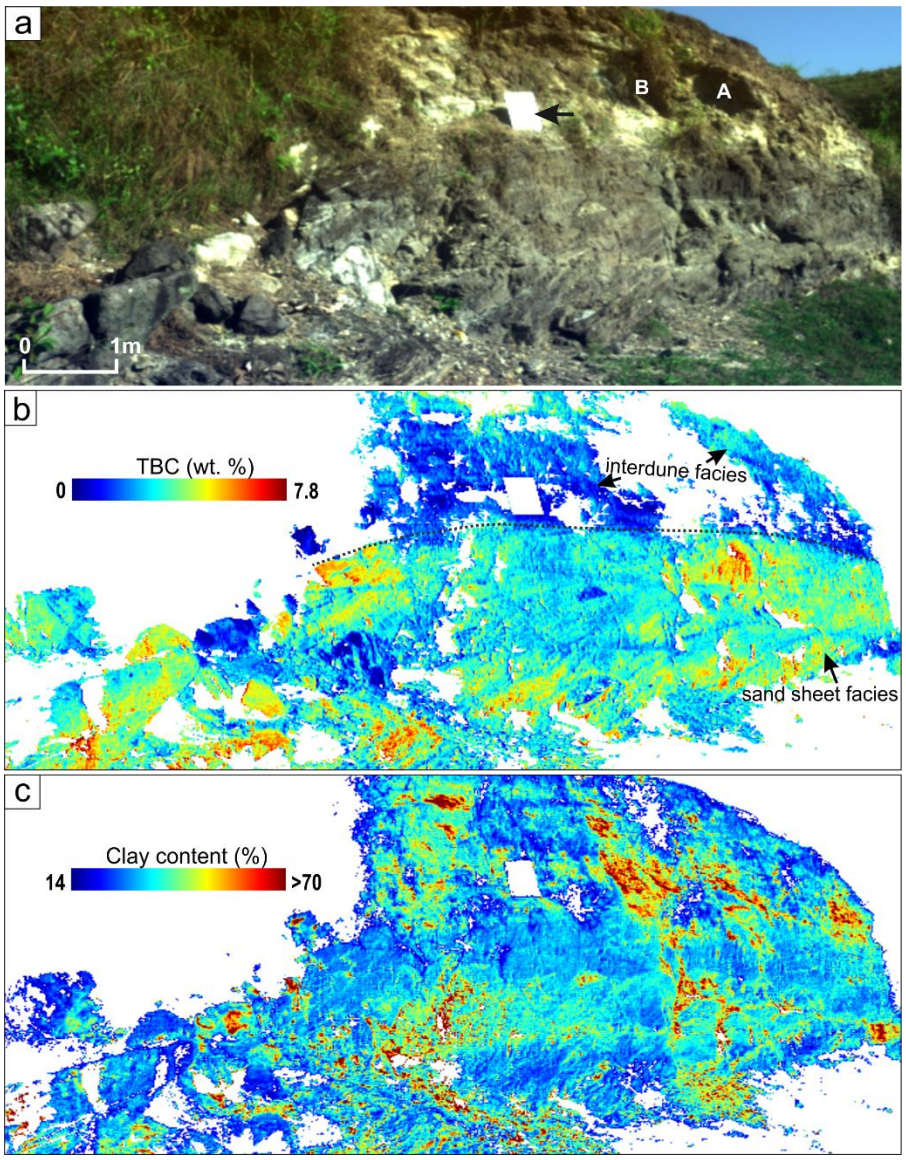

Fig. 18. The oil-sand Outcrop-1 in the Anhembi deposit. **a)** natural-color composite. **b)** total bitumen content (Wt.%) image calculated by continuum-removed depth (CRD) of the HC feature at 2303 nm and masked for dry vegetation and shadows. **c)** relative proportion of montmorillonitic clays normalized against the deepest Al–OH absorption within the scene. This image corresponds to the right rectangle in the photograph of Fig. 2e. A and B in (a) represent examples of shadowy pixels prone to TBC overestimation. The black arrow in (a) indicates the $BaSO_4$ reference panel.

Knowing that the original two band normalization method (Rivard et al. 2010) has a dispersion of ~1.5%, and considering the 0.77% dispersion of the model in Eq. 1, the accuracy of the displayed TBC maps is within ±2.27 wt. %. To lower the estimation error of spectral assaying (i.e. <1.5%), one needs to calibrate the ground-based images against a direct oil content determination process known as Dean-Stark analysis (see for example Speta et al. (2015) and references therein).

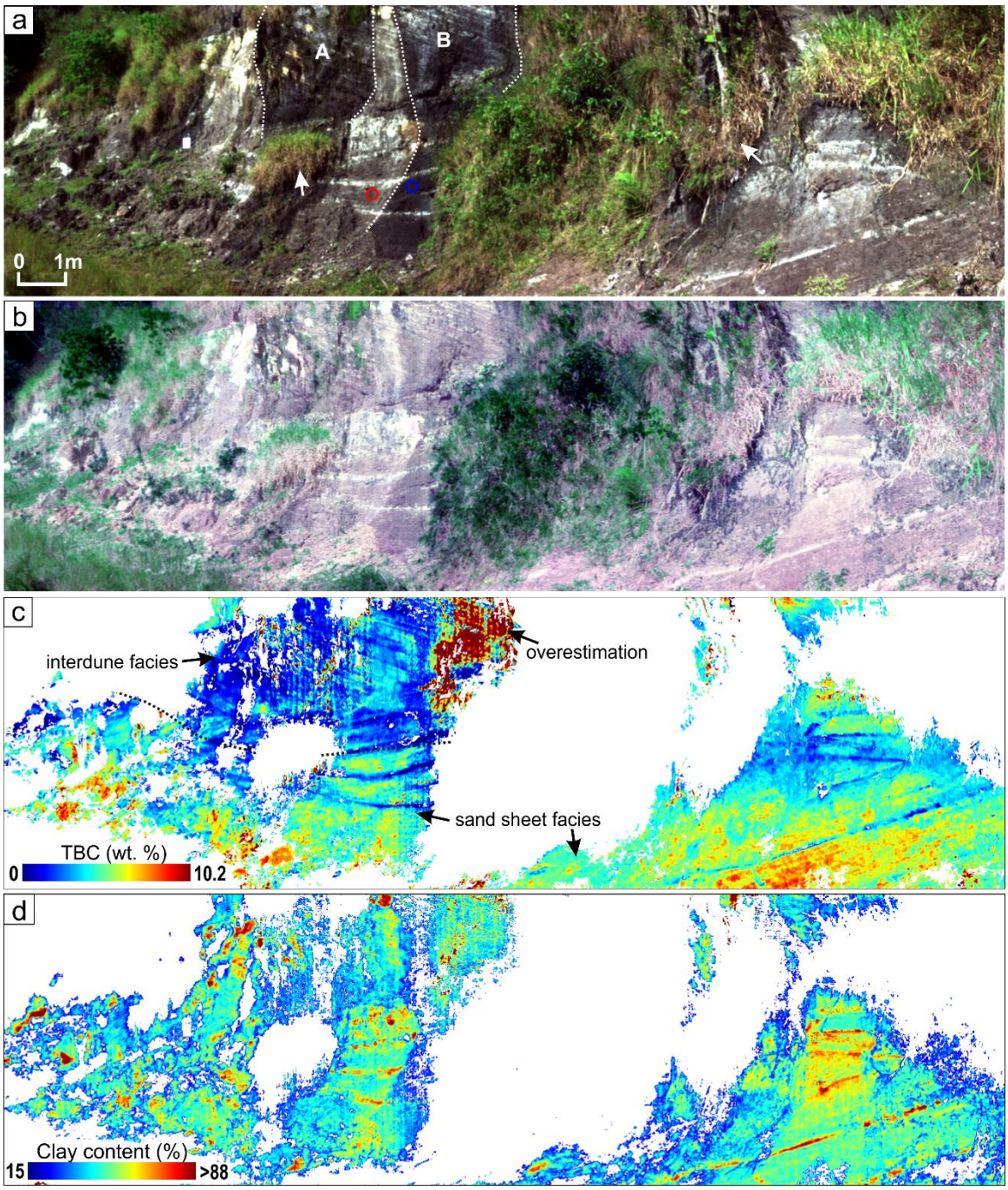

Fig. 19. The oil-sand Outcrop-2 in the Anhembi deposit. **a)** natural-color composite. **b)** continuum-removed natural-color composite. **c)** total bitumen content (Wt.%) image calculated by the continuum-removed depth of the HC feature at 2303 nm and masked for dry vegetation and shadows. **d)** relative proportion of montmorillonitic clays normalized against the deepest Al–OH absorption within the scene. The red and blue circles in (a) indicate, respectively the well-, and poorly-illuminated pixels incorporated in the plot of Fig. 14. The white arrows represent some of the non-photosynthetic vegetation excluded from the final TBC image. A and B represent the poorly-illuminated faces prone to TBC overestimation. This image corresponds to the left rectangle in the photograph of Fig. 2e.

The heterogeneity in the distribution of bitumen is also observed in the outcrop scale maps. Based on the images shown in Figs. 18 and 19, the oil-sand walls could be divided into two main facies: (i) the one richer in clays with low bitumen contents corresponding to interdune facies; (ii) the other one richer in bitumen with lower clay contents corresponding to sand sheet facies (see Table 4 for the statistics). The bitumen impregnation of the facies is controlled by permo-porosity of the strata and to a lesser extent, the clay content.

Table 4. Basic statistics of the estimated bitumen within the interdune and sand sheet facies, as indicated in Figs. 18 and 19.

Outcrop	Facies	TBC (wt. %)	
		μ	σ
1	Interdune	2.3	1.1
	sand sheet	3.8	1.2
2	Interdune	2.9	1.8
	sand sheet	4.9	1.9

Regarding the clay contents, a relative abundance map was calculated by normalizing the depth of the Al–OH feature against the deepest absorption within the scene. The results are shown in Figs. 18c and 19d. Evidently, montmorillonitic clay remains mixed with bitumen even at outcrop scale maps. The clay fraction at bituminous facies was determined to be at least 30% of the clay-rich zones by assuming a linear scale for the depth of the Al-OH absorption feature. This fact suggests that the clays were already formed (at least partially) at the time of bitumen emplacement and, as noted by Thomaz Filho et al. (2008), acted as a barrier against horizontal oil migrations. Since the bleaching (as the most distinctive mineralogic changes in the strata) is observed only in the oil-impregnated zones, then it is likely that the oil was originally very viscous, rather than breaking down to heavy oil by subsequent biodegradation processes.

The trend in the wavelength position of the Al–OH feature, however, was not as significant as close-range data (Fig. 16c) hance was discarded from ground-based image maps.

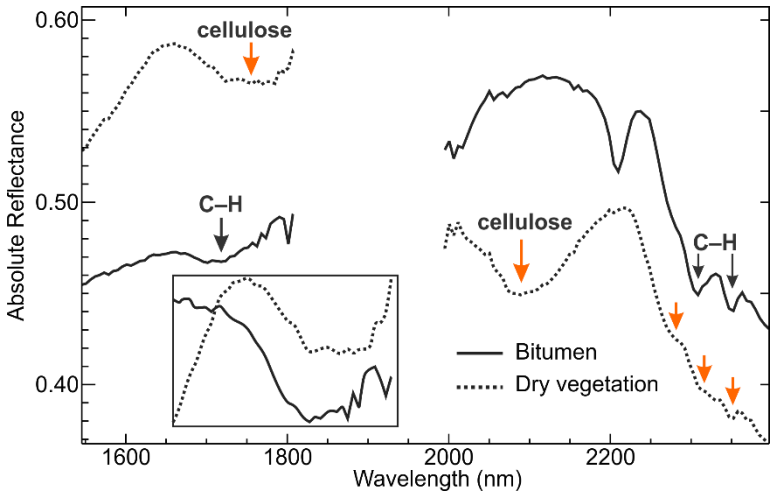

Fig. 20. Comparison of the spectrum of bitumen and non-photosynthetic vegetation extracted from the imagery of Outcrop-2 (arrows and red circle in Fig. 19a). Each spectrum is the average of several individual pixels. The inset graph shows the absorption feature between 1600–1800 nm after continuum removal.

8.4.6. Airborne spectral imaging

The distribution of oil-sands in the Anhembi deposit mapped by airborne imaging spectroscopy is shown in Fig. 21a-b. Obviously, the airborne system is capable of duplicating the HC (and clay) features as sensed by close-range (and ground-based) spectroscopy. The close association of bituminous outcrops with clays in the provided maps is because besides intimate mixing, the two components are areally mixed in the sub-meter scale of the maps (see, for example, Fig. 2f). The two-fold mixtures, however, could adveresly affect the accuracy of bitumen assying using the continuum-removed feature derived from the airborne data. The mean and standard deviation of the oil-sand outcrops in Fig. 21b was estimated to be 4.0 and 1.1 wt. %, respectively, which is lower than the contents estimated over vertical faces. The closest figures to the average reported in Thomaz Filho et al. (2008) is, in fact, provided by ground-based data (∼5 vs. 5.5%).

The lowest bitumen content detected (2.6 %) closely corresponds to the dispersion of the model in Eq. 1 (2.3 wt. %). This figure could also approximate the detection limit of the far-range data for estimation of the bitumen content over this deposit. The model used for TBC determination was developed originally for close-range spectroscopy in perfect illumination conditions. Further calibration data and investigation is required to appreciate the effects of scaling the laboratory results to far-range remote sensing data and establish a robust and scale-independent model for bitumen content estimations.

8.4.7. Space-borne spectral imaging

The oil-sands sensed by the WV-3 data were characterized by a conjunction of absorption features at SWIR band-4 (S4) and band-6 (S6) (Fig. 22a), respectively indicative of bitumen and Al–OH clays. During the processing, we noticed that the HC feature at S4 could be accentuated by continuum removal. In the continuum-removed image, the HC absorption was indicated by a minimum at S4 relative to S3 and S5 shoulders (Fig. 22a). The feature was present within a large part of the oil-sand pixels. By making use of this distinctive feature, the exposed oil-sands were successfully mapped at 7.5 m resolution. The results are shown in Fig. 21c. Evidently, there is a good match between this map and the one yielded from airborne data shown in Fig. 21b. In comparison, the WV-3 derived anomalies are in parts larger, but in terms of the numbers are limited, because some anomalies already mapped by airborne data are missing (marked by stars in Fig. 21c). The missing anomalies consist mainly of small oil-sand outcrops (∼2–5 m^2) not resolvable by the GSD (3.7 resampled to 7.5 m) of the sensor. The areal expansion of bituminous anomalies is likely due to the effect of larger pixel size (3.7 vs. <1.0 m) and subsequent resampling to 7.5 m resolution. The resampling (in conjunction with geolocation error) is also responsible for a shift (∼7 m equivalent to 1 pixel) in the locality of some anomalies (indicated by arrows in Fig. 21c) in the resulting map. It was not possible to estimate the TBC by the current model as it is based on the feature centered at 2300 nm. It was not also straightforward to estimate the bitumen contents of the pixles using the continuum-removed depth of the S4 band.

This experiment revealed that where bituminous targets occupy one-sixth (1/6) of the area of the resampled SWIR pixel (∼10 m^2), it could be unambiguously detected via WV-3 data. This overestimated fraction, however, should be explicit to Anhembi oil-sands of which due to the contrasting effect of coexisting clays, the HC targets were accentuated, thereby facilitating remote

sensing-based detection. A conservative estimate, however, could be around one-third of a pixel, as demonstrated using simulation data (Asadzadeh and Souza Filho 2016a).

The pixel spectra derived from WV-3's SWIR bands at two different sites conform to ground-truth data collected using a spectrometer in the lab. In comparison, the depth of the C–H and Al–OH features as well as the overall reflectance at S5 were lower in the imagery. In one case, the pixel spectrum was associated with an additional chlorophyll absorption (Fig. 22a).

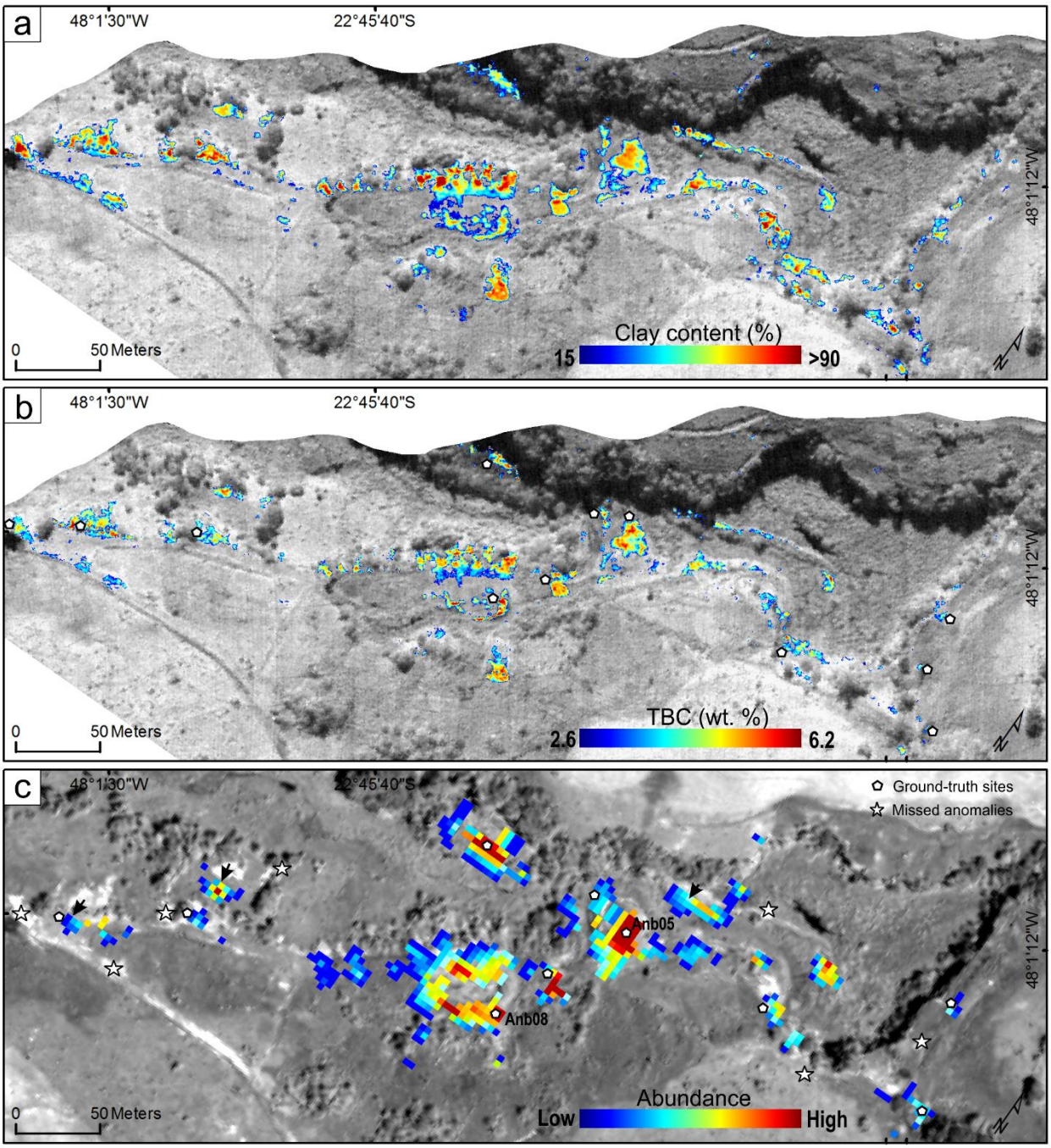

Fig. 21. The distribution and abundance of clays and bitumen over the Anhembi deposit derived from airborne hyperspectral and WV-3 superspectral data. **a)** relative proportion of montmorillonite normalized against the deepest Al–OH absorption within the imagery. **b)** the TBC map estimated from airborne data using Eq. 1. **c)** relative abundance of bitumen derived from WV-3 data using match filtering technique. The white pentagons and stars indicate the anomalies verified on the ground and missed in the WV-3 map, respectively. The labeled anomalies (Anb05 and 08) correspond to the photographs shown in Fig. 2f and 2g, respectively. The WV-3 spectra from these

sites are compared to ground spectra in Fig. 22a. Note that a subset of the image in Fig. 1 is displayed here and the points in (b) are shifted to not obscure the anomalies. The arrows in (c) indicate the offset in the locality of anomalies.

The finding that WV-3's SWIR bands can resolve the subtle HC feature at 1700 nm wavelength is very promising, because the bitumen in this area consists of extra heavy HCs with very low API° (~5). Certainly, in other cases in which the target consists of lighter oils (corresponding to higher API°), it displays a deeper absorption feature (Lammoglia and Souza Filho 2011); thus constitute an easier target to detect using orbital WV-3 imagery (Fig. 22b). In other words, oil-sands establish the toughest target (worst-case scenario) to detect via WV-3 data, as it retains one of the lowest signal intensities among HC-bearing materials at 1700 nm wavelengths.

A substance that could potentially interfere with HC's feature at 1700 nm is non-photosynthetic vegetation (see Fig. 20). In this area, the extent of which the non-photosynthetic vegetation interfered with bitumen map (as a false-positive anomaly) is unknown, because fieldwork was not simultaneous with data acquisition nor there were systematic attempts to outline it within the area.

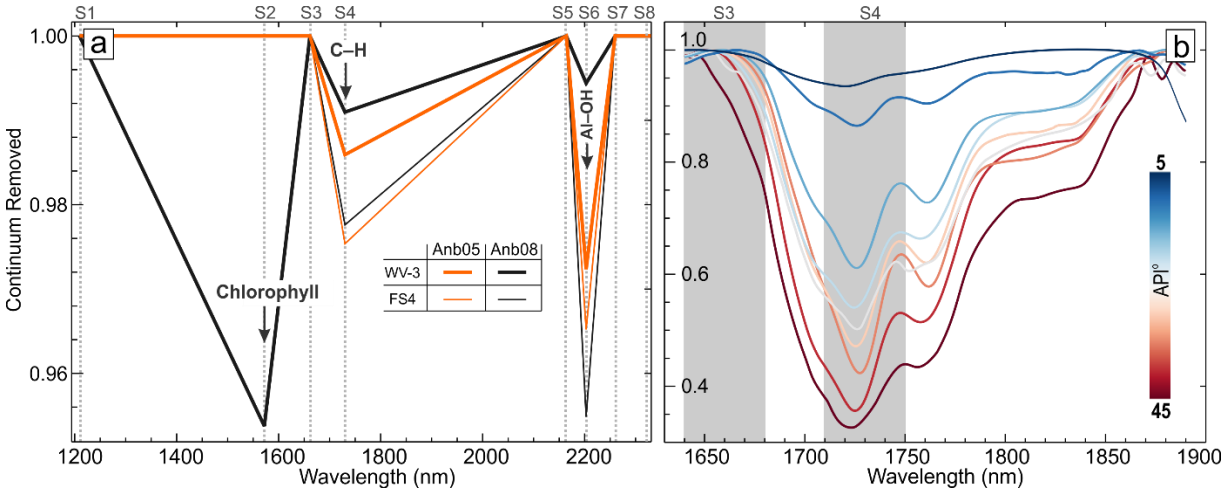

Fig. 22. a) Comparison of pixel spectra derived from WV-3 data with ground-truth data collected in the lab using a FieldSpec-4 (FS4) spectrometer over two sites namely Anb05 and Anb08 (see Fig. 21c for the localities). S_i on top of the vertical dashed lines indicates the central wavelength of the SWIR bands S1 to S8. **b)** the effect of API° on the intensity of absorption feature centered at ~1700 nm (adapted from Lammoglia and Souza Filho (2011)). The gray bars in (b) exhibit the wavelength range of SWIR bands S3 and S4.

8.5. Conclusion

Multi-scale studies of oil-sands in the Anhembi area via imaging and non-imaging spectroscopic techniques provided a complete picture of the changes in the content and composition of bitumen and accompanying alteration minerals. This work indicated that the HC signature, especially the one centered at 2300 nm, is consistent and comparable among scales, and upon employing proper calibration data and analysis technique, it could confidently map the bitumen content of the oil-sands at all imaging scales. The TBC of the deposit was determined to vary between 0–11.5 % with the mean of ~5 wt. %.

The induced alterations in the Pirambóia sandstones (i.e. montmorillonitic clay and iron oxides bleaching) bear some resemblance to microseepage alterations. The clay minerals were indicated to be ubiquitous and mixed intimately with bitumen at all studied scales. The restricted alteration diversity along with the heterogeneity in the distribution of bitumen, which is strongly controlled by

permo-porosity of the facies, were interpreted to be indicative of a single-phase flow of very viscous (API° ∼5) HCs to the reservoir.

Ground-based technology can effectively bridge the gap in bitumen content determination over vertical faces of oil-sand mines. For effective bitumen determination on the ground, besides the high SNR of the imaging system, a good illumination condition was demonstrated to be required. To recover HC's features in poorly illuminated conditions, at least two reference panels (white + dark ± gray) should be deployed within the field of view of the system. Spectral denoising also seems to be crucial and a prerequisite for spectroscopic analysis.

WV-3 satellite data was proved to resolve HC's feature at 1700 nm wavelength by its SWIR band-4 provided that the target is large enough to encompass ∼20–30% of the area of the resampled SWIR pixel.

Acknowledgement

We are indebted to Prof. Celio Pasquini for facilitating our access to sisuCHEMA instrument based in the Institute of Chemistry at UNICAMP. We are thankful to Emanuel Hernandez for assistance in the fieldworks. Raymond Kokaly is acknowledged for extensive help with WV-3 data ordering and advice on data calibration. We appreciate DigitalGlobe for providing the WV-3 data free of charge. SA and CRSF thank FAPESP and CNPq for the research grants No. 2015/06663-7 and 2008-7/303563, respectively. This work was partially supported by the GRSG student award (2016) given to SA.

References

Araújo, C.C., Yamamoto, J.K., & Rostirolla, S.P. (2006). Tar sandstones in the Paraná Basin: a study of the occurrences in the Anhembi Structural High (in Portuguese). *Petrobras Bulletin of Geosciences, 14*, 47-70

Araújo, C.C., Yamamoto, J.K., Rostirolla, S.P., Madrucci, V., & Tankard, A. (2005). Tar sandstones in the Paraná Basin of Brazil: structural and magmatic controls of hydrocarbon charge. *Marine and Petroleum Geology, 22*, 671-685

Araujo, L.M., Triguis, J.A., Cerqueira, J.R., & da S. Freitas, L.C. (2000). The Atypical Permian Petroleum System of the Parana Basin, Brazil. In M.R. Mello, & B.J. Katz (Eds.), *Petroleum systems of South Atlantic margins* (pp. 377-402). Tulsa, OK, U.S.A.: AAPG Memoir 73

Asadzadeh, S., & Souza Filho, C.R. (2016a). Investigating the capability of WorldView-3 superspectral data for direct hydrocarbon detection. *Remote Sensing of Environment, 173*, 162-173

Asadzadeh, S., & Souza Filho, C.R. (2016b). Iterative Curve Fitting: A Robust Technique to Estimate the Wavelength Position and Depth of Absorption Features From Spectral Data. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 5964-5974

Asadzadeh, S., & Souza Filho, C.R. (2016c). A review on spectral processing methods for geological remote sensing. *International Journal of Applied Earth Observation and Geoinformation, 47*, 69-90

Asadzadeh, S., & Souza Filho, C.R. (2017). Spectral remote sensing for onshore seepage characterization: A critical overview. *Earth-Science Reviews, 168*, 48-72

Bishop, J.L., Lane, M.D., Dyar, M.D., & Brown, A.J. (2008). Reflectance and emission spectroscopy study of four groups of phyllosilicates: smectites, kaolinite-serpentines, chlorites and micas. *Clay Minerals, 43*, 35-54

Cloutis, E.A. (1989). Spectral reflectance properties of hydrocarbons: Remote-sensing implications. *Science, 245*, 165-168

Cloutis, E.A., Gaffey, M.J., & Moslow, T.F. (1995). Characterization of minerals in oil sands by reflectance spectroscopy. *Fuel, 74*, 874-879

Correa Pabón, R.E., & Souza Filho, C.R. (2016). Spectroscopic characterization of red latosols contaminated by petroleum-hydrocarbon and empirical model to estimate pollutant content and type. *Remote Sensing of Environment, 175*, 323-336

Dias, K.D.N., & Scherer, C.M.S. (2008). Cross-bedding set thickness and stratigraphic architecture of aeolian systems: An example from the Upper Permian Pirambóia Formation (Paraná Basin), southern Brazil. *Journal of South American Earth Sciences, 25*, 405-415

Ellis, J.M., Davis, H.H., & Zamudio, J.A. (2001). Exploring for onshore oil seepas with hyperspectral imaging. *Oil and Gas Journal, 99*, 49-58

Greenberger, R.N., Ehlmann, B.L., Jewell, P.W., Birgenheier, L.P., & Green, R.O. (2016). Detection of organic-rich oil shales of the green river formation, Utah, with ground-based imaging spectroscopy. In, *8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)* (pp. 1-5). Piscataway, NJ: IEEE

Greenberger, R.N., Mustard, J.F., Ehlmann, B.L., Blaney, D.L., Cloutis, E.A., Wilson, J.H., Green, R.O., & Fraeman, A.A. (2015). Imaging spectroscopy of geological samples and outcrops: Novel insights from microns to meters. *GSA Today, 25*, 4-10

Hunt, G.R. (1977). Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics, 42*, 501-513

Kokaly, R.F., Couvillion, B.R., Holloway, J.M., Roberts, D.A., Ustin, S.L., Peterson, S.H., Khanna, S., & Piazza, S.C. (2013). Spectroscopic remote sensing of the distribution and persistence of oil from the Deepwater Horizon spill in Barataria Bay marshes. *Remote Sensing of Environment, 129*, 210-230

Kruse, F.A., Lefkoff, A.B., Boardman, J.B., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., & Goetz, A.F.H. (1993). The Spectral Image Processing System (SIPS) - Interactive Visualization and Analysis of Imaging Spectrometer Data. *Remote Sensing of Environment, 44*, 145–163

Kurz, T.H., Buckley, S.J., & Howell, J.A. (2013). Close-range hyperspectral imaging for geological field studies: workflow and methods. *International Journal of Remote Sensing, 34*, 1798-1822

Kurz, T.H., Dewit, J., Buckley, S.J., Thurmond, J.B., Hunt, D.W., & Swennen, R. (2012). Hyperspectral image analysis of different carbonate lithologies (limestone, karst and hydrothermal dolomites): the Pozalagua Quarry case study (Cantabria, North-west Spain). *Sedimentology, 59*, 623-645

Lammoglia, T., & Souza Filho, C.R. (2011). Spectroscopic characterization of oils yielded from Brazilian offshore basins: Potential applications of remote sensing. *Remote Sensing of Environment, 115*, 2525-2535

Martins, L.L., Cruz, G.F., Souza, E.S., & Severiano, R.H.J.P. (2015). Compositional variability in bitumens from eolian sedimentary facies of Pirambóia formation, Paraná basin, Brazil: characterizing biodegradation. *Geochimica Brasiliensis, 29*, 154-169

Murphy, R.J., & Monteiro, S.T. (2013). Mapping the distribution of ferric iron minerals on a vertical mine face using derivative analysis of hyperspectral imagery (430–970 nm). *ISPRS Journal of Photogrammetry and Remote Sensing, 75*, 29-39

Okyay, Ü., Khan, S., Lakshmikantha, M., & Sarmiento, S. (2016). Ground-Based Hyperspectral Image Analysis of the Lower Mississippian (Osagean) Reeds Spring Formation Rocks in Southwestern Missouri. *Remote Sensing, 8*, 1018

Rivard, B., Lyder, D., Feng, J., Gallie, A., Cloutis, E., Dougan, P., Gonzalez, S., Cox, D., & Lipsett, M.G. (2010). Bitumen content estimation of Athabasca oil sand from broad band infrared reflectance spectra. *The Canadian Journal of Chemical Engineering, 88*, 830-838

Rodger, A., Laukamp, C., Haest, M., & Cudahy, T. (2012). A simple quadratic method of absorption feature wavelength estimation in continuum removed spectra. *Remote Sensing of Environment, 118*, 273-283

Scafutto, R.D.P.M., & Souza Filho, C.R. (2016). Quantitative characterization of crude oils and fuels in mineral substrates using reflectance spectroscopy: Implications for remote sensing. *International Journal of Applied Earth Observation and Geoinformation, 50*, 221-242

Schwartz, G., Ben-Dor, E., & Eshel, G. (2013). Quantitative Assessment of Hydrocarbon Contamination in Soil Using Reflectance Spectroscopy: A “Multipath” Approach. *Applied Spectroscopy, 67*, 1323-1331

Shaw, R.C., & Kratochvil, B. (1990). Near-infrared diffuse reflectance analysis of Athabasca oil sand. *Analytical Chemistry, 62*, 167-174

Speta, M., Rivard, B., Feng, J., Lipsett, M., & Gingras, M. (2015). Hyperspectral imaging for the determination of bitumen content in Athabasca oil sands core samples. *AAPG Bulletin, 99*, 1245-1259

Thomaz Filho, A., Mizusaki, A.M.P., & Antonioli, L. (2008). Magmatism and petroleum exploration in the Brazilian Paleozoic basins. *Marine and Petroleum Geology, 25*, 143-151

Watson, K. (1993). Processing remote sensing images using the 2-D FFT–noise reduction and other applications. *Geophysics, 58*, 835-852.

Chapter 9

Concluding remarks and future studies

Spectral remote sensing techniques were successfully deployed to characterize different aspects of petroleum manifestations on the surface. It was used to target the footprints of microseepage systems and provided novel mineralogical indicators (in terms of abundance, composition, and crystallinity) for microseepage characterization. The techniqe was also successful in mapping oil indications corresponding to oil-sands, oil-seeps, and oil-contaminated soils by making use of HC's signatures at 2300 and 1700 nm wavelengths. Thus far, remote sensing of petroleum was confined to hyperspectral data; however, we substantiated that multispectral datasets from WV-3 and to a lesser extent ASTER instruments could be utilized for onshore HC detection. Our study indicated that WV-3's SWIR bands could resolve even the subtle feature arising from extra heavy HCs. The finding is encouraging as the oil-sands retain the weakest signal at ~1700 nm range. A particular concern about direct HC detection is the sub-pixel (and intimate) mixture of oil with the background in which it occurs. In vegetated areas, for instance, the feature might not be straightforward to distinguish from non-photosynthetic vegetation that bears some resemblance to HC signal. Careful spectral simulations in future works should be able to address this issue. Even though we indicated that HC targets in the order of ~25–30% of a pixel could be resolved using WV-3 data, the finding is not yet conclusive, particularly for oil-seepage detection. More case studies in different petroliferous terrains are required to account for the effect of seepage dimension (as well as background composition) on target detection.

The multi-scale spectroscopic study of oil-sands indicated that the feature at 2300 nm is consistent among scales and maintains a linear relationship with the TBC of samples/pixels. However, to obtain a robust spectral model for this aim, remote sensing data should be calibrated against geochemical assaying. It is worth mentioning that despite the merits of the feature in bitumen characterization, a standard technique for HC detection requires taking both HC's features into consideration.

The detection of microseeps in the arid environment was proved not to be a formidable challenge considering the wealth of soils and sediments exposition at the surface. However, in a tropical climate, where bedrock is partly exposed and the soil profile is deeply weathered, this methodology should be employed with caution. An effective remote sensing method in such environments should incorporate in-situ spectroscopy and measure the changes relative to a control area. Further studies should be conducted in different climates, bedrock covers, and geologic settings to provide insights into the full range of secondary changes in microseepage systems. Furthermore, the studies should be expanded to incorporate the longwave infrared (8–14 µm) wavelengths. We believe a new microseepage model capable of accounting for a large variety of secondary changes should be devised and employed in future investigations.

Foreseen enhancements in the AMISA package, including automated absorption feature detection and spectral deconvolution techniques, will turn it into a powerful tool for spectral analysis in the VNIR–SWIR–LWIR range in the near future.

References

Asadzadeh, S., & Souza Filho, C.R. (2016a). Investigating the capability of WorldView-3 superspectral data for direct hydrocarbon detection. *Remote Sensing of Environment, 173*, 162-173

Asadzadeh, S., & Souza Filho, C.R. (2016b). Iterative Curve Fitting: A Robust Technique to Estimate the Wavelength Position and Depth of Absorption Features From Spectral Data. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 5964-5974

Asadzadeh, S., & Souza Filho, C.R. (2016c). A review on spectral processing methods for geological remote sensing. *International Journal of Applied Earth Observation and Geoinformation, 47*, 69-90

Asadzadeh, S., & Souza Filho, C.R. (2017). Spectral remote sensing for onshore seepage characterization: A critical overview. *Earth-Science Reviews, 168*, 48-72

Appendix 1

Spectral analysis of fresh and weathered rocks overlying the Alborz hydrocarbon trap, Qom area, Iran: remote sensing implications[16]

1. Introduction

The near-vertical seepage of hydrocarbons (HC) from a trap to the surface gives rise to a diverse array of mineralogical and geochemical changes in the overlying soils and sediments (Schumacher 1996). To investigate the effect of weathering on microseepage-induced alterations in an arid environment, we surveyed the lithofacies overlying Alborz oilfield located in the Qom district, central Iran. The reservoir is covered in large part by sandstone, siltstone, marl, and conglomerate belonging to the Upper-Red Formation (URF) (Amini 2001). During fieldworks, we noticed that a thin but discernible weathering crust blankets the outcrops of the URF in the study area (Fig. 2). To characterize this pervasive crust, we collected > 160 samples from 11 traverses in a way to incorporate both fresh cuts and weathered coatings topping the URF bedrocks.

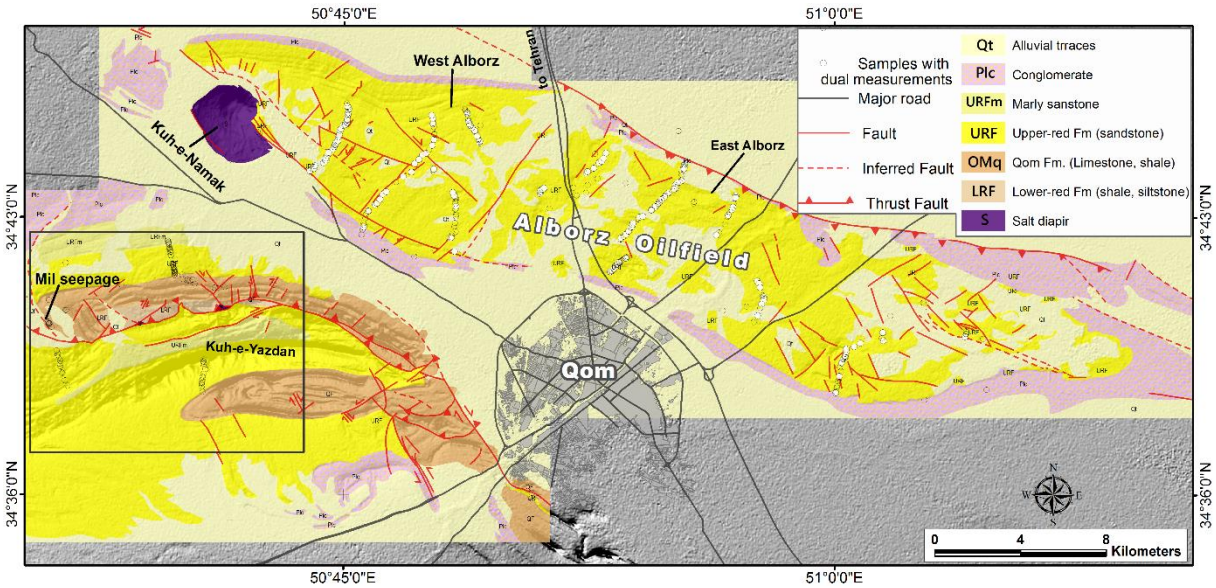

Fig. 1. The sampling sites with dual spectral measurements overlaied on the simplified geologic map of the area.

In the spectroscopic lab, both the fresh cuts and the weathered covers (intact surfaces) were systematically measured using a FieldSpec-4 spectrometer. The results were compiled into a particularly paired up spectral library (Fig. 2). The library was then analyzed by the AMISA package (Asadzadeh and Souza Filho 2016) to retreive the wavelength of minimum, depth, width, and asymmetry of the diagnostic absorption features centered at ~670, ~900, ~1900, ~2200, ~2290, and ~2340 nm wavelengths. These retreived spectral parameters were then used to assess the slight changes in the abundances and compositions of diagenetic minerals from fresh cuts to the corresponding weathered coatings.

[16] A version of this chapter was presented as a poster in the 35th International Geological Congress (IGC), Cape Town, South Africa.

Fig. 2. The weathering crust in which blankets the outcrops of the URF in the region.

2. Results and discussion

The results revealed that iron oxides/oxyhydroxides (including hematite, goethite, and ferrihydrite), clays (montmorillonite and nontronite), carbonates (calcite with traces of ankerite), and gypsum are present as alteration minerals in the sample suite. However, we recorded distinctive differences in the absorption features, the intrinsic composition, and the distribution of the mineral collections. The band absorption minimum for iron oxides in fresh cuts was centered between 885–915 nm interval. This absorption systematically shifted some 5 nm towards shorter wavelengths for the weathered surfaces (Fig. 4a). Goethite, evident by its absorption feature at ~670 nm (yellowness index), was absent in 15% of the weathered samples and where it occurred, the depth of the feature was 15% lower relative to fresh surfaces (Fig. 4b). A similar trend was observed in the abundance of nontronite by considering its subtle absorption band at 2290 nm (Fig. 4f). An inverse tendency was also observed in the abundance of carbonates and montmorillonitic (Al–OH) clays. The carbonate feature at ~2340 nm, on the other hand, showed to be twofold deeper in the fresh cuts. The reduction in montmorillonite proportion from fresh to weathered crust was around 25% (Fig. 4c-e).

The modification of the wavelength minimum of iron oxides signifies that goethite is transformed to hematite at the top of the surface. Hematite is known to favor positive Eh and neutral to alkaline pH that is characteristics of hot and semi-arid climates (Turner, 2003). Accordingly, the tranformation of goethite into hematite could indicate a change in the Eh of the environment at the topmost cover of the sediments. A similar transformation from nontronite to montmorillonite has likely taken place in the crust capping the fresh beds. Moreover, the weathered surface hosted more gypsum, which was marked by an increase in the width of the H_2O absorption feature at 1900 nm (Fig. 4f-h).

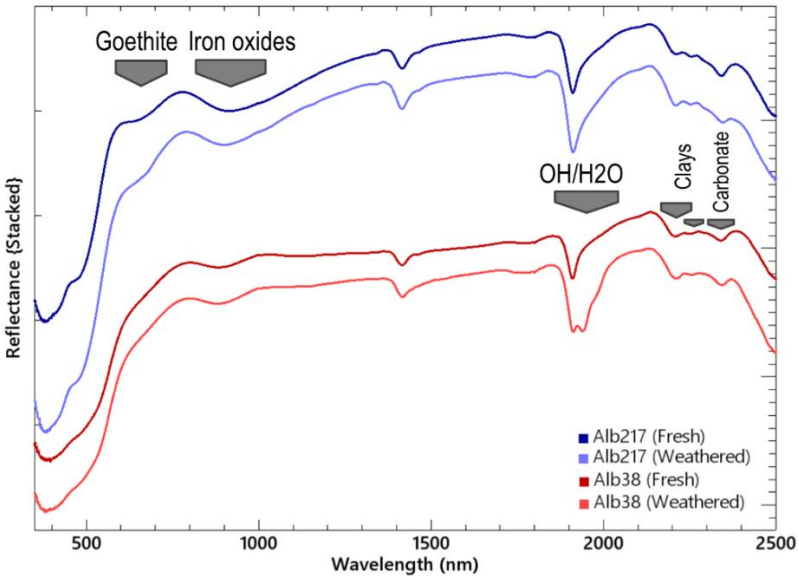

Fig. 3. Examples of the paired up spectra collected from fresh cuts and the weathered coating of the samples. Gray arrows show the assessed features.

3. Conclusion

A comparison between the mineralogic content of fresh surfaces against weathered counterparts over Alborz microseepage system in the Qom region indicated that in arid environments, a thin but pervasive weathering (oxidized) blanket could overly the bedrocks and partially modify the signal recorded by a remote sensing system. This signifies that a remote sensing survey over HC accumulations in arid regions requires taking the effect of surficial weathering into consideration for accurate mapping of the alteration footprints of underlying microseepage system.

References

Amini, A. (2001). Red colouring of the upper red formation in central part of its basin, Central zone, Iran. Journal of Sciences, Islamic Republic of Iran, 12, 145-156

Asadzadeh, S., & Souza Filho, C.R. (2016). Iterative Curve Fitting: A Robust Technique to Estimate the Wavelength Position and Depth of Absorption Features From Spectral Data. IEEE Transactions on Geoscience and Remote Sensing, 54, 5964-5974

Schumacher, D. (1996). Hydrocarbon-induced alteration of soils and sediments. In D. Schumacher, & M.A. Abrams (Eds.), Hydrocarbon migration and its near-surface expression (pp. 71-89). Tulsa, OK, U.S.A.: AAPG Memoir 66.

Turner, P., 2003. Red Beds, in: Middleton, G.V. (Ed.), Encyclopedia of Sediments and Sedimentary Rocks. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 555–557.

Fig. 4. Scatterplots of the wavelength, depth, width, and asymmetry of the absorption features exhibited in Fig. 3. **a)** Wavelength position of the ferric iron feature. **b)** depth of the goethite absorption feature centered at 680 nm. **c)** depth of the Al-OH absorption feature centered at 2200 nm. **d)** asymmetry of the Al–OH absorption feature. **e)** depth of the carbonate absorption feature centered at 2340 nm. **f)** depth of the Fe–OH (nontronite) absorption feature centered at 2290 nm. **g)** Wavelength position of the H_2O absorption feature at 1900 nm. **h)** width of the H_2O absorption feature at 1900 nm. The dashed red line indicates the 1-to-1 line.

Appendix 2

Hydrocarbon detection with spaceborne remote sensing platforms[17]

Hydrocarbons (HC) have several absorption bands in the short-wave infrared (SWIR) wavelengths due to the overtones and combinations of stretching fundamentals with the features happening at 2300 and 1700 nm comprising the most prominent of them. HCs have been successfully identified and mapped with hyperspectral imaging systems. However, the potentials of orbital multispectral instruments to replicate this capability has remained largely unknown. To address the capability of common multispectral data including WorldView-3 (WV-3), Sentinel-2, Landsat ETM/OLI, and ASTER for direct HC detection onshore, we conducted a series of experiments using simulated and real spectral datasets. In the lab, we recorded the spectra of several HCs yielded from different types of oils and oil-sand samples collected from Brazilian and Iranian basins using a FieldSpec-4 spectrometer and a sisuCHEMA hyperspectral imaging system. Also, we employed airborne ProspecTIR hyperspectral imagery over twelve synthetic oil-shows with different background soils. All datasets, which are described in details in Chapter 8, were resampled to the response of aforementioned sensors and processed with the aim of detecting HCs spectrally. Some known outcrops of oils and oil-contaminated soils currently exposed in the Qom area, Iran was also used as a control area to confirm (or not) the capabilities of operational multispectral data for HC detection in real exploration scenarios.

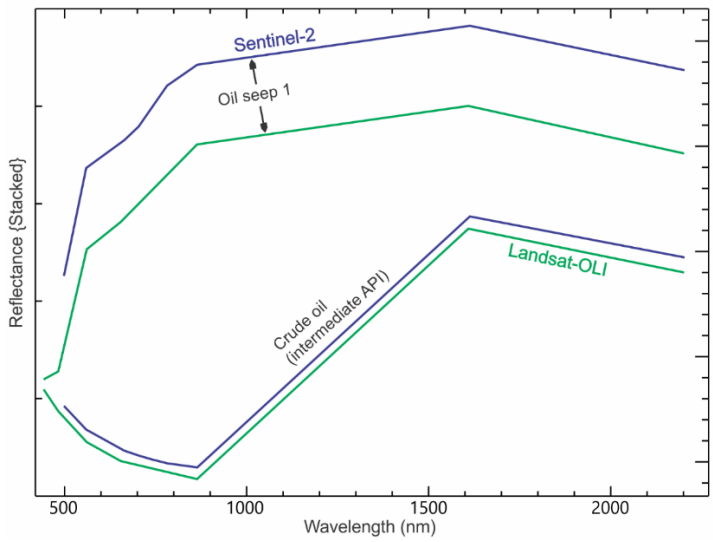

Fig. 1. Spectra of oil-bearing samples resampled to Sentinel-2 and Landsat-OLI instruments.

The results indicated that both Landsat and Sentinel-2 lack proper bands to resolve HC features (Fig. 1). The use of partial unmixing with these datasets led to the detection of a few boxes from the simulated oil-shows, albeit with a large error of commission (i.e., confusion with other dark substances in the scene). The sensors also failed to detect oil patches of the Qom area altogether.

[17] A version of this chapter was presented as an oral talk in the 35th International Geological Congress (IGC), Cape Town, South Africa.

In contrast, ASTER's band 8 and WV-3's band 12 coincide with intense HC features, respectively at 2300 nm and 1700 nm, thereby sounds to be appropriate for HC detection (Fig. 2). ASTER and WV-3 showed comparable results in detecting the extents of HC in the oil-sand sample when the sub-millimeter image cube was resampled to their SWIR's response (Fig. 3). However, their performance was different in the case of synthetic oil-show. Using the relative-absorption band depth (RBD) and three bands from each sensor (b6+b9/b8 for ASTER and b11+b13/b12 for WV-3), WV-3 was able to detect all the 12 targets with >2.5 times higher average score (231/91) relative to ASTER, which detected 11 targets, albeit with higher commission errors (see Table 1). ASTER had average to poor performance in detecting oil-contaminated soils in Qom (Fig. 4). The resultant map showed the excessive confusion of oil anomalies with limestone, shadowy areas, and some dark minerals (i.e. chlorites). However, through careful anomaly selection, we verified the occurrence of oil in at least seven of the proposed targets by this sensor (Fig. 5). Despite the exceptional performance of simulated WV-3 data in this study, however, its full potential for direct HC detection is yet to be tested with real datasets (this is addressed in Chapter 10).

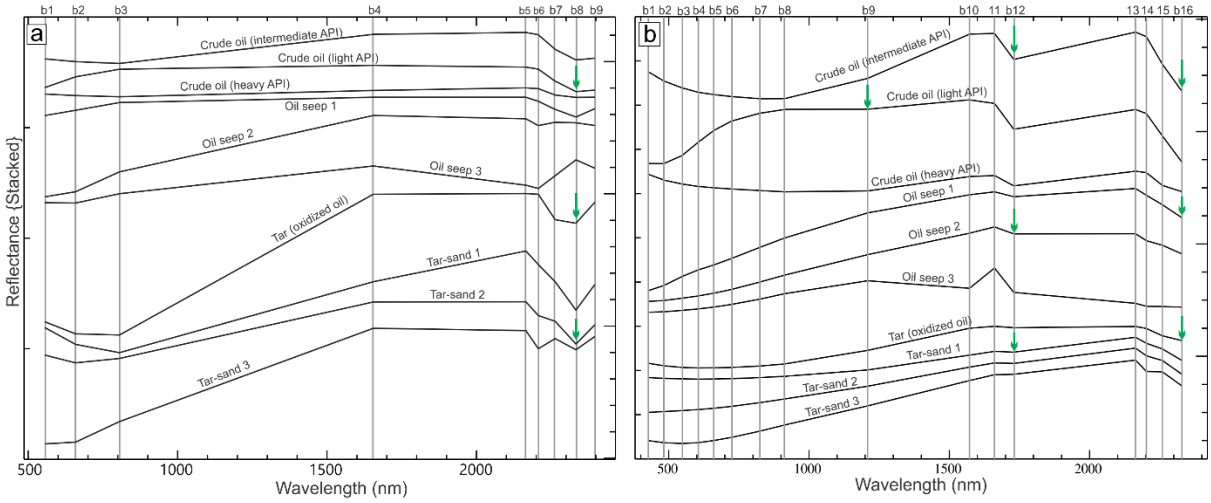

Fig. 2. Spectra of oil-bearing samples resampled to a) ASTER and b) WV-3 instruments. Arrows show the position of HC absorption features. The original spectral library is shown in Fig.3a in chapter 8.

Table. 1. The results of pixel counting derived from oil-show simulated data resampled to ASTER and WV-3. Note that the number of oil-bearing pixels detected by WV-3 is >2.5 times higher than ASTER.

Sensor	# hit pixels per box												Total	
	1	2	3	4	5	6	7	8	9	10	11	12	Pixels	Boxes
WV-3	9	20	18	22	13	34	27	36	4	19	16	13	231	12
ASTER	3	9	10	8	8	16	14	11	0	4	6	2	91	11

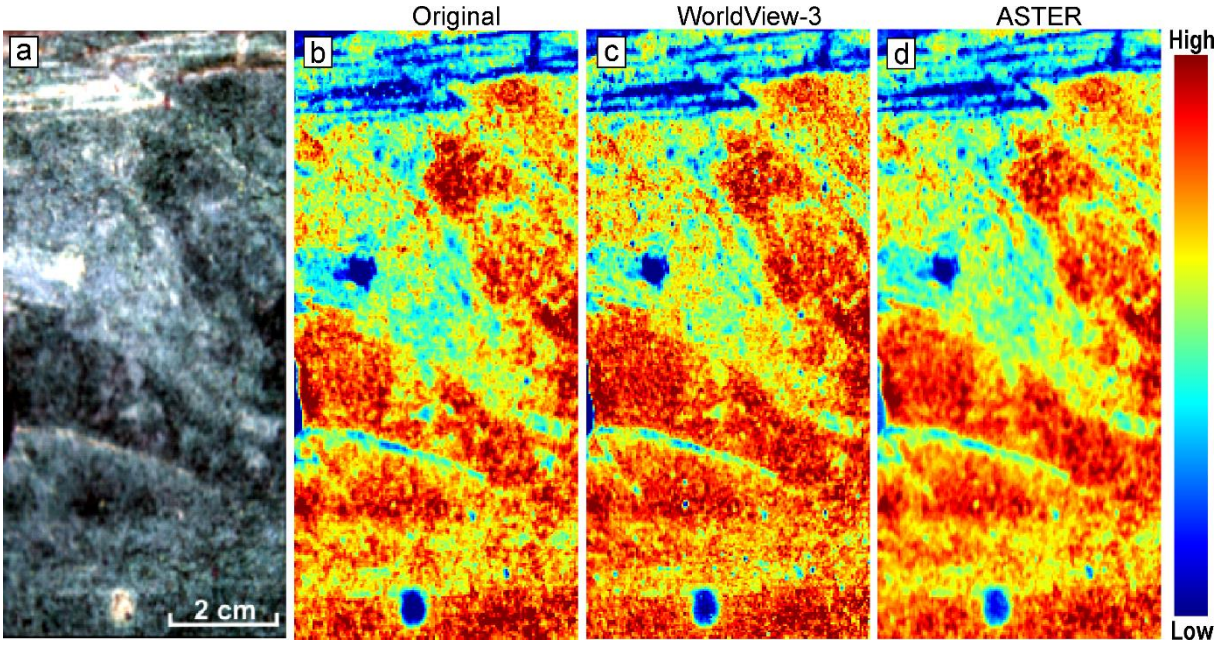

Fig. 3. The results of the oil-sand experiment. **a)** false-color composite imagery of hypercube data. **b)** RBD over 2300 nm feature of original data. **c)** RBD (b11+b13/b12) over 1700 nm feature after being resampled to WV-3 response. d) RBD (b6+b9/b8) over 2300 nm feature after being resampled to ASTER response. Images in (c) and (d) are, respectively 93 and 94% correlated to the original image in (b).

While ASTER SWIR detectors are no longer functional, the successful launch of WV-3 sensor opens up new opportunities for spaceborne detection of HCs. Such capability has important implications for oil and gas industry. The global coverage of WV-3 and its fine spatial resolution (7.5 m) could facilitate the monitoring of petroleum installations and help prospectors to screen frontier basins for oil seep indications.

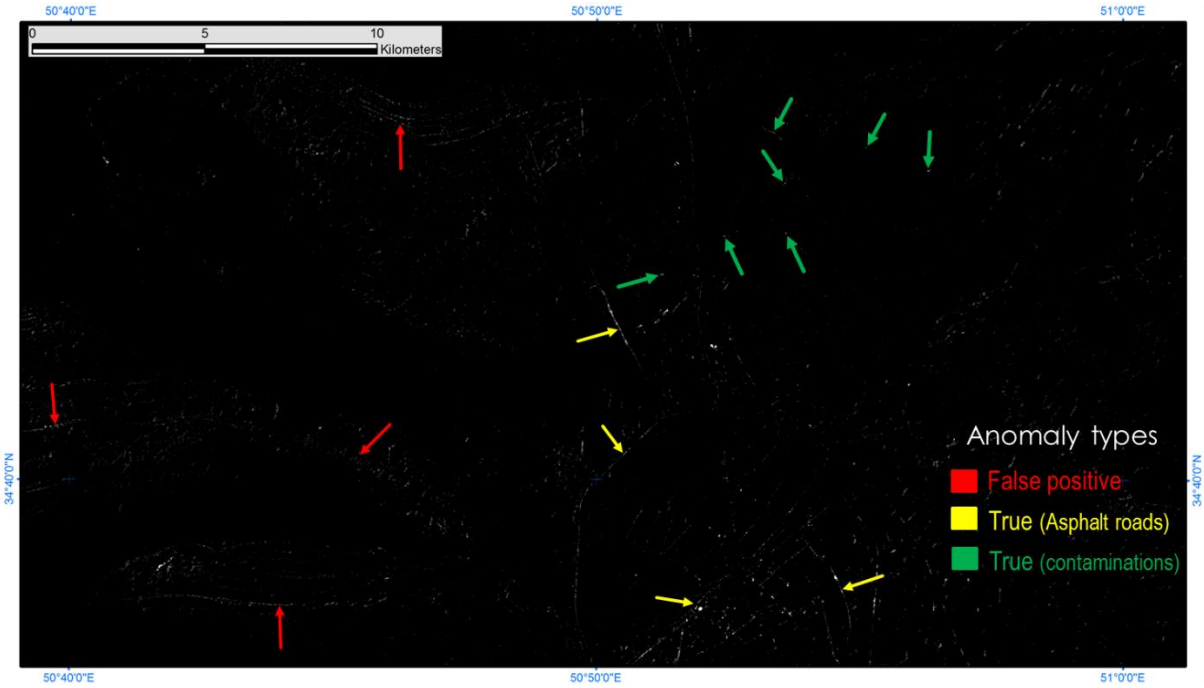

Fig. 4. HC anomaly map derived from ASTER data over Qom area. Three types of anomalies are defined: false (red arrows), true (green arrows), and true, but not related to oil contaminations (yellow arrows).

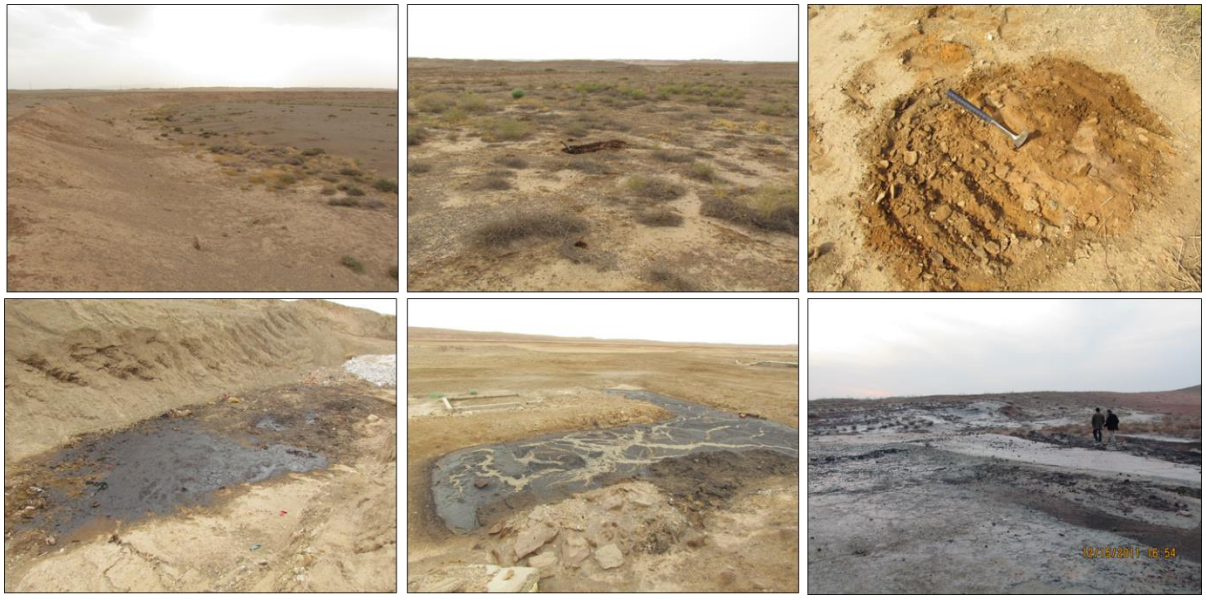

Fig. 5. Oil-contaminated soils detected by ASTER imagery and verified in the field.