



HAL
open science

Exploitation of hyperspectral data for assessing vegetation health under exposure to petroleum hydrocarbons

Guillaume Lassalle

► **To cite this version:**

Guillaume Lassalle. Exploitation of hyperspectral data for assessing vegetation health under exposure to petroleum hydrocarbons. Engineering Sciences [physics]. Toulouse, ISAE, 2019. English. NNT : 2019ESAE0030 . tel-04756747

HAL Id: tel-04756747

<https://hal.science/tel-04756747v1>

Submitted on 28 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut Supérieur de l'Aéronautique et de l'Espace

Présentée et soutenue par :

Guillaume LASSALLE

le jeudi 17 octobre 2019

Titre :

Exploitation of hyperspectral data for assessing vegetation health
under exposure to petroleum hydrocarbons

Exploitation de données hyperspectrales pour l'analyse de l'état de santé de la
végétation exposée aux hydrocarbures

École doctorale et discipline ou spécialité :

ED AA : Surfaces et interfaces continentales, Hydrologie Agrosystèmes, écosystèmes et environnement

Unité de recherche :

Équipe d'accueil ISAE-ONERA OLIMPES et ECOLAB

Directeur(s) de Thèse :

M. Arnaud ELGER (directeur de thèse)
Mme Sophie FABRE (co-directrice de thèse)

Jury :

Mme Isabelle LAFFONT-SCHWOB Professeure Aix-Marseille Université - Présidente
M. Arnaud ELGER Maître de Conférences Université Toulouse III - Directeur de thèse
M. Stéphane JACQUEMOUD Professeur Université Paris Diderot - Rapporteur
M. Carlos Roberto DE SOUZA FILHO Professeur University of Campinas Brésil - Rapporteur
Mme Agnès BÉGUÉ Directrice de Recherche CIRAD
Mme Sophie FABRE Ingénieure de Recherche ONERA - Co-directrice de thèse
M. Anthony CREDOZ Ingénieur TOTAL S.A.

REMERCIEMENTS

Mes premiers remerciements sont adressés à l'Office National d'Etudes et de Recherches Aérospatiales (ONERA) ainsi qu'au groupe TOTAL, à l'origine de cette thèse et de son financement. Je remercie plus précisément le Département d'Optique et Techniques Associées (DOTA) du centre ONERA de Toulouse et le service Environnement de TOTAL au Pôle d'Etudes et de Recherche de Lacq (PERL), pour m'avoir tour à tour accueilli au cours de ces trois années. J'adresse également toute ma reconnaissance aux unités mixtes de recherche Ecolab et DynaFor de l'Université Paul Sabatier (UPS) et de l'Ecole Nationale Supérieure Agronomique de Toulouse (ENSAT), pour leur implication dans cette thèse.

D'aucuns s'imaginent que trois années de thèse peuvent sembler longues et éprouvantes ; il n'en est rien en ce qui me concerne. Je m'étonne encore de la rapidité avec laquelle cette formidable aventure est arrivée à son terme et je reste convaincu que la bienveillance de mes encadrants a grandement contribué au plaisir que j'ai eu à la mener. Il m'est logiquement inconcevable de ne pas les remercier individuellement, aussi nombreux soient-ils (et ils le sont !).

Exception faite de leur expertise scientifique, la bienveillance fut sans nul doute la première qualité de mes directeurs de thèse, Sophie Fabre et Arnaud Elger. Ils ont su me guider et me suivre dans mes choix, m'accordant une confiance grandissante au cours de ces trois années. Je suis infiniment reconnaissant à Sophie pour avoir tenu un rôle allant bien au-delà de celui de co-directrice de thèse, se rendant toujours disponible, de près comme de loin. La rigueur et l'application qu'elle m'a imposées dans mon travail ont fait de cette thèse ce qu'elle est aujourd'hui. Je la remercie pour avoir trouvé la patience de me former dans un domaine qui m'était peu familier ; celui de la télédétection. Je suis tout autant reconnaissant à Arnaud pour m'avoir donné goût à la recherche et à la rédaction d'articles. D'une certaine manière, cette thèse a débuté il y a plus de trois ans, lorsqu'il m'a donné ma chance alors que j'étais étudiant de master. Depuis, sa croyance en moi n'a eu d'égal que sa bienveillance. Il m'était donc impensable d'effectuer une thèse autrement que sous sa direction.

Je ne saurais poursuivre mes remerciements autrement qu'en les adressant à mes encadrants industriels à TOTAL. Je remercie Anthony Credoz pour la confiance qu'il m'a accordée dès le stage précédent cette thèse et pendant les trois années qui ont suivi, et pour avoir su donner la priorité à l'intérêt scientifique lorsque cela était nécessaire. Je suis également reconnaissant à Dominique Dubucq pour son implication croissante dans cette thèse et pour son intérêt porté à l'égard de mes travaux, ainsi que pour les moyens qu'il a déployés pour que je les mène à bien. Enfin, je remercie Rémy Hédacq pour son aide indispensable au cours de ces mois passés en serre et sur le terrain, sous une chaleur accablante rendue largement supportable par nos discussions.

Mes remerciements suivants sont adressés à Georges Bertoni, qui a pris une place de plus en plus importante dans mon encadrement au cours de cette thèse. Son expérience et ses conseils ont largement contribué à améliorer la qualité de mon travail. Je suis également reconnaissant à Pierre Borderies et Jean-Pascal Monvoisin pour m'avoir initié à la télédétection radar.

J'exprime toute ma gratitude à mon comité de suivi de thèse, composé de Camille Larue et David Sheeren. Par deux fois ils ont répondu présent pour évaluer l'avancement de mes travaux et me faire part, avec complémentarité, de leurs remarques avisées. De même, je remercie les rapporteurs et les examinateurs du jury pour avoir accepté d'évaluer ma thèse.

Je souhaite également remercier, sans une quelconque hiérarchisation, les autres personnes impliquées dans cette thèse : Pierre-Yves Foucher, Véronique Achard, Karine Adeline, Laurent

Poutier, Phillipe Déliot, Thomas Rivière, Alain Michel, Xavier Dupuis, Alexandre Alakian, Pascale Dubois-Fernandez, Laurent Junca, Vincent Tardivat, Quentin Guarrigues, Nathalie Nief, Christophe Barreau, Evelyne Buffan-Dubau, Didier Lambrigot, Olivier Berseille et Annie Perrault.

Je ne pourrais oublier d'inclure les doctorants du DOTA à mes remerciements. Je remercie Thierry, Josselin et Etienne, parvenus au terme de leur 3^e année avant moi et dont les conseils m'ont été grandement utiles. Merci également aux doctorants arrivés plus récemment : Simon, Sandra, Thomas, Lucas, Simon (encore !) et Yohann, pour les échanges que nous avons eus autour d'un café lors de nos pauses journalières devenues immanquables. Enfin, je remercie tout particulièrement Florian, arrivé peu de temps après moi, pour nos discussions aussi bien à propos de l'exploitation de nos données, que de nos inquiétudes et de notre avenir professionnel. Je lui souhaite la meilleure fin de thèse qui soit, et de ne jamais perdre cette bonne humeur qu'il a su conserver malgré toutes ces années passées sous l'incessante pluie bordelaise.

Cette thèse m'aura mis à l'épreuve à plus d'un titre. Durant ces trois années, ma persévérance n'aurait jamais été aussi grande sans le soutien de mes proches et la présence de mes amis. Je poursuis donc mes remerciements en les adressant tout d'abord à Arnaud, Kéké et Nico. Bien que nos projets professionnels nous aient éloignés, nous avons toujours trouvé l'occasion de nous retrouver le temps d'une longue soirée passée à nous remémorer les anecdotes hélas parfois inoubliables de nos années passées. Je remercie également Alisson, Ludo, Mathieu, Mehdi et Alex, pour la très bonne ambiance qui régnait au PERL et en-dehors.

Je suis reconnaissant à mes parents, Dany et Eric, pour m'avoir encouragé tout au long de mes études. Par cette thèse, j'espère leur rendre en fierté tout le soutien dont ils m'ont fait part en huit ans. Je les remercie, ainsi que mon frère Kévin, ma belle-sœur Amandine et ma nièce Maëlys, pour leur présence inconditionnelle malgré la distance qui nous sépare.

Je ne saurais terminer mes remerciements autrement que par Clémence. Elle que j'ai embarquée dans ce périple de trois ans, alors même qu'elle venait d'entrer dans ma vie. Elle qui, armée de patience, m'a soutenu, encouragé et réconforté, se souciant en permanence de mon moral. Je lui dois tant de cette thèse, que je ne saurais la remercier assez.

TABLE OF CONTENTS

List of figures.....	12
List of tables.....	18
List of abbreviations	22
General introduction	24
The growing demand in global oil supply.....	24
Overview of existing technologies for oil and gas exploration	25
Oil contamination: sources, consequences and monitoring.....	26
Remote sensing: an emerging tool for oil exploration and contamination monitoring.....	27
Current limits of remote sensing for oil detection in vegetated regions.....	29
Objectives of the thesis.....	29
Chapter one: State-of-the-art of passive hyperspectral remote sensing for oil exploration and contamination monitoring in vegetated regions.....	31
Introduction	31
1. Vegetation optical properties in the reflective domain (400 – 2500 nm)	32
1.1. Influence of leaf pigments in the visible region (400 – 750 nm)	32
1.2. Influence of leaf anatomy in the near-infrared region (750 – 1300 nm).....	36
1.3. Influence of leaf water and dry matter contents in the near-infrared (750 – 1300 nm) and short-wave infrared (1300 – 2500 nm) regions	39
2. Effects of crude oil and petroleum products on vegetation health.....	42
2.1. Composition of crude oil and petroleum products.....	42
2.2. Effects on soil physico-chemical and biological properties and on plant roots	43
2.2.1. Effects on soil water regime	43
2.2.2. Effects on nutrient availability and root uptake capacities.....	43
2.3. Consequences on vegetation health.....	44
2.3.1. Effects on plant biochemical and biophysical parameters	45
2.3.2. Sources of variability	46
3. Detection of crude oil and petroleum products using vegetation optical properties.....	49
3.1. Effects of crude oil and petroleum products on vegetation optical properties measured under controlled conditions.....	49
3.1.1. Effects on reflectance in the visible region (400 – 750 nm)	49
3.1.2. Effects on reflectance in the near-infrared region (750 – 1300 nm).....	50

3.1.3. Effects on reflectance in the short-wave infrared region (1300 – 2500 nm).....	51
3.2. Methods developed for detecting crude oil and petroleum products under controlled conditions	52
3.3. Other methods used in vegetation stress assessment.....	53
3.4. Application in oil exploration and contamination monitoring using airborne and satellite multi- and hyperspectral imagery	55
4. Synthesis based on previous studies.....	57
Conclusion.....	59
Scientific problem and adopted approach	60
Chapter two: Development of methods for detecting and quantifying oil contamination based on vegetation optical properties, under controlled conditions	62
Introduction	62
5. First article: “Assessing soil contamination due to oil and gas production using vegetation hyperspectral reflectance”	64
Abstract	64
5.1. Introduction.....	65
5.2. Materials and Methods.....	66
5.2.1. Plant materials	66
5.2.2. Soil sampling and preparation.....	66
5.2.3. Plant treatments.....	67
5.2.4. Hyperspectral reflectance acquisitions	67
5.2.5. Data pre-processing and analysis	67
5.3. Results and Discussion.....	68
5.3.1. Visible symptoms.....	68
5.3.2. Spectral signatures.....	68
5.3.3. LDA classification.....	69
5.3.3.1. Leaf scale	69
5.3.3.2. Plant scale.....	70
5.3.4. Discriminant wavelengths	72
6. Second article: “Detection and discrimination of various oil-contaminated soils using vegetation reflectance”.....	76
Abstract	76
6.1. Introduction.....	77
6.2. Materials and Methods.....	78
6.2.1. Study site and species	78
6.2.2. Plant materials and treatments.....	78
6.2.3. Biochemical analysis	79

6.2.4. Spectral reflectance measurements and preprocessing.....	80
6.2.5. Vegetation indices.....	80
6.2.6. Field validation.....	83
6.3. Results and Discussion.....	83
6.3.1. Visible symptoms.....	83
6.3.2. Biochemical responses to the treatments.....	84
6.3.3. Reflectance modifications.....	86
6.3.4. Vegetation indices.....	88
6.3.4.1. Relationship between biochemical and spectral responses.....	88
6.3.4.2. Concordance of VI among acquisition scales.....	89
6.3.4.3. Discrimination of treatments using VI.....	90
6.3.5. Field validation.....	92
6.4. Conclusion and Perspectives.....	92
7. Third article: “Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression”.....	94
Abstract.....	94
7.1. Introduction.....	95
7.2. Materials and Methods.....	96
7.2.1. Study site and greenhouse experiment.....	96
7.2.2. Biochemical analyses and reflectance acquisitions.....	96
7.2.3. Vegetation indices.....	97
7.3. Results and discussion.....	98
7.3.1. Biochemical and spectral responses to oil contamination.....	98
7.3.2. Vegetation indices.....	100
7.3.2.1. Elastic net regressions.....	100
7.3.2.2. Random Forest regressions.....	102
7.4. Conclusion.....	104
Summary and conclusion.....	106

Chapter three: Application and evaluation of the methods under natural conditions, from field scale to airborne hyperspectral imagery 109

Introduction.....	109
8. Fourth article: “Application of PROSPECT for estimating Total Petroleum Hydrocarbons in contaminated soils from leaf optical properties”.....	110
Abstract.....	110
8.1. Introduction.....	111
8.2. Materials and Methods.....	112
8.2.1. Study site and species.....	112
8.2.2. Field sampling and measurements.....	112

8.2.3. PROSPECT inversion and validation	113
8.2.4. Variability of leaf pigment contents and TPH estimation.....	114
8.3. Results and Discussion.....	115
8.3.1. Measured spectral signatures.....	115
8.3.2. Evaluation of PROSPECT inversions outputs	116
8.3.2.1. Simulated spectral signatures	116
8.3.2.2. Retrieved biochemical parameters	117
8.3.3. Seasonal and interspecific variability of leaf pigments	117
8.3.4. Prediction of TPH from leaf pigments.....	119
8.4. Conclusion	122
9. Fifth article: “Toward quantifying oil contamination in vegetated areas using very high spatial and spectral resolution imagery”.....	124
Abstract	124
9.1. Introduction.....	125
9.2. Materials and Methods.....	126
9.2.1. Study area and soil sampling	126
9.2.2. Hyperspectral data acquisition.....	127
9.2.2.1. Airborne image.....	127
9.2.2.2. Field reflectance	128
9.2.3. First step: Detection of oil contamination	128
9.2.4. Second step: Quantification of soil TPH content	129
9.2.4.1. First method based on PROSPECT and PROSAIL	129
9.2.4.2. Second method based on Elastic net regression.....	130
9.3. Results and Discussion.....	131
9.3.1. Detection of oil contamination	131
9.3.2. Quantification of soil TPH content	131
9.3.2.1. PROSAIL-based method	131
9.3.2.2. Elastic net-based method.....	133
9.3.3. Validation of the methods.....	135
9.3.3.1. Validation of oil detection	135
9.3.3.2. Validation of TPH quantification	136
9.4. Perspectives	136
9.5. Conclusions.....	138
Summary and conclusion.....	140
General conclusion.....	144
Perspectives.....	146

Current limits for operational use on satellite imagery.....	146
Toward assessing mud pit contamination at large scale.....	146
Toward extending the scope of the methods to oil exploration and contamination monitoring	147
Toward improving oil detection and quantification using UAV platforms and other remote sensing technologies.....	148
Appendix: Evaluation of radar (SAR) remote sensing for assessing vegetation health under exposure to crude oil and petroleum products.....	151
Introduction	151
1. Overview of radar remote sensing for assessing vegetation health	153
1.1. Vegetation backscatter properties in the 1 – 40 GHz microwave domain.....	153
1.2. Application of radar remote sensing for assessing vegetation health.....	154
2. Radar remote sensing for detecting and quantifying crude oil and petroleum products ..	155
2.1. Application to oil detection under controlled conditions.....	155
2.1.1. Radar backscattering measurements	155
2.1.2. Results and discussion	157
2.2. Application to oil detection and quantification using airborne SAR imagery	158
2.2.1. SAR images acquisitions	158
2.2.2. Results and discussion	159
Summary and conclusion.....	161
Supporting information.....	163
Publications and communications.....	183
Publications (peer-reviewed articles).....	183
Communications.....	183
References	184

LIST OF FIGURES

Figure 1. Annual oil production and Proved and Probable (2P) discoveries since 1900 ⁴	24
Figure 2. (a) Schematic representation of oil macro- and microseepages (adapted from ²³). (HC: Hydrocarbons.) Macroseepages are characterized by apparent crude oil at the surface. (b) Microseepages are manifested by alterations of the surface lithology (c) and specific vegetation patterns (d) ^{20,24}	25
Figure 3. Principal sources of environmental contamination caused by oil activities. (a) Oil sludge pit ⁴⁵ , (b-c) vegetation and soil contaminated by crude oil leakage near a refining facility ⁴⁶ , (d) pipeline leakage ⁴⁷ , (e) crop contamination resulting from oil well blow out ⁴⁸ , (f) oil leakage from damaged storage tank following a storm ⁴⁹ and (g) contaminated wastewater near a production site ⁵⁰	26
Figure 4. Principle of passive hyperspectral imagery (adapted from ⁵⁸). This technology provides the reflectance of surfaces over a continuous spectrum in the optical reflective domain (<i>i.e.</i> the spectral signature).....	27
Figure 5. Typical spectral signature of healthy green leaf and most influential parameters in the different spectral regions.....	32
Figure 6. (a) Chemical structure and (b) absorption spectra of chlorophyll a and b ⁹²	33
Figure 7. Relationship between leaf chlorophyll content and the Red-Edge Position (REP) (modified from ⁹⁶).	33
Figure 8. (a-b) <i>Acer saccharum</i> Marsh. (sugar maple) during autumn senescence. Sections of (c) green, (d-e) yellow, (f) orange and (g) red sugar maple leaves, and (h) chlorophylls, (i) carotenoids, (j) carotenoid-chlorophyll ratio (Car/Chl) and (k) anthocyanins found in corresponding leaves. Non-significant differences ($p \geq 0.05$) are denoted by similar lowercase letters (modified from ⁸⁹).	34
Figure 9. (a-b) Absorption spectra of principal carotenoids ⁹² . (c) The xanthophyll cycle ⁸⁰	35
Figure 10. Relationship between the deepoxidation of the xanthophyll cycle and the Photochemical Reflectance Index (PRI) ¹⁰⁴	35
Figure 11. (a) Chemical structure of cyanidin, a red anthocyanin, and (b) absorption spectra of anthocyanins at pH 1.0 and 4.5 ¹⁰⁵	36
Figure 12. (a) Typical anatomy of a dicotyledonous leaf (cryo-scanning electron microscopy) and (b) relationship between spongy parenchyma void area (intercellular spaces) and leaf reflectance in the near-infrared (modified from ¹¹⁷).	37
Figure 13. Diversity of leaf anatomy in tropical region ¹¹⁴ . (Scale bars = 100 μm).	38
Figure 14. Influence of (a) Leaf Area Index (LAI) and (b) Mean Leaf Angle (MLA) on canopy reflectance ⁶⁹ . Spectral signatures were simulated using the PROSAIL model described in section 9.2.4.1.	39
Figure 15. (a) Absorption coefficient of water ¹³⁵ . and (b) spectral signatures of healthy and water-deficient plants (personal data).	40
Figure 16. Relationship between leaf Equivalent Water Thickness (EWT) and the Normalized Difference Infrared Index (NDII), which exploits reflectance at 850 and 1650 nm ¹⁴¹	40

Figure 17. Relative contribution of water, proteins and celluloses + lignins to light absorption of (a) fresh and (b) dry leaves in the NIR and SWIR regions ¹⁴⁸	41
Figure 18. Crude oil and petroleum products according to petroleum carbon ranges (modified from ¹⁵⁶).	42
Figure 19. Development of <i>Festuca arundinacea</i> grown for three months on (a) uncontaminated or (b) crude oil-contaminated (10.8 g.kg ⁻¹) soil. Root system and scanning-electron microscopy images of root sections from (c, e) uncontaminated and (d, f) crude-oil contaminated plants after 14 months. Comparison of selected root anatomical parameters between the two treatments (mean ± SE). Significant differences (p < 0.05) between soils are denoted by different superscript letters (modified from ¹⁶⁰).	44
Figure 20. <i>Canavalia ensiformis</i> (L.) DC grown on (a-b, e) diesel-contaminated (22.219 mg.kg ⁻¹) or (c-d) uncontaminated soil. White arrows indicate necrotic roots. (f-g) Symptoms of leaf discoloration observed on diesel-exposed plants. (h) Comparison of leaf anatomical parameters and pigment contents between the two treatments (mean ± SD). Significant differences (p < 0.05) between soils are denoted by different lowercase letters (modified from ¹⁸³).	46
Figure 21. Spectral signatures of leaves of <i>Zea mays</i> L. grown for 14 days on engine oil-contaminated (48 – 214 g.kg ⁻¹) or uncontaminated soils (modified from ¹⁸⁶).	50
Figure 22. Reflectance differences between <i>Brachiaria brizantha</i> H.S. plants grown on either (a-b) diesel- or (c-d) gasoline-contaminated and uncontaminated soil, at (a, c) leaf and (b, d) canopy scales. Reflectance differences were obtained by subtracting the spectral signature of uncontaminated plants to those from diesel- or gasoline-exposed plants. Curves represent different contamination levels (modified from ⁷⁷).	51
Figure 23. Temporal evolution of the Red-Edge Position (REP) of <i>Zea mays</i> L. grown on engine oil-contaminated (48 – 214 g.kg ⁻¹) or uncontaminated soils (modified from ¹⁸⁶).	53
Figure 24. (a) Graphical representation of the Linear Discriminant Analysis (LDA) principle ²²³ . This method relies on the linear combination of variables into a reduced projection space for separating classes. (b) Application of LDA to the discrimination of fungal disease severity (“Classes”) in <i>Musa</i> AAA cv. William using hyperspectral data. The first two discriminant axis are represented ²²⁴	54
Figure 25. (a) Crude oil seepage in vegetated area. (b) The seepage is characterized by bare ground mixed with crude oil at the center, surrounded by a particular vegetation distribution pattern, which allows being detected from hyperspectral airborne images using (c) vegetation indices and a spatial filter. However, false alarms (red pixels outside yellow circles) cannot be avoided while simultaneously detecting all seepages ²⁰	57
Figure 26. Flowchart of the adopted approach. The first step is presented in the chapter two and the second and third steps are presented in the chapter three.	60
Figure 27. Mean reflectance of <i>Rubus</i> grown for 61 days under different conditions, at (a) leaf and (b) plant scale. Data are shown prior to pre-processing.	69
Figure 28. Classification accuracy obtained from LDA performed on <i>Rubus</i> leaf (a) and plant (b) reflectance data, all acquisition dates gathered (1 (a) or 25 (b) to 61 days). (D1, first derivative; D2, second derivative; SNV, Standard Normal Variate; AUCN, Area Under Curve Normalization; CR, <i>Continuum</i> Removal).	71
Figure 29. Temporal evolution of classification accuracy obtained from LDA performed on <i>Rubus</i> reflectance data, at leaf (a-c) and plant (d-f) scales. Results are displayed from day 1 (‘J1’) or 25 (‘J25’) to day 61 (‘J61’) for leaf and plant scales respectively.	72

Figure 30. Discriminant wavelengths obtained from LDA performed on <i>Rubus</i> at leaf and plant scales. Frequencies are presented with untransformed and transformed data gathered, with bands grouped into 50 nm-wide intervals.	73
Figure 31. Flowchart of the approach using vegetation indices presented in this study. Sections and tables associated to the different steps are specified in brackets. The dotted line indicates the steps carried out at leaf scale only.....	83
Figure 32. Visible symptoms observed on <i>Rubus fruticosus</i> (bramble) leaves after 18 and 32 days of exposure to the treatments (see section 6.2.2 for further information about the treatments). These symptoms were confirmed later by biochemical analyses described in section 6.3.2.	84
Figure 33. (a) PCA performed on pigment and water content data from three dates (days 11, 18 and 32, n = 51). The first two Principal Components (PC) are represented. Correlations between each variable and PC are represented by annotated arrows. The length of the arrows was multiplied by five on each PC for graphical convenience. (b) Temporal evolution of treatments along the first principal component (PC1), from day 11 to 32.	85
Figure 34. Mean spectral signatures of <i>Rubus fruticosus</i> (bramble) measured after 32 days of exposure to the treatments at (a) leaf, (b) plant and (c) multi-plant scales. Reflectance data located in the 1350-1450 and 1800-1950 intervals were removed, because of significant atmospheric effects at plant and multi-plant scales.....	87
Figure 35. MTCI index computed from leaf, plant and multi-plant reflectance data acquired on day 32. This vegetation index was linked to total chlorophyll and β -carotene contents ($R^2 = 0.94$) and exhibited strong similarities among acquisition scales following Kendall's W selection procedure ($W = 0.90$, $p < 0.05$).....	90
Figure 36. Shoot dry weight and leaf chlorophyll and carotenoid contents (mean \pm SD) of <i>C. alopecuroides</i> after 42 days of exposure to various levels of Total Petroleum Hydrocarbons (TPH) (n = 11 and 5 samples per treatment, respectively). Differences among treatments are denoted by different lowercase letters (ANOVA and Tukey post-hoc tests, $p < 0.05$). (VAZ: Violaxanthin + Antheraxanthin + Zeaxanthin.).....	98
Figure 37. Mean reflectance of <i>C. alopecuroides</i> in the reflective domain after 42 days of exposure to various levels of Total Petroleum Hydrocarbons (TPH), at leaf and plant scales. Reflectance data from the 1350–1450 and 1800–1950 intervals were removed, because of low atmospheric transmission at plant scale.	99
Figure 38. Relationship between vegetation indices and leaf pigment contents. Figures include data from leaves sampled on all the treatments, after 21 and 42 days of experiment (n = 50).	102
Figure 39. Comparison between the measured and predicted concentrations of Total Petroleum Hydrocarbons (TPH) using the 33 vegetation indices and the random forest regression, at leaf and plant scales (top figures), and residuals of the predictions (bottom figures).....	103
Figure 40. Relative contribution of the vegetation indices to the estimation of Total Petroleum Hydrocarbons (TPH) using random forest regression, at leaf and plant scales. Only indices with non-zero contribution are displayed.....	104
Figure 41. Flowchart of the method presented in this study (N : structure parameter, C_{ab} : Chlorophyll content, C_{xc} : Carotenoid content, C_{bp} : Brown pigments, C_w : Water, C_m : Dry matter). The spectral signatures measured were used to retrieve leaf structural and biochemical parameters using PROSPECT-5 inversions. The estimated parameters were compared to chemical analyses, and the simulated spectral signatures were compared to the measured ones. The estimated pigment contents (C_{ab} and C_{xc}) obtained for each season were compared, using ANOVA and Tukey post-hoc tests, and	

were then used to predict TPH concentrations with the help of univariate models. The predicted TPH concentrations were then compared to those obtained from soil analyses.....115

Figure 42. Seasonal evolutions of the mean spectral signatures for the five species on the brownfield. 116

Figure 43. Comparisons of the predicted and measured biochemical parameters of *R. fruticosus* leaves. Water and dry matter contents were expressed in mg.cm⁻² for graphical convenience. 117

Figure 44. Temporal evolutions of the relative chlorophyll contents of leaves retrieved from PROSPECT inversions on the brownfield, for the five species. 118

Figure 45. Relationship between predicted leaf chlorophyll content and Total Petroleum Hydrocarbons (TPH) in the soil on the train datasets (top figures) and comparisons between predicted and measured TPH on the test datasets (bottom figures).....121

Figure 46. (a, d) Subsets of the airborne hyperspectral image acquired over (a) the contaminated and (d) the control sites. The calibration sites are illustrated by polygons. Circles correspond to the validation sites (red: Contaminated, white: Control). (b, e) Zoom on (b) the brownfield (10 plots) and (e) the control sites used for calibrating the methods, respectively. (c, f) *R. fruticosus* on (c) the brownfield and (f) the control site. (Note: Both photos c and f were taken on March.)..... 127

Figure 47. (a-b) Leaf Chlorophyll Content (LCC) (mean ± SD) retrieved from PROSPECT and PROSAIL model inversions on the 10 plots of the brownfield. (a) Comparison between leaf and canopy measurements from the field and (b) between leaf measurements and the airborne image.....132

Figure 48. (a-c) Relationship between LCC and soil TPH content observed on the train set on the 10 plots of the brownfield, at (a) leaf and (b) canopy scales and (c) on the airborne image. (d-f) Comparison between measured and predicted TPH concentrations on the test set, at (d) leaf and (e) canopy scales and (f) on the airborne image.....133

Figure 49. Comparison between measured and predicted TPH concentrations on the test set on the 10 plots of the brownfield using the Elastic net regression on the original and transformed spectral signatures (in columns). First, second and third rows show the results obtained at leaf and canopy scales and on the airborne image, respectively.....134

Figure 50. Coefficients obtained by the ENET method on the 10 plots of the brownfield using the original and transformed spectral signatures (in columns). First, second and third rows show the results obtained at leaf and canopy scales and on the airborne image, respectively. The spectral bands contributing to TPH predictions are denoted by non-zero coefficients.135

Figure 51. (top figures) Comparison between Leaf Chlorophyll Content (LCC) (mean ± SD) retrieved in the field (leaf-scale measurements) and on the airborne image, and (bottom figures) comparison between measured and predicted TPH concentrations on the airborne image at the spectral resolutions of HypXim and Worldview-3 sensors..... 142

Figure 52. (a) Main frequency bands used in radar remote sensing. (b) Synthetic Aperture Radar (SAR) imaging principles (adapted from ³⁹⁵). (S_{t1} and S_{t2} : satellite initial and final position, respectively; V_s : movement speed; H_s : satellite altitude; θ : viewing angle; R : distance between the satellite and the ground.)..... 152

Figure 53. Components of the backscattering signal emitted in the microwave domain (reproduced after ³⁹⁶). Direct scattering of vegetation is sometimes also considered as volume scattering.153

Figure 54. (top figures) Temporal evolution of Volumetric Water Content (VWC), Leaf Area Index (LAI) and Normalized Difference Vegetation Index (NDVI) of rice and soybean. (bottom figures) Temporal evolution of L-band backscattering coefficients (σ^0) in different polarizations (H: Horizontal;

V: Vertical) and Radar Vegetation Index ($RVI = 8\sigma_{HV}^0 / (\sigma_{HH}^0 + \sigma_{VV}^0 + 2\sigma_{HV}^0)$) for these species ³⁹⁷	155
Figure 55. Experimental setup used for radar backscattering acquisitions on vegetation during the experiment.....	157
Figure 56. Temporal evolution of <i>R. fruticosus</i> backscattering coefficients for the different treatments (see Second Article in Chapter II for details about the treatments).....	158
Figure 57. Relationship between <i>R. fruticosus</i> total dry biomass and backscattering coefficients after 32 days of exposure to the different treatments (see Second Article in Chapter II for details about the treatments).....	158
Figure 58. Subset of the airborne X-band SAR image acquired over the study area. The contaminated site (“brownfield”) is illustrated by the red polygon.....	159
Figure 59. (top figures) X- and (bottom figures) L-band’s backscattering coefficients observed for the control site (green) and the 10 plots of the brownfield (brown). (TPH: Total Petroleum Hydrocarbons.).....	160
Figure 60. Illustration of the hyperspectral acquisitions carried out at plant scale on the three species.	163
Figure 61. Example of spectral signatures (a) before and (b) after Area Under Curve Normalization (AUCN).....	164
Figure 62. Classification accuracy obtained from LDA performed on <i>Cenchrus</i> reflectance data, all acquisition dates combined (49-61 days).....	164
Figure 63. Classification accuracy obtained from LDA performed on <i>Panicum</i> reflectance data, all acquisition dates combined (49-61 days).....	165
Figure 64. Temporal evolution of classification accuracy obtained from LDA performed on <i>Cenchrus</i> reflectance data. Results are displayed from day 49 (‘J49’) to 61 (‘J61’).....	165
Figure 65. Temporal evolution of classification accuracy obtained from LDA performed on <i>Panicum</i> reflectance data. Results are displayed from day 49 (‘J49’) to 61 (‘J61’).....	166
Figure 66. Discriminant wavelengths obtained from LDA performed on <i>Panicum</i> and <i>Cenchrus</i> at plant scale. Frequencies are presented all untransformed and transformed data combined, with bands grouped in 50 nm-wide intervals.....	166
Figure 67. Examples of bramble plants joined for multi-plant scale measurements (left: Control treatment, right: Clay treatment). Individual black pots of 15 x 15 x 18 cm (L x W x H) were used for the experiment. These pots had no influence on measurements at multi-plant scale, since their reflectance was close to zero. Measurements were carried out with a 20-cm acquisition footprint, which included most of the four plants.....	167
Figure 68. RMSE, BIAS and SEPC computed between all simulated and measured spectral signatures.	172
Figure 69. Examples of simulated and measured spectral signatures.....	172
Figure 70. Temporal evolutions of the relative carotenoid contents of leaves retrieved from PROSPECT inversions on the brownfield, for the five species.....	176
Figure 71. Relationship between predicted leaf carotenoid content and Total Petroleum Hydrocarbons (TPH) in the soil on the train datasets (top figures) and comparisons between predicted and measured TPH on the test datasets (bottom figures).....	176

Figure 72. Original and transformed spectral signatures (in columns) acquired in the brownfield. First, second and third rows show the results obtained at leaf and canopy scales and on the airborne image, respectively. The bold curve represents the median spectral signature and the colored area, the [min:max] interval. 180

Figure 73. Flowchart of the approach proposed in this study. The method of oil detection was developed and tested at leaf and canopy scales in ³²⁴, so it was only evaluated on the airborne image in this study. The two methods of Total Petroleum Hydrocarbons (TPH) quantification were tested in the field at leaf and canopy scales and on the airborne image. All the methods were first calibrated on study sites and validated on additional ones (see Table 16 for details about the sites)..... 181

LIST OF TABLES

Table 1. Specifications of some of the current and future airborne and satellite-embedded sensors. For airborne sensors, the spatial resolution depends on the flight altitude. The name and specifications of future sensors may be modified until their operating.....	28
Table 2. Literature review of the effects induced by crude oil and petroleum products on vegetation health. Effects are sorted by biophysical and biochemical parameters related to vegetation optical properties in the different spectral regions. (↑ and ↓ denotes increase and decrease in the measured parameter, respectively; dose-dep. indicates dose-dependent effects; n.a.: not available.).....	48
Table 3. Literature review of the effects induced by crude oil and petroleum products on vegetation reflectance in the different spectral regions, at leaf and canopy scales. This review only includes studies carried out under greenhouse-controlled or field-semi-controlled conditions. (↑ and ↓ denotes reflectance increase and decrease, respectively; FC: Field capacity; dose-dep: dose-dependent effects; n.a.: not available; n.s.: non-significant effect.).....	52
Table 4. Review of studies on crude oil and petroleum products detection from multi- and hyperspectral airborne and satellite images. (Refl.: Reflectance; VI: Vegetation Indices; CR: <i>Continuum</i> Removal; RF: Random Forest; REP: Red-Edge Position; comp.: Comparison.).....	56
Table 5. Heavy metal and hydrocarbon concentrations found in the control and contaminated soils used in this study. Polycyclic Aromatic Hydrocarbons (PAH) include the 16 priority pollutants of the US Environmental Protection Agency.	66
Table 6. Composition of soil treatments used in this study. Concentrations are given in mg.kg ⁻¹ . Polycyclic Aromatic Hydrocarbons (PAH) represent the 16 priority pollutants of the U.S. Environmental Protection Agency (EPA ²⁷⁵). (< DL: below detection limit.). Only low concentrations of HM were found in the uncontaminated soil (Control and Water-str) and corresponded to the local geochemical background.	79
Table 7. Vegetation indices used in this study (R: Reflectance, D: First derivative).....	82
Table 8. Results of elastic net regressions and Kendall's W selection procedure obtained on vegetation indices. Biochemical parameters contributing to index variations are presented in order of importance, along with the resulting R ² . (Chl: Total chlorophylls, Lut: Lutein, B-car: β-carotene, Ant: Antheraxanthin, Zea: Zeaxanthin, Vio: Violaxanthin, LWC: Leaf Water Content.) Indices showing significant Kendall's W (p < 0.05) are denoted by a *. Those retained for classifications (R ² > 0.7, W > 0.75*) are marked bold.	89
Table 9. Mean Overall Accuracy (MOA) and mean Cohen's Kappa (± SD) obtained from RLR classifications performed on leaf, plant and multi-plant test sets with the 14 retained vegetation indices. The training step was exclusively performed at leaf scale using different samples on each of the 30 iterations. For each scale, significant differences of MOA and Kappa between dates are denoted by different superscript letters (t-test, p < 0.05).....	91
Table 10. Examples of relative confusion matrices (%) obtained from RLR classifications performed on leaf and plant test sets at day 18 and 32, following training at leaf scale. UA and PA stand for User Accuracy and Producer Accuracy (%), respectively.	92
Table 11. Results of the elastic net regressions performed on the 33 vegetation indices. For each index, the R ² of the model is presented, along with the contributing set of biochemical parameters (in order of importance). (Chl a: Chlorophyll a, Chl b: Chlorophyll b, B-car: β-carotene, Lut: Lutein, Ant: Antheraxanthin, Vio: Violaxanthin, Zea: Zeaxanthin, LWC: Leaf Water Content.)	101

Table 12. Summary of the experiments carried out in the first step of the approach, and the developed methods. (TPH: Total Petroleum Hydrocarbons; HM: Heavy Metals; LDA: Linear Discriminant Analysis; RLR: Regularized Logistic Regression; RF: Random Forest; RMSE: Root Mean Squared Error; RPD: Residual Predictive Deviation.).....	107
Table 13. Description of the sampling campaigns carried out on the brownfield. Only <i>R. fruticosus</i> and <i>B. davidii</i> were found on all 23 plots, except in March 2018.	113
Table 14. Parameter bounds used to perform PROSPECT inversions. C_{ab} and C_{xc} bounds were extended from those of literature to factor in seasonal variations of leaf pigment contents.....	114
Table 15. Assessment of Total Petroleum Hydrocarbons (TPH) in the soil predicted from leaf chlorophyll content on the train and test datasets. RMSE, BIAS and SEPC are given in $g.kg^{-1}$ (abs: absent from the site). The best prediction results are marked in bold.....	120
Table 16. Description of the sites used for testing and validating the methods. (< DL: Below detection limit.)	128
Table 17. Parameter bounds used to compute PROSPECT and PROSAIL (PROSPECT + SAIL) inversions.....	130
Table 18. Confusion matrix obtained from the RLR classification on the test set using the 14 vegetation indices. The <i>Brownfield</i> class denotes the pixels from the 10 zones of the brownfield, and the <i>Control</i> one, those from the large control site.	131
Table 19. Confusion matrix obtained from the RLR classification on the validation sites using the 14 vegetation indices.	136
Table 20. Summary of the studies carried out in the second and third steps of the approach, and the developed methods. (TPH: Total Petroleum Hydrocarbons; HM: Heavy Metals; ENET: Elastic net; RMSE: Root Mean Squared Error; RPD: Residual Predictive Deviation.).....	141
Table 21. BTEX, Polycyclic Aromatic Hydrocarbons (PAHs) and Total Petroleum Hydrocarbon concentrations ($g.kg^{-1}$) found in the soils used for the treatments applied to <i>Cenchrus alopecuroides</i> (L.). Polycyclic Aromatic Hydrocarbons (PAHs) include the 16 priority pollutants identified by the U.S. Environmental Protection Agency (EPA). (<DL: Below the detection limit.)	168
Table 22. Vegetation indices used in this study (R: Reflectance, D: First derivative).....	169
Table 23. Heavy metal and hydrocarbon concentrations ($g.kg^{-1}$) found in the brownfield and control soils. Polycyclic Aromatic Hydrocarbons (PAHs) include the 16 priority pollutants identified by the U.S. Environmental Protection Agency (EPA). Heavy metals found in the control site corresponded to the geochemical background. (<DL: below detection limit.).....	170
Table 24. Statistics of biochemical parameters obtained from leaf analysis ($n = 152$ leaves). These data were used for validating PROSPECT inversion outputs.	171
Table 25. Statistics of the datasets used for calibration (train) and validation (test) of the models. The same number of samples (n) was used as train and test sets on each date (except in March 2018 where only few <i>B. davidii</i> and <i>R. fruticosus</i> were found on the site).....	171
Table 26. Assessment of biochemical parameter retrieval using PROSPECT inversions.....	173
Table 27. Leaf chlorophyll content ($\mu g.cm^{-2}$) retrieved from PROSPECT inversions on the brownfield (abs: absent from the site). For each species, significant differences in mean chlorophyll content among sampling dates are denoted by different superscript letters (ANOVA and Tukey post-hoc test, $p < 0.05$).	174

Table 28. Leaf carotenoid content ($\mu\text{g}\cdot\text{cm}^{-2}$) retrieved from PROSPECT inversions on the brownfield (abs: absent from the site). For each species, significant differences in mean carotenoid content among sampling dates are denoted by different superscript letters (ANOVA and Tukey post-hoc test, $p < 0.05$). 175

Table 29. Assessment of Total Petroleum Hydrocarbons (TPH) in the soil predicted from leaf carotenoid content on the train and test datasets. RMSE, BIAS and SEPC are given in $\text{g}\cdot\text{kg}^{-1}$ (abs: absent from the site)..... 177

Table 30. Comparison of soil Total Petroleum Hydrocarbons (TPH) predicted on the test sets using the proposed approach (Chlorophyll content PROSPECT), vegetation indices (MTCI, mSR705) and Partial Least Square Regression (PLSR) on transformed spectral signatures. RMSE is given in $\text{g}\cdot\text{kg}^{-1}$. The tested vegetation indices and reflectance transformations can be found in ²⁹² and ²⁷¹. Those leading to the best predictions are presented. Important results are marked bold..... 178

Table 31. Vegetation indices used in the method of detection. For each index, the related biochemical parameters are specified. These parameters have been identified in a previous study carried out on *R. fruticosus*, under controlled conditions³²⁴. (Chl: Total chlorophylls, Lut: Lutein, B-car: β -carotene, Ant: Antheraxanthin, Zea: Zeaxanthin, Vio: Violaxanthin, LWC: Leaf Water Content.)..... 179

LIST OF ABBREVIATIONS^a

AUCN	Area Under Curve Normalization
CR	<i>Continuum</i> Removal
D1	First Derivative Spectrum
D2	Second Derivative Spectrum
ELM	Empirical Line Method
ENET	Elastic Net
EWT	Equivalent Water Thickness
HC	Hydrocarbons
HM	Heavy Metals
HPLC	High Pressure Liquid Chromatography
LAD	Leaf Angle Distribution
LAI	Leaf Area Index
LCC	Leaf Chlorophyll Content
LDA	Linear Discriminant Analysis
LWC	Leaf Water Content
MLA	Mean Leaf Angle
NIR	Near Infrared
OA / MOA	Overall Accuracy / Mean Overall Accuracy
PA	Producer's Accuracy
PAH	Polycyclic Aromatic Hydrocarbon
PCA	Principal Component Analysis
PLSR	Partial Least Square Regression
REP	Red-Edge Position
RF	Random Forest
RLR	Regularized Logistic Regression
RMSE	Root Mean Squared Error
RPD	Residual Predictive Deviation
RTM	Radiative Transfer Model
SAR	Synthetic Aperture Radar
SEPC	Standard Error of Prediction Corrected from the bias
SNR	Signal-to-Noise Ratio
SNV	Standard Normal Variate
SWIR	Short-Wave Infrared
TPH	Total Petroleum Hydrocarbons
UA	User's Accuracy
UAV	Unmanned Aerial Vehicle
UV	Ultraviolet
VI	Vegetation indices
VIF	Variable Importance Factor
VIS	Visible

^aSome abbreviations may vary according to the articles

GENERAL INTRODUCTION

The growing demand in global oil supply

Oil and gas industry currently holds a key role in the global energy mix¹⁻³. Since the beginning of the 20th century, crude oil supply has continuously increased to satisfy a growing demand, reaching over 35 billion barrels (Gb) produced in 2017⁴⁻⁶ (Figure 1). Although a global peak of production – followed by a decline – is expected in the future, its timing remains largely unprecise as it depends on several factors, such as reserve estimates, and on the scenario that will frame the energy mix⁷⁻¹⁰. According to the International Energy Agency (IEA), oil production will slow but still become 8 million barrels per day greater in 2040 than today under the New Policy Scenario, which considers current government goals and policies. This tendency is mainly explained by the growing demand from transportation and industry, sharing more than 95 % of the global oil consumption^{11,12}.

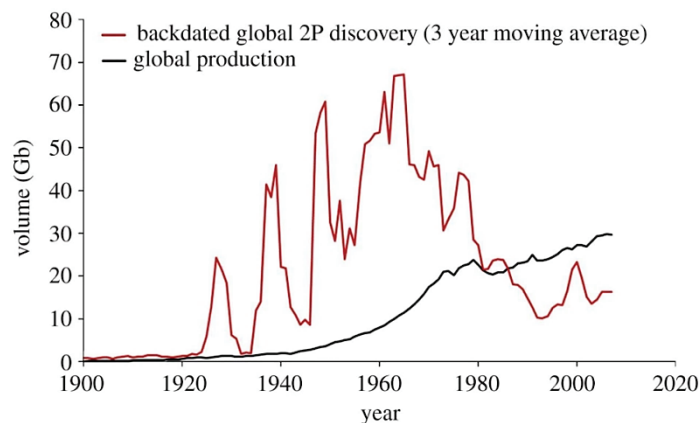


Figure 1. Annual oil production and Proved and Probable (2P) discoveries since 1900⁴.

Over 97 % of the global oil production is provided by conventional oil¹³, which refers to “*crude oil and natural gas that is produced by a well drilled into a geologic formation in which the reservoir and fluid characteristics permit the oil and natural gas to readily flow to the wellbore*”¹⁴. Once extracted from oil fields, crude oil is then refined to petroleum products, mainly gasoline and distillate fuel (including diesel)^{12,14,15}. Since oil is a finite resource, its exploitation is inevitably drawing down the reserves over time. The total proved and probable oil reserves were estimated to 1480 Gb in 2017⁵. This estimation is however based on known oil fields, and may increase with the discovery of new ones. Up to 70 000 oil fields are already exploited worldwide, but half of the production is ensured by only 110 of them¹⁰. The most important conventional oil fields were found in a relatively short period during the 20th century (1945 – 1980, Figure 1) and since then, global discoveries gradually decreased, as remaining oil fields tend to be smaller and thus harder to locate^{10,13}. Consequently, the prospection of new oil fields has become very challenging from a technical point a view. To face it, a constant effort to develop new advances in oil and gas exploration has been initiated in the past.

Overview of existing technologies for oil and gas exploration

Oil and gas exploration and production activities present multiple risks of being unsuccessful, which derive from technological, geological and economic uncertainties^{16–18}. During the exploration process, efficient prospecting technologies are required to minimize uncertainties about the presence of oil. This is of critical importance in the onshore domain, which represents more than 70 % of the global oil production. The mostly used technologies are presented in this section.

Oil macro- and microseepages bring evidence of mature source rock, which presence is an essential condition for oil accumulation in depth, as well as a reservoir and a cap rock. Macroseepages refer to visible oil and gas seeps caused by migration of hydrocarbons from porous rocks to the surface over a restricted area through faults and fractures (Figure 2a-b)¹⁹. Conversely, microseepages are more diffuse and are expressed as lithological and mineralogical alterations or vegetation changes over large areas (Figure 2a-b, d)^{20–22}. The field-based prospection of macro- and microseepages is the oldest and simplest way of detecting oil reservoirs from the surface.

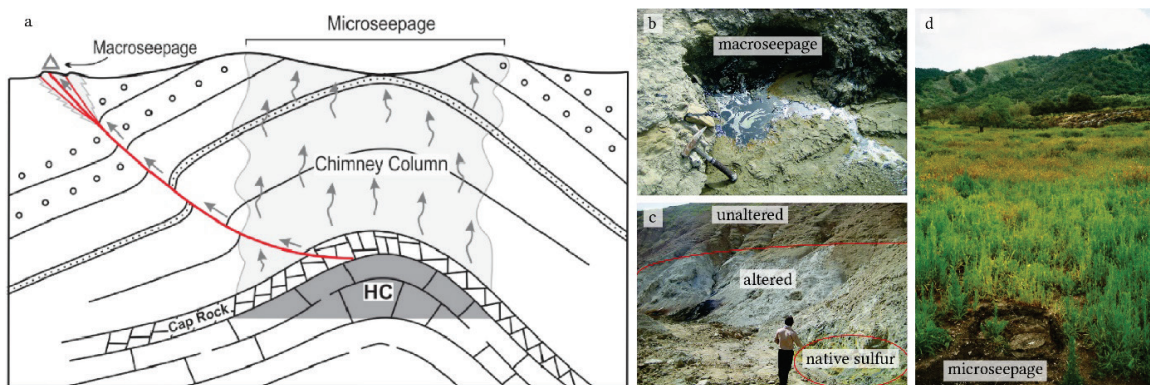


Figure 2. (a) Schematic representation of oil macro- and microseepages (adapted from ²³). (HC: Hydrocarbons.) Macroseepages are characterized by apparent crude oil at the surface. (b) Microseepages are manifested by alterations of the surface lithology (c) and specific vegetation patterns (d)^{20,24}.

Lots of geophysical technologies have later been developed for mapping the sub-surface. The use of reflection seismology started in the 1920's and has become the most reliable prospecting technique²⁵. This technique is based on the emission of seismic wave fields from the surface that are reflected and diffracted at each rock interface. The reflected waves are collected and analyzed to provide a detailed view of belowground geological formations and possible oil deposits. Reflection seismology has been complemented by gravimetric and magnetometric prospecting, which rely on the detection of anomalies in Earth gravity and magnetic field, respectively²⁶. When conjugated, all these techniques can prove oil accumulation under favorable conditions of depth and permeability, but in most cases, this has to be assumed from a known regional geological concept or from ancillary data, such as macro- and microseepage location. These techniques imply in return an important economic investment and workforce on site. Cost-effective and time-saving alternatives are therefore required. Surface oil detection is also in interest in other domains of oil industry. Among them, environmental monitoring and risk assessment are two major challenges for oil and gas companies.

Oil contamination: sources, consequences and monitoring

Soil and water contamination is the main ecological issue deriving from onshore oil production²⁷. Oil spills and leakages may occur at every step of the production process and affect ecosystems. They result from facility failures, bad practices and storm events (Figure 3a-g). For example, extraction wells, pipelines, refineries and mud pits are common sources of contaminant leaked in the environment²⁸⁻³². This includes crude oil, petroleum products, wastewaters and oil sludge³³⁻³⁵. All these contaminants cause severe ecological disturbances, such as landscape fragmentation and habitat loss or alteration, and affect human health³⁶⁻³⁹. Therefore, fast-detection is needed for assessing contamination and limiting its impacts. Lots of techniques have been developed for this purpose in response to major off-shore oil spills⁴⁰. However, the onshore domain did not receive the same attention. Main advances have been achieved in pipeline leak detection, which is one of the most important source of oil contaminants in the environment⁴¹⁻⁴⁴. Conversely, only little improvements have been made in assessing soil contamination deriving from extraction and refining activities or bad cessation management. Such operations are often made by field operators and do not guarantee an early detection of released contaminants, especially when it implies low and continuous quantities. They are time-consuming and lead to heavy ecological consequences when the contamination is not detected at early stage.



Figure 3. Principal sources of environmental contamination caused by oil activities. (a) Oil sludge pit⁴⁵, (b-c) vegetation and soil contaminated by crude oil leakage near a refining facility⁴⁶, (d) pipeline leakage⁴⁷, (e) crop contamination resulting from oil well blow out⁴⁸, (f) oil leakage from damaged storage tank following a storm⁴⁹ and (g) contaminated wastewater near a production site⁵⁰.

Common limits arise from the current technologies used in on-shore oil exploration and contamination monitoring. Among promising alternatives, remote sensing could achieve fast

detection of oil at large scale, fulfilling the needs of oil and gas companies. Encouraging perspectives of operational applications have emerged in this field, thanks to a growing interest over the last decades.

Remote sensing: an emerging tool for oil exploration and contamination monitoring

Active and passive remote sensing provide information about the composition of surfaces at large scale, by analyzing their radiometric properties in various domains of the electromagnetic spectrum^{51,52}. Active remote sensing is mainly used in the off-shore domain by taking advantage of surfaces' backscatter at various frequency bands and polarizations of the microwave domain (namely Synthetic Aperture Radar (SAR) remote sensing)^{53,54}. Focusing on the on-shore domain, applications in oil exploration mainly rely on passive optical remote sensing, which exploit the [400:2500] nm reflective domain⁵⁵. The use of optical remote sensing for oil seepage detection came along the emergence of airborne and satellite multispectral sensors. For example, several attempts of geological and mineralogical mapping have been made by photo-interpretation of LANDSAT imagery before 1990^{56,57}. However, a real interest has been given to remote sensing for oil exploration only a few decades ago, with the emergence of hyperspectral sensors (Figure 4, Table 1)²³.

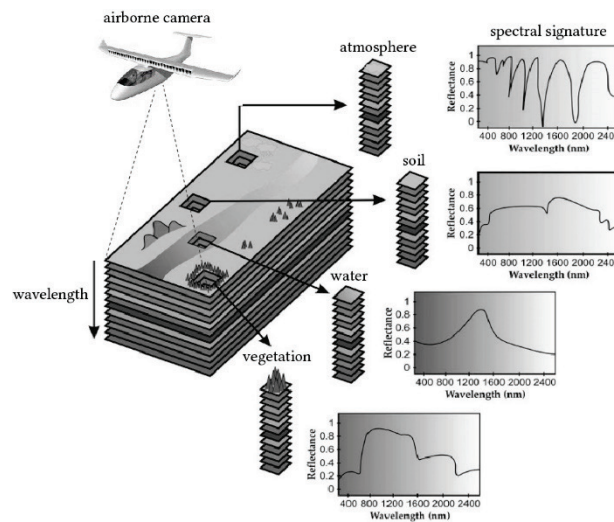


Figure 4. Principle of passive hyperspectral imagery (adapted from ⁵⁸). This technology provides the reflectance of surfaces over a continuous spectrum in the optical reflective domain (*i.e.* the spectral signature).

Hyperspectral sensors provide reflectance data over multiple and contiguous wavelengths of the optical reflective domain⁵². They give access to the spectral signature of surfaces (*e.g.* waterbodies, soils, vegetation), which helps determining their composition (Figure 4). Hyperspectral imaging sensors include drone-/UAV-, airborne- and satellite-embedded sensors⁵⁹. Along with a few multispectral sensors, some hyperspectral sensors provide high to very high spatial resolution, making possible to detect small targets. In complement, field portable spectroradiometers are usually used for characterizing the optical properties of surfaces in laboratory and for collecting reflectance data in the field⁶⁰. In oil exploration and contamination

monitoring, the use of hyperspectral sensors relies on exploiting the optical properties of petroleum hydrocarbons (*i.e.* direct detection) and changes in surface lithology, mineralogy and vegetation (*i.e.* indirect detection). Upcoming airborne- and satellite-embedded sensors will offer new opportunities in this field, so they are getting a growing attention by oil and gas companies (Table 1).

Table 1. Specifications of some of the current and future airborne and satellite-embedded sensors. For airborne sensors, the spatial resolution depends on the flight altitude. The name and specifications of future sensors may be modified until their operating.

Sensor name	Spectral domain (nm)	Bands	Spatial resolution (m)
Airborne sensors			
AisaFENIX	380 - 2500	348	0.5
AVIRIS	360 - 2500	224	20
CASI-1500	380 - 1050	48-288	0.2-1.5
DAIS	400 - 2500	72	1-10
HyMap	450 - 2500	128	3-10
HySpex ODIN-1024	400 - 2500	427	0.5
ROSI	430 - 860	103	1-2
Satellite sensors			
CHRIS	415 - 1050	19-63	18-36
EnMAP	420 - 2450	244	30
HISUI	400 - 2500	185	30
HJ-1A	450 - 950	115	100
Hyperion	357 - 2576	220	30
HypXim	400 - 2500	210	8
HySI	400 - 950	64	550
HyspIRI VSWIR	380 - 2500	212	30
PRISMA	400 - 2505	249	30
SHALOM	400 - 2500	275	10
TianGong-1	400 - 2500	128	10-20

The spectral signature of oil is characterized by two marked light absorption features in the [400:2500] nm reflective domain, located around 1730 and 2310 nm^{55,61}. These typical features come from the chemical properties of petroleum hydrocarbons⁶². Their amplitude varies with petroleum product type and concentration (crude oil, diesel, gasoline, etc.)⁶³. Consequently, petroleum hydrocarbons from macroseepages can be detected, characterized and quantified by analyzing the spectral signature of soils at specific wavelengths⁶¹. This refers to direct detection. Likewise, hyperspectral remote sensing can be used for indirect detection purposes, by identify mineralogical alterations related to the presence of oil, such as microseepages⁵⁹. Several studies succeeded in applying these principles in oil exploration, from experimental conditions to airborne imagery^{63,64}. Recent attempts also succeeded in detecting oil-contaminated soils around industrial facilities using hyperspectral airborne and satellite imagery^{46,65}. These studies emphasized the importance of high spectral resolution for achieving accurate detection of oil⁶⁶. Conversely, the importance of the spatial resolution highly differs according to the target. While microseepages and oil leakages spread over large areas (up to km²), macroseepages and chronic leakages are more localized (m²) and thus difficult to detect^{46,67}. Thus, very high spatial and spectral resolution imagery may satisfy direct and indirect detection of oil in all situations, making remote sensing a reliable alternative – or complement – to other existing techniques. From an operational point of view, hyperspectral imagery could provide a rapid diagnosis of surfaces at large scale, reducing the risk of exploration being unsuccessful or late detection of the contamination.

However, serious limits still compromise its use on non-bare soil surfaces, especially in vegetated regions.

Current limits of remote sensing for oil detection in vegetated regions

Oil is often apparent at the surface and therefore detectable directly using optical remote sensing that does not penetrate the ground. However, this becomes impossible in vegetated areas, because light penetration is strongly limited by the foliage and the spectral signature of soils is not accessible. This causes serious issues in oil exploration and contamination monitoring, especially since oil and gas companies concentrate an important part of their activities in vegetated regions. To overcome this limit, recent studies proposed an alternative technique to detect oil based on vegetation health assessment using multi- and hyperspectral remote sensing^{20,35,68}.

The optical properties of vegetation are strongly linked to its biophysical and biochemical parameters⁶⁹⁻⁷¹. This makes possible to detect changes in vegetation health from its spectral signature and thus to monitor stress using airborne and satellite optical imagery. A wide range of applications has emerged in this field during the last decades, especially in agriculture and ecology⁷²⁻⁷⁶. Lots of remote sensing methods have been developed for detecting and quantifying vegetation stress induced by biotic and abiotic factors (*e.g.* diseases, water-deficit). Most of them rely on tracking changes in leaf biochemistry or ground cover using reflectance data. Since crude oil and petroleum products affect vegetation health, these methods could serve for detecting oil indirectly in vegetated regions. Recent attempts have been made for this purpose^{20,68,77}. They have highlighted the need to develop supervised methods specifically dedicated to oil detection, which could be achieved using very high spatial and spectral resolution imagery. Important advances are therefore necessary to make hyperspectral remote sensing an operational tool for oil and gas companies.

OBJECTIVES OF THE THESIS

This work is part of the NAOMI project between TOTAL S.A. and the French Aerospace Lab (Office National d'Etudes et de Recherches Aéropatiales, ONERA), which aims to provide new advances in oil and gas exploration and contamination monitoring at large scale. It also benefited from the support of EcoLab and DynaFor research units of Toulouse. The present PhD thesis aims to develop new methods for detecting and quantifying oil in soils over temperate and tropical vegetated areas using hyperspectral airborne imagery.

The manuscript is structured as follows. The first chapter outlines a review of the current knowledge and limits in the research topic and introduces the approach adopted in the thesis. The following chapters are structured around scientific articles resulting from the thesis and submitted to or published in peer-reviewed journals. These articles describe the results obtained at each step of the approach.

This manuscript is written in English, in accordance with the authorization provided by the Institut Supérieur de l'Aéronautique et de l'Espace (ISAE – SUPAERO). It is accompanied by a synthesis in French that summarizes the main findings described in the articles.

CHAPTER ONE: STATE-OF-THE-ART OF PASSIVE HYPERSPECTRAL REMOTE SENSING FOR OIL EXPLORATION AND CONTAMINATION MONITORING IN VEGETATED REGIONS

INTRODUCTION

In regions covered by dense vegetation, hyperspectral remote sensing remains ineffective for detecting oil seepages and leakages directly (*i.e.* apparent oil), because of the masking effect of the foliage. The only information about soil composition can be provided by vegetation through its optical properties; in this case, its reflectance spectrum covering the reflective domain. This can be achieved because vegetation reflectance is closely linked to its biophysical and biochemical parameters, which are good indicators of environmental – especially stressful – conditions. Consequently, unfavorable growing conditions caused by oil might result in modifications of vegetation optical properties, which could be tracked by hyperspectral imagery. This would make possible to detect oil seepages and contaminated soil in vegetated areas at large scale. To succeed, this technique implies a good knowledge about vegetation optical properties and their modifications induced by exposure to oil, and being able to detect these modifications from airborne- or satellite-embedded sensors. Although a lot of work has been achieved experimentally in that direction, only few attempts have been made under natural conditions using hyperspectral imagery. This chapter proposes a state-of-the-art of advances in this field, and underlines its current limits that guided the work of this thesis. The chapter is organized as follows. The first section describes the biophysical and biochemical parameters influencing vegetation optical properties and the existing approaches used to retrieve them. An overview of their modifications induced by crude oil and petroleum products is then proposed in the second section. The last one focuses on the consequences of these modifications on vegetation optical properties and the supervised methods developed to detect them – and more generally vegetation stress – under controlled conditions and using airborne or satellite imagery.

1. Vegetation optical properties in the reflective domain (400 – 2500 nm)

Over the last 30 years, vegetation health assessment sparked an extensive attention by the remote sensing community. The development of field, airborne- and satellite-embedded sensors opened the way to various applications, especially in agriculture and ecology, thanks to a better comprehension of vegetation optical properties. In the reflective domain, these properties are driven by biophysical and biochemical parameters. They provide a singular shape to the spectral signature of healthy green vegetation, characterized by a peak of reflectance in the visible (VIS, 400 – 750 nm), a plateau in the near-infrared (NIR, 750-1300 nm) and two marked peaks in the short-wave-infrared (SWIR, 1300 – 2500 nm) (Figure 5). Leaf pigment and water contents and anatomy are the main parameters involved. This section describes their influence on leaf and canopy reflectance and introduces the main methods used for retrieving these parameters from vegetation reflectance.

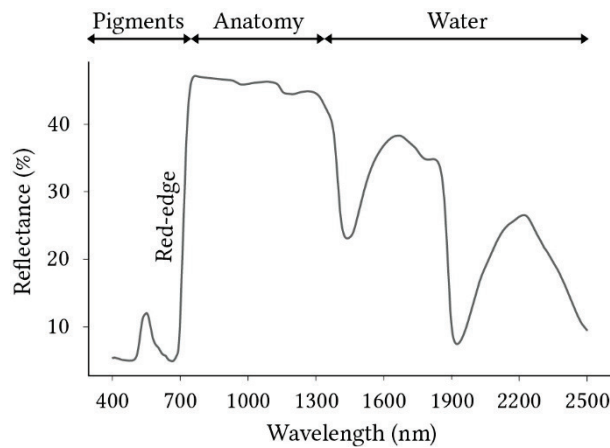


Figure 5. Typical spectral signature of healthy green leaf and most influential parameters in the different spectral regions.

1.1. Influence of leaf pigments in the visible region (400 – 750 nm)

A large diversity of pigments is present in plants⁷⁸⁻⁸⁰. Pigments are essential to the development of vegetation, because of their implications in photochemical reactions. They absorb light at various wavelengths in the ultraviolet (UV) and VIS regions, depending on their chemical properties. Consequently, the spectral signature of vegetation is strongly linked to leaf pigment content between 400 and 750 nm^{78,81,82}. This makes possible to track changes in pigments using multi- and hyperspectral sensors.

Chlorophylls a and b are the main pigments present in leaves. They are located in chloroplasts and are directly involved in the conversion of light into chemical energy through the photosynthesis, which makes them good indicators of vegetation health⁸³⁻⁸⁵. Therefore, they are largely studied in remote sensing^{73,86,87}. Chlorophyll concentration usually ranges from 0 to 80 $\mu\text{g}\cdot\text{cm}^{-2}$ in crops⁸⁸, of which only 20% are represented by chlorophyll b in healthy green leaves⁸⁹. These pigments show two light absorption peaks at 440-450 (blue) and 650-670 nm (red) that come from a magnesium ion (Mg^{2+}) surrounded by a chlorin ring^{90,91} (Figure 6a-b). Slight

spectral shifts between absorption spectra of chlorophyll a and b derive from differences in their chemical properties (a methyl group replaced by an aldehyde one).

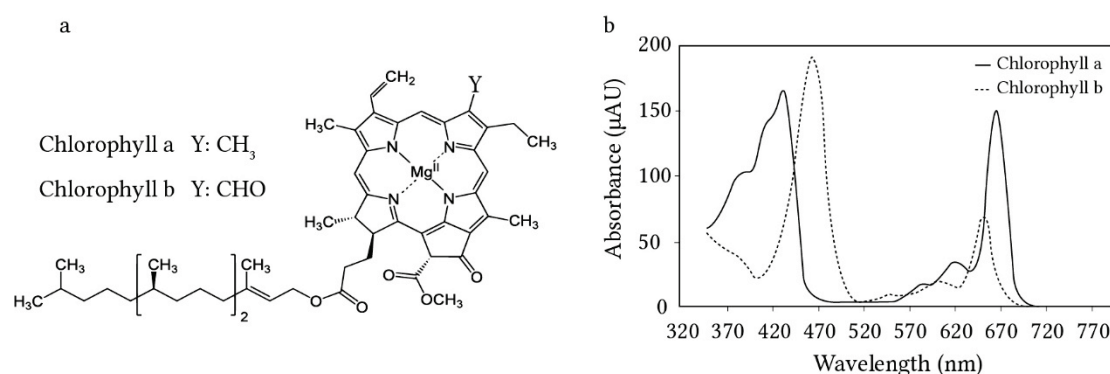


Figure 6. (a) Chemical structure and (b) absorption spectra of chlorophyll a and b⁹².

Due to their important concentration in leaves, chlorophylls have a strong influence on the spectral signature, so they are likely to hide the effects of other pigments sharing common absorption wavelengths. More precisely, the weak light absorption of chlorophylls around 550 (green) and 700 nm (red-edge) results in high correlation with leaf reflectance in these regions^{81,93}. Hence, remote sensing mostly exploits these wavelengths to quantify leaf chlorophyll content (LCC)⁸⁷. A large diversity of approaches have been developed for tracking changes in LCC, such as simple or normalized reflectance ratios (*vegetation indices*) and Radiative Transfer Models (RTM)^{70,78}. These approaches gave particular attention to the inflexion point of reflectance in the red-edge region – named the *Red-Edge Position* (REP), which is sensitive to little changes in LCC^{86,94,95} (Figure 7). Vegetation indices (VI) and RTM are developed later in this chapter, as well as their application for stress assessment purposes.

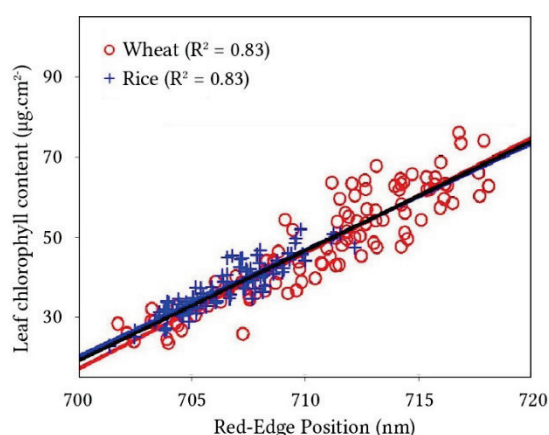


Figure 7. Relationship between leaf chlorophyll content and the Red-Edge Position (REP) (modified from ⁹⁶).

Leaf pigment contents naturally vary throughout the year, according to season changes^{89,97}. During autumn senescence, plant photosynthetic activity is highly reduced, so noticeable evolutions in leaf colors occur as a result of reduced chlorophyll content relative to other

pigments. Green leaves usually turn yellow-orange, since carotenoids – which are initially masked in leaves – are less degraded than chlorophylls and thus become “visible” (Figure 8a-e).

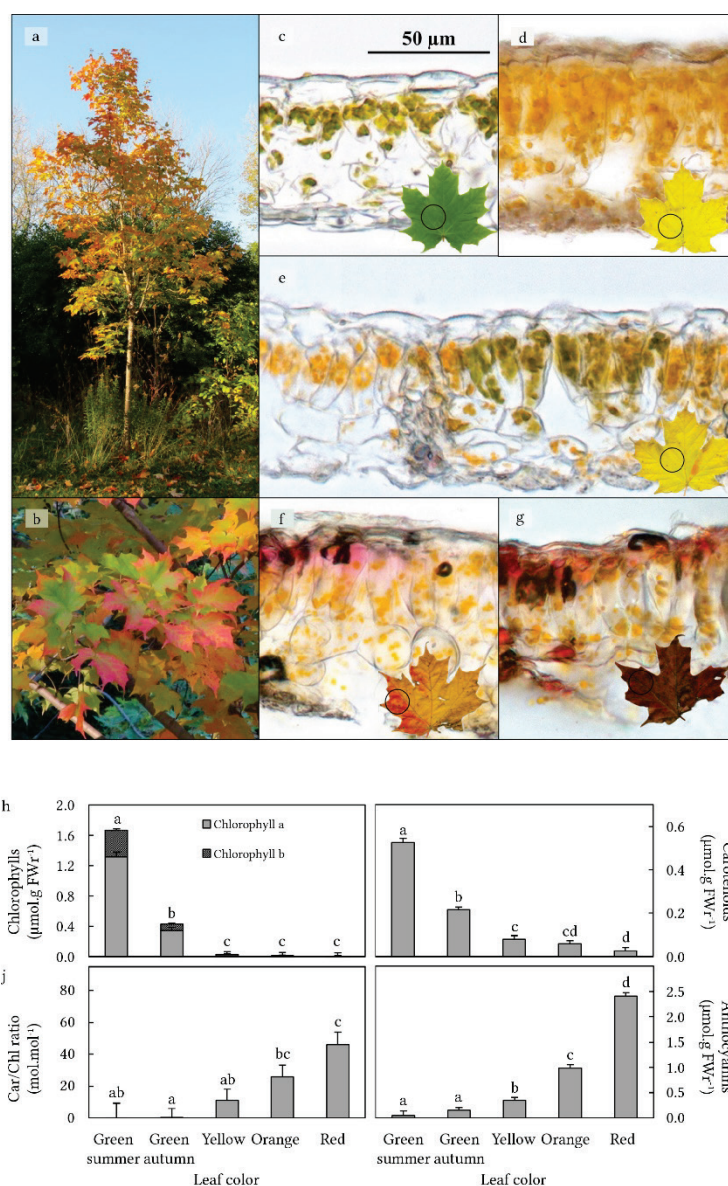


Figure 8. (a-b) *Acer saccharum* Marsh. (sugar maple) during autumn senescence. Sections of (c) green, (d-e) yellow, (f) orange and (g) red sugar maple leaves, and (h) chlorophylls, (i) carotenoids, (j) carotenoid-chlorophyll ratio (Car/Chl) and (k) anthocyanins found in corresponding leaves. Non-significant differences ($p \geq 0.05$) are denoted by similar lowercase letters (modified from ⁸⁹).

Carotenoids are the other photosynthetic pigments found in plants⁹⁸. They belong to C₄₀ isoprenoids and can be distinguished in two categories: carotenes and xanthophylls. α - and β -carotene are the principal carotenes present in leaf chloroplasts. Xanthophylls contain one or more oxygen functions and regroup lutein, violaxanthin, neoxanthin, antheraxanthin and zeaxanthin. The conjugated double bond of carotenoids is responsible for their ability to absorb light, mainly in the blue region (400 – 500 nm) (Figure 9a-b). This common feature with

chlorophylls explains their masking in healthy leaves, as their concentration rarely exceeds $25 \mu\text{g}\cdot\text{cm}^{-2}$ ⁸⁸. They are usually less influential on the spectral signature in the VIS and thus more difficult to quantify by remote sensing. However, the chlorophyll breakdown observed during leaf senescence increases the carotenoid-chlorophyll ratio (Figure 8j). Consequently, leaf reflectance rises between 500 and 750 nm (green – red), so carotenoids become more easily quantifiable.

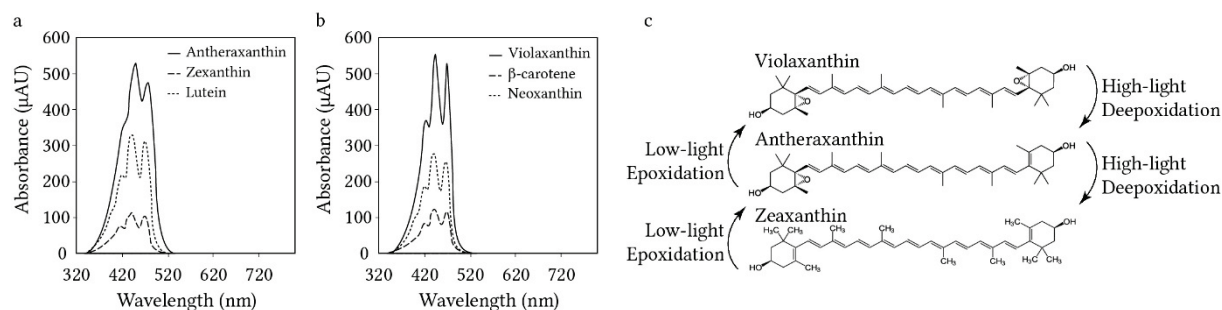


Figure 9. (a-b) Absorption spectra of principal carotenoids⁹². (c) The xanthophyll cycle⁸⁰.

Frequently described as accessory pigments, carotenoids ensure essential photoprotective functions in plants^{99,100}. They prevent leaf tissues from harmful effects of reactive oxygen species and photochemical stress that occur when absorbed light exceeds the photosynthetic capacity of leaves (*i.e.* energy assimilation and dissipation)^{97,98}. This function is particularly important in autumn, since photosynthesis is slowed by chlorophyll decrease and cold temperatures. It is ensured by violaxanthin, antheraxanthin and zeaxanthin, which form the xanthophyll cycle and interconvert, depending on light intensity (limiting or exceeding) (Figure 9). Therefore, the quantification of leaf carotenoid content is of great importance for monitoring vegetation health. Several VI have been designed for this purpose, such as the Photochemical Reflectance Index (PRI)^{101,102}. The PRI exploits reflectance at 531 and 570 nm to track the epoxidation state of the xanthophyll cycle and can be used for assessing variations of photosynthetic activity across seasons^{103,104} (Figure 10).

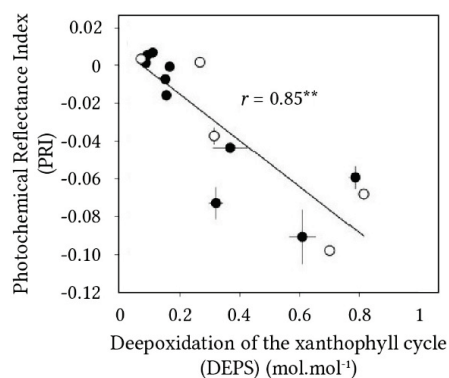


Figure 10. Relationship between the deepoxidation of the xanthophyll cycle and the Photochemical Reflectance Index (PRI)¹⁰⁴.

Leaves also contain non-photosynthetic pigments that are responsible for color changes in autumn. Several plants turn red during senescence, because of the accumulation of anthocyanins

in vacuoles (Figure 8f-g, k). Anthocyanins are water-soluble flavonoids that absorb light in the ultraviolet (UV, 250 – 350 nm) and green (500 – 560 nm) regions^{105,106} (Figure 11a-b). The pH strongly determines the proportion of the different forms of anthocyanins and thus their absorption spectrum. Anthocyanins accumulate in the upper region of palisade parenchyma (Figure 8g) and ensure a photoprotective function through UV screening. They are also antioxidant and are involved in leaf abscission timing and nutrient resorption^{89,107}. These pigments are thus important indicators of vegetation health.

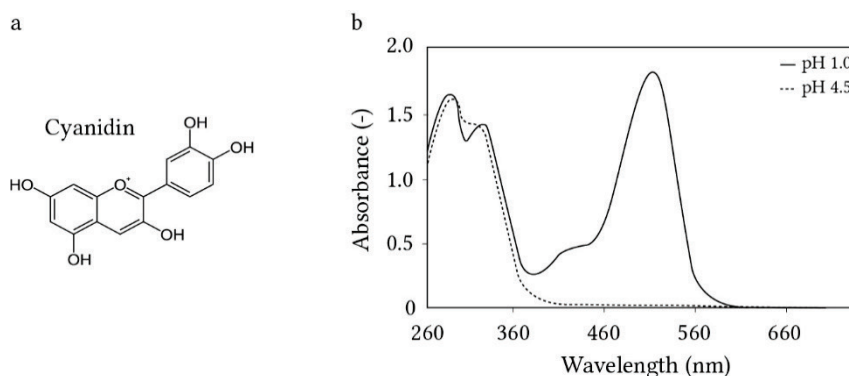


Figure 11. (a) Chemical structure of cyanidin, a red anthocyanin, and (b) absorption spectra of anthocyanins at pH 1.0 and 4.5¹⁰⁵.

There is evidence of anthocyanins *de novo* synthesis during autumn senescence^{89,106} (Figure 8k). Their accumulation leads to an increase of reflectance around 640 nm (red), providing the red coloration to leaves. Only little attention has been given to anthocyanins in remote sensing studies until recently. Specifically-designed VI and improved RTM have been proposed for estimating anthocyanins and emphasized the difficulty to achieve it for non-senescent leaves^{108,109}. Other compounds such as tannins are also found in leaves, but their influence on leaf optical properties is restricted to the late senescence – or pre-abscission – period⁹⁷. They are responsible for the browning of leaves.

Hence, leaf optical properties are mainly under influence of pigments – mainly chlorophylls – in the VIS. Evolutions in leaf color result from changes in leaf reflectance, which are caused by modifications of pigment composition. These modifications occur seasonally for most deciduous species, but they may also be induced by biotic and abiotic stressors, such as water and nutrient deficiencies. By exploiting vegetation optical properties, in particular around 550 nm and in the red-edge, multi- and hyperspectral remote sensing allows tracking pigment changes and assessing vegetation health.

1.2. Influence of leaf anatomy in the near-infrared region (750 – 1300 nm)

As pigments do in the VIS, leaf anatomy drives reflectance in the NIR region. Leaves of Angiosperms are formed by successive cellular layers structured in parenchyma – also called mesophyll – and protected by a cuticle and an epidermis on abaxial (lower) and adaxial (upper) faces (Figure 12a). This anatomy is at the origin of the plateau observed on leaf spectral signature in the NIR, which ranges from 30 to 80% reflectance^{71,110,111}. The upper cuticle and epidermis are the first barriers to the penetration of light. The cuticle is mostly made of cutin, a polymer

structure that protects a non-photosynthetic cellular layer (*i.e.* the epidermis). The epidermis also contains stomata by which gas exchanges occur (CO₂ and H₂O). Incident light follows diffuse and specular reflection at leaf surface, but most radiations go through it and are transmitted to lower layers^{112,113}. The presence of wax coat on the cuticle also modifies leaf optical properties, but this case remains poorly documented. Epidermis cells help converging light to parenchyma, where photosynthesis occurs.

The internal anatomy of leaves greatly contribute to their optical properties in the NIR, but differs between mono- and dicotyledonous species^{110,114,115}. In dicotyledonous leaves, cells are typically arranged in two distinct parenchyma. Both can be organized in multiple layers. The upper one – known as palisade parenchyma – is made of well-structured elongated cells with high chloroplast concentration (Figure 12a). Intercellular spaces are almost absent from this layer so light scattering remains limited. Conversely, the lower – spongy – parenchyma is characterized by irregularly-shaped and spaced cells with low chloroplast content. It also contains vascular bundles. The spongy parenchyma has an important function in leaves, as it sends back a fraction of incident light to the palisade parenchyma, thus increasing the photosynthetic activity¹¹⁶. In monocotyledonous leaves, parenchyma are undifferentiated. Cells form a unique layer similar to the spongy parenchyma of dicotyledonous leaves, although this one is more compact so intercellular spaces are reduced (Figure 13c).

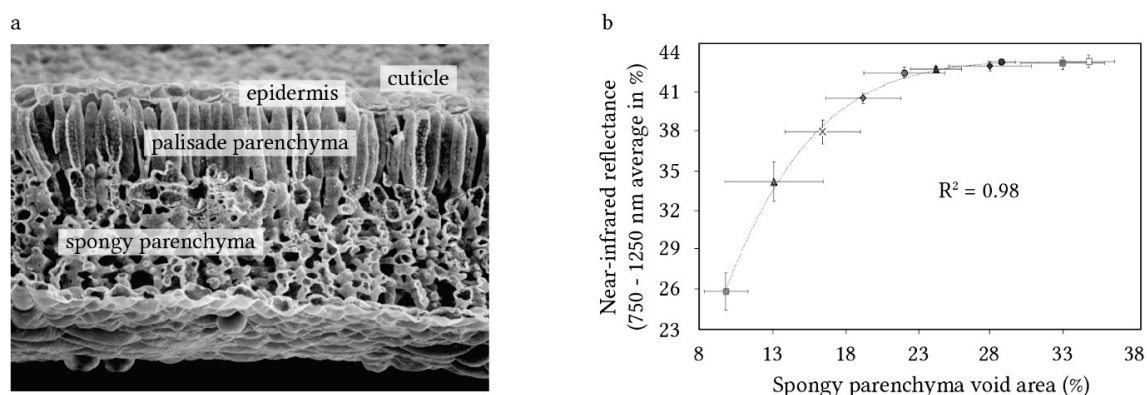


Figure 12. (a) Typical anatomy of a dicotyledonous leaf (cryo-scanning electron microscopy) and (b) relationship between spongy parenchyma void area (intercellular spaces) and leaf reflectance in the near-infrared (modified from ¹¹⁷).

Several studies focused on linking leaf anatomy and optical properties from laboratory or field measurements (*i.e.* using an integrating sphere or a leaf-clip connected to a spectroradiometer). It has been shown that the cuticle and parenchyma thickness, the proportion of intercellular spaces and the arrangement of chloroplasts greatly affect leaf reflectance in the NIR^{71,110,118,119} (Figure 12b). Additionally, these parameters may also influence light diffusion inside leaves in the VIS, as suggested by recent studies¹²⁰.

Leaf anatomy substantially varies among species, partly as a result of phylogeny and adaptation to light conditions (Figure 13a-i)^{121–123}. These differences can be easily observed in temperate and tropical forests, where light competition is important¹²⁴. They also appear for a single individual. For example, Castro & Sánchez-Azofeifa¹²⁵ noticed modifications of palisade and spongy parenchyma thickness and intercellular spaces of *Populus tremuloides* Michx. and *Populus balsamifera* L. leaves when exposed to shade. Additional factors also influence leaf anatomy and NIR reflectance, such as nutrient and water availability or soil contamination.

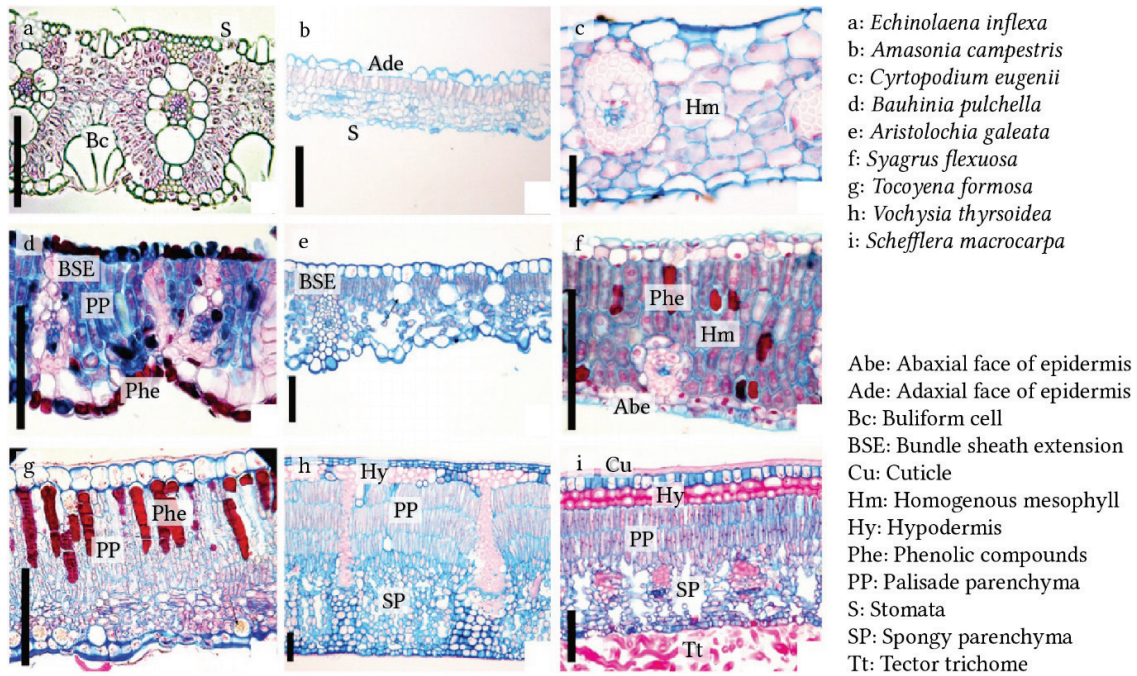


Figure 13. Diversity of leaf anatomy in tropical region¹¹⁴. (Scale bars = 100 μm .)

While the anatomy of leaves determines their reflectance in the NIR, other biophysical parameters prevail when measuring reflectance at canopy scale (*e.g.* using airborne- or satellite-embedded sensors). Leaf Area Index (LAI) and the Leaf Angle Distribution (LAD) are the most influential ones^{69,116,126}. The LAI is defined as the projected area of leaves per unit of ground surface ($\text{m}^2.\text{m}^{-2}$), so it is directly linked to plant development. Canopy reflectance is positively correlated to LAI in the NIR, because the influence of bare soil is reduced in this region as LAI increases¹²⁷ (Figure 14a). However, the reflectance reaches a plateau above very high LAI values¹¹⁶ (>6). LAD characterizes canopy architecture, *i.e.* the angular orientation of leaves. This parameter greatly influences light incidence angle at leaf surface and thus light trajectory inside leaves. Wang *et al.*¹²⁸ defined the LAD as the fraction of leaf area per unit leaf zenithal or azimuthal angle. In previous work, de Wit¹²⁹ suggested that leaves have no preferred azimuthal angle and proposed to classify species in the following six LAD types, according to leaf zenithal angle: Planophile, plagiophile, erectophile, extremophile, spherical and uniform. The contribution of the LAD to canopy reflectance in the NIR has been assessed in previous studies by varying Mean Leaf Angle (MLA) in RTM^{69,116}. As MLA is moving away from zero degrees (toward planophile LAD), canopy reflectance decreases in the NIR (Figure 14b).

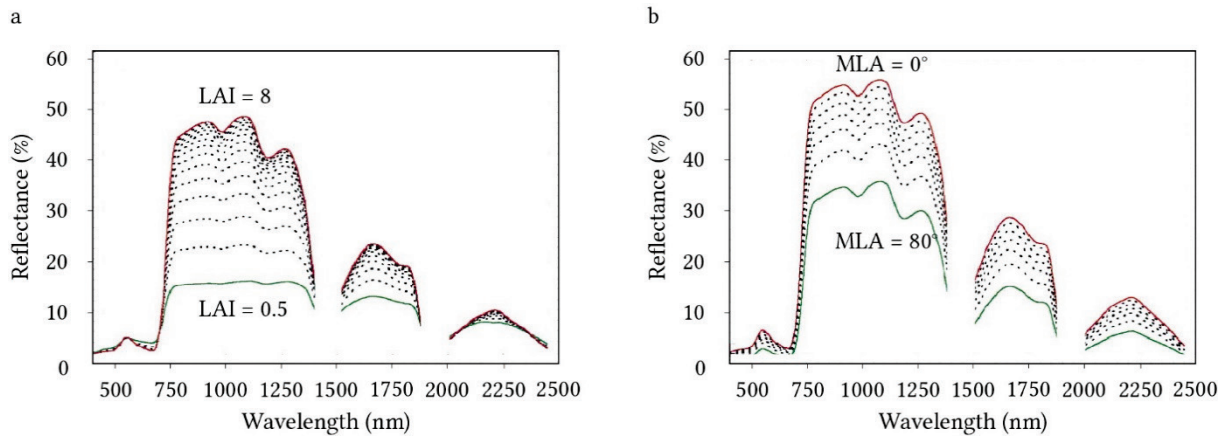


Figure 14. Influence of (a) Leaf Area Index (LAI) and (b) Mean Leaf Angle (MLA) on canopy reflectance⁶⁹. Spectral signatures were simulated using the PROSAIL model described in section 9.2.4.1.

Because of its relationship with vegetation biophysical parameters, reflectance in the NIR can be used to describe leaf anatomy, canopy architecture and ground cover^{71,130}. These parameters have in common to be directly or indirectly influenced by vegetation water status^{111,131,132}. Water availability is a key parameter for understanding vegetation optical properties, as it drives many physiological mechanisms.

1.3. Influence of leaf water and dry matter contents in the near-infrared (750 – 1300 nm) and short-wave infrared (1300 – 2500 nm) regions

Water is absorbed by roots and distributed in plant through vascular bundles. It is present under liquid and gaseous states in plants and found in free form in all compartments or chemically bound in molecules. Vegetation optical properties are directly influenced by free form water in leaves, which absorbs light around 970, 1200, 1450, 1950 and 2450 nm^{133–135} (Figure 15a). These features are easily observed on the spectral signature of healthy plants and are affected by changes in leaf water content¹¹¹ (Figure 15b). Hence, they are reliable indicators of vegetation water status¹³⁶. In addition, water is likely to affect reflectance indirectly in other spectral regions, as it is involved in many physiological mechanisms in plants, such as photosynthesis and leaf turgor. This is particularly marked for plants undergoing water-deficit stress. Water availability is a limiting factor to primary production. When it reaches critical levels, plants close stomata to prevent excessive loss of water by transpiration¹³². This causes a decrease in photosynthetic activity and thus in pigment content, which results in reflectance changes in the VIS. Likewise, changes in leaf turgor and tissue destructuring induced by water-deficit greatly affect light scattering and thus leaf reflectance in the whole NIR region¹¹¹. These effects are also observed at canopy scale, as plant LAI and LAD are also affected by water-deficit stress¹³⁷.

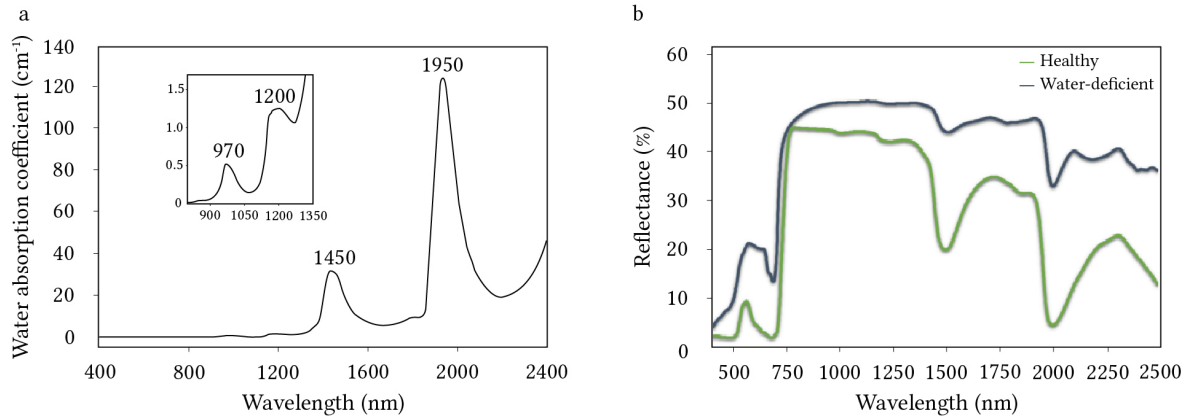


Figure 15. (a) Absorption coefficient of water¹³⁵, and (b) spectral signatures of healthy and water-deficient plants (personal data).

Several studies demonstrated the effectiveness of the NIR and SWIR reflectance to assess vegetation water status by estimating Leaf Water Content (LWC) or Equivalent Water Thickness (EWT)^{133,134,136}. VI and RTM have been widely used for this purpose¹³⁸⁻¹⁴¹. Although water absorption bands previously cited may be appropriate^{133,134}, their utilization remains limited in airborne or satellite imagery, because of important noise due to atmospheric effects of water vapor. This limit can be however overcome by exploiting other water-dependent and atmospherically-resistant wavelengths in the NIR and SWIR regions¹⁴¹⁻¹⁴³ (Figure 16).

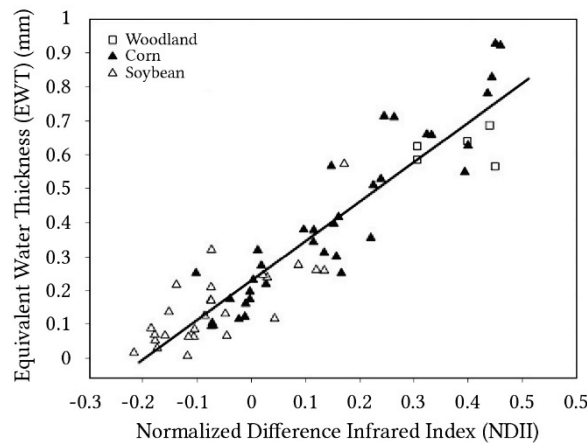


Figure 16. Relationship between leaf Equivalent Water Thickness (EWT) and the Normalized Difference Infrared Index (NDII), which exploits reflectance at 850 and 1650 nm¹⁴¹.

As described in this section, vegetation optical properties are strongly linked in the NIR and SWIR regions, because of direct and indirect influence of water. According to Ceccato *et al.*¹³⁹, water stands for approximately 55 to 75% of healthy leaf fresh weight for temperate species. More than two thirds of the remaining part come from hemicelluloses, celluloses, lignins and proteins, which are often grouped in the “dry matter” term^{144,145}. Celluloses are the most abundant organic compounds on earth and are found in all plants. Hemicelluloses and lignins are mostly represented in woody species (up to 20 – 35% of wood dry weight)^{146,147}. All these biochemical parameters enter in the composition of cell wall and provide rigidity to plants. They

share common light absorption features in the optical reflective domain. These are located in the NIR and SWIR regions, at 1200, 1450 – 1490, 1540, 1760, 2100 and 2340 nm, and mainly come from O–H and C–H stretches in chemical structures¹⁴⁸. Proteins are heteropolymers made of amino acids, and are essential to many physiological functions (*e.g.* photosynthesis) and cell membrane composition. N–H stretches and bends provide them quite different light absorption features, located at 1510 – 1520, 1730, 1980, 2060, 2165 – 2180 and 2300 nm.

The influence of hemicellulose, cellulose, lignins and proteins on vegetation optical properties remains poorly documented, since attention focused on other parameters to monitor vegetation health (*e.g.* chlorophylls and LAI). Moreover, they remain difficult to estimate, because their influence on reflectance in the NIR and SWIR regions is limited in comparison to water^{144,148} (Figure 17a). They become however more influential in dry leaves (Figure 17b). Few VI have been designed for retrieving celluloses and lignins content in leaves or decomposing litter^{149,150}. Likewise, attempts have been made with RTM and succeeded when grouping celluloses and lignins^{148,151}. Despite efforts in that sense, all these biochemical parameters remain difficult to discriminate, so they are preferably estimated all together as dry matter.

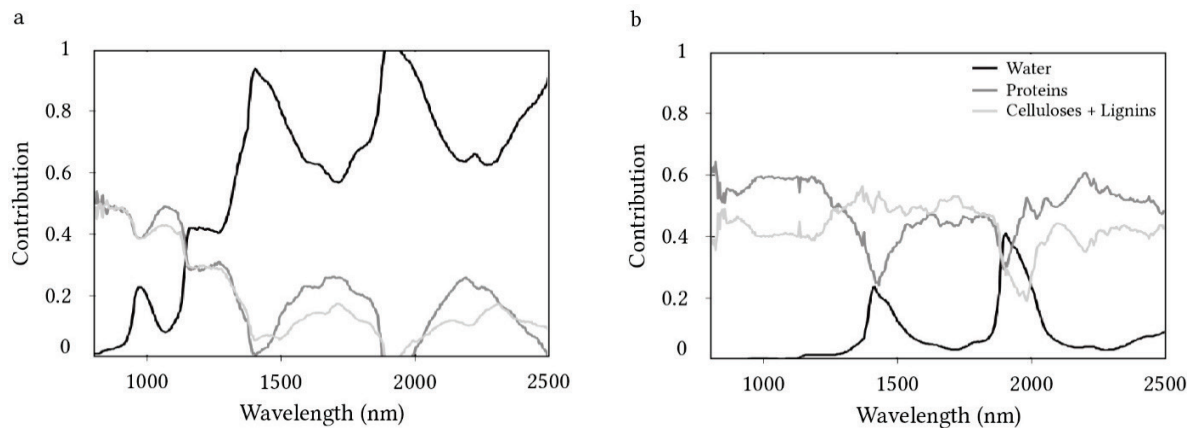


Figure 17. Relative contribution of water, proteins and celluloses + lignins to light absorption of (a) fresh and (b) dry leaves in the NIR and SWIR regions¹⁴⁸.

As outlined in this section, the biophysical and biochemical parameters driving vegetation optical properties differ according to the spectral region (VIS, NIR and SWIR). Modifications in these parameters are expressed as changes in the reflectance of leaves and canopies. This suggests making possible to detect oil-induced alterations in vegetation health using multi- and hyperspectral remote sensing. This purpose however requires identifying the most suitable (*i.e.* oil-sensitive) spectral regions. A good comprehension of the effects induced by crude oil and petroleum products on vegetation helps achieving it. These effects are described in the next section.

2. Effects of crude oil and petroleum products on vegetation health

Whether they are naturally present or leaked from industrial facilities, crude oil and petroleum products are likely to affect vegetation health and optical properties. Their particular nature and composition are greatly responsible for these effects.

2.1. Composition of crude oil and petroleum products

Crude oil refers to oil in its natural and extractible form, *i.e.* oil stored in geological formation and brought to the surface¹⁴. Petroleum products result from the refining of crude oil. They include fuels (diesel, gasoline, kerosene), lubricant, waxes and miscellaneous products used in various domains (*e.g.* transportation and industry)²⁶. The term oil refers both to crude oil and petroleum products. Wastewaters and oil sludge are produced during the refining process^{33,34}. Both crude oil and petroleum products are mixtures of volatile to dense hydrocarbons (called *Petroleum hydrocarbons*), heavy metals (HM, also termed *Trace Metal Elements*) and oxygen, sulfur and nitrogen compounds in various proportions^{152–154}. Petroleum hydrocarbons include Mono- and Polycyclic Aromatic Hydrocarbons (BTEX and PAH, respectively), and saturated (alkanes or paraffins) and unsaturated (alkenes and alkynes) hydrocarbons²⁶. *Total Petroleum Hydrocarbons* (TPH) is a generic term that encompasses all these compounds. Depending on the length of their carbon chain, petroleum hydrocarbons are refined to different petroleum products^{155,156}. An illustration is given in Figure 18.

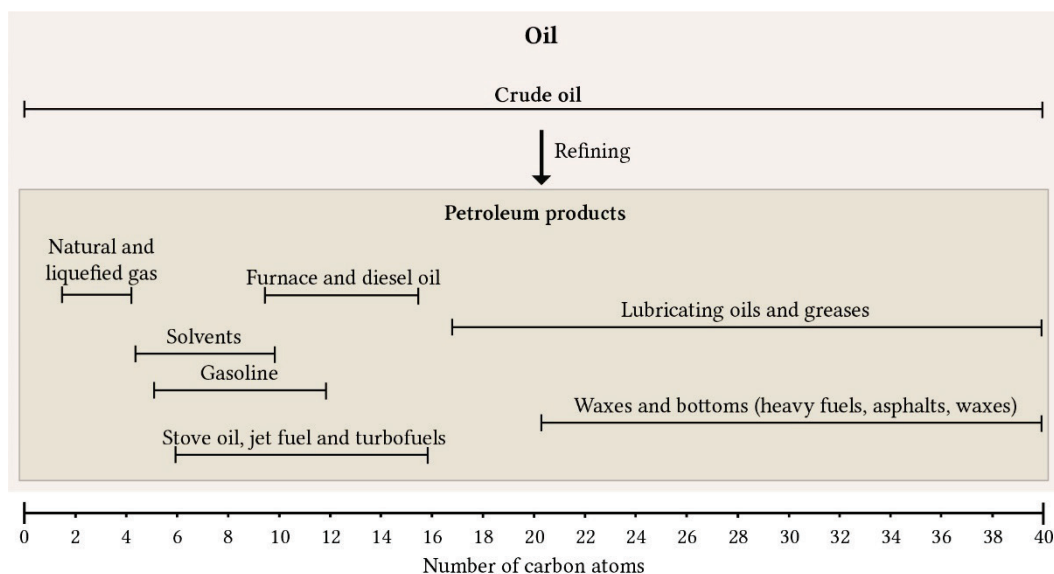


Figure 18. Crude oil and petroleum products according to petroleum carbon ranges (modified from ¹⁵⁶).

The composition of crude oil and petroleum products gives them a high toxicity towards vegetation¹⁵⁷. When considered separately, each hydrocarbon and HM type is likely to affect vegetation health¹⁵⁸. Since they are in mixture, it remains difficult to identify which of these compounds are responsible for the observed response. In addition, interactions can occur among hydrocarbons and HM and result in synergistic or antagonist effects on vegetation. However, the

influence of mixture composition is still misunderstood. Different mixtures such as crude oil, diesel or gasoline, lead to different responses of vegetation^{77,159}. These responses result from indirect effects caused by modifications of soil physico-chemical and biological properties, and from direct effects through contact with plant and assimilation in tissues^{160,161}. Both occur at root level and lead to anatomical and biochemical changes in leaves, so these direct and indirect effects remain difficult to differentiate^{162–164}. For that reason, they are described jointly in the following sections.

2.2. Effects on soil physico-chemical and biological properties and on plant roots

The phytotoxicity of petroleum hydrocarbons and HM has been subject to numerous studies. However, most of them focused on one compound and did not account for their effects in mixture^{165,166}. This topic has only been addressed recently and provides a better comprehension of how vegetation is affected by oil seepages and leakages.

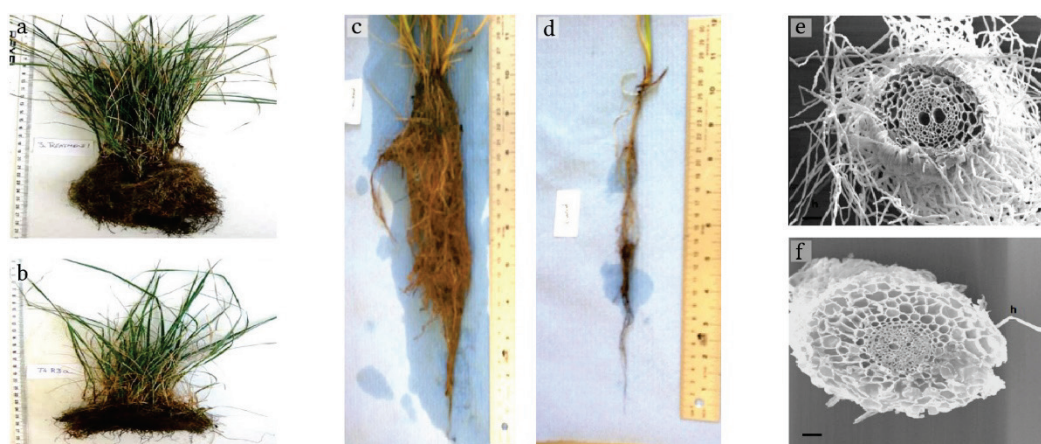
2.2.1. Effects on soil water regime

Because of their particular nature and composition, crude oil and petroleum products induce important modifications of soil physico-chemical and biological properties^{154,167,168}. Consequently, they impose selective growing conditions to plants²⁰. Soil water regime is one of the most impacted properties. Because of hydrocarbons, crude oil and petroleum products are in liquid – highly hydrophobic – form²⁶. When found in soils, they occupy a fraction of porosity that becomes unavailable to water. In addition, by interacting with soil materials (especially clay), oil forms a hydrophobic film at their surface, which forces water drainage toward deeper soil layers. These phenomena contribute to reducing the field capacity of soil (*i.e.* water remaining after infiltration) and plant water supply^{169–171}. This effect is amplified by HM, which affect soil water potential and water uptake by roots once transferred to soil solution¹⁷². Hence, crude oil and petroleum products can lead to water-deficient soils and eventually plant water stress¹⁵⁷. As described above, this type of stress is easily recognizable from the spectral signature of vegetation.

2.2.2. Effects on nutrient availability and root uptake capacities

Petroleum hydrocarbons represent a considerable enrichment in organic material, leading to an increase of soil carbon content and carbon / nitrogen ratio (C/N)^{154,171,173}. This stimulates the growth of microorganisms capable of degrading hydrocarbons, thus modifying organic matter mineralization cycles and reshaping microorganism communities^{174–176}. The biodegradation of hydrocarbons is accompanied by an elevation of soil CO₂ concentration, which is enhanced in the presence of vegetation¹⁷⁷. The proportion of light petroleum hydrocarbons (< 10 C), which are more easily biodegraded than the heavy ones (> 10 C), is strongly determining. Some of the HM found in oil are essential to vegetation growth (*e.g.* Fe, Zn, Cu), but their occurrence at high concentrations along with other HM (*e.g.* Cd, Mg, Pb) also affect microorganisms¹⁷⁸. However, they are not degradable and in the case of oil leakage, they concentrate in the first soil layers¹⁷⁹. The nitrogen cycle is particularly impacted by carbon enrichment: the availability of inorganic nitrogen decreases so vegetation nitrogen status is highly altered¹⁷³. Likewise, several studies revealed that petroleum hydrocarbons and HM reduce nutrient availability (P, K) and soil Cation Exchange Capacity (CEC)^{170,171,180}. The latter is indeed closely linked to soil organic matter content, C/N ratio and pH; so many parameters affected by oil^{152,154}. Through modifications of soil physico-chemical and biological properties, crude oil and petroleum products thus affect water

and nutrient availability for plants¹⁸¹. These effects are called indirect effects. In addition, direct effects occur when oil is in contact with roots¹⁸². As they do with soil materials, petroleum hydrocarbons are able to coat plant roots by adsorbing at their surface. As well as HM, their assimilation inhibits root growth and causes a thickening of root epidermis, endodermis and cortex, and a reduction of root hair diameter and density^{160,172,183,184} (Figure 19a-g). These anatomical changes heavily alter water and nutrient uptake capacities of plants. For some species, they are partly compensated by a higher allocation of resources to roots.



g

Root anatomical parameter	Uncontaminated soil	Crude oil-contaminated soil
Root diameter (μm)	457.78 (\pm 19.86) ^a	527.46 (\pm 29.11) ^b
Cortex zone area (mm)	0.12 (\pm 0.004) ^a	0.22 (\pm 0.02) ^b
Endodermis wall thickness (μm)	2.21 (\pm 0.20) ^a	5.99 (\pm 0.07) ^b
Root hair length (μm)	767,1 (\pm 94,58) ^a	241,86 (\pm 34,69) ^b

Figure 19. Development of *Festuca arundinacea* grown for three months on (a) uncontaminated or (b) crude oil-contaminated (10.8 g.kg^{-1}) soil. Root system and scanning-electron microscopy images of root sections from (c, e) uncontaminated and (d, f) crude-oil contaminated plants after 14 months. Comparison of selected root anatomical parameters between the two treatments (mean \pm SE). Significant differences ($p < 0.05$) between soils are denoted by different superscript letters (modified from ¹⁶⁰).

As soon as water or nutrient supply is no longer sufficient to ensure essential physiological functions, stressful conditions arise so plant undergoes anatomical and biochemical modifications that affect its reflectance. These effects are amplified by the accumulation of certain hydrocarbons and HMs in leaves^{158,161,162}. They are summarized in the following section.

2.3. Consequences on vegetation health

The biophysical and biochemical parameters affected by exposure to crude oil and petroleum products are involved in vegetation optical properties. A review of these effects is proposed in Table 2. This suggests reflectance changes at particular wavelengths, making possible to discriminate healthy and affected vegetation.

2.3.1. Effects on plant biochemical and biophysical parameters

The alteration of leaf pigment content is the most frequently described response of plant to crude oil and petroleum products^{77,162,185}. This alteration is induced by that of plant water and nitrogen status described above¹⁶⁹. It can be visually observed through symptoms of leaf discoloration after a short to prolonged time of exposure^{183,186} (Figure 20f-g). The discoloration is caused by a reduction of LCC and indicates a decrease in photosynthetic activity¹⁵⁹ (Figure 20h, Table 2). This response is very common for water-deficient plants. It is caused by a decrease of stomatal activity (*i.e.* stomatal conductance), which aims to preserve water by limiting transpiration^{48,132}. In return, this mechanism reduces CO₂ entrance and thus photosynthetic activity and chlorophyll content. Although they are naturally present at lower concentrations in leaves, carotenoids and anthocyanins are also affected⁷⁷. HM accumulation amplifies this effect^{172,187}. Judging by the influence of pigment on leaf optical properties, their alteration is likely to induce reflectance changes in the VIS region. As described in section 1.3, the SWIR region is greatly impacted by changes in plant water status. Since LWC and cell turgor are altered by exposure to crude oil and petroleum products, modifications of reflectance may occur in this region^{169,172}.

Alterations of biophysical parameters can be observed at different scales. At leaf scale, they are expressed as a reduction in the number and size of cells and an increase of intercellular spaces in parenchyma^{162,164} (Figure 20h, Table 2). The accumulation of certain hydrocarbons and HMs – especially Cd and Mg – also causes tissue destructuring^{157,158}. Consequently, important modifications of leaf spectral signature are expected in the NIR region, where reflectance is largely determined by leaf anatomy. At canopy scale, water and nutrient deficiency leads to a limited development (*i.e.* a reduction of leaf and stem length and density), reducing aboveground biomass and LAI (Table 2). In addition, changes in leaf anatomy and water content affect plant habit and consequently LAD. As for leaves, changes in canopy reflectance are thus possible in the NIR region.

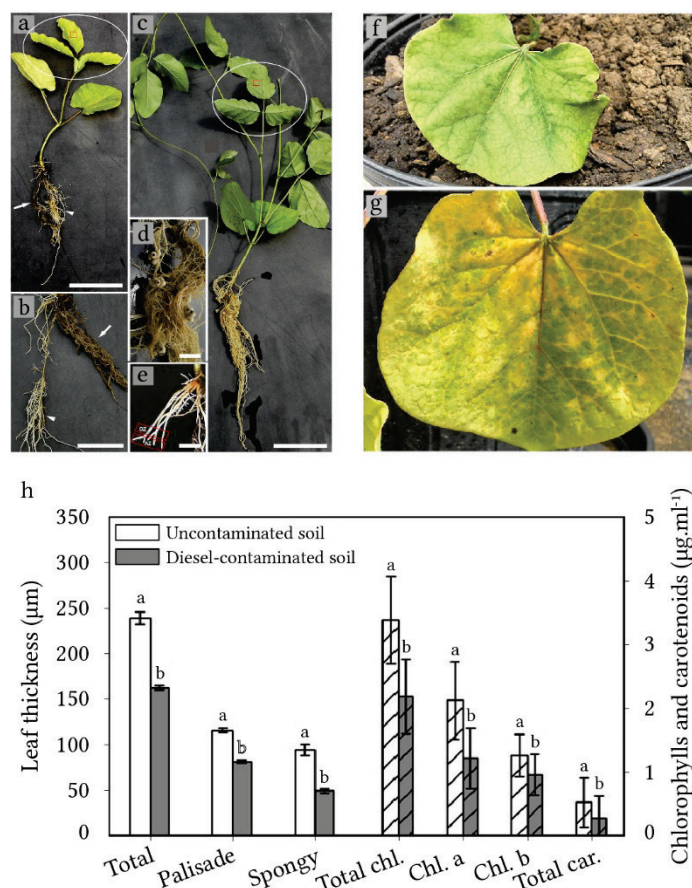


Figure 20. *Canavalia ensiformis* (L.) DC grown on (a-b, e) diesel-contaminated (22.219 mg.kg⁻¹) or (c-d) uncontaminated soil. White arrows indicate necrotic roots. (f-g) Symptoms of leaf discoloration observed on diesel-exposed plants. (h) Comparison of leaf anatomical parameters and pigment contents between the two treatments (mean ± SD). Significant differences ($p < 0.05$) between soils are denoted by different lowercase letters (modified from ¹⁸³).

2.3.2. Sources of variability

The effects described above correspond to global trends from the literature. An exhaustive review is proposed in Table 2. The severity of these effects highly varies according to the context, because they are influenced by many factors. The sensitivity of the species is a determining one^{188,189}. Since all species do not share similar ecological requirements, their tolerance to stressful conditions differs. Consequently, a decrease in soil water and nutrient availability caused by crude oil and petroleum products will not affect all species in the same way¹⁹⁰. Moreover, some species are capable of detoxifying hydrocarbons and HMs accumulated in leaves through mechanisms of sequestration, transportation and excretion^{165,191}. This prevents biochemical alterations and tissue destructuring. Few species are even stimulated by the enrichment of soil organic matter provided by crude oil and petroleum products, but this response remains uncommon¹⁹². This variability in species' sensitivity has strong implications under natural conditions. For example, only few species are established around micro- and macroseepages²⁰. Their presence is explained by a high tolerance to chronic crude oil exposure, so these species undergo no or little alterations. Mud pits contaminated by oil production residues (*e.g.* oil sludge) are similar cases, but they remain rarely studied¹⁸¹. Conversely, crude oil and petroleum products leaked from drilling well, storage tank and pipeline leakages consist in a rapid exposure of

probably oil-intolerant species. In those conditions, severe alterations and sometimes plant death are observed¹⁸⁵. Most of the previous studies focused on crops, which are poorly representative of the species found around seepages and industrial facilities and on mud pits.

Petroleum hydrocarbon and HM availability for plants strongly varies according to their chemical properties. For example, low-carbon PAHs and As are easily accumulated in leaves¹⁵⁸. Therefore, mixture composition influences plant response, so different crude oils or petroleum products (*e.g.* diesel, gasoline) do not affect leaf biophysical and biochemical parameters of a single species in the same extent^{157,161}. Apart from mixture composition, these effects are also conditioned by the level and time of exposure to oil^{77,169,189}. More precisely, the amplitude of pigment and water content alteration in leaves is positively correlated to the overall TPH concentration. Above a threshold concentration that depends on species' sensitivity (generally in g.kg^{-1}), plant death can be observed after only few days^{162,188}. In contrast, several weeks of exposure might be required to induce biophysical and biochemical alterations at low concentrations¹⁸³ (μg to mg.kg^{-1}).

Although the effects of petroleum hydrocarbons and HM mixtures on vegetation are well documented in the literature, they cannot be generalized to all contexts of seepages and leakages because their severity depends on many factors. Species' sensitivity, mixture composition and concentration and exposure time have been identified as the most influential ones. These factors have critical implications in remote sensing, since they determine the amplitude of reflectance changes in vegetation and thus hydrocarbon detectability using airborne and satellite-embedded sensors.

Table 2. Literature review of the effects induced by crude oil and petroleum products on vegetation health. Effects are sorted by biophysical and biochemical parameters related to vegetation optical properties in the different spectral regions. (↑ and ↓ denotes increase and decrease in the measured parameter, respectively; dose-dep. indicates dose-dependent effects; n.a.: not available.)

Species	Crude oil / petroleum product	TPH	Total time of exposure	Effect	Ref.
Anatomy / Development					
<i>Cyperus brevifolius</i>	Crude oil	10-50 g.kg ⁻¹	6 months	↑ Cuticle thickness (dose-dep.)	162
<i>Cyperus brevifolius</i>	Crude oil	10-50 g.kg ⁻¹	6 months	↓ parenchymatous cell length and diameter (dose-dep.)	162
<i>Cyperus brevifolius</i>	Crude oil	10-50 g.kg ⁻¹	6 months	↓ intercellular spaces length and diameter (dose-dep.)	162
<i>Terminalia catappa</i>	Crude oil (spill)	n.a	3 weeks	↑ Cuticle thickness	164
<i>Terminalia catappa</i>	Crude oil (spill)	n.a	3 weeks	↑ Epidermal cell diameter	164
<i>Terminalia catappa</i>	Crude oil (spill)	n.a	3 weeks	↑ Palisade and ↓ spongy parenchyma thickness	164
<i>Canavalia ensiformis</i>	Diesel	22,219 mg.kg ⁻¹	30 days	↓ Palisade and spongy parenchyma thickness	183
<i>Allophylus edulis</i>	Crude oil	13.65 g.kg ⁻¹	30-60 days	↑ Shoot length and biomass unchanged	182
<i>Amorpha fruticosa</i>	Crude oil	5-20 g.kg ⁻¹	6 months	↓ Shoot biomass (dose-dep.)	193
<i>Cedrela odorata</i>	Crude oil	18-47.10 g.kg ⁻¹	245 days	↓ Shoot length and biomass	189
<i>Cyperus brevifolius</i>	Crude oil	10-50 g.kg ⁻¹	6 months	↓ Shoot biomass (dose-dep.)	162
<i>Glycine hyspida</i>	Crude oil	1.3-3.1 g.kg ⁻¹	>6 months	↓ Shoot biomass (dose-dep.)	154
<i>Haematoxylum campechianum</i>	Crude oil	18-47.10 g.kg ⁻¹	245 days	↓ Shoot length and biomass	189
<i>Hordeum vulgare</i>	Crude oil	1.3-3.1 g.kg ⁻¹	>6 months	↓ Shoot biomass	154
<i>Phragmites australis</i>	Crude oil	2-12 g.kg ⁻¹	2 months	↓ Shoot biomass (dose-dep.)	173
<i>Swietenia macrophylla</i>	Crude oil	18-47.10 g.kg ⁻¹	245 days	↓ Shoot length and biomass	189
<i>Tabebuia rosea</i>	Crude oil	18-47.10 g.kg ⁻¹	245 days	↓ Shoot length and biomass	189
<i>Triticum aestivum</i>	Crude oil	1.3-3.1 g.kg ⁻¹	>6 months	↓ Shoot biomass	154
<i>Vicia faba</i>	Crude oil	9-18 g.kg ⁻¹	5 weeks	↓ Shoot length and biomass (dose-dep.)	161
<i>Vicia faba</i>	Crude oil	1.56-50%	30 days	↓ Shoot biomass	194
<i>Zea mays</i>	Crude oil	0.28-0.66%	6 weeks	↓ Shoot biomass	169
<i>Glycine hyspida</i>	Crude oil (spill)	1.1-3.8 g.kg ⁻¹	>6 months	↓ Shoot biomass (dose-dep.)	154
<i>Hordeum vulgare</i>	Crude oil (spill)	1.1-3.8 g.kg ⁻¹	>6 months	Shoot biomass unchanged	154
<i>Triticum aestivum</i>	Crude oil (spill)	1.1-3.8 g.kg ⁻¹	>6 months	↓ Shoot biomass (dose-dep.)	154
<i>Canavalia ensiformis</i>	Diesel	22,219 mg.kg ⁻¹	30 days	↓ Stem and leaf length and biomass	183
<i>Vicia faba</i>	Diesel	9-18 g.kg ⁻¹	5 weeks	↓ Shoot length and biomass (dose-dep.)	161
<i>Glycine hyspida</i>	Drilling fluids	1.6-76.1 g.kg ⁻¹	>6 months	↓ Shoot biomass (dose-dep.)	154
<i>Hordeum vulgare</i>	Drilling fluids	1.6-76.1 g.kg ⁻¹	>6 months	↓ Shoot biomass (dose-dep.)	154
<i>Triticum aestivum</i>	Drilling fluids	1.6-76.1 g.kg ⁻¹	>6 months	↓ Shoot biomass (dose-dep.)	154
<i>Corchorus olitorius</i>	Engine oil	0.2-3%	6 weeks	↓ Shoot length (dose-dep.)	195
<i>Corchorus olitorius</i>	Engine oil	0.2-3%	6 weeks	↓ Leaf area (dose-dep.)	195
<i>Vicia faba</i>	Engine oil	9-18 g.kg ⁻¹	5 weeks	↓ Shoot length and biomass (dose-dep.)	161
<i>Capsicum annum</i>	Lubricating oil	1-5%	84 days	↓ Shoot length (dose-dep.)	188
<i>Capsicum annum</i>	Lubricating oil	1-5%	84 days	↓ Leaf area (dose-dep.)	188
<i>Lycopersicon esculentum</i>	Lubricating oil	1-5%	84 days	↓ Shoot length (dose-dep.)	188
<i>Lycopersicon esculentum</i>	Lubricating oil	1-5%	84 days	↓ Leaf area (dose-dep.)	188
<i>Medicago sativa</i>	Oil sludge	4-5%	9 weeks	↓ Shoot length and biomass unchanged	181
<i>Deschampsia caespitosa</i>	Petroleum cokes	n.a	3 months	↓ Shoot length	196
<i>Triticum aestivum</i>	Petroleum cokes	n.a	2 months	↓ Shoot length	196
<i>Triticum aestivum</i>	Petroleum cokes	n.a	2 months	↓ Leaf area	196
Pigments					
<i>Amorpha fruticosa</i>	Crude oil	5-20 g.kg ⁻¹	6 months	↓ Leaf chlorophyll content (dose-dep.)	193
<i>Cyperus brevifolius</i>	Crude oil	10-50 g.kg ⁻¹	6 months	Light to very dark leaves	162
<i>Cyperus brevifolius</i>	Crude oil	10-50 g.kg ⁻¹	6 months	↓ Leaf chlorophyll content (dose-dep.)	162
<i>Vicia faba</i>	Crude oil	1.56-50%	30 days	↓ Leaf chlorophyll content	194
<i>Vicia faba</i>	Crude oil	1.56-50%	30 days	Leaf carotenoid content unchanged	194
<i>Zea mays</i>	Crude oil	0.28-0.66%	6 weeks	↓ Leaf chlorophyll content	169
<i>Canavalia ensiformis</i>	Diesel	22,219 mg.kg ⁻¹	30 days	↓ Leaf chlorophyll content	183
<i>Canavalia ensiformis</i>	Diesel	22,219 mg.kg ⁻¹	30 days	Leaf discoloration and necrosis	183
<i>Canavalia ensiformis</i>	Diesel	22,219 mg.kg ⁻¹	30 days	↓ Leaf carotenoid content	183
<i>Corchorus olitorius</i>	Engine oil	0.2-3%	6 weeks	↓ Leaf chlorophyll content (dose-dep.)	195
<i>Deschampsia caespitosa</i>	Petroleum cokes	n.a	3 months	↓ Leaf chlorophyll content	196
<i>Deschampsia caespitosa</i>	Petroleum cokes	n.a	3 months	↓ Leaf carotenoid content	196
<i>Triticum aestivum</i>	Petroleum cokes	n.a	2 months	↓ Leaf chlorophyll content	196
<i>Triticum aestivum</i>	Petroleum cokes	n.a	2 months	↓ Leaf carotenoid content	196
Water					
<i>Amorpha fruticosa</i>	Crude oil	5-20 g.kg ⁻¹	6 months	↓ Leaf water content	193
<i>Amorpha fruticosa</i>	Crude oil	5-20 g.kg ⁻¹	6 months	↓ Stomatal conductance and transpiration rate (dose-dep.)	193
<i>Vicia faba</i>	Crude oil	1.56-50%	30 days	↓ Leaf water content	194
<i>Zea mays</i>	Crude oil	0.28-0.66%	6 weeks	↓ Leaf water, osmotic and turgor potentials	169
<i>Corchorus olitorius</i>	Engine oil	0.2-3%	6 weeks	↓ Leaf water content (dose-dep.)	195
<i>Deschampsia caespitosa</i>	Petroleum cokes	n.a	3 months	↓ Transpiration rate and stomatal conductance	196
<i>Triticum aestivum</i>	Petroleum cokes	n.a	2 months	↓ Transpiration rate and stomatal conductance	196

3. Detection of crude oil and petroleum products using vegetation optical properties

Vegetation optical properties have been extensively used for tracking alterations in pigment or water content caused by biotic and abiotic factors^{75,76,197–199} (*e.g.* drought, pests). Conversely, their exploitation in oil seepage and leakage detection has been initiated very recently^{77,185,186,200,201}. Major progress have been made in this field by taking advantage of multi- and hyperspectral methods developed for assessing vegetation health in other contexts, such as crop and ecosystem monitoring. Some of them – especially VI – proved to be efficient for tracking oil-induced alterations in vegetation reflectance under controlled conditions, from spectroradiometer-acquired reflectance data^{186,200}. However, serious limits remain as to apply these methods on airborne or satellite imagery.

This section focuses on the modifications of vegetation optical properties induced by crude oil and petroleum products. The methods developed to detect these modifications are described as well as those from other context of vegetation stress monitoring, which were exploited in this thesis. Then, the section discusses the application of these methods for oil seepage and leakage detection on airborne and satellite imagery. More precisely, this section highlights the current limits of these methods, from which the approach proposed in this thesis resulted.

3.1. Effects of crude oil and petroleum products on vegetation optical properties measured under controlled conditions

As described above, crude oil and petroleum products affect the main biophysical and biochemical parameters driving vegetation optical properties. These effects suggest modifications in the spectral signature at leaf and canopy scales, which have been studied under green-house-controlled or field-semi-controlled conditions. They are summarized here.

3.1.1. Effects on reflectance in the visible region (400 – 750 nm)

The VIS has been mostly exploited for tracking the effects of crude oil and petroleum products from the spectral signature of vegetation, because of its strong link with pigments^{159,185,186,200}. The alteration of chlorophyll content described in the previous section immediately leads to an increase of reflectance in this region, at leaf and canopy scales (Table 3). This increase is essentially located in the green-red wavelengths (500 – 670 nm), where it can reach 20% (Figure 21), and is expressed as a shift of the REP toward shorter wavelengths around 700 nm. In comparison, the blue wavelengths (400 – 500 nm) are little affected. This response is observed after few days of exposure – even at low concentration – and becomes more pronounced in time, making crude oil and petroleum products more easily detectable. Once again, it is difficult to identify the most contributing hydrocarbons and HMs, since a single of these compounds is able to induce a similar response^{166,202,203}.

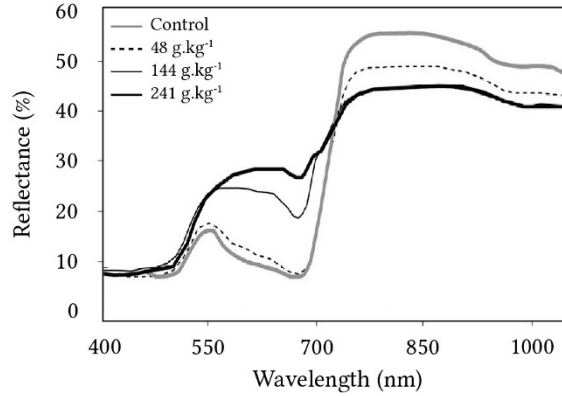


Figure 21. Spectral signatures of leaves of *Zea mays* L. grown for 14 days on engine oil-contaminated (48 – 214 g.kg⁻¹) or uncontaminated soils (modified from ¹⁸⁶).

Although the increase of green-red reflectance and the shift of the REP are frequent, an absence of reflectance change has been more rarely observed in studies^{29,159,204}. This underlines the variability of vegetation response to crude oil and petroleum products discussed in the previous section. Of the mentioned studies, only few attempted to clearly link the level of pigment content alteration to that of reflectance in the VIS²⁰⁰. They focused on leaf chlorophyll content, because of its major influence on reflectance in the 500 – 670 nm wavelengths²⁰⁰. Sanches *et al.*^{77,185} conducted an experiment on four species exposed to gasoline and diesel and concluded that carotenoid and anthocyanin contents had only few contribution to reflectance changes in the VIS, despite their implication in vegetation response to stress. Once again, this cannot be generalized to all contexts, since the species studied were particularly sensitive to petroleum products.

3.1.2. Effects on reflectance in the near-infrared region (750 – 1300 nm)

As described in section 1.2, reflectance in the NIR is highly dependent on the species – especially mono- and dicotyledonous – and on the acquisition scale (leaf, canopy). The same factors, as well as mixture composition, lead to contrasted response of vegetation in this region (Table 3). Whether they result from an increase or a decrease of reflectance, differences between healthy and affected vegetation can exceed 20% in the NIR⁷⁷ (Figure 22a-d). A decrease in reflectance is more likely to be observed at canopy scale, since plant development – and thus LAI – is strongly limited by hydrocarbons and HMs. However, several exceptions have been noticed in the literature. As pointed out by Rosso *et al.*¹⁵⁹ and Sanches *et al.*⁷⁷, a single species can undergo opposite reflectance changes in the NIR, depending on the crude oil or petroleum product to which it is exposed (Figure 22b, d). Likewise, two species exposed to a similar concentration of the same petroleum product can exhibit contrasted responses in this region²⁰¹. This causes serious detection limits in regions with high species diversity. Similar observations have been made at leaf scale (Table 3), where reflectance in the NIR mainly depends on anatomy. However, no study demonstrated the relationship between alterations of parenchyma and reflectance changes in this region.

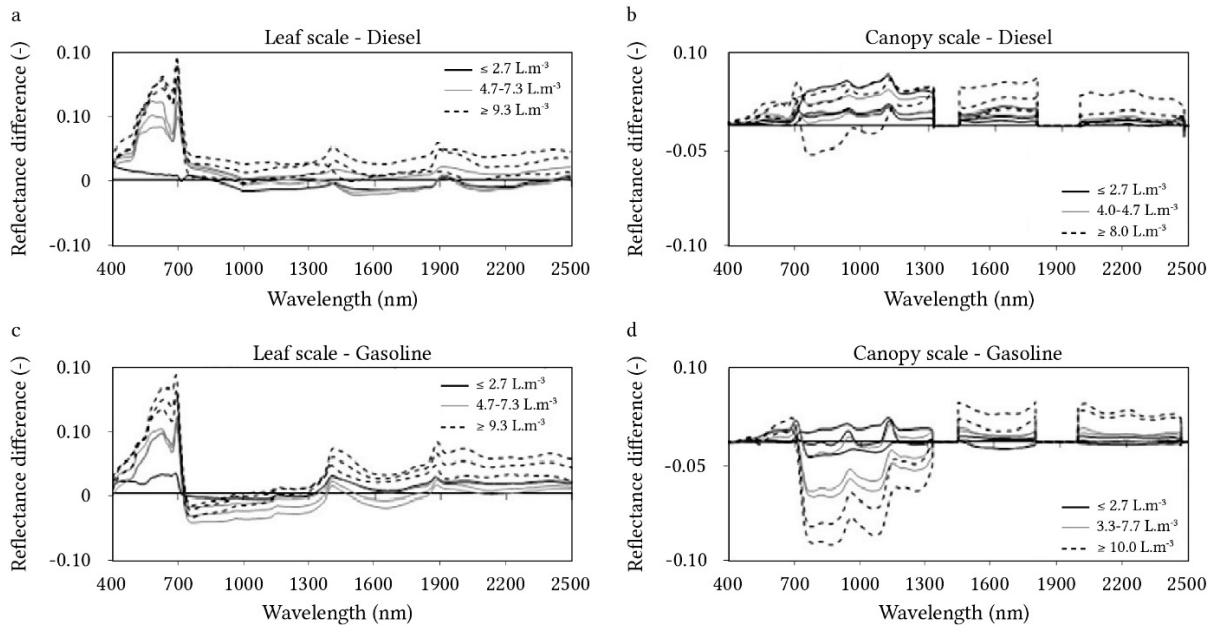


Figure 22. Reflectance differences between *Brachiaria brizantha* H.S. plants grown on either (a-b) diesel- or (c-d) gasoline-contaminated and uncontaminated soil, at (a, c) leaf and (b, d) canopy scales. Reflectance differences were obtained by subtracting the spectral signature of uncontaminated plants to those from diesel- or gasoline-exposed plants. Curves represent different contamination levels (modified from ⁷⁷).

3.1.3. Effects on reflectance in the short-wave infrared region (1300 – 2500 nm)

Because of modifications in vegetation water status, the SWIR is largely impacted by exposure to crude oil and petroleum products. However, this region has been rarely exploited in previous studies. As well as in the NIR, the response of vegetation in the SWIR varies among studies (Table 3). In the case where a decrease of reflectance is observed on exposed vegetation, it remains lower than 10% ²⁰⁵. Conversely, an increase of reflectance, which is more consistent with the reduction of leaf water content and canopy LAI, can exceed 20% for the most oil-sensitive species. In both cases, the response appears later than in the VIS and is thus a good indicator of a long-term exposure. As expected, the most affected wavelengths are located in water absorption features²⁰⁰. Because of strong atmospheric effects (low signal-to-noise ratio, SNR), these features are however unusable at canopy – and image – scale, but other ones (*e.g.*: 1600 and 2200 nm) proved to be good alternatives⁷⁷. This response of vegetation to crude oil and petroleum products is less pronounced than that induced by a “natural” water-deficit stress²⁰⁶, which suggests being able to differentiate these two stressors using the SWIR.

Vegetation reflectance in the SWIR also depends on celluloses, hemicelluloses, lignins and proteins, which have already been reported as slightly sensitive to petroleum products in one study⁷⁷. Because of the strong influence of LWC in this regions, it is unlikely that alterations in these biochemical compounds have major contribution to the modifications of reflectance described here.

Table 3. Literature review of the effects induced by crude oil and petroleum products on vegetation reflectance in the different spectral regions, at leaf and canopy scales. This review only includes studies carried out under greenhouse-controlled or field-semi-controlled conditions. (\uparrow and \downarrow denotes reflectance increase and decrease, respectively; FC: Field capacity; dose-dep: dose-dependent effects; n.a.: not available; n.s.: non-significant effect.)

Species	Crude oil / petroleum product	TPH Concentration	Total time of exposure	VIS (350 - 750 nm)	NIR (750 - 1300 nm)	SWIR (1300 - 2500 nm)	Ref.
Leaf							
<i>Brachiaria brizantha</i>	Diesel	12.7 L.m ⁻³	30 days	\uparrow (dose-dep)	\uparrow (dose-dep)	\uparrow (dose-dep)	77
<i>Brachiaria brizantha</i>	Gasoline	12.7 L.m ⁻³	30 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\uparrow (dose-dep)	77
<i>Neonotonia wightii</i>	Diesel	6.25 L.m ⁻³	184 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205
<i>Neonotonia wightii</i>	Gasoline	6.25 L.m ⁻³	184 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205
<i>Rubus fruticosus</i>	Mud pit	4-40 g.kg ⁻¹	100 days	\uparrow	\uparrow	\uparrow	29
<i>Salicornia virginica</i>	Alba ¹ crude oil	7.7-9.1 %	32 days	n.s.	\uparrow (high dose)	n.a.	159
<i>Salicornia virginica</i>	Escravos ¹ crude oil	0.7-1.4 %	32 days	\downarrow (high dose)	\downarrow	n.a.	159
<i>Triticum sp.</i>	Gasoline	10-100 ml.kg ⁻¹	106 days	n.s.	n.s.	n.s.	204
<i>Zea mays</i>	Diesel	6.25 L.m ⁻³	184 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205
<i>Zea mays</i>	Engine oil	48-241 g.kg ⁻¹	14 days	\uparrow (dose-dep)	\downarrow (dose-dep)	n.a.	186
<i>Zea mays</i>	Gasoline	6.25 L.m ⁻³	184 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205
<i>Zea mays</i>	Mud pit	4-40 g.kg ⁻¹	100 days	n.s.	\downarrow	\uparrow	29
Canopy							
<i>Brachiaria brizantha</i>	Diesel	12.7 L.m ⁻³	30 days	\uparrow (dose-dep)	\uparrow (dose-dep)	\uparrow (dose-dep)	77
<i>Brachiaria brizantha</i>	Gasoline	12.7 L.m ⁻³	30 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\uparrow (dose-dep)	77
<i>Forsythia suspensa</i>	Engine oil	20-60 % soil FC	28 days	\uparrow (dose-dep)	\uparrow	n.a.	201
<i>Neonotonia wightii</i>	Diesel	8.33 L.m ⁻³	203 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205
<i>Neonotonia wightii</i>	Gasoline	8.33 L.m ⁻³	203 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205
<i>Pennisetum alopecuroides</i>	Engine oil	20-60 % soil FC	28 days	\uparrow (dose-dep)	\downarrow (dose-dep)	n.a.	201
<i>Zea mays</i>	Diesel	8.33 L.m ⁻³	203 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205
<i>Zea mays</i>	Engine oil	96 g.kg ⁻¹	20 days	\uparrow	\uparrow	\uparrow	206
<i>Zea mays</i>	Gasoline	8.33 L.m ⁻³	203 days	\uparrow (dose-dep)	\downarrow (dose-dep)	\downarrow (dose-dep)	205

3.2. Methods developed for detecting crude oil and petroleum products under controlled conditions

As pointed out in Table 3, only few experiments have been carried out to characterize the spectral response of vegetation to crude oil and petroleum products. This gave rise to a limited number of methods developed for detection purpose. These methods are based on exploiting the modifications of reflectance induced by crude oil and petroleum products and described above. They are discussed in this section.

Most of existing methods rely on visual or statistical comparisons of spectral signatures between healthy and oil-exposed vegetation^{29,159}. These methods are however not suitable for application beyond the context studied. Other authors exploited reflectance at particular wavelengths by using VI, REP and spectrum transformations, and converged on the critical importance of VIS wavelengths^{77,186,200} (Figure 23). Gürtler *et al.*²⁰⁵ compared these methods for discriminating among healthy and gasoline- or diesel-exposed vegetation, at leaf and canopy scales, and concluded that their performance depends on the species. In the same experiment, Sanches *et al.*⁷⁷ combined first derivative and *continuum* removal spectra transformation to Principal Component Analysis (PCA) for similar purpose and identified the red-edge region as a good indicator of soil contamination. Strictly speaking, none of the above-mentioned methods aimed to predict whether vegetation is – or has been – exposed to crude oil or petroleum products from its spectral signature. This represents an important gap to fill for detecting seepages or contamination under natural conditions without *a priori* knowledge about their presence. In addition, studies listed in Table 3 showed that mixture composition and concentration strongly

influence the amplitude of reflectance modifications observed in the whole spectral signature. This suggests being able to both characterize oil (*i.e.* crude oil or petroleum product type) and quantify the contamination level (*i.e.* TPH concentration), which are two major pieces of information in contamination monitoring. Once again, no method has been developed yet for this purpose.

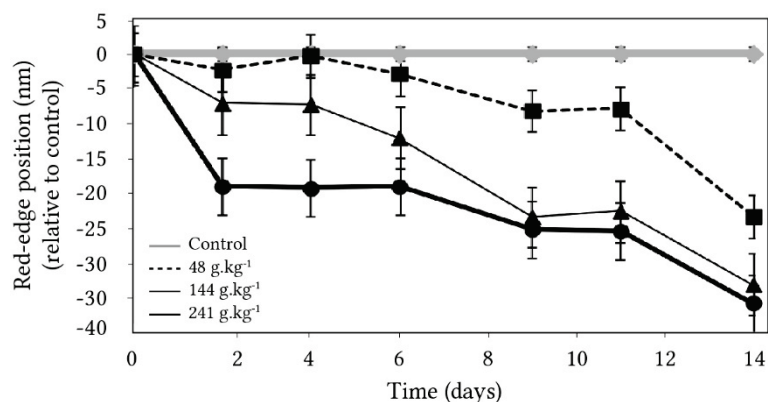


Figure 23. Temporal evolution of the Red-Edge Position (REP) of *Zea mays* L. grown on engine oil-contaminated (48 – 214 g.kg⁻¹) or uncontaminated soils (modified from ¹⁸⁶).

To ensure reliable detection, characterization and quantification of crude oil and petroleum products, such method must be adapted to its context of application; more precisely, to the targeted species' sensitivity. It is unlikely that a single method could be applied to any species, because of the variability in sensitivity to hydrocarbons and HMs (Table 2). For that reason, the existing studies are strongly limited, since almost all of them focused on crop species (Table 3). These species are particularly sensitive to oil and thus unrepresentative of those found on seepage sites and industrial facilities, for which little effects are expected on the spectral signature^{20,29}. Therefore, new methods are required for detecting subtle biophysical and biochemical alterations in oil-tolerant species using hyperspectral data. One possible way would be to draw on methods developed in other contexts of vegetation stress assessment.

3.3. Other methods used in vegetation stress assessment

VI, REP and spectrum transformations have been used for assessing stress-induced alterations in vegetation health in a wide range of contexts. For example, Filella & Peñuelas⁹⁵ used reflectance derived in the red-edge for tracking changes in LCC and LAI of *Capsicum annuum* and *Phaseolus vulgaris* resulting from nitrogen deficiencies. Likewise, VI exploiting the NIR and SWIR regions succeeded in tracking water-stress caused by insufficient irrigation or pests^{140,207}. When used in classification or simple and multiple regression methods, these reflectance data allow predicting stressed vegetation and quantifying biophysical and biochemical parameters automatically^{74,208–210}. Therefore, they could be used for detecting crude oil and petroleum products and quantifying TPH indirectly. The aim here is not to detail the mathematical formulation of classification and regression methods, which can be easily found in recent reviews, but to discuss their advantages and drawbacks regarding vegetation stress assessment.

Classification relies on the combination of several discrete or continuous variables (*e.g.* reflectance data, VI) to predict a categorical response variable (*e.g.* “healthy” or “stressed”) through a mathematical function²¹¹. In this thesis, we only consider supervised classification. In most cases, these methods are first calibrated on a set of data with known categories, called the *training set*, and tested on an independent set – the *test set* – by predicting categories and comparing them to the true ones²¹². Numerous classification methods have been proposed in the literature^{213–215}. When dealing with hyperspectral data, several constraints yet arise. Since reflectance is measured over multiple and contiguous wavelengths, it is not rare to have more variables than observations (*i.e.* reflectance wavelengths > sample size). This phenomenon, known as the Hughes’ effect²¹⁶, leads to overfitting of the training set, which negatively affects classification accuracy. This effect can be partly avoided by reducing the dimensionality of the variables using, for example, Principal Component Analysis (PCA). Focusing on vegetation studies, Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), Random Forest (RF) and Spectral-Angle-Mapper-based classification (SAM) revealed to be particularly efficient for discriminating healthy and stressed categories while avoiding overfitting^{74,210,217,218} (Figure 24a-b). However, an exhaustive review proposed by Lowe *et al.*¹⁹⁸ stated that no consensus is made in the choice of the method, since their performance highly depends on the purpose of the classification. Few methods allow identifying the most important (*i.e.* discriminant) variables through weighting or stepwise selection criteria^{217,219}. Stepwise Forward LDA²²⁰ has been specifically designed for this purpose, but remains poorly adapted to hyperspectral data because of high multicollinearity. In a spectral region, multicollinearity occurs when reflectance data are linear combination of each other²²¹. For example, correlation coefficients (r) among reflectance data from different red wavelengths can easily exceed 0.8, which indicates high redundancy. Variable selection becomes very difficult in this case. To achieve it, penalized methods have been developed, such as the Elastic net²²², but remain unexploited in vegetation studies. They are developed in this thesis.

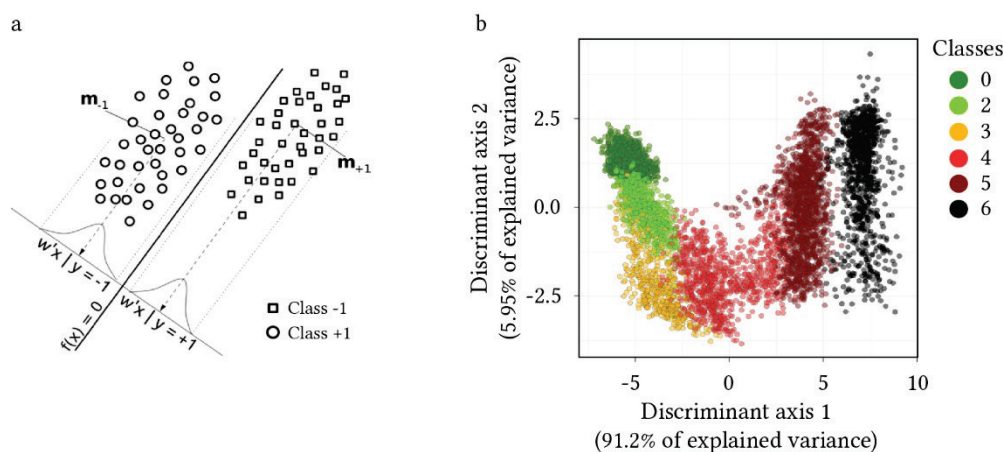


Figure 24. (a) Graphical representation of the Linear Discriminant Analysis (LDA) principle²²³. This method relies on the linear combination of variables into a reduced projection space for separating classes. (b) Application of LDA to the discrimination of fungal disease severity (“Classes”) in *Musa* AAA cv. William using hyperspectral data. The first two discriminant axis are represented²²⁴.

Regression methods are used to predict a continuous response variable from one (simple regression) or several (multiple regression) continuous input variables²²⁵. In practice, these

methods follow the same calibration – test procedure than for classification. Simple regression relies on the calibration of univariate models (*e.g.* polynomial, exponential, etc.). It has been especially used for predicting LCC from single VI in previous studies^{78,81,95}. Multiple regression regroups a wide range of methods that do not substantially differ from classification ones, and are constrained by the same overfitting and multicollinearity issues. Regarding vegetation studies, it has been shown that Stepwise LDA, Partial Least Square Regression (PLSR) and Support Vector Regression (SVR) are well-suited for retrieving biophysical and biochemical parameters from hyperspectral data^{208,209,226}. Once again, there is not preferable method since the performance varies according to the context.

Since both classification and regression methods perform well for detecting and quantifying stress-induced changes in vegetation health, they are promising for application in the context of the thesis. Other methods have been developed for the same purpose in the literature and could help achieving the detection of crude oil and petroleum products and the quantification of TPH. Among them, RTM are of great interest. RTM are physically-based methods aiming to model vegetation optical properties. They are typically classified in four categories: plate models, N-flux models, stochastic models and ray tracing models^{113,227}. Focusing at leaf scale, the plate model PROSPECT is probably the most widespread^{70,73}. In its direct mode, PROSPECT allows simulating leaf optical properties (reflectance and transmittance) in the optical reflective domain from its biophysical and biochemical parameters (structure and pigment, water and dry matter contents). Inversion of the model allows retrieving these parameters from reflectance and transmittance measurements performed on leaves¹⁴⁵. PROSPECT has been used in many studies dealing with environmental monitoring purposes^{73,228}. While LCC and LWC remain the most targeted parameters in vegetation stress assessment^{139,229}, recent improvements of the model allow separating chlorophylls, carotenoids and anthocyanins with good precision^{82,108}. In a recent study, Arellano *et al.*²³⁰ inverted the model to compare LCC of various tropical plant families among uncontaminated and oil-spill sites, and found significant alterations for some of them. This opened up promising perspectives in the context of the thesis.

Hence, although there is no method specifically developed for detecting and quantifying crude oil and petroleum products under controlled conditions, lots of ones inherited from other vegetation studies look promising. On that point, adapted methods are definitely needed for ensuring – at least – reliable detection under natural conditions at large scale, from airborne and satellite imagery. Simple methods using REP or VI have already been tested on multi- and hyperspectral images in various cases of oil detection (seepages and leakages), but suffered from serious limits. They are discussed here.

3.4. Application in oil exploration and contamination monitoring using airborne and satellite multi- and hyperspectral imagery

Few attempts have been made in detecting oil seepages or leakages in vegetated areas using remote sensing in the past (Table 4). In most cases, studies aimed to assess the impact of oil leakages on the environment using multi- (Landsat) or hyperspectral (Hyperion) satellite imagery at 30-m spatial resolution^{35,231–233}. In comparison, seepage detection only sparked a limited interest²⁰. A limited number of authors have used airborne hyperspectral images, and those who did have rarely exploited the entire spectral signature of vegetation (Table 4). Almost all the mentioned studies used REP or VI to detect changes in vegetation health induced by crude oil or petroleum products. As for experiments carried out under controlled conditions, these methods rely on mean comparison between sites with healthy and oil-exposed vegetation^{68,231,234}. They

proved to be efficient for identifying vegetation stress on seepage or leakage sites, but suffered from serious limits when applied outside the study area (Figure 25a-c).

Table 4. Review of studies on crude oil and petroleum products detection from multi- and hyperspectral airborne and satellite images. (Refl.: Reflectance; VI: Vegetation Indices; CR: *Continuum* Removal; RF: Random Forest; REP: Red-Edge Position; comp.: Comparison.)

Vegetation type	Target	Sensor name	Sensor type (resolution)	Bands (spectral domain)	Method	Ref.
Multispectral						
Mangrove	Crude oil leakage	Landsat-8	Satellite (30 m)	9 (435 - 2294 nm)	VI + Mean comp.	231
Crops, grassland & trees	Crude oil leakage	Landsat-8	Satellite (30 m)	9 (435 - 2294 nm)	VI + RF classification	232
Mangrove	Crude oil leakage	Landsat-5 & -7	Satellite (30 m)	6 (450 - 2350 nm)	VI + Simple regression	235
Mangrove	Crude oil leakage	Landsat-5 & -7	Satellite (30 m)	6 (450 - 2350 nm)	VI + Mean comp.	235
Hyperspectral						
Wetland	Crude oil leakage	AISA	Airborne (1.5 m)	286 (400 - 2400 nm)	Reflectance + Classification	236
Crops	Benzene pipeline leak	HyMap	Airborne (4 m)	128 (436 - 2485 nm)	REP & VI + Spatial filter	35
Mediterranean grassland	Crude oil microseepage	Probe-1	Airborne (8 m)	128 (436 - 2480 nm)	REP & VI + Spatial filter	20
Tropical forest	Crude oil leakage	Hyperion	Satellite (30 m)	242 (400 - 2500 nm)	VI + Threshold	68
Plain & rainforest	Crude oil leakage	Hyperion	Satellite (30 m)	242 (400 - 2500 nm)	CR + Mean comp.	234
Plain & rainforest	Crude oil leakage	Hyperion	Satellite (30 m)	242 (400 - 2500 nm)	Refl. & VI + Mean comp.	234

In contrast to experimental studies, REP and VI have already been exploited in classification or anomaly detection methods on multi- and hyperspectral images. However, the performance of these methods has been rarely quantified; their evaluation mostly relied on visual interpretation of detection mapping without ground validation. Among notable examples, Ozigis *et al.*²³² combined 10 VI in RF on 30-m resolution Landsat-8 images for detecting oil leakages and obtained an overall accuracy of maximum 70% on selected sites. Conversely, Arellano *et al.*⁶⁸ applied successive vegetation index thresholds to map oil-induced stress near production facilities using 30-m resolution Hyperion images, but did not quantify the detection accuracy. These methods were first calibrated on a study area, and then applied to the entire image. In all cases, they led to the apparition of false alarms, especially false positives (*i.e.* vegetation stress not induced by petroleum hydrocarbons and HMs). This phenomenon is observed under various contexts (*e.g.* temperate, tropical) and results from multiple factors. First, in most studies, the spatial resolution of the images was not adapted to the size of the target. In addition, as described in section 2.3.2, certain species are particularly tolerant to crude oil and petroleum products and undergo only little changes in their spectral signature, which make them difficult to discriminate from healthy vegetation. In that situation, high spectral resolution and SNR are needed to catch these changes in reflectance, so hyperspectral sensors are required. In addition, natural differences in optical properties among species and individuals – as well as in sensitivity to oil – make the detection particularly challenging in area with high species diversity. For instance, a species exposed to crude oil or petroleum products may exhibit a similar spectral signature than that of another unexposed species. This becomes a serious issue at decametric spatial resolution, where species are highly mixed inside pixels. A similar issue arises when exposed species are mixed with bare soil. Very high spatial resolution (1 – 2 m) is thus needed to overcome these limits.

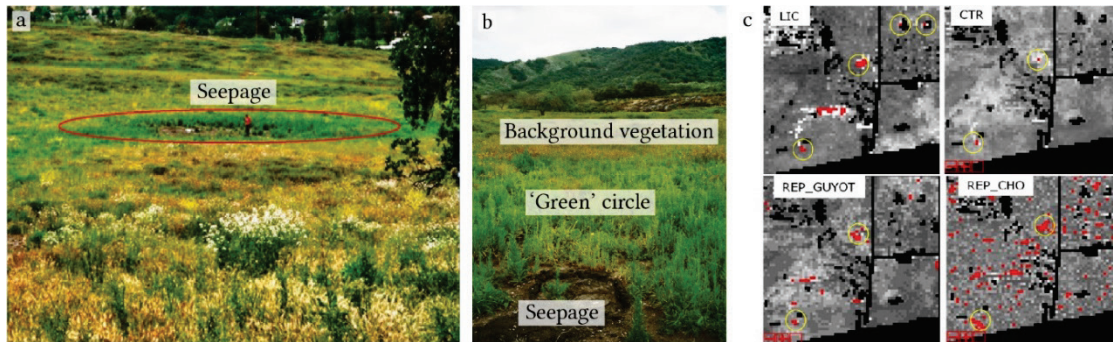


Figure 25. (a) Crude oil seepage in vegetated area. (b) The seepage is characterized by bare ground mixed with crude oil at the center, surrounded by a particular vegetation distribution pattern, which allows being detected from hyperspectral airborne images using (c) vegetation indices and a spatial filter. However, false alarms (red pixels outside yellow circles) cannot be avoided while simultaneously detecting all seepages²⁰.

As pointed out by Ozigis *et al.*²³², several sources of confusion also contribute to increasing false positives. It is obvious that under natural conditions, crude oil and petroleum products are not the only factor affecting vegetation health and optical properties. Some biotic or abiotic factors are likely to induce similar effects, thus introducing confusion. As described in section 2.2.1, crude oil and petroleum products reduce water availability for plants and can induce a water-deficit stress. Under natural conditions, this effect can be easily confused with that of a “natural” water-deficit (*i.e.* resulting from insufficient precipitation and/or highly-drained soils). Although it seems possible to discriminate these stressors for highly oil-sensitive species under controlled conditions^{186,200}, it is of another difficulty for oil-tolerant species and using airborne or satellite hyperspectral images. Stress confusion has been identified as one of the most important cause of misclassification in previous studies. It is therefore necessary to account for these sources of confusion in a given context, when applying detection methods over an entire region. Once again, very high spatial and spectral resolutions are needed to achieve efficient discrimination of oil and other stressors. Although no current satellite-embedded sensor offers such resolutions simultaneously, airborne imagery represents a good alternative²³⁶.

4. Synthesis based on previous studies

As highlighted in this section, the detection of crude oil and petroleum products from vegetation optical properties is a recent topic that requires further research before it can be applied with reliability to hyperspectral images. Previous studies allowed identifying the most influencing factors on vegetation spectral response and its detectability. In this way, they emphasized the necessity of taking the context into account and the importance of very high spatial and spectral resolutions. Although no method of detection and quantification have been proposed, promising perspectives arise from vegetation studies in the literature.

As concluded from previous studies, it is not the best option to develop detection methods directly from airborne or satellite images, especially without solid knowledge about the context (species' sensitivity, hydrocarbon and HM mixture, other potential stressors). As a first step, experiments carried out under controlled conditions are necessary, since they help determining the response of vegetation specifically induced by crude oil and petroleum products. However, a missing link remains between image-based and experimental studies in the literature. In order to

support imagery application, experiments have to meet essential criteria. They have to be representative of realistic field conditions (*i.e.* to the targeted context) and serve as basis for developing methods that are suitable for use on images (*e.g.* classification or regression). The upscaling of methods is the most important difficulty in this approach, so it is crucial to address it progressively; for example, from leaf to canopy scales and finally on images. The validation of the methods in the field is an intermediate – and critical – step prior to imagery application. Then, the methods should be progressively applied to imagery; first, on selected sites with known species' sensitivities, and thereafter at large scale.

To date, no study has followed a multiscale approach in the context of crude oil and petroleum product detection. Conversely, several studies emphasized its pertinence for assessing vegetation health in other contexts (*e.g.* agriculture, ecology)^{237–239}. VI and RTM show great advantages in the upscaling process. Some pigment-related indices are little affected by soil or canopy structure and are thus robust indicators of vegetation health from field to images^{240,241}. Likewise, leaf RTM can be adapted to field-based canopy measurements and airborne or satellite images for estimating vegetation biophysical and biochemical parameters at large scale. For this purpose, PROSPECT has been frequently coupled with the SAIL model – forming PROSAIL, which takes canopy LAI and structure, background soil reflectance and viewing geometry into account^{73,242}.

CONCLUSION

This chapter outlines the current knowledge in vegetation optical properties in the reflective domain, as well as their modifications induced by exposure to crude oil and petroleum products, which are exploited for detecting oil seepages and leakages from multi- and hyperspectral airborne and satellite imagery. There is a close link between vegetation optical properties and biophysical and biochemical parameters (anatomy, pigment and water content) in the different spectral regions (VIS, NIR and SWIR), which provides a singular shape to the spectral signature of leaves and canopies. Hydrocarbon and HM mixtures are likely to affect all these parameters and thus vegetation reflectance indirectly and directly, through accumulation in tissues and alterations in water and nutrient availability and plant uptake capacities. However, many factors enter into account, such as species' sensitivity or mixture composition. This represents a first difficulty to oil detection from images, which can be limited by faithfully reproducing the context of study under controlled conditions. Greenhouse experiments are particularly appropriate for this purpose. They are intended to (I) characterize the biophysical – biochemical and spectral responses of the identified species to the hydrocarbon and HM mixture of interest (crude oil or petroleum product), and (II) to develop oil detection and TPH quantification methods that are adapted to application on hyperspectral images with very high spatial and spectral resolutions, under natural conditions. Based on previous studies, spectra transformations and VI seem suitable to meet the second criteria if coupled to classification and regression. RTM have also been identified as promising. In a second time, the validation of methods in the field (*i.e.* on proximal reflectance measurements) is mandatory before image applications. With the view of operational use, the methods should allow detection and quantification outside the study area – but still in the same context. To achieve it, several difficulties must be overcome at image scale, such as confusion with other vegetation stressors. As highlighted in the last section, very high spatial and spectral resolutions are necessary.

The proposed three-step approach (greenhouse – field – images) derives from the current limits of the literature, and has thus never been adopted in previous study. This approach is developed in this thesis. The following sections formulate the scientific problem of the thesis, and describe the context of study and the research carried out at each step of the proposed approach.

SCIENTIFIC PROBLEM AND ADOPTED APPROACH

To date, there is only little evidence that crude oil and petroleum products can be detected and quantified from airborne or satellite images, by exploiting vegetation optical properties. This thesis aims to achieve it from airborne hyperspectral images with very high spatial and spectral resolutions, thanks to new methods developed under experimental and field conditions. For this purpose, the thesis focuses on vegetated mud pits, a particular case of soil contamination deriving from oil production. Mud pits can be defined as moderately large areas (50 – 5000 m²), where residues – typically oil sludge – have been accumulated near the surface during crude oil extraction and refining. These residues are often mixed with the soil matrix and contain high concentrations of TPH and HM. Mud pits sometimes remain after production cessation, so the contamination persists and represent an important environmental issue. They become progressively vegetated by oil-tolerant species over time. Therefore, mud pits were particularly interesting in the frame of this thesis. In this manuscript, they are also referred as *industrial brownfields*.

In this thesis, mud pits were studied under temperate and tropical contexts. According to conclusions of the first chapter, a three-step approach has been adopted to detect soil contamination and quantify TPH from vegetation optical properties (Figure 26). The first one relied on experiments carried out under controlled conditions, which aimed to characterize the sensitivity of the species to mud pit contamination and to develop the methods. A first feasibility study has been conducted for both context, and was completed by more important ones. These experiments are presented in the second chapter. Then, the developed methods were tested under natural conditions at field and image scales, in the temperate context. This forms the third chapter on the manuscript. All the methods were applied to airborne hyperspectral images. At each step of the approach, the performance of the methods were evaluated and compared to those of the previous step.

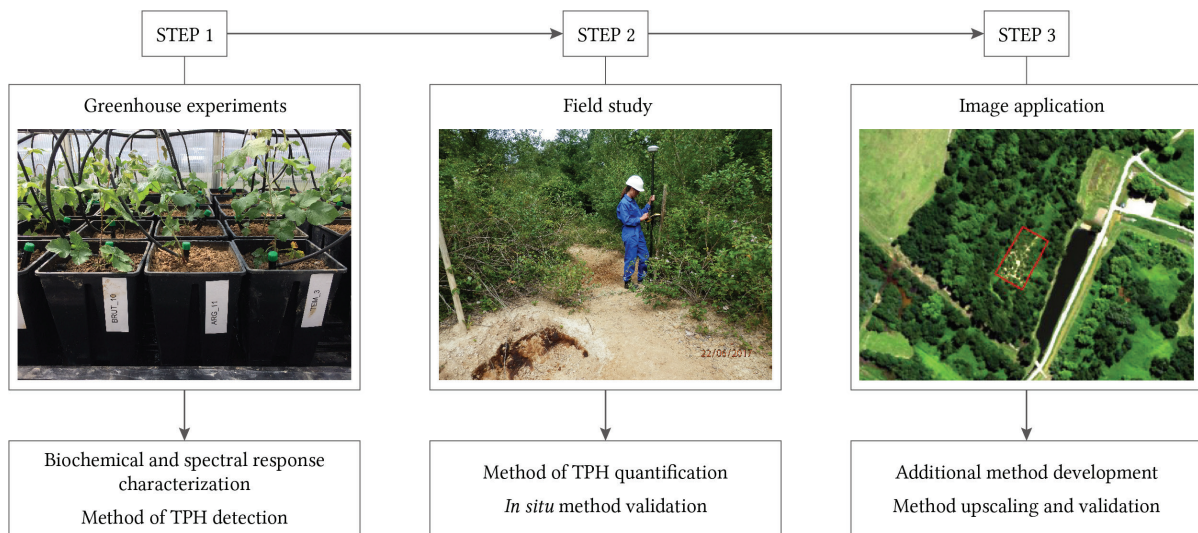


Figure 26. Flowchart of the adopted approach. The first step is presented in the chapter two and the second and third steps are presented in the chapter three.

CHAPTER TWO: DEVELOPMENT OF METHODS FOR DETECTING AND QUANTIFYING OIL CONTAMINATION BASED ON VEGETATION OPTICAL PROPERTIES, UNDER CONTROLLED CONDITIONS

INTRODUCTION

The first step of the approach adopted in this thesis aimed at developing methods for assessing oil contamination based on vegetation optical properties, under controlled conditions. This included the detection of the contamination, the identification of the oil type and the quantification of TPH, which is the main constituent of crude oil and petroleum products. The methods were developed in order to be applicable to airborne hyperspectral imagery. To achieve this, three experiments were carried out under greenhouse, by reproducing the contexts of mud pit contamination identified under natural conditions, in temperate and tropical regions. Each of them forms a section of this chapter. Two of these experiments were published and the third one was submitted in a peer-review journal. All these experiments relied on exposing temperate or tropical species to oil-contaminated soils and other growing conditions. Their reflectances were measured periodically in the optical reflective domain at increasing acquisition scales and linked to biochemical analyses. Because these species had never been studied before, the first experiment assessed their suitability for developing the methods. This experiment addressed the detection of the contamination and its discrimination with a water-deficit stress, which may induce similar effects on vegetation health. It helped identifying the most important spectral regions to focus on for developing the methods. Two more consequent experiments were then conducted under either temperate or tropical conditions. They followed the same measurement protocol and overall procedure based on vegetation indices, but served for developing methods that differed in their purpose. The first one, developed under temperate context, aimed to detect oil contamination and to characterize its type qualitatively (*i.e.* crude oil, mud pit). The risk of confusion with other stressors was once again assessed by introducing a water-deficit treatment. Conversely, the method developed under tropical context exploited VI for quantifying TPH. A single type of oil contamination was applied to the species, this time in various concentrations in mud pit soils. The procedure carried out for developing these methods first consisted in identifying the visible stress symptoms and linking the biochemical and spectral responses induced on the long-term by crude oil or petroleum products. VI and multiple regression were particularly useful to achieve it, thanks to their respective advantages presented in chapter one and discussed here. These indices were then exploited in classification or multiple regression for developing the methods of oil detection and quantification. Both methods helped approaching a detection limit of the contamination, in terms of TPH concentration. They were assessed at the different acquisition

scales, and in the field for the temperate context. This provided an indication of their suitability for application to airborne hyperspectral imagery under natural conditions, which is developed in chapter three.

5. First article: “Assessing soil contamination due to oil and gas production using vegetation hyperspectral reflectance”

Article reference:

Lassalle, G.; Credo, A.; Hédacq, R.; Fabre, S.; Dubucq, D.; Elger, A. Assessing Soil Contamination Due to Oil and Gas Production Using Vegetation Hyperspectral Reflectance. *Environ. Sci. Technol.* **2018**, 52 (4), 1756–1764. <https://doi.org/10.1021/acs.est.7b04618>.

Abstract

The remote assessment of soil contamination remains difficult in vegetated areas. Recent advances in hyperspectral spectroscopy suggest making use of plant reflectance to monitor oil and gas leakage from industrial facilities. However, knowledge about plant response to oil contamination is still limited, so only very few imaging applications are possible at this stage. We therefore conducted a greenhouse experiment on three species long-term exposed to either oil-contaminated or water-deficient soils. Reflectance measurements were regularly performed at leaf and plant scale over 61 days of exposure. Results showed an increase of reflectance in the visible (VIS), the red-edge and the short-wave infrared (SWIR) under both oil and water-deficit stress exposure. A contrasted response in the near-infrared (NIR) was also observed among species. Spectra underwent transformations to discriminate species' responses to the different treatments using linear discriminant analysis (LDA) with a stepwise procedure. Original and transformed spectra enabled to discriminate the plants' responses to the different treatments without confusion after 61 days. The discriminating wavelengths were consistent with the spectral differences observed. These results suggest differential changes in plant pigments, structure and water content as a response to various stressors, and open up promising perspectives for airborne and satellite applications.

5.1. Introduction

Among recent techniques developed to monitor oil and gas exploration and production activities, the use of hyperspectral spectroscopy for risk assessment has raised particular interest in during the last decade^{243,244}. Hyperspectral spectroscopy gives information about the composition of surfaces based on their reflectance and particular light absorption features over multiple wavelengths, from the visible to the infrared. Compared to other time-consuming and costly techniques, this one enables remote detection of chemical compounds of natural and anthropogenic origin present in soils, such as hydrocarbons and heavy metals resulting from oil and gas facility failures (exploration and production units, pipelines)^{61,243,245}. However, this technique remains unexploited in densely vegetated regions, where light penetration is severely limited by foliage. Very few radiations reach ground level and it is impossible to remotely detect compounds directly from soil spectra. An alternative approach proposed by recent studies aims to use vegetation as a spectral bioindicator of soil contaminated by oil and gas^{20,68,77}.

The main parameters involved in vegetation reflectance differ in the visible (VIS, 350-750 nm), the near-infrared (NIR, 750-1300 nm) and the short-wave infrared (SWIR, 1300-2500 nm). The optical properties of leaves are essentially determined by pigment composition in the VIS^{78,113}, parenchyma structure and cell morphology in the NIR^{71,119}, and water content in the NIR and SWIR^{134,246}. Consequently, vegetation health can be assessed through its reflectance spectra, which can be modified by soil contaminants that have an impact on plant biochemistry and anatomy^{29,76,77,203}.

Crude oil and refined products are complex mixtures of compounds, essentially volatile to dense hydrocarbons (HC) and heavy metals (HM). Both are likely to have direct effects on plants, through contact and the assimilation of contaminants^{160,162,163}, as well as indirect ones due to modifications of the soil's physico-chemical and biological properties^{174,247,248} and plant nutrition¹⁷³. They cause structural changes in leaves and roots^{162,163} and decreased pigment concentrations^{77,159,166} and water content^{159,169,174} in leaves. As a consequence, vegetation spectra can be affected by exposure to oil, resulting in reflectance modifications at corresponding wavelengths^{68,159,186}. Recent studies have thus attempted to detect soil contamination resulting from pipeline leaks and oil spills using reflectance data acquired on vegetation at multiple scales^{68,77,233,249}. However, only a few of these studies focused on the interspecific variability of plant spectral responses to crude oil and by-products, and how to differentiate them from responses induced by long-term exposure to other types of stressors, which currently significantly compromise the reliability of this technique. Regarding the latter point, none of these studies followed a statistical approach based on the classification of original and transformed spectra.

This paper outlines the first part in a series of researches carried out at multiple scales to assess the potential of hyperspectral remote sensing of vegetation for detecting oil-contaminated soils. The first step presented here involved small-scale greenhouse experiments conducted under controlled conditions, with a view to future airborne or satellite applications. This study aims to describe the spectral responses of various plant species to long-term oil exposure based on a statistical approach. Particular attention was focused on discriminating oil-induced and water-deficit stress responses, which can have similar effects on plants^{169,184,206}.

5.2. Materials and Methods

5.2.1. Plant materials

A greenhouse experiment was conducted at the Pôle d'Études et de Recherche de Lacq (Platform for Experimental Research in Lacq, PERL, France) on three different plant species to study their hyperspectral responses to oil exposure under controlled conditions. The species were chosen because of their marked presence in temperate and tropical mud pits resulting from oil and gas exploration and production activities. Cultivated young plants of *Cenchrus alopecuroides* (L.) Thunb. and *Panicum virgatum* L., two species typical of tropical regions, were acclimated in pots for 15 days in a greenhouse before being transplanted to the treatment soils. Wild plants of *Rubus fruticosus* L. were collected from an uncontaminated site, and went through the same acclimation period on their soil of origin. All the plants reached a height of 15 cm before the study began.

5.2.2. Soil sampling and preparation

Two different samples of contaminated soil were collected on the same brownfield used for oil and gas production activities and now covered by vegetation. The soil was organized into two layers, contaminated by various concentrations of HC and HM as shown in Table 5. The bottom layer was made up of 1.5 m of oil and gas waste mud covered with a 0.2-m layer of clay soil. C₁₀-C₄₀ hydrocarbons were the main compounds found in the soil, essentially dense hydrocarbons from the C₂₁-C₄₀ fraction, with concentrations up to 36 000 and 140 000 mg.kg⁻¹. In terms of HM, the soil, particularly the upper clay layer contained high concentrations of chromium (Cr) and zinc (Zn) of 2100 and 4500 mg.kg⁻¹. This layer also contained large amounts of mercury (Hg) (6 mg.kg⁻¹), whereas high concentrations of Polycyclic aromatic hydrocarbons (PAHs) (1600 mg.kg⁻¹) were found essentially in the oil and gas mud. Both layers were sampled and used for plant exposure to contaminants. The upper layer was sieved to remove root residues, and the oil and gas mud was placed under a fume hood for 10 days to remove volatile aromatic compounds and avoid contaminating the other treatments during the experiment. Each type of soil was homogenized in a cement mixer before the pots were filled. An uncontaminated soil was also sampled and used for both control and contamination dilution purposes.

Table 5. Heavy metal and hydrocarbon concentrations found in the control and contaminated soils used in this study. Polycyclic Aromatic Hydrocarbons (PAH) include the 16 priority pollutants of the US Environmental Protection Agency.

	Control	Contaminated soil	
		Clay layer	Bottom layer
<hr/>			
Heavy metals (mg.kg ⁻¹):			
arsenic	16	23	11
cadmium	0.66	1.7	2
chromium	30	2100	1300
copper	25	82	37
mercury	0.07	6	2.6
lead	56	85	44
nickel	28	84	16
zinc	290	4500	2100
<hr/>			
Hydrocarbons (mg.kg ⁻¹):			
Polycyclic Aromatic Hydrocarbons	Below	34	1600
Total hydrocarbons C ₁₀ -C ₄₀	limit of detection	36000	140000

5.2.3. Plant treatments

Different treatments were applied to the temperate and tropical species. *Rubus* plants were grown on a contaminated soil made up of 100% upper clay layer, as well as on an uncontaminated soil with a good water supply as control plants or under limited irrigation to induce water-deficit stress. *Panicum* and *Cenchrus* plants were exposed to a diluted contaminated soil made up of 90% uncontaminated soil and 10% oil and gas mud, and on a 100% uncontaminated soil as control plants. Five replicates per treatment were used for each species. Plant roots were cleaned with water before transplantation to 8-L individual glass pots filled with a 3-cm layer of clay balls and the corresponding treatment soil. A drainage system was installed under each pot to collect lixiviates from irrigation. The pots were randomly placed in a greenhouse and regularly moved to avoid a potential heterogeneity of ambient conditions. Plants were grown for 61 days between April and June at 27°C and 70% hygrometry, with an LD 12:12 photoperiod provided by natural and supplementary artificial light. Plants were fertilized weekly (N-P-K, 6-6-6) and irrigated daily, except for those subject to water-deficit which were watered only during the acclimation period.

5.2.4. Hyperspectral reflectance acquisitions

Hyperspectral reflectance data were collected using an ASD FieldSpec 4 Hi-Res spectroradiometer (Analytical Spectral Devices Inc. Boulder, Colorado, USA), with a spectral range of 350 to 2500 nm, a spectral resolution of 3 nm in the visible (VIS, 350-750 nm) and near-infrared (NIR, 750-1300 nm) and 8 nm in the short-wave infrared (SWIR, 1300-2500 nm), and a spectral sampling of 1.4 and 1.1 nm in the 350-1000 nm and 1001-2500 nm domains respectively. Data were collected in radiance and converted to reflectance using a white reference calibration panel²⁰³. Hyperspectral data were acquired at leaf scale every 12 days from day 1 to 61 of the experiment, with a leaf-clip fixed to a contact probe and an internal light source, on a small black background panel. The leaves of *Panicum* and *Cenchrus* were too thin to allow leaf acquisitions, so only *Rubus* leaves were measured. For each *Rubus* treatment, 3 young leaves per replicate were chosen at random, so 45 spectra were collected on each date.

Hyperspectral data were also acquired from the three species at plant scale. A 5° fore-optic lens placed 35 cm at nadir above the plant and providing a 3-cm acquisition footprint was used for taking the measurements. These acquisitions were performed under an external light source produced by an Aluminized Reflektor lamp (USHIO America Inc., Cypress, USA) set 70 cm away from the target at a 20° zenithal angle. To obtain several measurements, the acquisitions were made at different points of the plant. So, 3 and 6 acquisitions per replicate were made on *Rubus* and the tropical species respectively. Data were acquired every 12 days, from day 25 to 61 for *Rubus* and from day 49 to 61 for the tropical species. Foliage density was insufficient to consider measurements at plant scale before this.

5.2.5. Data pre-processing and analysis

In order to set our research in a context for larger-scale applications, reflectance measurements were concatenated to the spectral resolution of the airborne hyperspectral sensors HySpex VNIR-1600 and SWIR-320m-e (Norsk Elektro Optikk AS, Lørenskog, Norway), resulting in a spectral resolution of 3,7 nm in the 400-1000 nm domain and 6 nm in the 1000-2500 nm domain. Atmospheric water absorption bands in the SWIR (1350-1500 nm and 1800-1950 nm) were removed, because they are unusable at airborne scale due to significant atmospheric effects. A low signal-to-noise ratio was observed between 2400 and 2500 nm, so these wavelengths were

also removed. Consequently, 160 and 183 bands remained in the 400-1000 nm and 1000-2500 nm domains respectively for data analysis.

Different data transformations were applied in order to reduce the variability observed in the original spectra and to enhance the information contained in particular bands for statistical treatment. In similar studies^{185,202}, first and second derivatives appeared to be particularly sensitive to reflectance modifications at multiple scales so in our study, we calculated them using the reflectance data. The *Continuum Removal*²⁵⁰ (CR) transformation was also applied, because it showed real potential for analyzing plant response to soil contamination^{77,251}. In addition, two more transformations, Standard Normal Variate (SNV) and Area Under Curve Normalization (AUCN), were tested. SNV consists in subtracting the reflectance at each wavelength by the mean reflectance of the spectra, and dividing the result by the standard deviation²⁰³. AUCN computes the ratio of reflectance at each wavelength to the area under the entire reflectance curve. Both SNV and AUCN transformations were expected to reduce variability by normalizing signatures, thus allowing better detection of spectral differences between treatments.

All the transformed data were analyzed by linear discriminant analysis²²⁰ (LDA), following a stepwise forward procedure based on Wilk's lambda to discriminate between the two or three treatments applied to the species. In comparison with other classification methods²¹⁷, LDA appeared to be particularly efficient in hyperspectral remote sensing for discriminating between vegetation classes. One of the most advantageous aspects of the stepwise procedure is that it provides information on the most discriminant wavelengths, so it is possible to know which parts of the spectra and corresponding plant parameters (*e.g.* pigments, tissue structure, water content) are particularly affected by exposure to contamination or water-deficit. It also enables these discriminant wavelengths to be compared with those observed in similar studies. LDAs were carried out on the three species separately, with all acquisition dates gathered. Original (untransformed) and transformed reflectance data were each divided into 50% training and 50% test sets, with the same proportion of acquisition dates. Leaf and plant scales were addressed separately in the case of *Rubus*. LDAs were performed using Statistica Software (StatSoft Inc., Oklahoma, USA).

5.3. Results and Discussion

5.3.1. Visible symptoms

All species exhibited visible changes when exposed to contaminants. Symptoms of chlorosis appeared just a few days after the beginning of the experiment (1-5 days) and remained until day 61. Similar responses were observed after 8 days²⁰⁰ in previous studies on *Phaseolus vulgaris* exposed to 96 g.kg⁻¹ of 15/W40 oil. Likewise, *Zea mays* leaves also developed chlorosis associated with decreased photosynthesis after 6 to 14 days, depending on the concentration of 15/W40 oil or oil and gas mud in the soil^{29,186}. Chlorosis is due to a lack of chlorophyll in leaves, and can be induced by contaminant exposure, as frequently described by authors^{77,159,162}. A particular pigmentation was observed on *Rubus* leaves after 20 days, materialized by the appearance of persistent red spots on the upper surface and contour. None of the control and water-deficit treatments exhibited this type of pigmentation during the experiment.

5.3.2. Spectral signatures

The hyperspectral responses of plants to contamination exposure were consistent with those observed in similar studies. In comparison with the control plants, all species exhibited increased

reflectance in the VIS when exposed to contaminants (Figure 27), even in the early stages of the experiment (13-25 days). The wavelengths of plant pigment absorption (450-650 nm) and of red-edge region (~700 nm) appeared to be particularly sensitive to contamination, with a reflectance up to 0.1 greater than that of the control plants. A similar response was observed later in *Rubus* plants subject to water-deficit stress (49-61 days), but the reflectance was still higher for plants exposed to contaminants. The differences observed between treatments stayed almost constant in the VIS until the end of the experiment (61 days). Similar differences were previously observed in *Rubus fruticosus* and *Salicornia virginica* exposed to HC and HM^{29,159}, and a marked increase of reflectance of *Phaseolus vulgaris* was found at these wavelengths at high concentrations of oil²⁰⁰.

In the NIR and SWIR, the reflectance of *Rubus* increased from day 25 to 61 in the case of both treatments, but was nonetheless lower for plants exposed to contaminants than to water-deficit (Figure 27). However, an opposite trend was observed in *Panicum* and *Cenchrus*, which both exhibited decreasing in the NIR between days 49 and 61. This is consistent with a similar study²⁰¹ and confirms that the response of plants to contaminants in the NIR depends on the species^{29,77,159,201}. As described for *Hordeum vulgare* exposed to Zn and Cd²⁵², this could be explained by changes in leaf structure, which have a particular influence on NIR reflectance^{69,71,119}. Finally, the amplitude of differences in the NIR and SWIR remained less significant than in the VIS, and no marked variability was observed between the responses of tropical species to contamination in any part of the spectra.

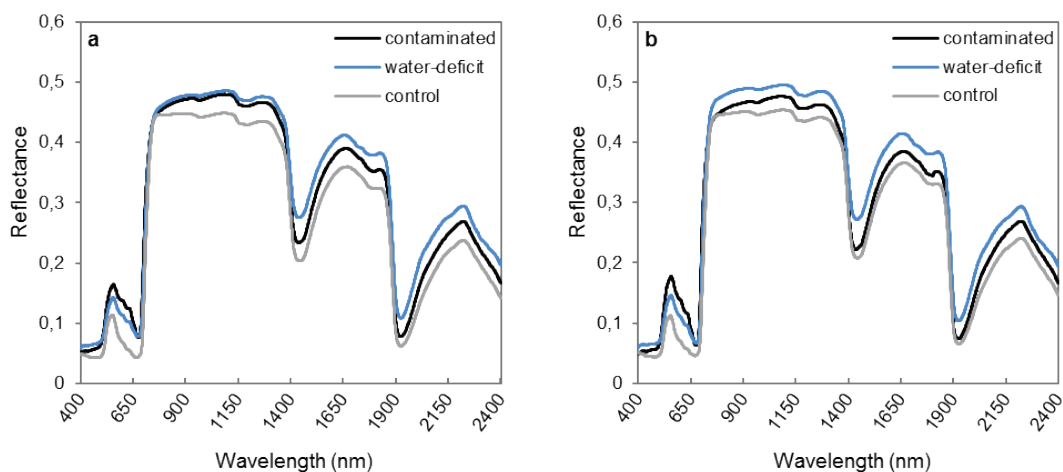


Figure 27. Mean reflectance of *Rubus* grown for 61 days under different conditions, at (a) leaf and (b) plant scale. Data are shown prior to pre-processing.

5.3.3. LDA classification

5.3.3.1. Leaf scale

The accuracy of the classification performed by LDA was assessed for the three species. At leaf scale, the classification of *Rubus* treatments showed an average accuracy of 75% for the original (untransformed) signatures and the six acquisition dates gathered. The best results (> 80%) were obtained with derivatives and CR transformations, with an accuracy of 93% in the case of the contaminated treatment (Figure 28a). The results obtained with other transformations were almost similar to or lower than that of the original data. The water-deficit treatment was the most difficult to classify, and accuracy never exceeded 79%. The spectra from the control

and water-deficit treatments were essentially confused with each other (up to 38%), but rarely with those of contaminants-exposed plants (<10%). This was particularly true for the most accurate transformations, where fewer than 2 and 9% of the spectra of the control and water-deficit treatments were confused with those of contaminated plants.

To better understand these results, it is important to note that LDAs were performed on all acquisition dates, including in the initial stages (1-13 days) when there were only a few noticeable differences between treatments. Looking at things from a temporal point of view, the classification gradually improved as the exposure time increased (Figure 29), which is consistent with results described in the previous section. Hence, the spectral signature of the plant is more greatly affected and more accurately classified the longer the plant is exposed to contamination or water-deficit. All treatment exhibited the same trend, with, on day 1, an accuracy of less than 75% for control plants, 72% for contaminated ones and 57% for those exposed to water-deficit. An accuracy of 100% was then achieved on day 61, or sometimes before, for all treatments and transformations. Control plants were easier to classify than the other treatments in the first stages. Consequently, after 61 days, there was no confusion between *Rubus* plants exposed to different treatments using both untransformed and transformed spectral signatures.

5.3.3.2. Plant scale

The reflectance data used for LDAs at plant scale included the last two acquisition dates (49 and 61 days) for *Panicum* and *Cenchrus* and two additional dates for *Rubus* (25 to 61 days). Based on the results obtained at leaf scale, those at plant scale were initially expected to be more accurate, because no data were available in the early stages (1-25 days). However, the fore optic lens acquisitions displayed greater variability than the leaf-clip ones, because they were influenced by many additional factors such as leaf angle distribution (LAD) and background and atmospheric effects. Despite this, the classification of *Rubus* test sets showed similar trends at plant scale, with the best overall accuracy obtained with derivatives and CR (> 82%), this time followed by AUCN (Figure 28b). However, transformed data produced better results (> 70%) than untransformed data (54%). This reflects the significant variability found at this scale, and the ability of data transformation to reduce it without any major loss of information. More precisely, the plants exposed to contaminants were always easier to classify than those undergoing other treatments, even when using untransformed data. This resulted in an accuracy of more than 90% with first derivative and CR. The water-deficit treatment was once again the most difficult to classify when all dates were included in the LDA. The confusion described at leaf scale was not observed at plant scale, which leads us to assume that confusion between water-deficit and control treatments is particularly marked during the first days of the experiment (1-13 days). Again at plant scale, the temporal evolution of LDA accuracy followed the same increasing trend as observed for leaves. Transformed spectra achieved an accuracy of more than 90% when classifying control and contaminated plants after 61 days (Figure 29). The first derivative, AUCN and CR were able to classify plants undergoing water-deficit without any confusion after the same amount of time.

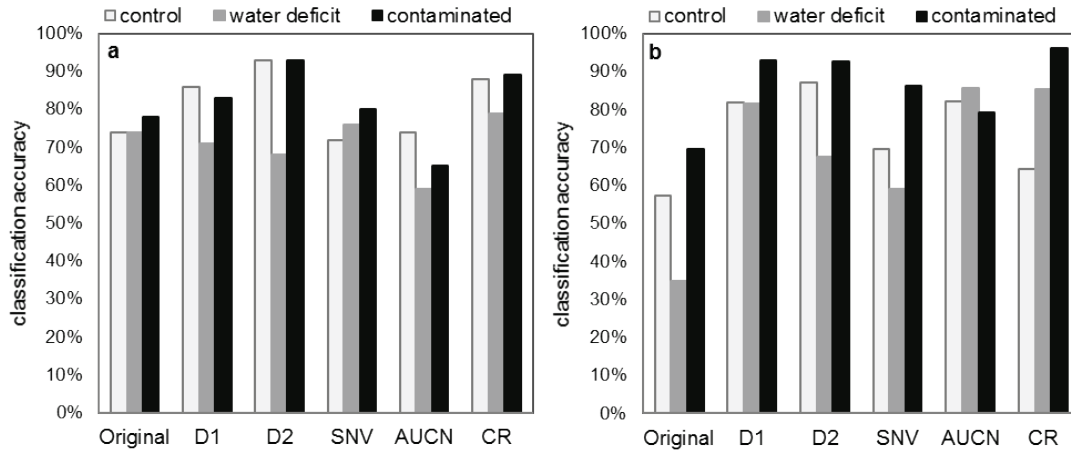


Figure 28. Classification accuracy obtained from LDA performed on *Rubus* leaf (a) and plant (b) reflectance data, all acquisition dates gathered (1 (a) or 25 (b) to 61 days). (D1, first derivative; D2, second derivative; SNV, Standard Normal Variate; AUCN, Area Under Curve Normalization; CR, *Continuum* Removal.)

Compared to *Rubus*, the tropical species *Panicum* and *Cenchrus* were exposed to two treatments only (no water-deficit). Because the leaves of these species were too thin to allow leaf-clip acquisitions, reflectance data were collected only at plant scale, making it impossible to compare the two scales. However, both species were exposed to the same contaminated soil in similar conditions, so it was interesting to study the interspecific variability of their response to contamination. For the *Cenchrus* plants, transformed and untransformed spectra produced substantially similar classification results, with an accuracy close to 80%. The results for *Panicum* plants were closer to those obtained with *Rubus*. Once more, CR gave the best classification accuracy, especially for plants exposed to contaminants, for which spectra were never confused with those of the control plants on all acquisition dates. It is important to note that although acquisitions were made at plant scale, the untransformed reflectance data were also able to discriminate treatments, sometimes better than the transformed data. As described above for *Rubus*, the accuracy increased with exposure time, and often reached more than 90% on day 61. Consequently, the convergent results observed for the three species at leaf and plant scales suggest that vegetation reflectance is a reliable tool for the detection of contaminated soils under long-term exposure. This opens up promising perspectives for airborne and satellite applications at discussed later on in this article.

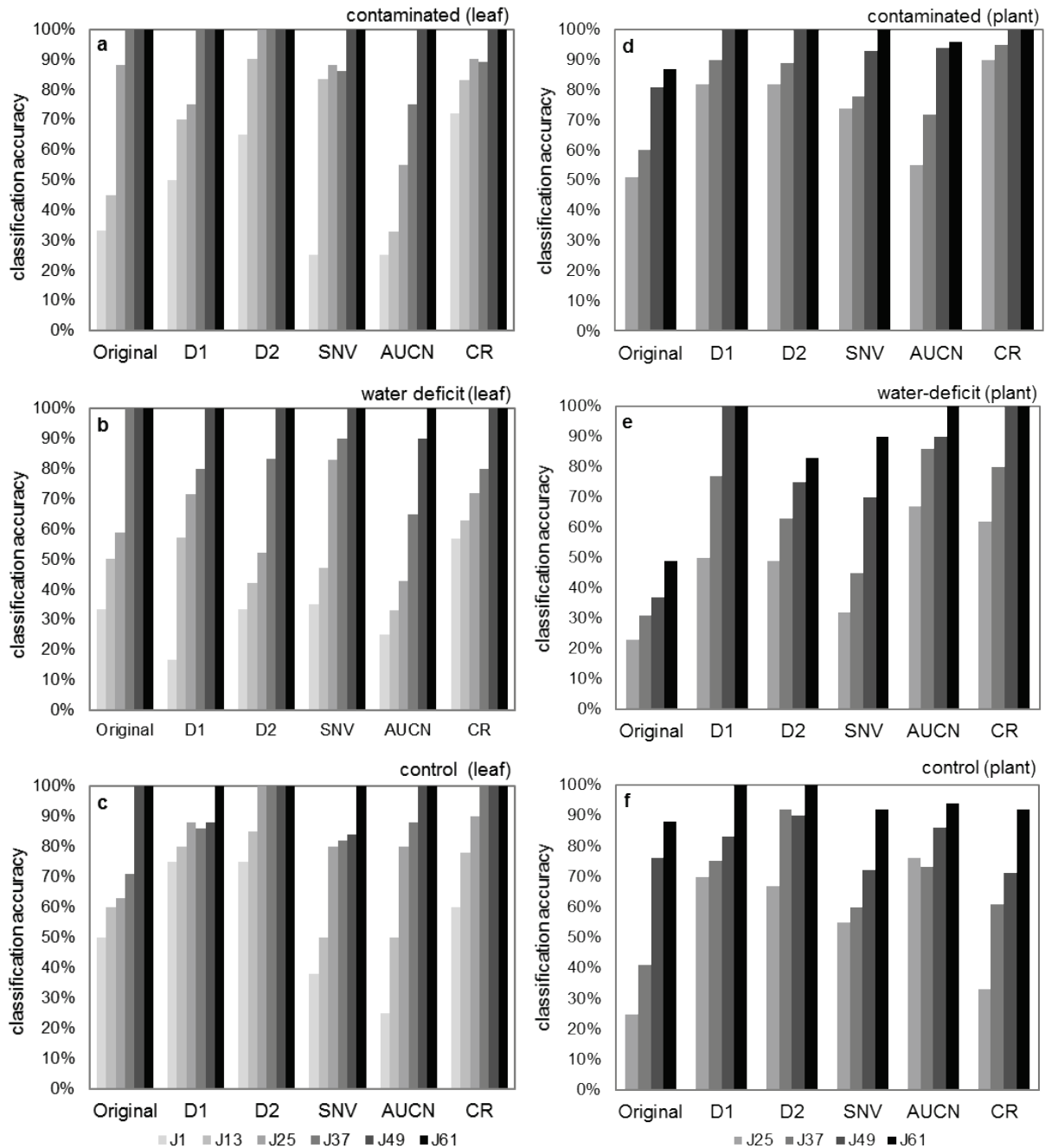


Figure 29. Temporal evolution of classification accuracy obtained from LDA performed on *Rubus* reflectance data, at leaf (a-c) and plant (d-f) scales. Results are displayed from day 1 ('J1') or 25 ('J25') to day 61 ('J61') for leaf and plant scales respectively.

5.3.4. Discriminant wavelengths

As described above, the main advantage of the stepwise procedure is that it identifies the most discriminant wavelengths between treatments. The relative frequency of these wavelengths obtained by LDA was analyzed with 50 nm-wide spectral sampling, with transformed and untransformed data gathered, to define the most discriminant domains of the spectra between treatments for each species and study scale. This step helped us to better identify the spectral domains that we should focus on in future studies for airborne and satellite applications.

At leaf scale, the main wavelengths used for classification were essentially located in the VIS between 400 and 500 nm (Figure 30), which corresponds to the absorption of plant pigments, especially chlorophylls^{90,240,253}. This result is consistent with those described in similar studies on vegetation responses to contaminated soils^{186,251}. It also reflects the differences observed in this region on spectral signatures, from day 13 until the end of the experiment. In addition, a large number of discriminant wavelengths appeared in the SWIR, from 2200 to 2300 nm. A smaller number of wavelengths were also found between 1500-1600 nm and 2000-2400 nm. Only a few authors have investigated these parts of the spectrum in similar studies^{29,77,251}. It appears, however, that the SWIR might also be significant for the remote detection of soil contamination based on vegetation⁷⁷. In our case, the presence of discriminant wavelengths at around 2200 nm could arise from differences between treatments observed toward the end of the experiment, as described above. Many authors have demonstrated that reflectance in the SWIR is particularly dependent on leaf water content and structure^{111,139,254}, which are affected by HC and HM contamination^{159,169,172,255}. It might explain the large number of discriminant wavelengths in the SWIR in our study. This is corroborated by the observations in the NIR, which exhibits discriminant wavelengths between 900-1100 nm, also consistent with water absorption^{134,256}. So, despite the intentional elimination of water absorption bands at around 1400 and 1900 nm that are particularly sensitive to the alteration of plant water content, the rest of the SWIR and the NIR were still significant for discriminating the responses of *Rubus* plants to the different treatments.

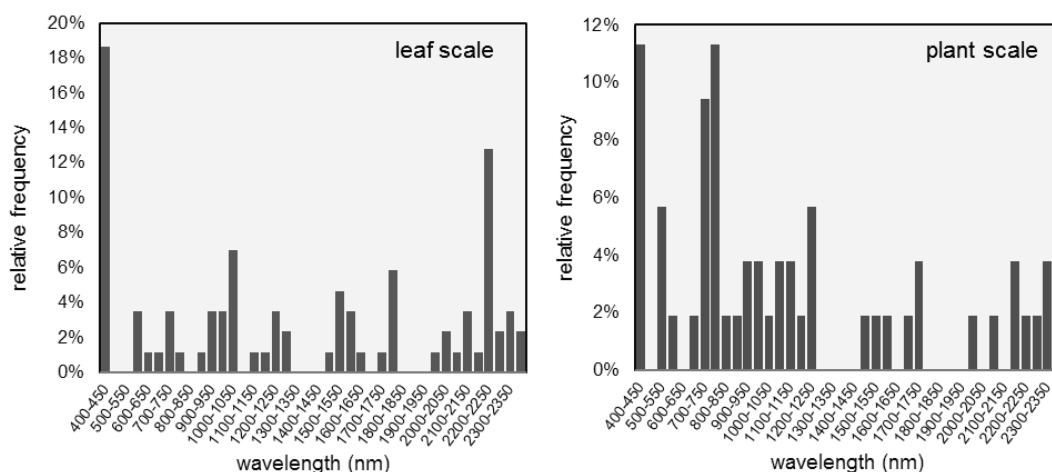


Figure 30. Discriminant wavelengths obtained from LDA performed on *Rubus* at leaf and plant scales. Frequencies are presented with untransformed and transformed data gathered, with bands grouped into 50 nm-wide intervals.

Discriminant wavelengths at plant scale were also found in the VIS, but they were essentially located in the red-edge (700-800 nm) (Figure 30). This result was observed for the three species, particularly *Rubus* and *Cenchrus*, and was described many times in similar studies as a shift of the red-edge position (REP) toward shorter or longer wavelengths, depending on the contaminant^{166,185,186,257-259}. This is consistent with the results described in the previous section for derivatives and CR, particularly sensitive to REP. These shifts might be induced by pigments variations as shown by some authors^{78,86,159,166,239}. A strong relationship between pigments concentrations and the red-edge area of *Suaeda salsa* exposed to phenanthrene was also observed at plant scale¹⁶⁶. In the rest of the spectrum, the discriminant wavelengths of *Panicum* and *Cenchrus* plants were closer to those of *Rubus* leaves, with a peak in the SWIR and bands of

interest in the NIR. So, the discriminant wavelengths were virtually identical among the three species at the same scale. They appear to be particularly sensitive to HC and HM contamination, so could be of great interest for upscaling purposes.

Some authors have already investigated the feasibility of detecting oil and gas seepage or leakage in soils by hyperspectral remote sensing based on vegetation reflectance. For example, the substantial potential of the REP applied to hyperspectral imaging has already been proven in particular cases of pipeline leakage, oil spill monitoring and natural oil and gas seepage detection^{20,35,68,233,249}. However, the lack of knowledge about the factors involved in vegetation response to HC and HM and the variability of this response compromise the reliability of this approach. In this respect, our study aimed to provide new contributions to help the understand these factors encountered in natural conditions at airborne and satellite scales.

The optical properties of vegetation are governed by its biochemistry and structure, which can be affected by many environmental factors such as water and nutrients or light availability^{258,260}. It is essential to consider these factors in natural conditions, because their effects on vegetation reflectance can be similar to those induced by HC and HM. Some authors have already discriminated plant spectral responses to various biotic and abiotic stressors of natural and anthropogenic origins^{74,76,258}, but only a few studies looked at oil contamination^{200,206}. Our results demonstrated the possibility of differentiating the effects of a complex mixture of contaminants from those of natural stress, that both have effects on vegetation water content^{169,206}. Results indicated that long-term exposure improves this discrimination. However, these factors were studied separately, so the next step after this study will be to investigate vegetation responses to multiple interacting factors, including soil contaminated by HC and HM.

Not all plant species have the same sensitivity to oil and gas product exposure^{163,261}. As a consequence, the amplitude of vegetation spectral response is highly dependent on the plants' tolerance to contamination^{77,201,230}. So, the more the plant is affected by oil, the easier it is to remotely detect soil contamination. Our results underlined the importance of the interspecific variability in vegetation response to HC and HM, with contrasting visible and spectral symptoms between the two monocotyledonous species *Panicum* and *Cenchrus* and the dicotyledonous species *Rubus*. Similar differences between these groups had already been observed in the spectral signatures of *Cenchrus alopecuroides* (formerly known as *Pennisetum*) and *Forsythia suspensa* after 28 days of exposure to oil²⁰¹, and on *Zea mays* and *Rubus fruticosus* after 90 days of exposure to mud pits soil²⁹. However, our study indicates that not all parts of the spectrum are affected by oil at the same time, and that the interspecific variability of the response is essentially located in the NIR and could arise from well-known differences between the internal leaf structures of monocotyledonous and dicotyledonous plants⁷⁰. Hence, a better understanding of these contrasted sensitivities is needed for upscaling purposes, especially in densely vegetated regions subject to oil and gas condensate seepage and leakage, where specific patterns of species distribution appear^{20,262}.

Another important requirement for the remote monitoring of soils based on vegetation is to have a precise description of the composition and concentration of the contamination, in order to ensure efficient response with the appropriate remediation solutions. However, contaminants with different chemical properties induce similar responses on plant reflectance^{77,159,166}, so it is not yet possible to clearly distinguish the presence of complex mixtures from that of a single contaminant. Concerning this last point, our upcoming research will focus on the variability of plant response to multiple soil contaminants, with a particular focus on different HC and HM mixtures. Promising studies on HM quantification based on vegetation reflectance^{202,203,259} also

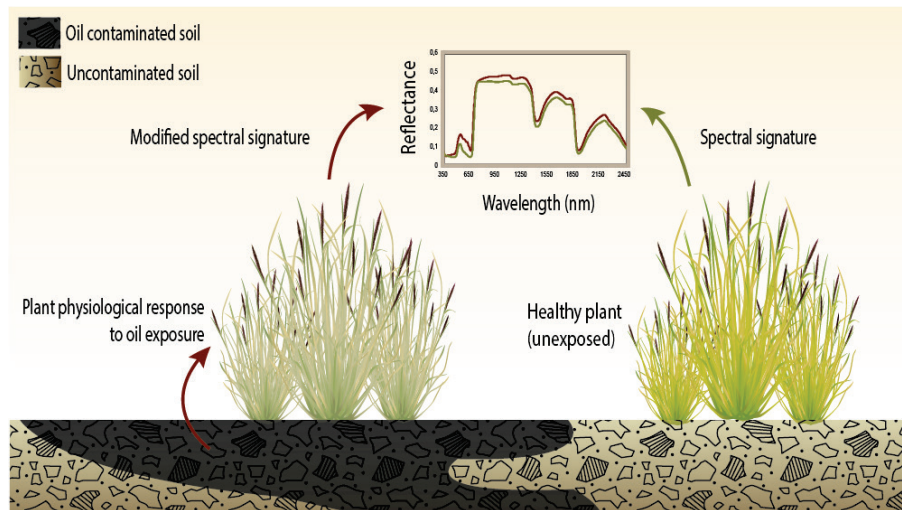
encouraged us to address the same issue with hydrocarbons, with studies recently initiated on bare soils^{61,245}.

The aim is not to replace other techniques used for oil spill detection²⁶³, but to make hyperspectral remote sensing an indispensable tool for monitoring soil contamination in densely vegetated oil and gas exploration and production regions. Hyperspectral remote sensing still needs to be vastly improved before it can be applied to actual case studies. In this respect, our study was the initial stage of a step-by-step multiscale approach, and will be followed by field and airborne acquisitions to assess this technique on temperate and tropical oil and gas production sites subject to soil contamination.

Acknowledgements

This work was performed in the frame of the NAOMI R&D project between TOTAL and the ONERA, with the support of the Ecolab research unit of Toulouse. Financial support was provided by TOTAL. We gratefully acknowledge Elaine Hannan, native English speaker, for correcting the manuscript.

Graphical abstract



6. Second article: “Detection and discrimination of various oil-contaminated soils using vegetation reflectance”

Article reference:

Lassalle, G.; Fabre, S.; Credoz, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Erudel, T.; Buffan-Dubau, E.; Dubucq, D.; Elger, A. Detection and Discrimination of Various Oil-Contaminated Soils Using Vegetation Reflectance. *Sci. Total Environ.* **2019**, 655, 1113–1124. <https://doi.org/10.1016/j.scitotenv.2018.11.314>.

Abstract

The use of hyperspectral spectroscopy for oil detection recently sparked a growing interest for risk assessment over vegetated areas. In a perspective of image applications, we conducted a greenhouse experiment on a brownfield-established species, *Rubus fruticosus* L. (bramble), to evaluate the potential of vegetation reflectance to detect and discriminate among various oil-contaminated soils. The species was grown for 32 days on four different soils with mixtures of petroleum hydrocarbons and heavy metals. Additional plants were grown on either uncontaminated control or water-deficient soils for comparison. Repeated reflectance measurements indicated modified spectral signatures under both oil and water-deficit exposure, from leaf to multi-plant scales. The amplitude of the response varied with mixture composition, exposure time, acquisition scale and spectrum region. Reflectance changes were linked to alterations in chlorophyll, carotenoid and water contents using vegetation indices. These indices were used to catch spectral similarities among acquisition scales and to discriminate among treatments using Kendall’s coefficient of concordance (W) and regularized logistic regression. Of the 33 vegetation indices tested, 14 were concordant from leaf to multi-plant scales ($W > 0.75$, $p < 0.05$) and strongly related to leaf biochemistry ($R^2 > 0.7$). The 14 indices allowed discriminating between each mixture and the control treatment with no or minor confusions ($\leq 5\%$) at all acquisition scales, depending on exposure time. Some of the mixtures remained difficult to discriminate among them and from the water-deficit treatment. The approach was tested at the canopy scale under natural conditions and performed well for identifying bramble exposed to either one of the experimentally-tested mixtures (90 % accuracy) or to uncontaminated soil (83 % accuracy). This study provided better understanding of vegetation spectral response to oil mixtures and opens up promising perspectives for future applications.

Keywords: Remote sensing, Hyperspectral spectroscopy, Total petroleum hydrocarbons, Heavy metals, Vegetation index, Pigment

6.1. Introduction

In oil and gas industry, exploration and production activities require efficient detection of total petroleum hydrocarbons (TPH) for assessing environmental risks from accidental leakages^{263,264}. For both purposes, active and passive remote sensing have recently raised a growing interest^{243,265,266}. Thanks to high spectral resolution, hyperspectral remote sensing revealed to be particularly efficient onshore⁶⁴. For example, direct oil detection can be achieved by analyzing the spectral signature of soils at wavelengths corresponding to light absorption of TPH^{61,267}. However, this technique implies the presence of apparent oil at the surface, so its application on-shore is restricted to bare soils and not effective in densely vegetated areas. As a solution, recent studies suggest to detect oil indirectly, through its effects on vegetation health^{20,29,77}.

TPH and heavy metals (HM) present in oil both affect vegetation biochemical and biophysical parameters related to leaf optical properties^{158,162,169,183}. Authors noticed alterations of leaf anatomy, pigment and water contents under TPH and HM exposure, resulting in reflectance modifications at corresponding wavelengths^{77,159,166,268}. By exploiting these modifications, it is thus possible to detect oil in soils from the spectral signature of vegetation. For such purpose, vegetation indices (VI) provided encouraging results but remain underexploited^{68,206,268}.

Although the reliability of TPH and HM detection has been demonstrated experimentally, its application from airborne and satellite hyperspectral images leads to the apparition of false alarms (*i.e.* a vegetation stress not related to the presence of oil)^{20,35,68,269}. As pointed out by authors, a better knowledge about the spectral response of vegetation to TPH and HM exposure is necessary for improving the detection. In that sense, studies carried out under controlled conditions may help to better identify the factors influencing this response.

The sensitivity of species to TPH and HM exposure is among the most influential factors^{163,189}, as it leads to an important interspecific variability of vegetation spectral response^{29,185,201,270}. Hence, the less a species is sensitive to TPH and HM, the harder it is to detect oil contamination. But even when considering a sensitive species, several sources of variability compromise the detection. For instance, the presence of other stressors that induce almost similar effects to TPH and HM (*e.g.* water-deficit) induce false alarms^{186,271}. Likewise, the study scale (leaf, plant or canopy) strongly influences the spectral response^{77,186}. Contrasted spectral signatures of a single species exposed to TPH and HM mixtures appear in the near-infrared (NIR) when measured at different scales⁷⁷. The composition and overall concentration of oil products also contribute to the response of vegetation, but have been rarely studied^{77,159,206}. However, they are of great importance, because they vary substantially among regions and contexts (natural oil seeps, pipelines, production mud pits, etc.)^{29,152,272}. In a perspective of airborne and satellite applications, it is therefore essential to account for all these factors under controlled conditions by reproducing as faithfully as possible the contexts of oil exposure encountered under natural conditions.

This study focuses on the detection of various oil products, with the aim to distinguish among them from the spectral signature of shrubby temperate vegetation, under controlled conditions. Their discrimination with a water deficit stress, which is very common under natural conditions, was also investigated. A particular attention has been given to the biochemical response of the species, and its consequences on reflectance at three scales of measurement. A concrete case of field application is also presented for validating the approach.

6.2. Materials and Methods

6.2.1. Study site and species

An experiment was carried out for 32 days at the Pôle d'Etudes et de Recherche de Lacq (*Platform for Experimental Research in Lacq* (PERL), France), to study the biochemical and spectral responses of vegetation to various TPH and HM contaminations under controlled conditions. This experiment consisted in reproducing under greenhouse the conditions of a realistic case of oil contamination. For this purpose, an industrial brownfield located in temperate region, also known as “mud pit”, was identified. The site served as a deposit of oil and gas production residues for a long time and is now contaminated by TPH and HM. The soil consists in two layers, made of muddy residues at the bottom, covered by a clayed coat at the surface. This brownfield has been colonized by vegetation, mainly *Rubus fruticosus* L. (bramble). The greenhouse experiment focused on this species, which is typically encountered on industrial sites under temperate climate^{29,273,274}.

6.2.2. Plant materials and treatments

The greenhouse experiment consisted in six treatments applied to bramble, including four different mixtures of TPH and HM (Table 6). Wild plants of bramble were sampled from an uncontaminated site and acclimated in their soil of origin for 15 days under greenhouse. All the plants reached 15 cm-height before the study began. Both soil layers from the brownfield were collected and used for the treatments applied to brambles. An uncontaminated soil was also sampled and has been analyzed to ensure similar texture than that of the brownfield. All the soils were sieved to 10 mm to remove residual roots and homogenized manually.

Bramble plants were exposed to six treatments, including four different mixtures of TPH and HM (Table 6). The first one (Clay) was made of 100% clayed layer from the brownfield. This soil also contained various heavy metals known for their phytotoxicity¹⁵⁸. Two other treatments were made of muddy residues (Mud_B) or “Dahlia” crude oil (Crude) diluted in an uncontaminated soil. All these treatments exhibited similar C₁₀-C₄₀ TPH concentrations (25 000 mg.kg⁻¹), but differed from each other by their composition in other contaminants. A fourth contaminated treatment (Mud_A) was applied to the plants. This involved muddy residues diluted to 6000 mg.kg⁻¹ C₁₀-C₄₀ TPH. The remaining treatments consisted in 100 % uncontaminated soil, with (Control) or without (Water_str) water supply.

Table 6. Composition of soil treatments used in this study. Concentrations are given in mg.kg⁻¹. Polycyclic Aromatic Hydrocarbons (PAH) represent the 16 priority pollutants of the U.S. Environmental Protection Agency (EPA²⁷⁵). (< DL: below detection limit.).

Only low concentrations of HM were found in the uncontaminated soil (Control and Water-str) and corresponded to the local geochemical background.

Origin of oil mixture	Control	Water-str	Clay	Mud_A	Mud_B	Crude
	None		Brownfield (clay)	Brownfield (muddy residues)		“Dahlia” crude oil
Heavy metals:						
arsenic	7.5		22	12	14	12
cadmium	< DL		1.7	0.20	0.67	< DL
chromium	23		2000	74	310	30
copper	4.5		73	9.9	17	9.1
mercury	< DL		11	0.14	0.72	0.05
lead	38		92	43	40	33
nickel	12		61	15	18	14
zinc	28		4000	170	660	57
Polycyclic Aromatic Hydrocarbons	< DL		19	63	260	25
Total petroleum hydrocarbons:						
C ₅ -C ₁₀	< DL		< DL	52	200	4000
C ₁₀ -C ₄₀	< DL		25000	6000	25000	25000

Bramble roots were cleaned with water before transplanting in individual black pots filled with a 3-cm height layer of clay balls and 3 L of the corresponding treatment soil. 13 plant replicates were used for each treatment. They were grown between April and May at 27 °C and 60% hygrometry, under natural and artificial light providing a 12:12 light:dark photoperiod. All the plants were fertilized weekly (N-K-P: 6-6-6) and irrigated to field capacity on a daily basis, except those subject to water-deficit (Water_str) which were not watered during the study.

6.2.3. Biochemical analysis

The biochemical response of plants to the treatments was analyzed during the experiment. For this purpose, three leaves were sampled per treatment, on which two 15-mm wide disks were collected. The first one was used to determine leaf fresh (FW) and dry (DW) weights, leaf water content (LWC) and equivalent water thickness (EWT)^{68,139,141}. Leaf pigments were extracted in methanol from the second disk using the protocol of Diepens *et al.*²⁷⁶ and analyzed by High Pressure Liquid Chromatography (HPLC) according to the method described by Barlow *et al.*²⁷⁷. Analyses included chlorophyll a, b and various carotenoids. LWC, EWT and pigment content were examined together through Principal Component Analysis (PCA)^{278,279}. The importance of the different variables was identified from their contribution to the main components²⁸⁰. In addition, the spectral signatures of all disks were measured using the protocol described for leaves in section 6.2.4 and linked to LWC, EWT and pigment content for further analysis described in section 6.2.5. Biochemical analyses were performed at intermediate stages (days 11 and 18) and at the end of the experiment (day 32). No leaf was sampled during early stages to avoid influencing plant growth.

6.2.4. Spectral reflectance measurements and preprocessing

The spectral signatures of the species were measured using an ASD FieldSpec 4 Hi-Res spectroradiometer (Analytical Spectral Device Inc. Boulder, Colorado, USA), with a spectral range of 350 to 2500 nm. Data were acquired in radiance and converted to reflectance from the formula of Milton²⁸¹, using a white reference calibration panel (Spectralon, Labsphere Inc., North Sutton, USA). Measurements were performed at leaf, plant and multi-plant scales^b, from early (day 4) to intermediate (days 11 and 18) stages and at the end of the experiment (day 32). Leaf reflectance was measured using a leaf-clip with an internal light source. For each of the six treatments, the measurements were made on 3 young leaves per replicate (n = 234 measurements per date). For each leaf, 5 consecutive spectral signatures were acquired and averaged to obtain a single one.

Additional measurements were carried out with a 10°-FOV fore-optic placed above the pots, providing a corresponding acquisition footprint of 5 and 20 cm at plant and multi-plant scales, respectively. Measurements were performed under natural light between 11.30 am. and 1.30 pm., in the absence of clouds. The reflectance of each replicate was thus measured at plant scale (n = 78 measurements per date). Four replicates of the same treatment were joined for measurements at multi-plant scale. The operation was repeated three times for each treatment, with different replicates (n = 18 measurements per date). 10 consecutive spectral signatures were acquired per measurement at plant and multi-plant scales and averaged to obtain a single one.

Reflectance data located below 400 and above 2400 nm were removed at all acquisition scales, because of a low signal-to-noise ratio (SNR). Likewise, we did not conserve those from the 1350-1450 and 1800-1950 intervals due to low atmospheric transmission at plant and multi-plant scales. At all acquisition scales, a Savitzky-Golay smoothing filter with a second order polynomial and a seven-point window was then applied to the spectral signatures to improve the SNR at the remaining wavelengths^{282,283}.

6.2.5. Vegetation indices

In order to catch similarities among acquisition scales, 33 simple and normalized vegetation indices (VI) were computed and used in a three-step approach. These indices are presented in Table 7. The first step consisted in identifying which biochemical parameters (pigments, LWC and EWT) were responsible for index changes among treatments (Figure 31). This was achieved by regression using the leaves sampled for biochemical analysis. A strong multicollinearity was detected among pigments by calculating Variance Inflation Factor ($VIF > 10$)²²¹. Multicollinearity occurs when predictors (*i.e.* pigments, LWC, EWT) are linear combination of each other, and leads to confusions when identifying those contributing most to the response variable²⁸⁴ (*i.e.* VI). To prevent such consequence, we used the elastic net regression²²², a penalized least squared method allowing efficient predictor selection under multicollinearity²⁸⁵. The elastic net regression was performed on each VI. Their best set of predictors and the resulting R^2 of the model are presented.

Once we linked VI to biochemical parameters, we tested for their ability to discriminate among treatments, at all acquisition scales. The 33 VI were computed from the spectral signatures acquired on (unsampled) leaves, plants and multi-plants. The second step of the approach consisted in detecting similarities in VI among acquisition scales, through the ranking of treatments. This was achieved for each VI using Kendall's coefficient of concordance²⁸⁶ (W) based on VI values of treatments at day 32, as it was the most representative of a long-term

^b In this study, the term “canopy scale” was restricted to *in situ* measurements to avoid confusions.

exposure of vegetation in our study. Kendall's W evaluates the degree of agreement of treatment ranking among acquisition scales for a given VI and is tested for significance using a 1000-fold permutation test^{287,288}. As a result, we only retained VI with best elastic net results ($R^2 > 0.7$) and high significant concordance ($W > 0.75$, $p < 0.05$), thus expressing changes in pigment and water content resulting from treatment exposure at leaf, plant and multi-plant scales (Figure 31). The selected VI were finally tested to discriminate among treatments using L^2 -Regularized Logistic Regression (RLR), also known as Ridge regression^{219,289}. RLR has been widely used for classification purposes in various domains including remote sensing^{290,291}, and revealed to be efficient when applied on vegetation²⁹². We only trained the RLR classifier at leaf scale to assess the robustness of VI at higher acquisition scales. For this purpose, half of the leaf dataset was used for training, and the remaining half as test set. In addition, the entire plant and multi-plant datasets were used to test the method. The classification was carried out on each date separately and iterated 30 times. The overall quality of predictions was evaluated by Mean Overall Accuracy (MOA) and mean Cohen's Kappa coefficient^{218,292}. Confusion matrices are also presented to illustrate the results. These matrices were obtained by comparing the predicted treatments to the true ones, and thus allowed identifying the confusions among treatments. User's and Producer's accuracies (UA and PA, respectively) were also computed as described in Story & Congalton²¹². UA indicates the probability for a sample being classified in a given treatment, to actually correspond to this treatment, whereas PA indicates the probability of a sample from this treatment to be correctly classified. All data analyses were performed under Python language using Statsmodels²⁹³, Scipy²⁹⁴ and Scikit-Learn²⁹⁵ libraries.

Table 7. Vegetation indices used in this study (R: Reflectance, D: First derivative).

Index	Formula	Reference
Chlorophyll Absorption Ratio Index	$CARI = R_{700} R_{670}a + R_{670} + b / R_{670} (a^2 + 1)^{0.5}$ with $a = (R_{700} - R_{550}) / 150$ and $b = R_{550} - R_{550}a$	296
Chlorophyll/Carotenoids Index	$CCI = D_{720} / D_{700}$	297
Carter Index 1	$CTR1 = R_{695} / R_{420}$	298
Carter Index 2	$CTR2 = R_{695} / R_{760}$	298
Gitelson & Merzlyak Index 1	$M1 = R_{750} / R_{550}$	299
Gitelson & Merzlyak Index 2	$M2 = R_{750} / R_{700}$	299
Lichtenthaler Index 1	$LI1 = (R_{800} - R_{680}) / (R_{800} + R_{680})$	253
Lichtenthaler Index 2	$LI2 = R_{440} / R_{690}$	253
Lichtenthaler Index 3	$LI3 = R_{440} / R_{740}$	253
modified Chlorophyll Absorption Ratio Index 1	$mCARI1 = 1.2[2.5(R_{800} - R_{670}) - 1.3(R_{800} - R_{550})]$	130
modified Chlorophyll Absorption Ratio Index 2	$mCARI2 = \frac{1.5[2.5(R_{800} - R_{670}) - 1.3(R_{800} - R_{550})]}{\sqrt{(2R_{800} + 1)^2 - (6R_{800} - 5\sqrt{R_{670}}) - 0.5}}$	130
modified Simple Ratio 705 nm	$mSR705 = (R_{750} - R_{445}) / (R_{750} + R_{445})$	78
MERIS Terrestrial Chlorophyll Index	$MTCI = (R_{754} - R_{709}) / (R_{709} + R_{681})$	300
Normalized Difference 705 nm	$ND705 \text{ or } NDVI705 = (R_{750} - R_{705}) / (R_{750} + R_{705})$	78
Normalized Pigment Chlorophyll Index	$NPCI = (R_{680} - R_{430}) / (R_{680} + R_{430})$	301
Optimized Soil-Adjusted Vegetation Index	$OSAVI = (1 + 0.16) (R_{800} - R_{670}) / (R_{800} + R_{670} + 0.16)$	302
Photochemical Reflectance Index 1	$PRI1 = (R_{528} - R_{567}) / (R_{528} + R_{567})$	101
Photochemical Reflectance Index 2	$PRI2 = (R_{531} - R_{570}) / (R_{531} + R_{570})$	101
Photochemical Reflectance Index 3	$PRI3 = (R_{570} - R_{539}) / (R_{570} + R_{539})$	101
Plant Senescence Reflectance Index	$PSRI = (R_{678} - R_{500}) / R_{750}$	303
Pigment Specific Simple Ratio a	$PSSRa = R_{800} / R_{680}$	92
Pigment Specific Simple Ratio b	$PSSRb = R_{800} / R_{635}$	240
Pigment Specific Simple Ratio c	$PSSRc = R_{800} / R_{470}$	240
Structure Intensive Pigment Index 1	$SIPI1 = (R_{800} - R_{445}) / (R_{800} + R_{680})$	81
Structure Intensive Pigment Index 2	$SIPI2 = (R_{800} - R_{505}) / (R_{800} + R_{690})$	81
Structure Intensive Pigment Index 3	$SIPI3 = (R_{800} - R_{470}) / (R_{800} + R_{680})$	81
Simple Ratio 705 nm	$SR705 = R_{750} / R_{705}$	78
Transformed CARI	$TCARI = 3[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550}) (R_{700} / R_{670})]$	304
Transformed CARI / OSAVI	$\frac{TCARI}{OSAVI} = \frac{3[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550}) (R_{700} / R_{670})]}{(1 + 0.16) (R_{800} - R_{670}) / (R_{800} + R_{670} + 0.16)}$	304
Vogelmann Index 1	$VO1 = R_{740} / R_{720}$	305
Vogelmann Index 2	$VO2 = (R_{734} - R_{747}) / (R_{715} + R_{726})$	306
Vogelmann Index 3	$VO3 = (R_{734} - R_{747}) / (R_{715} + R_{720})$	306
Disease Water Stress Index	$DWSI = (R_{800} - R_{550}) / (R_{1660} + R_{680})$	207

6.2.6. Field validation

In order to validate the approach under natural conditions, additional measurements were performed on brambles directly on both the brownfield and on uncontaminated sites. For this purpose, 30 patches were identified on the brownfield. Their soil was analyzed and showed similar TPH and HM concentrations to those of the Clay treatment from the experiment. 30 other patches from five natural sites were also identified. These sites were located a few kilometers away from the brownfield and had never been subject to industrial activities. They were considered as controls, because no contaminant was found in their soil. The spectral signatures of all patches were collected at canopy scale in June, under the same conditions of acquisition as those of multi-plant scale during the experiment (11.30 am. – 1.30 pm., 20-cm radiance acquisition footprint). VI were computed and tested with the RLR classifier (Figure 31), following training on the leaf dataset from the greenhouse experiment exclusively.

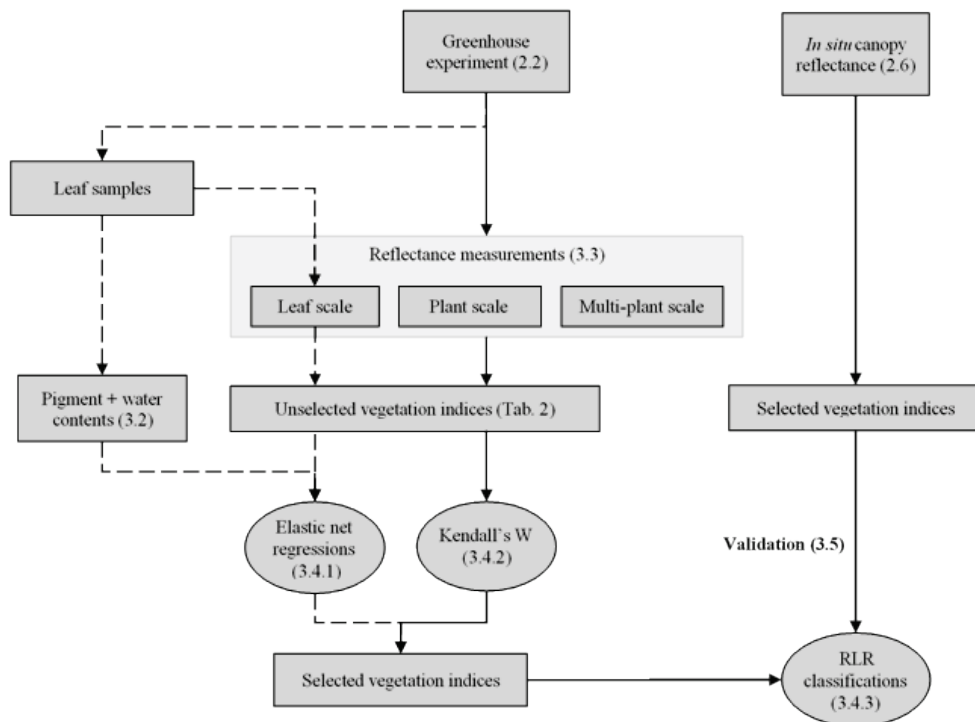


Figure 31. Flowchart of the approach using vegetation indices presented in this study. Sections and tables associated to the different steps are specified in brackets^c. The dotted line indicates the steps carried out at leaf scale only.

6.3. Results and Discussion

6.3.1. Visible symptoms

Visible symptoms of leaf discoloration appeared on all plants exposed to TPH and HM mixtures and to water-deficit after 18 days of experiment (Figure 32). Commonly described as chlorosis, these symptoms have been frequently observed on plant species exposed to crude oil or

^c See the original published article for corresponding section numbers.

refined products^{159,162,169} and various HM such as Ni, Pb and Zn¹⁵⁸, including *Rubus fruticosus*^{271,307}. In our study, it was particularly pronounced for plants exposed to crude oil or muddy residues. The latter also exhibited a singular red pigmentation on leaf margin, along with symmetric interveinal marks on leaf adaxial face at highest concentrations (Mud_B) (Figure 32). No similar symptom has already been described in previous experiments involving TPH and HM. However, several authors noticed that brambles are able to accumulate HM in leaves^{274,307,308}. Hence, the observed symptoms may be linked to HM toxicity^{191,309}. All these symptoms were associated to a reduced plant growth when compared to Control. Plants undergoing water-deficit also showed a drooping habit, which is typically observed in those conditions¹³⁷. On day 32, substantial evolutions appeared on plants from the Mud_A and Clay treatments. No or very few symptoms were observed on newly appeared leaves from these treatments and plant development was similar to that of the Control (Figure 32). None of the plants exposed to any of the TPH and HM mixtures died during the experiment. This tolerance of bramble to oil contamination explains its marked presence on industrial brownfields in temperate regions.

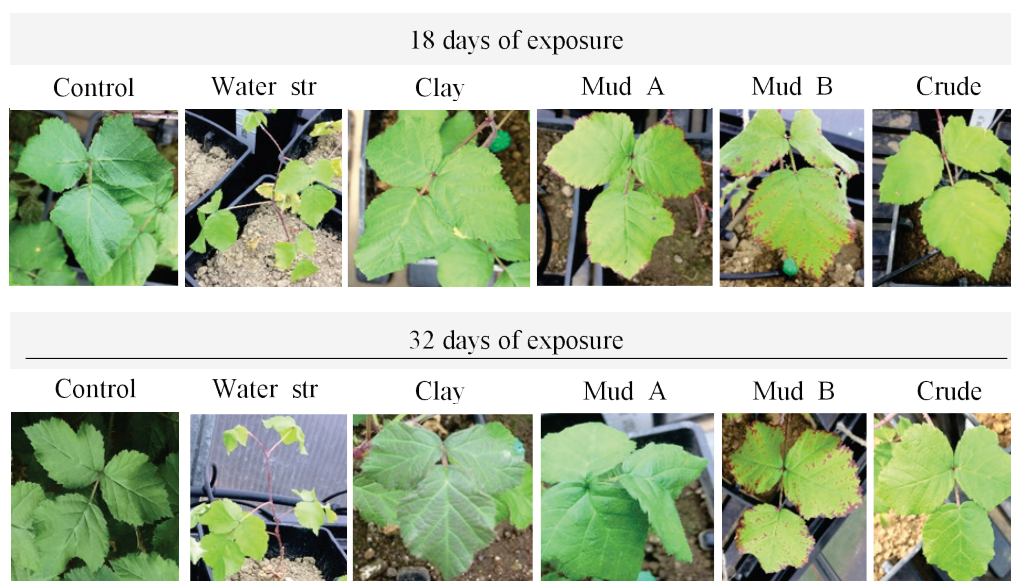


Figure 32. Visible symptoms observed on *Rubus fruticosus* (bramble) leaves after 18 and 32 days of exposure to the treatments (see section 6.2.2 for further information about the treatments). These symptoms were confirmed later by biochemical analyses described in section 6.3.2.

6.3.2. Biochemical responses to the treatments

The PCA performed on biochemical analyses of sampled leaves confirmed the differentiated effects of TPH and HM mixtures on plants. Most of biochemical parameters were positively correlated to the first Principal Component (PC1) (Figure 33a), with major contribution of chlorophylls and carotenoids. Antheraxanthin and zeaxanthin were the only pigments contributing to PC2. As PC1 explained most of the variability (72 %), we focused on the temporal evolution of plant treatments along this axis. Results indicated no difference after 11 days of exposure (Figure 33b). However, significant changes appeared on day 18, thus reflecting stress symptoms described in section 6.3.1. Control plants moved toward positive PC1 values, indicating an increase in pigment and water contents resulting from healthy plant growth. All other treatments remained close to the initial axis values. Plants exhibiting pronounced symptoms

(Crude, Mud_B and Water_str) slightly moved toward negative PC1 values, as their pigment and water contents began to decrease. This was confirmed at the end of the study (day 32), with an opposite trajectory compared to Control. In contrast, the disappearance of stress symptoms on Mud_A- and Clay-exposed plants on day 32 reflected an increase of pigment and water contents, expressed by positive PC1 values.

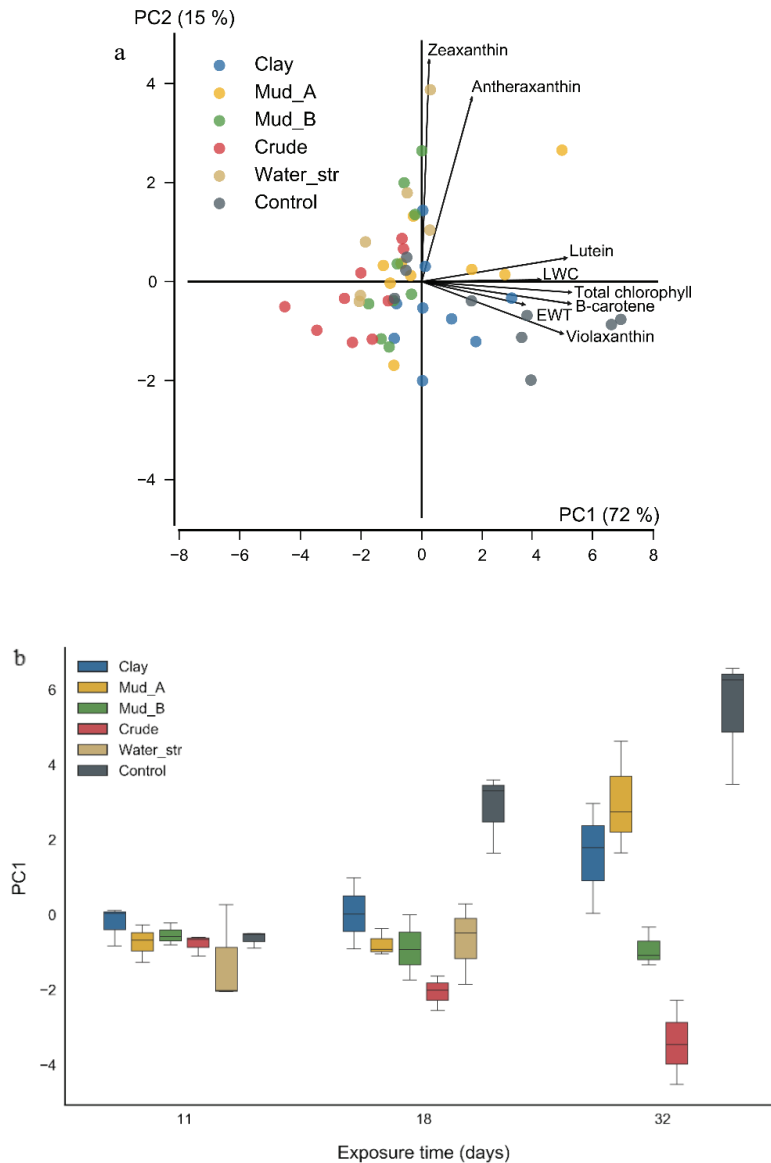


Figure 33. (a) PCA performed on pigment and water content data from three dates (days 11, 18 and 32, $n = 51$). The first two Principal Components (PC) are represented. Correlations between each variable and PC are represented by annotated arrows. The length of the arrows was multiplied by five on each PC for graphical convenience. (b) Temporal evolution of treatments along the first principal component (PC1), from day 11 to 32.

Alterations of pigment and water contents are frequently observed under TPH and HM exposure, alone or in mixture^{162,183,187,310}. They might be due to modifications of soil physico-chemical and biological properties^{154,167,248}, which lead to decreasing nutrient and water

availability for plants^{169,173}. PAH, TPH and HM can penetrate plant tissues and affect plant water status and photosynthetic capacity^{158,169,172,186}. However, their uptake by plant roots strongly varies with their chemical properties and association in mixtures^{191,311,312}. For example, low-carbon PAH and C₅-C₁₀ TPH penetrate roots more easily than larger chemical compounds^{313,314}, so they might be main responsible for biochemical alterations in our study. HM were also probably involved, since their bioavailability is enhanced in the presence of TPH^{158,172,311}. Judging by the evolution of Mud_A- and Clay-exposed plants, the effects of C₅-C₁₀ TPH and HM are however reversible at low concentrations, thanks to detoxification mechanisms^{191,315}.

None of the Control and contaminated treatments followed a clear trajectory on PC2. However, a noticeable evolution of water-deficient plants toward positive values appeared along this axis on day 18, indicating a peak of antheraxanthin and zeaxanthin. These pigments have already been reported as affected by TPH and HM exposure for the same species²⁶⁸. Their increase is however closely involved in plant response to water stress through the activation of the xanthophyll cycle³¹⁶⁻³¹⁸. This demonstrates differentiated effects of these two stressors on plant biochemistry.

6.3.3. Reflectance modifications

The spectral signatures of brambles were consistent with their biochemical response (Figure 34a-c). No difference appeared among treatments on days 4 and 11, and the spectral signatures were representative of healthy vegetation at all acquisition scales. These signatures were characterized by a low reflectance in the VIS with a peak around 550 nm, followed by a plateau in the NIR, and two marked peaks in the SWIR around 1650 and 2200 nm. On day 18, several reflectance changes appeared for brambles exposed to TPH and HM mixtures or to water deficit. Compared to Control, they showed an increase of reflectance in the VIS, which is consistent with previous studies^{29,77,159,206}. The affected wavelengths corresponded to those of chlorophyll and carotenoid absorption⁹². This increase was more pronounced on Crude, Mud_B and Water_str treatments, which shows that the spectral response intensity was in accordance to that of the biochemical response. Such changes in reflectance are commonly observed on stressed vegetation^{75,159,199}. However, because brambles are naturally tolerant to stressful conditions, their response to TPH and HM remained less pronounced than that of other species found in similar studies, especially crops^{29,77,206}. The most important differences among acquisition scales appeared in the NIR and short-wave infrared (SWIR). Because of reduced leaf density and ground cover^{69,126}, reflectance was lower than that of Control at plant and multi-plant scales in the NIR. This response was however not observed on leaf reflectance, which is mostly dependent on leaf anatomy in this region of the spectrum⁷¹. These results are consistent with those described in similar studies^{77,166}. In the short-wave infrared (SWIR), reflectance was also affected by biochemical alterations at leaf scale and by reduced ground cover at plant and multi-plant scales when compared to that of Control.

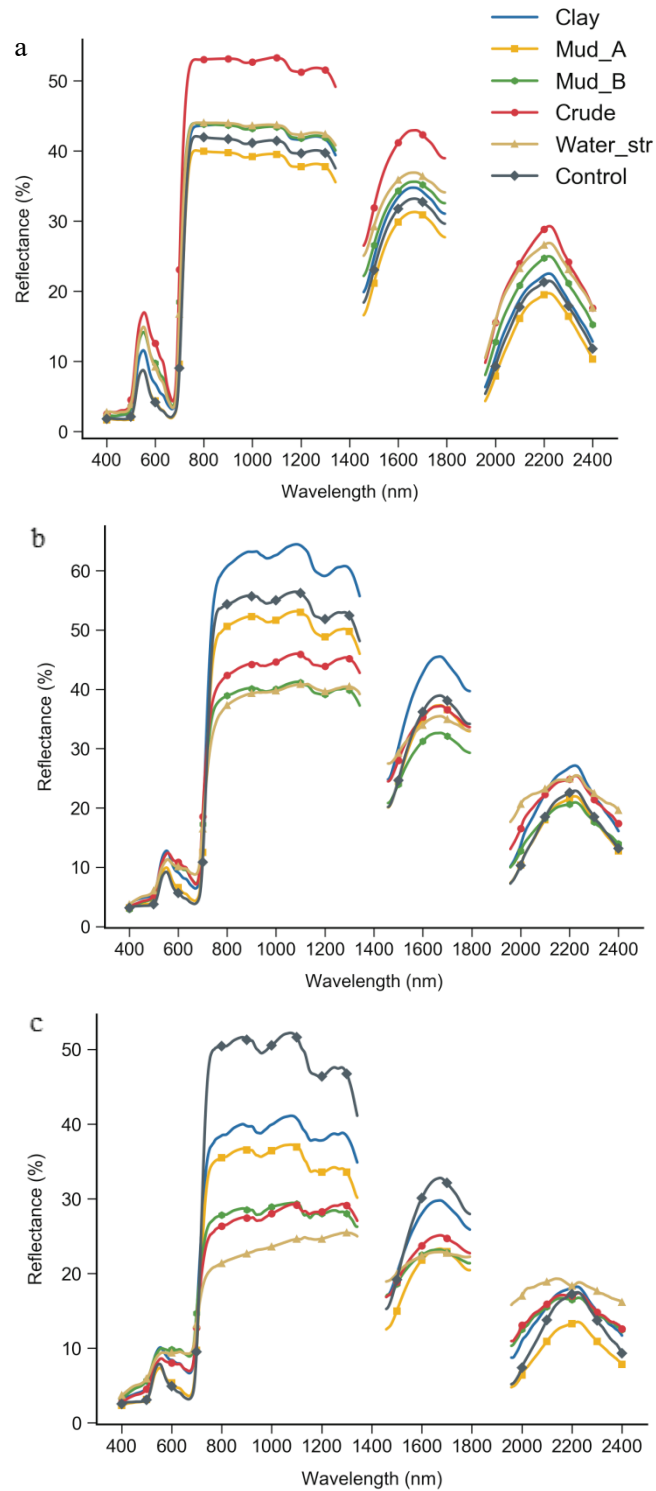


Figure 34. Mean spectral signatures of *Rubus fruticosus* (bramble) measured after 32 days of exposure to the treatments at (a) leaf, (b) plant and (c) multi-plant scales. Reflectance data located in the 1350-1450 and 1800-1950 intervals were removed, because of significant atmospheric effects at plant and multi-plant scales.

After 32 days, important changes occurred among treatments. Brambles growing on Mud_B and Crude soils exhibited lower pigment content (Figure 33b), so leaf reflectance continued to

increase in the VIS (Figure 34a). In contrast, the late pigment increase of brambles exposed to Clay and Mud_A treatments induced a reduction of reflectance in the VIS, compared to day 18. As a result, the spectral signatures of Mud_A became closely similar to Control in this region, at all acquisition scales. The late growth of brambles also modified plant and multi-plant reflectance. Plant reflectance from the Clay treatment exceeded that of Control in the NIR (Figure 34b), which confirmed previous results obtained with similar TPH and HM mixtures²⁷¹. Such changes were however not observed at multi-plant scale (Figure 34c), since ground cover was still limited. On the same date, brambles from other treatments remained severely impacted in the NIR and the SWIR, especially those exposed to the Crude treatment. A similar response was observed on *Brachiaria brizantha* H.S. exposed to 10 L.m⁻³ gasoline spill in a previous study⁷⁷. In addition, this treatment induced an important increase of reflectance in the NIR and SWIR at leaf scale, which might be related to the strong alterations of leaf anatomy and water content specifically induced by this mixture¹⁶².

6.3.4. Vegetation indices

6.3.4.1. Relationship between biochemical and spectral responses

Regressions performed on each VI from leaf samples led to the identification of at least one contributing pigment (Table 8), except for the DWSI which was only correlated to LWC. From the 33 indices, 15 were retained after this step ($R^2 > 0.7$). For almost all VI, variations were mainly explained by total chlorophylls (a + b), β -carotene and lutein, which were implicated in the biochemical response of brambles to the treatments (Figure 33a). These results confirmed that VI are suitable for tracking changes in pigment and water contents resulting from TPH and HM exposure, as suggested by previous studies^{68,268}. Other carotenoids were sometimes involved, especially antheraxanthin and zeaxanthin. Total carotenoid estimation has been widely assessed using VI, but only few studies attempted to separate carotenes and xanthophylls^{103,319}. Our study showed that it is of great importance for understanding the biochemical and spectral responses of vegetation to TPH and HM exposure.

Table 8. Results of elastic net regressions and Kendall's W selection procedure obtained on vegetation indices. Biochemical parameters contributing to index variations are presented in order of importance, along with the resulting R². (Chl: Total chlorophylls, Lut: Lutein, B-car: β -carotene, Ant: Antheraxanthin, Zea: Zeaxanthin, Vio: Violaxanthin, LWC: Leaf Water Content.) Indices showing significant Kendall's W ($p < 0.05$) are denoted by a *. Those retained for classifications ($R^2 > 0.7$, $W > 0.75^*$) are marked bold.

Index	Elastic net regression		Kendall's W
	Pigments	R ²	
CARI	Chl, Lut	0.77	0.29
CCI	Chl, B-car	0.94	0.92*
CTR1	Lut, Ant	0.49	0.90*
CTR2	Chl, Lut	0.87	0.90*
GM1	Chl, B-car	0.90	0.94*
GM2	Chl, B-car	0.93	0.90*
LI1	Chl	0.17	0.92*
LI2	Chl, B-car, Zea	0.60	0.97*
LI3	Lut, Ant	0.19	0.90*
MTCI	Chl, B-car	0.94	0.90*
ND705	Chl, B-car	0.93	0.90*
NPCI	Lut, Ant	0.41	0.76*
OSAVI	Chl	0.23	0.97*
PRI1	B-car, Vio	0.67	0.97*
PRI2	Chl, B-car	0.82	0.92*
PRI3	Lut, B-car	0.87	0.90*
PSRI	Chl, B-car	0.50	0.34
PSSRa	B-car, Zea, Vio	0.20	0.38
PSSRb	B-car, Zea, Vio	0.49	0.83*
PSSRc	B-car, Zea, Vio	0.28	0.64
SIPI1	Lut, Ant	0.51	0.62
SIPI2	Lut, Ant	0.75	0.92*
SIPI3	Lut, Ant	0.45	0.62
SR705	Chl, B-car	0.94	0.90*
TCARI	Lut, Zea	0.67	0.34
TCARI / OSAVI	Lut, Ant	0.08	0.23
VOG1	Chl, B-car	0.22	0.58
VOG2	Lut	0.93	0.90*
VOG3	Lut	0.93	0.90*
mCARI1	Lut, Ant	0.09	0.21
mCARI2	Zea	0.39	0.31
mSR705	Chl, B-car	0.92	0.90*
DWSI	LWC	0.84	0.85*

6.3.4.2. Concordance of VI among acquisition scales

When applied to all spectral signatures, the same VI showed similarities among acquisition scales, despite differences described in section 6.3.3. Most of them exploited the VIS, which is suitable for tracking alterations of vegetation pigments at varying scales of measurement^{102,240,241,320}. Based on treatment ranking at day 32, Kendall's W was significant and

greater than 0.75 for 22 of the 33 VI, indicating good concordance among acquisition scales (Table 8). This included 14 out of the 15 indices strongly linked to biochemical parameters ($R^2 > 0.7$), which were selected for classifications. For most of the retained VI, treatments were arranged as follows. As expected, Mud_B, Crude and Water_str treatments often exhibited similar VI values, which corresponded to low pigment and water contents (Figure 35). They were followed by Clay and Mud_A treatments, which values were close – and sometimes equal – to Control, because of pigment and growth recovery after 32 days. A lower variability was observed among replicates at multi-plant scale due to strong influence of bare soil fraction.

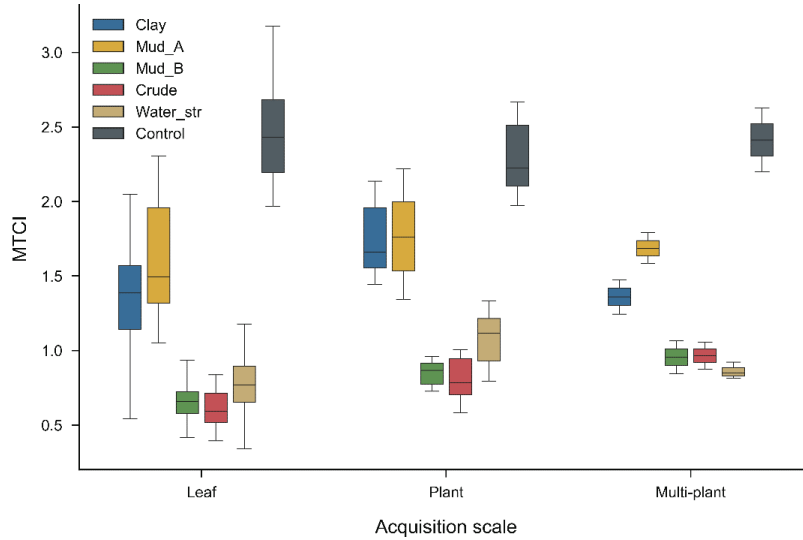


Figure 35. MTCI index computed from leaf, plant and multi-plant reflectance data acquired on day 32. This vegetation index was linked to total chlorophyll and β -carotene contents ($R^2 = 0.94$) and exhibited strong similarities among acquisition scales following Kendall's W selection procedure ($W = 0.90$, $p < 0.05$).

6.3.4.3. Discrimination of treatments using VI

RLR classifications performed on the 14 retained VI confirmed the concordance among acquisition scales. Same classifications using the original 33 VI led to worse results at day 18 and 32 and highlighted the improvements brought by the selection procedure (data not shown). Following training on half of leaf dataset exclusively, MOA and Kappa obtained from day 4 to day 32 on leaf, plant and multi-plant test sets were consistent with evolutions of spectral signatures. These results are summarized in Table 9. MOA and Kappa were significantly lower while increasing acquisition scale, since the alterations of leaf optical properties became more difficult to detect (Table 9). From day 4 to 11, MOA and Kappa stayed constant and lower than 40 % and 0.4, respectively, at all acquisition scales. On these dates, leaf biochemistry and plant development were similar among treatments, which explain the difficulty to discriminate among them.

Table 9. Mean Overall Accuracy (MOA) and mean Cohen’s Kappa (\pm SD) obtained from RLR classifications performed on leaf, plant and multi-plant test sets with the 14 retained vegetation indices. The training step was exclusively performed at leaf scale using different samples on each of the 30 iterations. For each scale, significant differences of MOA and Kappa between dates are denoted by different superscript letters (t-test, $p < 0.05$).

		Exposure time (days)			
		4	11	18	32
Leaf	MOA	36.3 (\pm 5) ^a	39.9 (\pm 6) ^a	70.8 (\pm 3) ^b	74 (\pm 4) ^b
	Kappa _a	0.30 (\pm 0.06) ^a	0.35 (\pm 0.06) ^a	0.65 (\pm 0.04) ^b	0.69 (\pm 0.05) ^b
Plant	MOA	26.8 (\pm 6) ^a	32.2 (\pm 5) ^a	61.8 (\pm 7) ^b	56.7 (\pm 7) ^b
	Kappa _a	0.13 (\pm 0.07) ^a	0.15 (\pm 0.08) ^a	0.54 (\pm 0.09) ^b	0.48 (\pm 0.09) ^b
Multi-plants	MOA	24.1 (\pm 6) ^a	28.9 (\pm 4) ^a	55.6 (\pm 6) ^b	51.7 (\pm 5) ^b
	Kappa _a	0.12 (\pm 0.07) ^a	0.14 (\pm 0.06) ^a	0.48 (\pm 0.08) ^b	0.42 (\pm 0.06) ^b

On day 18, MOA and Kappa became significantly higher than on previous dates. On this date, the levels of water and pigment alterations differed among the treatments (Figure 33b). These alterations were caught by VI at all acquisition scales, which led to a better discrimination among the treatments compared to the previous dates (Table 10). Since all VI were not correlated to the same set of biochemical parameters (Table 8), they were complementary in classifications. No significant difference of MOA and Kappa was observed at any acquisition scale between day 18 and 32, which firstly suggested no evolution in treatment discrimination. However, confusion matrices indicated critical changes between the two dates. On day 18, only few confusions were made between Control and other treatments at all acquisition scales ($\leq 5\%$), which agreed with biochemical and spectral responses (Table 10). Other treatments were however harder to discriminate among them, since their responses remained close on this date. Evolutions of VI values observed on Clay and Mud_A treatments toward those of Control then introduced confusions with the latter on day 32. More than 25 % of Mud_A samples were thus incorrectly labeled as Control at all acquisition scales. As a consequence, classification accuracy of Control decreased between day 18 and 32, as the prolonged exposure time to Clay and Mud_A mixtures negatively affected their discrimination with Control. In contrast, it was benefit for discriminating the Control with other mixtures. Mud_B and Crude treatments were never confused with Control on this date and rarely with Clay and Mud_A treatments, at all acquisition scales. This was also the case of water-deficient plants. These impacting treatments remained however difficult to discriminate among them, because of close VI values (Figure 35).

Table 10. Examples of relative confusion matrices (%) obtained from RLR classifications performed on leaf and plant test sets at day 18 and 32, following training at leaf scale. UA and PA stand for User Accuracy and Producer Accuracy (%), respectively.

18 days of exposure															
Leaf scale							Plant scale								
	Clay	Mud_A	Mud_B	Crude	Water_str	Control	UA		Clay	Mud_A	Mud_B	Crude	Water_str	Control	UA
Clay	68	11	4	7	7	3	68	Clay	28	38	7	9	14	4	28
Mud_A	17	54	11	12	2	4	54	Mud_A	24	34	16	14	9	3	34
Mud_B	2	11	79	3	5	0	79	Mud_B	2	0	76	22	0	0	76
Crude	2	13	4	64	13	4	64	Crude	8	2	18	58	11	3	58
Water_str	1	0	9	14	75	1	75	Water_str	4	0	0	15	81	0	81
Control	3	5	3	4	0	85	85	Control	2	1	1	2	0	94	94
PA	73	54	72	62	74	88		PA	41	45	64	48	70	90	

32 days of exposure															
Leaf scale							Plant scale								
	Clay	Mud_A	Mud_B	Crude	Water_str	Control	UA		Clay	Mud_A	Mud_B	Crude	Water_str	Control	UA
Clay	68	14	2	5	3	8	68	Clay	52	25	7	4	3	9	52
Mud_A	4	65	2	0	1	28	65	Mud_A	7	62	0	3	0	28	62
Mud_B	6	1	75	17	1	0	75	Mud_B	14	7	42	22	15	0	42
Crude	3	0	8	83	6	0	83	Crude	13	11	23	36	17	0	36
Water_str	1	0	6	5	88	0	88	Water_str	7	3	14	10	66	0	66
Control	9	26	0	0	0	65	65	Control	2	16	0	0	0	82	82
PA	75	61	81	75	89	64		PA	55	50	49	48	65	69	

6.3.5. Field validation

Following training on leaf dataset from the greenhouse experiment, RLR classifications performed on brownfield canopy measurements allowed identifying 27 out of the 30 bramble patches (90 %) as Clay-exposed vegetation. The remaining ones were confused with Mud_A (7 %) and Control (3 %) treatments. These results were better than those obtained on the same treatment at leaf and plant scales at the end of the experiment (Table 10) and suggest involvement of additional factors in vegetation long-term response to TPH and HM under natural conditions. For example, Emengini *et al.*²⁰⁶ showed that a combination of oil and water-deficit stress lead to stronger modifications of vegetation reflectance than oil stress alone. This might facilitate the discrimination with healthy vegetation. Of the 30 patches measured on uncontaminated sites, 25 were correctly labeled as Control (83 %). Confusions with Clay (3 %) and Mud_A (13 %) treatments were consistent with those observed for the Control after 32 days of experiment. These results thus demonstrated the potential of the approach developed under controlled conditions for discriminating among contaminated and uncontaminated soils under natural conditions.

6.4. Conclusion and Perspectives

This study aimed to demonstrate the potential of bramble reflectance for detecting and discriminating among various oil contaminations under controlled and natural conditions.

Modifications of reflectance occurred from leaf to multi-plant scales under exposure to TPH and HM mixtures. This study revealed that the composition of mixtures is largely responsible for these modifications and determines their temporal evolution. A clear link was established between modifications of reflectance and the biochemical response of brambles using VI. This confirmed that chlorophylls are not the only pigments involved in the spectral response of vegetation to TPH and HM. Following a leaf-scale training approach, VI enabled discriminating among mixtures and with uncontaminated control and water-deficient soils, from leaf to multi-plant scales. This highlighted the consequences of leaf biochemical alterations on the reflectance observed at higher scales. The results emphasized the importance of multitemporal acquisitions to ensure detection of various oil-contaminated soils at appropriate time. The soils contaminated by high levels of oil production residues or crude oil were particularly well detected after a prolonged time (32 days). Conversely, the best detection of low levels of production residues was obtained after a moderate time (18 days). VI showed great potential when applied under natural conditions.

The proposed method thus proved to be efficient for monitoring oil contamination at field scale. The diversity of soils being detected emphasizes the potential of this method for various operational applications, such as the detection of chronic pipeline leakages or soil contaminations resulting from past production activities. To expand its scope, the method needs to be assessed in other contexts, involving different plant species and oil contaminations. In that sense, upcoming study will focus on the tropical context, where the remote detection of oil remains a major challenge for environmental risk assessment. The method could be adapted to be applied to hyperspectral images with high spectral and spatial resolutions acquired from airborne or future satellite sensors. This represents a major challenge for a monitoring over large areas that will be addressed in future work.

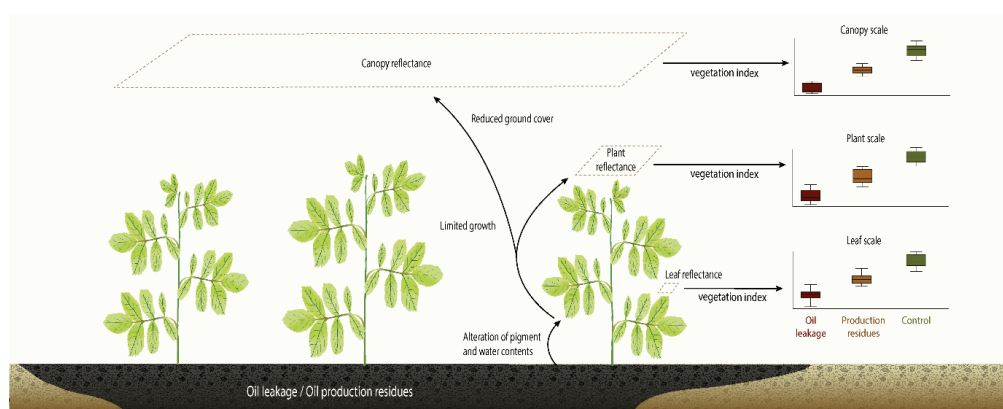
Acknowledgements

This collaborative work is part of the NAOMI R&D project between TOTAL and the ONERA, with the support of Ecolab and Dynafor research units of Toulouse. The authors gratefully acknowledge Didier Lambrigot and Olivier Berseille for their assistance in pigment analysis.

Declarations of interest: None.

Funding: Financial support of this work was provided by TOTAL.

Graphical abstract



7. Third article: “Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression”

Article reference:

Lassalle, G.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Fabre, S.; Elger, A. Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression. *Ecotox. Environ. Safe.* **2019**, 184, 109654. <https://doi.org/10.1016/j.ecoenv.2019.109654>

Abstract

The persistence of soil contamination after cessation of oil activities remains a major environmental issue in tropical regions. The assessment of the contamination is particularly difficult on vegetated sites, but promising advances in reflectance spectroscopy have recently emerged for this purpose. This study aimed to exploit vegetation reflectance for estimating low concentrations of Total Petroleum Hydrocarbons (TPH) in soils. A greenhouse experiment was carried out for 42 days on *Cenchrus alopecuroides* (L.) under realistic tropical conditions. The species was grown on oil-contaminated mud pit soils from industrial sites, with various concentrations of TPH. After 42 days, a hormesis phenomenon was observed. A significant decrease in plant growth and leaf chlorophyll and carotenoid contents was observed for plants exposed to 5 to 19 g.kg⁻¹ TPH in comparison to the controls ($p < 0.05$). Conversely, pigment contents were higher for plants exposed to 1 g.kg⁻¹ TPH. These modifications proportionally affected the reflectance of leaves and plants, especially in the visible region around 550 and 700 nm. 33 vegetation indices were used for linking the biochemical and spectral responses of the species to oil using elastic net regressions. The established models indicated that chlorophylls a and b and β -carotene were the main pigments involved in the modifications of reflectance ($R^2 > 0.7$). The same indices also succeeded in estimating the concentrations of TPH using random forest regression, at leaf and plant scales (RMSE = 1.46 and 1.63 g.kg⁻¹ and RPD = 5.09 and 4.44, respectively). Four out of the 33 indices contributed the most to the models (>75%). This study opens up encouraging perspectives for monitoring the cessation of oil activities in tropical regions. Further researches will focus on the application of our approach at larger scale, on airborne and satellite imagery.

Keywords: Reflectance spectroscopy, Soil contamination, Total petroleum hydrocarbons, Vegetation indices, Random forest

7.1. Introduction

Along with its development during the last century, oil and gas industry has become a major source of contamination in the environment^{36,321,322}. Crude oil and by-products (*e.g.* petroleum products, wastewater, oil sludge) are frequently released in soils, following accidental facility failures, bad practices and, more rarely, storm events^{28,33,34,46}. They cause important ecological disturbances, because oil contaminants – especially Total Petroleum Hydrocarbons (TPH) – are highly toxic toward organisms^{39,190,323}. The contamination of soil may persist in brownfields and mud pits after cessation of the oil production activity and affect ecosystems on the long term^{324,325}. The monitoring of oil activities has therefore become a critical environmental issue, and gave rise to an increasing need in reliable and cost-effective methods for assessing soil contamination. For this purpose, reflectance spectroscopy provided promising results when applied to bare soils^{61,63}. Its application to multi- and hyperspectral remote sensing imagery makes possible to detect and quantify soil TPH content at large scale using airborne sensors^{46,64}. More recently, a new approach has been proposed to extend its use to vegetated areas, where oil cannot be detected directly at the surface^{77,159,271}. This approach shows great interest in tropical regions, where vegetation is particularly dense^{68,231,326}.

TPH affect vegetation health – especially pigment and water contents – and optical properties in the reflective domain (400 – 2500 nm)^{159,183,186,327}. By exploiting the reflectance of leaves and canopies at particular wavelengths, it is possible to detect the changes induced by TPH in leaf biochemistry^{77,205,271}. This approach is however still under development and needs further improvements in order to be used operationally on imagery. Methods based on vegetation indices (VI) (*i.e.* reflectance ratios) have succeeded in distinguishing healthy and oil-exposed vegetation and in discriminating among various types of oil contamination, under controlled and natural conditions^{68,186,324}. Their ability to quantify TPH concentrations has however not really been exploited, although it is of great interest for assessing the environmental risks deriving from the contamination.

The quantification of TPH remains a major challenge, as it implies tracking little variations in leaf biochemistry from its reflectance. This might be particularly difficult on contaminated brownfields and mud pits, because the established species are particularly tolerant to oil exposure^{29,324}. This difficulty is moreover enhanced when dealing with low TPH concentrations, for two main reasons. In most cases, the lower the concentration, the lower the effects on vegetation health and reflectance^{77,186}. These effects become more difficult to detect below a certain concentration³²⁴. More rarely, some species undergo a stimulation of growth under exposure to low oil contamination^{328,329}. It is particularly frequent for tropical species. These phenomena make the quantification of TPH very challenging using vegetation reflectance. Previous studies focusing on heavy metals mostly used empirical univariate models and Partial Least Square Regression (PLSR) for linking vegetation reflectance to contamination^{203,245}. These methods achieved accurate quantification of the contamination, thanks to a pronounced and monotonous response of the species in the studied range of concentrations. However, they are not suitable for quantifying low TPH concentrations³²⁵, because of the abovementioned effects. Alternatives methods are therefore necessary. For this purpose, emerging machine learning methods proved efficient for solving complex – especially nonlinear – regression problems, so they could help quantifying TPH using reflectance spectroscopy^{330,331}.

This study aims to assess the potential of vegetation reflectance for quantifying low TPH concentrations in soils, under controlled conditions. The proposed approach relies on the combination of VI and a machine learning method – namely Random Forest (RF) regression.

This method was tested on reflectance measurements performed at various acquisition scales, under realistic tropical conditions.

7.2. Materials and Methods

7.2.1. Study site and greenhouse experiment

An experiment was carried out for 42 days at the Pôle d'Etudes et de Recherche de Lacq (Platform for Experimental Research in Lacq (PERL), France), for which a realistic case of oil contamination was reproduced under controlled tropical conditions. An industrial vegetated mud pit located in tropical region was identified for this purpose. Residues of oil and gas production have been accumulated on this site during the previous decades and resulted in persistent oil contamination in soils. After cessation of the activity, the site has been colonized by herbaceous vegetation, mainly *Cenchrus alopecuroides* (L.). The experiment focused on this species, which proved good indicator of oil contamination in previous studies^{201,271}. In order to reproduce as faithfully the conditions of the mud pit, the soil was sampled on two locations of the site and used in the experiment. Soil analyses revealed low concentrations of BTEX and Polycyclic Aromatic Hydrocarbons (PAHs) in both samples, and C₅-C₄₀ TPH concentrations of 5 and 21 g.kg⁻¹, respectively. The lower (C₅-C₁₀) and intermediate (C₁₀-C₂₁) TPH fractions have been partially degraded, so the remaining contamination mostly came from the dense fractions (C₂₁-C₄₀ hydrocarbons), which are poorly degradable and whose uptake by roots is very limited. Detailed soil analyses can be found in the Supporting Information section. No substantial change in TPH concentrations was observed for the different treatments at the end of the experiment (data not shown). In addition, an uncontaminated soil with similar texture was sampled on another site and used in the experiment. All the soils were homogenized manually and sieved to 10 mm to remove root residuals.

Five treatments were applied to *C. alopecuroides* for the greenhouse experiment, with TPH concentrations ranging from 0 to 19 g.kg⁻¹. For this purpose, the soils sampled on the mud pit and the control site were mixed in varying proportions to obtain five levels of contamination, respectively 0 (control), 1, 5, 13 and 19 g.kg⁻¹ TPH. The detailed soil analyses are presented in the Supporting Information section. Cultivated young plants were acclimated for 15 days in greenhouse and reached 15-cm height before being transplanted in individual pots filled with a 3-cm layer of clay balls and 3 L of the corresponding treatment soil. For each treatment, 11 replicates were grown for 42 days between May and June at 27° C and 70% hygrometry, with a 12:12 light:dark photoperiod provided by natural and artificial light. N-K-P (6-6-6) fertilization was applied weekly to the plants, which were irrigated to field capacity every day. At the end of the experiment, plant shoots were harvested, oven-dried for 48 h at 60° C and then weighed.

7.2.2. Biochemical analyses and reflectance acquisitions

The biochemical and spectral responses of *C. alopecuroides* to the treatments were measured during the study. For this purpose, we followed the procedure described in Lassalle *et al.*³²⁴. Young leaves were sampled on five different replicates per treatment and served for determining leaf water (LWC) and pigment contents^{68,276}. Leaf pigment content was quantified by High Pressure Liquid Chromatography (HPLC), and included chlorophylls a and b and various carotenoids, such as β-carotene, lutein, antheraxanthin, violaxanthin and zeaxanthin²⁷⁷. Differences among the treatments were analyzed through ANOVA and Tukey post-hoc tests. In addition, the spectral reflectance of the same sampled leaves was measured following the protocol

described below, and linked to biochemical analyses (see section 7.2.3). This procedure was carried out after 21 and 42 day of experiment. No leaf was sampled before in order to avoid influencing plant growth during early stages.

Reflectance measurements were also performed directly on the plants (*i.e.* without picking leaves) at day 0, 21 and 42, using an ASD FieldSpec 4 Hi-Res spectroradiometer (Malvern Panalytical, Malvern, UK). Data were acquired in radiance in the reflective domain (400 – 2400 nm) and converted to reflectance using a white reference calibration panel (Spectralon, Labsphere Inc., North Sutton, USA)²⁸¹. For each treatment, three young leaves per replicate were measured at random (n = 165 spectra per date) using a leaf-clip with an internal light source. Additional measurements were conducted at plant scale, by placing a 10°-FOV fore-optic above the pots to obtain a 5-cm wide acquisition footprint. These measurements were performed on each replicate (n = 55 spectra per date) under natural light between 11.30 am and 1.30 pm, under clear sky. The reflectance data from the 1350 – 1450 and 1800 – 1950 nm ranges were not conserved, because of low atmospheric transmission at plant scale. A Savitzky-Golay smoothing filter was then applied at the remaining wavelengths, improving the signal-to-noise ratio²⁸³.

7.2.3. Vegetation indices

33 VI were computed from the reflectance data and used in two ways. They are listed in the Supporting Information section. These indices have been exploited for detecting oil-induced changes in leaf biochemistry in previous studies, so they were tested in our context³²⁴. The 33 indices were first computed from the reflectance data of leaves sampled for biochemical analyses, and each of them was linked to leaf pigment and water contents using Elastic net (ENET) multiple regression²²², as described in Lassalle *et al.*³²⁴. This aimed to identify which biochemical parameters (pigment and water contents) contributed the most to index changes among treatments. ENET is a penalized least squared regression method that allows selecting predictors under multicollinearity. Multicollinearity appears when predictors – in our case, leaf pigment contents – are linear combination of each other, and results in confusions when selecting those contributing the most to the target variable (*i.e.* VI)^{221,284}. ENET regression prevents from such consequences and has already proved efficient for linking leaf biochemical and reflectance data³²⁴. For each VI, the best predictors were retained and the R² of the model was calculated.

In a second time, the same 33 indices were computed from the reflectance data acquired at leaf and plant scales and used for predicting TPH concentrations using RF regression³³¹. RF is an ensemble method that builds and average a lot of independent decision trees to model the relationship between predictors (*i.e.* the 33 VI) and a response variable – here, TPH concentration^{330,332,333}. Each tree is constructed from a set of decision rules that results from the recursive fragmentation of the original space into successively smaller sub-regions. To define these sub-regions, binary splits are applied independently to each vegetation index. A simple model is then adjusted between VI and TPH concentrations with respect to the splits, and those minimizing the mean squared error between the measured and predicted concentrations are retained. The succession of the selected index splits represents the branches of the decision tree, in such a way that each set of decision rules leads to an estimation of TPH concentration. A comprehensive description of the RF algorithm can be found in³³². RF accounts for non-linear relationships between the predictors and the response variable – which can occur for low concentrations of contaminants, so it was of great interest in our case. Moreover, RF informs on the relative contribution of VI to the quantification of TPH^{331,334}, which is essential for operational applications of the method. The RF regression was fitted to half of the data (50% training) and tested on the remaining part (50% test), at leaf and plant scales separately. For this purpose, we only used the data from day 42, because no differences among the treatments

were observed on the previous dates (see section 7.3.1). The predictions of TPH made on the test set were compared to those from initial soil analyses based on the R^2 , the Root Mean Squared Error (RMSE) and the Residual Predictive Deviation²⁰³ (RPD). The RPD denotes the ratio of the standard deviation of the measured TPH concentrations to the RMSE calculated between the measured and predicted concentrations.

7.3. Results and discussion

7.3.1. Biochemical and spectral responses to oil contamination

As expected, *C. alopecuroides* was highly tolerant to oil, as none of the plants from the contaminated treatments died during the experiment. This observation is consistent with previous studies carried out on the same species with various types of oil contamination (engine oil and oil and gas waste mud)^{201,271}. No visible stress symptom was observed on any of the replicates before the last stages of the experiment (35 – 42 days), so no difference in leaf pigment contents and in leaf and plant reflectance was observed among the treatments on the previous dates (0 and 21 days). On day 42, leaf discoloration was observed for the individuals exposed to 5 g.kg⁻¹ TPH or more, expressing a significant decrease in leaf chlorophyll and carotenoid contents when compared to the control ($p < 0.05$) (Figure 36). Plant growth was also significantly affected on this date. The alterations in pigment contents induced an increase of reflectance of 5 to 10% in the visible (VIS) region (400 – 750 nm), especially around 550 and 700 nm (Figure 37). Other regions of the spectrum were also affected in the same way by contamination. This response was more pronounced at leaf scale, because plant reflectance is also influenced by the plant architecture, the background soil and the illuminating and viewing geometry^{69,77}. As suggested in previous work^{201,271}, the late response of *C. alopecuroides* to oil highlights its interest for monitoring long-term contaminated sites, especially brownfields and mud pits remaining after cessation of oil and gas activities.

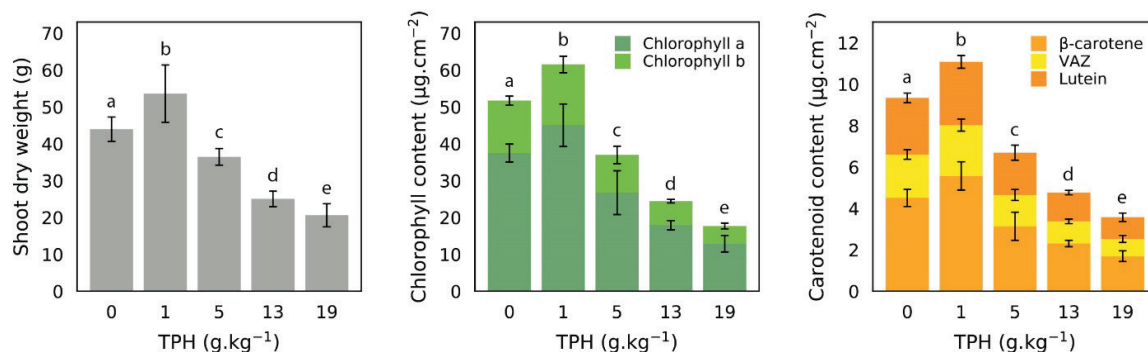


Figure 36. Shoot dry weight and leaf chlorophyll and carotenoid contents (mean \pm SD) of *C. alopecuroides* after 42 days of exposure to various levels of Total Petroleum Hydrocarbons (TPH) ($n = 11$ and 5 samples per treatment, respectively). Differences among treatments are denoted by different lowercase letters (ANOVA and Tukey post-hoc tests, $p < 0.05$). (VAZ: Violaxanthin + Antheraxanthin + Zeaxanthin.)

Alterations in leaf biochemistry and optical properties are very common for vegetation exposed to contaminants^{77,159,183,324}. They depend on multiple factors, especially the species, the

contamination type and concentration and the time of exposure, and have been previously observed on *C. alopecuroides* and other tropical species^{201,268,270}. In the case of oil, these alterations result from the reduction of water and nutrient availability in soils and uptake capacities of roots^{160,169,173}. Additional effects may come from the accumulation of hydrocarbons in tissues¹⁶², but they mostly involve low-carbon TPH, so they are unlikely to explain the responses observed in our experiment. The amplitude of these biochemical and spectral responses was determined by the level of contamination, which confirmed previous results obtained for higher levels of oil^{77,201}. The more the TPH, the lower the leaf pigment contents and the higher the reflectance (Figure 37). Consequently, leaf chlorophyll content fell to $17.7 (\pm 3) \mu\text{g}\cdot\text{cm}^{-2}$ under exposure to $19 \text{ g}\cdot\text{kg}^{-1}$ TPH, while it reached $51.7 (\pm 3.5) \mu\text{g}\cdot\text{cm}^{-2}$ for control at the end of the experiment (Figure 36). A similar trend was observed on carotenoids, and suggested being able to discriminate among the different levels of contamination using reflectance data.

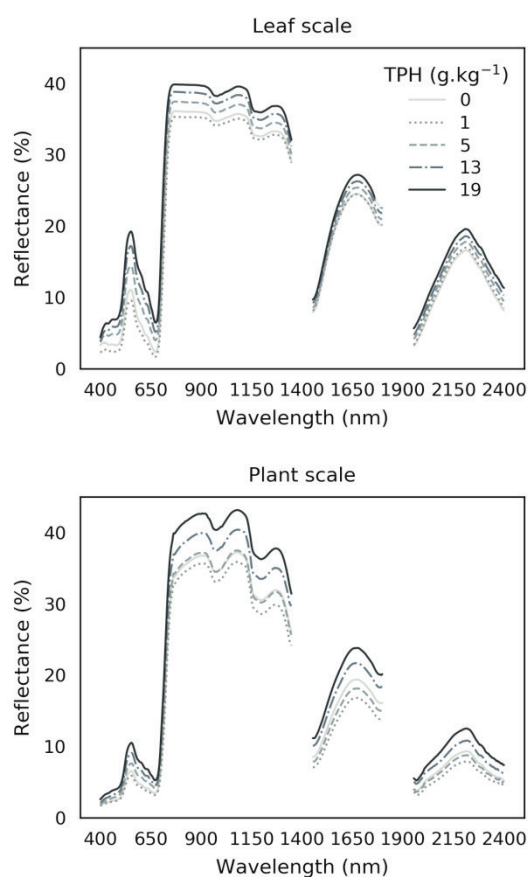


Figure 37. Mean reflectance of *C. alopecuroides* in the reflective domain after 42 days of exposure to various levels of Total Petroleum Hydrocarbons (TPH), at leaf and plant scales. Reflectance data from the 1350–1450 and 1800–1950 intervals were removed, because of low atmospheric transmission at plant scale.

In contrast to the other treatments, a stimulating effect was observed for plants exposed to $1 \text{ g}\cdot\text{kg}^{-1}$ TPH. Shoot dry weight and leaf chlorophyll and carotenoid contents were significantly higher for this treatment than those of the control ($p < 0.05$), and reflectance was slightly lower at leaf and plant scales. This particular response – called hormesis – has already been observed on the growth of other oil-exposed species in previous studies^{328,329,335}. However, its causes remain

poorly documented. To our knowledge, this study is the first to bring evidence of this phenomenon using reflectance measurements. In previous experiments, species were sometimes exposed to various levels of TPH^{77,186,205}. These experiments aimed to simulate accidental oil spills (*e.g.* pipeline leakage), by applying high doses of oil – mainly diesel or engine oil – to crop species, which are particularly sensitive to such contamination. Consequently, strong alterations in leaf biochemistry and reflectance have been observed regardless of the contamination level, so no hormesis has been noticed. The responses of *C. alopecuroides* observed in our study substantially differed, because we focused on a totally different context. When oil contamination persists in industrial brownfields and mud pits soils – such as that reproduced in our experiment, the most phytotoxic fractions of petroleum hydrocarbons (*e.g.* BTEX, PAHs) are almost absent from the mixture. The effects on oil-tolerant established vegetation are therefore less pronounced than those observed for other types of oil contamination with high proportion of these compounds (*e.g.* crude oil, diesel, gasoline), especially at low concentration³²⁴. Here, PAHs and BTEX were found below the detection limits (0.02 and 0.05 mg.kg⁻¹, respectively) at 1 g.kg⁻¹ of TPH and in the control soil, so they did not affect the plants negatively. Conversely, these compounds are likely to induce positive effects at very low concentration, but they cannot be the only contaminants responsible for the differences observed between the two treatments. Some other hydrocarbon fractions – which were absent from the control soil – enhance the availability of nutrients and organic matter for plants by stimulating microorganisms. Therefore, the very low concentrations of PAHs and BTEX combined to those of other hydrocarbon fractions might have induced the hormetic effect observed in this study.

7.3.2. Vegetation indices

7.3.2.1. Elastic net regressions

The VI succeeded in linking the biochemical and spectral responses of *C. alopecuroides* to oil contamination. 20 out of the 33 indices tested were strongly correlated to leaf pigment or water content ($R^2 \geq 0.7$) (Table 11). For most of them, at least three contributing pigments were identified, the others having not being retained by the ENET regression. Chlorophylls and β -carotene were often the most contributing ones, and provided some of the best models ($R^2 \geq 0.8$) when used together (*i.e.* without addition of other carotenoids). Therefore, these pigments explained most of the modifications of reflectance observed among the treatments in the VIS region. Other carotenoid pigments – especially lutein and violaxanthin – were frequently retained in the models (*e.g.* GM2 and PSSRb indices), thus confirming previous results obtained on another oil-tolerant species³²⁴. These pigments are usually of less contribution to leaf optical properties, because of the masking effect of chlorophylls, which are present at higher concentrations in leaves and share common light absorption features^{82,92}. Our study shows that they also contribute to the spectral response of vegetation to oil contamination.

Table 11. Results of the elastic net regressions performed on the 33 vegetation indices. For each index, the R^2 of the model is presented, along with the contributing set of biochemical parameters (in order of importance). (Chl a: Chlorophyll a, Chl b: Chlorophyll b, B-car: β -carotene, Lut: Lutein, Ant: Antheraxanthin, Vio: Violaxanthin, Zea: Zeaxanthin, LWC: Leaf Water Content.)

Index	R^2	Pigments
CARI	0.76	B-car, Chl b
CCI	0.73	B-car, Chl b, Chl a
CTR1	0.24	Vio, B-car, Chl b, Lut, Ant, Zea, Chl a
CTR2	0.81	Chl a, Chl b
GM1	0.84	B-car, Chl b, Chl a
GM2	0.87	Chl a, Chl b, B-car, Lut
LI1	0.65	Chl a, Chl b
LI2	0.18	Chl a
LI3	0.42	Vio, Lut, Ant, Zea, B-car, Chl b, Chl a
mCARI1	0.43	Ant, B-car, Lut, Chl b, Zea, Chl a
mCARI2	0.12	Chl a
mSR705	0.72	Chl a
MTCI	0.84	Chl a, Chl b
ND705	0.82	B-car, Chl b, Chl a
PRI1	0.54	B-car, Vio, Chl b, Zea, Ant, Chl a, Lut
PRI2	0.67	Chl b, Chl a
PRI3	0.63	Chl b, Chl a, Lut
PSRI	0.43	B-car, Chl b, Lut, Ant, Zea, Chl a
PSSRa	0.72	Chl a
PSSRb	0.88	Chl a, Chl b, Lut, B-car, Vio
PSSRc	0.83	Chl a
SIPI1	0.71	Chl a, Chl b
SIPI2	0.81	Chl a, Chl b
SIPI3	0.73	Chl a, Chl b
SR705	0.83	B-car, Chl b, Chl a
TCARI	0.68	B-car, Chl b, Chl a
TCARI_OSAVI	0.75	B-car, Chl b, Chl a
VOG1	0.75	B-car, Chl b, Chl a
VOG2	0.7	B-car, Chl b, Chl a
VOG3	0.71	B-car, Chl b, Vio, Chl a
DWSI	0.82	LWC

The close relationship between some indices and contributing pigments are illustrated in Figure 38. As observed, these indices exhibited strong link with single pigment content. ENET regression allowed identifying those contributing the most to index changes, in order of importance. These results confirmed the interest of combining pigment contents into multiple models for better understanding the effects induced by oil contamination on vegetation reflectance³²⁴. Most of the indices exploited reflectance around 550 and 700 nm, which was particularly affected by oil (Figure 37), so they were particularly adapted to our context of study. The same wavelengths also proved efficient for assessing soil contaminated by oil and by-products in other situations (oil extraction, pipeline leakages, etc.). Zhu *et al.*¹⁶⁶ exploited reflectance at 700 nm for distinguishing among various levels of alteration in pigment contents induced by phenanthrene contamination on *Suaeda salsa*. Likewise, Sanches *et al.*¹⁸⁵ used the same spectrum

region for assessing the effects of oil leakages on crops species. One of the main advantages of VI relies on their robustness. Some of them remain only little affected by bare soil and plant architecture, so they can be applied for tracking subtle alterations in leaf pigment contents induced by oil at different acquisition scales (*e.g.* leaf, plant and canopy)^{240,300}. For example, the same 33 indices succeeded in discriminating among various types of oil contamination from leaf to canopy scale in previous study³²⁴. Here, they were suitable for estimating TPH concentrations in soils.

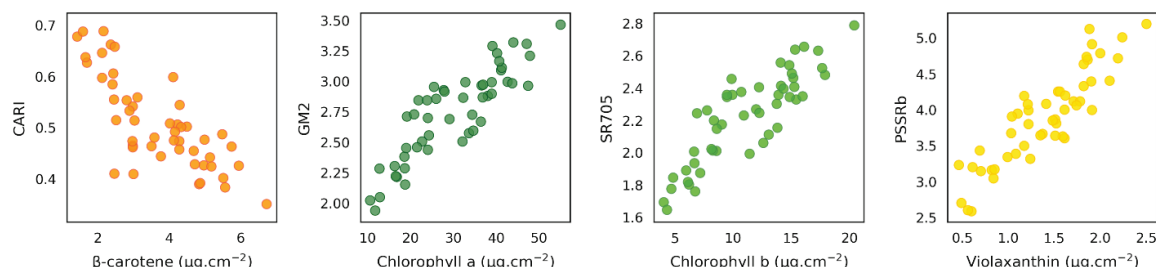


Figure 38. Relationship between vegetation indices and leaf pigment contents. Figures include data from leaves sampled on all the treatments, after 21 and 42 days of experiment ($n = 50$).

7.3.2.2. Random Forest regressions

RF regressions performed using the 33 VI provided accurate predictions of TPH concentrations, at leaf and plant scales. These results are presented in Figure 39. High R^2 values – respectively 0.96 and 0.95 at leaf and plant scales – indicated strong correlation between the measured and predicted concentrations on the test sets. Predictions of TPH were very close to the true ones, as indicated by the low RMSE values obtained at both scales. This accuracy was confirmed by RPD values greater than 4.4. RPD provides a better interpretation of predictions, because it compares the RMSE to the variability in the true TPH concentrations. RPD values above 2 or 3 are usually needed for considering the models reliable for use. The 33-VI based RF models were therefore particularly adapted for estimating TPH in our study. The analysis of residuals revealed that the highest level of contamination (19 g.kg^{-1}) was the most difficult to predict, as it was almost systematically underestimated (Figure 39). This may be caused by saturation in the spectral response of *C. alopecuroides* to this type of oil contamination, as described for other species in previous study. Confusions also arose at the lowest concentrations ($0 - 1 \text{ g.kg}^{-1}$). They highlighted the difficulty to detect the hormesis phenomenon using reflectance data, for which only little differences were observed among the two treatments (Figure 37). These results thus helped identifying the detection limit of our approach.

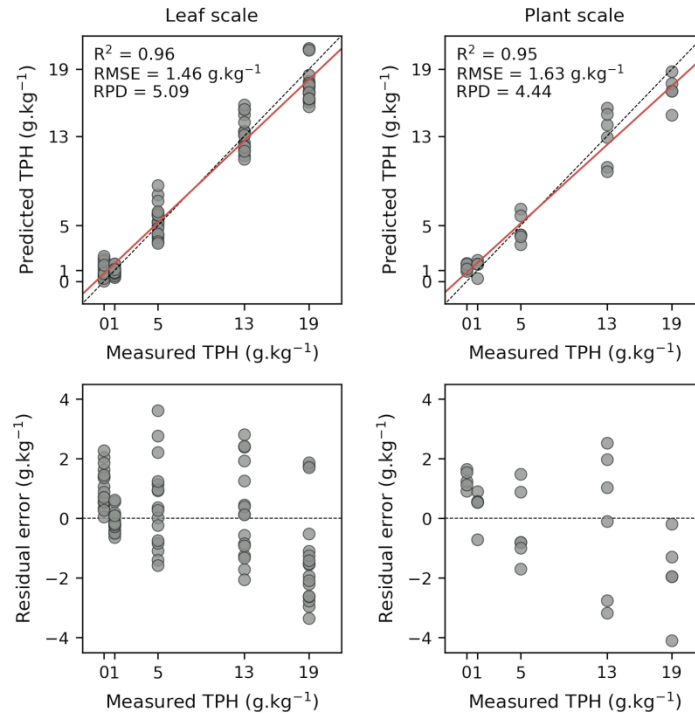


Figure 39. Comparison between the measured and predicted concentrations of Total Petroleum Hydrocarbons (TPH) using the 33 vegetation indices and the random forest regression, at leaf and plant scales (top figures), and residuals of the predictions (bottom figures).

The relative contribution of VI to the RF models was consistent from leaf to plant scale. Four out of the 33 indices contributed the most to TPH predictions ($> 75\%$, Figure 40). These indices were among those closely linked to leaf pigment contents ($R^2 > 0.7$, Table 11), and more precisely chlorophylls and β -carotene. They exploit reflectance around 550 and 700 nm, but also 670 – 680 and 750 nm, which helped enhancing the differences among the treatments. Only few wavelengths are therefore necessary for estimating TPH accurately using reflectance spectroscopy. Conversely, the other indices remained of little contribution to the models ($< 5\%$ each), especially those linked to additional carotenoids (*e.g.* lutein, violaxanthin), because they brought redundant information.

The consistence of the results between acquisition scales opens up promising perspectives for operational applications of our approach under natural conditions. Some of the main indices – especially the MTCI and TCARI / OSAVI – have been originally designed for imagery application purposes^{300,304}, and proved efficient for detecting oil contamination in tropical regions. For example, Arellano *et al.*⁶⁸ used the MTCI index to discriminate among control and oil-contaminated sites in the Amazon forest using hyperspectral satellite images. However, the application of these indices for estimating the level of contamination remained unexplored until now. This study was a first attempt in that direction, but further researches are needed to assess the reliability of our approach under natural conditions, using field measurements, and in the long term using multi- and hyperspectral imagery. High to very high spatial resolution might help achieving this, because the contamination can occur on a few square meters and be therefore difficult to detect using medium to low spatial resolution imagery. As discussed in section 7.3.1, the case of persistent low contamination in brownfields and mud pits makes the estimation of TPH very challenging, because of the composition of the contamination and the tolerance of the

species. Under natural conditions, additional difficulties should be considered before application of our approach. Vegetation established on brownfields and mud pits is exposed to a multitude of environmental factors that affect its health and reflectance (*e.g.* drought, waterlogging), especially in tropical regions with marked seasonality^{235,258}. These factors might overlap the effects induced by TPH, making their estimation more difficult. Thanks to recent advances, it seems however possible to discriminate among oil and other stressors using reflectance spectroscopy^{206,271,324}. Because of the selective growing conditions imposed, only few species are generally established on soils with oil²⁰, as observed on our study site. This makes the estimation of TPH feasible, provided that these species are oil-sensitive. Species richness is however very high in some tropical regions subject to oil contamination^{230,270}. Such diversity in plant species means important differences in sensitivity to oil among them. Some species are totally tolerant to oil, whereas others are affected regardless of the concentration. This makes the estimation of TPH more difficult, so it is necessary to identify the most suitable species before, as we did under controlled conditions.

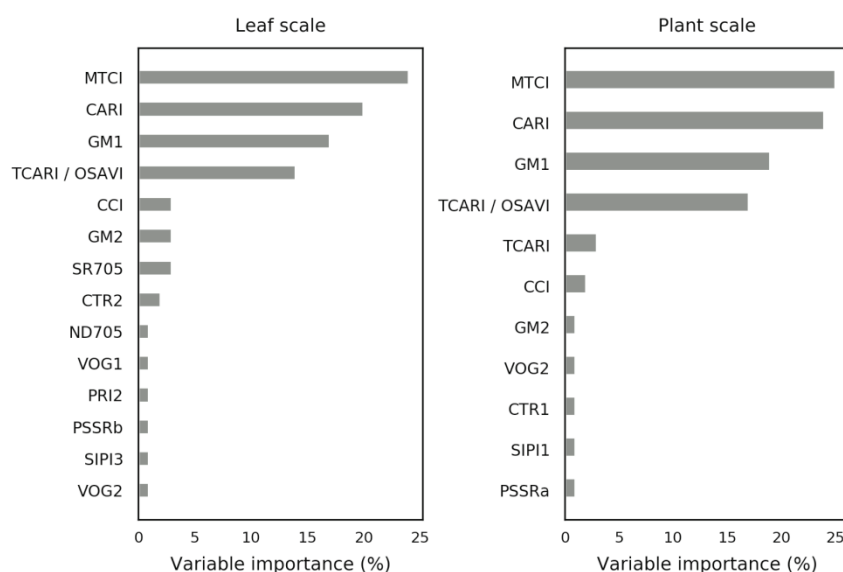


Figure 40. Relative contribution of the vegetation indices to the estimation of Total Petroleum Hydrocarbons (TPH) using random forest regression, at leaf and plant scales. Only indices with non-zero contribution are displayed.

7.4. Conclusion

This study aimed to quantify low TPH concentrations in soils using *C. alopecuroides* reflectance under controlled tropical conditions. Modifications in leaf biochemistry appeared after a long-term exposure, depending on the level of contamination. Pigment contents were reduced for TPH concentrations of 5 g.kg⁻¹ and above, whereas they increased at lower concentration when compared to control. These modifications were linked to the reflectance data through VI, which brought evidence of the implication of various carotenoids in the spectral response of the species to oil. The same indices succeeded in predicting TPH concentrations with good accuracy using RF regressions, at leaf and plant scales. Four out of the 33 indices tested were almost sufficient to achieve these predictions. This study emphasizes the potential of reflectance spectroscopy for assessing oil contamination in tropical regions with dense vegetation. More

specifically, the long-term response of *C. alopecuroides* to oil highlights its interest for assessing persistent contamination, for example after cessation of the oil production activity. Although our approach is at early stage and needs further improvements, we are convinced that it could soon result in operational imagery applications. The emergence of new satellite- and UAV-embedded hyperspectral sensors is sparking a growing interest by oil and gas companies, because they could help assessing oil contamination locally or at large scale. Our study under controlled conditions was the first necessary step prior to such applications. In its continuity, our upcoming research will focus on the adaptation of the approach to hyperspectral imagery with high spatial resolution, and its assessment in tropical region with heavy past oil and gas production activities.

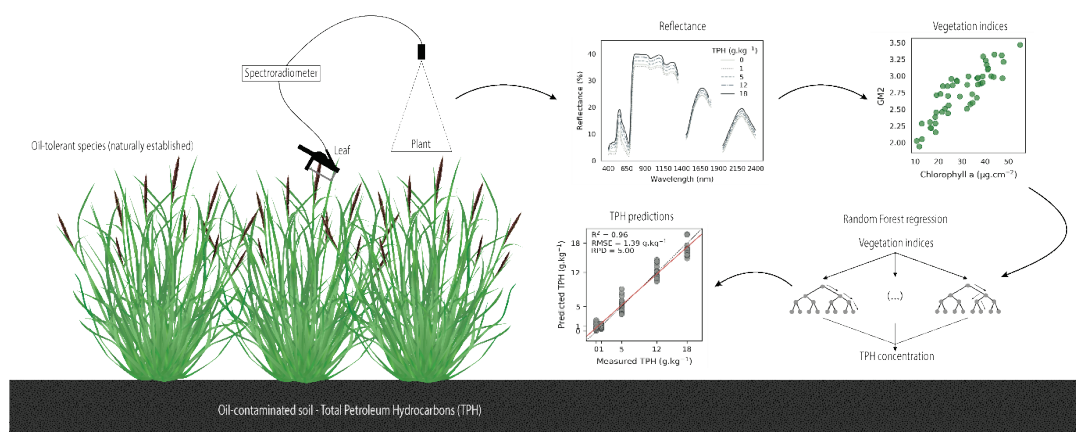
Acknowledgements

This interdisciplinary work was achieved in the frame of the NAOMI R&D project between TOTAL and the ONERA, with the support of Ecolab and Dynafor research units of Toulouse. The authors gratefully acknowledge E. Buffan-Dubau, D. Lambrigot and O. Berseille for their assistance in pigment analysis.

Declarations of interest: None.

Funding: Financial support of this work was provided by TOTAL.

Graphical abstract



SUMMARY AND CONCLUSION

This chapter outlined the experiments carried out in the first step of the thesis' approach. There are summarized in Table 12. These experiments distinguished themselves from those of the literature, because they were the first involving realistic oil mixtures and oil-tolerant species, and clearly linking the measured biochemical and spectral responses of the plants studied. In addition, the experimental design and the measurement protocol (*i.e.* leaf and plant scales) set up proved suitable for developing the methods under both temperate and tropical contexts. In the first experiment, a perfect detection of the contamination was achieved using stepwise LDA (accuracy of 100%) at leaf and plant scales after 61 days for both *R. fruticosus* and *C. alopecuroides*, in a simple two- or three-treatments experimental design. This experiment confirmed that the species identified in both temperate and tropical contexts were well-suited for the purpose of the thesis. It helped defining species' sensitivity to oil and the spectral regions to focus on for developing the methods. Further results obtained in the frame of this work can be found in ²⁶⁸. The following studies were dedicated to the development of the methods that could be applied to airborne hyperspectral imagery. Two methods relying on either RLR classification or RF regression were proposed for detecting and characterizing or quantifying contamination. These methods exploited VI closely linked to the biochemical alterations induced by crude oil and petroleum products. Among the major findings, it has been shown that chlorophylls were not the only pigment involved in the response of vegetation for both the temperate and the tropical species. While their contribution remained underestimated in previous studies, carotenoids – especially β -carotene and lutein – played an important role in the modifications of reflectance observed in the VIS under exposure to oil mixtures in both temperate and tropical contexts. The use of ENET regressions overcame the issue of multicollinearity among pigment contents, distinguishing the contribution of carotenes and xanthophylls to VI changes from that of chlorophylls. The methods exploiting VI provided encouraging results in detecting or quantifying the contamination for the next steps of this research. The first method (RLR) combining 14 selected VI and applied to *R. fruticosus* succeeded in discriminating control soils from various types of oil mixtures (accuracy $\geq 85\%$ after 18 days), and from a water-deficit condition. The second (RF), applied to *C. alopecuroides*, quantified low TPH concentrations (≤ 19 g.kg⁻¹) with good accuracy (RPD = 4.44 at plant scale) using 33 VI with no prior selection procedure. The performance of the methods was strongly determined by the time of exposure to the contamination, its composition (*i.e.* the oil mixture) and its concentration in the soil (in terms of TPH). For example, oil-contaminated mud pit soils were strongly impacting *R. fruticosus* at 25 g.kg⁻¹ TPH (Mud_B) and therefore easy to detect. Conversely, they were confused with control soils at 6 g.kg⁻¹ TPH (Mud_A) after a prolonged time of exposure (32 days). Regarding *C. alopecuroides*, 42 days of exposure were necessary to achieve accurate quantification of TPH. The quantification of TPH was not in the scope of the method from the temperate context (RLR), and it was not possible to apply that developed on *C. alopecuroides* to *R. fruticosus*, because these species have not been exposed to the same contamination type. A detection limit of 6 and 1 g.kg⁻¹ TPH was identified in the temperate and tropical contexts, respectively. Consistent results were observed among the different acquisition scales (*i.e.* leaf, plant, etc.) for the two species, by taking advantage of the robustness of VI to upscaling, with or without prior selection procedure. Focusing on the temperate context, a first validation of the detection method in the field was performed at canopy scale and provided an accurate detection of oil-contaminated mud pit soils (accuracy = 90%).

Table 12. Summary of the experiments carried out in the first step of the approach, and the developed methods. (TPH: Total Petroleum Hydrocarbons; HM: Heavy Metals; LDA: Linear Discriminant Analysis; RLR: Regularized Logistic Regression; RF: Random Forest; RMSE: Root Mean Squared Error; RPD: Residual Predictive Deviation.)

	Experiment 1		Experiment 2	Experiment 3
<u>Experimental design:</u>				
Context	Temperate	Tropical	Temperate	Tropical
Species	<i>R. fruticosus</i> L.	<i>C. alopecuroides</i> (L.) and <i>P. virgatum</i> L.	<i>R. fruticosus</i> L.	<i>C. alopecuroides</i> (L.)
Duration (days)	61	61	32	42
Number of treatments	3	2	6	5
TPH concentration ^d	36 g.kg ⁻¹	14 g.kg ⁻¹	6 and 25 g.kg ⁻¹	1, 5, 13 and 19 g.kg ⁻¹
<u>Scales of reflectance acquisitions:</u>				
	Leaf and plant	Leaf	Leaf, plant and multi-plant (+ canopy <i>in situ</i>)	Leaf and plant
<u>Proposed method:</u>				
Purpose	Feasibility of TPH detection		Detection of TPH + HM mixtures	Quantification of TPH
Principle	Spectrum transformation + Stepwise LDA classification		Vegetation index selection + L ² -RLR classification	Vegetation index + RF regression
Main results	Up to 100% classification accuracy after 61 days		Accuracy ≥ 85% for controls after 18 days Field validation	RMSE ≤ 1.63 g.kg ⁻¹ and RPD ≥ 4.44 after 42 days
Remarks	High contribution of pigment- and water-related wavelengths		Accuracy strongly influenced by mixture composition, exposure time and acquisition scale	Hormesis effect at 1 g.kg ⁻¹ of TPH

These results confirmed the suitability of the methods for assessing oil contamination, under controlled conditions. The next step of the approach aimed at assessing their performance on airborne hyperspectral images, in order to demonstrate the potential of hyperspectral imagery for monitoring oil contamination, from the detection to the quantification. For this purpose, we focused on the temperate context (*R. fruticosus*), because of the large amount of validation data collected on the mud pits. A method of quantification adapted to the temperate context was therefore implemented, completing that of detection developed under controlled conditions. The next chapter describes its principle and the evaluation of the methods on airborne hyperspectral imagery, as well as their perspectives of operational use in oil exploration and contamination monitoring.

^dFor the contaminated treatments only.

CHAPTER THREE: APPLICATION AND EVALUATION OF THE METHODS UNDER NATURAL CONDITIONS, FROM FIELD SCALE TO AIRBORNE HYPERSPECTRAL IMAGERY

INTRODUCTION

The detection and the quantification of oil contamination from airborne hyperspectral imagery represented the main aim of this thesis. Prior to this last step – which was addressed in the temperate context, a validation of the methods in the field was necessary, in order to complete the multiscale approach. A first successful validation was achieved using the RLR-based method in the field on *R. fruticosus* (see 6.3.5), and emphasized the potential of VI for detecting and characterizing oil. Thus, the next step of the approach was dedicated to developing a method for quantifying TPH adapted to this context. In order to anticipate its future application to airborne imagery, this method was developed directly in the field. For this purpose, an exhaustive 14-months study was carried out at leaf scale on the oil-contaminated mud pit site. This study relied on monitoring several species of the site, including *R. fruticosus*, and linking its optical properties to varying levels of TPH in the soil. The proposed method of quantification exploited the PROSPECT model for retrieving leaf biochemical parameters, which had never been used before for similar application – including those focusing on other contaminants. This method was tested over a complete seasonal cycle and compared to others deriving from the literature (VI, spectrum transformation and multiple regression). This study was published in a peer-reviewed journal. It is presented in the first section of this chapter. In a second time, the PROSPECT-based method was adapted to canopy-scale measurements performed in the field and to the airborne hyperspectral image using the PROSAIL model. A second method of TPH quantification was proposed, based on coupling spectrum transformations and ENET regression. The methods were assessed on the image along that of detection (RLR), on various uncontaminated and mud pit sites identified in the field. This consisted in first detecting (RLR-based method) and then quantifying oil (PROSAIL- and ENET-based methods). For this purpose, the methods were calibrated on two sites and applied to other spatially-independent sites. The location of *R. fruticosus* on the image was supposed to be known; its discrimination with other species being out of the scope of this thesis. This final step helped determining the limits of the methods at high spatial resolution, and the perspectives for future operational applications. This work will be submitted to a peer-reviewed journal, and forms the second section of this chapter.

8. Fourth article: “Application of PROSPECT for estimating Total Petroleum Hydrocarbons in contaminated soils from leaf optical properties”

Article reference:

Lassalle, G.; Fabre, S.; Credoz, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Application of PROSPECT for Estimating Total Petroleum Hydrocarbons in Contaminated Soils from Leaf Optical Properties. *J. Hazard. Mater.* **2019**, 377, 409–417. <https://doi.org/10.1016/j.jhazmat.2019.05.093>.

Abstract

Recent advances in hyperspectral spectroscopy suggest making use of leaf optical properties for monitoring soil contamination in oil production regions by detecting pigment alterations induced by Total Petroleum Hydrocarbons (TPH). However, this provides no quantitative information about the level of contamination. To achieve this, we propose an approach based on the inversion of the PROSPECT model. 1620 leaves from five species were collected on a site contaminated by 16 to 77 g.kg⁻¹ of TPH over a 14-month period. Their spectral signature was measured and used in PROSPECT model inversions to retrieve leaf biochemistry. The model performed well for simulating the spectral signatures (RMSE < 2%) and for estimating leaf pigment contents (RMSE ≤ 2.95 µg.cm⁻² for chlorophylls). Four out of the five species exhibited alterations in pigment contents when exposed to TPH. A strong correlation was established between leaf chlorophyll content and soil TPH concentrations ($R^2 \geq 0.74$) for three of them, allowing accurate predictions of TPH (RMSE = 3.20 g.kg⁻¹ and RPD = 5.17). The accuracy of predictions varied by season and improved after the growing period. This study demonstrates the capacity of PROSPECT to estimate oil contamination and opens up promising perspectives for larger-scale applications.

Keywords: Soil contamination, Leaf optical properties, PROSPECT, Pigment, Total petroleum hydrocarbons

8.1. Introduction

In the last decades, increasing efforts have been made to improve the monitoring of past and current industrial activities^{41,43,336}. The assessment of health and environmental risks remains a major issue as it requires efficient detection of contamination deriving from production facility failures and bad practices^{36,46,337}. For this purpose, new approaches using hyperspectral spectroscopy have recently emerged in the oil and gas industry^{338,339}. In exploiting the light absorption features of oil, these approaches have proved efficient for detecting and quantifying Total Petroleum Hydrocarbons (TPH) from the spectral signature of soils^{63,64}. However, their applicability remains limited in vegetated areas, since only few radiations get through the foliage. As an alternative, recent studies have proposed exploiting the spectral signature of vegetation to detect reflectance changes induced by the presence of oil in soils^{20,68,77,271}.

The biophysical and biochemical parameters driving leaf optical properties are likely to be affected when vegetation is exposed to TPH^{162,169,183}. This includes alterations in leaf anatomy (*e.g.* cuticle and parenchyma thickness), pigment and water contents^{169,183,193}. These effects mainly result from the modification of soil properties and the reduction of root uptake capacities^{154,160,173}. They vary according to the species, the level of contamination and the duration of exposure. Alterations in leaf biophysical and biochemical parameters lead to reflectance changes, which help discriminate between healthy and affected vegetation^{159,271}. Previous work focused on pigment alteration, since it is responsible for the increase in reflectance typically observed in the visible (VIS) when plants are exposed to TPH^{230,268,324}. The amplitude of this increase is influenced by the level of contamination^{77,340}. This suggests that leaf pigments could be used for detecting and quantifying oil contamination in vegetated areas. However, no method has been yet proposed in literature. Previous studies focused on quantifying heavy metal contamination deriving from mining or agricultural activities, a very different case of soil contamination^{203,341}.

Leaf pigments undergo significant changes throughout the year, depending on the season^{97,103,342}. Seasonal variations in pigment contents are therefore likely to hide the alterations induced by TPH. This can cause confusion and currently represents a major limitation for TPH detection under natural conditions. All the more so that not all species are good indicators of TPH, and their sensitivity to these contaminants is likely to affect the relative influence of TPH and seasons on leaf pigments and reflectance^{29,201,271}.

The species established around industrial facilities are particularly tolerant to high levels of TPH^{230,270}. This makes it difficult to detect and estimate TPH, as this implies being able to differentiate small alterations in pigment contents using leaf optical properties. Many approaches have been developed in hyperspectral spectroscopy for estimating leaf pigment contents^{78,240,343,344}. The physically-based model PROSPECT⁷⁰ proved to be one of the most reliable approaches across a wide range of ecological contexts^{73,228}. The model has been recently improved and now enables pigment separation^{82,108}. This is of major importance for detecting and estimating TPH, since not all pigments are affected in the same way by contamination^{268,324}. PROSPECT could therefore be used for detecting and estimating TPH-induced changes in leaf pigment contents and consequently for determining the level of soil contamination^{230,270}.

This study proposes to assess the potential of leaf optical properties for predicting TPH concentrations in the soil, under natural temperate conditions. The proposed method is based on the retrieval of leaf pigment contents using the PROSPECT model, which are thereafter linked to TPH concentrations. The study was carried out on a range of plant species over a complete

seasonal cycle in order to identify the most suitable conditions for predicting TPH. The proposed method was compared to other approaches typically used in similar studies.

8.2. Materials and Methods

8.2.1. Study site and species

The study was carried out on an industrial brownfield located in a temperate region. The site was subject to intensive oil and gas activities for 20 years, and production residues contaminated by TPH and heavy metals (HM) accumulated in the soil. Since then, it has been colonized by vegetation. The study focused on the predominant species, *Rubus fruticosus* L., which is generally well-established on industrial sites in temperate regions. Moreover, this species has already proved great potential in hyperspectral spectroscopy for detecting oil-induced pigment alterations^{29,271,324}. Other species found on the brownfield, including *Quercus pubescens* Wild., *Populus x canadensis* Moench., *Cornus sanguinea* L. and *Buddleja davidii* Franch., were also tested for their sensitivity to TPH. The same species were also studied on an uncontaminated site (hereafter described as “control site”) for comparison purposes. This site was located a few kilometers away from the brownfield and exhibited similar soil properties.

8.2.2. Field sampling and measurements

A mesh of 23 vegetated plots was defined, covering most of the brownfield. Plots consisted in areas of one meter in diameter and were spaced by 10 meters. The upper soil layer (0-20 cm) was analyzed for contaminants at the center of each plot. Analyses revealed C₁₀-C₄₀ TPH concentrations ranging from 16 to 77 g.kg⁻¹, the main contributors being in of the C₂₁-C₄₀ fraction (dense hydrocarbons). These concentrations correspond to a moderate to high level of exposure for plants, in comparison with other studies^{183,340}. High concentrations of HM and Polycyclic Aromatic Hydrocarbons (PAH) were also found. The detailed soil analyses are given in the Supporting Information.

Leaf samples were collected on site. For each species found in each plot, three young leaves were randomly sampled on different individuals. Their reflectance was measured in the [400:2500] nm domain using an ASD FieldSpec 4 Hi-Res spectroradiometer (Malvern Panalytical, Malvern, UK) and a leaf-clip with an internal light source. These measurements avoid the influence of leaf orientation and density and the effects of soil and atmosphere. Leaf radiance was acquired on a small black background panel and converted to reflectance as described in ³²⁴. Measurements were carried out from June 2017 to October 2018, on a seasonal basis (Table 13). This included six different sampling dates, with samples taken in contrasted environmental conditions and at different phenological stages. On the same dates, additional spectral signatures were acquired on the control site. There were no TPH detected in the control soil and low HM concentrations corresponded to the geochemical background. Only *B. davidii* and *R. fruticosus* were found on the brownfield in March 2018. A total of 1620 spectral signatures had been collected on the brownfield by the end of the study. 152 leaves were used for determining water (C_w), dry matter (C_m) and pigment contents, following the procedures described in ^{141,276}. Pigments were analyzed by High Pressure Liquid Chromatography (HPLC). The chlorophyll content (C_{ab}) was expressed as the sum of chlorophylls a and b, and the carotenoid content (C_{xc}) as the sum of β-carotene and xanthophylls. These analyses are detailed in the Supporting Information. They served to validate PROSPECT inversion outputs, as described in section 8.2.3.

Table 13. Description of the sampling campaigns carried out on the brownfield. Only *R. fruticosus* and *C. sanguinea* were found on all 23 plots, except in March 2018.

Date	Species found (n)	Spectral signatures (n)	Season
June 2017	5	309	Summer (heat waves)
October 2017	5	309	Autumn
March 2018	2	75	Late winter
April 2018	5	309	Spring
June 2018	5	309	Summer
October 2018	5	309	Autumn

8.2.3. PROSPECT inversion and validation

The PROSPECT model was developed for simulating the hemispherical reflectance and transmittance of leaves based on their known structural and biochemical parameters⁷⁰ (listed in Table 14). Inversion of the model allows these parameters to be retrieved from the leaf spectral signatures^{82,228}. Although PROSPECT was initially calibrated for measurements carried out with an integrating sphere, those performed with a leaf-clip can also be used^{345,346}. The inversion consists in finding the set of structural and biochemical parameters – denoted by the vector θ – that best simulates the leaf spectral signature measured⁸². This is achieved by successively testing sets of parameters and retaining the one that minimizes the cost function J , defined as:

$$J(\theta) = \sum_{\lambda_{\min}}^{\lambda_{\max}} (R_{\text{meas}}(\lambda) - R_{\text{sim}}(\lambda, \theta))^2 \quad (1)$$

where R_{sim} and R_{meas} stand for the simulated and measured reflectance, respectively^{345–347}. The cost function is computed in the $[\lambda_{\min}; \lambda_{\max}]$ spectral range⁸². Since only leaf reflectance was measured in our study, we simplified the original function that included transmittance⁷⁰.

PROSPECT inversions were first used to estimate the biochemical parameters of the 152 leaves sampled for analysis. Anthocyanins were recently added to the latest version of the model (PROSPECT-D), but their concentrations remained below the detection limit for all the species in our study. We therefore used the PROSPECT-5 version to perform the inversions as described in^{88,148,348}. The six leaf parameters (N , C_{ab} , C_{xc} , C_{bp} , C_{w} and C_{m}) were simultaneously determined through minimization of the cost function (1) over the entire reflective domain (400–2500 nm). (N corresponds to the leaf structure parameter and mainly influences reflectance in the near-infrared domain (750 – 1300 nm)^{71,119}). Inversions were performed using an iterative optimization approach. Many methods, such as look-up tables and conventional optimization algorithms have been tested for this purpose in previous studies^{349,350}. However, these methods can sometimes pose problems. For example, local minima and algorithm nonconvergence lead to biased or inaccurate estimation of biochemical parameters^{351,352}. To avoid this, we used the differential evolution (DE) algorithm³⁵³. The DE combines the advantages of evolutionary and genetic algorithms, which proved to be good candidates for solving inversion problems in remote sensing^{354–356}. To avoid extensive computational time, the parameters tested for optimization were constrained to ranges set out in Table 14. For each sample, the best set of leaf parameters was retained. Inversion outputs were validated in two ways. The Root Mean Squared Error

(RMSE), the BIAS and the Standard Error of Prediction Corrected from the bias (SEPC) were used to assess the fit between the simulated and the measured spectral signatures on a wavelength-by-wavelength basis⁸². The estimated parameters were compared to biochemical analysis using the same metrics and the R^2 . After validation of the model's performance, we applied the inversion of PROSPECT to the 1620 leaves sampled in order to determine their pigment contents. The same procedure was applied to the leaves sampled on the control site.

Table 14. Parameter bounds used to perform PROSPECT inversions. C_{ab} and C_{xc} bounds were extended from those of literature to factor in seasonal variations of leaf pigment contents.

Parameter	Unit	Range
Leaf structure (N)	-	1 - 5
Chlorophylls (C_{ab})	$\mu\text{g.cm}^{-2}$	1 - 100
Carotenoids (C_{xc})	$\mu\text{g.cm}^{-2}$	1 - 50
Brown pigments (C_{bp})	$\mu\text{g.cm}^{-2}$	0.01-1
Water (C_w)	g.cm^{-2}	0.001 - 0.1
Dry matter (C_m)	g.cm^{-2}	0.001 - 0.1

8.2.4. Variability of leaf pigment contents and TPH estimation

The study then focused on the chlorophyll and carotenoid contents retrieved from PROSPECT inversions. Their seasonal variability was examined for each species using one-way ANOVA and Tukey post-hoc tests to look for significant differences among sampling dates. Pigment contents were then linked to soil TPH concentrations in order to assess their ability to predict the level of contamination. For this purpose, data were divided into 50% train and 50% test sets using the Kennard-Stone algorithm, which ensures that representative samples are selected^{357,358}. Univariate regression models were adjusted between pigment contents and TPH concentrations on the train set^{340,358}. The models were then applied to the test set to predict TPH concentrations. This procedure was carried out on the five species successively, on each sampling date. The accuracy of the predictions was assessed by calculating the R^2 and the RMSE on the train and test sets. In addition, we computed the BIAS, the SEPC and the Residual Predictive Deviation (RPD) on the test sets to complete the accuracy assessment. The RPD is defined as the ratio of the standard deviation of the measured TPH concentration to the RMSE computed between the predicted and measured concentrations²⁰³. The interpretation of RPD values strongly depends on the context^{203,359}. RPD values above 2 and 3 are often considered as good and excellent predictions of TPH, respectively, and values below 2 as poor ones. A comprehensive flowchart of the method is presented in Figure 41.

In order to assess the accuracy of our approach, the results were compared to those obtained with approaches directly linking TPH concentrations to leaf reflectance in the VIS, including pigment-related vegetation indices (listed in ²⁹²) and partial least square regression (PLSR) with reflectance transformation (derivatives, *continuum* removal, etc.)^{203,271}. These approaches usually perform well for tracking reflectance changes induced by contaminants^{68,166,271,324}. The approaches were compared based on the RMSE and RPD values obtained on the test sets.

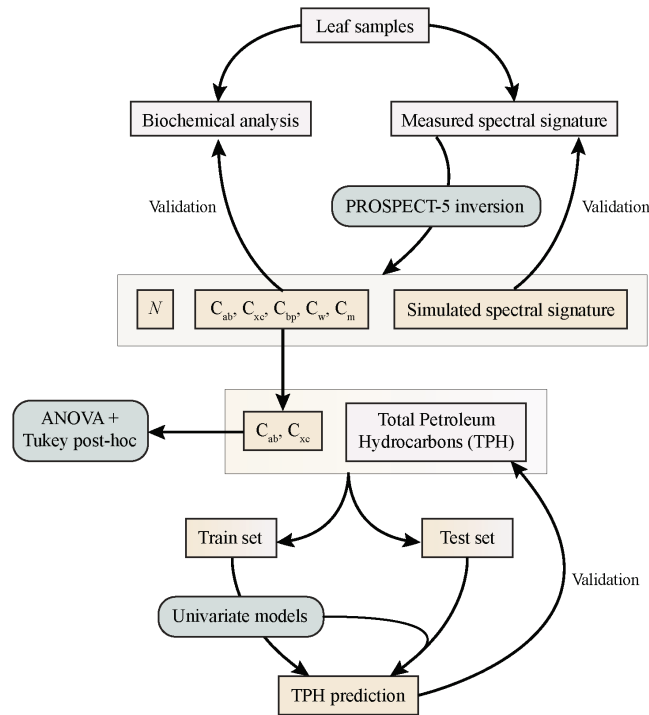


Figure 41. Flowchart of the method presented in this study (N : structure parameter, C_{ab} : Chlorophyll content, C_{xc} : Carotenoid content, C_{bp} : Brown pigments, C_w : Water, C_m : Dry matter). The spectral signatures measured were used to retrieve leaf structural and biochemical parameters using PROSPECT-5 inversions. The estimated parameters were compared to chemical analyses, and the simulated spectral signatures were compared to the measured ones. The estimated pigment contents (C_{ab} and C_{xc}) obtained for each season were compared, using ANOVA and Tukey post-hoc tests, and were then used to predict TPH concentrations with the help of univariate models. The predicted TPH concentrations were then compared to those obtained from soil analyses.

8.3. Results and Discussion

8.3.1. Measured spectral signatures

Since the five species are well-established on the brownfield, they are naturally tolerant to high levels of TPH contamination^{29,324,360}. Consequently, the leaves did not exhibit marked visible stress symptoms, and their spectral signatures strongly differed from those observed for species highly affected by oil^{77,159,186} (Figure 42). The spectral signatures did however vary temporally and among species during the study.

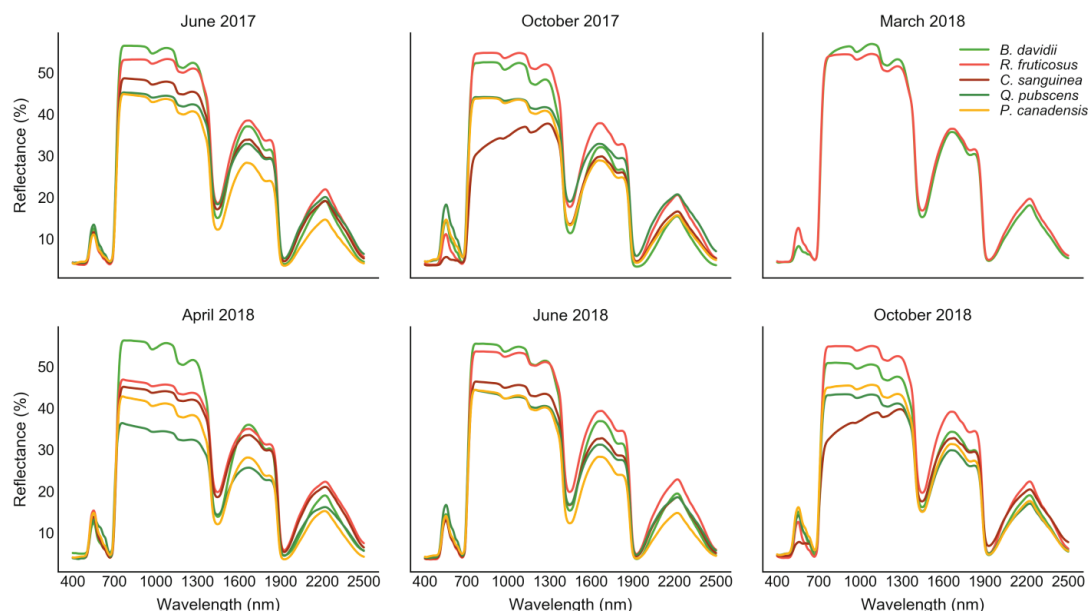


Figure 42. Seasonal evolutions of the mean spectral signatures for the five species on the brownfield.

The variability among the five species highlighted natural differences in leaf anatomy and biochemistry and in sensitivity to TPH exposure^{78,123,270}. When considering a single species, the spectral signatures also varied temporally with season changes, especially in the VIS and the red-edge regions. The summer of 2017 was characterized by repeated heat waves (June), and reflectance was particularly low for all species, varying between 550 and 680 nm and affecting the red-edge region on the following sampling dates, an indication of changes in leaf pigment contents^{166,185}. Likewise, reflectance in the near- and short-wave infrared regions was also affected by the seasons, a sign of changes in leaf anatomy and water content^{111,119,133}. All species do not share the same sensitivity to environmental changes from season to season^{103,342}. Only small seasonal differences in the leaf reflectance of *R. fruticosus* were observed in relation to the other species. In comparison, *C. sanguinea* underwent substantial changes in October, exhibiting dry blackish leaves with very low VIS reflectance. This unknown symptom differed from the autumn leaf coloration described on *Cornus* species¹⁰⁶.

The spectral signatures observed on the brownfield are thus the result of leaf anatomy and biochemistry, which are strongly influenced by the species' sensitivities to TPH and the seasons. This explains some of the differences observed among the species and is a significant factor in the prediction of TPH, as described in sections 8.3.3 to 8.3.4.

8.3.2. Evaluation of PROSPECT inversions outputs

8.3.2.1. Simulated spectral signatures

PROSPECT performed well for simulating the spectral signatures of leaves measured during the study. The DE algorithm always converged during inversions and the RMSE computed between the simulated and measured signatures ranged from 0.2 to 2% reflectance across the entire spectrum for the five species. RMSE, BIAS and SEPC values were consistent with those observed for other datasets^{82,228,361}. These results are presented in the Supporting Information, along with examples of simulated and measured signatures. Although spectral signatures varied

according to the sampling date, the RMSE remained almost unchanged throughout the study. PROSPECT therefore proved efficient for tracking seasonal changes in leaf biochemistry, as suggested in previous studies³⁴².

8.3.2.2. Retrieved biochemical parameters

Chemical analyses performed on the 152 leaves indicated that PROSPECT accurately estimated biochemical parameters, regardless of the species. An illustration is provided for *R. fruticosus* in Figure 43. The RMSE between the simulated and measured signatures was particularly low in the VIS (< 1.3%), and resulted in accurate retrieval of chlorophyll and carotenoid contents. Carotenoids were more difficult to estimate because of predominant absorption of chlorophylls in the VIS⁸². However, the RMSE remained below or equal to 1.82 $\mu\text{g}\cdot\text{cm}^{-2}$ for these pigments, which still indicated good accuracy. As suggested in previous work, separating carotenoids into carotenes and xanthophylls in PROSPECT might improve estimations, but this is not yet possible in the current version of the model^{82,92}. Leaf water and dry matter contents were also correctly estimated (Figure 43). All these results were consistent with those obtained on other datasets^{92,345–347,361}.

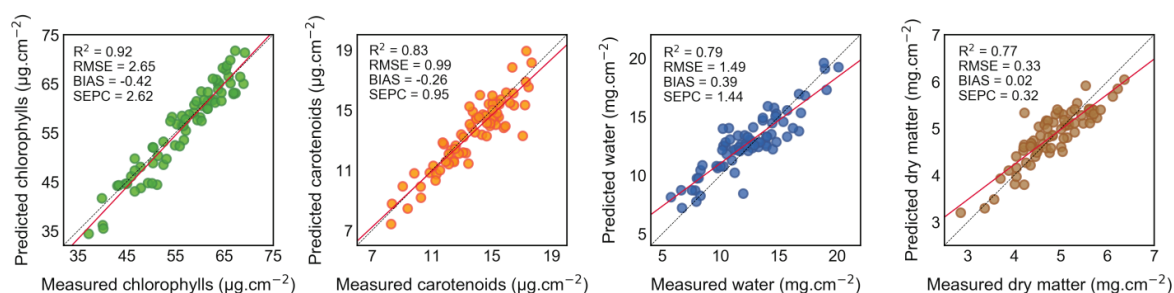


Figure 43. Comparisons of the predicted and measured biochemical parameters of *R. fruticosus* leaves. Water and dry matter contents were expressed in $\text{mg}\cdot\text{cm}^{-2}$ for graphical convenience.

8.3.3. Seasonal and interspecific variability of leaf pigments

The chlorophyll contents estimated throughout study ranged from 9.3 to 90 $\mu\text{g}\cdot\text{cm}^{-2}$ on the brownfield, and were higher for all species on the control site with the exception of *B. davidii*. This species was not affected by exposure to TPH and is therefore not a suitable indicator of oil contamination. Detailed results can be found in the Supporting Information section. The reduction in chlorophyll content observed for the other species is commonly described on oil-exposed vegetation^{183,185,196,270}. Oil contamination reduces water and nutrient availability for plants as well as root uptake capacities^{160,169,173}. This results in anatomical and biochemical alterations in leaves, including a decrease in pigment contents^{162,183}. On the brownfield, chlorophyll content significantly varied among species ($p < 0.05$). The chlorophyll contents of *R. fruticosus* and *B. davidii* systematically attained the highest values, whereas those of *Q. pubescens* and *P. canadensis* never exceeded 40 $\mu\text{g}\cdot\text{cm}^{-2}$ on average. Chlorophyll content also varied significantly for each species according to sampling dates ($p < 0.05$), as suggested by the temporal evolution of spectral signatures (Figure 42). These variations were however less marked than those of other temperate species³⁴². Some smaller differences among species were observed on the control site, indicating natural differences in pigment contents^{124,362,363}. Pigment contents

on the brownfield were thus expressed as percentages of those of the control to factor in these differences and to better compare species' sensitivities to TPH exposure (Figure 44).

The chlorophyll content of *C. sanguinea* leaves taken from the brownfield remained close to that of the control site on all dates, indicating this species' high tolerance to oil contamination (Figure 44). *Q. pubescens* was the most sensitive species, regardless of the season. The leaf chlorophyll content almost remained below 60% of that of the control for this species. *R. fruticosus* and *P. canadensis* were more tolerant to TPH but their pigment contents still exhibited severe alterations (around 70% in June and October). This is consistent with the findings of previous studies^{268,324}. Such differences in sensitivity among species have already been reported in tropical regions²⁷⁰, but the factors influencing this sensitivity remain poorly understood. Oil contamination, as found on our study site, is made of a complex cocktail of contaminants – mainly TPH and HM. All these contaminants are likely to affect leaf biochemistry, depending on each species' sensitivity to each of them. *R. fruticosus*, for example, is particularly tolerant to HM^{274,307}. Several species, including *P. canadensis*, developed detoxification mechanisms to protect themselves against the effects of HM and PAH^{165,191,360}.

Q. pubescens, *P. canadensis* and *R. fruticosus* exhibited a decrease in relative chlorophyll content in April (spring) ($p < 0.05$), after the appearance of new leaves (Figure 44). Their sensitivity to TPH exposure was therefore more pronounced during this period, but remained unchanged between June and October in 2017 and 2018, despite important differences in environmental conditions among seasons and years. Only *R. fruticosus* and *B. davidii* were found on the brownfield in March 2017, and displayed chlorophyll content very similar to that of the control site.

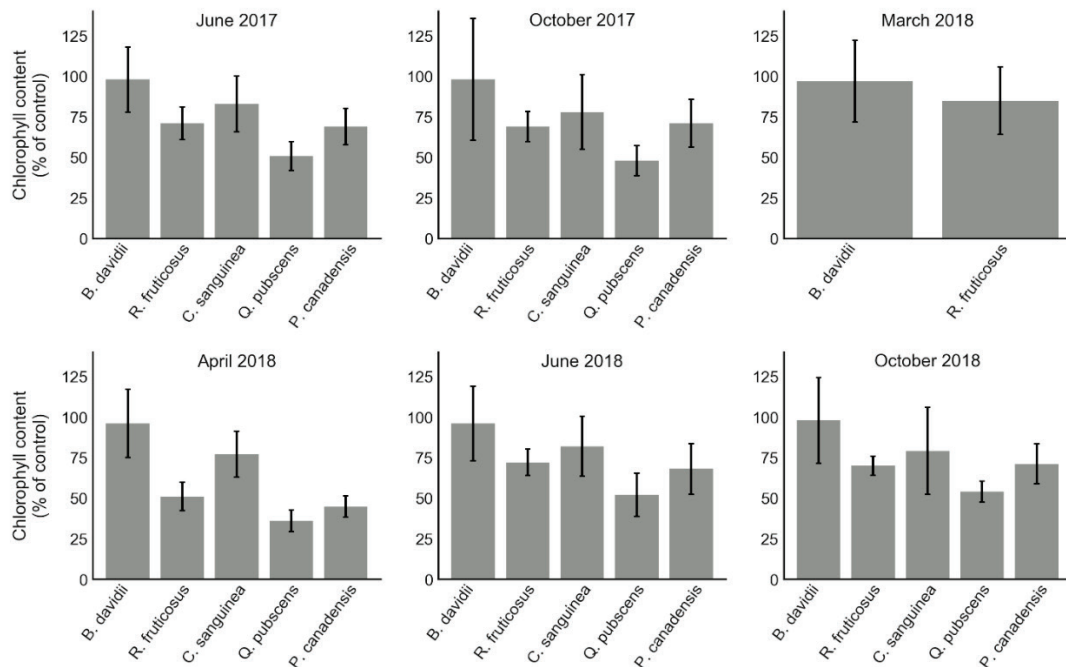


Figure 44. Temporal evolutions of the relative chlorophyll contents of leaves retrieved from PROSPECT inversions on the brownfield, for the five species.

Leaf carotenoid content also varied among species, but did not always vary significantly from one season to the next. These results are shown in the Supporting Information. According to

Archetti *et al.*⁹⁷, seasonal variations in leaf color – and thus in leaf reflectance – tend to result from changes in chlorophyll content rather than from carotenoid *de novo* synthesis. Carotenoid content was up to 60% lower on the brownfield than on the control site, indicating alterations induced by exposure to TPH. The observations made regarding relative chlorophyll content also apply to carotenoids. Previous studies mentioned that carotenoid pigments (*i.e.* carotenes and xanthophylls) are not affected in the same way by oil contaminants^{268,324}. For example, the lutein content might increase when plants are exposed to TPH, but the zeaxanthin content might be altered. Consequently, separating carotenoid pigments would benefit the detection of TPH-affected vegetation, but remains difficult when using hyperspectral spectroscopy^{82,92}.

The degree of pigment alteration therefore reflected species' sensitivity to TPH, as suggested by the spectral signatures observed on the brownfield (Figure 42). Results indicate that this sensitivity varies by season, which is likely to affect TPH predictions, as described in section 8.3.4.

8.3.4. Prediction of TPH from leaf pigments

Leaf pigment contents from the five species were tested to estimate TPH concentrations in the soil. The results are summarized in Table 15. A strong relationship was observed on the train sets between leaf chlorophyll content and TPH concentrations for *R. fruticosus*, *Q. pubescens* and *P. canadensis* in June and October (R^2 train ≥ 0.74). The more contaminated the soil, the lower the chlorophyll content (Figure 45). This type of plant response has frequently been reported under experimental conditions^{162,169}, but never in a natural context and using hyperspectral spectroscopy. Leaf chlorophyll content helped predict TPH concentrations with good (RPD > 2) or excellent (RPD > 3) accuracy, as indicated by the results obtained on the test sets (Table 15). Conversely, poor predictions were observed for *B. davidii*, confirming the absence of effects induced by TPH on this species. Over the entire study, the best predictions were obtained with *R. fruticosus* and *Q. pubescens*, for which the lowest RMSE – 3.63 and 3.20 g.kg⁻¹ respectively – and the highest RPD – 4.10 and 5.17 respectively – were observed (Figure 45). Successful predictions were also achieved with *P. canadensis*, sometimes even exceeding those of the other species, such as in October 2017 (RMSE = 5.86 g.kg⁻¹ and RPD = 3.04). Inaccuracies in pigment estimation also affected TPH predictions. For example, the RMSE observed for chlorophylls on *R. fruticosus* (2.62 $\mu\text{g}\cdot\text{cm}^{-2}$, Figure 43) was greater than the pigment variations for TPH concentrations between 40 and 60 g.kg⁻¹ (Figure 45). In that sense, making improvements to PROSPECT, including the separation of chlorophylls a and b, might help improve TPH predictions⁹².

Predictions were more accurate for TPH concentrations of less than 30 g.kg⁻¹, because small variations rapidly induced changes in chlorophyll content (Figure 45). Above 30 g.kg⁻¹, TPH were more difficult to predict as the chlorophyll content tended to stabilize. This type of difficulty might also arise outside of the range studied. Lethal concentrations of TPH would become impossible to predict for example. Likewise, predictions might be difficult with *R. fruticosus*, *Q. pubescens* and *P. canadensis* at very low concentrations (μg to mg.kg⁻¹), because little or no pigment alterations are expected for these species. At such low concentrations, the use of a species more sensitive to TPH might however make predictions possible³⁴⁰.

Table 15. Assessment of Total Petroleum Hydrocarbons (TPH) in the soil predicted from leaf chlorophyll content on the train and test datasets. RMSE, BIAS and SEPC are given in g.kg⁻¹ (abs: absent from the site). The best prediction results are marked in bold.

Species	Season	train		test				
		R ²	RMSE	R ²	RMSE	BIAS	SEPC	RPD
<i>B. davidii</i>	June 2017	0.06	16.78	0.04	16.24	3.04	15.95	1.00
	October 2017	0.05	19.27	0.07	18.14	-4.40	17.60	0.90
	March 2018	0.01	17.01	0.03	16.47	-10.35	12.81	0.99
	April 2018	0.01	21.13	0.02	20.01	-5.50	19.24	0.81
	June 2018	0.04	16.18	0.07	15.98	0.12	15.98	1.02
	October 2018	0.01	19.04	0.02	17.76	-1.31	17.71	0.92
<i>R. fruticosus</i>	June 2017	0.89	5.54	0.94	3.63	-0.21	3.62	4.10
	October 2017	0.61	10.44	0.74	7.64	1.56	7.48	1.95
	March 2018	0.02	15.47	0.19	17.05	-0.93	17.02	0.87
	April 2018	0.05	13.31	0.12	18.31	-4.68	17.70	0.81
	June 2018	0.75	8.56	0.78	6.64	-0.58	6.62	2.24
	October 2018	0.79	6.50	0.77	8.28	-0.35	8.27	1.80
<i>C. sanguinea</i>	June 2017	0.01	15.71	0.01	15.75	1.93	15.63	1.02
	October 2017	0.03	17.81	0.04	17.20	0.32	17.19	0.93
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	0.01	14.24	0.02	17.33	-1.34	17.28	0.92
	June 2018	0.01	15.88	0.01	15.98	-0.03	15.98	1.00
	October 2018	0.04	18.39	0.01	17.45	0.86	17.43	0.92
<i>Q. pubescens</i>	June 2017	0.78	8.41	0.85	5.70	2.24	5.24	2.89
	October 2017	0.66	10.09	0.86	5.82	0.65	5.79	2.83
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	0.01	15.80	0.03	17.94	0.84	17.92	0.92
	June 2018	0.72	8.93	0.96	3.20	-0.20	3.19	5.17
	October 2018	0.57	11.82	0.83	5.88	0.66	5.85	2.80
<i>P. canadensis</i>	June 2017	0.61	11.03	0.80	6.85	0.22	6.85	2.60
	October 2017	0.76	7.62	0.89	5.86	0.67	5.82	3.04
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	0.05	16.91	0.13	16.26	5.64	15.25	1.09
	June 2018	0.62	10.70	0.75	7.83	-1.16	7.74	2.27
	October 2018	0.74	8.93	0.81	6.62	-3.18	5.81	2.69

There was no observable relationship between leaf chlorophyll content and TPH concentrations in March and April for *R. fruticosus*, *Q. pubescens* and *P. canadensis* (Table 15). Consequently, the predictions of TPH on these dates were poor. The leaf chlorophyll content of *R. fruticosus* was the same as that of the control site in March (Figure 44), meaning that the species was thus not affected by TPH, which explains the inaccurate predictions. April was characterized by the appearance of new leaves in the three species, along with the most pronounced alteration in leaf chlorophyll content. The plants' sensitivity to TPH exposure is particularly pronounced in the growing stages^{190,271}. As a result, chlorophyll content in April 2017

displayed similar alterations regardless of the contamination level, which had an impact on TPH predictions. A similar response was observed for *C. sanguinea* leaves over the entire study (Table 15). In this case, chlorophyll content made it possible to detect but not to quantify TPH (RMSE > 15 g.kg⁻¹, RPD < 1.10). Consequently, the relationship between the leaf chlorophyll content and TPH concentrations observed for three out of the five species appeared after the growing period (April). This temporal pattern is of critical importance for the long-term monitoring of oil activities, as it shows that a single campaign carried out at the right time can be sufficient to assess the level of soil contamination.

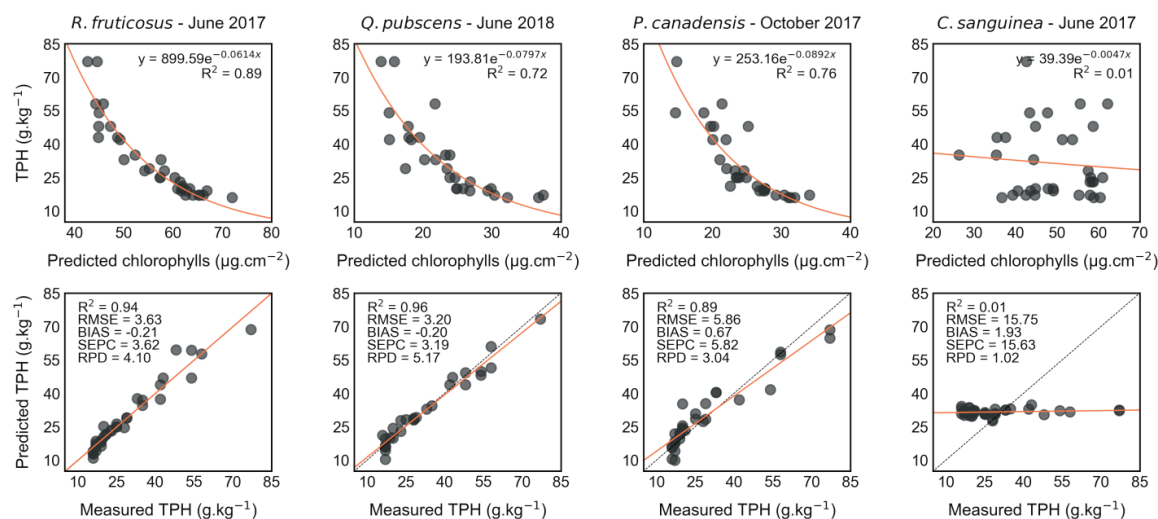


Figure 45. Relationship between predicted leaf chlorophyll content and Total Petroleum Hydrocarbons (TPH) in the soil on the train datasets (top figures) and comparisons between predicted and measured TPH on the test datasets (bottom figures).

Leaf chlorophyll and carotenoid contents are usually well-correlated^{82,228}. However, the exposure of plants to TPH altered this correlation. The relationship between carotenoid content and TPH was weaker than for chlorophyll content ($R^2 \leq 0.60$) and made predictions more difficult. The results obtained with *R. fruticosus*, *Q. pubescens* and *P. canadensis* were very similar. The best RMSE was observed with *R. fruticosus* (8.31 g.kg⁻¹), but represented poor predictions nonetheless (RPD < 2) (see Supporting Information). As discussed above, the contrasting effects of TPH on carotenoid pigments might have altered the relationship between leaf carotenoid content and TPH concentrations and therefore contributed to inaccuracies. Predictions might be greatly improved by separating carotenoid pigments and focusing on particular ones.

When compared to the proposed approach, vegetation indices and PLSR did not give any more accurate TPH predictions for *B. davidii* and *C. sanguinea*, confirming that these species are not suitable for predicting TPH. These results are presented in the Supporting Information. Predictions were better than those obtained with carotenoid content in the other species, but never reached those achieved with chlorophyll content. RMSE never dropped below 5.86 g.kg⁻¹ and RPD never exceeded 2.57. Consequently, our approach outperformed vegetation indices and PLSR. One possible explanation is that in the VIS, leaf optical properties are mainly influenced by chlorophylls, and to a lesser extent by carotenoids – and other pigments – that were less

correlated to TPH concentrations in our study^{78,240}. Therefore, TPH predictions using vegetation indices and PLSR suffered from the influence of carotenoids, which introduced inaccuracies. Conversely, the proposed approach using PROSPECT made it possible to isolate chlorophyll content, thereby improving TPH predictions.

8.4. Conclusion

This study demonstrates that TPH concentrations in oil-contaminated soils can be estimated from leaf optical properties using PROSPECT. The model performed well for simulating the spectral signatures and for retrieving pigment contents of leaves from exposed vegetation. Leaf chlorophyll content was strongly linked to TPH concentrations and allowed accurate predictions of these contaminants. The response of carotenoid pigments was complex and made predictions more difficult. Our approach was more effective than vegetation indices and PLSR, thanks to pigment separation. This study reveals that the accuracy of predictions strongly varies among species, because of their different sensitivities to TPH exposure. Three out of the five species studied (*R. fruticosus*, *Q. pubescens* and *P. canadensis*) were successful in predicting TPH. Moreover, the sensitivity of a single species also varies throughout the year, according to season changes. Changes in the species' sensitivity, during the growing period (April) for example, affect TPH predictions.

Considering these factors, the proposed approach can be used as a reliable tool for assessing soil contamination on oil production facilities. In the future, the approach could be applied to airborne images with high spatial and spectral resolutions for monitoring large industrial areas with dense vegetation. For this purpose, the combination of the PROSPECT and SAIL models (called PROSAIL⁷³), which takes into account canopy architecture, Leaf Area Index (LAI) and viewing geometry, will be addressed in further research.

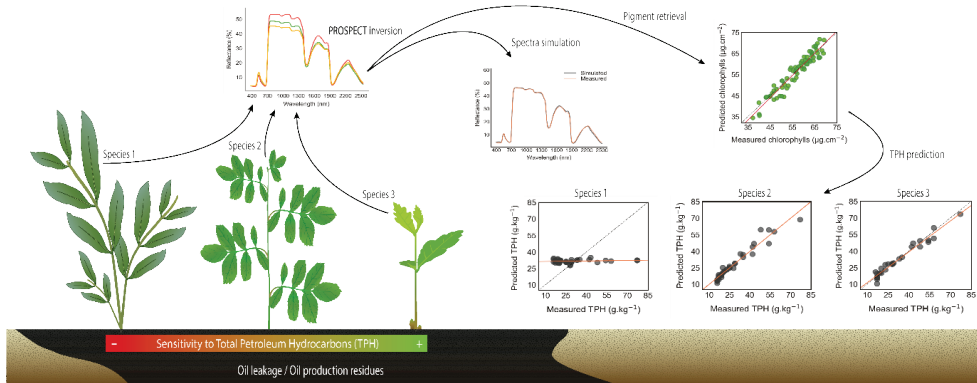
Acknowledgements

This work was performed in the frame of the NAOMI project between TOTAL and the ONERA, with the support of the Ecolab and Dynafor research units of Toulouse. The authors gratefully acknowledge E. Buffan-Dubau, D. Lambrigot and O. Berseille for their assistance in pigment analysis, and E. Hannan for correcting the manuscript.

Declarations of interest: None.

Funding: Financial support of this work was provided by TOTAL.

Graphical abstract



9. Fifth article: “Toward quantifying oil contamination in vegetated areas using very high spatial and spectral resolution imagery”

Article reference:

Lassalle, G.; Elger, A.; Credoz, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Fabre, S. Toward quantifying oil contamination in vegetated areas using high spatial and spectral resolution imagery. *Remote Sens.* **2019**, *11*, 2241. <https://doi.org/10.3390/rs11192241>

Abstract

Recent remote sensing studies have suggested exploiting vegetation optical properties for assessing oil contamination – especially Total Petroleum Hydrocarbons (TPH) – in vegetated areas. Methods based on the tracking of alterations in leaf biochemistry have been proposed for detecting and quantifying TPH under controlled and field conditions. In this study, we expand their use to airborne imagery, in order to monitor oil contamination at large areas. Airborne hyperspectral images with very high spatial and spectral resolutions were acquired over an industrial site with oil-contaminated (mud pits) and control sites both colonized by *Rubus fruticosus* L. The method of oil detection exploiting 14 vegetation indices succeeded in discriminating the sites in the case of high contamination (overall accuracy $\geq 91.8\%$). Two methods based on either the PROSAIL (PROSPECT + SAIL) radiative transfer model or Elastic net multiple regression were also developed for quantifying TPH. Both methods were tested on reflectance measurements in the field – at leaf and canopy scales – and on the image, and achieved accurate predictions of TPH concentrations (RMSE ≤ 3.28 g.kg⁻¹ and RPD ≥ 2). The methods were validated on additional sites and open up promising perspectives of operational application for oil and gas companies, with the emergence of new airborne and satellite sensors.

Keywords: Hyperspectral remote sensing, Vegetation, Soil contamination, Total petroleum hydrocarbons, radiative transfer model, Pigment, Elastic net regression

9.1. Introduction

Since the beginning of the 20th century, oil and gas supply has constantly increased to satisfy a growing demand worldwide^{4,10}. Along with the development of oil industry, important efforts have been made to mitigate the associated environmental risks^{36,43,65}. Oil spills and leakages are of major concern in the onshore domain. They are likely to occur at every step of the production (*i.e.* oil extraction, refining and transportation), contaminating the soil and groundwater and remaining as mud pits after their cessation^{28,32,35}. The resulting soil contamination causes important ecological alterations (*e.g.* landscape fragmentation and habitat loss)^{37,39,364,365}. To avoid such consequences, fast and accurate detection and quantification of oil contamination are therefore necessary.

Total Petroleum Hydrocarbons (TPH) are good indicators of oil contamination in soils^{154,263}. They are the main constituent of oil and show high toxicity toward organisms. (In this study, the terms *oil* and *TPH* were used indifferently.) Among the methods proposed for detecting and quantifying TPH contamination, hyperspectral remote sensing has become a promising one^{63,64}. Recent advances in this field allow monitoring industrial facilities (*e.g.* pipelines, storage tanks) at large scale using airborne or satellite imagery^{46,67}. However, this approach is based on the detection of apparent oil at the soil surface. This causes a critical issue in regions with dense vegetation and intensive oil production activity. In this context, alternative approaches exploiting vegetation optical properties have emerged during the last decade^{35,68,186,205}.

TPH induce strong alterations in leaf anatomy and biochemistry related to its optical properties in the reflective domain (400 – 2500 nm)^{78,93,162,183,196}. These alterations lead to modifications in the spectral signature of vegetation and suggest being able to detect oil in vegetated areas using optical remote sensing^{77,159,324}. Based on this assumption, few studies attempted to map oil spills and leakages from optical airborne or satellite imagery with four to 30 m spatial resolution (mainly Landsat, Hyperion and Hymap systems)^{35,68,231,232}. Most of them relied on comparisons of vegetation indices between healthy and contaminated sites, but only a few have proposed a method to detect oil contamination automatically, and even fewer have evaluated its performance. In addition, none of these studies attempted to quantify TPH. In most cases, the method was first calibrated on a known contaminated area and then applied to the entire image^{68,232}. This gave rise to the apparition of false alarms, caused by differences in species' sensitivity to oil, bare soil-mixed vegetation pixels, and the presence of other vegetation stressors (*e.g.*, water deficit). False alarms have been more pronounced using multispectral satellite imagery. Thus, several conclusions emerged from these studies. First, very high spatial (1–2 m) and spectral (<10 nm) resolutions are needed to achieve accurate detection of oil. Then, prior application to entire images, *i.e.*, it is necessary to calibrate the methods over restricted areas with known species and contamination. Then, the methods should be validated on spatially independent sites with the same species and, as a last step, applied to entire images, provided that the target species' location is known.

Species established around industrial oil facilities are naturally tolerant to soil contamination and difficult to distinguish from healthy vegetation^{270,271}. It is, therefore, challenging to develop methods for detecting and quantifying TPH adapted to these species and intended to be applied to airborne and satellite hyperspectral images. Detecting and quantifying TPH is even more difficult in the presence of other environmental factors (*e.g.*, water deficit). As suggested in previous work, this could be achieved by experiments carried out under controlled conditions^{204,324}. Several studies have aimed to characterize the effects of oil on vegetation reflectance under controlled conditions^{29,159,200,268}. Vegetation indices and spectrum transformations (first and second derivatives, *continuum* removal) have been frequently used for this purpose^{77,185,186,251,268}.

The wavelengths linked to pigment and water contents were particularly suitable for distinguishing healthy and oil-exposed vegetation with no or minor confusion^{205,271}. More recently, Lassalle *et al.*³²⁴ succeeded in discriminating among various types of oil contamination from leaf to canopy scale using vegetation indices. The same approach performed well for detecting mud pits contaminated by TPH under natural conditions, by exploiting the reflectance of an oil-tolerant species (*Rubus fruticosus* L.). This study opened up promising perspectives of applications to hyperspectral images. Radiative transfer models also sparked great interest for tracking alterations in leaf pigment contents caused by oil. For example, Arellano *et al.*²³⁰ inverted the PROSPECT model⁷⁰ to detect changes in Leaf Chlorophyll Content (LCC) resulting from an oil spill under tropical region. The same model was recently used for quantifying TPH in mud pit soils from leaf reflectance measurements³²⁵. As suggested, its coupling with the SAIL canopy model²⁴² (forming the PROSAIL model⁷³) might be of great interest for quantifying oil contamination in vegetated areas, from hyperspectral images. This challenge remains however very difficult, as it implies being able to track subtle changes in LCC from airborne or satellite sensors.

In the continuity of the abovementioned studies^{324,325}, this work aims to detect and quantify oil contamination (i.e., soil TPH content) in vegetated areas using airborne hyperspectral images with very high spatial and spectral resolutions. To achieve this, a method of detection previously developed under controlled conditions, hereafter summarized, was first applied to the images. Then, two approaches, either PROSAIL or multiple regression, are proposed to quantify TPH. Both were assessed in the field and on airborne hyperspectral images, on selected vegetation patches with known species.

9.2. Materials and Methods

9.2.1. Study area and soil sampling

The study was carried out under temperate region, on an industrial area subject to important oil and gas activities. During the past 50 years, oil production residues were accumulated in mud pits, which resulted in contamination of soil by TPH. Some of these mud pits remain and have been colonized by oil-tolerant vegetation, especially *Rubus fruticosus* L. (bramble) (Figure 46c). The study focused on this species, as it is generally well-established in temperate industrial areas. Bramble served for developing the methods of oil detection and quantification under controlled and field conditions^{271,324,325} and was suitable for application to hyperspectral imagery.

Five mud pits were identified for the study (Figure 46a), and the largest one (40 x 100 m, hereafter described as "brownfield") was used for calibrating the methods. Within this site, 10 homogeneous and spatially-correlated plots of 16 to 20 m², mainly covered by *R. fruticosus* (>95%), were defined in the field (Figure 46b). Their geographical coordinates were acquired and soil analyses were performed at two locations in each plot. Analyses revealed C₁₀-C₄₀ TPH concentrations ranging from 17 to 39 g.kg⁻¹ on the site. The same sampling procedure was carried out on the four other mud pits, which were not spatially correlated with the largest one. A single plot was defined on each of them, since they occupied a smaller area. One of these mud pits exhibited comparable TPH concentrations (24 g.kg⁻¹) and was used for validating the quantification method (Table 16). Low concentrations were found in the other ones (≤ 3.15 g.kg⁻¹). In addition, one large (20 x 20 m) and five smaller control sites colonized by *R. fruticosus* were identified in the same region and used for the calibration and validation of the detection method, respectively (Figure 46d-f). No TPH were detected in any of the control soils. These sites however slightly differed in soil properties – especially in texture – and surrounding

ecosystem (poplar plantation, grassland and riparian forest). They were thus suitable for assessing the robustness of the method of detection.

9.2.2. Hyperspectral data acquisition

9.2.2.1. Airborne image

Hyperspectral images were acquired over the study area on July 5, 2017 at 1.15 pm., at 6 900 ft. above sea level in the absence of clouds, using airborne VNIR-1600 and SWIR-320m-e HySpex cameras (Norsk Elektro Optikk AS, Lørenskog, Norway) embedded on ONERA's SETHI platform³⁶⁶ (Figure 46a, d). These images had a spatial resolution of 1 and 2.5 m and a spectral resolution of 5.2 and 7.8 nm in the visible-near-infrared (VNIR, 400 – 1000 nm) and short-wave infrared (SWIR, 1000 – 2500 nm) domains, respectively. The SWIR image was resampled to 1-m spatial resolution using a nearest neighborhood filter, in order to preserve the spatial information, and registered according to the VNIR image using the Gefolki algorithm³⁶⁷. The resulting spectral radiance image was then converted to spectral reflectance using the Empirical Line Method (ELM)^{368,369}, because of the lack of knowledge about the composition of the local industrial atmosphere. This provided a 1-m resampled reflectance image with 409 spectral bands covering the reflective domain (400 – 2500 nm). Because of the low signal-to-noise ratio (SNR), we did not conserve the bands with atmospheric transmission below 80%, as described in²⁹². A Savitzky-Golay smoothing filter²⁸³ was applied at the remaining bands, and the resulting image was used for calibrating and validating the methods of detection and quantification described hereafter.

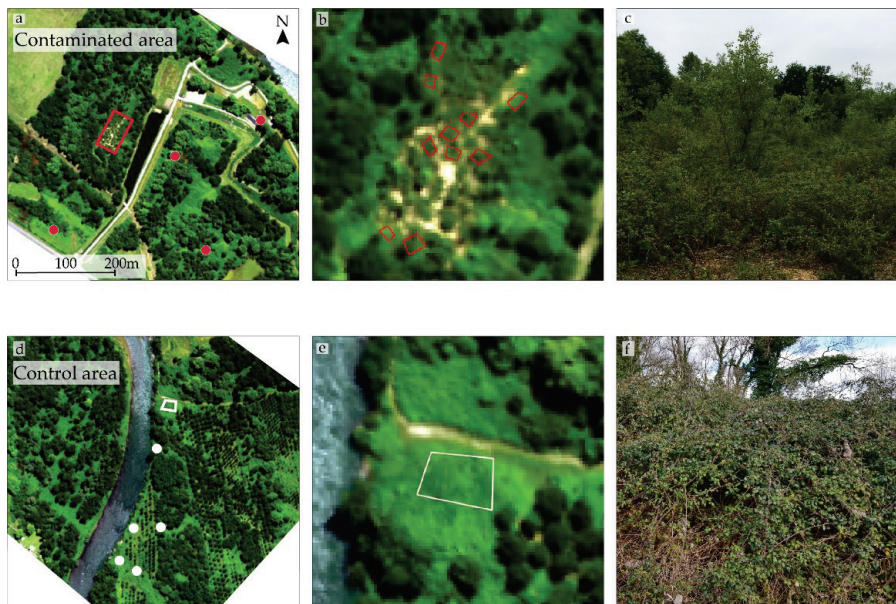


Figure 46. (a, d) Subsets of the airborne hyperspectral image acquired over (a) the contaminated and (d) the control sites. The calibration sites are illustrated by polygons. Circles correspond to the validation sites (red: Contaminated, white: Control). (b, e) Zoom on (b) the brownfield (10 plots) and (e) the control sites used for calibrating the methods, respectively. (c, f) *R. fruticosus* on (c) the brownfield and (f) the control site. (Note: Both photos c and f were taken on March.)

9.2.2.2. Field reflectance

On the day of image acquisitions, a field campaign was carried out on the 10 plots of the large mud pit (“brownfield”) (Figure 46b). On each plot, 18 and 9 spectral signatures of brambles were acquired at leaf and canopy scales respectively, using an ASD FieldSpec 4 Hi-Res spectroradiometer (Malvern Panalytical, Malvern, UK). Data were acquired in spectral radiance between 11.30 am and 2.00 pm and converted to spectral reflectance as described in ³²⁴. Leaf measurements were performed with a leaf-clip. Canopy reflectance was acquired using a 25-mm wide fore optics placed 45 cm above the target at nadir, providing a 20-cm wide acquisition footprint. All spectral signatures were then resampled to the spectral resolution of the airborne image and the same band removal and smoothing procedures (atmospheric transmittance < 80% and Savitzky-Golay smoothing filter) were applied. These data were used for calibrating the method of TPH quantification.

9.2.3. First step: Detection of oil contamination

The approach proposed in this study is divided in two successive steps: the detection of oil contamination and the quantification of TPH. The first one relies on the application of a method developed in the same context of study. The method exploits 14 vegetation indices – computed from the spectral signatures – and L²-Regularized Logistic Regression (RLR²⁸⁹) classifier to determine whether vegetation is – or has been exposed to oil (see ³²⁴ for a detailed description of the method). These indices have been linked to the effects specifically induced by oil on leaf pigment and water contents of bramble. They are listed in Table 31 in the Supporting Information section. The method was initially developed and tested under controlled conditions and succeeded in detecting oil contamination, from leaf to canopy scale³²⁴. It has been validated under natural conditions on the brownfield, so in this study, we focused on its application to the airborne hyperspectral image.

The method was first tested on the 10 plots of the brownfield and on the large control site (Figure 46b, e). The corresponding spectral signatures were extracted from the image and the 14 vegetation indices were computed. Data were split in 50% training and 50% test sets, with equal proportion of pixels from the brownfield and the control site in each. The RLR classifier was fitted on the training set and applied to the test set. Predictions made on the test set were evaluated using the Overall Accuracy (OA), Cohen’s Kappa coefficient and confusion matrices^{212,218,292}. The method was then validated on the other contaminated and control sites (n = 4 and 5 sites, respectively, Table 16).

Table 16. Description of the sites used for testing and validating the methods. (< DL: Below detection limit.)

Step	Zone type	Number of zones	Total pixel count	C ₁₀ -C ₄₀ TPH (g.kg ⁻¹)
train + test	control	1	388	< DL
train + test	mud pit (“brownfield”)	10	184	17 - 39
validation	control	5	211	< DL
			33	0.25
validation	mud pit	4	20	0.38
			48	3.15
			22	24

9.2.4. Second step: Quantification of soil TPH content

9.2.4.1. First method based on PROSPECT and PROSAIL

The first method proposed in our study for quantifying soil TPH content relies on the retrieval of LCC by inverting the PROSPECT and PROSAIL radiative transfer models^{70,73,242}. The PROSPECT model simulates the hemispherical reflectance and transmittance spectra of leaves in the reflective domain, knowing their biophysical and biochemical parameters (listed in Table 17). The model can be inverted to retrieve these parameters from reflectance measurements, using an iterative optimization approach^{82,108}. The latter consists in successively testing different sets of parameters (denoted by the vector θ) and retaining the one that best simulates the measured spectral signature^{347,352}. This is achieved by minimizing the cost function J , defined as:

$$J(\theta) = \sum_{\lambda_{\min}}^{\lambda_{\max}} (R_{\text{meas}}(\lambda) - R_{\text{sim}}(\lambda, \theta))^2, \quad (2)$$

where R_{sim} and R_{meas} denote the simulated and measured reflectance in the $[\lambda_{\min}:\lambda_{\max}]$ spectral range, respectively. In previous work, the LCC of *R. fruticosus* estimated using PROSPECT inversion was consistent with laboratory analyses and strongly linked to TPH concentrations³²⁵. This method was developed in the field and performed well for predicting the contamination level (*i.e.* TPH concentration) at leaf scale. In this study, we propose to assess its robustness on field measurements performed at canopy scale and on the airborne images, using the PROSAIL model. PROSAIL results from the coupling of PROSPECT and SAIL models⁷³. It expands the scope of PROSPECT to dense canopy reflectance data by taking the influence of canopy architecture, Leaf Area Index (LAI) and bare soil, and illumination and viewing geometry into account^{349,370}. The inversion of the model follows the same procedure as that of PROSPECT. In this study, PROSAIL was used for retrieving LCC.

The objective of this method is to quantify TPH from LCC, so we applied it to the contaminated sites. PROSPECT was inverted on leaf reflectance measurements performed on the 10 plots of the brownfield ($n = 180$). On the same plots, PROSAIL was inverted on both field canopy measurements ($n = 90$) and on the pixels from the airborne image ($n = 184$). Inversions were computed using the differential evolution algorithm^{325,353} and the model parameters listed in Table 17. The LCC values retrieved from canopy measurements and from the image were compared to those of leaves using the R^2 and the Root Mean Squared Error (RMSE)³²⁵.

Table 17. Parameter bounds used to compute PROSPECT and PROSAIL (PROSPECT + SAIL) inversions.

Parameter	Unit	Range	Reference
<u>PROSPECT:</u>			
Leaf structure (N)		1 - 5	
Chlorophyll a + b (C_{ab})	$\mu\text{g.cm}^{-2}$	1 - 100	
Carotenoids (C_{cx})	$\mu\text{g.cm}^{-2}$	1 - 50	325
Brown pigments (C_{bp})	$\mu\text{g.cm}^{-2}$	0.01 - 1	
Water (C_w)	g.cm^{-2}	0.001 - 0.1	
Dry matter (C_m)	g.cm^{-2}	0.001 - 0.1	
<u>SAIL:</u>			
Leaf Area Index (LAI)	$\text{m}^2.\text{m}^{-2}$	0.1 - 5	
Hotspot (hot)	m.m^{-1}	0.01 - 0.1	
Leaf Inclination Distribution Function (LIDF)		Planophile ^e	
Soil brightness ^f ($bright$)		0.5 - 1.5	88,349,370-373
Solar zenith angle (ϑ_s)	deg.	Fixed (30°)	
Viewing zenith angle (ϑ_v)	deg.	Fixed (0°)	
Relative azimuth angle (φ_{S_v})	deg.	Fixed (0°)	

LCC was then used for quantifying TPH. Each data type (leaf, canopy and image) was split into 50% training and 50% test sets. Negative exponential models were fitted between LCC and TPH concentrations on the train set and then applied to the test set³²⁵. TPH predictions were compared to the true concentrations by computing the R^2 , the RMSE and the Residual Predictive Deviation (RPD)^{85,203}. Since the method was calibrated for TPH concentrations ranging from 17 to 39 g.kg^{-1} , it was validated on the mud pit with 24 g.kg^{-1} TPH only.

9.2.4.2. Second method based on Elastic net regression

The second method of quantification consists in linking the spectral signatures to TPH concentrations directly using multiple regression. Several approaches have been proposed in previous studies aiming to quantify heavy metals from vegetation reflectance in the field, especially Partial Least Square Regression (PLSR)^{203,341}. PLSR is not suitable for TPH quantification³²⁵. So in this study, we proposed another regression approach, the Elastic Net (ENET²²²), which shows multiple advantages but remains underexploited in remote sensing of vegetation^{374,375}. ENET is a penalized least squared approach that allows efficient variable selection under multicollinearity, which is a major limitation of other multiple regression approaches when dealing with hyperspectral data. Moreover, ENET performed well for tracking alterations in leaf biochemistry induced by oil in previous study³²⁴.

In the present study, ENET was used for predicting TPH concentration from the spectral signature of vegetation while selecting the most contributing wavelengths (*i.e.* spectral bands). For this purpose, several spectrum transformations were tested: First and second derivatives, Standard Normal Variate (SNV), $\text{Log}(1/R)$ and Area Under Curve Normalization (AUCN) (see^{203,271,376} for a detailed description of the transformations). All these transformations are likely to enhance the information contained in the spectral signature and to improve TPH predictions. They are illustrated in Figure 72 in the Supporting Information section. The method was applied to the spectral signatures acquired in the field (leaf and canopy scales) and on the airborne image, following the same procedure as the PROSAIL method (*i.e.* training and test steps on the brownfield and validation on the other mud pits). The accuracy of TPH predictions were

^eAs specified for *Rubus fruticosus* L. in ³⁷³.

^fThe recent version of PROSAIL includes a simple lambertian soil reflectance model. A brightness parameter ($bright$) was added to account for soil moisture and roughness^{349,372}.

compared between the two methods, as well as their respective advantages. A comprehensive flowchart of this study is presented in Figure 73 in the Supporting Information section^g.

9.3. Results and Discussion

9.3.1. Detection of oil contamination

The RLR-based method exploiting vegetation indices performed well for discriminating between healthy and oil-exposed vegetation from the airborne image. The classification trained on the 10 zones of the brownfield and the large control site reached OA and Kappa of 97.9% and 0.95 on the test set, respectively. Only little confusions were made between healthy and oil-exposed vegetation (Table 18). Within the brownfield, the misclassified pixels came from the least contaminated plots (17 – 18 g.kg⁻¹). Almost all pixels from the large control site were correctly classified. These results were consistent with those of previous studies carried out under controlled and field conditions on the same species, at leaf and canopy scales^{271,324}.

Table 18. Confusion matrix obtained from the RLR classification on the test set using the 14 vegetation indices. The *Brownfield* class denotes the pixels from the 10 zones of the brownfield, and the *Control* one, those from the large control site.

	Brownfield	Control	User's accuracy (%)
Brownfield	88	4	95,7
Control	2	192	99,0
Producer's accuracy (%)	97,8	98,0	

The robustness of the method to upscaling (*i.e.* from leaf to canopy scale and imagery) is due to the limited influence of plant architecture, bare soil, and illuminating and viewing geometry on the selected vegetation indices at very high spatial resolution. These factors usually affect the ability of vegetation indices to detect changes in leaf biochemistry from canopy measurements or optical images^{240,377,378}. This phenomenon was likely to induce confusion between healthy and oil-exposed vegetation. However, the 14 indices used in the method have been specifically selected for their robustness to upscaling and their link with oil-induced effects on leaf biochemistry^{299,300,306}. These indices mainly depend on chlorophylls a and b, β -carotene and lutein contents for bramble³²⁴ (see Table 31 in the Supporting Information section). They are little affected by the abovementioned factors (especially bare soil), so their value remains almost unchanged from leaf to canopy scale and on the airborne image. Therefore, they allow tracking alterations in leaf pigment contents specifically induced by oil with good accuracy, regardless of the acquisition scale³²⁴.

9.3.2. Quantification of soil TPH content

9.3.2.1. PROSAIL-based method

LCC estimated from the inversions of PROSPECT and PROSAIL models on the 10 plots of the brownfield is presented in Figure 47. LCC estimations were strongly consistent among data types (*i.e.* leaf, canopy and airborne image), especially in the field ($R^2 = 0.95$ and $RMSE = 1.05 \mu\text{g.cm}^{-2}$, Figure 47a). They ranged from 53 to 69 $\mu\text{g.cm}^{-2}$, which is particularly high for vegetation

^gsee also the Graphical Abstract at the end of this manuscript

exposed to contamination. This highlights the tolerance of bramble to oil exposure and explains its establishment on industrial sites under temperate regions^{29,274,308,325}. PROSAIL is a one-layer model adapted to dense vegetation. Its application to the airborne image succeeded in retrieving LCC, thanks to very high spatial resolution. As expected, estimations were slightly less accurate and showed a higher variability within the plots than for field measurements ($R^2 = 0.85$, $RMSE = 3.70 \mu\text{g.cm}^{-2}$, Figure 47b).

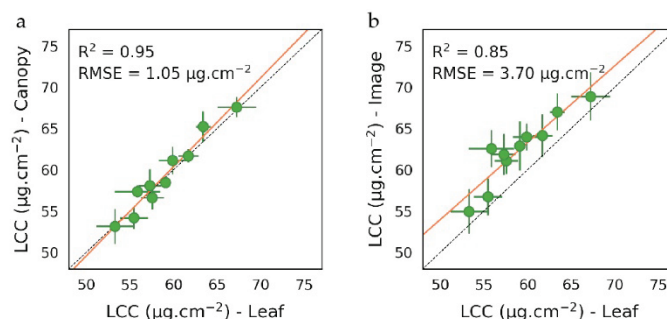


Figure 47. (a-b) Leaf Chlorophyll Content (LCC) (mean \pm SD) retrieved from PROSPECT and PROSAIL model inversions on the 10 plots of the brownfield. (a) Comparison between leaf and canopy measurements from the field and (b) between leaf measurements and the airborne image.

LCC exhibited a close negative relationship with soil TPH content on the train set ($0.74 < R^2 < 0.87$, Figure 48a-c). The more the soil was contaminated, the more LCC decreased. This relationship has been previously noticed under controlled conditions for oil-sensitive species, and has been linked to the reduction of water and nutrient availability in soil and the alteration of root uptake capacities^{162,169,193}. Similar observations have been made under natural conditions on bramble, at leaf scale only³²⁵. However, our method is the first to detect this relationship from leaf to canopy scale and on airborne images. Its application to the test set resulted in accurate predictions of TPH concentrations (Figure 48d-f). Results were very close at leaf and canopy scales ($RMSE = 2.62$ and 2.85 g.kg^{-1} and $RPD = 2.26$ and 2.15 , respectively), thanks to consistent LCC described above (Figure 48a). A higher variability in LCC was observed on the image and led to little inaccuracies in TPH predictions ($RMSE = 3.2 \text{ g.kg}^{-1}$ and $RPD = 2$), in comparison to field measurements. The highest level of contamination (39 g.kg^{-1}) was more difficult to predict and contributed a lot to inaccuracies. As previously observed at leaf scale, TPH are particularly difficult to quantify above a certain concentration³²⁵, probably because variations in LCC become too subtle to be caught from imagery. In contrast, LCC rapidly shifts for moderate levels of contamination ($20 - 30 \text{ g.kg}^{-1}$) and makes predictions more accurate.

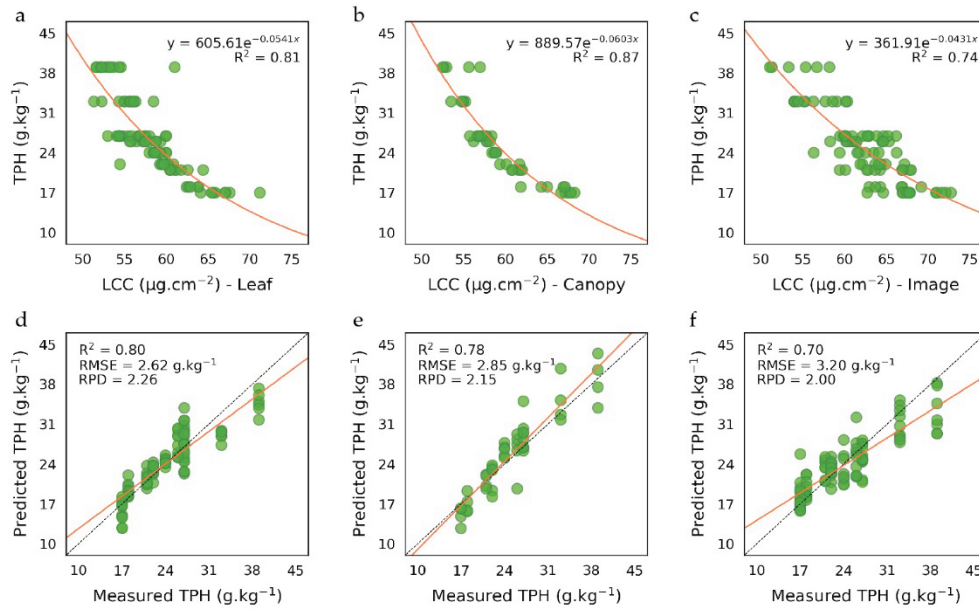


Figure 48. (a-c) Relationship between LCC and soil TPH content observed on the train set on the 10 plots of the brownfield, at (a) leaf and (b) canopy scales and (c) on the airborne image. (d-f) Comparison between measured and predicted TPH concentrations on the test set, at (d) leaf and (e) canopy scales and (f) on the airborne image.

9.3.2.2. Elastic net-based method

The ENET regression based on the original and transformed spectral signatures also succeeded in predicting TPH concentrations on the brownfield (Figure 49). However, the overall performance of the method was lower than that based on PROSPECT and PROSAIL inversions, regardless of the transformation. Although TPH predictions were good for the original – reflectance – spectra for all scales ($\text{RMSE} \leq 4.04 \text{ g.kg}^{-1}$), they were improved using first and second derivatives ($\text{RMSE} \leq 3.53 \text{ g.kg}^{-1}$). These improvements were mainly due to a better quantification of the highest level of contamination (39 g.kg^{-1}) at canopy scale and on the airborne image. Conversely, predictions were almost unchanged with the $\text{Log}(1/R)$ transformation and less accurate using SNV and AUCN ones. As for the PROSPECT and PROSAIL- based method, the accuracy of predictions was reduced while increasing acquisition scale (best RPD = 2.16, 2.07 and 1.9 at leaf and canopy scales and on the image, respectively, Figure 49).

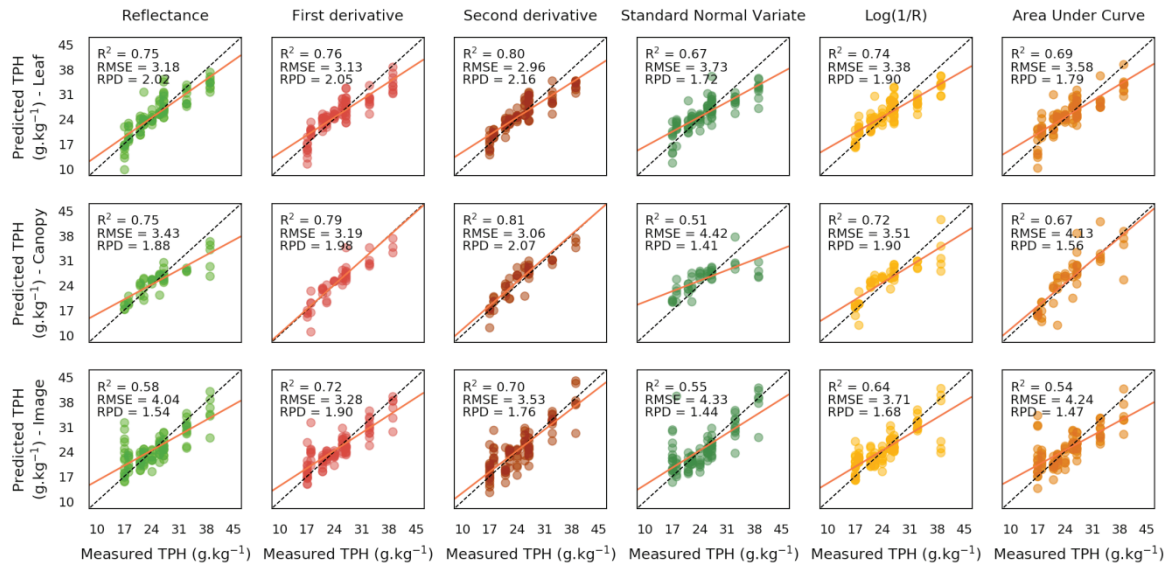


Figure 49. Comparison between measured and predicted TPH concentrations on the test set on the 10 plots of the brownfield using the Elastic net regression on the original and transformed spectral signatures (in columns). First, second and third rows show the results obtained at leaf and canopy scales and on the airborne image, respectively.

Spectrum transformations have been extensively used for detecting vegetation stress in various contexts, including oil contamination^{77,86,258,271}. The first derivative sparked particular interest, as it allows tracking shifts in the Red-Edge Position (REP) caused by chlorophyll alteration^{159,166,186}. Therefore, the REP has been many times depicted as a reliable indicator of alterations in vegetation health following exposure to TPH, however, many other environmental stressors also affect the REP, which might lead to inaccuracies in TPH quantification. In our case, the ENET method highlighted the importance of the red-edge – and more largely the VIS region – for quantifying TPH, since the most contributing bands were concentrated around 550 and 700 nm. This trend remained consistent among data types (leaf, canopy and image) and between reflectance and derivative spectra (Figure 49), and confirmed the involvement of chlorophylls in the response of vegetation to oil. The red-edge was almost sufficient alone for predicting TPH using first derivative spectra, especially at leaf and canopy scales. Additional spectral bands from the SWIR region were often selected on the airborne image. These bands have been linked to leaf water, cellulose and lignin contents in previous studies^{135,148}, but remained of less contribution to TPH predictions in our case. In contrast, SNV, Log(1/R) and AUCN transformation degraded TPH predictions, and did not provide a clear interpretation of the contributing spectral bands (Figure 50). These transformations are usually employed for reducing the high variability of spectral signatures while enhancing the information they contain^{203,271}. In our case, reflectance varied little along the gradient of TPH ($\leq 10\%$ at 550 nm and in the red-edge), because of the strong tolerance of bramble to oil exposure, in comparison to other species^{77,186}. Consequently, SNV, Log(1/R) and AUCN transformations were unable to conserve this information in the transformed spectra, which introduced inaccuracies in TPH predictions.

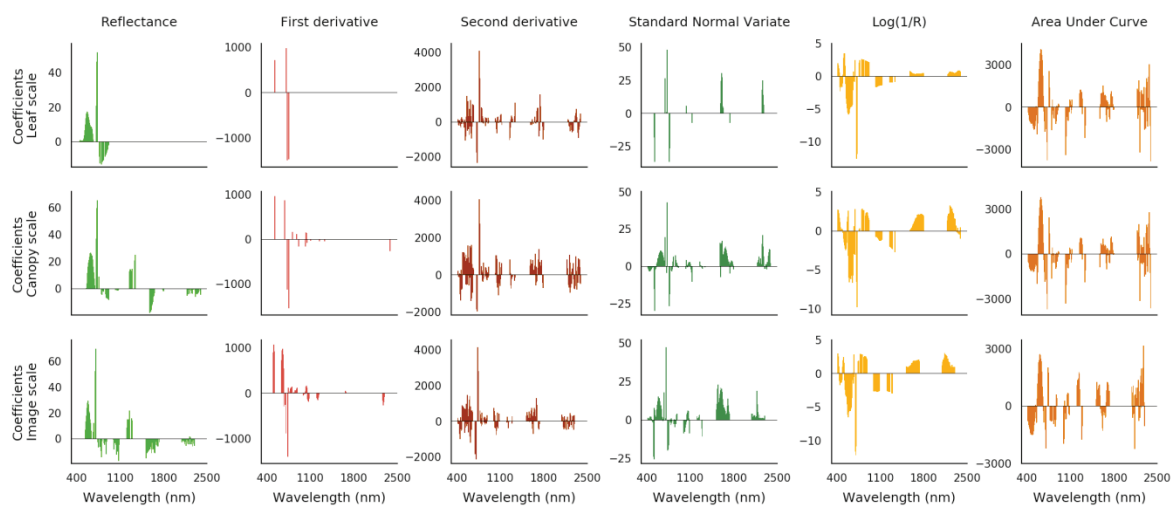


Figure 50. Coefficients obtained by the ENET method on the 10 plots of the brownfield using the original and transformed spectral signatures (in columns). First, second and third rows show the results obtained at leaf and canopy scales and on the airborne image, respectively. The spectral bands contributing to TPH predictions are denoted by non-zero coefficients.

9.3.3. Validation of the methods

9.3.3.1. Validation of oil detection

When applied to the validation sites, the method of oil detection based on the vegetation indices succeeded in detecting healthy vegetation. The ecological differences among the sites (soil properties and surrounding ecosystem) were challenging, because they introduced additional variability in reflectance – and thus in index values. However, this had only little influence on the performance of the method, since more than 91% of the pixels from the five control sites used for validation were correctly labelled (OA = 91.8%, Table 19). Conversely, the level of oil contamination was critical for the detection of mud pits. Accuracy averaged 50.5% on the four validation mud pits, and resulted from strong differences among sites. As expected, the most contaminated one (24 g.kg⁻¹) was well-classified (95.5%), as it was in the range of the method calibration (17 – 39 g.kg⁻¹). Conversely, accuracy rapidly fell for the other mud pits, for which low TPH concentrations were observed. More precisely, the two least contaminated sites (0.25 and 0.38 g.kg⁻¹ TPH) were poorly classified (accuracy of 6.1 and 15%, respectively, Table 19), because vegetation indices from these sites exhibited values similar to that of the control ones. These results emphasize the limits of the method. Below a certain concentration, oil no longer alters leaf pigment and water contents, so soil contamination becomes impossible to detect using vegetation reflectance^{77,205,324}. This detection limit strongly depends on the sensitivity of the species, and is higher for oil-tolerant ones^{186,201,268,270}. Thus, the closer the level of contamination to this limit, the less accurate the classification. In our case, this level was probably below the detection limit for bramble in the two least contaminated sites. Conversely, the method performed well for detecting the fourth validation mud pit contaminated by 3.15 g.kg⁻¹ TPH. More than 85% of the pixels from this site were correctly labelled, despite its contamination level being outside the calibration range. Hence, oil still induces important effects on leaf biochemistry of bramble at moderate contamination levels, which can be detected using the method proposed.

Table 19. Confusion matrix obtained from the RLR classification on the validation sites using the 14 vegetation indices.

	C ₁₀ -C ₄₀ TPH (g.kg ⁻¹)	Mud pit	Control	User's accuracy (%)
Control ^h	< DL	13	146	91.8
	0.25	2	31	6.1
Mud pit	0.38	3	17	15
	3.15	41	7	85.4
	24	21	1	95.5

9.3.3.2. Validation of TPH quantification

The two methods of quantification were applied to the mud pit contaminated by 24 g.kg⁻¹ TPH and confirmed their potential for application to airborne hyperspectral images. TPH concentrations were estimated to 23.69 (± 1.93) and 23.39 (± 2.41) g.kg⁻¹ using PROSAIL inversions and ENET with first derivative spectra, respectively. Predictions were slightly less accurate using other spectrum transformations (SNV, Log(1/R) and AUCN), but remained very close to the true level of contamination (data not shown). Both methods thus provide reliable quantification of TPH in our context of study. It is however important to note that the concentration observed on this mud pit corresponded to that for which best predictions were obtained on the brownfield (Figure 48, Figure 49). This explains the high accuracy of predictions observed here. Unfortunately, none of the four validation mud pits exhibited TPH concentrations above 35 g.kg⁻¹, which are probably the most difficult to predict. It was thus not possible to test the methods for such concentrations, but an underestimation would be probably observed, as on the brownfield. More generally, some inaccuracies in both methods of quantification might result from local variations in soil TPH content. These local variations introduce variability in bramble reflectance and might have affected TPH predictions, as suspected at leaf scale in previous study³²⁵.

9.4. Perspectives

To date, there was no remote sensing method for detecting and quantifying oil contamination in vegetated areas^{68,231,232}. Our study is the first to achieve it, from leaf to canopy scale and on airborne hyperspectral images in a temperate region. This opens up encouraging perspectives for application of the methods over oil industrial areas. However, prior to operational use in a broader context than that of our study, several points must be considered. They are discussed in this section.

The proposed methods proved to be efficient when applied to bramble. In accordance to our previous studies^{271,324,325}, this species can be considered as a reliable indicator of oil contamination. Moreover, bramble is highly tolerant to other environmental stressors (*e.g.* water-deficit), which avoid confusions with oil contamination under natural conditions using the proposed methods^{271,324}. This species thus shows great interest for environmental monitoring, especially since it is common on industrial facilities in temperate regions^{29,274}. Nonetheless, other species may be established on those sites and do not share the same sensitivity to oil. Some of them may be completely oil-tolerant (*i.e.* they do not exhibit reflectance modifications), whereas other ones may be affected to a greater extent than bramble^{77,159,201}. For example, the relationship between TPH and LCC has also been observed at leaf scale on oak and poplar species in a similar

^h The five control sites were grouped for more concision.

range of TPH concentrations, but in a lower range of LCC than that of bramble³²⁵. Since our methods were developed for "pure" homogeneous bramble covers, other species may alter their accuracy on images when mixed with bramble in pixels. This may rise false positives and negative and lead to an under- or overestimation of TPH. Species discrimination was out of the scope of this study, so it is discussed here as a perspective. Very high spatial resolution could be a good way of better discriminating species and adapting our methods to each of them^{379,380}. Unmixing spectral methods might also help achieving it^{381,382}. Likewise, Dehaan *et al.*³⁸³ proposed a Mixture-Tuned Matched Filter algorithm for mapping bramble that could serve as a preliminary step to the application of our methods. We are aware that the use of our methods cannot be generalized without being adapted to a wide range of ecological contexts, which involve different species and contamination types. For this purpose, our upcoming researches will focus on the tropical context, where the monitoring of oil and gas activities remains a major environmental issue and a great challenge for our methods, because of high species richness^{68,235,270,365}.

Regardless of the context, oil production facilities require permanent monitoring, since contamination can occur at any time. This implies being able to detect and quantify oil throughout the year, regardless of the vegetation stage (*e.g.* growing, flowering and senescence). However, it is well-known that seasons strongly affect leaf biochemistry – especially pigments – and optical properties^{97,342,384}. These changes can influence the performance of the methods through masking the effects of contamination on vegetation³⁴¹. Using the PROSPECT-based method, our previous study demonstrated that oil can be accurately quantified at leaf scale in summer and autumn for several species (including bramble) in temperate region³²⁵. Conversely, this remains more difficult in winter and spring because of the low development of vegetation and the emergence of leaves, respectively. These variations would certainly affect our methods to a greater extent on airborne images. As suggested, other species could be identified for achieving accurate oil detection and quantification during "unfavorable" seasons (*i.e.* winter and spring). This would make possible to monitor production facilities throughout the year in temperate regions, thanks to multitemporal acquisitions.

Our study provided a better understanding of the spatial and spectral resolutions needed to monitor oil contamination, which has strong implications for future applications. Before this study, very high spatial resolution (1 – 2 m) had never been considered for similar purposes. As discussed above, such resolution helps focusing on homogeneous patches of vegetation, discriminating species and limiting inaccuracies due to local variations in TPH concentrations, which cannot be accurately achieved at high to medium spatial resolutions (4 – 30 m)^{35,68,231,232}. Moreover, one of the aims of our methods is to detect and quantify oil at early stages of the contamination, before it spreads in the environment. Very high spatial resolution is mandatory, because oil is much localized in the soil at these stages (*e.g.* just after oil leaks and for recent mud pits). Conversely, the importance of high spectral resolution differs according to the operational need (*i.e.* detection or quantification). Our method of detection is based on vegetation indices, whose computation does not require a continuous spectrum. Therefore, it could be applied to multispectral sensors, provided that they include the appropriate spectral bands and very high spatial resolution. Likewise, PROSAIL is frequently applied to multispectral images for retrieving LCC^{73,88}, so our method of quantification could perform well at such spectral resolutions. However, this implies being able to track subtle changes in LCC (few $\mu\text{g}\cdot\text{cm}^{-2}$), which remains difficult using PROSAIL without high spectral resolution^{349,371}. In that case, the second method (ENET) could outperform PROSAIL using reflectance at only few spectral bands around 550 nm and in the red-edge.

Satellite sensors cover large areas with good revisit period, so they fulfill the requirements of oil and gas companies^{20,68}. However, none of those currently available offer the spatial and

spectral resolutions needed for future application of our methods^{385,386}. This justified the use of airborne imagery instead of satellite in our study. It is likely that promising perspectives will arise in the future with the emergence of new satellite sensors. From an operational point of view, a growing interest is given to Unmanned Aerial Vehicles (UAVs). UAV-embedded sensors provide very high spatial resolution (up to cm scale) and allow multitemporal acquisitions, so they show many advantages for monitoring oil contamination^{387,388}. Although hyperspectral UAV-embedded sensors are still under development, multispectral sensors are commonly employed for assessing vegetation health in other contexts^{389,390}. Their use could be therefore expanded to that of oil contamination monitoring. More precisely, the spectral bands required for applying our methods could serve for designing sensors specifically dedicated to the detection and quantification of oil in homogenous vegetated industrial areas.

9.5. Conclusions

This study aimed to detect and quantify oil contamination (*i.e.* TPH) in soils in temperate vegetated areas using airborne imagery with very high spatial and spectral resolutions. A two-steps approach was proposed. As a first step, a method of detection exploiting 14 vegetation indices performed well for discriminating between healthy and oil-exposed bramble plants (OA = 98% and Kappa = 0.95). This method was tested on additional sites and succeeded in detecting oil contamination levels similar to those of the calibration site (17 – 39 g.kg⁻¹). The performance of the methods then decreased for lower levels of contamination, especially below 1 g.kg⁻¹. As a second step, two methods of TPH quantification were proposed and assessed on the same contaminated sites. The first one, which relied on LCC retrieval using PROSPECT and PROSAIL models, achieved accurate predictions of TPH concentrations both in the field, at leaf and canopy scales, and on the airborne image (RMSE \leq 3.20 g.kg⁻¹ and RPD \geq 2), thanks to consistent LCC estimations. The second method combined spectrum transformations with ENET regression. This method provided slightly less accurate quantification of TPH (best RMSE \leq 3.28 g.kg⁻¹ and RPD \geq 1.9), but required only few spectral bands in the VIS and the red-edge regions. Both methods were validated on another contaminated site and their performance was compared.

In the continuity of previous work, this multiscale study highlighted the importance of controlled and field conditions for developing reliable methods that can be applied to hyperspectral imagery. Our methods focused on tracking oil-induced alterations in leaf biochemistry while preventing from undesired effects (plant architecture, bare soil, and illuminating and viewing geometry), which made them accurate regardless of the acquisition scale. Very high spatial resolution helped achieving it. As discussed above, further research is needed to adapt our methods to various ecological contexts subject to oil contamination. In addition, promising perspectives of operational use will arise in the future with the emergence of new satellite and UAV-embedded sensors.

Author Contributions: Conceptualization, G.L., S.F., A.C., D.D. and A.E.; methodology, software and validation, G.L.; formal analysis, G.L.; investigation, G.L., S.F., A.C., R.H., G.B., D.D. and A.E.; resources, S.F., A.C. and R.H.; data curation, G.L.; writing—original draft preparation, G.L.; writing—review and editing, G.L., S.F., G.B. D.D. and A.E.; visualization, G.L.; supervision, S.F., A.E. and D.D.; project administration, G.L., S.F., A.C., D.D. and A.E.; funding acquisition, A.C. and D.D.

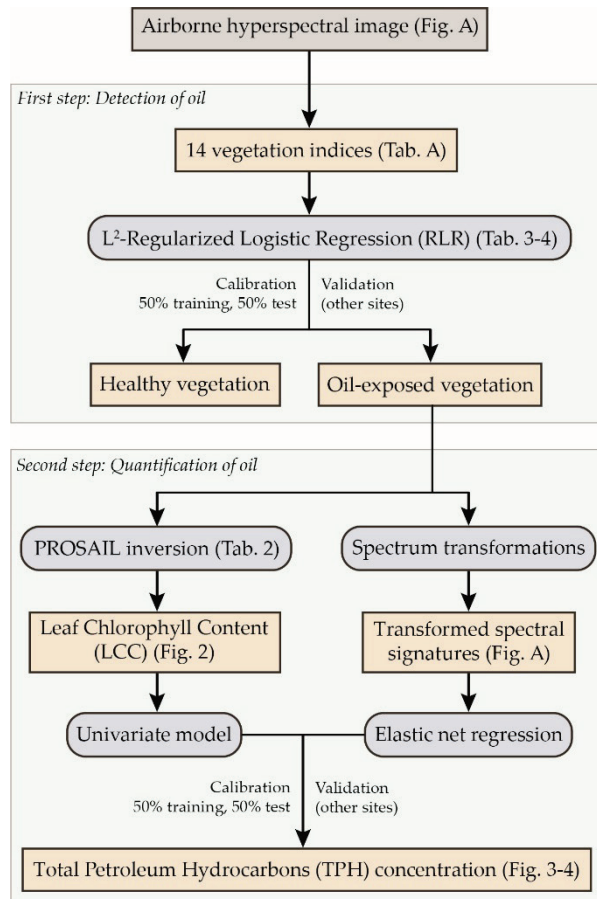
Funding: This research was funded by TOTAL.

Acknowledgements

This work was performed in the frame of the NAOMI project between TOTAL and the ONERA, with the support of Ecolab and Dynafor research units of Toulouse. The authors gratefully acknowledge V. Tardivat for the field access authorizations, V. Achard, K. Adeline and P. Borderies for their assistance in field sampling and P. Déliot and L. Poutier for their implication in image acquisitions and atmospheric correction.

Conflicts of Interest: The authors declare no conflict of interest.

Graphical abstract



SUMMARY AND CONCLUSION

This chapter described the research conducted in the field and on the airborne hyperspectral image, in the temperate context. The two studies carried out for this purpose completed the multiscale approach of the thesis initiated under controlled conditions. They are summarized in Table 20. While the method of oil detection was already tested in the field on known species, it proved efficient when applied to the airborne image, detecting almost all the pixels from the control and highly oil-contaminated mud pit sites (OA > 90%). The method remained robust to changes in ecological conditions – especially soil properties – and provided encouraging results for detecting low contamination (accuracy = 85.4% at 3.15 g.kg⁻¹ TPH). The detection limit of the method, which was estimated to 6 g.kg⁻¹ TPH under controlled conditions, was actually closer to 1 g.kg⁻¹ for long-term-established *R. fruticosus* plants under natural conditions, which may be due to additional factors enhancing the effects of oil. Conversely, the quantification of TPH had not been addressed for *R. fruticosus* in the first – experimental – step of the approach. The development of methods dedicated to that aim represented the main contribution of the studies presented in this chapter. Two methods were proposed, each having its advantages. The first one based on PROSPECT – and then PROSAIL – inversion has been subject to a first separate study at leaf scale. It has been shown that leaf pigment contents retrieved from the model – especially LCC – were strongly linked to TPH concentrations and could serve for quantification purpose. *R. fruticosus* was particularly well-suited for using of this approach, achieving accurate predictions of TPH (best RPD = 4.1). Other species were also suitable for applying the approach and sometimes outperformed *R. fruticosus* (RPD = 5.17 for *Q. pubescens* in summer 2017). The accuracy of the quantification was highly influenced by season changes, and was better using PROSPECT than vegetation indices and PLSR. This method provided satisfying TPH predictions when scaling to canopy measurements and the airborne image using PROSAIL (RPD ≥ 2), during a favorable season (summer). The second method based on spectrum transformation and ENET regression also succeeded in doing the same with few spectral bands. Both were successfully validated on a spatially-independent mud pit.

Chapters two and three provide a complete demonstration of using hyperspectral remote sensing for monitoring oil contamination in temperate vegetated areas. Although the proposed methods are not yet ready for operational use on imagery, promising ways of improvements have emerged toward applications to entire images. For example, it would be necessary to extend the scope of the methods on imagery; first, by testing them on other relevant species identified in the field (*Q. pubescens*, *P. canadensis*) and then, by adapting them to other contexts (*e.g.* tropical region, seepage prospecting, etc.). In addition, further studies should focus on using airborne hyperspectral imagery for assessing the influence of spatial and spectral resolution degradation on the performance of the methods. More precisely, the methods should be tested at the resolutions of future satellite-embedded hyperspectral sensors, which show great interest for operational use. In this perspective, multispectral sensors should be also considered. A first attempt was carried out for quantifying TPH at the spectral resolutions of planned HypXim and operating Worldview-3 satellite-embedded sensors (210 and 16 spectral bands, respectively), on the 10 plots of the brownfield. This was achieved by degrading the spectral resolution of the airborne hyperspectral image, while keeping the original 1-m spatial resolution. The method of quantification based on LCC retrieval – by inverting PROSAIL – was tested on the resulting 210- and 16-band images, following the same procedure as described in the fifth article in this chapter. In comparison to field-based estimations, LCC was still accurately retrieved at the spectral resolution of HypXim ($R^2 = 0.82$, RMSE = 3.78 $\mu\text{g.cm}^{-2}$), but this became more difficult at that of Worldview-3 ($R^2 = 0.70$, RMSE = 4.23 $\mu\text{g.cm}^{-2}$) (Figure 51). Consequently, the predictions of TPH were satisfying in the first case ($R^2 = 0.68$, RPD = 1.79), but suffered from

a more important variability of LCC within plots using Worldview-3 bands ($R^2 = 0.43$, $RPD = 1.33$).

Table 20. Summary of the studies carried out in the second and third steps of the approach, and the developed methods. (TPH: Total Petroleum Hydrocarbons; HM: Heavy Metals; ENET: Elastic net; RMSE: Root Mean Squared Error; RPD: Residual Predictive Deviation.)

	Field (<i>in situ</i>)	Field + Airborne images	
<u>Experimental design:</u>			
Context	Temperate	Temperate	
Species	<i>R. fruticosus</i> L., <i>Quercus pubescens</i> Wild., <i>Populus x canadensis</i> Moench., <i>Cornus sanguinea</i> L. and <i>Buddleja davidii</i> Franch.	<i>R. fruticosus</i> L.	
TPH concentration ⁱ	16 to 77 g.kg ⁻¹	17 to 39 g.kg ⁻¹	
Scales of reflectance acquisitions:	Leaf	Leaf, canopy and 1-m spatial resolution images	
<u>Proposed method:</u>			
Purpose	Quantification of TPH		Quantification of TPH ^j
Principle	PROSPECT model inversion + Univariate model	PROSAIL model inversion + Univariate model	Spectrum transformation + ENET regression
Main results	RMSE ≤ 5.86 g.kg ⁻¹ and RPD ≥ 3.04 for three out of the five species	RMSE ≤ 3.20 g.kg ⁻¹ and RPD ≥ 2 at all acquisition scales	RMSE ≤ 3.28 g.kg ⁻¹ and RPD ≥ 1.90 at all acquisition scales
Remarks	Accuracy strongly influenced by the species and the season	Accuracy slightly decreased with upscaling	Accuracy slightly decreased with upscaling

These results support the potential of future satellite-embedded hyperspectral sensors for quantifying oil contamination, as well as the limits of multispectral sensors. Further research should be conducted in that direction, by also taking changes in the spatial resolution and the sensor noise into account. Most of future satellite-embedded sensors will provide 8- or 30-m spatial resolution images, which has strong implications when applying the methods. This perspective is largely discussed in the following section.

ⁱFor the contaminated sites only.

^jThe method of TPH detection developed under controlled conditions and applied on the airborne images is not included here, as it was not properly developed during this step of the approach.

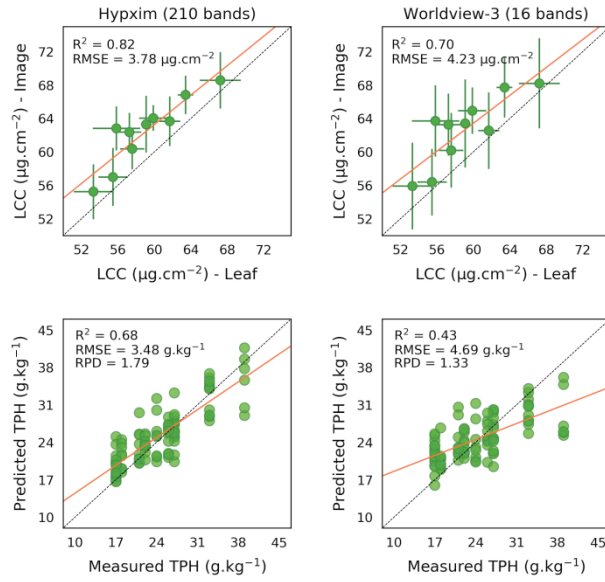


Figure 51. (top figures) Comparison between Leaf Chlorophyll Content (LCC) (mean \pm SD) retrieved in the field (leaf-scale measurements) and on the airborne image, and (bottom figures) comparison between measured and predicted TPH concentrations on the airborne image at the spectral resolutions of HypXim and Worldview-3 sensors.

GENERAL CONCLUSION

This thesis aimed at detecting and quantifying crude oil and petroleum products in vegetated regions using airborne hyperspectral imagery with very high spatial and spectral resolutions in the [400:2500] nm spectral domain. No reliable method had been previously proposed for this purpose. To achieve it, this thesis focused on mud pits contaminated by oil production residues with long-term-established, oil-tolerant, plant species. Based on the limits of previous studies – depicted in chapter one, a three-step, multiscale approach was adopted for developing the detection and quantification methods. These steps included greenhouse experiments (first step), field campaigns (second step) and airborne imagery acquisitions (third step) in a temperate context, involving *R. fruticosus* as a model species. At the end of the approach, three methods have been proposed and validated for use in the field, at leaf and canopy scales, and on airborne hyperspectral images; one for detecting and characterizing and two for quantifying oil contamination.

The first step of the approach served for developing the method for detecting and characterizing various types of oil contamination – and also water-deficit, including mud pit soils. For this purpose, relevant information about the sensitivity of the species to oil provided by the first – feasibility – experiment was exploited in the method. More precisely, stepwise LDA identified particular regions of the VIS and SWIR – related to leaf pigment and water contents – helpful for detecting oil (accuracy = 100 %), especially when combined in VI. Some of these VI were retained, based on their ability to track biochemical changes in leaves from leaf to multi-plant scale measurements. The proposed method, combining these VI in RLR classification, succeeded in distinguishing control soils from oil contaminated mud pit and water-deficient soils (accuracy $\geq 85\%$ after 18 days). The performance of the method varied with the contamination type and concentration and the time of exposure, and remained consistent among all acquisition scales. This first step of the approach ended with the successful validation of the method in the field, on naturally-established *R. fruticosus* plants (accuracy = 83 and 90% for control and contaminated mud pit soils, respectively).

The second step was intended to quantify TPH concentrations in mud pit soils from leaf optical properties. A method based on the retrieval of leaf pigment contents by inverting the PROSPECT model was proposed. It relied on predicting TPH concentrations by estimating the level of pigment alteration in leaves, after establishing the relationship between the two. The method provided accurate predictions of TPH from LCC when applied to *R. fruticosus* and two other species (RPD ≥ 3.04), but failed for other species not affected by the contamination or regardless of its concentration (RPD ≤ 1.02). The performance of the method was reduced using carotenoids instead of chlorophylls, and greatly impacted by season changes, making it reliable on summer and autumn only.

The third step of the approach was dedicated to assessing the method of detection – developed under controlled conditions – on airborne hyperspectral images; then to adapting the method of quantification to field-based canopy-scale measurements and to the images using PROSAIL. In this thesis, we did not consider applying the methods to large areas with mixed species or sparse vegetation. This point is discussed as a perspective in the following section. We assumed the location of *R. fruticosus* known in advance and we focused on homogenous covers. The method of detection using VI performed well for discriminating the control and highly-contaminated (TPH ≥ 17 g.kg⁻¹) mud pits soils colonized by *R. fruticosus* on the images, at high spatial (1 m) and spectral (5.2 – 7.8 nm) resolutions (accuracy $\geq 90\%$). Its scope was extended to low-contaminated soils (accuracy = 85.4% at 3.15 g.kg⁻¹ TPH) and reached a limit below a certain level (accuracy $\leq 15\%$ below 1 g.kg⁻¹ TPH). Regarding the method of quantification, accurate

predictions of TPH ($RPD \geq 2$) were obtained using LCC retrieved from both field measurements (leaf and canopy scales) and images using either PROSPECT or PROSAIL. During this step, an additional method of quantification coupling spectrum transformations to ENET regression was developed and achieved accurate predictions of TPH concentrations ($RPD \geq 1.90$). A first attempt using the spectral bands of HypXim and Worldview-3 satellite-embedded sensors indicated that high spectral resolution is necessary for this purpose.

Thus, the work performed in this thesis is the first to demonstrate that oil – and more specifically TPH – can be detected and quantified from vegetation optical properties using airborne hyperspectral imagery. None of the methods previously proposed allowed assessing oil contamination automatically and accurately, by predicting whether vegetation is – or has been – exposed. Our methods achieved it by taking advantage of VI in classification and regression and of RTM, which are commonly used in other contexts of vegetation stress assessment. The three-step multiscale approach adopted for this purpose proved suitable for developing methods adapted to the context of study. The same approach was initiated under tropical context, involving *C. alopecuroides* as a model species. By reproducing the realistic conditions of exposure to contaminated mud pit soils, LDA succeeded once again in detecting oil for this species (accuracy $>90\%$). The VI used for developing the method with *R. fruticosus* were also suitable for use with *C. alopecuroides*. A method for quantifying TPH was developed based on combining these VI in RF regression, and proved efficient at leaf and plant scales ($RPD \geq 4.44$).

This thesis demonstrated the feasibility of assessing mud pit contamination using very high spatial and spectral resolution imagery. At this stage, the methods are still under development, making them usable, provided that several conditions are fulfilled (species with known sensitivity and location, mud pit contamination, very high spatial and spectral resolutions). As a perspective, they are intended to be applied operationally in a broader range of contexts encountered in oil exploration and contamination monitoring (*e.g.* microseepage and pipeline leak detection). This implies extending the scope of the methods and overcoming their current limits in regard to operating and future sensors.

PERSPECTIVES

Current limits for operational use on satellite imagery

In a perspective of operational use, it is essential that remote sensing provides reliable detection and quantification of oil over large vegetated areas. Until now, the proposed methods remained poorly effective – or often unassessed – outside the initial study site^{20,68,232}, making them unusable operationally. Conversely, those developed in this thesis were successfully validated, by detecting and quantifying oil on independent mud pits that did not serve for method calibration. Our methods were however applied under several assumptions. Since we focused on *R. fruticosus* in the temperate context, the methods were only adapted to this species. Thus, they could be applied for identifying new contaminated mud pits and for quantifying TPH, provided the sites are colonized by dense *R. fruticosus* covers with known location. This is of great interest for monitoring contamination, because this species is widespread on industrial sites under temperate regions^{29,274,307}. These constraints are however very restrictive, because the methods can be applied only locally and to pre-selected vegetated sites. In an operational perspective, the methods should be applicable at large scale (*i.e.* to entire images) in a wide range of contexts (in terms of species and contamination type and level). It is however not conceivable to use airborne hyperspectral imagery – especially for daily monitoring, because it implies an important economic cost. Conversely, satellite imagery is already used operationally by oil and gas companies for mineralogical mapping and marine oil spill tracking^{21,243}. In that sense, the work performed in this thesis aimed at anticipating the use of future satellite-embedded hyperspectral sensors. Thanks to a good revisit time, satellite-embedded sensors show great interest for monitoring oil contamination continuously over industrial facilities. However, to date, none of the operating and planned hyperspectral satellite-embedded sensors covering the reflective domain offers a spatial resolution lower than 8 m (planned Hypxim resolution), and most provide 30-m resolution images with less than 250 spectral bands. In contrast, the methods developed in this thesis were assessed using airborne imagery with high spatial (1 m) and spectral (> 400 bands) resolutions. It is highly probable that, when applying these methods to satellite imagery, changes in spatial and spectral resolutions will affect their performance. Thus, in order to support oil exploration and contamination monitoring, the methods should improve in two ways: first, by extending their scope, and then, by adapting them to satellite imagery. These two perspectives are discussed jointly in this section.

Toward assessing mud pit contamination at large scale

At this stage, the need to know the location of *R. fruticosus* is one of the most important limits to the application of the methods at large scale. In an operational frame, an automatic mapping of this species would be helpful. Without this preliminary step, the methods would lead to false-detection alarms and inaccurate quantification of TPH if applied to other species, which differ in optical properties and sensitivity to oil. As discussed in chapter three, the mapping of *R. fruticosus* could be achieved quite easily for homogenous and dense covers, such as those studied in this thesis. Conversely, this species might become harder to locate when sparse or mixed with other species. It is particularly true when using satellite imagery, as “pure” pixels of dense vegetation (*i.e.* including only *R. fruticosus* and no bare soil) become even rarer at satellite spatial resolution. *R. fruticosus* is more likely to be mixed with other species and bare soil in

pixels, affecting the values of the 14 VI used in the detection method and consequently reducing its accuracy by introducing confusions. Likewise, PROSAIL inversions might experience difficulties in retrieving LCC accurately for sparse or plurispecific vegetation, because of the influence of bare soil and differences in optical properties, leaf orientation and sensitivity to oil among species in 8- or 30-m pixels^{73,88}. This would negatively affect TPH predictions. The use of PROSAIL variants and other PROSPECT-coupled RTM, such as the Discrete Anisotropic Radiative Transfer³⁹¹ (DART), might be more adapted for retrieving LCC in this case. The second method of quantification (ENET) might be also impacted at lower spatial and spectral resolutions. This method relies on exploiting reflectance or derivative reflectance over few spectral bands – located around 550 and 700 nm – for predicting TPH concentrations. Bare soil is very influential on these bands at 8- or 30-m spatial resolution, and might significantly alter the performance of the method. Moreover, it is likely that the useful information contained in these bands will be partly lost when degrading the spectral resolution, because of their concatenation with other less-informative, neighboring, bands.

In response to this limit, spectral unmixing⁵⁸ seems a promising solution. Unmixing aims at identifying the different materials – in our case, the different species and bare soil – inside pixels using spectral libraries. Lots of unmixing methods have been proposed in previous studies^{58,356,382,383}. Focusing on vegetation studies, they have been developed for two main purposes: mapping a single target species and discriminating among various ones. Thus, they could be used for mapping *R. fruticosus* – even mixed with bare soil or other species and vegetation types – before applying our methods of oil detection and quantification. Future studies should focus on assessing unmixing methods on known sites with this species, for example by degrading 1-m spatial resolution airborne images to 8- and 30-m resolution. A successful attempt has been achieved for mapping *R. fruticosus* using unmixing methods on HyMap images (spatial resolution < 10 m) in a previous study³⁸³. However, it might be interesting not to limit to this species in our case. First, this would be spatially too restrictive from an operational point of view. Moreover, other temperate species might serve for detecting and quantifying oil – such as *Q. pubescens* and *P. canadensis*, as highlighted in chapter three. Our methods were not tested on the airborne images for these species, but accurate predictions of TPH were obtained in the field using PROSPECT-retrieved LCC. Thus, they could be used along with *R. fruticosus* for assessing mud pit contamination at large scale, after being identified by spectral unmixing. Likewise, other suitable species – which allow both detecting and quantifying oil – might be found out in the field, by performing reflectance measurements on known control and contaminated mud pits. All the measured spectra could be exploited for unmixing purpose, in order to map the target species on hyperspectral images and then for detecting and quantifying TPH. In addition, multitemporal acquisitions could help discriminating species, thanks to the good revisit time of satellite-embedded sensors. A similar procedure should be adopted in a tropical context, based on the method of quantification developed for *C. alopecuroides*. This would extend the scope of our methods and fulfill operational needs.

Toward extending the scope of the methods to oil exploration and contamination monitoring

Even if the methods can be applied to entire images – provided that the target species have been identified, it is important to note that their performance will depend on the level of contamination. Focusing on *R. fruticosus*, the method of detection was initially adapted to high TPH concentrations ($\geq 17 \text{ g.kg}^{-1}$). Satisfying accuracy was obtained for low-contaminated mud

pit (3.15 g.kg⁻¹ TPH), but accuracy fell below 1 g.kg⁻¹ TPH. This helped determining a lower detection limit. Conversely, the method of quantification based on LCC retrieval was accurate within the range of TPH studied, and even above in the field (up to 77 g.kg⁻¹ TPH at leaf scale.). Its exact range of effectiveness remains however unknown. This information is essential for operational applications, because oil contamination can extend to a wide range of concentrations. Thus, further research should focus on determining the exact limits of detection and quantification of the methods. This might be achieved by first testing the methods on known mud pits contaminated by low to moderate (1 – 15 g.kg⁻¹) or very high (> 77 g.kg⁻¹) concentrations of TPH, using both field measurements and hyperspectral imagery. These limits may however vary among species, depending on their sensitivity; all of them do not allow detecting and quantifying mud pit contamination in the same range. Species with different sensitivities could be complementary for quantifying TPH over a wide range of concentrations. Spatial resolution might be also crucial, especially if TPH concentrations vary locally. In case of 8- or 30-m satellite imagery, pixels may include vegetation exposed to varying levels of contamination, making it difficult to quantify accurately. An important effort still remains to determine which species are suitable for applying the methods and their respective range of effectiveness at different spatial resolutions.

Although the scope of the methods is restricted to assessing mud pit contamination (*i.e.* production residues, oil sludge) at this stage, it should extend to other scenarios. Crude oil or petroleum product leaks deriving from pipeline or storage tank failures are priority. Along with wastewater, they represent the main sources of contaminant release from oil industry^{28,31}. Moreover, the detection of crude oil is of great interest in microseepage prospecting. This represents a major way of improvements toward operational applications. The method of detection and characterization based on VI proved promising for detecting crude oil-contaminated soils under controlled conditions (no confusion with control soils). Thus, it would be interesting to assess its performance using hyperspectral imagery over microseepages, vegetated pipeline routes and near storage tanks where *R. fruticosus* – or other suitable species – is established, by testing the method first to field measurements and then to hyperspectral images. To achieve the quantification of TPH for crude oil, it may however be necessary to adapt the methods based on LCC retrieval and ENET regression. Both methods rely on tracking alterations in biochemistry and reflectance, which differed for mud pits and crude oil-contaminated soils under experimental conditions. Consequently, their application remains limited to mud pit soils at this stage. To extend this scope to crude oil in soils, it is necessary to establish the relationship between LCC or reflectance and TPH specific to this type of contamination for each species. From the perspective of satellite imagery application, one possible limit to applying our methods may arise at the spatial resolution of satellite images for punctual oil patches. Microseepages and mud pits generally occupy large areas (>30 m), but pipeline and storage tank leaks can occur on a few square meters^{23,46}. Their detection might be very difficult without very high spatial resolution, because pixels will rarely include only oil-exposed vegetation. Thus, it is important to keep in mind that the needed spatial resolution also depends on the purpose of the detection.

Toward improving oil detection and quantification using UAV platforms and other remote sensing technologies

Hence, the exploratory work initiated in this thesis should continue in further research, in order to extend the scope of the methods and to assess their operational maturity. More precisely, future studies should first focus on identifying more relevant species and, for each of them, the

types of oil (*i.e.* crude oil and petroleum products) and the range of concentrations that can be detected or quantified. This would be helpful for remote sensing operators of oil and gas companies, as the methods could be used for a wide range of purposes in oil exploration and contamination monitoring. Prior to operational applications, the methods should be evaluated at the spatial and spectral resolutions of future satellite-embedded hyperspectral sensors, along with species unmixing. On the long term, oil and gas companies may spark growing interest in UAV-embedded hyperspectral sensors. Although they are still under development, they represent a promising alternative to satellite imagery. UAV-embedded sensors allow multitemporal, localized, monitoring, while providing very high spatial (up to cm scale) and spectral resolutions^{380,387}, therefore overcoming some of the above-mentioned limits. In addition, active remote sensing technologies could be used to improve oil detection and quantification, by providing complementary information about vegetation. For example, radar and LiDAR imagery are useful for estimating canopy height and biomass, which are affected by oil. Radar remote sensing is light-independent and atmospherically-resistant, which is a considerable advantage in wet tropical regions. An overview of its use for assessing oil contamination is presented in the Appendix section. Thus, by combining various technologies (active and passive) and sensor platforms (satellite, drone), remote sensing will undoubtedly become an indispensable support to oil exploration and contamination monitoring in vegetated areas in the coming decades.

APPENDIX: EVALUATION OF RADAR (SAR) REMOTE SENSING FOR ASSESSING VEGETATION HEALTH UNDER EXPOSURE TO CRUDE OIL AND PETROLEUM PRODUCTS

INTRODUCTION

In this thesis, hyperspectral imagery proved suitable for detecting and quantifying oil from vegetation optical properties. As a perspective, a first evaluation of its complementarity with radar remote sensing was performed. In oil and gas industry, radar remote sensing is mainly used in the off-shore domain. This active remote sensing technology relies on exploiting surfaces' backscatter at various frequency bands and polarizations of the microwave domain for detecting oil directly^{53,54}. It remains however unexploited for indirect detection purpose in vegetated areas, despite it already proved suitable for tracking changes in vegetation health in various agricultural and ecological contexts³⁹²⁻³⁹⁴. Radar remote sensing provides additional information about vegetation health to hyperspectral imagery, because it exploits longer wavelengths (Figure 52a). Moreover, it shows multiple advantages from an operational point of view. Radar remote sensing can be used day and night and is little sensitive to weather conditions (*i.e.* clouds, rain) and atmospheric constituents. Thus, it might be used in a complementarity way in oil exploration and contamination monitoring, in order to improve oil detection and quantification. To date, no research has been however carried out in this field.

This appendix aims at investigating the potential and the limits of radar remote sensing for detecting and quantifying oil using vegetation backscatter properties. For this purpose, we considered Synthetic Aperture Radar (SAR) imagery, which principle is illustrated in Figure 52b. This work was out of the initial scope of this thesis, so it is presented here as a perspective for future research. It relied on the experiments carried out under controlled conditions and on the analysis of airborne SAR images acquired over the mud pit in temperate region. The appendix is organized as follows. The first section provides an overview of vegetation backscatter properties and applications of radar remote sensing in vegetation health assessment. The second section describes the results of experiments carried out under controlled conditions, as well as those of SAR image analysis, in the temperate context.

a

Frequency band	Frequency range (GHz)	Wavelength range (cm)
L band	1 - 2	15 - 30
S band	2 - 4	7,5 - 15
C band	4 - 8	3,75 - 7,5
X band	8 - 12	2,5 - 3,75
Ku band	12 - 18	1,67 - 2,5
K band	18 - 27	1,11 - 1,37
Ka band	27 - 40	0,75 - 1,11
V band	40 - 75	0,4 - 0,75
W band	75 - 110	0,27 - 0,4

b

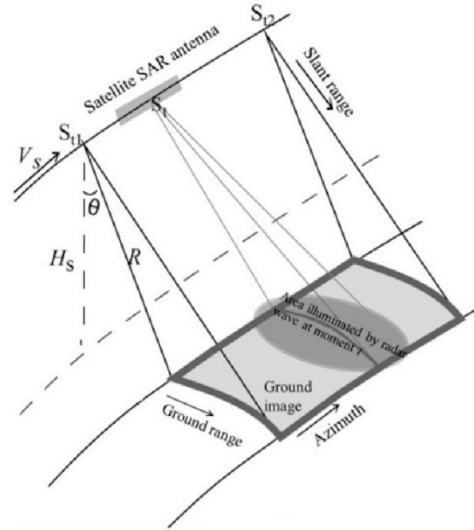


Figure 52. (a) Main frequency bands used in radar remote sensing. (b) Synthetic Aperture Radar (SAR) imaging principles (adapted from ³⁹⁵). (S_{i1} and S_{i2} : satellite initial and final position, respectively; V_s : movement speed; H_s : satellite altitude; θ : viewing angle; R : distance between the satellite and the ground.)

1. Overview of radar remote sensing for assessing vegetation health

Radar remote sensing provides information about surfaces from their backscatter properties in the microwave domain³⁹⁵. It is an active remote sensing technology. It relies on emitting a signal to a surface and analyzing the backscattered signal, in various frequency ranges (*i.e.* bands) and polarizations. The backscattering coefficient (σ^0 , in db) is then extracted from the backscattered signal. The use of radar remote sensing for assessing vegetation health is mostly based on centimeter wavelengths^{51,393,394}. Thus, it does not provide information about leaf anatomy or pigment contents, but it is useful for assessing vegetation development at canopy scale.

1.1. Vegetation backscatter properties in the 1 – 40 GHz microwave domain

Many factors influence vegetation backscatter properties in the microwave domain. The penetration of the emitted signal in the foliage is strongly determined by the frequency band, the polarization (horizontal or vertical) and various vegetation parameters, such as leaf and stem orientation, density and water content^{392,396,397}. Direct backscattering is observed either when the signal passes through the foliage and reaches the soil or when is backscattered by the canopy surface without penetrating (Figure 53). The signal may also bounce on the soil before reaching the vegetation – and vice versa. This phenomenon is called “double bounce”. Whether the signal reaches the canopy directly or through double bounce, it follows multiple interactions with leaves and stems, which generates “volume scattering” and attenuation (extinction) inside the volume. Thus, it may be difficult to distinguish the signal of the vegetation from that of the soil.

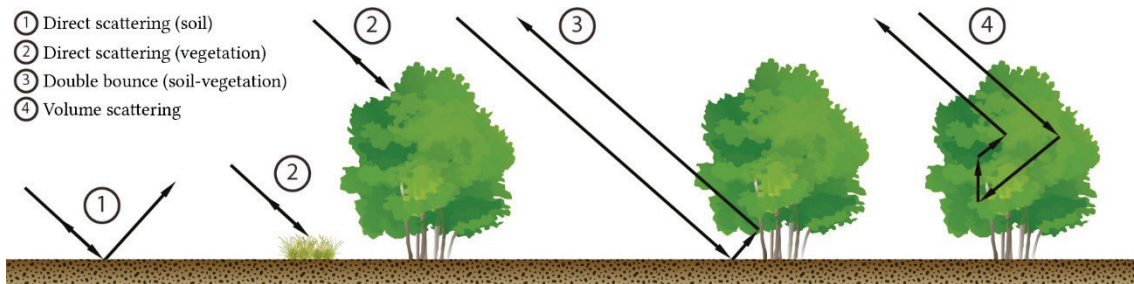


Figure 53. Components of the backscattering signal emitted in the microwave domain (reproduced after ³⁹⁶). Direct scattering of vegetation is sometimes also considered as volume scattering.

Plant biomass, LAI and LAD are three key parameters involved in vegetation backscatter properties. They are very influential on volume scattering, making possible to monitor vegetation growth and to discriminate among species with different canopy architecture (*e.g.* herbaceous, shrubby, trees)³⁹⁸. This suggest being able to detect oil-induced alteration of plant growth. The dielectric properties of vegetation – which are strongly linked to its water status – are also determining³⁹⁷. Changes in water availability induced, for example, by crude oil and petroleum products, affect plant dielectric properties and architecture. Thus, radar remote sensing can be used for monitoring vegetation water status and, more generally, for assessing vegetation health.

1.2. Application of radar remote sensing for assessing vegetation health

Focusing on vegetation studies, radar remote sensing has been used for a wide range of purposes in previous studies³⁹⁸. Plant growth and water status assessment were the most interesting ones in the frame of this thesis, because these parameters are usually affected by crude oil and petroleum products. L, C, X and P bands are usually exploited for these purposes^{392,393,397}. Most of studies focused on forests or crop species^{392,396,397}. Shrubby vegetation – such as *R. fruticosus* – has been subject to a very limited number of studies, all carried out in particular contexts^{399–402} (e.g. wetlands, arctic regions). Although they are of great interest in oil exploration and contamination monitoring, forests – and more generally trees – were out of the scope of this thesis. Moreover, their backscatter properties strongly differ from those of other vegetation types – especially herbaceous and shrubby ones. Crops are particularly adapted to radar remote sensing, because they are made of a monospecific, regularly- and homogeneously-distributed, plant cover. Shrubby vegetation should be somewhere in between crops and trees, in terms of backscattering properties.

Regardless of the vegetation type, radar remote sensing proved efficient for assessing plant development and water status. Previous studies relying on either field-based measurements using a fixed radar antenna or airborne and satellite imagery showed that plant's backscattering coefficient in L, C and X bands and several polarizations were highly correlated to LAI and plant biomass and water content^{392,397,403} (Figure 54). For example, X-band's backscattering coefficient increases with plant biomass, but reaches a plateau above a certain value – which depends on the species. The Radar Vegetation Index⁴⁰³ (RVI) has been specified for similar purpose. Thus, radar remote sensing might be helpful for tracking oil-induced alterations in plant growth and water status.

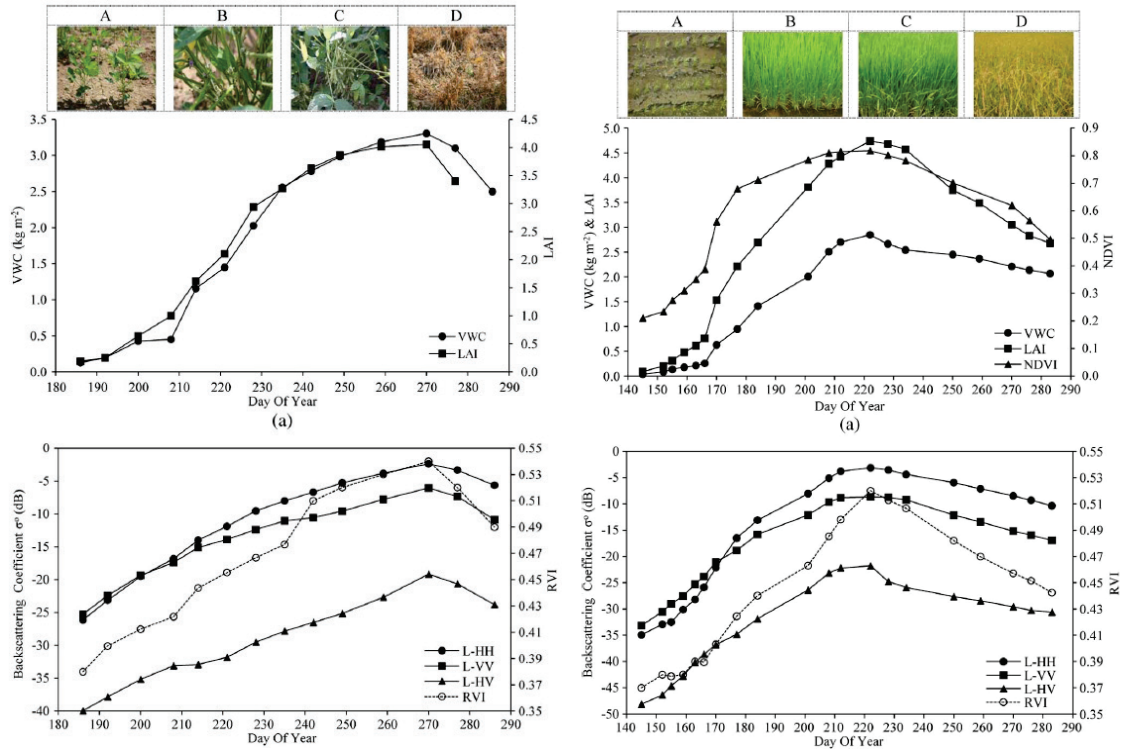


Figure 54. (top figures) Temporal evolution of Volumetric Water Content (VWC), Leaf Area Index (LAI) and Normalized Difference Vegetation Index (NDVI) of rice and soybean. (bottom figures) Temporal evolution of L-band backscattering coefficients (σ^0) in different polarizations (H: Horizontal; V: Vertical) and Radar Vegetation Index ($RVI = 8\sigma_{HV}^0 / (\sigma_{HH}^0 + \sigma_{VV}^0 + 2\sigma_{HV}^0)$) for these species³⁹⁷.

2. Radar remote sensing for detecting and quantifying crude oil and petroleum products

In this thesis, radar remote sensing was used for detecting and quantifying crude oil and petroleum products through their effects on vegetation growth and water status. For this purpose, we focused on the temperate context. Radar remote sensing was assessed in two steps; first, during the experiment carried out on *R. fruticosus* under controlled conditions, and then on airborne SAR images acquired over the same study site as that of hyperspectral images. The main results are presented in this section.

2.1. Application to oil detection under controlled conditions

2.1.1. Radar backscattering measurements

Based on the experiment carried out on *R. fruticosus* under controlled conditions (see Second Article in Chapter II), radar remote sensing was tested for detecting and discriminating among the control, water-deficient and various oil-contaminated soils to which the species was exposed. X-band (8 – 12 GHz) backscattering measurements have been performed in vertical-emitted vertical-received polarization (VV) at multi-plant scale during the experiment. For this purpose, nine individuals (3 x 3) of the same treatment were joined on a rotative platform, forming a 45

cm square (Figure 55). The configuration of the acquisition system was quite similar to that described by Boisot⁴⁰⁴. To be representative of airborne or satellite measurement configuration with a plane wave illumination, the foliage has to be located inside the near field zone of the radiating aperture. In practice, a 40-cm wide reflector antenna using a broadband horn as primary source was fixed 2 m away from the center of the platform on a support in cross shape, at 20° zenith angle, providing a 40-cm wide acquisition footprint. Measurements were performed outside after 5, 18 and 32 days of experiment, as follows.

The signal was emitted and received using a vector network analyzer in S11 mode. Frequency measurements were repeated $n = 360$ times by varying the acquisition angle from 1° between each measurement, thanks to the rotative platform. Calibration measurements were carried out first without the plants, using a trihedral corner reflector. In order to isolate the signal of the plants from that of the environment (*i.e.* the pots, the platform and the soil), a range gating was applied. The impulse response of calibration helped determining the range in which the plants foliage were located on the impulse responses. Plants' impulse responses were thus isolated and averaged on this distance range before computing the backscattering coefficient⁴⁰⁵. This isolation procedure was fundamental, because the ground response was parasitic in our case.

The backscattering coefficients were computed as follows. We considered a frequency ramp of width Δ_f being the frequency step ∂_f . The Fourier transform of the complex data acquired gave the impulse response, where the range resolution was $c/2\Delta_f$ and the unambiguous range was $c/2\partial_f$, where c is the light velocity. The mean backscattering coefficient corresponded to the incoherent field. The latter was extracted from the total field obtained over the n successive measurements after the complex subtraction of the coherent field, which was the complex average of these n measurements. Consequently, the fixed echoes (*e.g.* the various reflections on the cables and at the antenna aperture) were confined to the coherent field. In the impulse responses, a range gating was performed so that only the bins corresponding to the minimum and the maximum of the foliage distance were considered. They were summed in square amplitude and then averaged. The precision of the backscattering coefficient increased as \sqrt{n} .

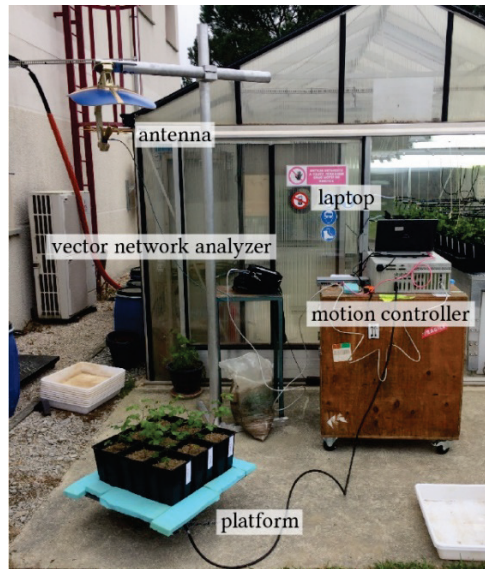


Figure 55. Experimental setup used for radar backscattering acquisitions on vegetation during the experiment.

At the end of the experiment, plant shoots were harvested, oven-dried for 48 h at 60° C, and then weighed. The backscattering coefficient of plants extracted from the processing were compared graphically among treatments, as only one measurement could be performed for each of them, so no statistical comparison was possible. Backscattering coefficients from day 32 were linked to total dry biomass and water content through simple regression.

2.1.2. Results and discussion

The temporal evolution of plants' backscattering coefficient was consistent with that of the biochemical and optical responses previously described. No difference was observed among the treatments on day 5 as plants remained no or little affected on this date (Figure 56). Then, the limited development of plants exposed to water-deficient soils ("Water_str"), muddy residues with 25 g.kg⁻¹ TPH ("Mud_B") or crude oil ("Crude") reduced their backscattering coefficient, so that it became possible to discriminate between them and those from the other treatments. This response was observed to a greater extent for plants exposed to water-deficient soils, which might come from their dropping habit. Conversely, the backscatter of control plants slightly increased on day 18 and 32, making possible to distinguish them from all other treatments at the end of the experiment.

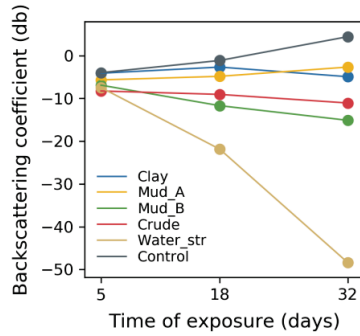


Figure 56. Temporal evolution of *R. fruticosus* backscattering coefficients for the different treatments (see Second Article in Chapter II for details about the treatments).

The backscattering coefficient of *R. fruticosus* measured at day 32 was strongly correlated with total dry biomass and plant water content (Figure 57), which differed among treatments. These results were consistent with those observed on other vegetation types in previous studies (*e.g.* herbaceous vegetation, crops), and emphasized the capability of radar remote sensing to track alterations in vegetation health. All the results described here remained little affected by the acquisition angle, as leaf orientation was similar among the treatments – except for plants exposed to water-deficient soils.

Thus, this study confirmed the suitability of radar remote sensing for assessing vegetation health under exposure to crude oil and mud pit contamination. The next step of this work aimed at investigating its interest for detecting and quantifying oil under natural conditions using airborne SAR imagery. This was very challenging in comparison to experimentally-controlled acquisitions.

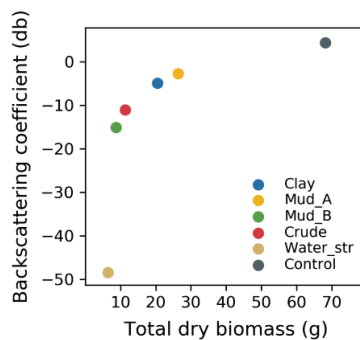


Figure 57. Relationship between *R. fruticosus* total dry biomass and backscattering coefficients after 32 days of exposure to the different treatments (see Second Article in Chapter II for details about the treatments).

2.2. Application to oil detection and quantification using airborne SAR imagery

2.2.1. SAR images acquisitions

On the same day as hyperspectral images acquisitions, SAR images were acquired from the same airborne platform over the study area in temperate region (see Fifth Article in Chapter III

for details about the study area) (Figure 58). The images were acquired in L and X bands and full polarization (*i.e.* HH, VV, HV and VH emitted-received polarizations), at 6900 ft. above sea level and 45° incidence angle (West – East orientation). After preprocessing, the resulting 20-cm spatial resolution images were used for evaluating the capability of radar remote sensing to detect and quantify oil contamination on the large mud pit (*i.e.* the “brownfield”) colonized by *R. fruticosus*. For this purpose, we focused on the control site and the 10 plots of the brownfield contaminated by 17 to 39 g.kg⁻¹ TPH that served for applying the methods based on optical properties. The backscattering coefficients of the pixels corresponding to each site were extracted from the SAR images using field GPS coordinates. The mean of the backscattering coefficients was computed for each site, compared graphically between the control and contaminated ones, and linked to TPH concentrations on the 10 plots of the brownfield.

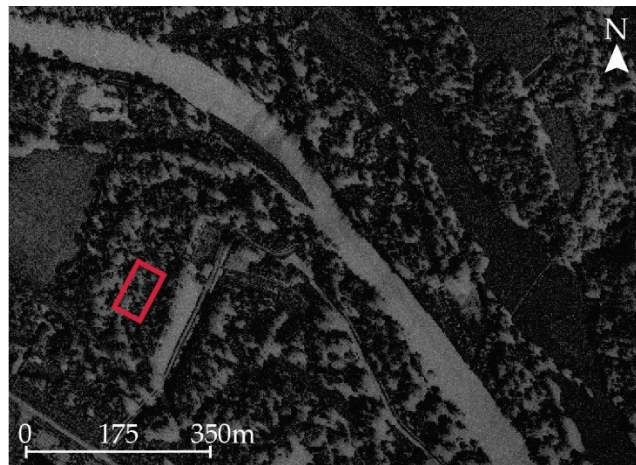


Figure 58. Subset of the airborne X-band SAR image acquired over the study area. The contaminated site (“brownfield”) is illustrated by the red polygon.

2.2.2. Results and discussion

X and L bands usually provide different information about shrubby vegetation. While the penetration of X-band (8 – 12 GHz) waves is limited through the foliage (few cm, depending on the geometry), L-band (1 – 2 GHz) waves reach the lower canopy – and eventually the soil – layers, so it is more subject to volume scattering. Because of the large incidence angle of acquisitions (45°), shadowing, layover effects (foreshortening) have affected the average X-band’s backscattering coefficient of two out of the 10 plots of the brownfield. Projected shadows from surrounding tree have been observed for these plots – which were located at the border of the site, so they were not considered in analyses.

Regardless of the band and polarization, no difference in averaged backscattering coefficient was observed between the control and contaminated sites on the images (Figure 59). Although the control site exhibited a higher backscattering coefficient than most of contaminated plots (17 – 28 g.kg⁻¹ TPH), it was not possible to distinguish it from that contaminated by 33 g.kg⁻¹ TPH. In addition, no clear relationship was established between either X- or L-band’s backscattering coefficient and soil TPH concentrations. Similar results were obtained by computing the RVI.

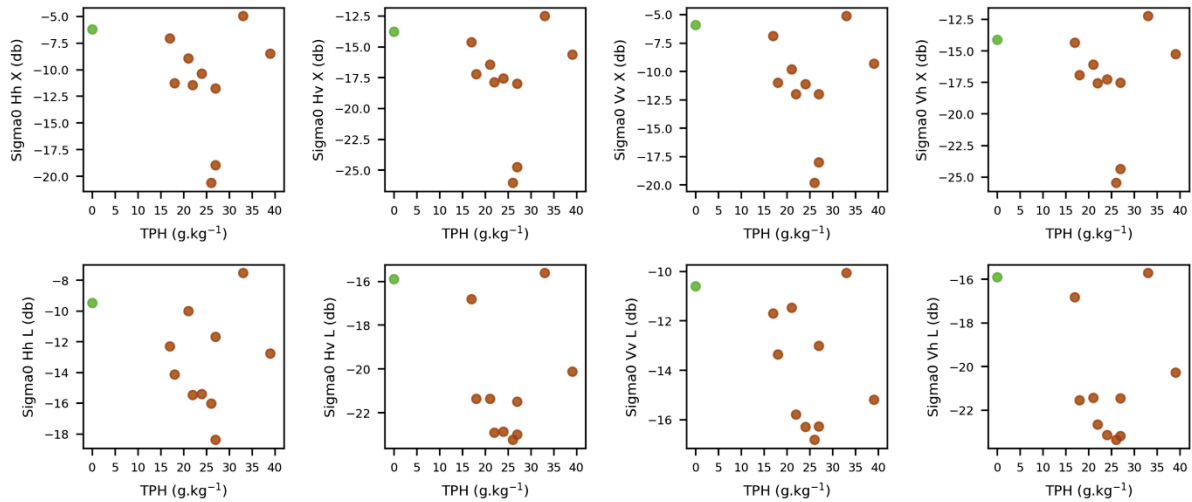


Figure 59. (top figures) X- and (bottom figures) L-band's backscattering coefficients observed for the control site (green) and the 10 plots of the brownfield (brown). (TPH: Total Petroleum Hydrocarbons.)

Based on visible observations made in the field, control and contaminated sites differed by LAI (> 2 and 1.2 , respectively). This was confirmed later by PROSAIL inversion outputs provided from the hyperspectral images. Leaf water content was also higher on the control site, whereas no difference in leaf orientation was observed between those sites. The sensitivity of the X band may have been insufficient for detecting these differences on SAR images. The contribution of the soil may have been too important in L-band's backscattering coefficient, as volume scattering is limited in bramble, which is mainly organized in one or two layers of leaves, overlying a stem network. Additional factors might have also contributed to the observations, such as local shadowing or radar clutter. Thus, radar remote sensing was unable to detect and quantify oil contamination in our context. Further data analysis was necessary to limit the influence of the undesired effects – especially shadowing and layover – on the signal.

SUMMARY AND CONCLUSION

This exploratory study highlighted the potential and the limits of radar remote sensing for detecting and quantifying oil contamination in vegetated mud pits, in temperate context. The X-band backscattering properties of *R. fruticosus* were affected by exposure to crude oil and petroleum products under experimental conditions. This effect was explained by alterations in plant growth and water status. Conversely, L- and X-band SAR imagery did not allow detecting oil contamination in mud pit under natural conditions. Based on this work, the complementarity of radar remote sensing to hyperspectral imagery for oil exploration and contamination monitoring was limited for shrubby vegetation. Further approaches from the literature should be tested in complement to the simple analysis performed here, in order to confirm or disprove this conclusion. Other bands – especially C and P bands – and multitemporal acquisitions should be also considered. Moreover, it is important to note that even radar remote sensing was not helpful for detecting mud pit contamination in our case, it might be more efficient in other contexts with homogeneous ground cover. For example, it could be used for monitoring oil leakages, which are more likely to affect highly sensitive species. SAR imagery might be particularly adapted in cloudy tropical regions with dense vegetation, where volume scattering is important and limits the contribution of the soil.

SUPPORTING INFORMATION

First article: Assessing soil contamination due to oil and gas production using vegetation hyperspectral reflectance

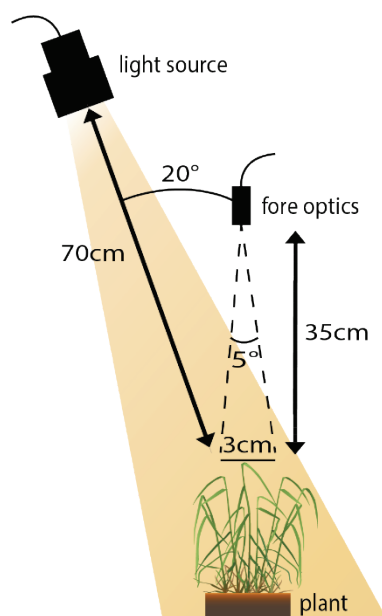


Figure 60. Illustration of the hyperspectral acquisitions carried out at plant scale on the three species.

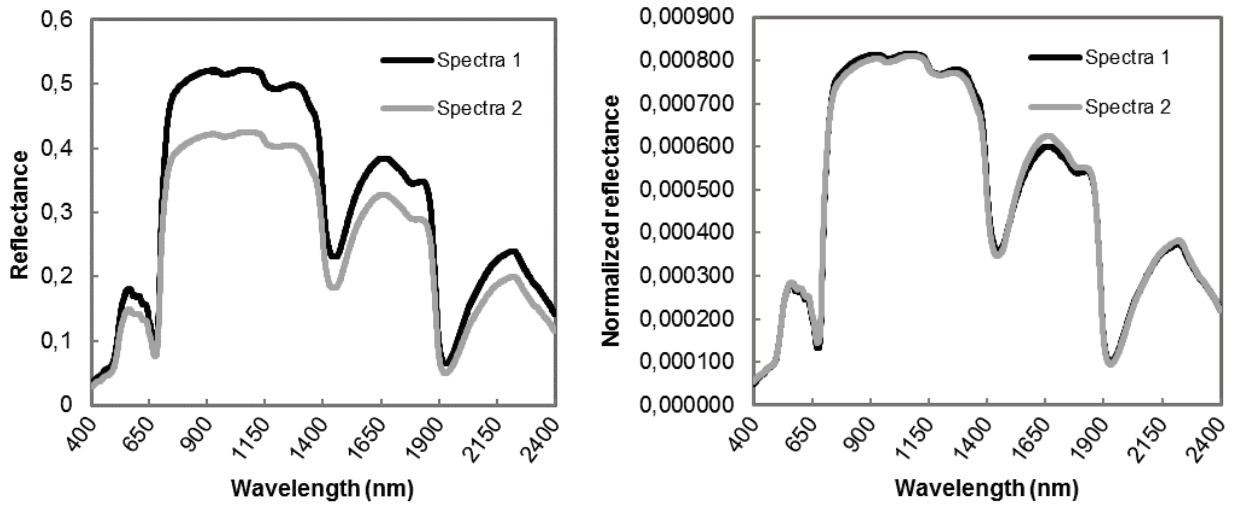


Figure 61. Example of spectral signatures (a) before and (b) after Area Under Curve Normalization (AUCN).

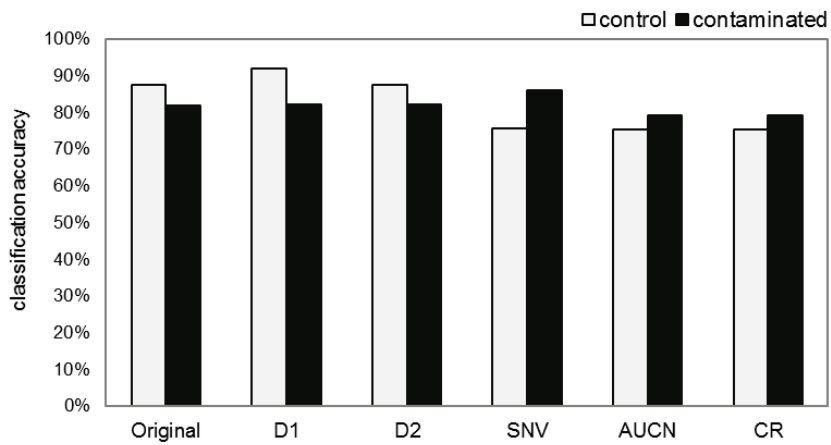


Figure 62. Classification accuracy obtained from LDA performed on Cenchrus reflectance data, all acquisition dates combined (49-61 days).

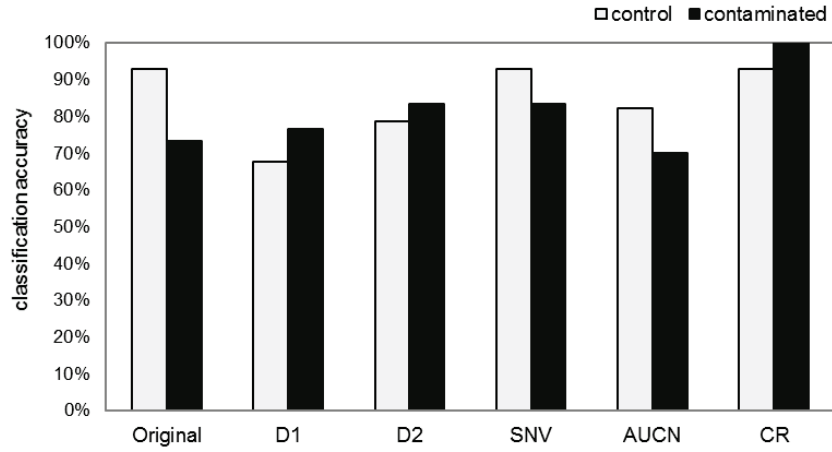


Figure 63. Classification accuracy obtained from LDA performed on *Panicum* reflectance data, all acquisition dates combined (49-61 days).

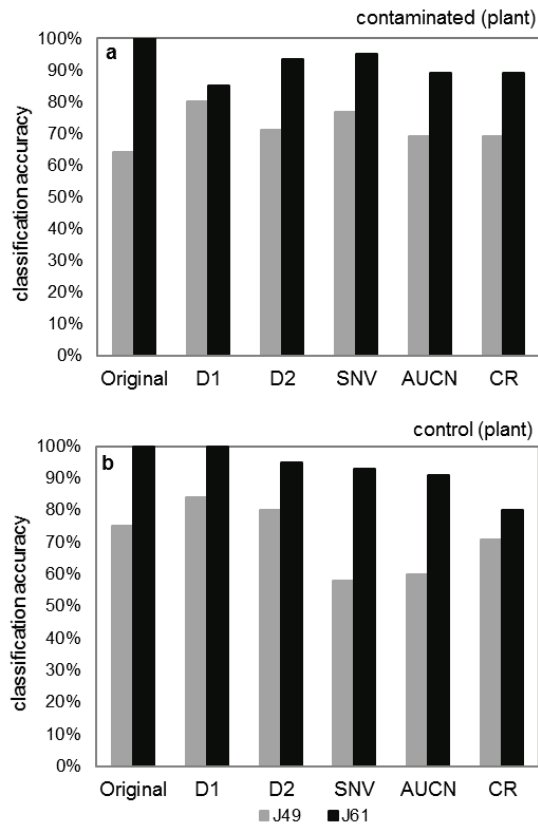


Figure 64. Temporal evolution of classification accuracy obtained from LDA performed on *Cenchrus* reflectance data. Results are displayed from day 49 ('J49') to 61 ('J61').

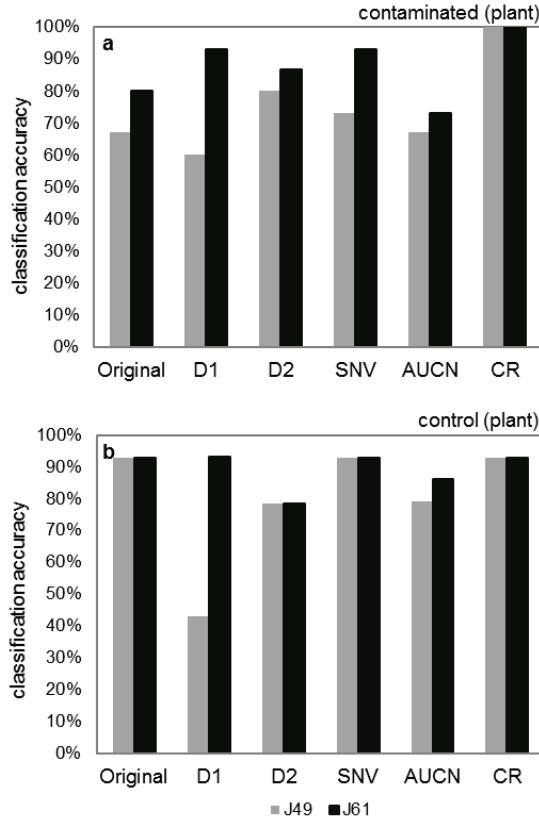


Figure 65. Temporal evolution of classification accuracy obtained from LDA performed on Panicum reflectance data. Results are displayed from day 49 ('J49') to 61 ('J61').

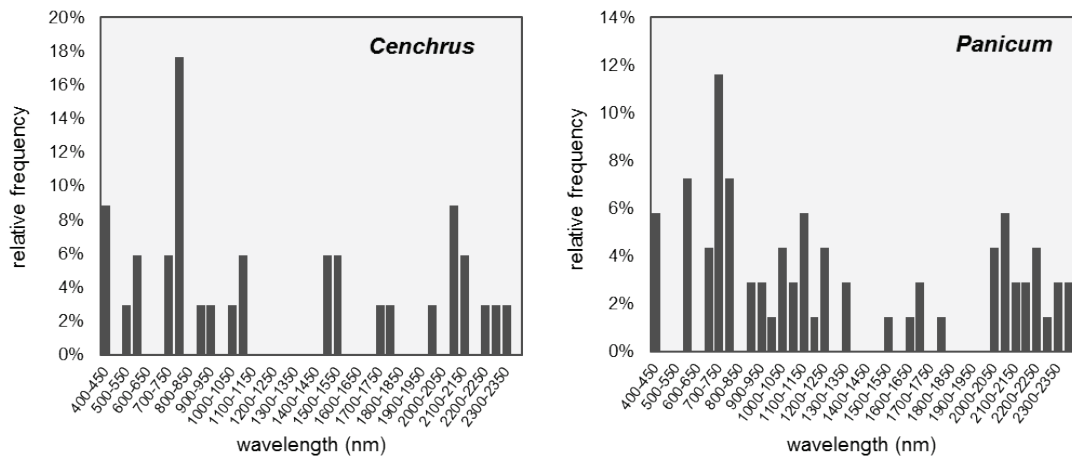


Figure 66. Discriminant wavelengths obtained from LDA performed on Panicum and Cenchrus at plant scale. Frequencies are presented all untransformed and transformed data combined, with bands grouped in 50 nm-wide intervals.

Second article: Detection and discrimination of various oil-contaminated soils using vegetation reflectance



Figure 67. Examples of bramble plants joined for multi-plant scale measurements (left: Control treatment, right: Clay treatment). Individual black pots of 15 x 15 x 18 cm (L x W x H) were used for the experiment. These pots had no influence on measurements at multi-plant scale, since their reflectance was close to zero. Measurements were carried out with a 20-cm acquisition footprint, which included most of the four plants.

Third article: Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression

Table 21. BTEX, Polycyclic Aromatic Hydrocarbons (PAHs) and Total Petroleum Hydrocarbon concentrations (g.kg^{-1}) found in the soils used for the treatments applied to *Cenchrus alopecuroides* (L.). Polycyclic Aromatic Hydrocarbons (PAHs) include the 16 priority pollutants identified by the U.S. Environmental Protection Agency (EPA). (<DL: Below the detection limit.)

	Control	Oil-contaminated mud pit soils			
<u>BTEX:</u>	<DL	<DL	<DL	0.0027	0.0046
<u>Polycyclic Aromatic Hydrocarbons (PAHs):</u>	<DL	<DL	<DL	0.045	0.099
<u>Total Petroleum Hydrocarbons (TPH):</u>					
C ₅ -C ₁₀	<DL	<DL	<DL	0.065	0.150
C ₁₀ -C ₂₁	<DL	0.140	0.850	4.635	7.850
C ₂₁ -C ₄₀	<DL	0.860	4.150	8.300	11.000
C ₅ -C ₄₀	<DL	1	5	13	19

Table 22. Vegetation indices used in this study (R: Reflectance, D: First derivative).

Index	Formula	Reference
Chlorophyll Absorption Ratio Index	$CARI = R_{700} R_{670}a + R_{670} + b / R_{670} (a^2 + 1)^{0.5}$ with $a = (R_{700} - R_{550}) / 150$ and $b = R_{550} - R_{550}a$	296
Chlorophyll/Carotenoids Index	$CCI = D_{720} / D_{700}$	297
Carter Index 1	$CTR1 = R_{695} / R_{420}$	298
Carter Index 2	$CTR2 = R_{695} / R_{760}$	298
Gitelson & Merzlyak Index 1	$M1 = R_{750} / R_{550}$	299
Gitelson & Merzlyak Index 2	$M2 = R_{750} / R_{700}$	299
Lichtenthaler Index 1	$LI1 = (R_{800} - R_{680}) / (R_{800} + R_{680})$	253
Lichtenthaler Index 2	$LI2 = R_{440} / R_{690}$	253
Lichtenthaler Index 3	$LI3 = R_{440} / R_{740}$	253
modified Chlorophyll Absorption Ratio Index 1	$mCARI1 = 1.2[2.5(R_{800} - R_{670}) - 1.3(R_{800} - R_{550})]$	130
modified Chlorophyll Absorption Ratio Index 2	$mCARI2 = \frac{1.5[2.5(R_{800} - R_{670}) - 1.3(R_{800} - R_{550})]}{\sqrt{(2R_{800} + 1)^2 - (6R_{800} - 5\sqrt{R_{670}}) - 0.5}}$	130
modified Simple Ratio 705 nm	$mSR705 = (R_{750} - R_{445}) / (R_{750} + R_{445})$	78
MERIS Terrestrial Chlorophyll Index	$MTCI = (R_{754} - R_{709}) / (R_{709} + R_{681})$	300
Normalized Difference 705 nm	$ND705 \text{ or } NDVI705 = (R_{750} - R_{705}) / (R_{750} + R_{705})$	78
Normalized Pigment Chlorophyll Index	$NPCI = (R_{680} - R_{430}) / (R_{680} + R_{430})$	301
Optimized Soil-Adjusted Vegetation Index	$OSAVI = (1 + 0.16) (R_{800} - R_{670}) / (R_{800} + R_{670} + 0.16)$	302
Photochemical Reflectance Index 1	$PRI1 = (R_{528} - R_{567}) / (R_{528} + R_{567})$	101
Photochemical Reflectance Index 2	$PRI2 = (R_{531} - R_{570}) / (R_{531} + R_{570})$	101
Photochemical Reflectance Index 3	$PRI3 = (R_{570} - R_{539}) / (R_{570} + R_{539})$	101
Plant Senescence Reflectance Index	$PSRI = (R_{678} - R_{500}) / R_{750}$	303
Pigment Specific Simple Ratio a	$PSSRa = R_{800} / R_{680}$	92
Pigment Specific Simple Ratio b	$PSSRb = R_{800} / R_{635}$	240
Pigment Specific Simple Ratio c	$PSSRc = R_{800} / R_{470}$	240
Structure Intensive Pigment Index 1	$SIP11 = (R_{800} - R_{445}) / (R_{800} + R_{680})$	81
Structure Intensive Pigment Index 2	$SIP12 = (R_{800} - R_{505}) / (R_{800} + R_{690})$	81
Structure Intensive Pigment Index 3	$SIP13 = (R_{800} - R_{470}) / (R_{800} + R_{680})$	81
Simple Ratio 705 nm	$SR705 = R_{750} / R_{705}$	78
Transformed CARI	$TCARI = 3[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550}) (R_{700} / R_{670})]$	304
Transformed CARI / OSAVI	$\frac{TCARI}{OSAVI} = \frac{3[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550}) (R_{700} / R_{670})]}{(1 + 0.16) (R_{800} - R_{670}) / (R_{800} + R_{670} + 0.16)}$	304
Vogelmann Index 1	$VO1 = R_{740} / R_{720}$	305
Vogelmann Index 2	$VO2 = (R_{734} - R_{747}) / (R_{715} + R_{726})$	306
Vogelmann Index 3	$VO3 = (R_{734} - R_{747}) / (R_{715} + R_{720})$	306
Disease Water Stress Index	$DWSI = (R_{800} - R_{550}) / (R_{1660} + R_{680})$	207

Fourth article: Application of PROSPECT for estimating Total Petroleum Hydrocarbons in contaminated soils from leaf optical properties

Table 23. Heavy metal and hydrocarbon concentrations ($\text{g}\cdot\text{kg}^{-1}$) found in the brownfield and control soils. Polycyclic Aromatic Hydrocarbons (PAHs) include the 16 priority pollutants identified by the U.S. Environmental Protection Agency (EPA). Heavy metals found in the control site corresponded to the geochemical background. (<DL: below detection limit.)

	Brownfield		Control	
	mean	sd	mean	sd
<u>Heavy metals (HM):</u>				
arsenic	0.023	0.004	0.008	0.002
cadmium	0.002	0.000	<DL	<DL
chromium	2.022	0.330	0.023	0.005
copper	0.074	0.015	0.009	0.005
mercury	0.006	0.002	<DL	<DL
lead	0.087	0.020	0.024	0.007
nickel	0.095	0.038	0.0135	0.002
zinc	4.339	0.817	0.0395	0.12
<u>Polycyclic Aromatic Hydrocarbons (PAHs):</u>				
	0.035	0.013	<DL	<DL
<u>Total Petroleum Hydrocarbons (TPH):</u>				
C ₁₀ -C ₁₂	0.42	0.44	<DL	<DL
C ₁₂ -C ₁₆	2.73	1.99	<DL	<DL
C ₁₆ -C ₂₁	3.85	1.80	<DL	<DL
C ₂₁ -C ₄₀	23.48	12.42	<DL	<DL
C ₁₀ -C ₄₀	30.52	16.18	<DL	<DL

Table 24. Statistics of biochemical parameters obtained from leaf analysis (n = 152 leaves). These data were used for validating PROSPECT inversion outputs.

	<i>B. davidii</i>	<i>R. fruticosus</i>	<i>C. sanguinea</i>	<i>Q. pubescens</i>	<i>P. canadensis</i>
<u>Chlorophylls (C_{ab}, µg.cm⁻²)</u>					
min	17.4	37.2	12.5	16.9	21.4
med	31.1	58.2	18.1	32.4	37.1
max	48.4	69.2	44.0	43.2	45.9
mean	32.0	56.7	19.8	31.7	34.8
sd	10.1	8.2	6.8	7.6	8.4
<u>Carotenoids (C_{xc}, µg.cm⁻²)</u>					
min	4.7	8.3	3.7	3.1	4.9
med	10.3	15.0	6.9	7.5	7.3
max	16.0	17.9	15.2	10.1	9.3
mean	10.2	14.5	8.6	7.2	7.3
sd	3.1	2.5	4.1	1.7	1.2
<u>Water (C_w, g.cm⁻²):</u>					
min	0.0072	0.0057	0.0017	0.0050	0.0090
med	0.0181	0.0122	0.0055	0.0069	0.0125
max	0.0278	0.0201	0.0087	0.0093	0.0180
mean	0.0177	0.0124	0.0056	0.0071	0.0124
sd	0.0068	0.0031	0.0019	0.0013	0.0020
<u>Dry matter (C_m, g.cm⁻²):</u>					
min	0.0031	0.0029	0.0002	0.0031	0.0050
med	0.0040	0.0048	0.0023	0.0050	0.0059
max	0.0053	0.0064	0.0047	0.0067	0.0067
mean	0.0041	0.0048	0.0023	0.0049	0.0059
sd	0.0006	0.0007	0.0015	0.0009	0.0005

Table 25. Statistics of the datasets used for calibration (train) and validation (test) of the models. The same number of samples (n) was used as train and test sets on each date (except in March 2018 where only few *B. davidii* and *R. fruticosus* were found on the site).

	<i>B. davidii</i>		<i>R. fruticosus</i>		<i>C. sanguinea</i>		<i>Q. pubescens</i>		<i>P. canadensis</i>	
	train	test	train	test	train	test	train	test	train	test
	n = 27	n = 27	n = 34	n = 35	n = 34	n = 35	n = 28	n = 29	n = 30	n = 30
<u>C₁₀-C₄₀ TPH (g.kg⁻¹):</u>										
mean	33.6	30.1	32	29	31.4	29.7	33.4	32	30.9	29.2
sd	17.3	16.3	15.9	14.9	16	16	17.1	16.5	16.9	17.8

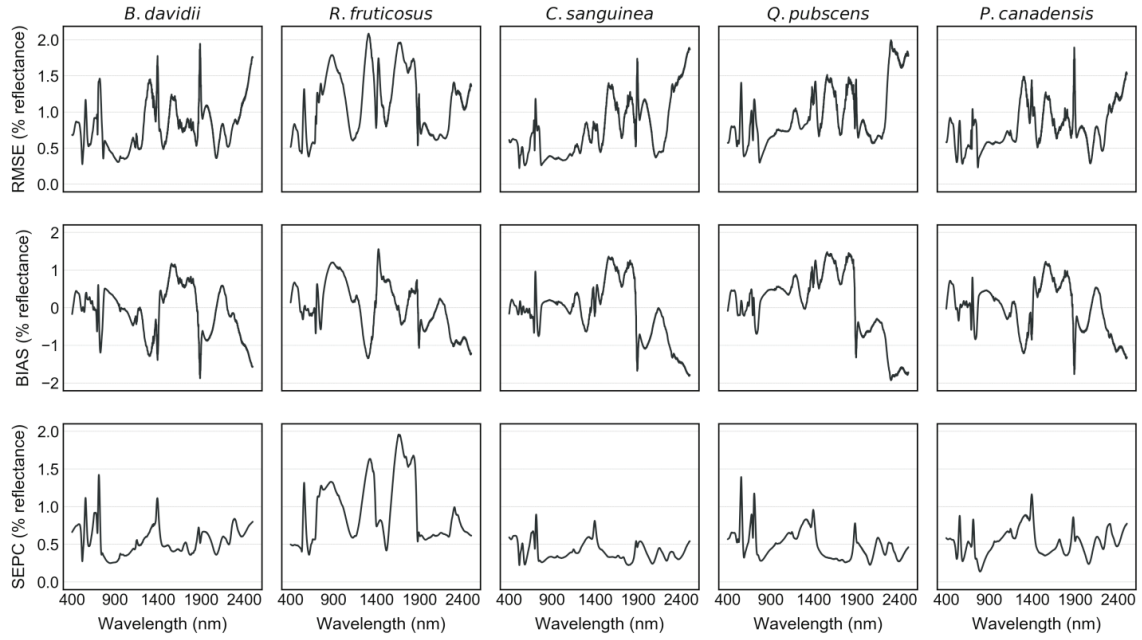


Figure 68. RMSE, BIAS and SEPC computed between all simulated and measured spectral signatures.

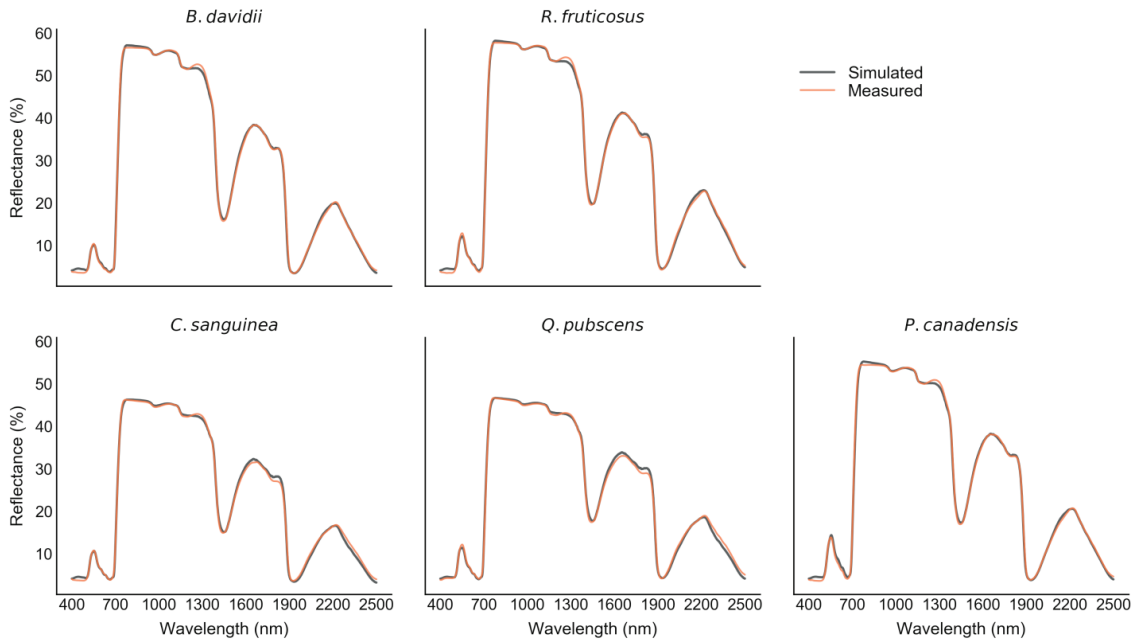


Figure 69. Examples of simulated and measured spectral signatures.

Table 26. Assessment of biochemical parameter retrieval using PROSPECT inversions.

	<i>B. davidii</i>	<i>R. fruticosus</i>	<i>C. sanguinea</i>	<i>Q. pubescens</i>	<i>P. canadensis</i>
<u>Chlorophylls (C_{chl}, $\mu\text{g}\cdot\text{cm}^{-2}$):</u>					
R ²	0.97	0.92	0.91	0.91	0.96
RMSE	2.59	2.65	2.67	2.65	2.95
BIAS	2.03	-0.42	-1.52	-1.43	-2.39
SEPC	1.61	2.62	2.19	2.23	1.73
<u>Carotenoids (C_{xc}, $\mu\text{g}\cdot\text{cm}^{-2}$):</u>					
R ²	0.92	0.83	0.85	0.78	0.81
RMSE	1.43	0.99	1.82	0.83	0.54
BIAS	1.00	-0.26	0.34	0.29	-0.02
SEPC	1.02	0.95	1.78	0.78	0.54
<u>Water (C_w, $\text{g}\cdot\text{cm}^{-2}$):</u>					
R ²	0.94	0.78	0.85	0.85	0.86
RMSE	0.00206	0.00149	0.00106	0.00060	0.00076
BIAS	0.00120	0.00039	-0.00060	0.00000	-0.00004
SEPC	0.00167	0.00144	0.00087	0.00060	0.00076
<u>Dry matter (C_m, $\text{g}\cdot\text{cm}^{-2}$):</u>					
R ²	0.78	0.77	0.89	0.92	0.83
RMSE	0.00031	0.00033	0.00049	0.00025	0.00020
BIAS	-0.00012	0.00002	-0.00001	-0.00001	-0.00002
SEPC	0.00029	0.00032	0.00049	0.00025	0.00020

Table 27. Leaf chlorophyll content ($\mu\text{g}\cdot\text{cm}^{-2}$) retrieved from PROSPECT inversions on the brownfield (abs: absent from the site). For each species, significant differences in mean chlorophyll content among sampling dates are denoted by different superscript letters (ANOVA and Tukey post-hoc test, $p < 0.05$).

Species	Season	Plots (n)	Spectral signatures (3n)	Leaf chlorophyll content				
				min	max	med	mean	sd
<i>B. davidii</i>	June 2017	18	54	20.4	76.2	60.3	58.9 ^a	12.0
	October 2017	18	54	10.9	84.2	39.8	41.9 ^b	16.1
	March 2018	6	18	26.8	76.4	52.0	50.6 ^{ab}	13.2
	April 2018	18	54	32.5	90.0	58.6	57.4 ^a	12.5
	June 2018	18	54	26.4	72.0	44.4	45.1 ^b	10.8
	October 2018	18	54	21.0	60.1	34.9	35.4 ^c	9.5
<i>R. fruticosus</i>	June 2017	23	69	41.9	72	58.4	57.2 ^a	8.1
	October 2017	23	69	49.1	80.2	58.4	60.4 ^a	8.1
	March 2018	19	57	14.4	62.8	44.3	42.5 ^b	10.4
	April 2018	23	69	19.4	42.1	28.9	29.8 ^c	5.1
	June 2018	23	69	37	56.8	44.5	46.0 ^b	5.2
	October 2018	23	69	43.6	61.4	53.4	53.0 ^a	4.4
<i>C. sanguinea</i>	June 2017	23	69	26.3	75.9	47.8	48.1 ^a	9.9
	October 2017	23	69	10.8	54.4	28.2	29.9 ^b	8.8
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	23	69	18.9	46.6	31.6	32.1 ^b	5.9
	June 2018	23	69	21	57.4	34.1	34.7 ^b	7.8
	October 2018	23	69	6.8	32.6	15.6	15.9 ^c	5.4
<i>Q. pubescens</i>	June 2017	19	57	22.9	41.1	29.6	31.7 ^a	5.5
	October 2017	19	57	9.7	26.9	19.6	19.3 ^b	3.7
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	19	57	9.3	27.1	14.8	14.8 ^c	2.8
	June 2018	19	57	12.2	37.4	24	23.8 ^b	6.2
	October 2018	19	57	18.9	32.7	23.6	24.1 ^b	2.9
<i>P. canadensis</i>	June 2017	20	60	28.2	53.2	38.7	39.7 ^a	6.4
	October 2017	20	60	14.6	36.3	24.9	24.9 ^b	5.2
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	20	60	19.3	33.5	24.7	25.2 ^b	3.7
	June 2018	20	60	19	43.7	27.3	28.7 ^b	6.6
	October 2018	20	60	16.2	30.4	22.9	23.1 ^b	4

Table 28. Leaf carotenoid content ($\mu\text{g}\cdot\text{cm}^{-2}$) retrieved from PROSPECT inversions on the brownfield (abs: absent from the site). For each species, significant differences in mean carotenoid content among sampling dates are denoted by different superscript letters (ANOVA and Tukey post-hoc test, $p < 0.05$).

Species	Season	Plots (n)	Spectral signatures (3n)	Leaf carotenoid content				
				min	max	med	mean	sd
<i>B. davidii</i>	June 2017	18	54	5.6	21.8	15.8	15.3 ^a	3.2
	October 2017	18	54	6.8	22.6	12.6	12.7 ^a	3.6
	March 2018	6	18	10.7	29.1	21.4	20.4 ^b	5.4
	April 2018	18	54	7.8	22.7	13.6	13.5 ^a	3.1
	June 2018	18	54	8.0	24.4	14.0	14.3 ^a	3.5
	October 2018	18	54	5.4	17.5	9.9	10.4 ^a	2.7
<i>R. fruticosus</i>	June 2017	23	69	8.8	19.3	14.0	13.7 ^{ab}	2.1
	October 2017	23	69	11.9	24.9	14.8	15.7 ^a	2.7
	March 2018	19	57	5.7	29.6	13.3	13.9 ^{ab}	4.6
	April 2018	23	69	4.7	10.5	7.0	7.2 ^c	1.3
	June 2018	23	69	9.6	15.4	12.6	12.7 ^b	1.6
	October 2018	23	69	11.7	15.6	13.5	13.6 ^b	0.9
<i>C. sanguinea</i>	June 2017	23	69	6.3	12.6	9.4	9.6 ^a	1.5
	October 2017	23	69	1.2	29.0	10.2	11.9 ^{ab}	6.3
	March 2018	abs.	abs.	abs	abs	abs	abs	abs
	April 2018	23	69	3.8	10.5	6.8	7.0 ^b	1.3
	June 2018	23	69	5.7	22.8	9.2	9.5 ^{ab}	3.2
	October 2018	23	69	1.2	27.0	7.2	8.9 ^{ab}	5.9
<i>Q. pubescens</i>	June 2017	19	57	4.8	13.5	7.6	7.9 ^a	1.8
	October 2017	19	57	3.0	8.7	5.5	5.5 ^b	1.4
	March 2018	abs.	abs.	abs	abs	abs	abs	abs
	April 2018	19	57	2.9	19.3	6.1	6.7 ^{ab}	3.1
	June 2018	19	57	3.3	10.2	7.0	6.7 ^{ab}	1.6
	October 2018	19	57	5.0	11.8	6.5	6.9 ^{ab}	1.2
<i>P. canadensis</i>	June 2017	20	60	6.9	11.3	8.8	8.7 ^a	1.1
	October 2017	20	60	4.7	8.4	6.8	6.7 ^b	0.9
	March 2018	abs.	abs.	abs	abs	abs	abs	abs
	April 2018	20	60	5.0	8.5	6.7	6.8 ^b	1.0
	June 2018	20	60	4.7	14.8	7.3	7.6 ^{ab}	1.6
	October 2018	20	60	4.2	11.8	6.5	6.6 ^b	1.3

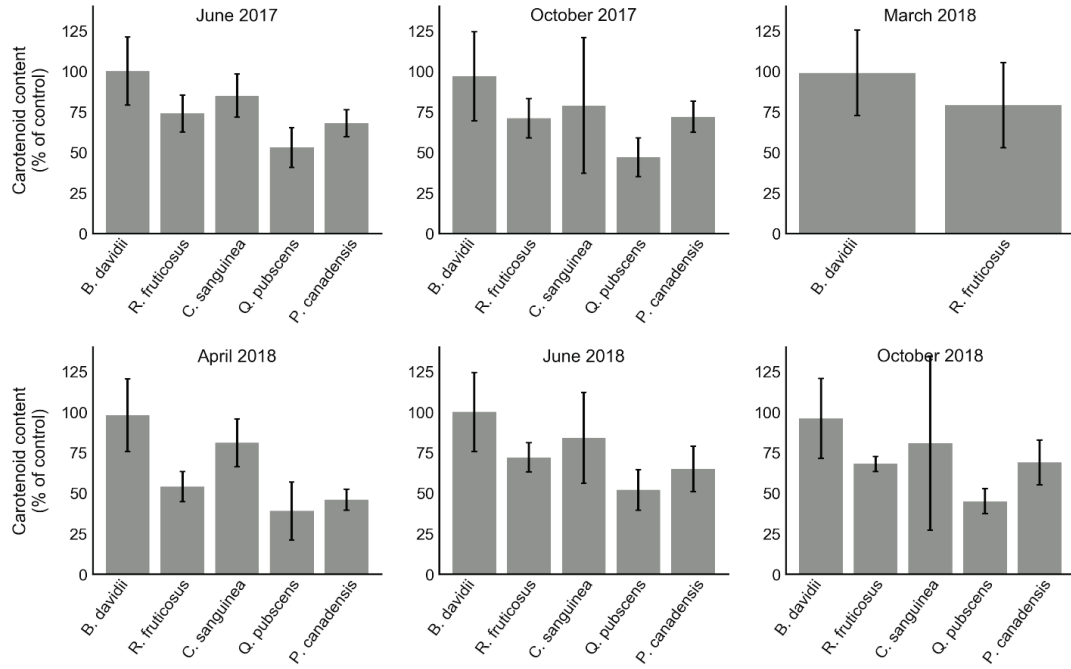


Figure 70. Temporal evolutions of the relative carotenoid contents of leaves retrieved from PROSPECT inversions on the brownfield, for the five species.

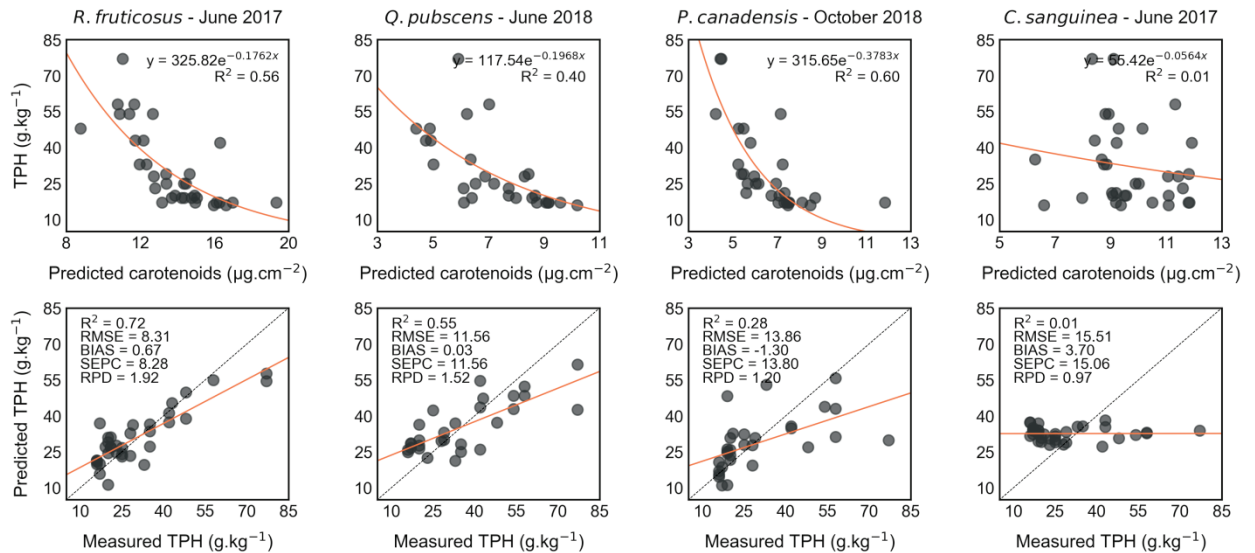


Figure 71. Relationship between predicted leaf carotenoid content and Total Petroleum Hydrocarbons (TPH) in the soil on the train datasets (top figures) and comparisons between predicted and measured TPH on the test datasets (bottom figures).

Table 29. Assessment of Total Petroleum Hydrocarbons (TPH) in the soil predicted from leaf carotenoid content on the train and test datasets. RMSE, BIAS and SEPC are given in g.kg⁻¹ (abs: absent from the site).

Species	Season	train		test				
		R ²	RMSE	R ²	RMSE	BIAS	SEPC	RPD
<i>B. davidii</i>	June 2017	0.02	16.27	0.05	15.94	1.94	15.82	1.02
	October 2017	0.01	16.52	0.03	16.93	-4.27	16.38	0.96
	March 2018	0.03	14.31	0.05	12.86	-2.60	12.59	1.27
	April 2018	0.02	21.44	0.01	20.07	-4.88	19.47	0.81
	June 2018	0.04	19.17	0.04	19.03	-0.58	19.03	0.86
	October 2018	0.06	17.63	0.08	17.09	6.10	15.96	0.95
<i>R. fruticosus</i>	June 2017	0.56	10.54	0.72	8.31	0.67	8.28	1.79
	October 2017	0.01	17.02	0.04	14.73	2.35	14.54	1.01
	March 2018	0.03	15.44	0.11	14.63	-0.77	14.61	1.02
	April 2018	0.02	17.80	0.21	14.21	5.55	13.08	1.05
	June 2018	0.58	11.07	0.45	10.28	1.13	10.21	1.45
	October 2018	0.31	12.72	0.15	14.95	-0.21	14.95	1.00
<i>C. sanguinea</i>	June 2017	0.03	16.42	0.10	15.51	3.70	15.06	1.03
	October 2017	0.04	20.89	0.14	20.02	-7.36	18.62	0.80
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	0.01	17.25	0.34	15.37	7.26	13.55	1.04
	June 2018	0.01	16.63	0.02	15.04	-1.24	14.99	1.06
	October 2018	0.08	18.01	0.40	16.58	8.87	14.00	0.97
<i>Q. pubescens</i>	June 2017	0.24	13.87	0.34	13.84	-2.86	13.54	1.19
	October 2017	0.15	16.45	0.24	13.29	-2.09	13.12	1.24
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	0.01	15.75	0.01	18.20	2.13	18.08	0.91
	June 2018	0.40	11.81	0.55	11.56	0.03	11.56	1.43
	October 2018	0.04	17.26	0.15	14.10	2.05	13.95	1.17
<i>P. canadensis</i>	June 2017	0.16	16.40	0.44	11.36	1.74	11.22	1.57
	October 2017	0.32	14.29	0.19	14.07	1.41	14.00	1.27
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	0.06	17.77	0.13	14.86	4.74	14.09	1.20
	June 2018	0.35	13.59	0.38	12.85	2.40	12.62	1.39
	October 2018	0.60	10.65	0.28	13.86	-1.30	13.80	1.28

Table 30. Comparison of soil Total Petroleum Hydrocarbons (TPH) predicted on the test sets using the proposed approach (Chlorophyll content PROSPECT), vegetation indices (MTCI, mSR705) and Partial Least Square Regression (PLSR) on transformed spectral signatures. RMSE is given in g.kg⁻¹. The tested vegetation indices and reflectance transformations can be found in ²⁹² and ²⁷¹. Those leading to the best predictions are presented. Important results are marked bold.

Species	Season	Chlorophyll content PROSPECT		Vegetation indices MTCI		Vegetation indices mSR705		PLSR first derivative		PLSR <i>continuum</i> removal	
		RMSE	RPD	RMSE	RPD	RMSE	RPD	RMS E	RPD	RMSE	RPD
<i>B. davidii</i>	June 2017	16.24	1.00	19.46	0.84	17.39	0.94	18.57	0.88	17.77	0.92
	October 2017	18.14	0.90	18.01	0.91	17.80	0.92	19.64	0.83	18.89	0.86
	March 2018	16.47	0.99	17.17	0.95	16.21	1.01	16.70	0.98	16.91	0.96
	April 2018	20.01	0.81	21.43	0.76	20.59	0.79	20.12	0.81	21.62	0.75
	June 2018	15.98	1.02	15.24	1.07	16.01	1.02	16.31	1.00	15.31	1.06
	October 2018	17.76	0.92	17.71	0.92	16.44	0.99	18.12	0.90	17.14	0.95
<i>R. fruticosus</i>	June 2017	3.63	4.10	6.17	2.41	6.98	2.13	5.86	2.54	6.00	2.48
	October 2017	7.64	1.95	10.21	1.46	11.17	1.33	9.29	1.60	10.30	1.45
	March 2018	17.05	0.87	17.44	0.85	18.09	0.82	18.03	0.83	15.57	0.96
	April 2018	18.31	0.81	19.02	0.78	18.70	0.80	19.61	0.76	20.82	0.72
	June 2018	6.64	2.24	9.31	1.60	9.22	1.62	8.99	1.66	9.24	1.61
	October 2018	8.28	1.80	12.66	1.18	11.88	1.25	11.12	1.34	11.40	1.31
<i>C. sanguinea</i>	June 2017	15.75	1.02	16.67	0.96	15.53	1.03	15.90	1.01	17.01	0.94
	October 2017	17.20	0.93	17.04	0.94	16.97	0.94	18.11	0.88	17.40	0.92
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	17.33	0.92	16.84	0.95	18.20	0.88	16.86	0.95	17.88	0.89
	June 2018	15.98	1.00	16.61	0.96	17.76	0.90	16.00	1.00	16.43	0.97
	October 2018	17.45	0.92	18.49	0.87	17.81	0.90	17.27	0.93	17.92	0.89
<i>Q. pubescens</i>	June 2017	5.70	2.89	8.97	1.84	9.11	1.81	8.68	1.90	8.85	1.86
	October 2017	5.82	2.83	9.24	1.79	9.37	1.76	9.05	1.82	9.05	1.82
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	17.94	0.92	17.31	0.95	19.21	0.86	19.17	0.86	17.59	0.94
	June 2018	3.20	5.17	6.93	2.38	7.49	2.20	6.74	2.45	6.41	2.57
	October 2018	5.88	2.80	9.00	1.83	9.20	1.79	8.88	1.86	8.77	1.88
<i>P. canadensis</i>	June 2017	6.85	2.60	10.82	1.65	10.27	1.73	9.41	1.89	10.10	1.76
	October 2017	5.86	3.04	11.08	1.61	10.45	1.70	10.58	1.68	9.24	1.93
	March 2018	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.	abs.
	April 2018	16.26	1.09	15.96	1.12	15.92	1.12	17.12	1.04	16.99	1.05
	June 2018	7.83	2.27	11.69	1.52	11.13	1.60	11.09	1.61	11.47	1.55
	October 2018	6.62	2.69	10.72	1.66	10.76	1.65	10.19	1.75	10.27	1.73

Fifth article: Toward quantifying oil contamination in vegetated areas using high spatial and spectral resolution imagery

Table 31. Vegetation indices used in the method of detection. For each index, the related biochemical parameters are specified. These parameters have been identified in a previous study carried out on *R. fruticosus*, under controlled conditions³²⁴. (Chl: Total chlorophylls, Lut: Lutein, B-car: β -carotene, Ant: Antheraxanthin, Zea: Zeaxanthin, Vio: Violaxanthin, LWC: Leaf Water Content.)

Index	Formula	Reference	Related parameter
Chlorophyll/Carotenoids Index	$CCI = D_{720}/D_{700}$	297	Chl, B-car
Carter Index 2	$CTR2 = R_{695}/R_{760}$	298	Chl, Lut
Gitelson & Merzlyak Index 1	$M1 = R_{750}/R_{550}$	299	Chl, B-car
Gitelson & Merzlyak Index 2	$M2 = R_{750}/R_{700}$	299	Chl, B-car
modified Simple Ratio 705 nm	$mSR705 = (R_{750} - R_{445})/(R_{750} + R_{445})$	78	Chl, B-car
MERIS Terrestrial Chlorophyll Index	$MTCI = (R_{754} - R_{709})/(R_{709} + R_{681})$	300	Chl, B-car
Normalized Difference 705 nm	$ND705 \text{ or } NDVI705 = (R_{750} - R_{705})/(R_{750} + R_{705})$	78	Chl, B-car
Photochemical Reflectance Index 2	$PRI2 = (R_{531} - R_{570})/(R_{531} + R_{570})$	101	Chl, B-car
Photochemical Reflectance Index 3	$PRI3 = (R_{570} - R_{539})/(R_{570} + R_{539})$	101	Lut, B-car
Structure Intensive Pigment Index 2	$SIP2 = (R_{800} - R_{505})/(R_{800} + R_{690})$	240	Lut, Ant
Simple Ratio 705 nm	$SR705 = R_{750}/R_{705}$	78	Chl, B-car
Vogelmann Index 2	$VO2 = (R_{734} - R_{747})/(R_{715} + R_{726})$	306	Lut
Vogelmann Index 3	$VO3 = (R_{734} - R_{747})/(R_{715} + R_{720})$	306	Lut
Disease Water Stress Index	$DWSI = (R_{800} - R_{550})/(R_{1660} + R_{680})$	207	LWC

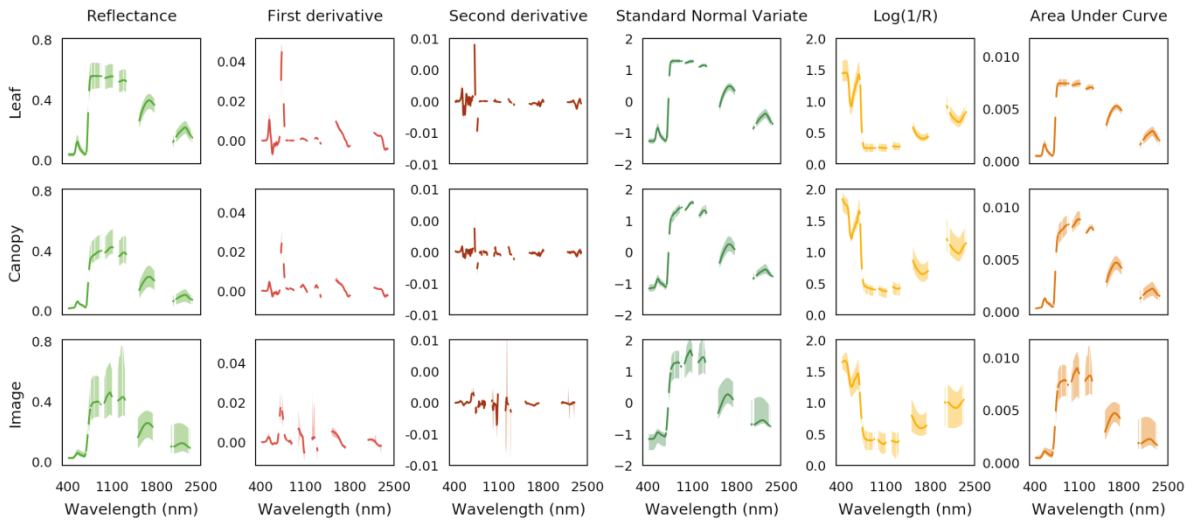


Figure 72. Original and transformed spectral signatures (in columns) acquired in the brownfield. First, second and third rows show the results obtained at leaf and canopy scales and on the airborne image, respectively. The bold curve represents the median spectral signature and the colored area, the [min:max] interval.

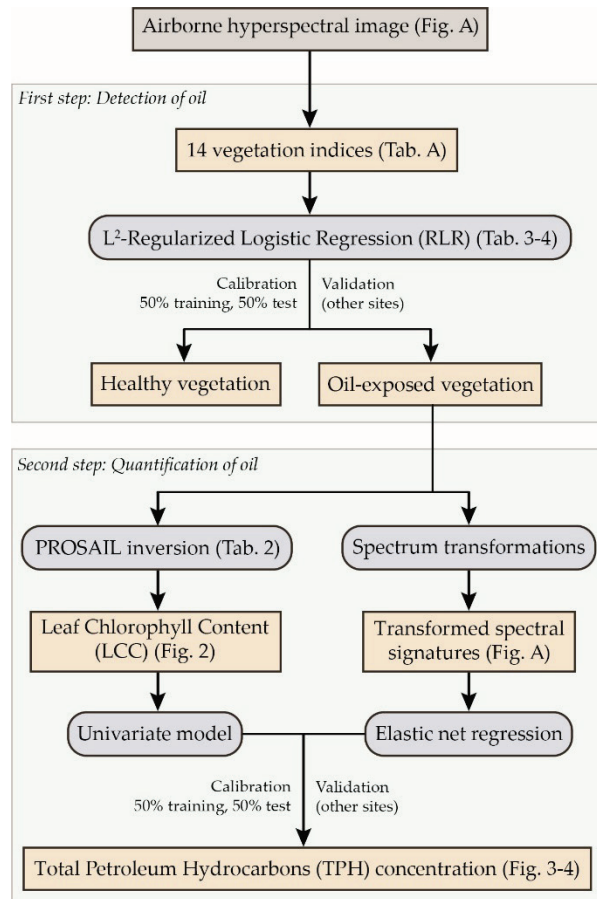


Figure 73. Flowchart of the approach proposed in this study. The method of oil detection was developed and tested at leaf and canopy scales in ³²⁴, so it was only evaluated on the airborne image in this study. The two methods of Total Petroleum Hydrocarbons (TPH) quantification were tested in the field at leaf and canopy scales and on the airborne image. All the methods were first calibrated on study sites and validated on additional ones (see Table 16 for details about the sites).

PUBLICATIONS AND COMMUNICATIONS

PUBLICATIONS (PEER-REVIEWED ARTICLES)

- Lassalle, G.; Elger, A.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Fabre, S. Toward quantifying oil contamination in vegetated areas using high spatial and spectral resolution imagery. *Remote Sens.* **2019**, 11, 2241. <https://doi.org/10.3390/rs11192241>
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression. *Ecotox. Environ. Safe.* **2019**, 184, 109654. <https://doi.org/10.1016/j.ecoenv.2019.109654>.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Application of PROSPECT for Estimating Total Petroleum Hydrocarbons in Contaminated Soils from Leaf Optical Properties. *J. Hazard. Mater.* **2019**, 377, 409–417. <https://doi.org/10.1016/j.jhazmat.2019.05.093>.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Erudel, T.; Buffan-Dubau, E.; Dubucq, D.; Elger, A. Detection and Discrimination of Various Oil-Contaminated Soils Using Vegetation Reflectance. *Sci. Total Environ.* **2019**, 655, 1113–1124. <https://doi.org/10.1016/j.scitotenv.2018.11.314>.
- Lassalle, G.; Credoza, A.; Hédacq, R.; Fabre, S.; Dubucq, D.; Elger, A. Assessing Soil Contamination Due to Oil and Gas Production Using Vegetation Hyperspectral Reflectance. *Environ. Sci. Technol.* **2018**, 52 (4), 1756–1764. <https://doi.org/10.1021/acs.est.7b04618>.

COMMUNICATIONS

- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Utilisation de données hyperspectrales pour la détection et la quantification des hydrocarbures pétroliers en régions végétalisées. SFPT-GH **2019**, Toulouse, France.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Dubucq, D.; Elger, A. Detection and quantification of total petroleum hydrocarbons in soils using vegetation optical properties. IEEE IGARSS **2019**, Yokohama, Japan.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Dubucq, D.; Elger, A. Detection and quantification of oil contamination in vegetated areas using hyperspectral remote sensing. SETAC **2018**, Roma, Italy.
- Lassalle, G.; Credoza, A.; Hédacq, R.; Fabre, S.; Dubucq, D.; Elger, A. Hyperspectral signature analysis of three plant species to long-term hydrocarbon and heavy metal exposure. Proceedings of the SPIE Remote Sensing **2017**, Warsaw, Poland. <https://doi.org/10.1117/12.2277709>

REFERENCES

- (1) Marrero, G. A. Greenhouse Gases Emissions, Growth and the Energy Mix in Europe. *Energy Econ.* **2010**, *32* (6), 1356–1363. <https://doi.org/10.1016/j.eneco.2010.09.007>.
- (2) Suganthi, L.; Samuel, A. A. Energy Models for Demand Forecasting - A Review. *Renew. Sustain. Energy Rev.* **2012**, *16* (2), 1223–1240. <https://doi.org/10.1016/j.rser.2011.08.014>.
- (3) Jones, G. A.; Warner, K. J. The 21st Century Population-Energy-Climate Nexus. *Energy Policy* **2016**, *93*, 206–212. <https://doi.org/10.1016/j.enpol.2016.02.044>.
- (4) Miller, R. G.; Sorrell, S. R.; Lane, H.; Kt, S. The Future of Oil Supply. *Philos. Trans. R. Soc. A* **2014**, *372*, 20130179. <https://doi.org/10.1098/rsta.2013.0179>.
- (5) OPEC. The Organization of the Petroleum Exporting Countries Annual Report 2017. **2018**.
- (6) Zou, C.; Zhao, Q.; Zhang, G.; Xiong, B. Energy Revolution: From a Fossil Energy Era to a New Energy Era. *Nat. Gas Ind. B* **2016**, *3*, 1–11. <https://doi.org/10.1016/j.ngib.2016.02.001>.
- (7) Bentley, R. W. Oil Forecasts, Past and Present. *Energy Explor. Exploit.* **2002**, *20* (6), 481–491. <https://doi.org/10.1260/014459802321615108>.
- (8) Jackson, P. M.; Smith, L. K. Exploring the Undulating Plateau: The Future of Global Oil Supply. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2013**, *372* (20120491). <https://doi.org/10.1098/rsta.2012.0491>.
- (9) Sorrell, S.; Speirs, J.; Bentley, R.; Brandt, A.; Miller, R. Global Oil Depletion: A Review of the Evidence. *Energy Policy* **2010**, *38* (9), 5290–5295. <https://doi.org/10.1016/j.enpol.2010.04.046>.
- (10) Sorrell, S.; Miller, R.; Bentley, R.; Speirs, J. Oil Futures: A Comparison of Global Supply Forecasts. *Energy Policy* **2010**, *38*, 4990–5003. <https://doi.org/10.1016/j.enpol.2010.04.020>.
- (11) He, K.; Huo, H.; Zhang, Q.; He, D.; An, F.; Wang, M.; Walsh, M. P. Oil Consumption and CO₂ Emissions in China's Road Transport: Current Status, Future Trends, and Policy Implications. *Energy Policy* **2005**, *33* (12), 1499–1507. <https://doi.org/10.1016/j.enpol.2004.01.007>.
- (12) Lean, H. H.; Smyth, R. Long Memory in US Disaggregated Petroleum Consumption: Evidence from Univariate and Multivariate LM Tests for Fractional Integration. *Energy Policy* **2009**, *37* (8), 3205–3211. <https://doi.org/10.1016/j.enpol.2009.04.017>.
- (13) Sorrell, S.; Speirs, J.; Bentley, R.; Miller, R.; Thompson, E. Shaping the Global Oil Peak: A Review of the Evidence on Field Sizes, Reserve Growth, Decline Rates and Depletion Rates. *Energy* **2012**, *37* (1), 709–724. <https://doi.org/10.1016/j.energy.2011.10.010>.
- (14) U.S. Energy Information Administration (EIA). *Definitions of Petroleum Products and Other Terms*, 2010.
- (15) Solé, J.; García-Olivares, A.; Turiel, A.; Ballabrera-Poy, J. Renewable Transitions and the Net Energy from Oil Liquids: A Scenarios Study. *Renew. Energy* **2018**, *116*, 258–271. <https://doi.org/10.1016/j.renene.2017.09.035>.
- (16) Hayashi, S. H. D.; Ligeró, E. L.; Schiozer, D. J. Risk Mitigation in Petroleum Field Development by Modular Implantation. *J. Pet. Sci. Eng.* **2010**, *75* (1–2), 105–113. <https://doi.org/10.1016/j.petrol.2010.10.013>.
- (17) Kreuzer, O. K.; Etheridge, M. A.; Guj, P.; McMahon, M. E.; Holden, D. J. Linking Mineral

- Deposit Models to Quantitative Risk Analysis and Decision-Making in Exploration. *Econ. Geol.* **2008**, *103* (4), 829–850. <https://doi.org/10.2113/gsecongeo.103.4.829>.
- (18) Suslick, S. B.; Schiozer, D. J. Risk Analysis Applied to Petroleum Exploration and Production: An Overview. *J. Pet. Sci. Eng.* **2004**, *44* (1–2), 1–9. <https://doi.org/10.1016/j.petrol.2004.02.001>.
- (19) Tedesco, S. A. *Surface Geochemistry in Petroleum Exploration*; Springer Science Business Media, Ed.; Springer US: Boston, MA, 1995. <https://doi.org/10.1007/978-1-4615-2660-5>.
- (20) Noomen, M. F.; van der Werff, H. M. A.; van der Meer, F. D. Spectral and Spatial Indicators of Botanical Changes Caused by Long-Term Hydrocarbon Seepage. *Ecol. Inform.* **2012**, *8* (February), 55–64. <https://doi.org/10.1016/j.ecoinf.2012.01.001>.
- (21) Tangestani, M. H.; Validabadi, K. Mineralogy and Geochemistry of Alteration Induced by Hydrocarbon Seepage in an Evaporite Formation; a Case Study from the Zagros Fold Belt, SW Iran. *Appl. Geochemistry* **2014**, *41*, 189–195. <https://doi.org/10.1016/j.apgeochem.2013.12.015>.
- (22) Shi, P.; Fu, B.; Ninomiya, Y.; Sun, J.; Li, Y. Multispectral Remote Sensing Mapping for Hydrocarbon Seepage-Induced Lithologic Anomalies in the Kuqa Foreland Basin, South Tian Shan. *J. Asian Earth Sci.* **2012**, *46*, 70–77. <https://doi.org/10.1016/j.jseaes.2011.10.019>.
- (23) Asadzadeh, S.; de Souza Filho, C. R. Spectral Remote Sensing for Onshore Seepage Characterization: A Critical Overview. *Earth-Science Rev.* **2017**, *168*, 48–72. <https://doi.org/10.1016/j.earscirev.2017.03.004>.
- (24) Salati, S.; Van Ruitenbeek, F. J. A.; De Smeth, J. B.; Van Der Meer, F. D. Spectral and Geochemical Characterization of Onshore Hydrocarbon Seep-Induced Alteration in the Dezful Embayment, Southwest Iran. *Am. Assoc. Pet. Geol. Bull.* **2014**, *98* (9), 1837–1857. <https://doi.org/10.1306/04051413105>.
- (25) Sheriff, R. E.; Geldart, L. P. *Exploration Seismology*, 2nd ed.; Cambridge University Press, 1995. <https://doi.org/10.1017/CBO9781139168359>.
- (26) Chaudhuri, U. R. *Fundamentals of Petroleum and Petrochemical Engineering*; Taylor & Francis Group, Ed.; CRC Press, 2016. <https://doi.org/10.1201/b10486>.
- (27) Ite, A. E.; Ibok, U. J.; Ite, M. U.; Petters, S. W. Petroleum Exploration and Production: Past and Present Environmental Issues in the Nigeria's Niger Delta. *Am. J. Environ. Prot.* **2013**, *1* (4), 78–90. <https://doi.org/10.12691/env-1-4-2>.
- (28) Chang, J. I.; Lin, C. C. A Study of Storage Tank Accidents. *J. Loss Prev. Process Ind.* **2006**, *19*, 51–59. <https://doi.org/10.1016/j.jlp.2005.05.015>.
- (29) Credoza, A.; Hédacq, R.; Barreau, C.; Dubucq, D. Experimental Study of Hyperspectral Responses of Plants Grown on Mud Pit Soil. In *Earth Resources and Environmental Remote Sensing/GIS Applications VII*; 2016; Vol. 10005, p 100051E. <https://doi.org/10.1117/12.2239606>.
- (30) Gogoi, B. K.; Dutta, N. N.; Goswami, P.; Krishna Mohan, T. R. A Case Study of Bioremediation of Petroleum-Hydrocarbon Contaminated Soil at a Crude Oil Spill Site. *Adv. Environ. Res.* **2003**, *7* (4), 767–782. [https://doi.org/10.1016/S1093-0191\(02\)00029-1](https://doi.org/10.1016/S1093-0191(02)00029-1).
- (31) Cunha, S. B. da. A Review of Quantitative Risk Assessment of Onshore Pipelines. *J. Loss Prev. Process Ind.* **2016**, *44*, 282–298. <https://doi.org/10.1016/j.jlp.2016.09.016>.
- (32) Shadizadeh, S. R.; Zoveidavianpoor, M. A Drilling Reserve Mud Pit Assessment in Iran:

- Environmental Impacts and Awareness. *Pet. Sci. Technol.* **2010**, *28* (14), 1513–1526. <https://doi.org/10.1080/10916460903117545>.
- (33) Ahmadun, F.-R.; Pendashteh, A.; Abdullah, L. C.; Awang Biak, D. R.; Madaeni, S. S.; Abidin, Z. Z. Review of Technologies for Oil and Gas Produced Water Treatment. *J. Hazard. Mater.* **2009**, *170*, 530–551. <https://doi.org/10.1016/j.jhazmat.2009.05.044>.
- (34) Hu, G.; Li, J.; Zeng, G. Recent Development in the Treatment of Oily Sludge from Petroleum Industry: A Review. *J. Hazard. Mater.* **2013**, *261*, 470–490. <https://doi.org/10.1016/j.jhazmat.2013.07.069>.
- (35) van der Werff, H.; van der Meijde, M.; Jansma, F.; van der Meer, F.; Groothuis, G. J. A Spatial-Spectral Approach for Visualization of Vegetation Stress Resulting from Pipeline Leakage. *Sensors* **2008**, *8*, 3733–3743. <https://doi.org/10.3390/s8063733>.
- (36) Barraza, F.; Maurice, L.; Uzu, G.; Becerra, S.; López, F.; Ochoa-Herrera, V.; Ruales, J.; Schreck, E. Distribution, Contents and Health Risk Assessment of Metal(Loid)s in Small-Scale Farms in the Ecuadorian Amazon: An Insight into Impacts of Oil Activities. *Sci. Total Environ.* **2018**, *622–623*, 106–120. <https://doi.org/10.1016/j.scitotenv.2017.11.246>.
- (37) Bi, X.; Wang, B.; Lu, Q. Fragmentation Effects of Oil Wells and Roads on the Yellow River Delta, North China. *Ocean Coast. Manag.* **2011**, *54* (3), 256–264. <https://doi.org/10.1016/j.ocecoaman.2010.12.005>.
- (38) Duke, N. C. Oil Spill Impacts on Mangroves: Recommendations for Operational Planning and Action Based on a Global Review. *Mar. Pollut. Bull.* **2016**, *109* (2), 700–715. <https://doi.org/10.1016/j.marpolbul.2016.06.082>.
- (39) Finer, M.; Jenkins, C. N.; Pimm, S. L.; Keane, B.; Ross, C. Oil and Gas Projects in the Western Amazon: Threats to Wilderness, Biodiversity, and Indigenous Peoples. *PLoS One* **2008**, *3* (8). <https://doi.org/10.1371/journal.pone.0002932>.
- (40) Ivshina, I. B.; Kuyukina, M. S.; Krivoruchko, A. V.; Elkin, A. A.; Makarov, S. O.; Cunningham, C. J.; Peshkur, T. A.; Atlas, R. M.; Philp, J. C. Oil Spill Problems and Sustainable Response Strategies through New Technologies. *Environ. Sci. Process. Impacts* **2015**, *17* (7), 1201–1219. <https://doi.org/10.1039/c5em00070j>.
- (41) Datta, S.; Sarkar, S. A Review on Different Pipeline Fault Detection Methods. *J. Loss Prev. Process Ind.* **2016**, *41*, 97–106. <https://doi.org/10.1016/j.jlp.2016.03.010>.
- (42) Kishawy, H. A.; Gabbar, H. A. Review of Pipeline Integrity Management Practices. *Int. J. Press. Vessel. Pip.* **2010**, *87* (7), 373–380. <https://doi.org/10.1016/j.ijpvp.2010.04.003>.
- (43) Shukla, A.; Karki, H. Application of Robotics in Onshore Oil and Gas Industry—A Review Part I. *Rob. Auton. Syst.* **2016**, *75*, 490–507. <https://doi.org/10.1016/j.robot.2015.09.012>.
- (44) Shukla, A.; Karki, H. Application of Robotics in Offshore Oil and Gas Industry— A Review Part II. *Rob. Auton. Syst.* **2016**, *75*, 508–524. <https://doi.org/10.1016/j.robot.2015.09.013>.
- (45) Al-Sayegh, A. Enhanced Oil Recovery Using Biotransformation Technique on Heavy Crude Oil. *Int. J. GEOMATE* **2017**, *13* (36). <https://doi.org/10.21660/2017.36.2842>.
- (46) Correa Pabón, R. E.; de Souza Filho, C. R.; Oliveira, W. J. de. Reflectance and Imaging Spectroscopy Applied to Detection of Petroleum Hydrocarbon Pollution in Bare Soils. *Sci. Total Environ.* **2019**, *649*, 1224–1236. <https://doi.org/10.1016/j.scitotenv.2018.08.231>.
- (47) Gudmestad, O. T. Sustainable Oil and Gas Production in the 21st Century with Emphasis on Offshore Fields. *WIT Trans. Ecol. Environ.* **2014**, *190*, 777–788.

<https://doi.org/10.2495/EQ140722>.

- (48) van der Lelie, D.; Schwitzguébel, J.-P.; Glass, D. J.; Vangronsveld, J.; Baker, A. Phytoremediation: European and American Trends Successes, Obstacles and Needs. *J. Soils Sediments* **2008**, *2* (2), 91–99. <https://doi.org/10.1007/bf02987877>.
- (49) Necci, A.; Girgin, S.; Krausmann, E. *Understanding Natech Risk Due to Storms; Analysis, Lessons Learned and Recommendations*; 2018. <https://doi.org/10.2760/21366>.
- (50) Rehman, K.; Imran, A.; Amin, I.; Afzal, M. Inoculation with Bacteria in Floating Treatment Wetlands Positively Modulates the Phytoremediation of Oil Field Wastewater. *J. Hazard. Mater.* **2018**, *349* (February), 242–251. <https://doi.org/10.1016/j.jhazmat.2018.02.013>.
- (51) Jin, Y.-Q.; Xu, F. *Polarimetric Scattering and SAR Information Retrieval*; John Wiley & Sons (Asia) Pte Ltd: Singapore, 2013. <https://doi.org/10.1002/9781118188149>.
- (52) Prasad, S.; Bruce, L. M.; Chanussot, J. *Optical Remote Sensing: Advances in Signal Processing and Exploitation Techniques*; Springer Singapore: Singapore, 2019; Vol. 526. <https://doi.org/10.1007/978-981-13-2553-3>.
- (53) Alpers, W.; Holt, B.; Zeng, K. Oil Spill Detection by Imaging Radars: Challenges and Pitfalls. *Remote Sens. Environ.* **2017**, *201*, 133–147. <https://doi.org/10.1016/j.rse.2017.09.002>.
- (54) Brekke, C.; Solberg, A. H. S. Oil Spill Detection by Satellite Remote Sensing. *Remote Sens. Environ.* **2005**, *95* (1), 1–13. <https://doi.org/10.1016/j.rse.2004.11.015>.
- (55) Kühn, F.; Oppermann, K.; Hörig, B. Hydrocarbon Index - An Algorithm for Hyperspectral Detection of Hydrocarbons. *Int. J. Remote Sens.* **2004**, *25* (12), 2467–2473. <https://doi.org/10.1080/01431160310001642287>.
- (56) Deutsch, M.; Strong, A. E.; Estes, J. E. Use Of Landsat Data For The Detection Of Marine Oil Slicks. In *Offshore Technology Conference*, 1977. <https://doi.org/10.4043/2763-ms>.
- (57) Halbouty, M. T. Geologic Significance of Landsat Data for 15 Giant Oil and Gas Fields. *Am. Assoc. Pet. Geol. Bull.* **1980**, *64* (1), 8–36.
- (58) Bioucas-Dias, J. M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5* (2), 354–379. <https://doi.org/10.1109/JSTARS.2012.2194696>.
- (59) 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.; de Smeth, J. B.; Woldai, T. Multi- and Hyperspectral Geologic Remote Sensing: A Review. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *14* (1), 112–128. <https://doi.org/10.1016/j.jag.2011.08.002>.
- (60) Milton, E. J.; Schaepman, M. E.; Anderson, K.; Kneubühler, M.; Fox, N. Progress in Field Spectroscopy. *Remote Sens. Environ.* **2009**, *113*, S92–S109. <https://doi.org/10.1016/j.rse.2007.08.001>.
- (61) Scafutto, R. D. M.; de Souza Filho, C. R. Quantitative Characterization of Crude Oils and Fuels in Mineral Substrates Using Reflectance Spectroscopy: Implications for Remote Sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 221–242. <https://doi.org/10.1016/j.jag.2016.03.017>.
- (62) Correa Pabón, R. E.; de Souza Filho, C. R. Crude Oil Spectral Signatures and Empirical Models to Derive API Gravity. *Fuel* **2019**, *237* (May 2018), 1119–1131.

- <https://doi.org/10.1016/j.fuel.2018.09.098>.
- (63) Correa Pabón, R. E.; de Souza Filho, C. R. Spectroscopic Characterization of Red Latosols Contaminated by Petroleum-Hydrocarbon and Empirical Model to Estimate Pollutant Content and Type. *Remote Sens. Environ.* **2016**, *175*, 323–336. <https://doi.org/10.1016/j.rse.2016.01.005>.
- (64) Scafutto, R. D. P. M.; de Souza Filho, C. R.; de Oliveira, W. J. Hyperspectral Remote Sensing Detection of Petroleum Hydrocarbons in Mixtures with Mineral Substrates: Implications for Onshore Exploration and Monitoring. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 146–157. <https://doi.org/10.1016/j.isprsjprs.2017.03.009>.
- (65) Slonecker, T.; Fisher, G. B.; Aiello, D. P.; Haack, B. Visible and Infrared Remote Imaging of Hazardous Waste: A Review. *Remote Sens.* **2010**, *2* (11), 2474–2508. <https://doi.org/10.3390/rs2112474>.
- (66) Lammoglia, T.; de Souza Filho, C. R. Spectroscopic Characterization of Oils Yielded from Brazilian Offshore Basins: Potential Applications of Remote Sensing. *Remote Sens. Environ.* **2011**, *115* (10), 2525–2535. <https://doi.org/10.1016/j.rse.2011.04.038>.
- (67) Hese, S.; Schmullius, C. High Spatial Resolution Image Object Classification for Terrestrial Oil Spill Contamination Mapping in West Siberia. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11* (2), 130–141. <https://doi.org/10.1016/j.jag.2008.12.002>.
- (68) Arellano, P.; Tansey, K.; Balzter, H.; Boyd, D. S. Detecting the Effects of Hydrocarbon Pollution in the Amazon Forest Using Hyperspectral Satellite Images. *Environ. Pollut.* **2015**, *205*, 225–239. <https://doi.org/10.1016/j.envpol.2015.05.041>.
- (69) Asner, G. P. Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sens. Environ.* **1997**, *64* (3), 234–253.
- (70) Jacquemoud, S.; Baret, F. PROSPECT: A Model of Leaf Optical Properties Spectra. *Remote Sens. Environ.* **1990**, *34*, 75–91. [https://doi.org/10.1016/0034-4257\(90\)90100-Z](https://doi.org/10.1016/0034-4257(90)90100-Z).
- (71) Slaton, M. R.; Hunt, E. R.; Smith, W. K. Estimating Near-Infrared Leaf Reflectance from Leaf Structural Characteristics. *Am. J. Bot.* **2001**, *88* (2), 278–284. <https://doi.org/10.2307/2657019>.
- (72) Calderón, R.; Navas-Cortés, J. A.; Lucena, C.; Zarco-Tejada, P. J. High-Resolution Airborne Hyperspectral and Thermal Imagery for Early Detection of Verticillium Wilt of Olive Using Fluorescence, Temperature and Narrow-Band Spectral Indices. *Remote Sens. Environ.* **2013**, *139*, 231–245. <https://doi.org/10.1016/j.rse.2013.07.031>.
- (73) Jacquemoud, S.; Verhoef, W.; Baret, F.; Bacour, C.; Zarco-Tejada, P. J.; Asner, G. P.; François, C.; Ustin, S. L. PROSPECT+SAIL Models: A Review of Use for Vegetation Characterization. *Remote Sens. Environ.* **2009**, *113*, S56–S66. <https://doi.org/10.1016/j.rse.2008.01.026>.
- (74) Rumpf, T.; Mahlein, A.-K.; Steiner, U.; Oerke, E.-C.; Dehne, H.-W.; Plümer, L. Early Detection and Classification of Plant Diseases with Support Vector Machines Based on Hyperspectral Reflectance. *Comput. Electron. Agric.* **2010**, *74* (1), 91–99. <https://doi.org/10.1016/j.compag.2010.06.009>.
- (75) Smith, K. L.; Steven, M. D.; Colls, J. J. Plant Spectral Responses to Gas Leaks and Other Stresses. *Int. J. Remote Sens.* **2005**, *26* (18), 4067–4081. <https://doi.org/10.1080/01431160500165625>.

- (76) Zinnert, J. C.; Via, S. M.; Young, D. R. Distinguishing Natural from Anthropogenic Stress in Plants: Physiology, Fluorescence and Hyperspectral Reflectance. *Plant Soil* **2013**, *366* (1–2), 133–141. <https://doi.org/10.1007/s11104-012-1414-1>.
- (77) Sanches, I. D.; de Souza Filho, C. R.; Magalhães, L. A.; Quitério, G. C. M.; Alves, M. N.; Oliveira, W. J. Assessing the Impact of Hydrocarbon Leakages on Vegetation Using Reflectance Spectroscopy. *ISPRS J. Photogramm. Remote Sens.* **2013**, *78*, 85–101. <https://doi.org/10.1016/j.isprsjprs.2013.01.007>.
- (78) Sims, D. A.; Gamon, J. A. Relationships between Leaf Pigment Content and Spectral Reflectance across a Wide Range of Species, Leaf Structures and Developmental Stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X).
- (79) Vershinin, A. Biological Functions of Carotenoids - Diversity and Evolution. *BioFactors* **1999**, *10*, 99–104. <https://doi.org/10.1002/biof.5520100203>.
- (80) Tanabe, Y.; Shitara, T.; Kashino, Y.; Hara, Y.; Kudoh, S. Utilizing the Effective Xanthophyll Cycle for Blooming of *Ochromonas Smithii* and *O. Itoi* (Chrysophyceae) on the Snow Surface. *PLoS One* **2011**, *6* (2). <https://doi.org/10.1371/journal.pone.0014690>.
- (81) Blackburn, G. A. Spectral Indices for Estimating Photosynthetic Pigment Concentrations: A Test Using Senescent Tree Leaves. *Int. J. Remote Sens.* **1998**, *19* (4), 657–675. <https://doi.org/10.1080/014311698215919>.
- (82) Feret, J. B.; François, C.; Asner, G. P.; Gitelson, A. A.; Martin, R. E.; Bidel, L. P. R.; Ustin, S. L.; le Maire, G.; Jacquemoud, S. PROSPECT-4 and 5: Advances in the Leaf Optical Properties Model Separating Photosynthetic Pigments. *Remote Sens. Environ.* **2008**, *112*, 3030–3043. <https://doi.org/10.1016/j.rse.2008.02.012>.
- (83) Cartelat, A.; Cerovic, Z. G.; Goulas, Y.; Meyer, S.; Lelarge, C.; Prioul, J. L.; Barbottin, A.; Jeuffroy, M. H.; Gate, P.; Agati, G.; *et al.* Optically Assessed Contents of Leaf Polyphenolics and Chlorophyll as Indicators of Nitrogen Deficiency in Wheat (*Triticum Aestivum* L.). *F. Crop. Res.* **2005**, *91* (1), 35–49. <https://doi.org/10.1016/j.fcr.2004.05.002>.
- (84) Muchecheti, F.; Madakadze, C.; Soundy, P. Leaf Chlorophyll Readings as an Indicator of Nitrogen Status and Yield of Spinach (*Spinacia Oleracea* L.) Grown in Soils Amended with *Luecaena Leucocephala* Prunings. *J. Plant Nutr.* **2016**, *39* (4), 539–561. <https://doi.org/10.1080/01904167.2016.1143488>.
- (85) Wang, J.; Wang, T.; Skidmore, A. K.; Shi, T.; Wu, G. Evaluating Different Methods for Grass Nutrient Estimation from Canopy Hyperspectral Reflectance. *Remote Sens.* **2015**, *7*, 5901–5917. <https://doi.org/10.3390/rs70505901>.
- (86) Zarco-Tejada, P. J.; Pushnik, J. C.; Dobrowski, S.; Ustin, S. L. Steady-State Chlorophyll a Fluorescence Detection from Canopy Derivative Reflectance and Double-Peak Red-Edge Effects. *Remote Sens. Environ.* **2003**, *84* (2), 283–294. [https://doi.org/10.1016/S0034-4257\(02\)00113-X](https://doi.org/10.1016/S0034-4257(02)00113-X).
- (87) Huang, J.; Wei, C.; Zhang, Y.; Blackburn, G. A.; Wang, X.; Wei, C.; Wang, J. Meta-Analysis of the Detection of Plant Pigment Concentrations Using Hyperspectral Remotely Sensed Data. *PLoS One* **2015**, *10* (9), 1–26. <https://doi.org/10.1371/journal.pone.0137029>.
- (88) Berger, K.; Atzberger, C.; Danner, M.; D'Urso, G.; Mauser, W.; Vuolo, F.; Hank, T. Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study. *Remote Sens.* **2018**, *10* (1). <https://doi.org/10.3390/rs10010085>.

- (89) Junker, L. V.; Ensminger, I. Relationship between Leaf Optical Properties, Chlorophyll Fluorescence and Pigment Changes in Senescing *Acer Saccharum* Leaves. *Tree Physiol.* **2016**, *36* (6), 694–711. <https://doi.org/10.1093/treephys/tpv148>.
- (90) Curran, P. J. Remote Sensing of Foliar Chemistry. *Remote Sens. Environ.* **1989**, *278*, 271–278. [https://doi.org/https://doi.org/10.1016/0034-4257\(89\)90069-2](https://doi.org/https://doi.org/10.1016/0034-4257(89)90069-2).
- (91) Lichtenthaler, H. K.; Buschmann, C. Chlorophylls and Carotenoids: Measurement and Characterization by UV-VIS. In *Current Protocols in Food Analytical Chemistry*; John Wiley & Sons, Inc., 2001. <https://doi.org/doi.org/10.1002/0471142913.faf0403s01>.
- (92) Zhang, Y.; Huang, J.; Wang, F.; Blackburn, G. A.; Zhang, H. K.; Wang, X.; Wei, C.; Zhang, K.; Wei, C. An Extended PROSPECT: Advance in the Leaf Optical Properties Model Separating Total Chlorophylls into Chlorophyll a and B. *Sci. Rep.* **2017**, *7* (6429). <https://doi.org/10.1038/s41598-017-06694-y>.
- (93) Ustin, S. L.; Schaepman, M. E.; Gitelson, A. A.; Jacquemoud, S.; Schaepman, M.; Asner, G. P.; Gamon, J. A.; Zarco-Tejada, P. Retrieval of Foliar Information about Plant Pigment Systems from High Resolution Spectroscopy. *Remote Sens. Environ.* **2009**, *113*, S67–S77.
- (94) Cho, M. A.; Skidmore, A. K. A New Technique for Extracting the Red Edge Position from Hyperspectral Data: The Linear Extrapolation Method. *Remote Sens. Environ.* **2006**, *101*, 181–193. <https://doi.org/10.1016/j.rse.2005.12.011>.
- (95) Filella, I.; Peñuelas, J. The Red Edge Position and Shape as Indicators of Plant Chlorophyll Content, Biomass and Hydric Status. *Int. J. Remote Sens.* **1994**, *15* (7), 1459–1470. <https://doi.org/10.1080/01431169408954177>.
- (96) Li, D.; Cheng, T.; Zhou, K.; Zheng, H.; Yao, X.; Tian, Y.; Zhu, Y.; Cao, W. WREP: A Wavelet-Based Technique for Extracting the Red Edge Position from Reflectance Spectra for Estimating Leaf and Canopy Chlorophyll Contents of Cereal Crops. *ISPRS J. Photogramm. Remote Sens.* **2017**, *129*, 103–117. <https://doi.org/10.1016/j.isprsjprs.2017.04.024>.
- (97) Archetti, M.; Döring, T. F.; Hagen, S. B.; Hughes, N. M.; Leather, S. R.; Lee, D. W.; Lev-Yadun, S.; Manetas, Y.; Ougham, H. J.; Schaberg, P. G.; *et al.* Unravelling the Evolution of Autumn Colours: An Interdisciplinary Approach. *Trends Ecol. Evol.* **2009**, *24* (3), 166–173. <https://doi.org/10.1016/j.tree.2008.10.006>.
- (98) Juvany, M.; Müller, M.; Munné-Bosch, S. Photo-Oxidative Stress in Emerging and Senescing Leaves: A Mirror Image. *J. Exp. Bot.* **2013**, *64* (11), 3087–3098. <https://doi.org/10.1093/jxb/ert174>.
- (99) Demmig-Adams, B.; Adams, W. W. The Role of Xanthophyll Cycle Carotenoids in the Protection of Photosynthesis. *Trends Plant Sci.* **1996**, *1* (1), 21–26. [https://doi.org/10.1016/S1360-1385\(96\)80019-7](https://doi.org/10.1016/S1360-1385(96)80019-7).
- (100) Niyogi, K. K.; Bjorkman, O.; Grossman, A. R. The Roles of Specific Xanthophylls in Photoprotection. *Proc. Natl. Acad. Sci.* **2002**, *94* (25), 14162–14167. <https://doi.org/10.1073/pnas.94.25.14162>.
- (101) Gamon, J. A. J. A.; Peñuelas, J.; Field, C. B. B.; Penuelas, J.; Field, C. B. B. A Narrow-Waveband Spectral Index That Tracks Diurnal Changes in Photosynthetic Efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [https://doi.org/10.1016/0034-4257\(92\)90059-S](https://doi.org/10.1016/0034-4257(92)90059-S).
- (102) Garbulsky, M. F.; Peñuelas, J.; Gamon, J.; Inoue, Y.; Filella, I. The Photochemical Reflectance Index (PRI) and the Remote Sensing of Leaf, Canopy and Ecosystem Radiation Use Efficiencies. A Review and Meta-Analysis. *Remote Sens. Environ.* **2011**, *115* (2), 281–

297. <https://doi.org/10.1016/j.rse.2010.08.023>.
- (103) Styliniski, C. D.; Gamon, J. A.; Oechel, W. C. Seasonal Patterns of Reflectance Indices, Carotenoid Pigments and Photosynthesis of Evergreen Chaparral Species. *Oecologia* **2002**, *131* (3), 366–374. <https://doi.org/10.1007/s00442-002-0905-9>.
- (104) Porcar-Castell, A.; Garcia-Plazaola, J. I.; Nichol, C. J.; Kolari, P.; Olascoaga, B.; Kuusinen, N.; Fernández-Marín, B.; Pulkkinen, M.; Juurola, E.; Nikinmaa, E. Physiology of the Seasonal Relationship between the Photochemical Reflectance Index and Photosynthetic Light Use Efficiency. *Oecologia* **2012**, *170*, 313–323. <https://doi.org/10.1007/s00442-012-2317-9>.
- (105) Mónica Giusti, M.; Wrolstad, R. E. Characterization and Measurement of Anthocyanins by UV-Visible Spectroscopy. *Handb. Food Anal. Chem.* **2005**, 2–2, 19–31. <https://doi.org/10.1002/0471709085.ch18>.
- (106) Dogwood, R.; Feild, T. S.; Lee, D. W.; Holbrook, N. M. Why Leaves Turn Red in Autumn . The Role of Anthocyanins in Senescing Leaves Of. **2001**, *127* (October), 566–574. <https://doi.org/10.1104/pp.010063.566>.
- (107) Hoch, W. A. Resorption Protection. Anthocyanins Facilitate Nutrient Recovery in Autumn by Shielding Leaves from Potentially Damaging Light Levels. *Plant Physiol.* **2003**, *133* (3), 1296–1305. <https://doi.org/10.1104/pp.103.027631>.
- (108) Féret, J. B.; Gitelson, A. A.; Noble, S. D.; Jacquemoud, S. PROSPECT-D: Towards Modeling Leaf Optical Properties through a Complete Lifecycle. *Remote Sens. Environ.* **2017**, *193*, 204–215. <https://doi.org/10.1016/j.rse.2017.03.004>.
- (109) Gitelson, A. A.; Chivkunova, O. B.; Merzlyak, M. N. Nondestructive Estimation of Anthocyanins and Chlorophylls in Anthocyanic Leaves. *Am. J. Bot.* **2009**, *96* (10), 1861–1868. <https://doi.org/10.3732/ajb.0800395>.
- (110) Baldini, E.; Facini, O.; Nerozzi, F.; Rossi, F.; Rotondi, A. Leaf Characteristics and Optical Properties of Different Woody Species. *Trees* **1997**, *12*, 73. <https://doi.org/10.1007/s004680050124>.
- (111) Foley, S.; Rivard, B.; Sanchez-Azofeifa, G. A.; Calvo, J. Foliar Spectral Properties Following Leaf Clipping and Implications for Handling Techniques. *Remote Sens. Environ.* **2006**, *103* (3), 265–275. <https://doi.org/10.1016/j.rse.2005.06.014>.
- (112) Grant, L. Diffuse and Specular Characteristics of Leaf Reflectance. *Remote Sens. Environ.* **1987**, *22* (2), 309–322. [https://doi.org/10.1016/0034-4257\(87\)90064-2](https://doi.org/10.1016/0034-4257(87)90064-2).
- (113) Jacquemoud, S. Utilisation de La Haute Résolution Spectrale Pour l'étude Des Couverts Végétaux: Développement d'un Modèle de Réflectance Spectrale, Université Paris VII, 1992.
- (114) Rossatto, D. R.; Kolb, R. M.; Franco, A. C. Leaf Anatomy Is Associated with the Type of Growth Form in Neotropical Savanna Plants. *Botanique* **2015**, *93* (8), 507–508. <https://doi.org/https://doi.org/10.1139/cjb-2015-0001>.
- (115) Boren, E. J.; Boschetti, L.; Johnson, D. M. Characterizing the Variability of the Structure Parameter in the PROSPECT Leaf Optical Properties Model. *Remote Sens.* **2019**, *11* (10), 1236. <https://doi.org/10.3390/rs11101236>.
- (116) Ollinger, S. V. Sources of Variability in Canopy Reflectance and the Convergent Properties of Plants. *New Phytol.* **2011**, *189*, 375–394. <https://doi.org/10.1111/j.1469-8137.2010.03536.x>.

- (117) Rapaport, T.; Hochberg, U.; Rachmilevitch, S.; Karnieli, A. The Effect of Differential Growth Rates across Plants on Spectral Predictions of Physiological Parameters. *PLoS One* **2014**, *9* (2). <https://doi.org/10.1371/journal.pone.0088930>.
- (118) Baránková, B.; Lazár, D.; Nauš, J. Analysis of the Effect of Chloroplast Arrangement on Optical Properties of Green Tobacco Leaves. *Remote Sens. Environ.* **2016**, *174*, 181–196. <https://doi.org/10.1016/j.rse.2015.12.011>.
- (119) Ourcival, J. M.; Joffre, R.; Rambal, S. Exploring the Relationships between Reflectance and Anatomical and Biochemical Properties in *Quercus Ilex* Leaves. *New Phytol.* **1999**, *143* (2), 351–364. <https://doi.org/10.1046/j.1469-8137.1999.00456.x>.
- (120) Xiao, Y.; Tholen, D.; Zhu, X. G. The Influence of Leaf Anatomy on the Internal Light Environment and Photosynthetic Electron Transport Rate: Exploration with a New Leaf Ray Tracing Model. *J. Exp. Bot.* **2016**, *67* (21), 6021–6035. <https://doi.org/10.1093/jxb/erw359>.
- (121) Hanba, Y. T.; Kogami, H.; Terashima, I. The Effect of Growth Irradiance on Leaf Anatomy and Photosynthesis in *Acer* Species Differing in Light Demand. *Plant, Cell Environ.* **2002**, *25* (8), 1021–1030. <https://doi.org/10.1046/j.1365-3040.2002.00881.x>.
- (122) Li, L.; Ma, Z.; Niinemets, Ü.; Guo, D. Three Key Sub-Leaf Modules and the Diversity of Leaf Designs. *Front. Plant Sci.* **2017**, *8* (September), 1–8. <https://doi.org/10.3389/fpls.2017.01542>.
- (123) Knapp, A. K. A. K.; Carter, G. A. G. A. Variability in Leaf Optical Properties among 26 Species From A Broad Range Of Habitats. *Am. J. Bot.* **1998**, *85* (7), 940–946. <https://doi.org/10.2307/2446360>.
- (124) Sánchez-Azofeifa, G. A.; Castro, K.; Wright, S. J.; Gamon, J.; Kalacska, M.; Rivard, B.; Schnitzer, S. A.; Feng, J. L. Differences in Leaf Traits, Leaf Internal Structure, and Spectral Reflectance between Two Communities of Lianas and Trees: Implications for Remote Sensing in Tropical Environments. *Remote Sens. Environ.* **2009**, *113* (10), 2076–2088. <https://doi.org/10.1016/j.rse.2009.05.013>.
- (125) Castro, K.; Sanchez-Azofeifa, G. Changes in Spectral Properties, Chlorophyll Content and Internal Mesophyll Structure of Senescing *Populus Balsamifera* and *Populus Tremuloides* Leaves. *Sensors* **2008**, *8*, 51–69. <https://doi.org/10.3390/s8010051>.
- (126) Asner, G. P.; Martin, R. E. Spectral and Chemical Analysis of Tropical Forests: Scaling from Leaf to Canopy Levels. *Remote Sens. Environ.* **2008**, *112* (10), 3958–3970. <https://doi.org/10.1016/j.rse.2008.07.003>.
- (127) Houborg, R.; Boegh, E. Mapping Leaf Chlorophyll and Leaf Area Index Using Inverse and Forward Canopy Reflectance Modeling and SPOT Reflectance Data. *Remote Sens. Environ.* **2008**, *112* (1), 186–202. <https://doi.org/10.1016/j.rse.2007.04.012>.
- (128) Wang, W. M.; Li, Z. L.; Su, H. B. Comparison of Leaf Angle Distribution Functions: Effects on Extinction Coefficient and Fraction of Sunlit Foliage. *Agric. For. Meteorol.* **2007**, *143* (1–2), 106–122. <https://doi.org/10.1016/j.agrformet.2006.12.003>.
- (129) de Wit, C. T. *Photosynthesis of Leaf Canopies*; 1965.
- (130) Haboudane, D.; Miller, J. R.; Pattey, E.; Zarco-Tejada, P. J.; Strachan, I. B. Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture. *Remote Sens. Environ.* **2004**, *90* (3), 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>.

- (131) Ge, T.; Sui, F.; Bai, L.; Tong, C.; Sun, N. Effects of Water Stress on Growth, Biomass Partitioning, and Water-Use Efficiency in Summer Maize (*Zea Mays* L.) throughout the Growth Cycle. *Acta Physiol. Plant.* **2012**, *34* (3), 1043–1053. <https://doi.org/10.1007/s11738-011-0901-y>.
- (132) Shao, H. B.; Chu, L. Y.; Jaleel, C. A.; Zhao, C. X. Water-Deficit Stress-Induced Anatomical Changes in Higher Plants. *Comptes Rendus - Biol.* **2008**, *331* (3), 215–225. <https://doi.org/10.1016/j.crvi.2008.01.002>.
- (133) Clevers, J. G. P. W.; Kooistra, L.; Schaepman, M. E. Using Spectral Information from the NIR Water Absorption Features for the Retrieval of Canopy Water Content. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10* (3), 388–397. <https://doi.org/10.1016/j.jag.2008.03.003>.
- (134) Peñuelas, J.; Filella, I.; Biel, C.; Serrano, L.; Savé, R. The Reflectance at the 950–970 nm Region as an Indicator of Plant Water Status. *Int. J. Remote Sens.* **1993**, *14* (10), 1887–1905. <https://doi.org/10.1080/01431169308954010>.
- (135) Sims, D. A.; Gamon, J. A. Estimation of Vegetation Water Content and Photosynthetic Tissue Area from Spectral Reflectance: A Comparison of Indices Based on Liquid Water and Chlorophyll Absorption Features. *Remote Sens. Environ.* **2003**, *84* (4), 526–537. [https://doi.org/10.1016/S0034-4257\(02\)00151-7](https://doi.org/10.1016/S0034-4257(02)00151-7).
- (136) Katsoulas, N.; Elvanidi, A.; Ferentinos, K. P.; Kacira, M.; Bartzanas, T.; Kittas, C. Crop Reflectance Monitoring as a Tool for Water Stress Detection in Greenhouses: A Review. *Biosyst. Eng.* **2016**, *151*, 374–398. <https://doi.org/10.1016/j.biosystemseng.2016.10.003>.
- (137) Percival, D. C.; Proctor, J. T. A.; Privé, J. P. Gas Exchange, Stem Water Potential and Leaf Orientation of *Rubus Idaeus* L. Are Influenced by Drought Stress. *J. Hortic. Sci. Biotechnol.* **1998**, *73* (6), 831–840. <https://doi.org/10.1080/14620316.1998.11511056>.
- (138) Baret, F.; Fourty, T. Estimation of Leaf Water Content and Specific Leaf Weight from Reflectance and Transmittance Measurements. *Agronomie* **2007**, *17* (9–10), 455–464. <https://doi.org/10.1051/agro:19970903>.
- (139) Ceccato, P.; Flasse, S.; Tarantola, S.; Jacquemoud, S.; Grégoire, J. M. Detecting Vegetation Leaf Water Content Using Reflectance in the Optical Domain. *Remote Sens. Environ.* **2001**, *77* (1), 22–33. [https://doi.org/10.1016/S0034-4257\(01\)00191-2](https://doi.org/10.1016/S0034-4257(01)00191-2).
- (140) Gao, B. NDWI—A Normalized Difference Water Index For. *Remote Sens. Environ.* **1996**, *266* (April), 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3).
- (141) Yilmaz, M. T.; Hunt, E. R.; Jackson, T. J. Remote Sensing of Vegetation Water Content from Equivalent Water Thickness Using Satellite Imagery. *Remote Sens. Environ.* **2008**, *112* (5), 2514–2522. <https://doi.org/10.1016/j.rse.2007.11.014>.
- (142) Tian, Q.; Tong, Q.; Pu, R.; Guo, X.; Zhao, C. Spectroscopic Determination of Wheat Water Status Using 1650–1850 nm Spectral Absorption Features. *Int. J. Remote Sens.* **2001**, *22* (12), 2329–2338. <https://doi.org/10.1080/01431160118199>.
- (143) Tucker, C. J. Remote Sensing of Leaf Water Content in the near Infrared. *Remote Sens. Environ.* **1980**, *10*, 23–32. [https://doi.org/10.1016/0034-4257\(80\)90096-6](https://doi.org/10.1016/0034-4257(80)90096-6).
- (144) Card, D. H.; Peterson, D. L.; Matson, P. A.; Aber, J. D. Prediction of Leaf Chemistry by the Use of Visible and near Infrared Reflectance Spectroscopy. *Remote Sens. Environ.* **1988**, *26* (2), 123–147. [https://doi.org/10.1016/0034-4257\(88\)90092-2](https://doi.org/10.1016/0034-4257(88)90092-2).
- (145) Jacquemoud, S.; Ustin, S. L.; Verdebout, J.; Schmuck, G.; Andreoli, G.; Hosgood, B.

- Estimating Leaf Biochemistry Using the PROSPECT Leaf Optical Properties Model. *Remote Sens. Environ.* **1996**, *56* (3), 194–202. [https://doi.org/10.1016/0034-4257\(95\)00238-3](https://doi.org/10.1016/0034-4257(95)00238-3).
- (146) Asif, M. Sustainability of Timber, Wood and Bamboo in Construction. In *Sustainability of Construction Materials*; Elsevier, 2009; pp 31–54. <https://doi.org/10.1533/9781845695842.31>.
- (147) Liu, Q.; Luo, L.; Zheng, L. Lignins: Biosynthesis and Biological Functions in Plants. *Int. J. Mol. Sci.* **2018**, *19* (2). <https://doi.org/10.3390/ijms19020335>.
- (148) Wang, Z.; Skidmore, A. K.; Wang, T.; Darvishzadeh, R.; Hearne, J. Applicability of the PROSPECT Model for Estimating Protein and Cellulose+lignin in Fresh Leaves. *Remote Sens. Environ.* **2015**, *168*, 205–218. <https://doi.org/10.1016/j.rse.2015.07.007>.
- (149) Nagler, P. L.; Inoue, Y.; Glenn, E. P.; Russ, A. L.; Daughtry, C. S. T. Cellulose Absorption Index (CAI) to Quantify Mixed Soil-Plant Litter Scenes. *Remote Sens. Environ.* **2003**, *87* (2–3), 310–325. <https://doi.org/10.1016/j.rse.2003.06.001>.
- (150) Serrano, L.; Peñuelas, J.; Ustin, S. L. Remote Sensing of Nitrogen and Lignin in Mediterranean Vegetation from AVIRIS Data. *Remote Sens. Environ.* **2002**, *81* (2–3), 355–364. [https://doi.org/10.1016/S0034-4257\(02\)00011-1](https://doi.org/10.1016/S0034-4257(02)00011-1).
- (151) Gerber, F.; Marion, R.; Olioso, A.; Jacquemoud, S.; Ribeiro da Luz, B.; Fabre, S. Modeling Directional-Hemispherical Reflectance and Transmittance of Fresh and Dry Leaves from 0.4 μm to 5.7 μm with the PROSPECT-VISIR Model. *Remote Sens. Environ.* **2011**, *115* (2), 404–414. <https://doi.org/10.1016/j.rse.2010.09.011>.
- (152) Adeniyi, A. .; Afolabi, J. . Determination of Total Petroleum Hydrocarbons and Heavy Metals in Soils within the Vicinity of Facilities Handling Refined Petroleum Products in Lagos Metropolis. *Environ. Int.* **2002**, *28*, 79–82. [https://doi.org/10.1016/S0160-4120\(02\)00007-7](https://doi.org/10.1016/S0160-4120(02)00007-7).
- (153) Doble, M.; Kumar, A. Petroleum Hydrocarbon Pollution. *Biotreat. Ind. Effluents* **2007**, 241–253. <https://doi.org/10.1016/B978-075067838-4/50025-7>.
- (154) Kisić, I.; Mešić, S.; Basić, F.; Brkić, V.; Mešić, M.; Durn, G.; Zgorelec, Z.; Bertović, L. The Effect of Drilling Fluids and Crude Oil on Some Chemical Characteristics of Soil and Crops. *Geoderma* **2009**, *149* (3–4), 209–216. <https://doi.org/10.1016/j.geoderma.2008.11.041>.
- (155) Brewer, R.; Nagashima, J.; Kelley, M.; Heskett, M.; Rigby, M. Risk-Based Evaluation of Total Petroleum Hydrocarbons in Vapor Intrusion Studies. *Int. J. Environ. Res. Public Health* **2013**, *10* (6), 2441–2467. <https://doi.org/10.3390/ijerph10062441>.
- (156) Turle, R.; Nason, T.; Malle, H.; Fowlie, P. Development and Implementation of the CCME Reference Method for the Canada-Wide Standard for Petroleum Hydrocarbons (PHC) in Soil: A Case Study. *Anal. Bioanal. Chem.* **2007**, *387* (3), 957–964. <https://doi.org/10.1007/s00216-006-0989-x>.
- (157) Baker, J. M. The Effects of Oils on Plants. *Environ. Pollut.* **1970**, *1*, 27–44. [https://doi.org/10.1016/0013-9327\(70\)90004-2](https://doi.org/10.1016/0013-9327(70)90004-2).
- (158) Nagajyoti, P. C.; Lee, K. D.; Sreekanth, T. V. M. Heavy Metals, Occurrence and Toxicity for Plants: A Review. *Environ. Chem. Lett.* **2010**, *8* (3), 199–216. <https://doi.org/10.1007/s10311-010-0297-8>.
- (159) Rosso, P. H.; Pushnik, J. C.; Lay, M.; Ustin, S. L. Reflectance Properties and Physiological Responses of *Salicornia Virginica* to Heavy Metal and Petroleum Contamination. *Environ. Pollut.* **2005**, *137*, 241–252. <https://doi.org/10.1016/j.envpol.2005.02.025>.

- (160) Balasubramaniyam, A.; Harvey, P. J. Scanning Electron Microscopic Investigations of Root Structural Modifications Arising from Growth in Crude Oil-Contaminated Sand. *Environ. Sci. Pollut. Res.* **2014**, *21* (22), 12651–12661. <https://doi.org/10.1007/s11356-014-3138-7>.
- (161) Rusin, M.; Gospodarek, J.; Nadgórska-Socha, A.; Barczyk, G. Effect of Petroleum-Derived Substances on Life History Traits of Black Bean Aphid (*Aphis Fabae* Scop.) and on the Growth and Chemical Composition of Broad Bean. *Ecotoxicology* **2017**, *26* (3), 308–319. <https://doi.org/10.1007/s10646-017-1764-9>.
- (162) Baruah, P.; Saikia, R. R.; Baruah, P. P.; Deka, S. Effect of Crude Oil Contamination on the Chlorophyll Content and Morpho-Anatomy of *Cyperus Brevifolius* (Rottb.) Hassk. *Environ. Sci. Pollut. Res.* **2014**, *21* (21), 12530–12538. <https://doi.org/10.1007/s11356-014-3195-y>.
- (163) Merkl, N.; Schultze-Kraft, R.; Infante, C. Phytoremediation in the Tropics - Influence of Heavy Crude Oil on Root Morphological Characteristics of Graminoids. *Environ. Pollut.* **2005**, *138* (1), 86–91. <https://doi.org/10.1016/j.envpol.2005.02.023>.
- (164) Punwong, P.; Juprasong, Y.; Traiperm, P. Effects of an Oil Spill on the Leaf Anatomical Characteristics of a Beach Plant (*Terminalia Catappa* L.). *Environ. Sci. Pollut. Res.* **2017**, *24*, 21821–21828. <https://doi.org/10.1007/s11356-017-9814-7>.
- (165) Shahid, M.; Dumat, C.; Khalid, S.; Schreck, E.; Xiong, T.; Niazi, N. K. Foliar Heavy Metal Uptake, Toxicity and Detoxification in Plants: A Comparison of Foliar and Root Metal Uptake. *J. Hazard. Mater.* **2017**, *325*, 36–58. <https://doi.org/10.1016/j.jhazmat.2016.11.063>.
- (166) Zhu, L.; Chen, Z.; Wang, J.; Ding, J.; Yu, Y.; Li, J.; Xiao, N.; Jiang, L.; Zheng, Y.; Rimmington, G. M. Monitoring Plant Response to Phenanthrene Using the Red Edge of Canopy Hyperspectral Reflectance. *Mar. Pollut. Bull.* **2014**, *86*, 332–341. <https://doi.org/10.1016/j.marpolbul.2014.06.046>.
- (167) Khamchhiyan, M.; Charkhabi, A. H.; Tajik, M. Effects of Crude Oil Contamination on Geotechnical Properties of Clayey and Sandy Soils Clayey and Sandy Soils. *Eng. Geol.* **2007**, *89*, 220–229. <https://doi.org/10.1016/j.enggeo.2006.10.009>.
- (168) Klamerus-Iwan, A.; Błońska, E.; Lasota, J.; Kalandyk, A.; Waligórski, P. Influence of Oil Contamination on Physical and Biological Properties of Forest Soil after Chainsaw Use. *Water. Air. Soil Pollut.* **2015**, *226* (11). <https://doi.org/10.1007/s11270-015-2649-2>.
- (169) Athar, H. ur R.; Ambreen, S.; Javed, M.; Hina, M.; Rasul, S.; Zafar, Z. U.; Manzoor, H.; Ogbaga, C. C.; Afzal, M.; Al-Qurainy, F.; *et al.* Influence of Sub-Lethal Crude Oil Concentration on Growth, Water Relations and Photosynthetic Capacity of Maize (*Zea Mays* L.) Plants. *Environ. Sci. Pollut. Res.* **2016**, *23* (18), 18320–18331. <https://doi.org/10.1007/s11356-016-6976-7>.
- (170) Ogboghodo, I. A.; Iruaga, E. K.; Osemwota, I. O.; Chokor, J. U. An Assessment of the Effects of Crude Oil Pollution on Soil Properties, Germination and Growth of Maize (*Zea Mays*) Using Two Crude Types – Forcados Light and Escravos Light. *Environ. Monit. Assess.* **2004**, *96*, 143–152. <https://doi.org/10.1023/B:EMAS.0000031723.62736.24>.
- (171) Wang, Y.; Feng, J.; Lin, Q.; Lyu, X.; Wang, X.; Wang, G. Effects of Crude Oil Contamination on Soil Physical and Chemical Properties in Momoge Wetland of China. *Chinese Geogr. Sci.* **2013**, *23* (6), 708–715. <https://doi.org/10.1007/s11769-013-0641-6>.
- (172) Barceló, J.; Poschenrieder, C. Plant Water Relations as Affected by Heavy Metal Stress: A Review. *J. Plant Nutr.* **1990**, *13* (1), 1–37. <https://doi.org/10.1080/01904169009364057>.
- (173) Nie, M.; Lu, M.; Yang, Q.; Zhang, X.; Xiao, M.; Jiang, L.; Yang, J.; Fang, C.; Chen, J.; Li,

- B. Plants ' Use of Different Nitrogen Forms in Response to Crude Oil Contamination. **2011**, *159*, 157–163. <https://doi.org/10.1016/j.envpol.2010.09.013>.
- (174) Hawrot-Paw, M.; Wijatkowski, A.; Mikiciuk, M. Influence of Diesel and Biodiesel Fuel-Contaminated Soil on Microorganisms, Growth and Development of Plants. *Plant Soil Environ.* **2015**, *61* (5), 189–194. <https://doi.org/10.17221/974/2014-PSE>.
- (175) Jiao, S.; Liu, Z.; Lin, Y.; Yang, J.; Chen, W.; Wei, G. Bacterial Communities in Oil Contaminated Soils: Biogeography and Co- Occurrence Patterns Occurrence Patterns. *Soil Biol. Biochem.* **2018**, *98*, 64–73. <https://doi.org/10.1016/j.soilbio.2016.04.005>.
- (176) Liao, J.; Wang, J.; Jiang, D.; Wang, M. C.; Huang, Y. Long-Term Oil Contamination Causes Similar Changes in Microbial Communities of Two Distinct Soils. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 10299–10310. <https://doi.org/10.1007/s00253-015-6880-y>.
- (177) Labud, V.; Garcia, C.; Hernandez, T. Effect of Hydrocarbon Pollution on the Microbial Properties of a Sandy and a Clay Soil. *Chemosphere* **2007**, *66* (10), 1863–1871. <https://doi.org/10.1016/j.chemosphere.2006.08.021>.
- (178) Xie, Y.; Fan, J.; Zhu, W.; Amombo, E.; Lou, Y.; Chen, L.; Fu, J. Effect of Heavy Metals Pollution on Soil Microbial Diversity and Bermudagrass Genetic Variation. *Front. Plant Sci.* **2016**, *7* (May), 1–12. <https://doi.org/10.3389/fpls.2016.00755>.
- (179) Fu, X.; Cui, Z.; Zang, G. Environmental Science Processes & Impacts Extraction Processes. *Environ. Sci. Process. Impacts* **2014**, *16*, 1737–1744. <https://doi.org/10.1039/c3em00618b>.
- (180) Zuofa, K.; Loganathan, P.; Isirimah, N. O. Effects of Crude Oil Applications to Soil on the Growth and Yield of Maize, Okro and Cassava in Nigeria. *Oil Chem. Pollut.* **1988**, *4* (4), 249–259. [https://doi.org/10.1016/S0269-8579\(88\)80001-7](https://doi.org/10.1016/S0269-8579(88)80001-7).
- (181) Martí, M. C.; Camejo, D.; Fernández-García, N.; Rellán-Álvarez, R.; Marques, S.; Sevilla, F.; Jiménez, A. Effect of Oil Refinery Sludges on the Growth and Antioxidant System of *Alfalfa* Plants. *J. Hazard. Mater.* **2009**, *171* (1–3), 879–885. <https://doi.org/10.1016/j.jhazmat.2009.06.083>.
- (182) Nogueira, L.; Inckot, C.; Santos, G. D. O. Phytotoxicity of Petroleum-Contaminated Soil and Bioremediated Soil on *Allophylus Edulis*. *Rodriguésia* **2011**, *62* (3), 459–466.
- (183) Balliana, A. G.; Moura, B. B.; Inckot, R. C.; Bona, C. Development of *Canavalia Ensiformis* in Soil Contaminated with Diesel Oil. *Environ. Sci. Pollut. Res.* **2017**, *24*, 979–986. <https://doi.org/10.1007/s11356-016-7674-1>.
- (184) Dupuy, J.; Ouvrard, S.; Leglize, P.; Sterckeman, T. Morphological and Physiological Responses of Maize (*Zea Mays*) Exposed to Sand Contaminated by Phenanthrene. *Chemosphere* **2015**, *124* (1), 110–115. <https://doi.org/10.1016/j.chemosphere.2014.11.051>.
- (185) Sanches, I. D.; Souza Filho, C. R.; Magalhães, L. A.; Quitério, G. C. M.; Alves, M. N.; Oliveira, W. J. Unravelling Remote Sensing Signatures of Plants Contaminated with Gasoline and Diesel: An Approach Using the Red Edge Spectral Feature. *Environ. Pollut.* **2013**, *174*, 16–27. <https://doi.org/10.1016/j.envpol.2012.10.029>.
- (186) Emengini, E. J.; Blackburn, G. A.; Theobald, J. C. Early Detection of Oil-Induced Stress in Crops Using Spectral and Thermal Responses. *J. Appl. Remote Sens.* **2013**, *7*. <https://doi.org/10.1117/1.jrs.7.073596>.
- (187) Shanker, A. K.; Cervantes, C.; Loza-Tavera, H.; Avudainayagam, S. Chromium Toxicity in Plants. *Environ. Int.* **2005**, *31* (5), 739–753. <https://doi.org/10.1016/j.envint.2005.02.003>.

- (188) Anoliefo, G. O.; Vwioko, D. E. Effects of Spent Lubricating Oil on the Growth of *Capsicum Annum* L. and *Lycopersicon Esculentum* Miller. *Environ. Pollut.* **1995**, *88* (3), 361–364. [https://doi.org/10.1016/0269-7491\(95\)93451-5](https://doi.org/10.1016/0269-7491(95)93451-5).
- (189) Pérez-Hernández, I.; Ochoa-Gaona, S.; Adams, R. H.; Rivera-Cruz, M. C.; Pérez-Hernández, V.; Jarquín-Sánchez, A.; Geissen, V.; Martínez-Zurimendi, P. Growth of Four Tropical Tree Species in Petroleum-Contaminated Soil and Effects of Crude Oil Contamination. *Environ. Sci. Pollut. Res.* **2017**, *24* (2), 1769–1783. <https://doi.org/10.1007/s11356-016-7877-5>.
- (190) Merkl, N.; Schultze-Kraft, R.; Infante, C. Phytoremediation in the Tropics—the Effect of Crude Oil on the Growth of Tropical Plants. *Bioremediat. J.* **2004**, *8* (3–4), 177–184. <https://doi.org/10.1080/10889860490887527>.
- (191) Kvesitadze, G.; Khatisashvili, G.; Sadunishvili, T.; Ramsden, J. J. *Biochemical Mechanisms of Detoxification in Higher Plants*; Springer-Verlag Berlin Heidelberg, 2006.
- (192) Li, Y.; Morris, J. T.; Yoch, D. C. Chronic Low Level Hydrocarbon Amendments Stimulate Plant Growth and Microbial Activity in Salt-Marsh Microcosms. *J. Appl. Ecol.* **1990**, *27* (1), 159–171. <https://doi.org/10.2307/2403575>.
- (193) Han, G.; Cui, B. X.; Zhang, X. X.; Li, K. R. The Effects of Petroleum-Contaminated Soil on Photosynthesis of *Amorpha Fruticosa* Seedlings. *Int. J. Environ. Sci. Technol.* **2016**, *13*, 2383–2392. <https://doi.org/10.1007/s13762-016-1071-7>.
- (194) Malallah, G.; Afzal, M.; Gulshan, S.; Abraham, D.; Kurian, M.; Dhami, M. S. I. *Vicia Faba* as a Bioindicator of Oil Pollution. *Environ. Pollut.* **1996**, *92* (2), 213–217. [https://doi.org/10.1016/0269-7491\(95\)00085-2](https://doi.org/10.1016/0269-7491(95)00085-2).
- (195) Adenipekun, C. O.; Oyetunji, O. J.; Kassim, L. S. Effect of Spent Engine Oil on the Growth Parameters and Chlorophyll Content of *Corchorus Olitorius* Linn. *Environmentalist* **2008**, *28* (4), 446–450. <https://doi.org/10.1007/s10669-008-9165-5>.
- (196) Nakata, C.; Qualizza, C.; MacKinnon, M.; Renault, S. Growth and Physiological Responses of *Triticum Aestivum* and *Deschampsia Caespitosa* Exposed to Petroleum Coke. *Water. Air. Soil Pollut.* **2011**, *216* (1–4), 59–72. <https://doi.org/10.1007/s11270-010-0514-x>.
- (197) De Jong, S. M.; Addink, E. A.; Hoogenboom, P.; Nijland, W. The Spectral Response of *Buxus Sempervirens* to Different Types of Environmental Stress - A Laboratory Experiment. *ISPRS J. Photogramm. Remote Sens.* **2012**, *74*, 56–65. <https://doi.org/10.1016/j.isprsjprs.2012.08.005>.
- (198) Lowe, A.; Harrison, N.; French, A. P. Hyperspectral Image Analysis Techniques for the Detection and Classification of the Early Onset of Plant Disease and Stress. *Plant Methods* **2017**, *13* (1), 1–12. <https://doi.org/10.1186/s13007-017-0233-z>.
- (199) Stimson, H. C.; Breshears, D. D.; Ustin, S. L.; Kefauver, S. C. Spectral Sensing of Foliar Water Conditions in Two Co-Occurring Conifer Species: *Pinus Edulis* and *Juniperus Monosperma*. *Remote Sens. Environ.* **2005**, *96* (1), 108–118. <https://doi.org/10.1016/j.rse.2004.12.007>.
- (200) Emengini, E. J.; Blackburn, G. A.; Theobald, J. C. Discrimination of Plant Stress Caused by Oil Pollution and Waterlogging Using Hyperspectral and Thermal Remote Sensing. *J. Appl. Remote Sens.* **2013**, *7*. <https://doi.org/10.1117/1.jrs.7.073476>.
- (201) Emengini, E. J.; Ezech, F. C.; Chigbu, N. Comparative Analysis of Spectral Responses of Varied Plant Species to Oil Stress. *Int. J. Sci. Eng. Res.* **2013**, *4* (6), 1421–1427.

- (202) Clevers, J. G. P. W. P. W.; Kooistra, L.; Salas, E. A. L. L. Study of Heavy Metal Contamination in River Floodplains Using the Red-Edge Position in Spectroscopic Data. *Int. J. Remote Sens.* **2004**, *25* (19), 3883–3895. <https://doi.org/10.1080/01431160310001654473>.
- (203) Shi, T.; Liu, H.; Wang, J.; Chen, Y.; Fei, T.; Wu, G. Monitoring Arsenic Contamination in Agricultural Soils with Reflectance Spectroscopy of Rice Plants. *Environ. Sci. Technol.* **2014**, *48* (11), 6264–6272. <https://doi.org/10.1021/es405361n>.
- (204) Huang, S.; Chen, S.; Wang, D.; Zhou, C.; van der Meer, F.; Zhang, Y. Hydrocarbon Micro-Seepage Detection from Airborne Hyper-Spectral Images by Plant Stress Spectra Based on the PROSPECT Model. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *74* (July 2018), 180–190. <https://doi.org/10.1016/j.jag.2018.09.012>.
- (205) Gürtler, S.; de Souza Filho, C. R.; Sanches, I. D.; Alves, M. N.; Oliveira, W. J. Determination of Changes in Leaf and Canopy Spectra of Plants Grown in Soils Contaminated with Petroleum Hydrocarbons. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146* (April), 272–288. <https://doi.org/10.1016/j.isprsjprs.2018.09.011>.
- (206) Emengini, E. J.; Blackburn, G. A.; Theobald, J. C. Detection and Discrimination of Oil and Water Deficit-Induced Stress in Maize (*Zea Mays* L.) Using Spectral and Thermal Responses. *IOSR J. Environ. Sci. Toxicol. Food Technol.* **2013**, *3* (3), 53–57.
- (207) Apan, A.; Held, A.; Phinn, S.; Markley, J. Detecting Sugarcane ‘Orange Rust’ Disease Using EO-1 Hyperion Hyperspectral Imagery. *Int. J. Remote Sens.* **2004**, *25* (2), 489–498. <https://doi.org/10.1080/01431160310001618031>.
- (208) Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C.; Corsi, F.; Cho, M. LAI and Chlorophyll Estimation for a Heterogeneous Grassland Using Hyperspectral Measurements. *ISPRS J. Photogramm. Remote Sens.* **2008**, *63* (4), 409–426. <https://doi.org/10.1016/j.isprsjprs.2008.01.001>.
- (209) Axelsson, C.; Skidmore, A. K.; Schlerf, M.; Fauzi, A.; Verhoef, W. Hyperspectral Analysis of Mangrove Foliar Chemistry Using PLSR and Support Vector Regression. *Int. J. Remote Sens.* **2013**, *34* (5), 1724–1743. <https://doi.org/10.1080/01431161.2012.725958>.
- (210) Hermosilla, T.; Wulder, M. A.; White, J. C.; Coops, N. C.; Hobart, G. W. Regional Detection, Characterization, and Attribution of Annual Forest Change from 1984 to 2012 Using Landsat-Derived Time-Series Metrics. *Remote Sens. Environ.* **2015**, *170*, 121–132. <https://doi.org/10.1016/j.rse.2015.09.004>.
- (211) Schowengerdt, R. A. *Remote Sensing: Models and Methods for Image Processing*, 3rd Ed.; Elsevier Inc., 2006. <https://doi.org/10.1016/C2009-0-21902-7>.
- (212) Story, M.; Congalton, R. G. Remote Sensing Brief Accuracy Assessment: A User’s Perspective. *Photogramm. Eng. Remote Sensing* **1986**, *52* (3), 397–399.
- (213) Belgiu, M.; Drăguț, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- (214) Tuia, D.; Verrelst, J.; Alonso, L.; Perez-Cruz, F.; Camps-Valls, G. Multioutput Support Vector Regression for Remote Sensing Biophysical Parameter Estimation. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8* (4), 804–808. <https://doi.org/10.1109/LGRS.2011.2109934>.
- (215) Melgani, F.; Bruzzone, L. Classification of Hyperspectral Remote Sensing Images with Support Vector Machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42* (8), 1778–1790. <https://doi.org/10.1109/TGRS.2004.831865>.

- (216) Hughes, G. On the Mean Accuracy of Statistical Pattern Recognizers. *IEEE Trans. Inf. Theory* **1968**, *14*, 55–63. <https://doi.org/10.1109/TIT.1968.1054102>.
- (217) Thenkabail, P. S.; Enclona, E. A.; Ashton, M. S.; Van Der Meer, B. Accuracy Assessments of Hyperspectral Waveband Performance for Vegetation Analysis Applications. *Remote Sens. Environ.* **2004**, *91* (3–4), 354–376. <https://doi.org/10.1016/j.rse.2004.03.013>.
- (218) Wei, C.; Huang, J.; Wang, X.; Blackburn, G. A.; Zhang, Y.; Wang, S.; Mansaray, L. R. Hyperspectral Characterization of Freezing Injury and Its Biochemical Impacts in Oilseed Rape Leaves. *Remote Sens. Environ.* **2017**, *195*, 56–66. <https://doi.org/10.1016/j.rse.2017.03.042>.
- (219) Friedman, J. H.; Popescu, B. E. *Gradient Directed Regularization for Linear Regression and Classification*; 2004.
- (220) Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. *Ann. Eugen.* **1936**, *7* (2), 179–188. <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>.
- (221) Belsley, D. A.; Kuh, E.; Wemsch, R. E. Detecting and Assessing Collinearity. In *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*; John Wiley & Sons, I., Ed.; 1980.
- (222) Zou, H.; Hastie, T. Regression and Variable Selection via the Elastic Net. *J. R. Stat. Soc. Ser. B (Statistical Methodol.* **2005**, *67* (2), 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>.
- (223) Ireland, G.; Volpi, M.; Petropoulos, G. P. Examining the Capability of Supervised Machine Learning Classifiers in Extracting Flooded Areas from Landsat Tm Imagery: A Case Study from a Mediterranean Flood. *Remote Sens.* **2015**, *7* (3), 3372–3399. <https://doi.org/10.3390/rs70303372>.
- (224) Morel, J.; Jay, S.; Féret, J. B.; Bakache, A.; Bendoula, R.; Carreel, F.; Gorretta, N. Exploring the Potential of PROCOSINE and Close-Range Hyperspectral Imaging to Study the Effects of Fungal Diseases on Leaf Physiology. *Sci. Rep.* **2018**, *8* (1), 1–13. <https://doi.org/10.1038/s41598-018-34429-0>.
- (225) Keith, T. Z. *Multiple Regression and Beyond*, 2nd Ed.; Taylor & Francis, Ed.; Routledge, 2014. <https://doi.org/10.4324/9781315749099>.
- (226) Mountrakis, G.; Im, J.; Ogole, C. Support Vector Machines in Remote Sensing: A Review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66* (3), 247–259. <https://doi.org/10.1016/j.isprsjprs.2010.11.001>.
- (227) Jacquemoud, S.; Ustin, S. L. *Leaf Optical Properties: A State of the Art*; 2001.
- (228) Jiang, J.; Comar, A.; Burger, P.; Bancal, P.; Weiss, M.; Baret, F. Estimation of Leaf Traits from Reflectance Measurements: Comparison between Methods Based on Vegetation Indices and Several Versions of the PROSPECT Model. *Plant Methods* **2018**, *14* (23). <https://doi.org/10.1186/s13007-018-0291-x>.
- (229) Barry, K. M.; Newnham, G. J.; Stone, C. Estimation of Chlorophyll Content in *Eucalyptus Globulus* Foliage with the Leaf Reflectance Model PROSPECT. *Agric. For. Meteorol.* **2009**, *149* (6–7), 1209–1213. <https://doi.org/10.1016/j.agrformet.2009.01.005>.
- (230) Arellano, P.; Tansey, K.; Balzter, H.; Boyd, D. S. Field Spectroscopy and Radiative Transfer Modelling to Assess Impacts of Petroleum Pollution on Biophysical and Biochemical Parameters of the Amazon Rainforest. *Environ. Earth Sci.* **2017**, *76* (5), 1–14.

<https://doi.org/10.1007/s12665-017-6536-6>.

- (231) Adamu, B.; Tansey, K.; Ogutu, B. Remote Sensing for Detection and Monitoring of Vegetation Affected by Oil Spills. *Int. J. Remote Sens.* **2018**, *39* (11), 3628–3645. <https://doi.org/10.1080/01431161.2018.1448483>.
- (232) Ozigis, M. S.; Kaduk, J. D.; Jarvis, C. H. Mapping Terrestrial Oil Spill Impact Using Machine Learning Random Forest and Landsat 8 OLI Imagery: A Case Site within the Niger Delta Region of Nigeria. *Environ. Sci. Pollut. Res.* **2019**, *26*, 3621–3635. <https://doi.org/10.1007/s11356-018-3824-y>.
- (233) van der Meijde, M.; van der Werff, H. M. A.; Jansma, P. F.; van der Meer, F. D.; Groothuis, G. J. A Spectral-Geophysical Approach for Detecting Pipeline Leakage. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11* (1), 77–82. <https://doi.org/10.1016/j.jag.2008.08.002>.
- (234) Onyia, N. N.; Balzter, H.; Berrio, J. C. Detecting Vegetation Response to Oil Pollution Using Hyperspectral Indices. *Int. Geosci. Remote Sens. Symp.* **2018**, *2018-July*, 3963–3966. <https://doi.org/10.1109/IGARSS.2018.8519398>.
- (235) Adamu, B.; Tansey, K.; Ogutu, B. An Investigation into the Factors Influencing the Detectability of Oil Spills Using Spectral Indices in an Oil-Polluted Environment. *Int. J. Remote Sens.* **2016**, *37* (10), 2338–2357. <https://doi.org/10.1080/01431161.2016.1176271>.
- (236) Salem, F.; Kafatos, M.; El-Ghazawi, T.; Gomez, R.; Yang, R. Hyperspectral Image Assessment of Oil-contaminated Wetland. *Int. J. Remote Sens.* **2005**, *26* (4), 811–821. <https://doi.org/10.1080/01431160512331316883>.
- (237) Croft, H.; Chen, J. M.; Zhang, Y.; Simic, A. Modelling Leaf Chlorophyll Content in Broadleaf and Needle Leaf Canopies from Ground, CASI, Landsat TM 5 and MERIS Reflectance Data. *Remote Sens. Environ.* **2013**, *133*, 128–140. <https://doi.org/10.1016/j.rse.2013.02.006>.
- (238) Lausch, A.; Pause, M.; Merbach, I.; Zacharias, S.; Doktor, D.; Volk, M.; Seppelt, R. A New Multiscale Approach for Monitoring Vegetation Using Remote Sensing-Based Indicators in Laboratory, Field, and Landscape. *Environ. Monit. Assess.* **2013**, *185* (2), 1215–1235. <https://doi.org/10.1007/s10661-012-2627-8>.
- (239) Zarco-Tejada, P. J. J.; Berjón, A.; López-Lozano, R.; Miller, J. R. R.; Martín, P.; Cachorro, V.; González, M. R. R.; de Frutos, A. Assessing Vineyard Condition with Hyperspectral Indices: Leaf and Canopy Reflectance Simulation in a Row-Structured Discontinuous Canopy. *Remote Sens. Environ.* **2005**, *99* (3), 271–287. <https://doi.org/10.1016/j.rse.2005.09.002>.
- (240) Blackburn, G. A. Quantifying Chlorophylls and Carotenoids at Leaf and Canopy Scales: An Evaluation of Some Hyperspectral Approaches. *Remote Sens. Environ.* **1998**, *66* (3), 273–285. [https://doi.org/10.1016/S0034-4257\(98\)00059-5](https://doi.org/10.1016/S0034-4257(98)00059-5).
- (241) Daughtry, C. S. T.; Walthall, C. L.; Kim, M. S.; Brown de Colstoun, E.; McMurtrey III, J. E. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9).
- (242) Verhoef, W. Light Scattering by Leaf Layers with Application to Canopy Reflectance Modeling: The SAIL Model. *Remote Sens. Environ.* **1984**, *16*, 125–141. [https://doi.org/10.1016/0034-4257\(84\)90057-9](https://doi.org/10.1016/0034-4257(84)90057-9).
- (243) Leifer, I.; Lehr, W. J.; Simecek-Beatty, D.; Bradley, E.; Clark, R.; Dennison, P.; Hu, Y.; Matheson, S.; Jones, C. E.; Holt, B.; *et al.* State of the Art Satellite and Airborne Marine Oil

- Spill Remote Sensing: Application to the BP Deepwater Horizon Oil Spill. *Remote Sens. Environ.* **2012**, *124*, 185–209. <https://doi.org/10.1016/J.RSE.2012.03.024>.
- (244) Kokaly, R. F.; Couvillion, B. R.; Holloway, J. M.; Roberts, D. A.; Ustin, S. L.; Peterson, S. H.; Khanna, S.; Piazza, S. C. Spectroscopic Remote Sensing of the Distribution and Persistence of Oil from the Deepwater Horizon Spill in Barataria Bay Marshes. *Remote Sens. Environ.* **2013**, *129*, 210–230. <https://doi.org/10.1016/J.RSE.2012.10.028>.
- (245) Shi, T.; Chen, Y.; Liu, Y.; Wu, G. Visible and Near-Infrared Reflectance Spectroscopy — An Alternative for Monitoring Soil Contamination by Heavy Metals. *J. Hazard. Mater.* **2014**, *265*, 166–176. <https://doi.org/10.1016/j.jhazmat.2013.11.059>.
- (246) Hunt, E. R.; Rock, B. N. Detection of Changes in Leaf Water Content Using Near- and Middle-Infrared Reflectances. *Remote Sens. Environ.* **1989**, *30* (1), 43–54. [https://doi.org/10.1016/0034-4257\(89\)90046-1](https://doi.org/10.1016/0034-4257(89)90046-1).
- (247) Sihota, N. J.; Singurindy, O.; Mayer, K. U. CO₂-Efflux Measurements for Evaluating Source Zone Natural Attenuation Rates in a Petroleum Hydrocarbon Contaminated Aquifer. *Environ. Sci. Technol.* **2011**, *45* (2), 482–488. <https://doi.org/10.1021/es1032585>.
- (248) Dindar, E.; Topaç Şağban, F. O.; Başkaya, H. S. Variations of Soil Enzyme Activities in Petroleum-Hydrocarbon Contaminated Soil. *Int. Biodeterior. Biodegradation* **2015**, *105*, 268–275. <https://doi.org/10.1016/J.IBIOD.2015.09.011>.
- (249) Li, L.; Ustin, S. L.; Lay, M. Application of AVIRIS Data in Detection of Oil-Induced Vegetation Stress and Cover Change at Jornada, New Mexico. *Remote Sens. Environ.* **2005**, *94* (1), 1–16. <https://doi.org/10.1016/j.rse.2004.08.010>.
- (250) Clark, R. N.; King, T. V. V.; Gorelick, N. S.; Geological, U. S. Automatic *Continuum* Analysis of Reflectance Spectra. In *Airborne Imaging Spectrometer Data Analysis Workshop*, 1985; pp 138–142.
- (251) Noomen, M. F.; Skidmore, A. K.; van der Meer, F. D.; Prins, H. H. T. *Continuum* Removed Band Depth Analysis for Detecting the Effects of Natural Gas, Methane and Ethane on Maize Reflectance. *Remote Sens. Environ.* **2006**, *105* (3), 262–270. <https://doi.org/10.1016/j.rse.2006.07.009>.
- (252) Sridhar, B. B. M. M.; Han, F. X.; Diehl, S. V.; Monts, D. L.; Su, Y. Spectral Reflectance and Leaf Internal Structure Changes of Barley Plants Due to Phytoextraction of Zinc and Cadmium. *Int. J. Remote Sens.* **2007**, *28* (5), 1041–1054. <https://doi.org/10.1080/01431160500075832>.
- (253) Lichtenthaler, H. K. K.; Lang, M.; Sowinska, M.; Heisel, F.; Miehé, J. A. A. Detection of Vegetation Stress via a New High Resolution Fluorescence Imaging System. *J. Plant Physiol.* **1996**, *148* (5), 599–612. [https://doi.org/10.1016/S0176-1617\(96\)80081-2](https://doi.org/10.1016/S0176-1617(96)80081-2).
- (254) Raymond Hunt, E.; Rock, B. N.; Nobel, P. S. Measurement of Leaf Relative Water Content by Infrared Reflectance. *Remote Sens. Environ.* **1987**, *22*, 429–435. [https://doi.org/10.1016/0034-4257\(87\)90094-0](https://doi.org/10.1016/0034-4257(87)90094-0).
- (255) Swapna, K. S.; Salim, N.; Chandra, R.; Puthur, J. T. Structural Changes in Response to Bioaccumulation of Iron and Mercury in *Chromolaena Odorata* (L.) King & Robins. *Environ. Monit. Assess.* **2015**, *187* (9), 551. <https://doi.org/10.1007/s10661-015-4732-y>.
- (256) Peñuelas, J.; Pinol, J.; Ogaya, R.; Filella, I. Estimation of Plant Water Concentration by the Reflectance Water Index WI (R900/R970). *Int. J. Remote Sens.* **1997**, *18* (13), 2869–2875. <https://doi.org/10.1080/014311697217396>.

- (257) Milton, N. .; Ager, C. .; Eiswerth, B. .; Power, M. . Arsenic- and Selenium-Induced Changes in Spectral Reflectance and Morphology of Soybean Plants. *Remote Sens. Environ.* **1989**, *30* (3), 263–269. [https://doi.org/10.1016/0034-4257\(89\)90068-0](https://doi.org/10.1016/0034-4257(89)90068-0).
- (258) Smith, K. L. L.; Steven, M. D. D.; Colls, J. J. J. Use of Hyperspectral Derivative Ratios in the Red-Edge Region to Identify Plant Stress Responses to Gas Leaks. *Remote Sens. Environ.* **2004**, *92* (2), 207–217. <https://doi.org/10.1016/j.rse.2004.06.002>.
- (259) Kooistra, L.; Salas, E. A. .; Clevers, J. G. P. .; Wehrens, R.; Leuven, R. S. E. .; Nienhuis, P. .; Buydens, L. M. . Exploring Field Vegetation Reflectance as an Indicator of Soil Contamination in River Floodplains. *Environ. Pollut.* **2004**, *127* (2), 281–290. [https://doi.org/10.1016/S0269-7491\(03\)00266-5](https://doi.org/10.1016/S0269-7491(03)00266-5).
- (260) Pacumbaba, R. O.; Beyl, C. A. Changes in Hyperspectral Reflectance Signatures of Lettuce Leaves in Response to Macronutrient Deficiencies. *Adv. Sp. Res.* **2011**, *48* (1), 32–42. <https://doi.org/10.1016/J.ASR.2011.02.020>.
- (261) Adam, G.; Duncan, H. J. Effect of Diesel Fuel on Growth of Selected Plant Species. *Environ. Geochem. Health* **1999**, *21* (4), 353–357. <https://doi.org/10.1023/A:1006744603461>.
- (262) Beland, M.; Roberts, D. A.; Peterson, S. H.; Biggs, T. W.; Kokaly, R. F.; Piazza, S.; Roth, K. L.; Khanna, S.; Ustin, S. L. Mapping Changing Distributions of Dominant Species in Oil-Contaminated Salt Marshes of Louisiana Using Imaging Spectroscopy. *Remote Sens. Environ.* **2016**, *182*, 192–207. <https://doi.org/10.1016/j.rse.2016.04.024>.
- (263) Wang, Z.; Fingas, M.; Page, D. S. Oil Spill Identification. *J. Chromatogr. A* **1999**, *843*, 369–411. <https://doi.org/10.1002/9781118989982.ch6>.
- (264) Logan, G. A.; Jones, A. T.; Kennard, J. M.; Ryan, G. J.; Rollet, N. Australian Offshore Natural Hydrocarbon Seepage Studies, a Review and Re-Evaluation. *Mar. Pet. Geol.* **2010**, *27* (1), 26–45. <https://doi.org/10.1016/j.marpetgeo.2009.07.002>.
- (265) Angelliaume, S.; Ceamanos, X.; Viallefont-Robinet, F.; Baqué, R.; Déliot, P.; Miegbielle, V. Hyperspectral and Radar Airborne Imagery over Controlled Release of Oil at Sea. *Sensors* **2017**, *17* (8), 1772. <https://doi.org/10.3390/s17081772>.
- (266) Konik, M.; Bradtke, K. Object-Oriented Approach to Oil Spill Detection Using ENVISAT ASAR Images. *ISPRS J. Photogramm. Remote Sens.* **2016**, *118*, 37–52. <https://doi.org/10.1016/J.ISPRSJPRS.2016.04.006>.
- (267) Asadzadeh, S.; de Souza Filho, C. R. A Review on Spectral Processing Methods for Geological Remote Sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *47*, 69–90. <https://doi.org/10.1016/J.JAG.2015.12.004>.
- (268) Lassalle, G.; Credoza, A.; Fabre, S.; Elger, A.; Hédacq, R.; Dubucq, D. Hyperspectral Signature Analysis of Three Plant Species to Long-Term Hydrocarbon and Heavy Metal Exposure. In *Earth Resources and Environmental Remote Sensing/GIS Applications VIII*; Michel, U., Schulz, K., Eds.; SPIE, 2017; Vol. 10428, p 33. <https://doi.org/10.1117/12.2277709>.
- (269) Wang, F.; Gao, J.; Zha, Y. Hyperspectral Sensing of Heavy Metals in Soil and Vegetation: Feasibility and Challenges. *ISPRS J. Photogramm. Remote Sens.* **2018**, *136*, 73–84. <https://doi.org/10.1016/j.isprsjprs.2017.12.003>.
- (270) Arellano, P.; Tansey, K.; Balzter, H.; Tellkamp, M. Plant Family-Specific Impacts of Petroleum Pollution on Biodiversity and Leaf Chlorophyll Content in the Amazon Rainforest of Ecuador. *PLoS One* **2017**, *12* (1). <https://doi.org/10.1371/journal.pone.0169867>.

- (271) Lassalle, G.; Credoza, A.; Hédacq, R.; Fabre, S.; Dubucq, D.; Elger, A. Assessing Soil Contamination Due to Oil and Gas Production Using Vegetation Hyperspectral Reflectance. *Environ. Sci. Technol.* **2018**, *52* (4), 1756–1764. <https://doi.org/10.1021/acs.est.7b04618>.
- (272) Metwally, M. E. S.; Al-Muzaini, S.; Jacob, P. G.; Bahloul, M.; Urushigawa, Y.; Sato, S.; Matsumura, A. Petroleum Hydrocarbons and Related Heavy Metals in the Near-Shore Marine Sediments of Kuwait. *Environ. Int.* **1997**, *23* (1), 115–121. [https://doi.org/10.1016/S0160-4120\(96\)00082-7](https://doi.org/10.1016/S0160-4120(96)00082-7).
- (273) Pini, R.; Pedron, F.; Petruzzelli, G.; Scatena, M.; Guidi, G. V. Modifications of the Structural Characteristics of New Soil Forming on Industrial Waste Colonized by Woody Plants. *Geoderma* **2009**, *149* (3–4), 373–378. <https://doi.org/10.1016/J.GEODERMA.2008.12.017>.
- (274) Nujkić, M. M.; Dimitrijević, M. M.; Alagić, S. Č.; Tošić, S. B.; Petrović, J. V. Impact of Metallurgical Activities on the Content of Trace Elements in the Spatial Soil and Plant Parts of *Rubus Fruticosus* L. *Environ. Sci. Process. Impacts* **2016**, *18* (3), 350–360. <https://doi.org/10.1039/C5EM00646E>.
- (275) United States Environmental Protection Agency (EPA). *Priority Pollutant List*, 2014.
- (276) Diepens, N. J.; Buffan-Dubau, E.; Budzinski, H.; Kallerhoff, J.; Merlina, G.; Silvestre, J.; Aubry, I.; Nathalie Tapie; Elger, A. Toxicity Effects of an Environmental Realistic Herbicide Mixture on the Seagrass *Zostera Noltei*. *Environ. Pollut.* **2017**, *222*, 393–403. <https://doi.org/10.1016/j.envpol.2016.12.021>.
- (277) Barlow, R. G.; Cummings, D. G.; Gibb, S. W. Improved Resolution of Mono- and Divinyl Chlorophylls a and b and Zeaxanthin and Lutein in Phytoplankton Extracts Using Reverse Phase C-8 HPLC. *Mar. Ecol. Prog. Ser.* **1996**, *161*, 303–307.
- (278) Harold Hotelling. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.* **1933**, *24* (6), 417–441. <https://doi.org/10.1037/h0071325>.
- (279) Pearson, K. On Lines and Planes of Closest Fit to Systems of Points in Space. *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **1901**, *2* (11), 559–572. <https://doi.org/10.1080/14786440109462720>.
- (280) Jolliffe, I. T. *Principal Component Analysis*; Springer Series in Statistics; Springer-Verlag New York Berlin Heidelberg, 2002. <https://doi.org/10.1007/b98835>.
- (281) Milton, E. J. Principles of Field Spectroscopy. *Int. J. Remote Sens.* **1987**, *8* (12), 1807–1827. <https://doi.org/https://doi.org/10.1080/01431168708954818>.
- (282) Gomez, C.; Lagacherie, P.; Coulouma, G. *Continuum* Removal versus PLSR Method for Clay and Calcium Carbonate Content Estimation from Laboratory and Airborne Hyperspectral Measurements. *Geoderma* **2008**, *148* (2), 141–148. <https://doi.org/10.1016/j.geoderma.2008.09.016>.
- (283) Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36* (8), 1627–1639. <https://doi.org/10.1021/ac60214a047>.
- (284) Dormann, C. F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J. R. G.; Gruber, B.; Lafourcade, B.; Leitão, P. J.; *et al.* Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance. *Ecography (Cop.)*. **2013**, *36*, 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.

- (285) Chakraborty, S.; Weindorf, D. C.; Deb, S.; Li, B.; Paul, S.; Choudhury, A.; Ray, D. P. Rapid Assessment of Regional Soil Arsenic Pollution Risk via Diffuse Reflectance Spectroscopy. *Geoderma* **2017**, *289*, 72–81. <https://doi.org/10.1016/J.GEODERMA.2016.11.024>.
- (286) Kendall, M. G.; Babington Smith, B. The Problem of m Rankings. *Ann. Math. Stat.* **1939**, *10* (3), 275–287. <https://doi.org/10.1214/aoms/1177732186>.
- (287) Elger, A.; Barrat-Segretain, M. H. Plant Palatability Can Be Inferred from a Single-Date Feeding Trial. *Funct. Ecol.* **2004**, *18* (3), 483–488. <https://doi.org/10.1111/j.0269-8463.2004.00846.x>.
- (288) Legendre, P. Species Associations: The Kendall Coefficient of Concordance Revisited. *J. Agric. Biol. Environ. Stat.* **2005**, *10* (2), 226–245. <https://doi.org/10.1198/108571105X46642>.
- (289) Hoerl, A. E.; Kennard, R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12* (1), 55–67.
- (290) Tuia, D.; Flamary, R.; Barlaud, M. Nonconvex Regularization in Remote Sensing. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54* (11), 6470–6480. <https://doi.org/10.1109/TGRS.2016.2585201>.
- (291) Zhang, X.; Hu, B.; Ma, X.; Xu, L. Resting-State Whole-Brain Functional Connectivity Networks for MCI Classification Using L2-Regularized Logistic Regression. *IEEE Trans. Nanobioscience* **2015**, *14* (2), 237–247. <https://doi.org/10.1109/TNB.2015.2403274>.
- (292) Erudel, T.; Fabre, S.; Houet, T.; Mazier, F.; Briottet, X. Criteria Comparison for Classifying Peatland Vegetation Types Using *in situ* Hyperspectral Measurements. *Remote Sens.* **2017**, *9* (748). <https://doi.org/10.3390/rs9070748>.
- (293) Seabold, S.; Perktold, J. *Statsmodels: Econometric and Statistical Modeling with Python*; 2010.
- (294) Oliphant, T. E. Python for Scientific Computing. *Comput. Sci. Eng.* **2007**, *9* (3), 10–20. <https://doi.org/10.1109/MCSE.2007.58>.
- (295) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; *et al.* Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- (296) Kim, M. S.; Daughtry, C. S. T.; Chappelle, E. W.; McMurtrey, J. E.; Walthall, C. L. The Use of High Spectral Resolution Bands for Estimating Absorbed Photosynthetically Active Radiation (Apar). In *CNES, Proceedings of 6th International Symposium on Physical Measurements and Signatures in Remote Sensing*; 1994; pp 299–306.
- (297) Sims, D. A.; Luo, H.; Hastings, S.; Oechel, W. C.; Rahman, A. F.; Gamon, J. A. Parallel Adjustments in Vegetation Greenness and Ecosystem CO₂ Exchange in Response to Drought in a Southern California Chaparral Ecosystem. *Remote Sens. Environ.* **2006**, *103* (3), 289–303. <https://doi.org/10.1016/j.rse.2005.01.020>.
- (298) Carter, G. A. Ratios of Leaf Reflectances in Narrow Wavebands as Indicators of Plant Stress. *Int. J. Remote Sens.* **1994**, *15* (3), 517–520. <https://doi.org/10.1080/01431169408954109>.
- (299) Gitelson, A. A.; Merzlyak, M. N. Remote Estimation of Chlorophyll Content in Higher Plant Leaves. *Int. J. Remote Sens.* **1997**, *18* (12), 2691–2697.

- (300) Dash, J.; Curran, P. J. Evaluation of the MERIS Terrestrial Chlorophyll Index (MTCI). *Adv. Sp. Res.* **2007**, *39* (1), 100–104. <https://doi.org/10.1016/j.asr.2006.02.034>.
- (301) Peñuelas, J.; Gamon, J. A.; Fredeen, A. L.; Merino, J.; Field, C. B. Reflectance Indices Associated with Physiological Changes in Nitrogen- and Water-Limited Sunflower Leaves. *Remote Sens. Environ.* **1994**, *48*, 135–146. [https://doi.org/10.1016/0034-4257\(94\)90136-8](https://doi.org/10.1016/0034-4257(94)90136-8).
- (302) Rondeaux, G.; Steven, M.; Baret, F. Optimization of Soil-Adjusted Vegetation Indices. *Remote Sens. Environ.* **1996**, *55* (2), 95–107. [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7).
- (303) Merzlyak, M. N.; Gitelson, A. A.; Chivkunova, O. B.; Rakitin, V. Y. Non-Destructive Optical Detection of Leaf Senescence and Fruit Ripening. *Physiol. Plant.* **1999**, *106* (1), 135–141.
- (304) Haboudane, D.; Miller, J. R.; Tremblay, N.; Zarco-Tejada, P. J.; Dextraze, L. Integrated Narrow-Band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [https://doi.org/10.1016/S0034-4257\(02\)00018-4](https://doi.org/10.1016/S0034-4257(02)00018-4).
- (305) Vogelmann, J. E.; Rock, B. N.; Moss, D. M. Red Edge Spectral Measurements from Sugar Maple Leaves. *Int. J. Remote Sens.* **1993**, *14* (8), 1563–1575. <https://doi.org/10.1080/01431169308953986>.
- (306) Zarco-Tejada, P. J.; Miller, J. R.; Noland, T. L.; Mohammed, G. H.; Sampson, P. H. Scaling-up and Model Inversion Methods with Narrowband Optical Indices for Chlorophyll Content Estimation in Closed Forest Canopies with Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39* (7), 1491–1507. <https://doi.org/10.1109/36.934080>.
- (307) Dorrington, V. H.; Pyatt, F. B. Some Aspects of Tissue Accumulation and Tolerance to Available Heavy Metal Ions by *Rubus Fruticosus* L., A Colonizer of Spoil Tips in S.W. England. *Int. J. Environ. Stud.* **1983**, *20*, 229–237. <https://doi.org/10.1080/00207238308710039>.
- (308) Yoon, J.; Cao, X.; Zhou, Q.; Ma, L. Q. Accumulation of Pb, Cu, and Zn in Native Plants Growing on a Contaminated Florida Site. *Sci. Total Environ.* **2006**, *368* (2–3), 456–464. <https://doi.org/10.1016/j.scitotenv.2006.01.016>.
- (309) Hagemeyer, J. Ecophysiology of Plant Growth Under Heavy Metal Stress. In *Heavy Metal Stress in Plants*; Springer Berlin Heidelberg, 1999. https://doi.org/10.1007/978-3-662-07745-0_8.
- (310) Garg, N.; Singla, P. Arsenic Toxicity in Crop Plants: Physiological Effects and Tolerance Mechanisms. *Environ. Chem. Lett.* **2011**, *9* (3), 303–321. <https://doi.org/10.1007/s10311-011-0313-7>.
- (311) Nie, M.; Xian, N.; Fu, X.; Chen, X.; Li, B. The Interactive Effects of Petroleum-Hydrocarbon Spillage and Plant Rhizosphere on Concentrations and Distribution of Heavy Metals in Sediments in the Yellow River Delta, China. *J. Hazard. Mater.* **2010**, *174* (1–3), 156–161. <https://doi.org/10.1016/j.jhazmat.2009.09.030>.
- (312) T., S. K.; J., M. A. W.; I., P. G. Bioavailability of Hydrophobic Organic Contaminants in Soils: Fundamental Concepts and Techniques for Analysis. *Eur. J. Soil Sci.* **2003**, *54* (4), 809–818. <https://doi.org/doi:10.1046/j.1351-0754.2003.0564.x>.
- (313) Su, Y.-H.; Zhu, Y.-G. Uptake of Selected PAHs from Contaminated Soils by Rice Seedlings (*Oryza Sativa*) and Influence of Rhizosphere on PAH Distribution. *Environ. Pollut.* **2008**, *155* (2), 359–365. <https://doi.org/10.1016/j.envpol.2007.11.008>.

- (314) Tao, S.; Cui, Y. H.; Xu, F. L.; Li, B. G.; Cao, J.; Liu, W. X.; Schmitt, G.; Wang, X. J.; Shen, W. R.; Qing, B. P.; *et al.* Polycyclic Aromatic Hydrocarbons (PAHs) in Agricultural Soil and Vegetables from Tianjin. *Sci. Total Environ.* **2004**, *320*, 11–24. [https://doi.org/10.1016/S0048-9697\(03\)00453-4](https://doi.org/10.1016/S0048-9697(03)00453-4).
- (315) Parrish, Z. D.; White, J. C.; Isleyen, M.; Gent, M. P. N.; Iannucci-Berger, W.; Eitzer, B. D.; Kelsey, J. W.; Mattina, M. I. Accumulation of Weathered Polycyclic Aromatic Hydrocarbons (PAHs) by Plant and Earthworm Species. *Chemosphere* **2006**, *64* (4), 609–618. <https://doi.org/10.1016/j.chemosphere.2005.11.003>.
- (316) Alonso, R.; Elvira, S.; Castillo, F. J.; Gimeno, B. S. Interactive Effects of Ozone and Drought Stress on Pigments and Activities of Antioxidative Enzymes in *Pinus Halepensis*. *Plant, Cell Environ.* **2001**, *24* (9), 905–916. <https://doi.org/10.1046/j.0016-8025.2001.00738.x>.
- (317) Galmés, J.; Abadía, A.; Cifre, J.; Medrano, H.; Flexas, J. Photoprotection Processes under Water Stress and Recovery in Mediterranean Plants with Different Growth Forms and Leaf Habits. *Physiol. Plant.* **2007**, *130* (4), 495–510. <https://doi.org/10.1111/j.1399-3054.2007.00919.x>.
- (318) Shvaleva, A. L.; Silva, F. C. E.; Breia, E.; Jouve, J.; Hausman, J. F.; Almeida, M. H.; Maroco, J. P.; Rodrigues, M. L.; Pereira, J. S.; Chaves, M. M. Metabolic Responses to Water Deficit in Two *Eucalyptus Globulus* Clones with Contrasting Drought Sensitivity. *Tree Physiol.* **2006**, *26* (2), 239–248. <https://doi.org/10.1093/treephys/26.2.239>.
- (319) Boelman, N. T.; Magney, T. S.; Logan, B. A.; Griffin, K. L.; Eitel, J. U. H.; Greaves, H.; Prager, C. M.; Vierling, L. A. Spectral Determination of Concentrations of Functionally Diverse Pigments in Increasingly Complex Arctic Tundra Canopies. *Oecologia* **2016**, *182* (1), 85–97. <https://doi.org/10.1007/s00442-016-3646-x>.
- (320) Xiao, Y.; Zhao, W.; Zhou, D.; Gong, H. Sensitivity Analysis of Vegetation Reflectance to Biochemical and Biophysical Variables at Leaf, Canopy, and Regional Scales. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52* (7), 4014–4024. <https://doi.org/10.1109/TGRS.2013.2278838>.
- (321) Ogrí, O. R. A Review of the Nigerian Petroleum Industry and the Associated Environmental Problems. *Environmentalist* **2001**, *21*, 11–21. <https://doi.org/https://doi.org/10.1023/A:1010633903226>.
- (322) Romero, I. C.; Toro-Farmer, G.; Diercks, A. R.; Schwing, P.; Muller-Karger, F.; Murawski, S.; Hollander, D. J. Large-Scale Deposition of Weathered Oil in the Gulf of Mexico Following a Deep-Water Oil Spill. *Environ. Pollut.* **2017**, *228*, 179–189. <https://doi.org/10.1016/j.envpol.2017.05.019>.
- (323) Bejarano, A. C.; Michel, J. Oil Spills and Their Impacts on Sand Beach Invertebrate Communities: A Literature Review. *Environ. Pollut.* **2016**, *218*, 709–722. <https://doi.org/10.1016/j.envpol.2016.07.065>.
- (324) Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Erudel, T.; Buffan-Dubau, E.; Dubucq, D.; Elger, A. Detection and Discrimination of Various Oil-Contaminated Soils Using Vegetation Reflectance. *Sci. Total Environ.* **2019**, *655*, 1113–1124. <https://doi.org/10.1016/j.scitotenv.2018.11.314>.
- (325) Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Application of PROSPECT for Estimating Total Petroleum Hydrocarbons in Contaminated Soils from Leaf Optical Properties. *J. Hazard. Mater.* **2019**, *377*, 409–417. <https://doi.org/10.1016/j.jhazmat.2019.05.093>.

- (326) Achard, V.; Fabre, S.; Alakian, A.; Dubucq, D.; Déliot, P. Direct or Indirect On-Shore Hydrocarbon Detection Methods Applied to Hyperspectral Data in Tropical Area. In *Earth Resources and Environmental Remote Sensing/GIS Applications IX*; Michel, U., Schulz, K., Eds.; SPIE, 2018; p 22. <https://doi.org/10.1117/12.2325097>.
- (327) Tran, T. H.; Gati, E. M.; Eshel, A.; Winter, G. Germination, Physiological and Biochemical Responses of Acacia Seedlings (*Acacia Raddiana* and *Acacia Tortilis*) to Petroleum Contaminated Soils. *Environ. Pollut.* **2018**, *234*, 642–655. <https://doi.org/10.1016/j.envpol.2017.11.067>.
- (328) Lin, Q.; Mendelsohn, I. A.; Suidan, M. T.; Lee, K.; Venosa, A. D. The Dose-Response Relationship between No. 2 Fuel Oil and the Growth of the Salt Marsh Grass, *Spartina Alterniflora*. *Mar. Pollut. Bull.* **2002**, *44* (9), 897–902. [https://doi.org/10.1016/S0025-326X\(02\)00118-2](https://doi.org/10.1016/S0025-326X(02)00118-2).
- (329) Kirk, J. L.; Klironomos, J. N.; Lee, H.; Trevors, J. T. Phytotoxicity Assay to Assess Plant Species for Phytoremediation of Petroleum-Contaminated Soil. *Bioremediat. J.* **2002**, *6* (1), 57–63. <https://doi.org/10.1080/10889860290777477>.
- (330) Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning The Elements of Statistical Learning*. **2017**.
- (331) L. Breiman. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. <https://doi.org/https://doi.org/10.1023/A:1010933404324>.
- (332) Hutengs, C.; Vohland, M. Downscaling Land Surface Temperatures at Regional Scales with Random Forest Regression. *Remote Sens. Environ.* **2016**, *178*, 127–141. <https://doi.org/10.1016/j.rse.2016.03.006>.
- (333) Mutanga, O.; Adam, E.; Cho, M. A. High Density Biomass Estimation for Wetland Vegetation Using Worldview-2 Imagery and Random Forest Regression Algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18* (1), 399–406. <https://doi.org/10.1016/j.jag.2012.03.012>.
- (334) Grömping, U. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *Am. Stat.* **2009**, *63* (4), 308–319. <https://doi.org/10.1198/tast.2009.08199>.
- (335) Salanitro, J. P.; Dorn, P. B.; Huesemann, M. H.; Moore, K. O.; Rhodes, I. A.; Rice Jackson, L. M.; Vipond, T. E.; Western, M. M.; Wisniewski, H. L. Crude Oil Hydrocarbon Bioremediation and Soil Ecotoxicity Assessment. *Environ. Sci. Technol.* **1997**, *31* (6), 1769–1776. <https://doi.org/10.1021/es960793i>.
- (336) Jha, M.; Levy, J.; Gao, Y. Advances in Remote Sensing for Oil Spill Disaster Management: State-of-the-Art Sensors Technology for Oil Spill Surveillance. *Sensors* **2008**, *8* (1), 236–255. <https://doi.org/10.3390/s8010236>.
- (337) Perrodin, Y.; Boillot, C.; Angerville, R.; Donguy, G.; Emmanuel, E. Ecological Risk Assessment of Urban and Industrial Systems: A Review. *Sci. Total Environ.* **2011**, *409* (24), 5162–5176. <https://doi.org/10.1016/j.scitotenv.2011.08.053>.
- (338) Khanna, S.; Santos, M.; Ustin, S.; Shapiro, K.; Haverkamp, P.; Lay, M. Comparing the Potential of Multispectral and Hyperspectral Data for Monitoring Oil Spill Impact. *Sensors* **2018**, *18* (2), 558. <https://doi.org/10.3390/s18020558>.
- (339) Scafutto, R. D. P. M.; de Souza Filho, C. R.; Rivard, B. Characterization of Mineral Substrates Impregnated with Crude Oils Using Proximal Infrared Hyperspectral Imaging. *Remote Sens. Environ.* **2016**, *179*, 116–130. <https://doi.org/10.1016/j.rse.2016.03.033>.

- (340) Zhu, L.; Zhao, X.; Lai, L.; Wang, J.; Jiang, L.; Ding, J.; Liu, N.; Yu, Y.; Li, J.; Xiao, N.; *et al.* Soil TPH Concentration Estimation Using Vegetation Indices in an Oil Polluted Area of Eastern China. *PLoS One* **2013**, *8* (1). <https://doi.org/10.1371/journal.pone.0054028>.
- (341) Zhou, C.; Chen, S.; Zhang, Y.; Zhao, J.; Song, D.; Liu, D. Evaluating Metal Effects on the Reflectance Spectra of Plant Leaves during Different Seasons in Post-Mining Areas, China. *Remote Sens.* **2018**, *10* (1211). <https://doi.org/10.3390/rs10081211>.
- (342) Demarez, V.; Gastellu-Etchegorry, V.; Mougín, E.; Marty, G.; Proisy, C.; Dufrêne, E.; Ledantec, V. Seasonal Variation of Leaf Chlorophyll Content of a Temperate Forest. Inversion of the Prospect Model. *Int. J. Remote Sens.* **1999**, *20* (5), 879–894. <https://doi.org/10.1080/014311699212975>.
- (343) Gitelson, A. A.; Merzlyak, M. N.; Lichtenthaler, H. K. Detection of Red Edge Position and Chlorophyll Content by Reflectance Measurements near 700 nm. *J. Plant Physiol.* **1996**, *148* (3–4), 501–508. [https://doi.org/10.1016/S0176-1617\(96\)80285-9](https://doi.org/10.1016/S0176-1617(96)80285-9).
- (344) Verrelst, J.; Camps-Valls, G.; Muñoz-Marí, J.; Rivera, J. P.; Veroustraete, F.; Clevers, J. G. P. W.; Moreno, J. Optical Remote Sensing and the Retrieval of Terrestrial Vegetation Biophysical Properties - A Review. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 273–290. <https://doi.org/10.1016/j.isprsjprs.2015.05.005>.
- (345) Li, P.; Wang, Q. Retrieval of Leaf Biochemical Parameters Using PROSPECT Inversion: A New Approach for Alleviating Ill-Posed Problems. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49* (7), 2499–2506. <https://doi.org/10.1109/TGRS.2011.2109390>.
- (346) Sonobe, R.; Sano, T.; Horie, H. Using Spectral Reflectance to Estimate Leaf Chlorophyll Content of Tea with Shading Treatments. *Biosyst. Eng.* **2018**, *175*, 168–182. <https://doi.org/10.1016/j.biosystemseng.2018.09.018>.
- (347) Sun, J.; Shi, S.; Yang, J.; Du, L.; Gong, W.; Chen, B.; Song, S. Analyzing the Performance of PROSPECT Model Inversion Based on Different Spectral Information for Leaf Biochemical Properties Retrieval. *ISPRS J. Photogramm. Remote Sens.* **2018**, *135*, 74–83. <https://doi.org/10.1016/j.isprsjprs.2017.11.010>.
- (348) Qiu, F.; Chen, J. M.; Ju, W.; Wang, J.; Zhang, Q.; Fang, M. Improving the PROSPECT Model to Consider Anisotropic Scattering of Leaf Internal Materials and Its Use for Retrieving Leaf Biomass in Fresh Leaves. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56* (6), 3119–3136. <https://doi.org/10.1109/TGRS.2018.2791930>.
- (349) Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C. Inversion of a Radiative Transfer Model for Estimating Vegetation LAI and Chlorophyll in a Heterogeneous Grassland. *Remote Sens. Environ.* **2008**, *112* (5), 2592–2604. <https://doi.org/10.1016/j.rse.2007.12.003>.
- (350) Zarco-Tejada, P. J.; Miller, J. R.; Harron, J.; Hu, B.; Noland, T. L.; Goel, N.; Mohammed, G. H.; Sampson, P. Needle Chlorophyll Content Estimation through Model Inversion Using Hyperspectral Data from Boreal Conifer Forest Canopies. *Remote Sens. Environ.* **2004**, *89* (2), 189–199. <https://doi.org/10.1016/j.rse.2002.06.002>.
- (351) Altmann, Y.; Halimi, A.; Dobigeon, N.; Tournier, J. Y. Supervised Nonlinear Spectral Unmixing Using a Postnonlinear Mixing Model for Hyperspectral Imagery. *IEEE Trans. Image Process.* **2012**, *21* (6), 3017–3025. <https://doi.org/10.1109/TIP.2012.2187668>.
- (352) Jacquemoud, S.; Baret, F.; Andrieu, B.; Danson, F. M.; Jaggard, K. Extraction of Vegetation Biophysical Parameters by Inversion of the PROSPECT + SAIL Models on Sugar Beet Canopy Reflectance Data. Application to TM and AVIRIS Sensors. *Remote Sens. Environ.* **1995**, *52*, 163–172.

- (353) Storn, R.; Price, K. Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *J. of Global Optim.* **1997**, *11*, 341–359. <https://doi.org/https://doi.org/10.1023/A:1008202821328>.
- (354) Das, S.; Mullick, S. S.; Suganthan, P. N. Recent Advances in Differential Evolution-An Updated Survey. *Swarm Evol. Comput.* **2016**, *27*, 1–30. <https://doi.org/10.1016/j.swevo.2016.01.004>.
- (355) Ghosh, A.; Datta, A.; Ghosh, S. Self-Adaptive Differential Evolution for Feature Selection in Hyperspectral Image Data. *Appl. Soft Comput. J.* **2013**, *13* (4), 1969–1977. <https://doi.org/10.1016/j.asoc.2012.11.042>.
- (356) Zhong, Y.; Zhao, L.; Zhang, L. An Adaptive Differential Evolution Endmember Extraction Algorithm for Hyperspectral Remote Sensing Imagery. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11* (6), 1061–1065. <https://doi.org/10.1109/LGRS.2013.2285476>.
- (357) Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. In *Technometrics*; Palgrave Macmillan UK: London, 1969; Vol. 11, pp 137–148. https://doi.org/10.1007/978-1-349-95810-8_109.
- (358) Shi, T.; Liu, H.; Chen, Y.; Wang, J.; Wu, G. Estimation of Arsenic in Agricultural Soils Using Hyperspectral Vegetation Indices of Rice. *J. Hazard. Mater.* **2016**, *308*, 243–252. <https://doi.org/10.1016/j.jhazmat.2016.01.022>.
- (359) Kawamura, K.; Tsujimoto, Y.; Rabenarivo, M.; Asai, H.; Andriamananjara, A.; Rakotoson, T. Vis-NIR Spectroscopy and PLS Regression with Waveband Selection for Estimating the Total C and N of Paddy Soils in Madagascar. *Remote Sens.* **2017**, *9* (10). <https://doi.org/10.3390/rs9101081>.
- (360) Benyó, D.; Horváth, E.; Németh, E.; Leviczky, T.; Takács, K.; Lehotai, N.; Feigl, G.; Kolbert, Z.; Ördög, A.; Gallé, R.; *et al.* Physiological and Molecular Responses to Heavy Metal Stresses Suggest Different Detoxification Mechanism of *Populus Deltoides* and *P. x Canadensis*. *J. Plant Physiol.* **2016**, *201*, 62–70. <https://doi.org/10.1016/j.jplph.2016.05.025>.
- (361) Liu, L.; Song, B.; Zhang, S.; Liu, X. A Novel Principal Component Analysis Method for the Reconstruction of Leaf Reflectance Spectra and Retrieval of Leaf Biochemical Contents. *Remote Sens.* **2017**, *9* (11). <https://doi.org/10.3390/rs9111113>.
- (362) Li, Y.; He, N.; Hou, J.; Xu, L.; Liu, C.; Zhang, J.; Wang, Q.; Zhang, X.; Wu, X. Factors Influencing Leaf Chlorophyll Content in Natural Forests at the Biome Scale. *Front. Ecol. Evol.* **2018**, *6* (June), 1–10. <https://doi.org/10.3389/fevo.2018.00064>.
- (363) Martin, R. E.; Asner, G. P.; Sack, L. Genetic Variation in Leaf Pigment, Optical and Photosynthetic Function among Diverse Phenotypes of *Metrosideros Polymorpha* Grown in a Common Garden. *Oecologia* **2007**, *151* (3), 387–400. <https://doi.org/10.1007/s00442-006-0604-z>.
- (364) Jones, N. F.; Pejchar, L.; Kiesecker, J. M. The Energy Footprint: How Oil, Natural Gas, and Wind Energy Affect Land for Biodiversity and the Flow of Ecosystem Services. *Bioscience* **2015**, *65* (3), 290–301. <https://doi.org/10.1093/biosci/biu224>.
- (365) Onyia, N.; Balzter, H.; Berrio, J.-C. Normalized Difference Vegetation Vigour Index: A New Remote Sensing Approach to Biodiversity Monitoring in Oil Polluted Regions. *Remote Sens.* **2018**, *10* (6), 897. <https://doi.org/10.3390/rs10060897>.
- (366) Dubois-Fernandez, P.; Ruault du Plessis, O.; Arnaubec, A.; Angelliaume, S.; Baque, R.; Bonin, G.; Briottet, X.; Cantalloube, H.; Chervet, P.; Coulombeix, C.; *et al.* The SETHI

Remote Sensing Airborne Platform and the Related Science Activities. *Rev. Française Photogrammétrie Télédétection* **2014**, *200*, 36–47.

- (367) Brigot, G.; Colin-Koeniguer, E.; Plyer, A.; Janez, F. Adaptation and Evaluation of an Optical Flow Method Applied to Coregistration of Forest Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9* (7), 2923–2939. <https://doi.org/10.1109/JSTARS.2016.2578362>.
- (368) Smith, G. M.; Milton, E. J. The Use of the Empirical Line Method to Calibrate Remotely Sensed Data to Reflectance. *Int. J. Remote Sens.* **1999**, *20* (13), 2653–2662. <https://doi.org/10.1080/014311699211994>.
- (369) Roberts, D. A.; Smith, M. O.; Adams, J. B. Green Vegetation, Nonphotosynthetic Vegetation, and Soils in AVIRIS Data. *Remote Sens. Environ.* **1993**, *44* (2–3), 255–269. [https://doi.org/10.1016/0034-4257\(93\)90020-X](https://doi.org/10.1016/0034-4257(93)90020-X).
- (370) Zarco-Tejada, P. J.; Camino, C.; Beck, P. S. A.; Calderon, R.; Hornero, A.; Hernández-Clemente, R.; Kattenborn, T.; Montes-Borrego, M.; Susca, L.; Morelli, M.; *et al.* Previsual Symptoms of *Xylella Fastidiosa* Infection Revealed in Spectral Plant-Trait Alterations. *Nat. Plants* **2018**, *4*, 432–439. <https://doi.org/10.1038/s41477-018-0189-7>.
- (371) Botha, E. J.; Leblon, B.; Zebarth, B.; Watmough, J. Non-Destructive Estimation of Potato Leaf Chlorophyll from Canopy Hyperspectral Reflectance Using the Inverted PROSAIL Model. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9* (4), 360–374. <https://doi.org/10.1016/j.jag.2006.11.003>.
- (372) Si, Y.; Schlerf, M.; Zurita-Milla, R.; Skidmore, A.; Wang, T. Mapping Spatio-Temporal Variation of Grassland Quantity and Quality Using MERIS Data and the PROSAIL Model. *Remote Sens. Environ.* **2012**, *121*, 415–425. <https://doi.org/10.1016/j.rse.2012.02.011>.
- (373) Balandier, P.; Marquier, A.; Casella, E.; Kiewitt, A.; Coll, L.; Wehrlen, L.; Harmer, R. Architecture, Cover and Light Interception by Bramble (*Rubus Fruticosus*): A Common Understorey Weed in Temperate Forests. *Forestry* **2013**, *86*, 39–46. <https://doi.org/10.1093/forestry/cps066>.
- (374) Hong, Y.; Chen, Y.; Yu, L.; Liu, Y.; Liu, Y.; Zhang, Y.; Liu, Y.; Cheng, H. Combining Fractional Order Derivative and Spectral Variable Selection for Organic Matter Estimation of Homogeneous Soil Samples by VIS-NIR Spectroscopy. *Remote Sens.* **2018**, *10* (3). <https://doi.org/10.3390/rs10030479>.
- (375) Prospere, K.; McLaren, K.; Wilson, B. Plant Species Discrimination in a Tropical Wetland Using *in situ* Hyperspectral Data. *Remote Sens.* **2014**, *6* (9), 8494–8523. <https://doi.org/10.3390/rs6098494>.
- (376) Peng, X.; Shi, T.; Song, A.; Chen, Y.; Gao, W. Estimating Soil Organic Carbon Using VIS/NIR Spectroscopy with SVMR and SPA Methods. *Remote Sens.* **2014**, *6* (4), 2699–2717. <https://doi.org/10.3390/rs6042699>.
- (377) Broge, N. H.; Leblanc, E. Comparing Prediction Power and Stability of Broadband and Hyperspectral Vegetation Indices for Estimation of Green Leaf Area Index and Canopy Chlorophyll Density. *Remote Sens. Environ.* **2000**, *76*, 156–172. [https://doi.org/10.1016/S0034-4257\(00\)00197-8](https://doi.org/10.1016/S0034-4257(00)00197-8).
- (378) Verrelst, J.; Schaepman, M. E.; Koetz, B.; Kneubühler, M. Angular Sensitivity Analysis of Vegetation Indices Derived from CHRIS/PROBA Data. *Remote Sens. Environ.* **2008**, *112* (5), 2341–2353. <https://doi.org/10.1016/j.rse.2007.11.001>.

- (379) Immitzer, M.; Atzberger, C.; Koukal, T. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sens.* **2012**, *4* (9), 2661–2693. <https://doi.org/10.3390/rs4092661>.
- (380) Lu, B.; He, Y. Species Classification Using Unmanned Aerial Vehicle (UAV)-Acquired High Spatial Resolution Imagery in a Heterogeneous Grassland. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 73–85. <https://doi.org/10.1016/j.isprsjprs.2017.03.011>.
- (381) Roth, K. L.; Dennison, P. E.; Roberts, D. A. Comparing Endmember Selection Techniques for Accurate Mapping of Plant Species and Land Cover Using Imaging Spectrometer Data. *Remote Sens. Environ.* **2012**, *127*, 139–152. <https://doi.org/10.1016/j.rse.2012.08.030>.
- (382) Stagakis, S.; Vanikiotis, T.; Sykioti, O. Estimating Forest Species Abundance through Linear Unmixing of CHRIS/PROBA Imagery. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 79–89. <https://doi.org/10.1016/j.isprsjprs.2016.05.013>.
- (383) Dehaan, R.; Louis, J.; Wilson, A.; Hall, A.; Rumbachs, R. Discrimination of Blackberry (*Rubus Fruticosus* Sp. Agg.) Using Hyperspectral Imagery in Kosciuszko National Park, NSW, Australia. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62* (1), 13–24. <https://doi.org/10.1016/j.isprsjprs.2007.01.004>.
- (384) Dillen, S. Y.; de Beek, M. O.; Hufkens, K.; Buonanduci, M.; Phillips, N. G. Seasonal Patterns of Foliar Reflectance in Relation to Photosynthetic Capacity and Color Index in Two Co-Occurring Tree Species, *Quercus Rubra* and *Betula Papyrifera*. *Agric. For. Meteorol.* **2012**, *160*, 60–68. <https://doi.org/10.1016/j.agrformet.2012.03.001>.
- (385) Guanter, L.; Kaufmann, H.; Segl, K.; Foerster, S.; Rogass, C.; Chabrillat, S.; Kuester, T.; Hollstein, A.; Rossner, G.; Chlebek, C.; *et al.* The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. *Remote Sens.* **2015**, *7* (7), 8830–8857. <https://doi.org/10.3390/rs70708830>.
- (386) Transon, J.; D'Andrimont, R.; Maignard, A.; Defourny, P. Survey of Hyperspectral Earth Observation Applications from Space in the Sentinel-2 Context. *Remote Sens.* **2018**, *10* (157). <https://doi.org/10.3390/rs10020157>.
- (387) Colomina, I.; Molina, P. Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. <https://doi.org/10.1016/j.isprsjprs.2014.02.013>.
- (388) Zhang, C.; Kovacs, J. M. The Application of Small Unmanned Aerial Systems for Precision Agriculture: A Review. *Precis. Agric.* **2012**, *13* (6), 693–712. <https://doi.org/10.1007/s11119-012-9274-5>.
- (389) Albetis, J.; Jacquin, A.; Goulard, M.; Poilvé, H.; Rousseau, J.; Clenet, H.; Dedieu, G.; Duthoit, S. On the Potentiality of UAV Multispectral Imagery to Detect Flavescence Dorée and Grapevine Trunk Diseases. *Remote Sens.* **2019**, *11*. <https://doi.org/10.3390/rs11010023>.
- (390) Candiago, S.; Remondino, F.; De Giglio, M.; Dubbini, M.; Gattelli, M. Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. *Remote Sens.* **2015**, *7*, 4026–4047. <https://doi.org/10.3390/rs70404026>.
- (391) Gastellu-Etchegorry, J. P.; Martin, E.; Gascon, F. DART: A 3D Model for Simulating Satellite Images and Studying Surface Radiation Budget. *Int. J. Remote Sens.* **2004**, *25* (1), 73–96. <https://doi.org/10.1080/0143116031000115166>.
- (392) Inoue, Y.; Sakaiya, E.; Wang, C. Capability of C-Band Backscattering Coefficients from High-Resolution Satellite SAR Sensors to Assess Biophysical Variables in Paddy Rice.

- Remote Sens. Environ.* **2014**, *140*, 257–266. <https://doi.org/10.1016/j.rse.2013.09.001>.
- (393) Notarnicola, C.; Posa, F. Inferring Vegetation Water Content From C- and L-Band SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45* (10), 3165–3171. <https://doi.org/10.1109/TGRS.2007.903698>.
- (394) Solberg, S.; Astrup, R.; Gobakken, T.; Næsset, E.; Weydahl, D. J. Estimating Spruce and Pine Biomass with Interferometric X-Band SAR. *Remote Sens. Environ.* **2010**, *114* (10), 2353–2360. <https://doi.org/10.1016/j.rse.2010.05.011>.
- (395) Pinel, V.; Raucoules, D. The Contribution of SAR Data to Volcanology and Subsidence Studies. *L. Surf. Remote Sens. Environ. Risks* **2016**, 221–262. <https://doi.org/10.1016/B978-1-78548-105-5.50007-4>.
- (396) Fernandez-Ordóñez, Y.; Soria-Ruiz, J.; Leblo, B. Forest Inventory Using Optical and Radar Remote Sensing. In *Advances in Geoscience and Remote Sensing*; IntechOpen, Ed.; InTech, 2009; pp 539–556. <https://doi.org/10.5772/8330>.
- (397) Yi Hyun Kim; Jackson, T.; Bindlish, R.; Hoonyol Lee; Sukyoung Hong. Radar Vegetation Index for Estimating the Vegetation Water Content of Rice and Soybean. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9* (4), 564–568. <https://doi.org/10.1109/LGRS.2011.2174772>.
- (398) Joshi, N.; Baumann, M.; Ehammer, A.; Fensholt, R.; Grogan, K.; Hostert, P.; Jepsen, M.; Kuemmerle, T.; Meyfroidt, P.; Mitchard, E.; *et al.* A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Remote Sens.* **2016**, *8* (70). <https://doi.org/10.3390/rs8010070>.
- (399) Duguay, Y.; Bernier, M.; Lévesque, E.; Domine, F.; Duguay, Y.; Bernier, M.; Lévesque, E.; Domine, F. Land Cover Classification in SubArctic Regions Using Fully Polarimetric RADARSAT-2 Data. *Remote Sens.* **2016**, *8* (9), 697. <https://doi.org/10.3390/rs8090697>.
- (400) Ullmann, T.; Schmitt, A.; Roth, A.; Duffe, J.; Dech, S.; Hubberten, H.-W.; Baumhauer, R. Land Cover Characterization and Classification of Arctic Tundra Environments by Means of Polarized Synthetic Aperture X- and C-Band Radar (PolSAR) and Landsat 8 Multispectral Imagery — Richards Island, Canada. *Remote Sens.* **2014**, *6* (9), 8565–8593. <https://doi.org/10.3390/rs6098565>.
- (401) Henderson, F. M.; Lewis, A. J. Radar Detection of Wetland Ecosystems: A Review. *Int. J. Remote Sens.* **2008**, *29* (20), 5809–5835. <https://doi.org/10.1080/01431160801958405>.
- (402) White, L.; Brisco, B.; Daboor, M.; Schmitt, A.; Pratt, A.; White, L.; Brisco, B.; Daboor, M.; Schmitt, A.; Pratt, A. A Collection of SAR Methodologies for Monitoring Wetlands. *Remote Sens.* **2015**, *7* (6), 7615–7645. <https://doi.org/10.3390/rs70607615>.
- (403) Szigarski, C.; Jagdhuber, T.; Baur, M.; Thiel, C.; Parrens, M.; Wigneron, J.-P.; Piles, M.; Entekhabi, D. Analysis of the Radar Vegetation Index and Potential Improvements. *Remote Sens.* **2018**, *10* (11), 1776. <https://doi.org/10.3390/rs10111776>.
- (404) Boisot, O.; Pioch, S.; Fatras, C.; Caulliez, G.; Bringer, A.; Borderies, P.; Lalaurie, J.-C.; Guérin, C.-A. Ka-Band Backscattering from Water Surface at Small Incidence: A Wind-Wave Tank Study. *J. Geophys. Res. Ocean.* **2015**, *120* (5), 3261–3285. <https://doi.org/10.1002/2014JC010338>.
- (405) Zhou, Z.-S.; Boerner, W.-M.; Sato, M. Development of a Ground-Based Polarimetric Broadband SAR System for Noninvasive Ground-Truth Validation in Vegetation Monitoring. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42* (9), 1803–1810. <https://doi.org/10.1109/TGRS.2004.832248>.



THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Institut Supérieur de l'Aéronautique et de l'Espace

Présentée et soutenue par

Guillaume LASSALLE

Le 17 octobre 2019

Exploitation de données hyperspectrales pour l'analyse de l'état
de santé de la végétation exposée aux hydrocarbures

SYNTHÈSE EN FRANÇAIS

Ecole doctorale :

AA - Aéronautique, Astronautique

Spécialités :

**Surfaces et interfaces continentales, Hydrologie
Agrosystèmes, écosystèmes et environnement**

Unité de recherche :

DOTA/POS - Propriétés Optiques des Scènes

Thèse dirigée par :

Arnaud Elger et Sophie Fabre

Jury

M. Stéphane Jacquemoud, Professeur des Universités, Rapporteur

M. Carlos Roberto de Souza Filho, Professeur, Rapporteur

Mme Isabelle Laffont-Schwob, Professeure des Universités, Examinatrice

Mme Agnès Bégué, Directrice de Recherche, Examinatrice

M. Anthony Credoz, Ingénieur de Recherche, Examineur

M. Arnaud Elger, Maître de Conférences, Directeur de thèse

Mme Sophie Fabre, Ingénieure de Recherche, Co-directrice de thèse

TABLE DES MATIERES

Liste des figures.....	7
Liste des tableaux.....	8
Liste des abréviations	9
Introduction générale	11
Objectifs de la thèse	13
Chapitre un : état de l’art de l’utilisation de la télédétection passive hyperspectrale pour l’exploration pétrolière et le monitoring de la contamination en régions végétalisées	15
Introduction	15
1. Les propriétés optiques de la végétation dans le domaine optique réflectif (400 – 2500 nm)	15
1.1. Influence des pigments dans la région du visible (400 – 750 nm)	15
1.2. Influence de l’anatomie des feuilles dans le proche-infrarouge (750 – 1300 nm).....	16
1.3. Influence du contenu en eau et en matière sèche dans le proche-infrarouge (750 – 1300 nm) et l’infrarouge aux courtes longueurs d’ondes (1300 – 2500 nm)	17
2. Effets du pétrole brut et des produits pétroliers sur l’état de santé de la végétation	18
2.1. Composition du pétrole brut et des produits pétroliers ; effets directs et indirects sur la végétation.....	18
2.2. Réponse biophysique et biochimique de la végétation et sources de variabilité.....	19
3. Détection du pétrole brut et des produits pétroliers par exploitation des propriétés optiques de la végétation	21
3.1. Modifications des propriétés optiques observées en conditions contrôlées	21
3.2. Méthodes existantes pour détecter le pétrole brut, les produits pétroliers et d’autres sources de stress de la végétation	22
3.3. Application en imagerie multi- et hyperspectrale aéroportée et satellitaire.....	22
Conclusion.....	24
Problématique scientifique et approche mise en œuvre	25
Chapitre deux : développement de méthodes de détection et de quantification de la contamination pétrolière exploitant les propriétés optiques de la végétation, en conditions contrôlées	27

Introduction	27
4. Premier article : “Assessing soil contamination due to oil and gas production using vegetation hyperspectral reflectance”	28
4.1. Contexte et objectif de l’étude	28
4.2. Principaux résultats	28
4.3. Conclusion et perspectives.....	30
5. Deuxième article : “Detection and discrimination of various oil-contaminated soils using vegetation reflectance”	31
5.1. Contexte et objectif de l’étude	31
5.2. Principaux résultats	31
5.3. Conclusion et perspectives.....	33
6. Troisième article : “Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression”	34
6.1. Contexte et objectif de l’étude	34
6.2. Principaux résultats	34
6.3. Conclusion et perspectives.....	36
Conclusion.....	37
Chapitre trois : application et évaluation des méthodes en conditions naturelles, sur le terrain et les images hyperspectrales aéroportées	39
Introduction	39
7. Quatrième article : “Application of PROSPECT for estimating Total Petroleum Hydrocarbons in contaminated soils from leaf optical properties”.....	40
7.1. Contexte et objectif de l’étude	40
7.2. Principaux résultats	40
7.3. Conclusion et perspectives.....	42
8. Cinquième article : “Toward quantifying oil contamination in vegetated areas using very high spatial and spectral resolution imagery”	43
8.1. Contexte et objectif de l’étude	43
8.2. Principaux résultats	44
8.3. Conclusion et perspectives.....	45
Conclusion.....	46
Conclusion générale	49
Perspectives.....	50
Valorisation de la thèse	53
Publications (revues à comité de lecture).....	53
Communications.....	53

References55

LISTE DES FIGURES

- Figure 1. (a) Représentation schématique d'un micro- et macrosuintement pétrolier (adapté de ⁹). (HC: hydrocarbures.) Les macrosuintements sont caractérisés par des remontées de pétrole apparentes en surface (b). Les microsuintements s'expriment au travers d'altérations de la lithologie (c) et de la répartition de la végétation en surface (d)^{4,10}. 11
- Figure 2. Principe de l'imagerie hyperspectrale (adapté de ¹⁶). Cette technologie fournit des données réflectance spectrale des surfaces sur un spectre continu du domaine optique réflectif (i.e. la signature spectrale). 12
- Figure 3. Signature spectrale d'une feuille de plante en bonne santé. Pour chaque région du spectre, le paramètre biophysique ou biochimique le plus influent est indiqué. 16
- Figure 4. Influence (a) du *Leaf Area Index* (LAI) et (b) de l'orientation moyenne des feuilles (*Mean Leaf Angle*, MLA) sur la réflectance d'un couvert végétal⁴⁹. 17
- Figure 5. Signatures spectrales de plantes en bonne santé (« Healthy ») ou en situation de stress hydrique (« Water-déficient ») (données personnelles). 18
- Figure 6. *Canavalia ensiformis* (L.) DC développée sur sol (a-b, e) contaminé au diesel (22,219 mg.kg⁻¹) ou (c-d) non-contaminé. Les flèches blanches indiquent les racines nécrosées. (f-g) Symptômes de décoloration des feuilles observés sur les plantes exposées au diesel. (h) Comparaison des paramètres anatomiques et du contenu en pigments des feuilles entre les deux traitements (moyenne \pm écart-type). Les différences significatives ($p < 0,05$) sont représentées par des lettres différentes (modifié de ⁷⁷). 20
- Figure 7. Différences de réflectance de *Brachiaria brizantha* H.S. développée sur sol contaminé (a-b) au diesel ou (c-d) à l'essence, à l'échelle (a, c) de la feuille et (b, d) du couvert. Les différences de réflectance ont été obtenues en soustrayant la signature spectrale des plantes témoins à celles de plantes exposées au diesel ou à l'essence. Les courbes d'une même figure représentent différents niveaux de contamination (modifié de ⁸¹). 21
- Figure 8. (a) Microsuintement végétalisé. (b) Le suintement est caractérisé par un mélange de sol nu et de pétrole brut au centre, entouré d'une végétation à la distribution spatiale particulière, le rendant détectable par imagerie hyperspectrale aéroportée à l'aide (c) d'indices de végétation et d'un filtre spatial. Toutefois, des fausses alarmes (pixels rouges en-dehors des cercles jaunes) apparaissent lors de la détection⁴. 23
- Figure 9. Réflectance moyenne de *R. fruticosus* après 61 jours d'exposition aux différents traitements, à l'échelle (a) de la feuille et (b) de l'individu. 29
- Figure 10. Évolution temporelle de la précision de la classification des plantes exposées à la contamination obtenue pour *R. fruticosus*, à l'échelle (a) de la feuille (J1 – J61) et (b) de l'individu (J25 – J61). 30
- Figure 11. Indice MTCI obtenu à l'échelle de la feuille (*Leaf*), de l'individu (*Plant*) et du couvert (*Multi-plant*) pour les différents traitements après 32 jours. Cet indice a été mis en lien avec les teneurs en chlorophylles et en β -carotène des feuilles ($R^2 = 0,94$) et s'est révélé fortement concordants entre les échelles ($W = 0,90$, $p < 0,05$). (cf. article original pour la signification des traitements.) 32
- Figure 12. Biomasse sèche des parties aériennes et contenus en chlorophylles et en caroténoïdes des feuilles (moyenne \pm écart-type) pour les différents traitements après 42 jours. Des lettres différentes indiquent les différences significatives entre traitements (ANOVA et test post-hoc de Tukey, $p <$

0,05). (VAZ : violaxanthine + anthéroxanthine + zéaxanthine ; TPH ; Total Petroleum Hydrocarbons.)	35
Figure 13. Comparaison entre les concentrations en TPH prédites par la méthode et celles mesurées dans les sols, à l'échelle de la feuille et de l'individu.	36
Figure 14. Evolutions temporelles du contenu en chlorophylles des feuilles relatif au témoin pour les différentes espèces présentes sur le borbier.	41
Figure 15. (figures du haut) Relation entre le contenu en chlorophylles des feuilles estimé par inversion de PROSPECT et la concentration en TPH du sol observée sur le lot d'apprentissage, et (figures du bas) comparaison entre les concentrations en TPH prédites et mesurées sur le lot de test.....	42
Figure 16. (a, d) Extraits des images hyperspectrales aéroportées acquises sur les sites (a) contaminés et (b) témoins. Les sites ayant servi à la calibration des méthodes sont illustrés par des polygones. Les cercles correspondent aux sites de validation (en rouge : borbier contaminé, en blanc : témoin). (b, e) Zoom sur (b) le borbier principal (10 parcelles) et (e) le site témoin, utilisés pour calibrer les méthodes. (c, f) <i>R. fruticosus</i> (c) sur le borbier et (f) sur le site témoin (photos prises en mars)...	43
Figure 17. (a-c) Relation entre le contenu en chlorophylles des feuilles et la concentration en TPH dans le sol (lot d'apprentissage) observée sur les 10 parcelles du borbier principal, à l'échelle (a) de la feuille, (b) de la canopée et (c) des images aéroportées. (d-f) Comparaison entre les concentrations en TPH prédites et mesurées (lot de test), à l'échelle (d) de la feuille, (e) de la canopée et (f) des images aéroportées.....	45
Figure 18. (figures du haut) Comparaison entre le contenu en chlorophylles des feuilles (moyenne ± écart-type) estimé sur le terrain (mesures à l'échelle de la feuille) et sur les images aéroportées, et (figures du bas) comparaison entre les concentrations en TPH prédites et mesurées, aux résolutions spectrales des capteurs HypXim et Worldview-3.	46

LISTE DES TABLEAUX

Tableau 1. Matrices de confusion (%) obtenues après 32 jours, à l'échelle de la feuille (<i>Leaf</i>) et de l'individu (<i>Plant</i>). (UA: <i>User Accuracy</i> , PA: <i>Producer Accuracy</i>).....	33
--	----

LISTE DES ABREVIATIONS^a

AUCN	Area Under Curve Normalization
CR	<i>Continuum</i> Removal
D1	First Derivative Spectrum
D2	Second Derivative Spectrum
ELM	Empirical Line Method
ENET	Elastic Net
EWT	Equivalent Water Thickness
HC	Hydrocarbons
HM	Heavy Metals
HPLC	High Pressure Liquid Chromatography
LAD	Leaf Angle Distribution
LAI	Leaf Area Index
LCC	Leaf Chlorophyll Content
LDA	Linear Discriminant Analysis
LWC	Leaf Water Content
MLA	Mean Leaf Angle
NIR	Near Infrared
OA / MOA	Overall Accuracy / Mean Overall Accuracy
PA	Producer's Accuracy
PAH	Polycyclic Aromatic Hydrocarbon
PCA	Principal Component Analysis
PLSR	Partial Least Square Regression
REP	Red-Edge Position
RF	Random Forest
RLR	Regularized Logistic Regression
RMSE	Root Mean Squared Error
RPD	Residual Predictive Deviation
RTM	Radiative Transfer Model
SAR	Synthetic Aperture Radar
SEPC	Standard Error of Prediction Corrected from the bias
SNR	Signal-to-Noise Ratio
SNV	Standard Normal Variate
SWIR	Short-Wave Infrared
TPH	Total Petroleum Hydrocarbons
UA	User's Accuracy
UAV	Unmanned Aerial Vehicle
UV	Ultraviolet
VI	Vegetation indices
VIF	Variable Importance Factor
VIS	Visible

^aCertaines abréviations peuvent varier entre les articles

INTRODUCTION GENERALE

Depuis le début du xx^e siècle, la production de pétrole et de gaz n'a cessé de croître (35 Gb/an en 2017), répondant à une demande mondiale croissante¹⁻³, en particulier dans les secteurs de l'industrie et du transport (> 95%). La quasi-totalité de cette production provient du pétrole dit « conventionnel » qui, une fois extrait, est raffiné en produits pétroliers tels que les carburants (essence, diesel, etc.). Son exploitation entraîne toutefois une diminution des réserves, estimées à 1480 Gb en 2017². Les principaux champs pétrolifères, dont seulement 110 assurent la production mondiale, ont été découverts au cours du xx^e siècle¹ (1945 – 1980). Par conséquent, la prospection de nouveaux champs, de moins en moins étendus, est devenue particulièrement difficile. Elle représente aujourd'hui un enjeu majeur pour les compagnies pétrolières, nécessitant un effort constant de leur part pour développer de nouvelles avancées technologiques. L'exploration pétrolière présente en effet de nombreux risques d'échec, et nécessite donc des technologies de prospection efficaces, en particulier dans le domaine *onshore* d'où provient 70% de la production mondiale.

La prospection pétrolière consiste à mettre en évidence la présence de pétrole en profondeur, laquelle peut s'exprimer en surface sous forme de micro- et macrosuintements⁴⁻⁷ (Figure 1a-d). Une simple reconnaissance de terrain peut ainsi suffire à détecter ces formations. De nombreuses technologies ont également été mises au point afin de sonder la croûte terrestre, parmi lesquelles la sismique par réflexion reste la plus utilisée. L'utilisation des technologies géophysiques réduit le risque de prospection infructueuse⁸, mais implique en retour un coût très important en temps et en main d'œuvre. Des alternatives plus rentables sont donc nécessaires pour détecter le pétrole, d'autant plus que des besoins similaires apparaissent dans d'autres domaines de l'industrie pétrolière. C'est notamment le cas en matière de monitoring de la contamination.

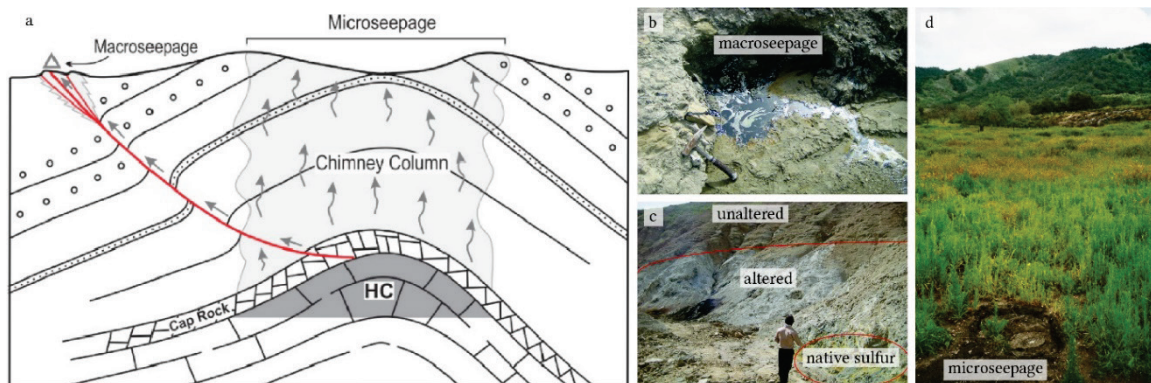


Figure 1. (a) Représentation schématique d'un micro- et macrosuintement pétrolier (adapté de ⁹). (HC: hydrocarbures.) Les macrosuintements sont caractérisés par des remontées de pétrole apparentes en surface (b). Les microsuintements s'expriment au travers d'altérations de la lithologie (c) et de la répartition de la végétation en surface (d)^{4,10}.

La contamination pétrolière demeure une préoccupation écologique majeure, car elle peut survenir à toutes les étapes de la production et causer des dommages irréversibles aux écosystèmes¹¹⁻¹⁵. Le déversement accidentel de pétrole brut, de produits raffinés ou de résidus de production (boues et eaux contaminées) sont autant de sources de contamination qu'il est nécessaire de détecter rapidement, afin de prémunir contre ces dommages, sinon les limiter. Peu

de technologies ont été développées pour détecter des déversements pétroliers dans le domaine *onshore*. Cette opération demeure le plus souvent réalisée par des opérateurs sur site et permet difficilement une détection précoce de la contamination, en particulier lorsque celle-ci est diffuse. Elle implique par ailleurs une main d'œuvre conséquente et une surveillance continue, tout comme la prospection pétrolière.

Face aux limites actuelles des technologies utilisées en prospection pétrolière et en monitoring de la contamination, la télédétection se positionne comme une alternative prometteuse. Plus précisément, la télédétection optique passive pourrait permettre de détecter rapidement et à grande échelle les suintements et les déversements pétroliers dans le domaine *onshore*, grâce à l'émergence des capteurs imageurs aéroportés et satellitaires hyperspectraux⁹. La télédétection hyperspectrale fournit des données de réflectance spectrale des surfaces sur un spectre continu dans le domaine optique réflectif [400 :2500] nm, renseignant ainsi sur leur composition (Figure 2). Dans le secteur pétrolier, son utilisation repose sur l'exploitation des propriétés optiques des hydrocarbures dans le sol (détection directe) et la détection de changements dans la lithologie, la minéralogie et la répartition de la végétation en surface (détection indirecte). La présence naturelle ou accidentelle de pétrole étant parfois très localisée, il est toutefois nécessaire de disposer d'images à très haute résolution spatiale pour parvenir à la détecter.

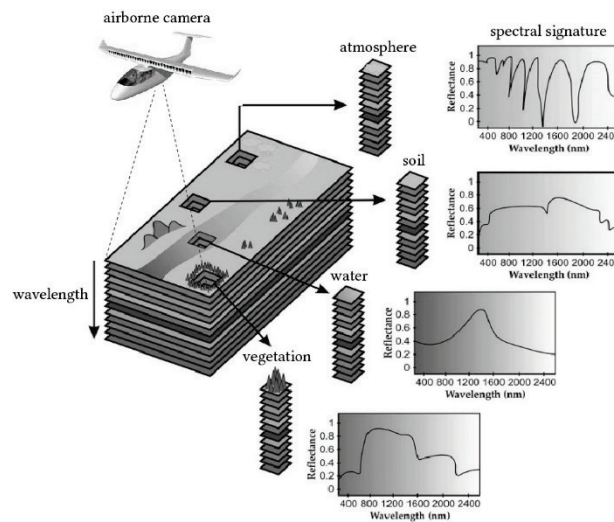


Figure 2. Principe de l'imagerie hyperspectrale (adapté de ¹⁶). Cette technologie fournit des données réflectance spectrale des surfaces sur un spectre continu du domaine optique réflectif (i.e. la signature spectrale).

Plusieurs études ont démontré l'intérêt de la télédétection hyperspectrale pour détecter des suintements naturels et des fuites d'installations pétrolières^{17,18}. Bien qu'elle constitue une alternative très intéressante aux technologies actuelles, son utilisation demeure très limitée dans les régions végétalisées, où les compagnies pétrolières concentrent une part importante de leurs activités. En présence de végétation, le rayonnement solaire ne pénètre que très peu le feuillage, rendant impossible l'accès à la signature spectral du sol et donc la détection directe du pétrole. Afin de contourner cette limite, de récentes études suggèrent d'exploiter les propriétés optiques de la végétation afin de détecter indirectement le pétrole par télédétection hyperspectrale^{4,19,20}. Ces propriétés sont en effet étroitement liées aux paramètres biophysiques et biochimiques des feuilles, lesquels renseignent sur l'état de santé de la plante. Ce lien rend possible le suivi de la végétation par télédétection hyperspectrale aéroportée et satellitaires. De nombreuses

applications ont vu le jour en agriculture et en écologie dans ce domaine, en particulier pour détecter et quantifier des réponses à divers stress biotiques et abiotiques²¹⁻²⁵ (e.g. le déficit en eau). Le pétrole brut, les produits raffinés et les résidus de production sont susceptibles d'affecter l'état de santé de la végétation, modifiant ainsi ses propriétés optiques. L'utilisation de la télédétection hyperspectrale permettrait donc de détecter ces modifications et, indirectement, les suintements naturels et la contamination pétrolière en régions végétalisées. Bien que de récentes tentatives aient été effectuées à cette fin, il n'existe aujourd'hui aucune méthode dédiée applicable à l'imagerie aéroportée ou satellitaire.

OBJECTIFS DE LA THESE

La présente thèse entre dans le cadre du projet NAOMI entre TOTAL et l'Office National d'Études et de Recherches Aérospatiales (ONERA). Elle vise à développer de nouvelles méthodes exploitant la télédétection hyperspectrale et permettant de détecter et de quantifier le pétrole dans les sols en régions tempérées et tropicales végétalisées. Cette thèse a également bénéficié des compétences complémentaires des unités mixtes de recherche EcoLab et DynaFor de l'Université Paul Sabatier et de l'École Nationale Supérieure Agronomique de Toulouse (INP-ENSAT).

Ainsi que le prévoit la dérogation accordée par l'Institut Supérieur de l'Aéronautique et de l'Espace (ISAE – SUPAERO), ce document de synthèse vient en complément du manuscrit de thèse rédigé en anglais à partir des articles publiés ou soumis pour publication. Il contient les éléments essentiels à la compréhension des travaux menés dans le cadre de cette thèse.

CHAPITRE UN : ETAT DE L'ART DE L'UTILISATION DE LA TELEDETECTION PASSIVE HYPERSPÉCTRALE POUR L'EXPLORATION PÉTROLIÈRE ET LE MONITORING DE LA CONTAMINATION EN RÉGIONS VÉGÉTALISÉES

INTRODUCTION

La détection indirecte du pétrole en régions végétalisées par imagerie hyperspectrale repose sur l'analyse de l'état de santé de la végétation, au travers de ses propriétés optiques. Elle requiert une solide connaissance de ces propriétés et de leurs modifications induites par l'exposition de la plante au pétrole brut et aux produits pétroliers. D'autre part, elle implique de disposer de méthodes permettant de détecter ces modifications à partir d'images optiques aéroportées et satellitaires. Ce chapitre dresse un état de l'art des avancées dans ce domaine émanant de la littérature, depuis l'étude des propriétés optiques de la végétation et leurs modifications en présence de pétrole, jusqu'aux cas d'application en exploration pétrolière et en monitoring de la contamination.

1. Les propriétés optiques de la végétation dans le domaine optique réflectif (400 – 2500 nm)

1.1. Influence des pigments dans la région du visible (400 – 750 nm)

Dans le domaine réflectif, les propriétés optiques de la végétation sont gouvernées par ses paramètres biophysiques et biochimiques. Leur influence relative diffère toutefois selon la région du spectre considérée (Figure 3). Dans le visible (VIS), l'influence des pigments, et plus particulièrement des chlorophylles, prédomine²⁶⁻²⁸. Les chlorophylles a et b, pigments directement impliqués dans le processus de photosynthèse, absorbent le rayonnement autour de 440-450 et 650-670 nm. Elles sont responsables du pic de réflectance à 550 nm et de l'augmentation abrupte de cette dernière vers 700 nm, aussi appelée *red-edge*^{29,30} (Figure 3). En exploitant ces longueurs d'ondes, il est alors possible de quantifier le contenu en chlorophylles des feuilles (*Leaf Chlorophyll Content*, LCC) par télédétection³¹. De nombreuses approches ont été développées à ce propos^{28,32}, tels que les rapports de réflectance (*Vegetation Indices*, VI) et les modèles de transfert radiatif (*Radiative Transfer Model*, RTM) présentés plus loin.

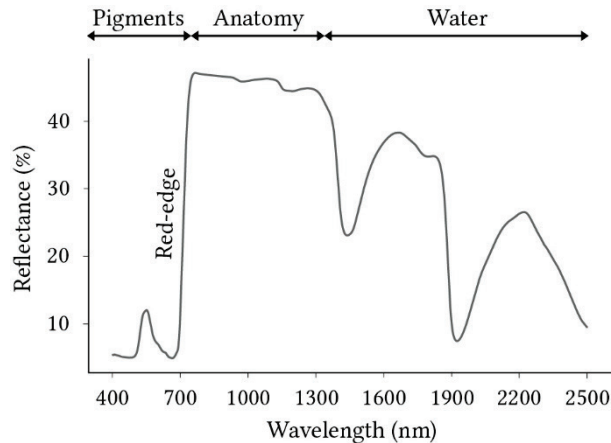


Figure 3. Signature spectrale d'une feuille de plante en bonne santé. Pour chaque région du spectre, le paramètre biophysique ou biochimique le plus influent est indiqué.

Également impliqués dans la photosynthèse, les caroténoïdes, qui regroupent les carotènes et les xanthophylles, absorbent la lumière essentiellement dans le bleu^{27,33} (400 – 500 nm). Habituellement masqués par les chlorophylles, ces pigments s'expriment à la senescence par une augmentation de la réflectance entre 500 et 700 nm³⁴ (vert – rouge). Ils assurent une fonction de photoprotection lorsque l'énergie absorbée excède les capacités photosynthétiques des feuilles et sont impliqués dans les processus de réponse au stress^{35,36}. Ces pigments présentent donc un fort intérêt pour le suivi de l'état de santé de la végétation. De même que pour les chlorophylles, l'utilisation de VI et de RTM permet de quantifier les caroténoïdes par télédétection^{27,37,38}.

Les feuilles contiennent également d'autres pigments non-photosynthétiques, notamment les anthocyanes. Situées dans les vacuoles, les anthocyanes absorbent la lumière dans l'ultraviolet (UV, 250 – 350 nm) et le vert^{39,40} (500 – 560 nm). Leur accumulation dans les feuilles induit une augmentation de la réflectance vers 640 nm, responsable de la coloration rouge à la sénescence. Les anthocyanes forment une barrière contre les rayonnements UV, source de stress pour la plante. Elles demeurent toutefois difficiles à quantifier par télédétection, car elles s'expriment peu en dehors de la sénescence. D'autres composés, tels que les tanins, sont également présents dans les feuilles. Ils sont responsables du brunissement ; leur influence sur les propriétés optiques des feuilles reste très limitée en-dehors de la senescence³⁵.

1.2. Influence de l'anatomie des feuilles dans le proche-infrarouge (750 – 1300 nm)

L'anatomie des feuilles influence leurs propriétés optiques dans le proche-infrarouge (*Near-Infrared*, NIR). Elle est responsable du plateau observé dans cette région du spectre (Figure 3), s'étalant de 30 à 80% de réflectance⁴¹⁻⁴³. L'épaisseur de la cuticule et des parenchymes, la proportion d'espaces intercellulaires et l'arrangement des chloroplastes dans les cellules sont parmi les facteurs les plus influents dans le NIR. L'anatomie des feuilles diffère fortement entre espèces mono- et dicotylédones, et apparaît très diversifiée au sein même de ces groupes^{41,44,45}. Elle traduit notamment l'adaptation aux conditions d'éclairement, en particulier dans les écosystèmes où la compétition pour la lumière est importante. Cette diversité anatomique contribue à la diversité des propriétés optiques entre espèces, mais également entre individus d'une même espèce. En effet, selon les conditions environnementales auxquelles une espèce est

exposée (disponibilité en lumière, en eau et en nutriments), sa signature spectrale dans le NIR s'en trouvera affectée⁴⁶⁻⁴⁸.

À l'échelle de l'individu et de la canopée, d'autres paramètres biophysiques influencent la réflectance de la végétation dans le NIR, notamment la densité de feuilles (*Leaf Area Index*, LAI) et leur orientation (*Leaf Angle Distribution*, LAD)⁴⁹⁻⁵¹. Le LAI, défini comme la surface projetée de feuilles par unité de surface au sol (en $m^2.m^{-2}$), est directement lié au développement de la végétation. À mesure que sa valeur augmente, la fraction de sol nu – dont la réflectance excède rarement 30% – diminue. Le LAI est donc positivement corrélé à la réflectance dans le NIR (Figure 4a). La LAD décrit quant à elle l'architecture du couvert et plus précisément l'orientation angulaire des feuilles. Ce paramètre détermine l'angle d'incidence du rayonnement à la surface des feuilles et donc la trajectoire de ce dernier dans les tissus. Selon de Wit⁵², les espèces peuvent être classées en six catégories selon la distribution de leurs angles foliaires : planophile, plagiophile, erectophile, extrémophile, sphérique et uniforme. Lorsque l'orientation moyenne des feuilles se rapproche de zéro (distribution planophile), la réflectance du couvert augmente dans le NIR (Figure 4b).

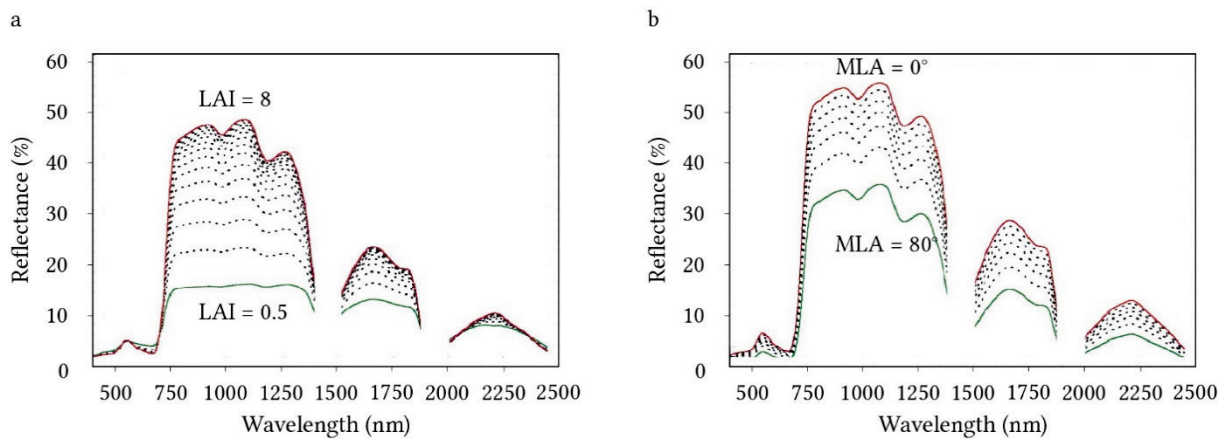


Figure 4. Influence (a) du *Leaf Area Index* (LAI) et (b) de l'orientation moyenne des feuilles (*Mean Leaf Angle*, MLA) sur la réflectance d'un couvert végétal⁴⁹.

1.3. Influence du contenu en eau et en matière sèche dans le proche-infrarouge (750 – 1300 nm) et l'infrarouge aux courtes longueurs d'ondes (1300 – 2500 nm)

Présente sous forme liquide et gazeuse dans la plante, l'eau absorbe la lumière autour de 970, 1200, 1450, 1950 et 2250 nm⁵³⁻⁵⁵. Son influence sur les propriétés optiques de la végétation s'étend ainsi du NIR à l'infrarouge aux courtes longueurs d'ondes (*Short-Wave Infrared*, SWIR). Des changements dans le statut hydrique de la végétation s'expriment directement par des modifications de réflectance dans ces longueurs d'ondes (Figure 5). En raison de son implication dans la photosynthèse et le mécanisme de turgescence cellulaire, l'eau affecte également la réflectance dans le reste du spectre. Bien que les longueurs d'ondes d'absorption de l'eau demeurent peu exploitables en raison d'effets atmosphériques, d'autres situées dans le NIR et le SWIR permettent de suivre le statut hydrique de la végétation par imagerie aéroportée et satellitaire⁵⁶⁻⁵⁸. L'eau représente environ 55 à 75% du poids frais des feuilles⁵⁹. Plus des deux tiers restants proviennent des hémicelluloses, celluloses, lignines et protéines^{60,61}. Souvent

regroupés sous le terme de « matière sèche », ces paramètres biochimiques entrent dans la composition des parois cellulaires et confèrent aux végétaux leur rigidité. Leur influence sur les propriétés optiques de la végétation se concentre dans le NIR et le SWIR, autour de 1200, 1450 – 1490, 1540, 1760, 2100 et 2340 nm, mais celle-ci reste très limitée en comparaison avec l'eau pour des feuilles fraîches⁶².

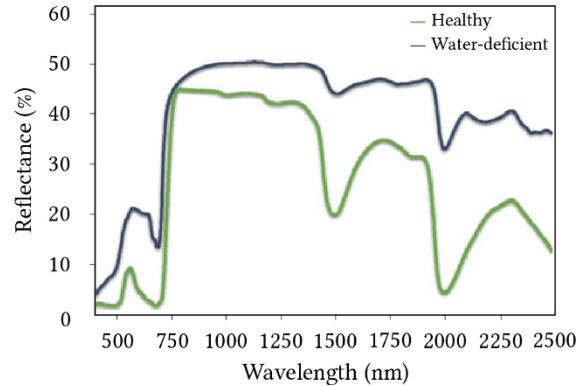


Figure 5. Signatures spectrales de plantes en bonne santé (« Healthy ») ou en situation de stress hydrique (« Water-deficient ») (données personnelles).

Selon la région du spectre, les propriétés optiques de la végétation sont ainsi déterminées par différents paramètres biophysiques et biochimiques. Toute modification de ces paramètres engendre des changements de réflectance des feuilles et des couverts aux longueurs d'ondes correspondantes. C'est notamment le cas en présence de pétrole brut ou de produits pétroliers dans le sol.

2. Effets du pétrole brut et des produits pétroliers sur l'état de santé de la végétation

2.1. Composition du pétrole brut et des produits pétroliers ; effets directs et indirects sur la végétation

Le pétrole brut correspond au pétrole sous sa forme naturelle extractible, tandis que les produits pétroliers (carburants, lubrifiants, etc.) résultent de son raffinage, au cours duquel sont également produits les boues et les eaux contaminées⁶³⁻⁶⁵. Le pétrole brut et les produits pétroliers sont des mélanges complexes d'hydrocarbures volatils à denses – le plus souvent regroupés sous le terme d'hydrocarbures pétroliers totaux (*Total Petroleum Hydrocarbons*, TPH), de métaux lourds – ou éléments traces métalliques, et de composés organiques à base d'oxygène, de soufre et d'azote⁶⁶⁻⁶⁸. Seul, chacun de ces composés est susceptible d'affecter l'état de santé de la végétation. Leur présence en mélange peut également induire des effets combinés, synergiques ou antagonistes.

Le pétrole brut et les produits pétroliers affectent la végétation de manière indirecte, en modifiant les propriétés physico-chimiques et biologiques du sol, et de manière directe au contact de la plante et par accumulation dans les tissus. Au niveau du sol, le régime hydrique est

fortement impacté par l'hydrophobicité des hydrocarbures⁶⁹. Ce phénomène entraîne une diminution de la disponibilité en eau du sol, pouvant induire un stress hydrique chez la plante⁷⁰⁻⁷². La présence naturelle ou accidentelle d'hydrocarbures pétroliers dans le sol représente également un apport considérable de matière organique, modifiant les communautés de microorganismes et les cycles de minéralisation, en particulier celui de l'azote^{68,72,73}. Ces changements réduisent la disponibilité en nutriments (N, P, K) pour les végétaux et leur imposent alors des conditions de croissance très sélectives^{71,72,74}. De manière directe, le pétrole brut et les produits pétroliers réduisent les capacités d'absorption des racines à leur contact⁷⁵. L'assimilation d'hydrocarbures et de métaux lourds dans les racines inhibe leur croissance et provoque un épaississement de l'épiderme, de l'endoderme et du cortex et une réduction du diamètre et de la densité des poils racinaires⁷⁶⁻⁷⁹. L'ensemble de ces effets conduit à des modifications des paramètres biophysiques et biochimiques de la végétation en lien avec ses propriétés optiques.

2.2. Réponse biophysique et biochimique de la végétation et sources de variabilité

L'altération du contenu en pigments des feuilles, en particulier les chlorophylles, est la réponse la plus fréquente à l'exposition au pétrole brut et aux produits pétroliers⁸⁰⁻⁸² (Figure 6a-b). Elle traduit une diminution de l'activité photosynthétique. Bien que présents à plus faible concentration dans les feuilles, les caroténoïdes et les anthocyanes sont également affectés⁸¹. L'altération du statut hydrique de la plante réduit la teneur en eau et donc la turgescence des cellules, modifiant l'anatomie interne des feuilles^{70,78}. À l'échelle du couvert, ces effets s'observent par un développement limité des parties aériennes et donc un faible LAI⁸¹. Les changements opérés au niveau de l'anatomie des feuilles affectent également le port de la plante et par conséquent le LAD.

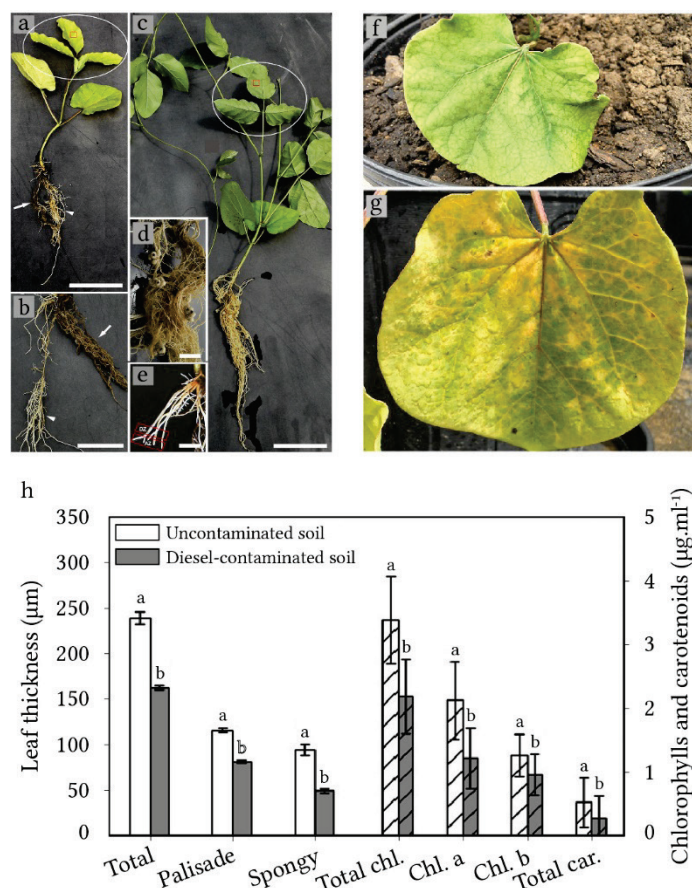


Figure 6. *Canavalia ensiformis* (L.) DC développée sur sol (a-b, e) contaminé au diesel (22,219 mg.kg⁻¹) ou (c-d) non-contaminé. Les flèches blanches indiquent les racines nécrosées. (f-g) Symptômes de décoloration des feuilles observés sur les plantes exposées au diesel. (h) Comparaison des paramètres anatomiques et du contenu en pigments des feuilles entre les deux traitements (moyenne ± écart-type). Les différences significatives ($p < 0,05$) sont représentées par des lettres différentes (modifié de ⁷⁷).

La sévérité des effets décrits ici est très variable selon la sensibilité de l'espèce aux hydrocarbures et aux métaux lourds^{83,84} et selon le scénario d'exposition^{12,81} (cf. Table 2 dans le manuscrit en anglais). Les conditions de croissance très sélectives imposées autour des suintements pétroliers et sur les bourniers contaminés rendent possible l'établissement d'un nombre très restreint d'espèces dites « tolérantes ». À l'inverse, les déversements pétroliers accidentels (e.g. fuite de pipeline) exposent des espèces bien souvent sensibles, induisant des effets très prononcés sur leur état de santé. La composition du mélange pétrolier est également déterminante^{80,83}. Selon leur composition en hydrocarbures et en métaux, les pétroles et les produits pétroliers n'affectent pas la végétation dans la même mesure. De même, leur concentration dans le sol détermine bien souvent l'amplitude des effets induits.

Les paramètres biophysiques et biochimiques en lien avec les propriétés optiques de la végétation sont ainsi affectés par l'exposition au pétrole brut et aux produits pétroliers. Ces changements se traduisent par des modifications de réflectance, rendant possible la détection de suintements et de sol contaminés par télédétection.

3. Détection du pétrole brut et des produits pétroliers par exploitation des propriétés optiques de la végétation

3.1. Modifications des propriétés optiques observées en conditions contrôlées

Le VIS reste la région du spectre la plus exploitée pour étudier les effets du pétrole brut et des produits pétroliers sur les propriétés optiques de la végétation^{82,85-87} (cf. Table 3 dans le manuscrit en anglais). L'altération des chlorophylles se traduit par une augmentation de la réflectance pouvant atteindre 20% entre 500 et 670 nm et un décalage de la position du red-edge vers des longueurs d'ondes plus courtes autour de 700 nm (Figure 7a-d). La réponse dans le NIR est en revanche variable selon l'espèce, la composition du mélange pétrolier et l'échelle de mesure⁸¹. Les différences entre plantes saines et exposées à la contamination excèdent parfois 20% de réflectance dans cette région. À l'échelle de la feuille, elles traduisent des changements anatomiques chez les plantes exposées, tandis qu'à l'échelle du couvert, ces différences proviennent bien souvent d'une réduction du LAI. Dans le SWIR, l'altération du statut hydrique provoquée par les hydrocarbures s'exprime par des changements de réflectance particulièrement marqués autour des longueurs d'ondes d'absorption de l'eau^{81,88}. Une augmentation de réflectance atteignant 20% est parfois observée dans cette région du spectre chez les espèces les plus sensibles.

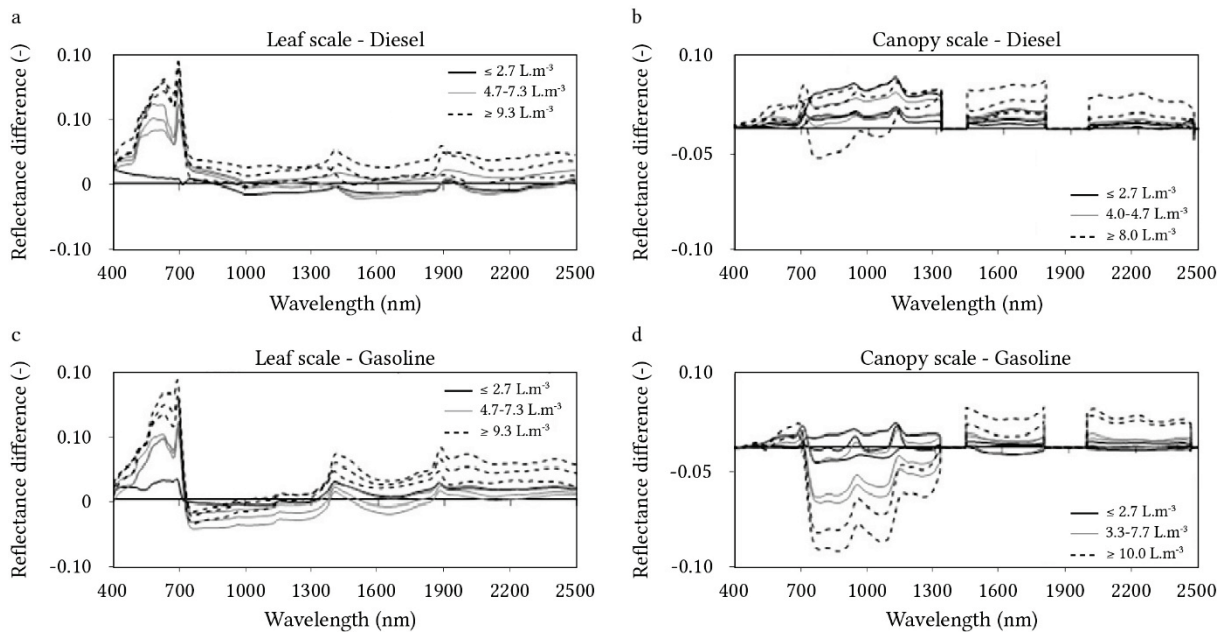


Figure 7. Différences de réflectance de *Brachiaria brizantha* H.S. développée sur sol contaminé (a-b) au diesel ou (c-d) à l'essence, à l'échelle (a, c) de la feuille et (b, d) du couvert. Les différences de réflectance ont été obtenues en soustrayant la signature spectrale des plantes témoins à celles de plantes exposées au diesel ou à l'essence. Les

courbes d'une même figure représentent différents niveaux de contamination (modifié de ⁸¹).

3.2. Méthodes existantes pour détecter le pétrole brut, les produits pétroliers et d'autres sources de stress de la végétation

En conditions contrôlées, la plupart des méthodes proposées pour détecter le pétrole brut et les produits pétroliers à partir des propriétés optiques de la végétation reposent sur des comparaisons statistiques de réflectances entre plantes saines et exposées à la contamination^{12,87}. Elles s'appuient sur les VI et les transformations de signatures spectrales^{81,85,86} (e.g. dérivées première), mais aucune d'entre elles ne permet véritablement de prédire si un sol contient ou non du pétrole ou un produit pétrolier, de manière automatique. L'utilisation de ces méthodes reste donc très limitée, d'autant plus qu'elles ont été développées sur des espèces sensibles et donc peu représentatives de celles établies autour des suintements et sur les bourniers contaminés. De nouvelles méthodes adaptées aux scénarios d'exposition réalistes (i.e. rencontrés en conditions naturelles) sont donc nécessaires. Elles pourraient s'appuyer sur les études portant sur l'évaluation de l'état de santé de la végétation dans d'autres contextes. Les VI et les transformations de signatures spectrales ont été largement exploités, soit au travers d'approches par classification visant à prédire le stress de la végétation, soit en régressions simples ou multiples pour quantifier ses paramètres biophysiques et biochimiques^{23,89-91}. L'analyse linéaire discriminante (*Linear Discriminant Analysis*, LDA), les machines à vecteurs de support (*Support Vector Machines*, SVM) et la régression partielle des moindres carrés (*Partial Least Square Regression*, PLSR) sont parmi les plus utilisées^{23,91-93}. D'autres approches ont également été développées pour des applications similaires, notamment les RTM. Des modèles physiques tels que PROSPECT permettent de simuler les propriétés optiques des feuilles et, par inversion, d'en estimer les paramètres biophysiques et biochimiques^{22,32}. Ces différentes approches pourraient ainsi servir au développement de méthodes spécifiquement dédiées à la détection et à la quantification – jusque-là non-abordée – du pétrole brut et des produits pétroliers en conditions naturelles.

3.3. Application en imagerie multi- et hyperspectrale aéroportée et satellitaire

Peu d'études ont tenté de détecter des suintements ou des déversements pétroliers par imagerie aéroportée et satellitaire en régions végétalisées^{20,94-96} (Figure 8a-b). La plupart ont exploité des données multispectrales pour détecter des altérations de l'état de santé de la végétation à l'aide de VI, par comparaison avec un site témoin (i.e. non-contaminé) (cf. Table 4 dans le manuscrit en anglais). Seules quelques-unes ont eu recours à la classification ou à la détection d'anomalies⁹⁵, et très peu d'entre elles ont quantifié les performances de ces approches. Ces dernières demeurent limitées par de nombreuses sources de confusion^{4,19} (Figure 8d). Par exemple, la présence de facteurs biotiques ou abiotiques induisant des effets similaires aux mélanges pétroliers. De plus, la variabilité interspécifique des propriétés optiques et de la sensibilité aux mélanges pétroliers rend la détection particulièrement difficile dans les régions à forte richesse spécifique. Cette difficulté est d'autant plus importante à des résolutions spatiales de quelques mètres ou décamétriques, où les espèces sont mélangées au sein d'un pixel. Une difficulté analogue se pose lorsque ces espèces sont mélangées avec du sol nu. La prise en compte de ces sources de confusion dans les méthodes, ainsi que l'utilisation de la très haute résolution spatiale (< 2 m), pourraient améliorer sensiblement la détection.

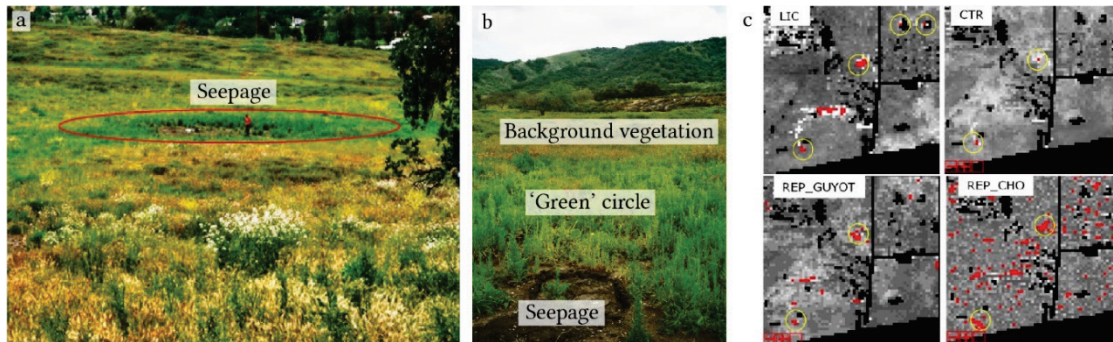


Figure 8. (a) Microsuintement végétalisé. (b) Le suintement est caractérisé par un mélange de sol nu et de pétrole brut au centre, entouré d'une végétation à la distribution spatiale particulière, le rendant détectable par imagerie hyperspectrale aéroportée à l'aide (c) d'indices de végétation et d'un filtre spatial. Toutefois, des fausses alarmes (pixels rouges en-dehors des cercles jaunes) apparaissent lors de la détection⁴.

La connaissance du contexte d'étude (i.e. espèces, mélange pétrolier, sources de confusion) est essentielle au développement de méthodes fiables de détection – et de quantification – du pétrole brut et des produits pétroliers applicables en imagerie hyperspectrale. Les expérimentations en conditions contrôlées représentatives de ce contexte constituent pour cela une première étape nécessaire au développement des méthodes, avant un passage en conditions naturelles, d'abord sur le terrain puis en imagerie aéroportée et satellitaire. Toutefois, aucune étude n'a à ce jour suivi une telle approche.

CONCLUSION

Les propriétés optiques de la végétation sont étroitement liées à ses paramètres biophysiques et biochimiques. L'altération de ces derniers en présence de pétrole brut ou de produits pétroliers se traduit par des modifications de réflectance des feuilles et des couverts végétaux, plus ou moins marquées selon l'espèce, la composition et la concentration du mélange et le temps d'exposition. L'exploitation de ces modifications à des fins de prospection pétrolière ou de monitoring de la contamination demeure toutefois limitée en imagerie aéroportée et satellitaire, en raison de nombreuses sources de confusion (variabilité interspécifique, autres stress, influence du sol nu). Il apparaît alors nécessaire de caractériser le contexte d'étude ciblé (i.e. espèces, mélange pétrolier, etc.) au préalable, au travers d'expérimentations en conditions contrôlées. Celles-ci doivent permettre de développer des méthodes de détection et de quantification du pétrole brut et des produits pétroliers applicables aux images à hautes résolutions spatiale et spectrale, en conditions naturelles. L'utilisation de VI, de transformations de signatures spectrales et de RTM en classifications et régressions pourrait servir de base au développement de ces méthodes.

PROBLEMATIQUE SCIENTIFIQUE ET APPROCHE MISE EN ŒUVRE

Il n'existe à l'heure actuelle aucune méthode permettant de détecter et de quantifier le pétrole brut et les produits pétroliers à partir d'images aéroportées ou satellitaires en régions végétalisées. Le développement de telles méthodes constitue l'objectif de la présente thèse. Pour y parvenir, la thèse s'appuie sur une démarche en trois étapes, depuis des expérimentations en conditions contrôlées jusqu'à la validation des méthodes sur le terrain puis sur des images aéroportées à très hautes résolution spatiale et spectrale. Pour cela, cette thèse se concentre sur l'étude des bourbiers contaminés végétalisés, en régions tempérée et tropicale. Ceux-ci peuvent être définis comme des friches industrielles (50 – 5000 m²) contaminées par des résidus de production pétrolière (TPH et métaux principalement). Lorsqu'ils subsistent – i.e. après cessation de l'activité de production, les bourbiers sont progressivement colonisés par des espèces tolérantes à la contamination ; dans notre cas, *Rubus fruticosus* L. et *Cenchrus alopecuroides* (L.) pour les contextes tempéré et tropical, respectivement.

Une approche en trois étapes a été adoptée dans le cadre de cette thèse, afin de détecter et quantifier la contamination du sol en exploitant les propriétés optiques de la végétation. La première étape a reposé sur l'étude des bourbiers en conditions contrôlées et a eu pour but de développer des méthodes de détection et de quantification. Pour cela, une étude expérimentale de faisabilité, suivie de deux études plus complètes, ont été conduites. Ces expérimentations sont présentées dans le chapitre deux. Les méthodes ont ensuite été évaluées et complétées en conditions naturelles, sur le terrain et les images aéroportées. Ces seconde et troisième étapes constituent le chapitre trois. À chaque étape de l'approche, les performances des méthodes ont été évaluées et comparées à celles de l'étape précédente.

CHAPITRE DEUX : DEVELOPPEMENT DE METHODES DE DETECTION ET DE QUANTIFICATION DE LA CONTAMINATION PETROLIERE EXPLOITANT LES PROPRIETES OPTIQUES DE LA VEGETATION, EN CONDITIONS CONTROLEES

INTRODUCTION

La première étape de l'approche adoptée dans cette thèse a eu pour objectif de développer des méthodes d'évaluation de la contamination (i.e. détection, caractérisation et quantification) exploitant les propriétés optiques de la végétation, en conditions contrôlées. Pour cela, trois expérimentations en serre, reproduisant les contextes des borbiers en régions tempérée et tropicale, ont été conduites. Elles ont consisté à exposer les espèces végétales à des sols contaminés et à mesurer régulièrement leur réflectance dans le domaine optique réflectif à plusieurs échelles. La première expérimentation a permis d'évaluer la sensibilité des espèces tempérée et tropicale à la contamination et de déterminer les longueurs d'ondes d'intérêt pour le développement des méthodes. Les contextes tempéré et tropical ont ensuite été traités séparément lors de deux expérimentations plus conséquentes. La première, en contexte tempéré, a eu pour but de développer une méthode de détection et de caractérisation de la contamination (en termes de TPH et de métaux). La seconde, en contexte tropical, a eu pour objet de proposer une méthode de quantification des TPH. Une procédure analogue a été suivie pour développer ces méthodes. Celle-ci a consisté à mettre en lien les réponses biochimique et spectrale de la végétation induites sur le long terme (≥ 1 mois) par l'exposition à la contamination, à l'aide de VI. Ces indices ont été ensuite utilisés en classification et en régression multiple pour développer les méthodes. Ces dernières ont été évaluées aux différentes échelles de mesure, puis sur le terrain pour le contexte tempéré.

4. Premier article : “Assessing soil contamination due to oil and gas production using vegetation hyperspectral reflectance”

4.1. Contexte et objectif de l'étude

La première expérimentation menée en serre a constitué une étude de faisabilité pour le développement des méthodes. Elle a permis d'évaluer la sensibilité des espèces identifiées en contextes tempéré et tropical à la contamination des bourniers ; *Rubus fruticosus* L. pour le contexte tempéré et *Cenchrus alopecuroides* (L.) et *Panicum virgatum* L. pour le contexte tropical. Les trois espèces (5 répliques/traitement) ont été exposées soit à un sol témoin, non contaminé, soit aux sols de bourniers contaminés par 36 (contexte tempéré) ou 14 (contexte tropical) g.kg⁻¹ de TPH et par des métaux lourds. Un troisième traitement constitué d'un sol déficient en eau (i.e. non irrigué), potentielle source de confusion avec la contamination pétrolière, a été appliqué à *R. fruticosus*. Des mesures de réflectance ont été conduites régulièrement à l'échelle de la feuille (*R. fruticosus* seulement, n = 45 mesures/date) et de l'individu (i.e. du « plant », n = 30 ou 45 mesures/espèce/date), à l'aide d'un spectroradiomètre et d'une source lumineuse contrôlée couvrant le domaine optique réflectif (350 – 2500 nm). Dans une perspective d'application en imagerie, les signatures spectrales obtenues ont été ré-échantillonnées à la résolution spectrale de l'instrument utilisé par la suite pour les acquisitions aéroportées (cf. cinquième article). Ces données ont ensuite été utilisées en classification pour discriminer les différents traitements *via* une démarche par apprentissage et test (50% des données utilisés à chaque étape). Une procédure pas-à-pas, l'analyse linéaire discriminante (*Linear Discriminant Analysis*, LDA), a été retenue, car elle permet de sélectionner les longueurs d'ondes contribuant à la discrimination *via* un « critère d'amélioration » (le lambda de Wilk). La LDA a été appliquée sur les signatures spectrales originales et transformées. Cinq transformations ont été évaluées : les dérivées première (D1) et seconde (D2), le *Standard Normal Variate* (SNV), la normalisation par l'aire (*Area Under Curve Normalization*, AUCN) et le *continuum removal* (CR).

4.2. Principaux résultats

Chez les trois espèces, l'exposition à la contamination a induit une augmentation modérée de la réflectance dans le VIS et le SWIR en comparaison avec les témoins après seulement 13 et 25 jours respectivement, traduisant une réduction des contenus en pigments et en eau des feuilles (Figure 9a-b). En revanche, une réponse opposée a été constatée dans le NIR après 49 jours entre *R. fruticosus*, dicotylédone, et les espèces tropicales, monocotylédones, provenant possiblement des différences anatomiques entre ces espèces. Ces modifications se sont accentuées ($\leq 10\%$) jusqu'à la fin de l'étude et sont apparues plus tardivement (49 jours) – et moins prononcées – pour les plantes exposées aux sols déficients en eau. Elles confirment par ailleurs les résultats obtenus dans la littérature, bien que l'amplitude des modifications observées ici demeure moins importante, du fait de la tolérance élevée des espèces à la contamination pétrolière.

La discrimination des traitements par LDA s'est révélée nettement améliorée (précision globale $> 80\%$) après transformation des signatures spectrales par les dérivées et le CR à l'échelle de la feuille (*R. fruticosus* uniquement), toutes dates confondues (13 – 61 jours). Les confusions entre plantes témoins et exposées aux sols déficients en eau ont constitué la principale source d'erreurs de classification. De manière globale, l'exposition prolongée des plantes aux traitements a permis d'améliorer leur discrimination, jusqu'à atteindre une précision globale de 100% après

61 jours – et parfois avant, selon la transformation appliquée aux signatures spectrales (Figure 10a).

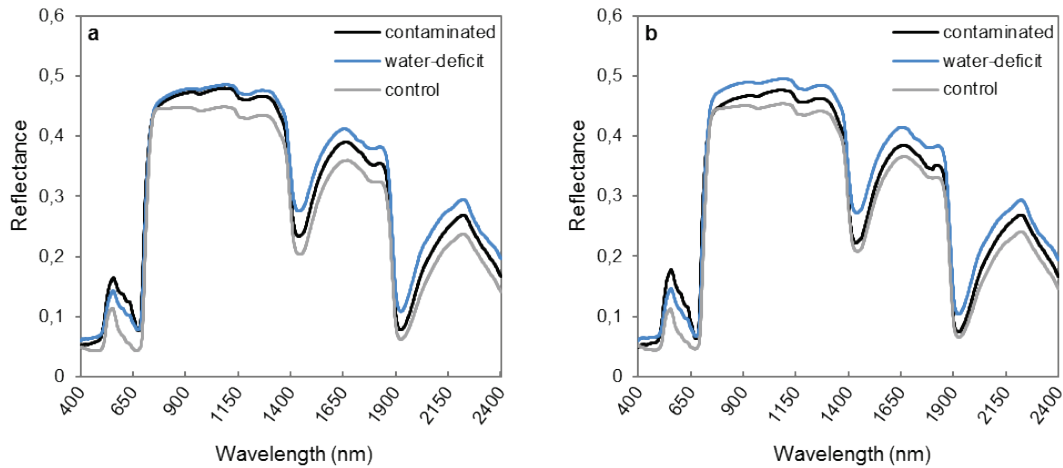


Figure 9. Réflectance moyenne de *R. fruticosus* après 61 jours d'exposition aux différents traitements, à l'échelle (a) de la feuille et (b) de l'individu.

Toujours concernant *R. fruticosus*, Une tendance similaire a été observée à l'échelle de l'individu. Toutefois, les signatures spectrales originales (i.e. non-transformées) ont conduit à une classification moins précise (précision globale = 54%) qu'avec l'ensemble des transformations testées (précision globale > 70%), soulignant davantage l'intérêt de ces dernières. Ces résultats s'expliquent par une meilleure discrimination des plantes exposées aux sols déficients en eau. Une précision globale supérieure à 90% a été atteinte après 61 jours avec les dérivées, le CR et l'AUCN (Figure 10b).

Chez les espèces tropicales (échelle de l'individu uniquement), l'utilisation des transformations de signatures spectrales n'a eu que peu d'incidence sur les résultats de classification de *C. alopecuroides* (précision globale ~ 80%). À l'inverse, les résultats obtenus pour *P. virgatum* se sont révélés plus proches de ceux de *R. fruticosus*. Pour ces deux espèces tropicales, la discrimination des traitements s'est vue une fois encore améliorée après un temps d'exposition prolongé (précision globale > 90% après 61 jours), en particulier celle des plantes exposées à la contamination (précision = 100%).

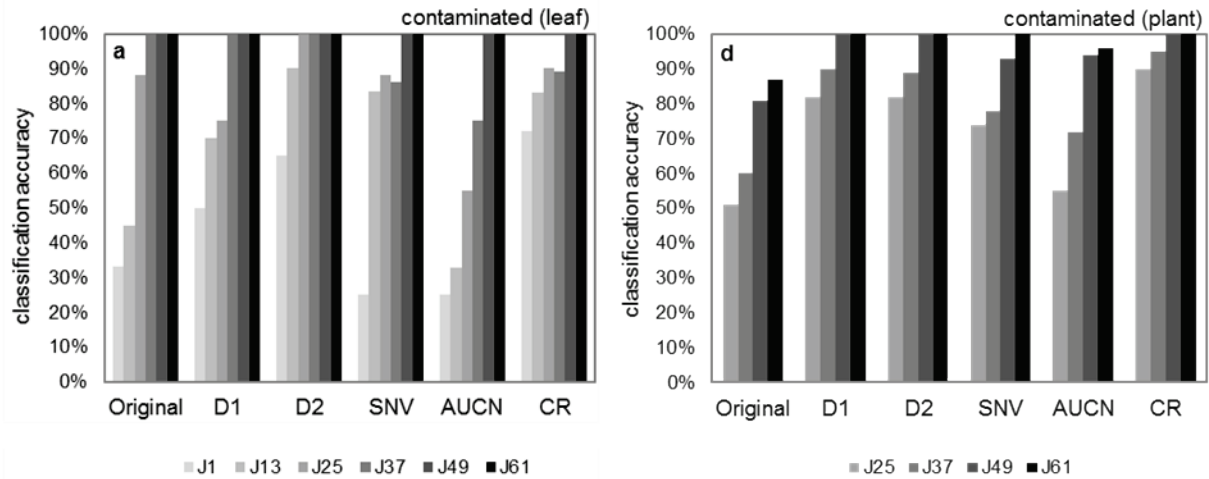


Figure 10. Évolution temporelle de la précision de la classification des plantes exposées à la contamination obtenue pour *R. fruticosus*, à l'échelle (a) de la feuille (J1 – J61) et (b) de l'individu (J25 – J61).

Les longueurs d'ondes du VIS situées dans les intervalles 400 – 500 nm et 700 – 750 nm (*red-edge*) se sont révélées être les plus importantes pour discriminer les traitements chez les trois espèces. D'autres localisées dans le NIR (900 – 110 nm) et le SWIR (1500 – 1600 et 2200 – 2300 nm) ont également été retenues. Ces résultats soulignent les possibles altérations du contenu en pigments et en eau des feuilles, déjà suspectées sur les signatures spectrales (Figure 10a-b).

4.3. Conclusion et perspectives

Cette première expérimentation a démontré la possibilité de détecter la contamination pétrolière des bourières à partir des propriétés optiques des espèces identifiées en contextes tempéré et tropical. Les résultats cohérents obtenus entre échelles de mesure pour *R. fruticosus*, ainsi que la discrimination de la contamination avec les sols déficients en eau, ont confirmé son potentiel à plus grande échelle en conditions naturelles. De même, l'absence presque totale de confusion entre les traitements observée pour les espèces tropicales a ouvert des perspectives encourageantes pour la suite des travaux de thèse dans ce contexte. La classification pas-à-pas a permis d'identifier les longueurs d'ondes – et les paramètres biochimiques associés – intéressantes pour le développement des méthodes abordé par la suite. Dans cette optique, une analyse complémentaire, basée sur des VI exploitant ces longueurs d'ondes, a été menée dans le cadre de cette même expérimentation. Elle a souligné l'intérêt de combiner plusieurs de ces longueurs d'ondes pour discriminer les plantes témoins de celles exposées à la contamination. L'utilisation des VI en classification et en régression multiple est amplement développée dans les articles suivants.

5. Deuxième article : “Detection and discrimination of various oil-contaminated soils using vegetation reflectance”

5.1. Contexte et objectif de l'étude

La seconde expérimentation menée en conditions contrôlées s'est concentrée sur le contexte tempéré. Elle avait pour objectif de permettre le développement d'une méthode de détection et de caractérisation de la contamination pétrolière applicable en conditions naturelles, sur le terrain et les images hyperspectrales aéroportées. L'expérimentation a consisté en l'exposition de *R. fruticosus* à quatre sols contaminés (trois sols de borbiers et un mélange de sol et de pétrole brut) et deux sols témoins (avec ou sans déficit en eau), pendant 32 jours (13 répliques/traitement). Les quatre types de sols contaminés différaient dans leur composition en TPH (6 ou 25 g.kg⁻¹ et coupes C₅-C₁₀ et C₁₀-C₄₀ en proportions variables) et en métaux lourds (Cr et Zn principalement), simulant ainsi plusieurs cas réalistes de contamination pétrolière. Les contenus en pigments et en eau des feuilles ont été suivis au cours de l'expérimentation. Des mesures de réflectance ont été effectuées régulièrement à l'échelle de la feuille (n = 234 mesures/date) et sous éclairage naturel à l'échelle de l'individu (n = 78 mesures/date) et du couvert (4 plants accolés, n = 18 mesures/date). 33 VI ont été calculés à partir de ces données et dans un premier temps mis en lien avec les teneurs en pigments et en eau par régression Elastic net (ENET), à l'échelle de la feuille. Une sélection des indices concordants entre les trois échelles de mesure (feuille, individu et couvert) a ensuite été conduite par analyse de concordance, basée sur le calcul du coefficient de Kendall (W). Les meilleurs indices – i.e. les plus corrélés aux paramètres biochimiques ($R^2 > 0,7$) et concordants – ont été retenus pour discriminer les traitements par régression logistique régularisée (*L²-Regularized Logistic Regression*, RLR). Pour cela, la méthode a été entraînée uniquement à l'échelle de la feuille, puis testée aux trois échelles de mesure, afin d'évaluer la robustesse de la méthode. Enfin, cette dernière a été évaluée sur le terrain, à partir de mesures effectuées sur des couverts de *R. fruticosus* sur un site témoin et un borbier contaminé.

5.2. Principaux résultats

Les premiers symptômes visibles de stress ont été observés après 18 jours d'exposition aux traitements, et différaient entre les traitements contaminés (décoloration, pigmentation rougeâtre). Les plantes exposées au borbier riche en coupes C₅-C₁₀, au pétrole brut et aux sols déficients en eau ont été les plus affectées, tandis que les autres, exposées aux borbiers pauvres en coupes C₅-C₁₀, ont vu leur état de santé s'améliorer après 32 jours. Une analyse en composantes principales des contenus en pigments et en eau a révélé des altérations plus ou moins marquées entre les traitements en comparaison avec les témoins à 18 jours, confirmant les symptômes. Cette réponse a été observée pour la plupart des pigments analysés (i.e. chlorophylles et caroténoïdes). La disparition des symptômes constatée pour deux des traitements contaminés à 32 jours s'est traduite par une augmentation des teneurs en pigments et en eau des feuilles, tandis que les altérations devenaient plus prononcées pour les autres traitements.

Les altérations biochimiques induites par l'exposition à la contamination et au déficit en eau ont entraîné d'importantes modifications de la réflectance des feuilles, en comparaison avec celle des témoins. Une augmentation de la réflectance est ainsi apparue à 18 jours autour de 500 – 700 nm dans le VIS, différant en amplitude (entre 2 et 10%) selon la composition du mélange pétrolier. D'importants changements ont également été constatés dans le NIR et le SWIR, suite à

l'altération du contenu en eau – et très probablement de l'anatomie – des feuilles. Une réduction importante de la réflectance ($> 10\%$) est apparue à l'échelle de l'individu et du couvert dans ces mêmes régions du spectre, en raison de la croissance limitée des plantes. L'ensemble de ces modifications s'est atténuée après 32 jours pour les plantes dont les contenus en pigments et en eau des feuilles avaient augmenté. Elles sont en revanche devenues plus marquées pour les autres traitements.

La régression ENET appliquée aux VI a permis de mettre en lien les modifications biochimiques et de réflectance observées. Sur les 33 indices testés, 15 se sont montrés très corrélés ($R^2 \geq 0,7$) aux teneurs en pigments ou en eau des feuilles. Les chlorophylles, le β -carotène et la lutéine ont en grande partie contribué aux variations de ces VI, bien que d'autres caroténoïdes aient été également impliqués. L'analyse de concordance des VI a permis de dégager des similarités entre les trois échelles de mesure, malgré les différences observées sur les signatures spectrales pour un même traitement et notamment l'influence du sol nu à l'échelle du couvert. 14 des 15 indices liés aux contenus en pigments ou en eau des feuilles se sont montrés significativement concordants entre échelles ($W > 0,75$, $p < 0,05$) et ont été retenus pour la classification (Figure 11).

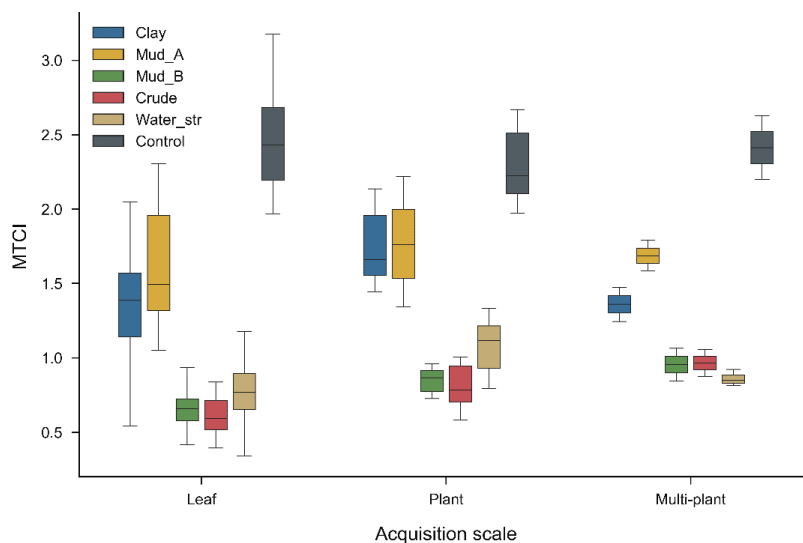


Figure 11. Indice MTCI obtenu à l'échelle de la feuille (*Leaf*), de l'individu (*Plant*) et du couvert (*Multi-plant*) pour les différents traitements après 32 jours. Cet indice a été mis en lien avec les teneurs en chlorophylles et en β -carotène des feuilles ($R^2 = 0,94$) et s'est révélé fortement concordants entre les échelles ($W = 0,90$, $p < 0,05$). (cf. article original pour la signification des traitements.)

La classification menée à partir des 14 indices retenus et entraînée à l'échelle de la feuille uniquement a permis de discriminer les plantes témoins des autres traitements avec précision ($\geq 85\%$) après 18 jours (Tableau 1). Deux types de contamination – regroupant chacun plusieurs traitements – ont également pu être distingués, comme le suggéraient les VI. Les évolutions de l'état de santé des plantes observées pour certains traitements ont ensuite fortement impacté la classification. Après 32 jours, d'importantes confusions ($> 25\%$) sont apparues entre plantes témoins et celles exposées aux bourbiers pauvres en coupes C₅-C₁₀. À l'inverse, l'exposition prolongée des plantes aux autres traitements (bourbier riche en coupes C₅-C₁₀, pétrole brut et déficit en eau) a amélioré leur discrimination avec les plantes témoins (aucune confusion après

32 jours). De manière plus globale, la discrimination des traitements est restée peu affectée par le changement d'échelle de mesure, soulignant ainsi la robustesse de la méthode. À ce titre, son application en conditions naturelles a conduit à la reconnaissance de 83 et 90% des patches de végétation témoins et contaminés, respectivement.

Tableau 1. Matrices de confusion (%) obtenues après 32 jours, à l'échelle de la feuille (*Leaf*) et de l'individu (*Plant*). (UA: *User Accuracy*, PA: *Producer Accuracy*.)

Leaf scale							Plant scale								
	Clay	Mud_A	Mud_B	Crude	Water_str	Control	UA		Clay	Mud_A	Mud_B	Crude	Water_str	Control	UA
Clay	68	14	2	5	3	8	68	Clay	52	25	7	4	3	9	52
Mud_A	4	65	2	0	1	28	65	Mud_A	7	62	0	3	0	28	62
Mud_B	6	1	75	17	1	0	75	Mud_B	14	7	42	22	15	0	42
Crude	3	0	8	83	6	0	83	Crude	13	11	23	36	17	0	36
Water_str	1	0	6	5	88	0	88	Water_str	7	3	14	10	66	0	66
Control	9	26	0	0	0	65	65	Control	2	16	0	0	0	82	82
PA	75	61	81	75	89	64		PA	55	50	49	48	65	69	

5.3. Conclusion et perspectives

Cette expérimentation a confirmé et enrichi les résultats préliminaires obtenus pour *R. fruticosus* lors de la première étude. Elle a démontré l'influence de la composition du mélange pétrolier sur la réponse spectrale de la végétation, soulignant de ce fait l'intérêt du suivi multitemporel pour améliorer la reconnaissance de certains types de contamination. Cette expérimentation a conduit au développement d'une méthode de détection et de caractérisation de la contamination pétrolière applicable en conditions contrôlées et naturelles, de l'échelle de la feuille et du couvert. La sélection de VI capables de suivre les altérations biochimiques des feuilles quelle que soit l'échelle de mesure a grandement contribué à la robustesse de cette méthode. La validation de cette dernière en conditions naturelles, sur un cas de borbier contaminé réaliste, a marqué la fin de la première étape de l'approche adoptée dans la thèse pour le contexte tempéré. Une démarche en partie analogue a été suivie pour l'expérimentation menée en contexte tropical.

6. Troisième article : “Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression”

6.1. Contexte et objectif de l'étude

La troisième et dernière expérimentation s'est focalisée sur *C. alopecuroides*, dans le but de développer une méthode de quantification de la contamination pétrolière en contexte tropical. Pour ce faire, l'espèce a été cultivée en serre sur des sols de borbiers tropicaux contaminés à hauteur de 1 à 19 g.kg⁻¹ de TPH (coupes C₂₁-C₄₀ principalement), ainsi que sur un sol témoin, pendant 42 jours (11 réplicas/traitement). Les contenus en pigments et en eau des feuilles ont été suivis au cours de l'expérimentation et des mesures de réflectance ont été effectuées régulièrement à l'échelle de la feuille (n = 165 mesures/date) et de l'individu (n = 55 mesures/date). La régression ENET a été à nouveau utilisée pour mettre en lien ces données via les 33 VI testés précédemment. Dans un second temps, ces mêmes indices ont servi à développer la méthode de quantification des TPH, basée sur la régression *Random Forest* (RF). Cette dernière repose sur la construction d'arbres de décision combinant les VI par seuillage et renseigne sur la contribution relative de chaque VI à la quantification des TPH. Comme pour les études précédentes, la méthode a été évaluée via une démarche par apprentissage et test (50% des données utilisés à chaque étape), à l'échelle de la feuille et de l'individu. Les prédictions de concentrations en TPH ont été comparées aux analyses de sols à l'aide du R², de l'erreur quadratique moyenne (*Root Mean Squared Error*, RMSE) et de la *Residual Predictive Deviation* (RPD).

6.2. Principaux résultats

Après 42 jours, une réduction significative des contenus en chlorophylles et en caroténoïdes ($p < 0,05$) est apparue chez les plantes exposées à 5 g.kg⁻¹ de TPH ou plus, en comparaison avec les plantes témoins (Figure 12). Celle-ci a induit une augmentation de réflectance à l'échelle de la feuille et de l'individu (5 à 10%) dans le VIS, autour de 550 et 700 nm, et dans une moindre mesure dans le reste du spectre. Plus le sol était contaminé, plus ces modifications étaient marquées. Ces altérations biochimiques ont limité significativement la croissance des parties aériennes. À l'inverse, un effet stimulant a été observé chez les plantes exposées à 1 g.kg⁻¹ de TPH. Le contenu en pigments des feuilles et la biomasse sèche des parties aériennes sont apparus significativement plus élevés ($p < 0,05$) pour ce traitement, conduisant à une diminution de la réflectance à l'échelle de la feuille et de l'individu.

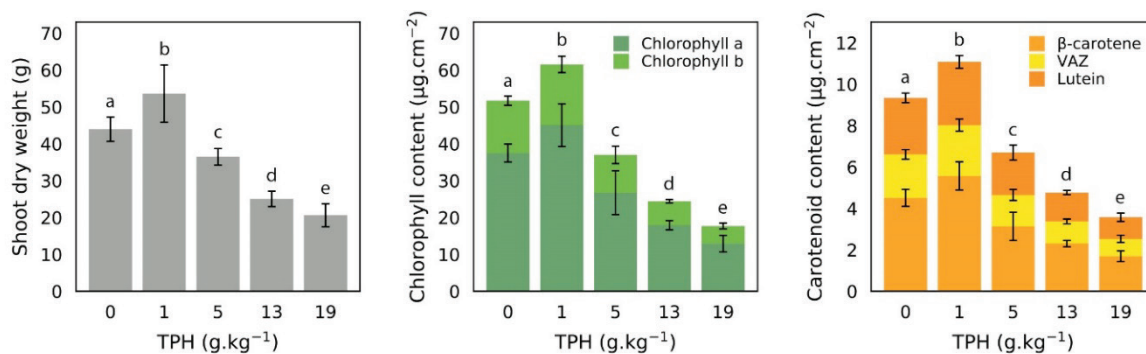


Figure 12. Biomasse sèche des parties aériennes et contenus en chlorophylles et en caroténoïdes des feuilles (moyenne \pm écart-type) pour les différents traitements après 42 jours. Des lettres différentes indiquent les différences significatives entre traitements (ANOVA et test post-hoc de Tukey, $p < 0,05$). (VAZ : violaxanthine + anthéroxanthine + zéaxanthine ; TPH ; Total Petroleum Hydrocarbons.)

20 des 33 VI testés se sont révélés très corrélés aux contenus en pigments et en eau des feuilles ($R^2 \geq 0,7$), en particulier les chlorophylles et le β -carotène. Ces pigments ont conduit à eux seuls aux meilleurs modèles ($R^2 \geq 0,8$) ; d'autres caroténoïdes, notamment la lutéine et la violaxanthine, ont également été retenus par la régression ENET dans les modèles.

La régression RF exploitant les 33 VI a permis de prédire les concentrations en TPH du sol avec précision, aux échelles de la feuille et de l'individu (Figure 13). Les concentrations prédites sur l'échantillon de test se sont montrées très corrélées aux concentrations de référence (i.e. aux analyses de sols) ($R^2 = 0,96$ et $0,95$ à l'échelle de la feuille et de l'individu, respectivement). Les faibles RMSE ($\leq 1,63$ g.kg⁻¹) et les RPD élevées ($\geq 4,44$) ont confirmé les performances de la méthode aux deux échelles de mesure. L'analyse des résidus de la prédiction a révélé que le niveau de contamination le plus élevé (19 g.kg⁻¹) était le plus difficile à prédire et systématiquement sous-estimé. Des confusions sont apparues entre les plantes témoins et celles exposées à 1 g.kg⁻¹ de TPH, en raison des faibles différences de réflectance – et donc de VI – observées aux échelles de la feuille et de l'individu. Quatre des 33 VI utilisés ont contribué à eux seuls à plus de 75% aux prédictions de TPH, aux deux échelles de mesure. Ces indices se trouvaient parmi ceux étroitement liés au contenu en pigments des feuilles. À l'inverse, la contribution des autres indices aux prédictions de TPH est demeurée très limitée ($\leq 5\%$).

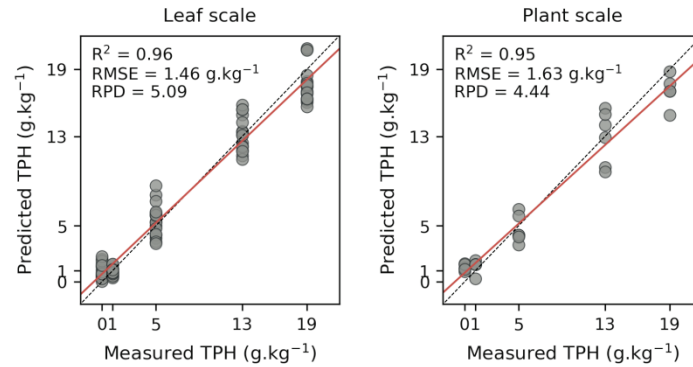


Figure 13. Comparaison entre les concentrations en TPH prédites par la méthode et celles mesurées dans les sols, à l'échelle de la feuille et de l'individu.

6.3. Conclusion et perspectives

Cette troisième expérimentation menée en contexte tropical a démontré la possibilité de quantifier la contamination pétrolière des bourbiers à partir des propriétés optiques de *C. alopecuroides*. Elle a permis de mettre en place une méthode de quantification des TPH, applicable aux échelles de la feuille et de l'individu. Les résultats obtenus en exploitant les 33 VI précédemment utilisés pour *R. fruticosus* ont montré leur robustesse et leur polyvalence en termes de contexte d'application (tempéré, tropical) et de monitoring de la contamination (détection, caractérisation et quantification). Ces résultats ont souligné la difficulté de prédire des niveaux de contamination élevés et de déterminer une limite inférieure de détection (1 g.kg^{-1}) de la méthode. Cette expérimentation ouvre des perspectives encourageantes d'application de la méthode en conditions naturelles, sur le terrain et les images hyperspectrales aéroportées.

CONCLUSION

Les travaux synthétisés dans ce chapitre ont constitué la première étape de l'approche adoptée dans cette thèse. Ils ont permis de développer des méthodes d'évaluation de la contamination pétrolière, en reproduisant fidèlement les contextes d'étude identifiés en conditions naturelles. La première expérimentation a montré l'intérêt de *R. fruticosus* et *C. alopecuroides* pour détecter les sols de borbiers contaminés, tout en identifiant les longueurs d'ondes à exploiter pour le développement des méthodes. Pour le contexte tempéré, une méthode basée sur une sélection de 14 VI utilisés en classification a été proposée à l'issue de la seconde expérimentation. Elle s'est montrée efficace pour détecter et discriminer plusieurs types de contamination, non seulement en conditions contrôlées, mais également sur le terrain. La méthode suivante, développée en contexte tropical, a rendu possible la quantification des TPH dans les sols de borbiers contaminés en combinant 33 VI en régression. La robustesse des méthodes au changement d'échelle de mesure (i.e. feuille, individu et couvert) a souligné leur potentiel en vue de leur application sur les images hyperspectrales aéroportées, abordée en contexte tempéré dans le chapitre suivant.

CHAPITRE TROIS : APPLICATION ET EVALUATION DES METHODES EN CONDITIONS NATURELLES, SUR LE TERRAIN ET LES IMAGES HYPERSPECTRALES AEROPORTEES

INTRODUCTION

La détection et la quantification de la contamination pétrolière par imagerie hyperspectrale aéroportée a constitué l'objectif final de la thèse. Pour cela, la suite des travaux s'est concentrée sur le contexte tempéré, pour lequel une méthode de détection basée sur une sélection de VI avait été proposée et validée sur le terrain. Afin de compléter celle-ci, une méthode de quantification des TPH adaptée à ce contexte a été développée lors de la seconde étape de l'approche ; cette fois, directement sur le terrain. Un suivi des propriétés optiques de la végétation a pour cela été mené à l'échelle de la feuille sur un bournier contaminé par des concentrations variables en TPH. Les données collectées ont servi au développement de la méthode, basée sur l'utilisation du modèle PROSPECT. Celle-ci a ensuite été adaptée à l'échelle du couvert *via* le modèle PROSAIL. Une seconde méthode de quantification des TPH combinant les transformations de signatures spectrales et la régression multiple a également été mise au point. La dernière étape de l'approche a consisté à tester et valider et comparer l'ensemble des méthodes (détection et quantification) sur les images hyperspectrales aéroportées, sur divers sites témoins et bourniers contaminés identifiés sur le terrain. Cette dernière étape a permis de déterminer les performances et les limites des méthodes et leurs perspectives d'applications opérationnelles.

7. Quatrième article : “Application of PROSPECT for estimating Total Petroleum Hydrocarbons in contaminated soils from leaf optical properties”

7.1. Contexte et objectif de l'étude

La seconde étape de l'approche de la thèse s'est déroulée sur un ancien bourbier de production contaminé et colonisé par la végétation. Elle devait permettre de développer une méthode de quantification des TPH exploitant les propriétés optiques de la végétation, applicable par la suite aux images aéroportées. La méthode proposée a donc été tout d'abord évaluée à l'échelle de la feuille, à partir de mesures de réflectance conduites sur les espèces présentes sur le bourbier : *R. fruticosus* L., *Quercus pubescens* Wild., *Populus x canadensis* Moench., *Cornus sanguinea* L. et *Buddleja davidii* Franch (la ronce étant l'espèce prépondérante). Pour ce faire, un maillage de 24 points a été défini sur le site. En chacun de ces points, les sols ont été analysés, révélant des concentrations en TPH s'étalant de 16 à 77 g.kg⁻¹. Trois jeunes feuilles ont été échantillonnées aléatoirement pour chaque espèce présente dans un rayon de 1 m autour de chaque point, puis leur réflectance a été mesurée. Cet échantillonnage a été répété chaque saison durant de 14 mois, soit à six dates. Un suivi analogue a été conduit aux mêmes dates sur un site témoin, non contaminé. Un total de 1620 feuilles a été collecté et mesuré, et 152 d'entre elles ont fait l'objet d'analyses de contenus en pigments, en eau et en matière sèche. Ces données ont servi au développement de la méthode de quantification des TPH, basée sur l'inversion du modèle PROSPECT-5. Dans son sens direct, PROSPECT permet de simuler la signature spectrale d'une feuille à partir de ses paramètres biophysiques et biochimiques, notamment les pigments. En inversant le modèle, il est possible de retrouver ces paramètres à partir de signatures spectrales mesurées. La procédure d'inversion est détaillée dans le quatrième article du chapitre III. PROSPECT a ainsi été inversé sur les 1620 signatures spectrales mesurées sur le bourbier et le site témoin. Les estimations des paramètres biochimiques ont été évaluées sur les 152 feuilles ayant fait l'objet d'analyses, à l'aide du R² et de la RMSE. Dans un deuxième temps, l'étude s'est concentrée sur les contenus en pigments – chlorophylles et caroténoïdes – estimés par ces inversions. Ces contenus ont tout d'abord été comparés entre le site témoin et le bourbier par ANOVA et test post-hoc de Tukey, en tenant compte de l'espèce et de la date de prélèvement. Ces mêmes données ont été ensuite utilisées pour prédire les concentrations en TPH dans les sols du bourbier. Pour cela, les données ont été divisées en deux lots (50% apprentissage et 50% test). Des modèles univariés de régression ont été ajustés entre les contenus en pigments des feuilles et les concentrations en TPH sur le premier lot (apprentissage), puis appliqués sur le second (test) pour prédire ces concentrations. Les prédictions de TPH ont ensuite été comparées aux concentrations issues des analyses de sols, à l'aide du R², de la RMSE et de la RPD.

7.2. Principaux résultats

Les signatures spectrales des feuilles mesurées sur le bourbier sont apparues très différentes entre les espèces, soulignant leurs différences d'anatomie et de biochimie. Leur évolution saisonnière très marquée, en particulier autour de 550, 680 et 700 nm, a suggéré des changements de contenus en pigments, confirmés par la suite grâce aux inversions de PROSPECT. Ce dernier s'est montré très efficace pour simuler les signatures spectrales mesurées sur le terrain (RMSE ≤ 2% de réflectance) et pour déterminer les contenus en chlorophylles (RMSE ≤ 2,95 µg.cm⁻²), en caroténoïdes (RMSE ≤ 1.82 µg.cm⁻²), en eau et en matière sèche correspondants. Le contenu en pigments des feuilles s'est révélé significativement différent entre espèces et saisons, non

seulement sur le bournier, mais également sur le site témoin. Afin de séparer la part de variabilité naturelle de ces contenus en pigments de celle due à l'exposition à la contamination sur le bournier, ceux-ci ont été exprimés relativement aux contenus observés sur le site témoin (Figure 14). Cela a permis de mieux comparer la sensibilité à la contamination entre les espèces, ainsi que son évolution saisonnière. Ainsi, quatre des cinq espèces étudiées – dont *R. fruticosus* – ont montré des contenus en chlorophylles et en caroténoïdes significativement plus faibles sur le bournier que sur le site témoin ; *Q. pubescens* étant la plus affectée. L'amplitude de ces altérations a varié entre saisons, mais est restée identique entre les deux années pour une même saison (juin et octobre).

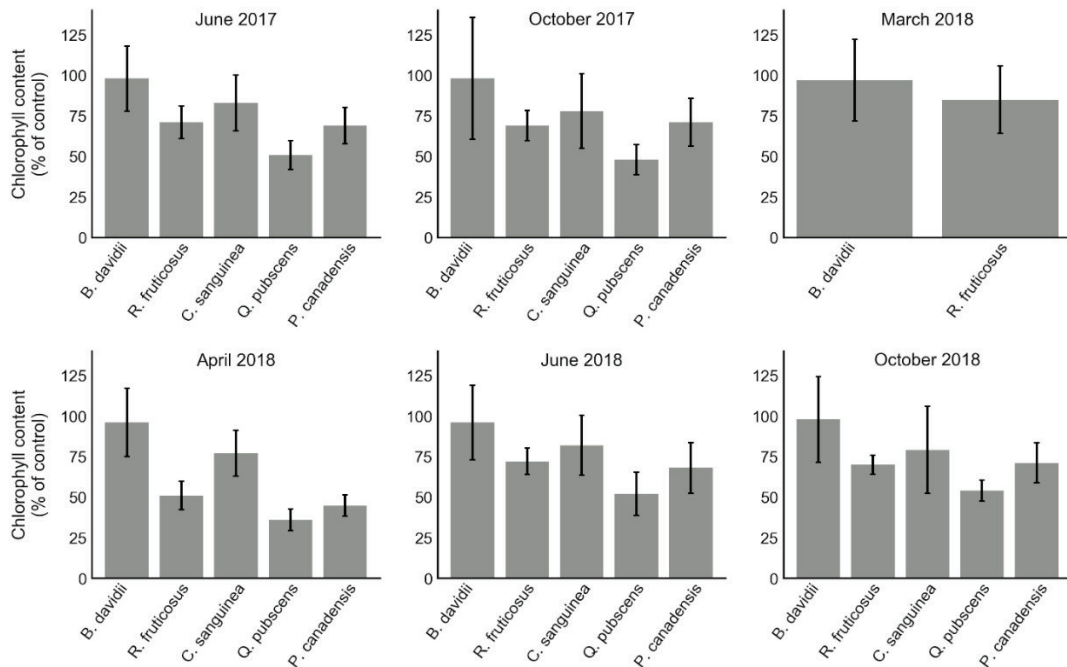


Figure 14. Evolutions temporelles du contenu en chlorophylles des feuilles relatif au témoin pour les différentes espèces présentes sur le bournier.

Trois des quatre espèces affectées par la contamination ont rendu possible la quantification des TPH à partir des contenus en pigments des feuilles : *R. fruticosus*, *Q. pubescens* et *P. canadensis*. Pour ces trois espèces, une forte corrélation négative a été observée en juin et en octobre entre le contenu en chlorophylles estimés par PROSPECT et la concentration en TPH du sol ($R^2 \geq 0,74$) (Figure 15). En appliquant cette relation sur le lot de données test, il a été possible de prédire avec précision les concentrations en TPH ($RPD \geq 2$). Les niveaux de contamination les plus élevés ($\geq 30 \text{ g.kg}^{-1}$) sont apparus les plus difficiles à prédire, car le contenu en chlorophylles tendait à se stabiliser à partir d'une certaine concentration en TPH. Le passage de l'été à l'automne n'a eu que peu d'incidence sur la précision de ces prédictions, en comparaison avec les autres saisons. Les feuilles issues de la croissance printanière (avril) n'ont en effet pas permis de quantifier les TPH, probablement en raison d'une sensibilité trop importante des espèces à la contamination. De même, les prédictions de concentrations en TPH obtenues à partir du contenu en caroténoïdes sont apparues bien moins précises que celles obtenues à partir du contenu en chlorophylles ($RPD < 2$) et ce, quelle que soit l'espèce et la saison. Enfin, les résultats ont été comparés à ceux obtenus par deux autres méthodes issues de la littérature : les

transformations de signatures spectrales combinées à la PLSR et les VI. De manière systématique, la méthode proposée a conduit à de meilleures prédictions des concentrations en TPH.

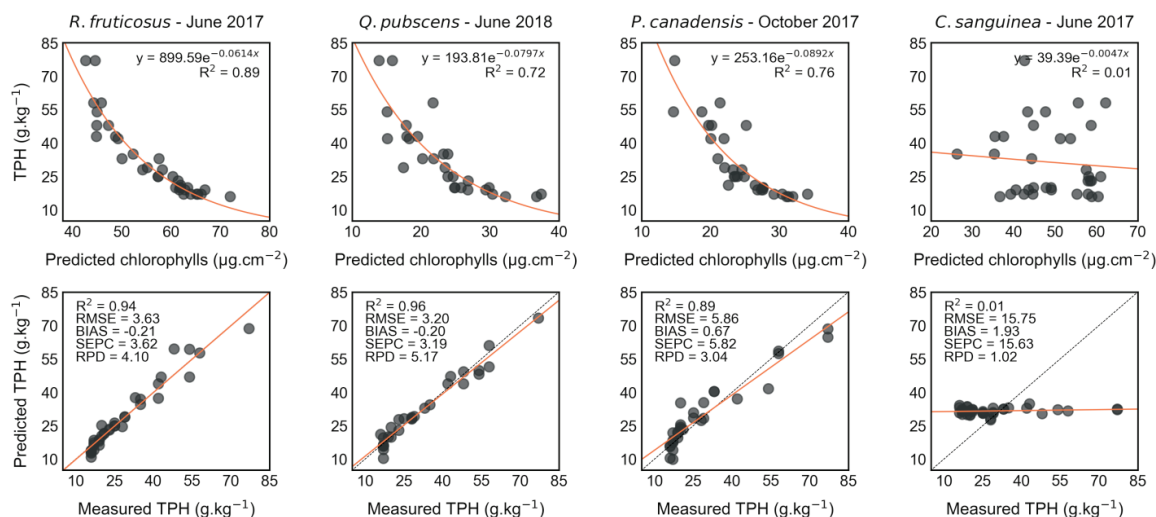


Figure 15. (figures du haut) Relation entre le contenu en chlorophylles des feuilles estimé par inversion de PROSPECT et la concentration en TPH du sol observée sur le lot d'apprentissage, et (figures du bas) comparaison entre les concentrations en TPH prédites et mesurées sur le lot de test.

7.3. Conclusion et perspectives

Cette seconde étape de l'approche suivie dans la thèse a démontré qu'il est possible de quantifier la contamination des sols de borbiers à partir de la réponse spectrale de la végétation, en conditions naturelles. Elle a conduit au développement d'une méthode de quantification basée sur l'inversion du modèle PROSPECT adaptée non seulement à *R. fruticosus*, mais également à d'autres espèces à fort potentiel. A ce titre, cette étude a souligné l'importance de la variabilité interspécifique et saisonnière de la sensibilité à la contamination ; celle-ci affectant grandement les contenus en pigments des feuilles et par conséquent les prédictions de concentrations en TPH. La méthode de quantification proposée est venue compléter la méthode de détection basée sur les VI. Les deux méthodes ayant été validées sur le terrain, la dernière étape de l'approche a consisté à évaluer leurs performances sur les images hyperspectrales aéroportées, toujours en contexte tempéré.

8. Cinquième article : “Toward quantifying oil contamination in vegetated areas using very high spatial and spectral resolution imagery”

8.1. Contexte et objectif de l'étude

La troisième et dernière étape de l'approche de cette thèse a eu pour objectif d'appliquer les méthodes de détection et de quantification de la contamination sur les images hyperspectrales aéroportées. Celles-ci ont été acquises en région tempérée, sur des sites témoins et des broussiers contaminés colonisés par *R. fruticosus* (Figure 16a-f). Les images en luminance spectrale ont été acquises en juillet 2017 à l'aide des caméras Hyspex VNIR-1600 et SWIR-230m-e couvrant le domaine 400 – 2500 nm avec des résolutions spectrales de 5,2 et 7,8 nm et spatiales de 1 et 2,5 m, respectivement. L'image SWIR a été ré-échantillonnée à la résolution spatiale de l'image VNIR (1 m), puis ces images ont été recalées à l'aide de l'algorithme GEFOLKI⁹⁷ et corrigées de l'atmosphère pour obtenir une image en réflectance spectrale par l'*Empirical Line Method*⁹⁸ (ELM). Chaque méthode (détection et quantification) a été calibrée puis testée (50% apprentissage et 50% test) sur les images, sur un site témoin et le broussier principal – celui ayant servi à la validation des méthodes sur le terrain, et enfin validée sur des sites spatialement indépendants. Sur le broussier principal, 10 parcelles de 16 à 20 m² contaminées par 17 à 39 g.kg⁻¹ de TPH ont été définies. Pour l'étape de validation, cinq sites témoins et quatre broussiers contaminés par 0,25, 0,38, 3,15 et 24 g.kg⁻¹ de TPH ont été identifiés.

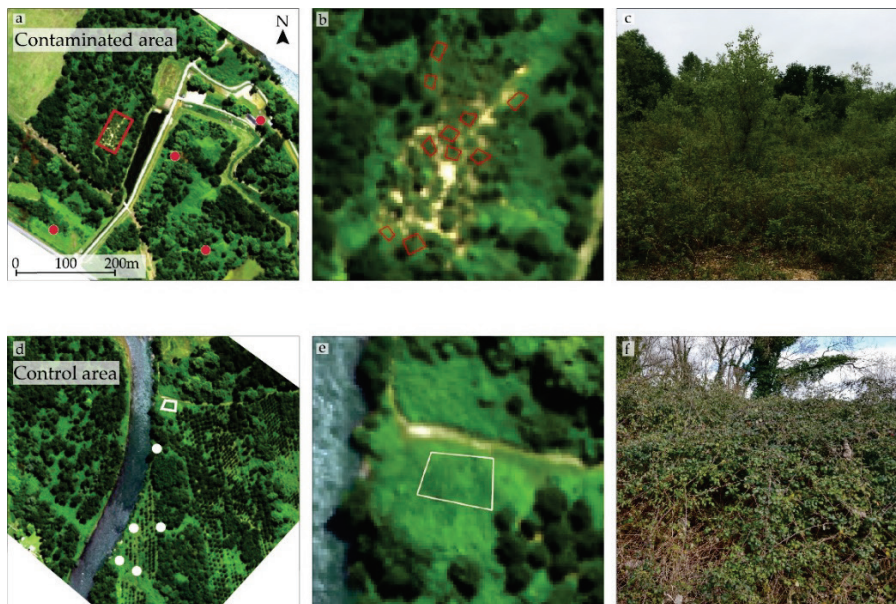


Figure 16. (a, d) Extraits des images hyperspectrales aéroportées acquises sur les sites (a) contaminés et (b) témoins. Les sites ayant servi à la calibration des méthodes sont illustrés par des polygones. Les cercles correspondent aux sites de validation (en rouge : broussier contaminé, en blanc : témoin). (b, e) Zoom sur (b) le broussier principal (10 parcelles) et (e) le site témoin, utilisés pour calibrer les méthodes. (c, f) *R. fruticosus* (c) sur le broussier et (f) sur le site témoin (photos prises en mars).

La méthode de détection – basée sur la classification RLR – a été appliquée sur les pixels correspondants aux sites, après calcul des 14 VI. La méthode de quantification des TPH –

initialement développée à l'échelle de la feuille – a quant à elle été adaptée dans un premier temps à l'échelle de la canopée, puis des images, à l'aide du modèle PROSAIL²² (PROSPECT + SAIL). Des mesures de réflectance ont pour cela été conduites aux échelles de la feuille et de la canopée sur les 10 parcelles du bournier principal. Les modèles PROSPECT et PROSAIL ont ensuite été inversés sur ces données et sur les pixels correspondants sur les images, afin d'estimer le contenu en chlorophylles des feuilles aux différentes échelles et de les comparer. Les concentrations en TPH dans les sols de bourniers ont ensuite été déterminées à partir de ces contenus en chlorophylles, selon une démarche analogue à l'étude précédente. Enfin, la méthode a été appliquée sur le bournier contaminé par 24 g.kg⁻¹ de TPH pour validation. D'autre part, une seconde méthode de quantification des TPH a été développée. Basée sur les transformations de signatures spectrales utilisées dans la toute première expérimentation, cette méthode a fait appel à la régression multiple ENET afin de prédire les concentrations en TPH à partir des données de réflectance – transformées ou non, tout en identifiant les bandes spectrales pertinentes.

8.2. Principaux résultats

La méthode de détection de la contamination a permis de discriminer très précisément les pixels de végétation témoins de ceux du bournier principal contaminé (précision = 99 et 95,7%, respectivement). Les parcelles les moins contaminées (17 – 18 g.kg⁻¹ de TPH) se sont montrées logiquement les moins évidentes à classifier. Ces résultats ont confirmé la robustesse de la méthode au changement d'échelle de mesure, et notamment au passage sur les images aéroportées. La méthode est également parvenue à identifier la végétation témoin sur les cinq sites de validation (précision = 91,8%). En revanche, le taux de bonne classification sur les quatre bourniers de validation s'est révélé très dépendant du niveau de contamination. Un seul de ces quatre bourniers présentait une concentration en TPH (24 g.kg⁻¹) située dans la gamme de celles de la méthode (17 – 39 g.kg⁻¹) ; celui-ci a été très bien reconnu (précision = 95,5%). Des résultats encourageants ont également été obtenus à plus faible concentration, en-dehors de cette gamme (précision = 85,4% à 3,15 g.kg⁻¹ de TPH). Les limites de la méthode ont toutefois été atteintes sur les bourniers très faiblement contaminés (précision \leq 15% à 0,25 et 0,38 g.kg⁻¹ de TPH) ; *R. fruticosus* étant peu affectée.

Sur le bournier principal, les inversions des modèles PROSPECT et PROSAIL ont conduit à des estimations très cohérentes du contenu en chlorophylles des feuilles entre les échelles de mesure, sur les 10 parcelles du site. Ainsi, le contenu estimé à l'échelle de la feuille a été facilement retrouvé à l'échelle de la canopée ($R^2 = 0,95$ et $RMSE = 1,05 \mu\text{g.cm}^{-2}$) et des images aéroportées ($R^2 = 0,85$ et $RMSE = 3,70 \mu\text{g.cm}^{-2}$). Une variabilité plus importante a toutefois été observée au sein des parcelles sur les images, en comparaison avec les données de terrain. La corrélation négative entre le contenu en chlorophylles des feuilles et la concentration en TPH dans les sols décrite dans l'étude précédente a également été observée, aux trois échelles de mesure ($R^2 \geq 0,74$) (Figure 17a-c). Elle a permis de prédire avec précision les concentrations en TPH sur le bournier principal, y compris à partir des images aéroportées ($RPD \geq 2$) (Figure 17d-f). Le changement d'échelle a toutefois affecté les prédictions et, une fois encore, les niveaux de contamination les plus élevés ($\geq 30 \text{ g.kg}^{-1}$ de TPH) ont été les plus difficiles à quantifier. L'application de la méthode sur le bournier contaminé par 24 g.kg⁻¹ de TPH a conduit à des prédictions très précises ($23,69 \pm 1,93 \text{ g.kg}^{-1}$ de TPH), validant son utilisation sur les images aéroportées.

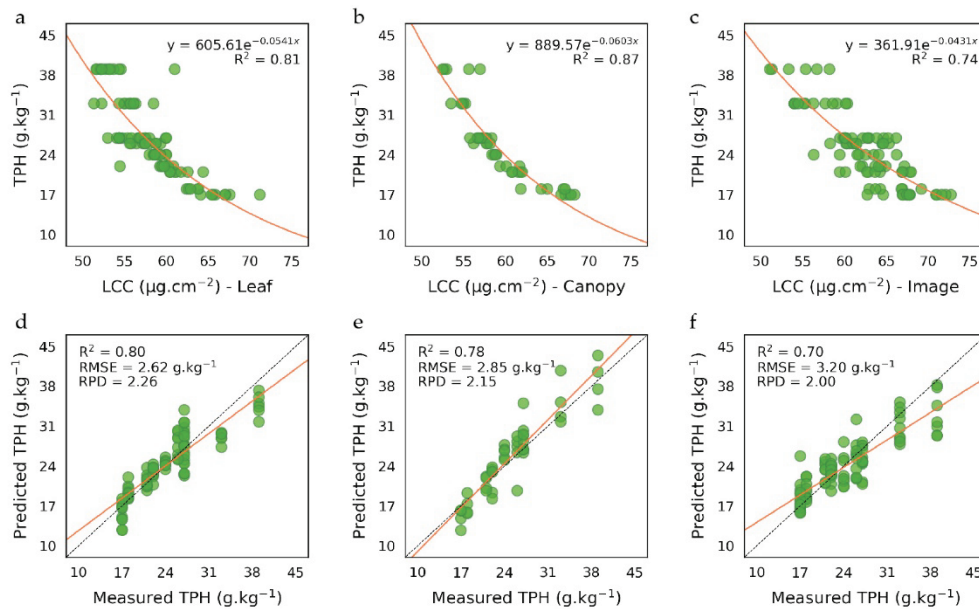


Figure 17. (a-c) Relation entre le contenu en chlorophylles des feuilles et la concentration en TPH dans le sol (lot d'apprentissage) observée sur les 10 parcelles du bournier principal, à l'échelle (a) de la feuille, (b) de la canopée et (c) des images aéroportées. (d-f) Comparaison entre les concentrations en TPH prédites et mesurées (lot de test), à l'échelle (d) de la feuille, (e) de la canopée et (f) des images aéroportées.

La méthode de quantification basée sur la régression ENET a également prédit des concentrations en TPH très proches de celles mesurées, aux trois échelles de mesure, à partir des données en réflectance non-transformées ou dérivées – première et seconde (meilleures $\text{RPD} \geq 1,90$). Les autres transformations de signatures spectrales (SNV, AUCN, etc.) n'ont pas permis d'améliorer la quantification. La méthode a également été validée sur le bournier contaminé par 24 g.kg^{-1} de TPH, grâce à des prédictions très précises ($23,39 \pm 2,41 \text{ g.kg}^{-1}$ de TPH). L'analyse des bandes spectrales ayant contribué aux meilleures prédictions a révélé que celles-ci se concentrent essentiellement autour de 550 et 700 nm, confirmant l'intérêt des pigments pour quantifier la contamination en comparaison avec les autres paramètres biochimiques comme le contenu en eau ou en matière sèche, peu représentés par les bandes du NIR et du SWIR.

8.3. Conclusion et perspectives

En validant les méthodes de détection et de quantification de la contamination sur les images hyperspectrales aéroportées, cette étude a marqué la fin de l'approche de la thèse. Elle a ainsi démontré la robustesse des méthodes développées en conditions contrôlées et sur le terrain ; qu'elles exploitent la réflectance directement (régression ENET), ou indirectement au travers des VI (classification RLR) et des modèles PROSPECT et PROSAIL (estimation du contenu en chlorophylles). En supposant un couvert homogène de *R. fruticosus* de localisation connue, il est donc possible de détecter et quantifier la contamination des bourniers à partir d'images à très hautes résolutions spatiale et spectrale. Cette étude a également permis d'approcher une limite de détection de la contamination ($< 3,15 \text{ g.kg}^{-1}$ de TPH), en-dessous de laquelle la plante n'est pas affectée ou du moins trop peu pour que cela soit détectable par imagerie.

CONCLUSION

Les deux études présentées dans ce chapitre ont constitué la deuxième et la troisième étape de l'approche adoptée dans cette thèse. Elles ont conduit à la validation de la méthode de détection de la contamination – basée sur les 14 VI – sur les images hyperspectrales aéroportées, sur des sites indépendants de ceux ayant servi à la calibrer. Ces études ont également permis de proposer et de valider deux méthodes de quantification des TPH dans les sols de bourbiers ; l'une basée sur l'inversion du modèle PROSAIL et l'autre sur les transformations de signatures spectrales et la régression ENET. Au terme des trois étapes de l'approche, l'ensemble des méthodes développées ont été validées pour une utilisation à l'échelle de la feuille, de la canopée et des images hyperspectrales aéroportées à haute résolution spatiale. Dans une perspective d'application opérationnelle, il serait nécessaire d'étendre le champ d'application des méthodes à d'autres espèces et, plus largement, à d'autres contextes (tropical notamment). De même, l'influence des résolutions spatiale et spectrale devrait être étudiée en vue d'appliquer les méthodes sur les capteurs satellitaires multi- et hyperspectraux actuels et futurs. A cet effet, la méthode de quantification des TPH basée sur l'inversion du modèle PROSAIL a été testée à la résolution spectrale du futur capteur satellitaire hyperspectral HypXim (210 bandes spectrales prévues), ainsi qu'à celle du capteur satellitaire multispectral Worldview-3 (16 bandes spectrales). La résolution spatiale initiale (1 m) est restée inchangée. Le contenu en chlorophylles des feuilles s'est montré très proche de celui obtenu sur le terrain à la résolution d'HypXim, mais beaucoup moins à celle de Worldview-3 (Figure 18). Par conséquent, les prédictions de TPH se sont révélées satisfaisantes dans le premier cas, mais beaucoup moins précises dans le second.

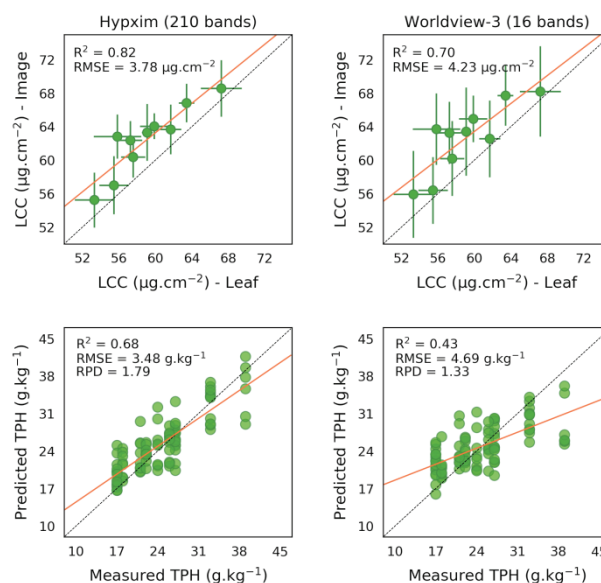


Figure 18. (figures du haut) Comparaison entre le contenu en chlorophylles des feuilles (moyenne \pm écart-type) estimé sur le terrain (mesures à l'échelle de la feuille) et sur les images aéroportées, et (figures du bas) comparaison entre les concentrations en TPH prédites et mesurées, aux résolutions spectrales des capteurs HypXim et Worldview-3.

Ces résultats démontrent le potentiel des futurs capteurs satellitaires hyperspectraux pour quantifier la contamination des bourbiers, ainsi que les limites du capteur multispectral WorldView-3. Dans la continuité de ce travail, il serait nécessaire de prendre en compte la

résolution spatiale et le bruit du capteur, très certainement déterminante dans les performances des méthodes.

CONCLUSION GENERALE

Cette thèse avait pour objectif de détecter et quantifier le pétrole brut et les produits pétroliers à partir d'images hyperspectrales aéroportées en régions végétalisées. Pour y parvenir, une approche multi-échelles en trois étapes a été appliquée à l'étude des bourbiers contaminés en contexte tempéré, impliquant *R. fruticosus* comme espèce-modèle. À l'issue de ces étapes, trois méthodes ont été développées pour une utilisation sur des mesures de terrain (échelles de la feuille et du couvert) et sur des images hyperspectrales aéroportées ; une méthode servant à détecter et caractériser la contamination et deux autres permettant de la quantifier.

Lors de l'étape initiale de l'approche, une première expérimentation en conditions contrôlées a permis de déterminer à la fois la sensibilité de l'espèce à la contamination et les régions du spectre affectées. Ces informations ont ensuite été exploitées lors d'une seconde expérimentation, afin de développer la méthode de détection et de caractérisation de la contamination. Basée sur la combinaison de 14 VI en classification (RLR), cette méthode a permis de discriminer des plantes témoins de plantes exposées à divers contaminants pétroliers ou à un déficit en eau (précision $\geq 85\%$ après 18 jours). La première étape de l'approche a pris fin avec la validation de la méthode sur le terrain. Lors de l'étape suivante, une méthode de quantification des TPH basée sur l'inversion du modèle PROSPECT a été mise en place à partir d'un suivi multitemporel de cinq espèces. Elle repose sur l'estimation du contenu en pigments des feuilles à partir de leur signature spectrale, lequel permet de prédire la concentration en TPH du sol. La méthode s'est montrée très performante pour *R. fruticosus* et deux autres espèces ($RPD \geq 3,04$) en saisons estivale et automnale. Enfin, la troisième étape de l'approche a permis de valider l'utilisation des méthodes de détection (précision $\geq 90\%$) et de quantification ($RPD \geq 2$) de la contamination sur les images aéroportées. Une seconde méthode combinant les transformations de signatures spectrales et la régression ENET a été proposée. Elle a également conduit à des prédictions précises de concentrations en TPH ($RPD \geq 1,90$).

L'approche suivie au cours de cette thèse a démontré qu'il est possible de détecter et quantifier la contamination pétrolière en exploitant les propriétés optiques de la végétation, à partir d'images hyperspectrales aéroportées. Une approche analogue a été adoptée en contexte tropical, sur *C. alopecuroides*. Elle a conduit au développement d'une méthode de quantification des TPH basée sur la combinaison de VI en régression RF ($RPD \geq 4,44$).

PERSPECTIVES

À ce stade, l'utilisation des méthodes proposées pour détecter et quantifier la contamination par imagerie aéroportée nécessite de connaître la sensibilité de l'espèce ciblée, ainsi que sa localisation. Dans une perspective d'application opérationnelle, il est essentiel que ces méthodes soient applicables à grande échelle et couvrent un large éventail de contextes, aussi bien en termes d'espèces que de types et de niveaux de contamination pétrolière. Les futurs capteurs hyperspectraux satellitaires présentent un fort intérêt pour l'exploration pétrolière et le monitoring continu de la contamination à grande échelle, en raison de leur importantes couvertures spatiale et temporelle. En revanche, ces capteurs fourniront des images à des résolutions spatiales de 8 ou 30 m, sur moins de 250 bandes spectrales. À de telles résolutions, les méthodes de détection et de quantification verraient leurs performances largement affectées.

La première voie d'amélioration des méthodes proposées en contexte tempéré, en vue d'une application à grande échelle, serait d'intégrer une cartographie automatique de *R. fruticosus*. Cela permettrait d'appliquer les méthodes sur des images entières, tout en limitant l'apparition de fausses alarmes dues à la présence d'autres espèces de sensibilité à la contamination inconnue. Cette étape pourrait toutefois constituer une difficulté majeure aux résolutions spatiales des futurs capteurs satellitaires, où les espèces sont bien souvent mélangées entre elles – et à du sol nu – au sein des pixels. Dans ce cas, l'utilisation de méthodes de démixage spectral^{16,99} apparaît indispensable, mais reste à évaluer dans le contexte de la contamination pétrolière. Le démixage spectral permettrait d'identifier les différentes espèces présentes au sein des pixels et donc d'isoler ceux dominés par les espèces-cibles. Il serait alors judicieux d'étendre le champ d'application des méthodes à d'autres espèces – telles que *Q. pubescens* et *P. canadensis*, également indicatrices de la contamination. Une caractérisation de leur sensibilité aux mélanges pétroliers en serre ou sur le terrain, telle que présentée dans cette thèse, permettrait d'identifier ces espèces indicatrices et d'adapter les méthodes de détection et de quantification de la contamination. L'utilisation d'une bibliothèque spectrale recensant les espèces indicatrices en démixage permettrait de localiser celles-ci sur les images satellitaires et d'appliquer les méthodes de détection et de quantification correspondantes. Des méthodes classification des espèces, exploitant notamment des données multitemporelles, pourraient également être adaptées. Une procédure analogue basée sur la méthode de quantification développée avec *C. alopecuroides* permettrait d'évaluer la contamination des brousses à grande échelle en régions tropicales.

À ce stade, les méthodes demeurent précises dans une gamme de concentrations en TPH limitée par celle des brousses étudiés et dépendante de la sensibilité des espèces présentes (17 – 39 ou 77 g.kg⁻¹ de TPH). En vue d'une application opérationnelle, il serait nécessaire d'évaluer leurs performances à des concentrations à la fois inférieures et supérieures à cette gamme, tel qu'initié avec la méthode de détection (TPH < 1 et = 3,15 g.kg⁻¹). À ce titre, l'identification d'espèces indicatrices aux sensibilités variables aux TPH permettrait de couvrir une gamme très large de concentrations. Il reste toutefois difficilement envisageable de quantifier précisément les TPH aux résolutions spatiales des capteurs satellitaires, car les pixels pourraient inclure de la végétation exposée à des niveaux variables de contamination. Un important effort demeure nécessaire pour identifier les espèces indicatrices, ainsi que la gamme de concentrations en TPH que chacune permet de détecter et de quantifier, à des résolutions spatiales variables.

L'application des méthodes à l'évaluation de la contamination des brousses ne saurait satisfaire l'ensemble des besoins opérationnels des compagnies pétrolières. Celle-ci doit s'étendre à d'autres scénarios de l'exploration pétrolière et du monitoring de la contamination. Par exemple, la détection du pétrole brut présente un enjeu majeur pour la prospection des suintements et le suivi d'installations (e.g. oléoducs, cuves de stockage). À cet effet, la méthode

exploitant les 14 VI s'est montrée prometteuse en conditions contrôlées. Afin de quantifier les concentrations en TPH du pétrole, il serait toutefois nécessaire d'adapter les méthodes basées sur l'inversion du modèle PROSAIL et de la régression ENET à ce type de mélange pétrolier. Il est toutefois probable que les performances de ces méthodes soient limitées aux résolutions spatiales des capteurs satellitaires en cas de contamination très localisée (i.e. quelques m²), car les pixels n'incluraient que rarement uniquement de la végétation exposée.

En perspective des travaux menés dans cette thèse, il apparaît ainsi nécessaire d'élargir le champ d'application des méthodes proposées et d'évaluer leur maturité opérationnelle en imagerie satellitaire. À long terme, l'utilisation de capteurs hyperspectraux embarqués sur drone est également envisagée par les compagnies pétrolières^{100,101}. Actuellement en développement, ces capteurs permettraient un monitoring permanent et localisé, grâce à des très hautes résolutions spatiale (de l'ordre du cm) et spectrale et à une répétitivité temporelle des acquisitions sur mesure. Enfin, la combinaison de l'imagerie passive hyperspectrale avec des technologies de télédétection active, notamment le radar et le LiDAR, pourrait améliorer la détection et la quantification du pétrole brut et des produits pétroliers en apportant des informations complémentaires sur la végétation. En combinant ainsi les technologies active et passive et en exploitant les différentes plateformes de capteurs (satellite, drone), il est fort probable que la télédétection deviendra un outil indispensable à l'exploration pétrolière et au monitoring de la contamination en régions végétalisées dans les prochaines décennies.

VALORISATION DE LA THESE

PUBLICATIONS (REVUES A COMITE DE LECTURE)

- Lassalle, G.; Elger, A.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Fabre, S. Toward quantifying oil contamination in vegetated areas using high spatial and spectral resolution imagery. *Remote Sens.* **2019**, 11, 2241. <https://doi.org/10.3390/rs11192241>
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Estimating persistent oil contamination in tropical region using vegetation indices and Random Forest regression. *Ecotox. Environ. Safe.* **2019**, 184, 109654. <https://doi.org/10.1016/j.ecoenv.2019.109654>.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Application of PROSPECT for Estimating Total Petroleum Hydrocarbons in Contaminated Soils from Leaf Optical Properties. *J. Hazard. Mater.* **2019**, 377, 409–417. <https://doi.org/10.1016/j.jhazmat.2019.05.093>.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Erudel, T.; Buffan-Dubau, E.; Dubucq, D.; Elger, A. Detection and Discrimination of Various Oil-Contaminated Soils Using Vegetation Reflectance. *Sci. Total Environ.* **2019**, 655, 1113–1124. <https://doi.org/10.1016/j.scitotenv.2018.11.314>.
- Lassalle, G.; Credoza, A.; Hédacq, R.; Fabre, S.; Dubucq, D.; Elger, A. Assessing Soil Contamination Due to Oil and Gas Production Using Vegetation Hyperspectral Reflectance. *Environ. Sci. Technol.* **2018**, 52 (4), 1756–1764. <https://doi.org/10.1021/acs.est.7b04618>.

COMMUNICATIONS

- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Bertoni, G.; Dubucq, D.; Elger, A. Utilisation de données hyperspectrales pour la détection et la quantification des hydrocarbures pétroliers en régions végétalisées. SFPT-GH **2019**, Toulouse, France.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Dubucq, D.; Elger, A. Detection and quantification of total petroleum hydrocarbons in soils using vegetation optical properties. IEEE IGARSS **2019**, Yokohama, Japan.
- Lassalle, G.; Fabre, S.; Credoza, A.; Hédacq, R.; Borderies, P.; Bertoni, G.; Dubucq, D.; Elger, A. Detection and quantification of oil contamination in vegetated areas using hyperspectral remote sensing. SETAC **2018**, Roma, Italy.
- Lassalle, G.; Credoza, A.; Hédacq, R.; Fabre, S.; Dubucq, D.; Elger, A. Hyperspectral signature analysis of three plant species to long-term hydrocarbon and heavy metal exposure. Proceedings of the SPIE Remote Sensing 2017, Warsaw, Poland. <https://doi.org/10.1117/12.2277709>

REFERENCES

- (1) Miller, R. G.; Sorrell, S. R.; Lane, H.; Kt, S. The Future of Oil Supply. *Philos. Trans. R. Soc. A* **2014**, *372*, 20130179. <https://doi.org/10.1098/rsta.2013.0179>.
- (2) OPEC. The Organization of the Petroleum Exporting Countries Annual Report 2017. **2018**.
- (3) Zou, C.; Zhao, Q.; Zhang, G.; Xiong, B. Energy Revolution: From a Fossil Energy Era to a New Energy Era. *Nat. Gas Ind. B* **2016**, *3*, 1–11. <https://doi.org/10.1016/j.ngib.2016.02.001>.
- (4) Noomen, M. F.; van der Werff, H. M. A.; van der Meer, F. D. Spectral and Spatial Indicators of Botanical Changes Caused by Long-Term Hydrocarbon Seepage. *Ecol. Inform.* **2012**, *8* (February), 55–64. <https://doi.org/10.1016/j.ecoinf.2012.01.001>.
- (5) Tangestani, M. H.; Validabadi, K. Mineralogy and Geochemistry of Alteration Induced by Hydrocarbon Seepage in an Evaporite Formation; a Case Study from the Zagros Fold Belt, SW Iran. *Appl. Geochemistry* **2014**, *41*, 189–195. <https://doi.org/10.1016/j.apgeochem.2013.12.015>.
- (6) Shi, P.; Fu, B.; Ninomiya, Y.; Sun, J.; Li, Y. Multispectral Remote Sensing Mapping for Hydrocarbon Seepage-Induced Lithologic Anomalies in the Kuqa Foreland Basin, South Tian Shan. *J. Asian Earth Sci.* **2012**, *46*, 70–77. <https://doi.org/10.1016/j.jseaes.2011.10.019>.
- (7) Tedesco, S. A. *Surface Geochemistry in Petroleum Exploration*; Springer Science Business Media, Ed.; Springer US: Boston, MA, 1995. <https://doi.org/10.1007/978-1-4615-2660-5>.
- (8) Sheriff, R. E.; Geldart, L. P. *Exploration Seismology*, 2nd ed.; Cambridge University Press, 1995. <https://doi.org/10.1017/CBO9781139168359>.
- (9) Asadzadeh, S.; de Souza Filho, C. R. Spectral Remote Sensing for Onshore Seepage Characterization: A Critical Overview. *Earth-Science Rev.* **2017**, *168*, 48–72. <https://doi.org/10.1016/j.earscirev.2017.03.004>.
- (10) Salati, S.; Van Ruitenbeek, F. J. A.; De Smeth, J. B.; Van Der Meer, F. D. Spectral and Geochemical Characterization of Onshore Hydrocarbon Seep-Induced Alteration in the Dezful Embayment, Southwest Iran. *Am. Assoc. Pet. Geol. Bull.* **2014**, *98* (9), 1837–1857. <https://doi.org/10.1306/04051413105>.
- (11) Chang, J. I.; Lin, C. C. A Study of Storage Tank Accidents. *J. Loss Prev. Process Ind.* **2006**, *19*, 51–59. <https://doi.org/10.1016/j.jlp.2005.05.015>.
- (12) Credoz, A.; Hédacq, R.; Barreau, C.; Dubucq, D. Experimental Study of Hyperspectral Responses of Plants Grown on Mud Pit Soil. In *Earth Resources and Environmental Remote Sensing/GIS Applications VII*; 2016; Vol. 10005, p 100051E. <https://doi.org/10.1117/12.2239606>.
- (13) Gogoi, B. K.; Dutta, N. N.; Goswami, P.; Krishna Mohan, T. R. A Case Study of Bioremediation of Petroleum-Hydrocarbon Contaminated Soil at a Crude Oil Spill Site. *Adv. Environ. Res.* **2003**, *7* (4), 767–782. [https://doi.org/10.1016/S1093-0191\(02\)00029-1](https://doi.org/10.1016/S1093-0191(02)00029-1).
- (14) Cunha, S. B. da. A Review of Quantitative Risk Assessment of Onshore Pipelines. *J. Loss Prev. Process Ind.* **2016**, *44*, 282–298. <https://doi.org/10.1016/j.jlp.2016.09.016>.
- (15) Shadizadeh, S. R.; Zoveidavianpoor, M. A Drilling Reserve Mud Pit Assessment in Iran: Environmental Impacts and Awareness. *Pet. Sci. Technol.* **2010**, *28* (14), 1513–1526.

<https://doi.org/10.1080/10916460903117545>.

- (16) Bioucas-Dias, J. M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5* (2), 354–379. <https://doi.org/10.1109/JSTARS.2012.2194696>.
- (17) Correa Pabón, R. E.; de Souza Filho, C. R.; Oliveira, W. J. de. Reflectance and Imaging Spectroscopy Applied to Detection of Petroleum Hydrocarbon Pollution in Bare Soils. *Sci. Total Environ.* **2019**, *649*, 1224–1236. <https://doi.org/10.1016/j.scitotenv.2018.08.231>.
- (18) Slonecker, T.; Fisher, G. B.; Aiello, D. P.; Haack, B. Visible and Infrared Remote Imaging of Hazardous Waste: A Review. *Remote Sens.* **2010**, *2* (11), 2474–2508. <https://doi.org/10.3390/rs2112474>.
- (19) Arellano, P.; Tansey, K.; Balzter, H.; Boyd, D. S. Detecting the Effects of Hydrocarbon Pollution in the Amazon Forest Using Hyperspectral Satellite Images. *Environ. Pollut.* **2015**, *205*, 225–239. <https://doi.org/10.1016/j.envpol.2015.05.041>.
- (20) van der Werff, H.; van der Meijde, M.; Jansma, F.; van der Meer, F.; Groothuis, G. J. A Spatial-Spectral Approach for Visualization of Vegetation Stress Resulting from Pipeline Leakage. *Sensors* **2008**, *8*, 3733–3743. <https://doi.org/10.3390/s8063733>.
- (21) Calderón, R.; Navas-Cortés, J. A.; Lucena, C.; Zarco-Tejada, P. J. High-Resolution Airborne Hyperspectral and Thermal Imagery for Early Detection of Verticillium Wilt of Olive Using Fluorescence, Temperature and Narrow-Band Spectral Indices. *Remote Sens. Environ.* **2013**, *139*, 231–245. <https://doi.org/10.1016/j.rse.2013.07.031>.
- (22) Jacquemoud, S.; Verhoef, W.; Baret, F.; Bacour, C.; Zarco-Tejada, P. J.; Asner, G. P.; François, C.; Ustin, S. L. PROSPECT+SAIL Models: A Review of Use for Vegetation Characterization. *Remote Sens. Environ.* **2009**, *113*, S56–S66. <https://doi.org/10.1016/j.rse.2008.01.026>.
- (23) Rumpf, T.; Mahlein, A.-K.; Steiner, U.; Oerke, E.-C.; Dehne, H.-W.; Plümer, L. Early Detection and Classification of Plant Diseases with Support Vector Machines Based on Hyperspectral Reflectance. *Comput. Electron. Agric.* **2010**, *74* (1), 91–99. <https://doi.org/10.1016/j.compag.2010.06.009>.
- (24) Smith, K. L.; Steven, M. D.; Colls, J. J. Plant Spectral Responses to Gas Leaks and Other Stresses. *Int. J. Remote Sens.* **2005**, *26* (18), 4067–4081. <https://doi.org/10.1080/01431160500165625>.
- (25) Zinnert, J. C.; Via, S. M.; Young, D. R. Distinguishing Natural from Anthropogenic Stress in Plants: Physiology, Fluorescence and Hyperspectral Reflectance. *Plant Soil* **2013**, *366* (1–2), 133–141. <https://doi.org/10.1007/s11104-012-1414-1>.
- (26) Blackburn, G. A. Spectral Indices for Estimating Photosynthetic Pigment Concentrations: A Test Using Senescent Tree Leaves. *Int. J. Remote Sens.* **1998**, *19* (4), 657–675. <https://doi.org/10.1080/014311698215919>.
- (27) Feret, J. B.; François, C.; Asner, G. P.; Gitelson, A. A.; Martin, R. E.; Bidel, L. P. R.; Ustin, S. L.; le Maire, G.; Jacquemoud, S. PROSPECT-4 and 5: Advances in the Leaf Optical Properties Model Separating Photosynthetic Pigments. *Remote Sens. Environ.* **2008**, *112*, 3030–3043. <https://doi.org/10.1016/j.rse.2008.02.012>.
- (28) Sims, D. A.; Gamon, J. A. Relationships between Leaf Pigment Content and Spectral Reflectance across a Wide Range of Species, Leaf Structures and Developmental Stages.

- Remote Sens. Environ.* **2002**, *81*, 337–354. [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X).
- (29) Curran, P. J. Remote Sensing of Foliar Chemistry. *Remote Sens. Environ.* **1989**, *278*, 271–278. [https://doi.org/https://doi.org/10.1016/0034-4257\(89\)90069-2](https://doi.org/https://doi.org/10.1016/0034-4257(89)90069-2).
- (30) Lichtenthaler, H. K.; Buschmann, C. Chlorophylls and Carotenoids: Measurement and Characterization by UV-VIS. In *Current Protocols in Food Analytical Chemistry*; John Wiley & Sons, Inc., 2001. <https://doi.org/doi.org/10.1002/0471142913.faf0403s01>.
- (31) Huang, J.; Wei, C.; Zhang, Y.; Blackburn, G. A.; Wang, X.; Wei, C.; Wang, J. Meta-Analysis of the Detection of Plant Pigment Concentrations Using Hyperspectral Remotely Sensed Data. *PLoS One* **2015**, *10* (9), 1–26. <https://doi.org/10.1371/journal.pone.0137029>.
- (32) Jacquemoud, S.; Baret, F. PROSPECT: A Model of Leaf Optical Properties Spectra. *Remote Sens. Environ.* **1990**, *34*, 75–91. [https://doi.org/10.1016/0034-4257\(90\)90100-Z](https://doi.org/10.1016/0034-4257(90)90100-Z).
- (33) Zhang, Y.; Huang, J.; Wang, F.; Blackburn, G. A.; Zhang, H. K.; Wang, X.; Wei, C.; Zhang, K.; Wei, C. An Extended PROSPECT: Advance in the Leaf Optical Properties Model Separating Total Chlorophylls into Chlorophyll a and B. *Sci. Rep.* **2017**, *7* (6429). <https://doi.org/10.1038/s41598-017-06694-y>.
- (34) Junker, L. V.; Ensminger, I. Relationship between Leaf Optical Properties, Chlorophyll Fluorescence and Pigment Changes in Senescing *Acer Saccharum* Leaves. *Tree Physiol.* **2016**, *36* (6), 694–711. <https://doi.org/10.1093/treephys/tpv148>.
- (35) Archetti, M.; Döring, T. F.; Hagen, S. B.; Hughes, N. M.; Leather, S. R.; Lee, D. W.; Lev-Yadun, S.; Manetas, Y.; Ougham, H. J.; Schaberg, P. G.; et al. Unravelling the Evolution of Autumn Colours: An Interdisciplinary Approach. *Trends Ecol. Evol.* **2009**, *24* (3), 166–173. <https://doi.org/10.1016/j.tree.2008.10.006>.
- (36) Juvany, M.; Müller, M.; Munné-Bosch, S. Photo-Oxidative Stress in Emerging and Senescing Leaves: A Mirror Image. *J. Exp. Bot.* **2013**, *64* (11), 3087–3098. <https://doi.org/10.1093/jxb/ert174>.
- (37) Gamon, J. A. J. A.; Peñuelas, J.; Field, C. B. B.; Penuelas, J.; Field, C. B. B. A Narrow-Waveband Spectral Index That Tracks Diurnal Changes in Photosynthetic Efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [https://doi.org/10.1016/0034-4257\(92\)90059-S](https://doi.org/10.1016/0034-4257(92)90059-S).
- (38) Garbulsky, M. F.; Peñuelas, J.; Gamon, J.; Inoue, Y.; Filella, I. The Photochemical Reflectance Index (PRI) and the Remote Sensing of Leaf, Canopy and Ecosystem Radiation Use Efficiencies. A Review and Meta-Analysis. *Remote Sens. Environ.* **2011**, *115* (2), 281–297. <https://doi.org/10.1016/j.rse.2010.08.023>.
- (39) Mónica Giusti, M.; Wrolstad, R. E. Characterization and Measurement of Anthocyanins by UV-Visible Spectroscopy. *Handb. Food Anal. Chem.* **2005**, *2-2*, 19–31. <https://doi.org/10.1002/0471709085.ch18>.
- (40) Dogwood, R.; Feild, T. S.; Lee, D. W.; Holbrook, N. M. Why Leaves Turn Red in Autumn . The Role of Anthocyanins in Senescing Leaves Of. **2001**, *127* (October), 566–574. <https://doi.org/10.1104/pp.010063.566>.
- (41) Baldini, E.; Facini, O.; Nerozzi, F.; Rossi, F.; Rotondi, A. Leaf Characteristics and Optical Properties of Different Woody Species. *Trees* **1997**, *12*, 73. <https://doi.org/10.1007/s004680050124>.
- (42) Foley, S.; Rivard, B.; Sanchez-Azofeifa, G. A.; Calvo, J. Foliar Spectral Properties Following Leaf Clipping and Implications for Handling Techniques. *Remote Sens. Environ.* **2006**, *103*

- (3), 265–275. <https://doi.org/10.1016/j.rse.2005.06.014>.
- (43) Slaton, M. R.; Hunt, E. R.; Smith, W. K. Estimating Near-Infrared Leaf Reflectance from Leaf Structural Characteristics. *Am. J. Bot.* **2001**, *88* (2), 278–284. <https://doi.org/10.2307/2657019>.
- (44) Rossatto, D. R.; Kolb, R. M.; Franco, A. C. Leaf Anatomy Is Associated with the Type of Growth Form in Neotropical Savanna Plants. *Botanique* **2015**, *93* (8), 507–508. <https://doi.org/https://doi.org/10.1139/cjb-2015-0001>.
- (45) Boren, E. J.; Boschetti, L.; Johnson, D. M. Characterizing the Variability of the Structure Parameter in the PROSPECT Leaf Optical Properties Model. *Remote Sens.* **2019**, *11* (10), 1236. <https://doi.org/10.3390/rs11101236>.
- (46) Hanba, Y. T.; Kogami, H.; Terashima, I. The Effect of Growth Irradiance on Leaf Anatomy and Photosynthesis in Acer Species Differing in Light Demand. *Plant, Cell Environ.* **2002**, *25* (8), 1021–1030. <https://doi.org/10.1046/j.1365-3040.2002.00881.x>.
- (47) Li, L.; Ma, Z.; Niinemets, Ü.; Guo, D. Three Key Sub-Leaf Modules and the Diversity of Leaf Designs. *Front. Plant Sci.* **2017**, *8* (September), 1–8. <https://doi.org/10.3389/fpls.2017.01542>.
- (48) Knapp, A. K. A. K.; Carter, G. A. G. A. Variability in Leaf Optical Properties among 26 Species From A Broad Range Of Habitats. *Am. J. Bot.* **1998**, *85* (7), 940–946. <https://doi.org/10.2307/2446360>.
- (49) Asner, G. P. Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sens. Environ.* **1997**, *64* (3), 234–253.
- (50) Asner, G. P.; Martin, R. E. Spectral and Chemical Analysis of Tropical Forests: Scaling from Leaf to Canopy Levels. *Remote Sens. Environ.* **2008**, *112* (10), 3958–3970. <https://doi.org/10.1016/j.rse.2008.07.003>.
- (51) Ollinger, S. V. Sources of Variability in Canopy Reflectance and the Convergent Properties of Plants. *New Phytol.* **2011**, *189*, 375–394. <https://doi.org/10.1111/j.1469-8137.2010.03536.x>.
- (52) de Wit, C. T. *Photosynthesis of Leaf Canopies*; 1965.
- (53) Clevers, J. G. P. W.; Kooistra, L.; Schaepman, M. E. Using Spectral Information from the NIR Water Absorption Features for the Retrieval of Canopy Water Content. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10* (3), 388–397. <https://doi.org/10.1016/j.jag.2008.03.003>.
- (54) Peñuelas, J.; Filella, I.; Biel, C.; Serrano, L.; Savé, R. The Reflectance at the 950–970 nm Region as an Indicator of Plant Water Status. *Int. J. Remote Sens.* **1993**, *14* (10), 1887–1905. <https://doi.org/10.1080/01431169308954010>.
- (55) Sims, D. A.; Gamon, J. A. Estimation of Vegetation Water Content and Photosynthetic Tissue Area from Spectral Reflectance: A Comparison of Indices Based on Liquid Water and Chlorophyll Absorption Features. *Remote Sens. Environ.* **2003**, *84* (4), 526–537. [https://doi.org/10.1016/S0034-4257\(02\)00151-7](https://doi.org/10.1016/S0034-4257(02)00151-7).
- (56) Tian, Q.; Tong, Q.; Pu, R.; Guo, X.; Zhao, C. Spectroscopic Determination of Wheat Water Status Using 1650–1850 nm Spectral Absorption Features. *Int. J. Remote Sens.* **2001**, *22* (12), 2329–2338. <https://doi.org/10.1080/01431160118199>.
- (57) Tucker, C. J. Remote Sensing of Leaf Water Content in the near Infrared. *Remote Sens. Environ.* **1980**, *10*, 23–32. [https://doi.org/10.1016/0034-4257\(80\)90096-6](https://doi.org/10.1016/0034-4257(80)90096-6).

- (58) Yilmaz, M. T.; Hunt, E. R.; Jackson, T. J. Remote Sensing of Vegetation Water Content from Equivalent Water Thickness Using Satellite Imagery. *Remote Sens. Environ.* **2008**, *112* (5), 2514–2522. <https://doi.org/10.1016/j.rse.2007.11.014>.
- (59) Ceccato, P.; Flasse, S.; Tarantola, S.; Jacquemoud, S.; Grégoire, J. M. Detecting Vegetation Leaf Water Content Using Reflectance in the Optical Domain. *Remote Sens. Environ.* **2001**, *77* (1), 22–33. [https://doi.org/10.1016/S0034-4257\(01\)00191-2](https://doi.org/10.1016/S0034-4257(01)00191-2).
- (60) Asif, M. Sustainability of Timber, Wood and Bamboo in Construction. In *Sustainability of Construction Materials*; Elsevier, 2009; pp 31–54. <https://doi.org/10.1533/9781845695842.31>.
- (61) Liu, Q.; Luo, L.; Zheng, L. Lignins: Biosynthesis and Biological Functions in Plants. *Int. J. Mol. Sci.* **2018**, *19* (2). <https://doi.org/10.3390/ijms19020335>.
- (62) Wang, Z.; Skidmore, A. K.; Wang, T.; Darvishzadeh, R.; Hearne, J. Applicability of the PROSPECT Model for Estimating Protein and Cellulose+lignin in Fresh Leaves. *Remote Sens. Environ.* **2015**, *168*, 205–218. <https://doi.org/10.1016/j.rse.2015.07.007>.
- (63) U.S. Energy Information Administration (EIA). *Definitions of Petroleum Products and Other Terms*, 2010.
- (64) Ahmadun, F.-R.; Pendashteh, A.; Abdullah, L. C.; Awang Biak, D. R.; Madaeni, S. S.; Abidin, Z. Z. Review of Technologies for Oil and Gas Produced Water Treatment. *J. Hazard. Mater.* **2009**, *170*, 530–551. <https://doi.org/10.1016/j.jhazmat.2009.05.044>.
- (65) Hu, G.; Li, J.; Zeng, G. Recent Development in the Treatment of Oily Sludge from Petroleum Industry: A Review. *J. Hazard. Mater.* **2013**, *261*, 470–490. <https://doi.org/10.1016/j.jhazmat.2013.07.069>.
- (66) Adeniyi, A. .; Afolabi, J. . Determination of Total Petroleum Hydrocarbons and Heavy Metals in Soils within the Vicinity of Facilities Handling Refined Petroleum Products in Lagos Metropolis. *Environ. Int.* **2002**, *28*, 79–82. [https://doi.org/10.1016/S0160-4120\(02\)00007-7](https://doi.org/10.1016/S0160-4120(02)00007-7).
- (67) Doble, M.; Kumar, A. Petroleum Hydrocarbon Pollution. *Biotreat. Ind. Effluents* **2007**, 241–253. <https://doi.org/10.1016/b978-075067838-4/50025-7>.
- (68) Kistic, I.; Mesic, S.; Basic, F.; Brkic, V.; Mesic, M.; Durn, G.; Zgorelec, Z.; Bertovic, L. The Effect of Drilling Fluids and Crude Oil on Some Chemical Characteristics of Soil and Crops. *Geoderma* **2009**, *149* (3–4), 209–216. <https://doi.org/10.1016/j.geoderma.2008.11.041>.
- (69) Chaudhuri, U. R. *Fundamentals of Petroleum and Petrochemical Engineering*; Taylor & Francis Group, Ed.; CRC Press, 2016. <https://doi.org/10.1201/b10486>.
- (70) Athar, H. ur R.; Ambreen, S.; Javed, M.; Hina, M.; Rasul, S.; Zafar, Z. U.; Manzoor, H.; Ogbaga, C. C.; Afzal, M.; Al-Qurainy, F.; et al. Influence of Sub-Lethal Crude Oil Concentration on Growth, Water Relations and Photosynthetic Capacity of Maize (*Zea Mays* L.) Plants. *Environ. Sci. Pollut. Res.* **2016**, *23* (18), 18320–18331. <https://doi.org/10.1007/s11356-016-6976-7>.
- (71) Ogboghodo, I. A.; Iruaga, E. K.; Osemwota, I. O.; Chokor, J. U. An Assessment of the Effects of Crude Oil Pollution on Soil Properties, Germination and Growth of Maize (*Zea Mays*) Using Two Crude Types – Forcados Light and Escravos Light. *Environ. Monit. Assess.* **2004**, *96*, 143–152. <https://doi.org/10.1023/B:EMAS.0000031723.62736.24>.
- (72) Wang, Y.; Feng, J.; Lin, Q.; Lyu, X.; Wang, X.; Wang, G. Effects of Crude Oil Contamination on Soil Physical and Chemical Properties in Momoge Wetland of China.

- Chinese Geogr. Sci.* **2013**, *23* (6), 708–715. <https://doi.org/10.1007/s11769-013-0641-6>.
- (73) Nie, M.; Lu, M.; Yang, Q.; Zhang, X.; Xiao, M.; Jiang, L.; Yang, J.; Fang, C.; Chen, J.; Li, B. Plants' Use of Different Nitrogen Forms in Response to Crude Oil Contamination. **2011**, *159*, 157–163. <https://doi.org/10.1016/j.envpol.2010.09.013>.
- (74) Zuofa, K.; Loganathan, P.; Isirimah, N. O. Effects of Crude Oil Applications to Soil on the Growth and Yield of Maize, Okro and Cassava in Nigeria. *Oil Chem. Pollut.* **1988**, *4* (4), 249–259. [https://doi.org/10.1016/S0269-8579\(88\)80001-7](https://doi.org/10.1016/S0269-8579(88)80001-7).
- (75) Nogueira, L.; Inckot, C.; Santos, G. D. O. Phytotoxicity of Petroleum-Contaminated Soil and Bioremediated Soil on *Allophylus Edulis*. *Rodriguésia* **2011**, *62* (3), 459–466.
- (76) Balasubramaniyam, A.; Harvey, P. J. Scanning Electron Microscopic Investigations of Root Structural Modifications Arising from Growth in Crude Oil-Contaminated Sand. *Environ. Sci. Pollut. Res.* **2014**, *21* (22), 12651–12661. <https://doi.org/10.1007/s11356-014-3138-7>.
- (77) Balliana, A. G.; Moura, B. B.; Inckot, R. C.; Bona, C. Development of *Canavalia Ensiformis* in Soil Contaminated with Diesel Oil. *Environ. Sci. Pollut. Res.* **2017**, *24*, 979–986. <https://doi.org/10.1007/s11356-016-7674-1>.
- (78) Barceló, J.; Poschenrieder, C. Plant Water Relations as Affected by Heavy Metal Stress: A Review. *J. Plant Nutr.* **1990**, *13* (1), 1–37. <https://doi.org/10.1080/01904169009364057>.
- (79) Dupuy, J.; Ouvrard, S.; Leglize, P.; Sterckeman, T. Morphological and Physiological Responses of Maize (*Zea Mays*) Exposed to Sand Contaminated by Phenanthrene. *Chemosphere* **2015**, *124* (1), 110–115. <https://doi.org/10.1016/j.chemosphere.2014.11.051>.
- (80) Baruah, P.; Saikia, R. R.; Baruah, P. P.; Deka, S. Effect of Crude Oil Contamination on the Chlorophyll Content and Morpho-Anatomy of *Cyperus Brevifolius* (Rottb.) Hassk. *Environ. Sci. Pollut. Res.* **2014**, *21* (21), 12530–12538. <https://doi.org/10.1007/s11356-014-3195-y>.
- (81) Sanches, I. D.; de Souza Filho, C. R.; Magalhães, L. A.; Quitério, G. C. M.; Alves, M. N.; Oliveira, W. J. Assessing the Impact of Hydrocarbon Leakages on Vegetation Using Reflectance Spectroscopy. *ISPRS J. Photogramm. Remote Sens.* **2013**, *78*, 85–101. <https://doi.org/10.1016/j.isprsjprs.2013.01.007>.
- (82) Sanches, I. D.; Souza Filho, C. R.; Magalhães, L. A.; Quitério, G. C. M.; Alves, M. N.; Oliveira, W. J. Unravelling Remote Sensing Signatures of Plants Contaminated with Gasoline and Diesel: An Approach Using the Red Edge Spectral Feature. *Environ. Pollut.* **2013**, *174*, 16–27. <https://doi.org/10.1016/j.envpol.2012.10.029>.
- (83) Anoliefo, G. O.; Vwioko, D. E. Effects of Spent Lubricating Oil on the Growth of *Capsicum Annum* L. and *Lycopersicon Esculentum* Miller. *Environ. Pollut.* **1995**, *88* (3), 361–364. [https://doi.org/10.1016/0269-7491\(95\)93451-5](https://doi.org/10.1016/0269-7491(95)93451-5).
- (84) Pérez-Hernández, I.; Ochoa-Gaona, S.; Adams, R. H.; Rivera-Cruz, M. C.; Pérez-Hernández, V.; Jarquín-Sánchez, A.; Geissen, V.; Martínez-Zurimendi, P. Growth of Four Tropical Tree Species in Petroleum-Contaminated Soil and Effects of Crude Oil Contamination. *Environ. Sci. Pollut. Res.* **2017**, *24* (2), 1769–1783. <https://doi.org/10.1007/s11356-016-7877-5>.
- (85) Emengini, E. J.; Blackburn, G. A.; Theobald, J. C. Early Detection of Oil-Induced Stress in Crops Using Spectral and Thermal Responses. *J. Appl. Remote Sens.* **2013**, *7*. <https://doi.org/10.1117/1.jrs.7.073596>.
- (86) Emengini, E. J.; Blackburn, G. A.; Theobald, J. C. Discrimination of Plant Stress Caused by Oil Pollution and Waterlogging Using Hyperspectral and Thermal Remote Sensing. *J. Appl.*

Remote Sens. **2013**, *7*. <https://doi.org/10.1117/1.jrs.7.073476>.

- (87) Rosso, P. H.; Pushnik, J. C.; Lay, M.; Ustin, S. L. Reflectance Properties and Physiological Responses of *Salicornia Virginica* to Heavy Metal and Petroleum Contamination. *Environ. Pollut.* **2005**, *137*, 241–252. <https://doi.org/10.1016/j.envpol.2005.02.025>.
- (88) Gürtler, S.; de Souza Filho, C. R.; Sanches, I. D.; Alves, M. N.; Oliveira, W. J. Determination of Changes in Leaf and Canopy Spectra of Plants Grown in Soils Contaminated with Petroleum Hydrocarbons. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146* (April), 272–288. <https://doi.org/10.1016/j.isprsjprs.2018.09.011>.
- (89) Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C.; Corsi, F.; Cho, M. LAI and Chlorophyll Estimation for a Heterogeneous Grassland Using Hyperspectral Measurements. *ISPRS J. Photogramm. Remote Sens.* **2008**, *63* (4), 409–426. <https://doi.org/10.1016/j.isprsjprs.2008.01.001>.
- (90) Axelsson, C.; Skidmore, A. K.; Schlerf, M.; Fauzi, A.; Verhoef, W. Hyperspectral Analysis of Mangrove Foliar Chemistry Using PLSR and Support Vector Regression. *Int. J. Remote Sens.* **2013**, *34* (5), 1724–1743. <https://doi.org/10.1080/01431161.2012.725958>.
- (91) Hermosilla, T.; Wulder, M. A.; White, J. C.; Coops, N. C.; Hobart, G. W. Regional Detection, Characterization, and Attribution of Annual Forest Change from 1984 to 2012 Using Landsat-Derived Time-Series Metrics. *Remote Sens. Environ.* **2015**, *170*, 121–132. <https://doi.org/10.1016/j.rse.2015.09.004>.
- (92) Thenkabail, P. S.; Enclona, E. A.; Ashton, M. S.; Van Der Meer, B. Accuracy Assessments of Hyperspectral Waveband Performance for Vegetation Analysis Applications. *Remote Sens. Environ.* **2004**, *91* (3–4), 354–376. <https://doi.org/10.1016/j.rse.2004.03.013>.
- (93) Wei, C.; Huang, J.; Wang, X.; Blackburn, G. A.; Zhang, Y.; Wang, S.; Mansaray, L. R. Hyperspectral Characterization of Freezing Injury and Its Biochemical Impacts in Oilseed Rape Leaves. *Remote Sens. Environ.* **2017**, *195*, 56–66. <https://doi.org/10.1016/j.rse.2017.03.042>.
- (94) Adamu, B.; Tansey, K.; Ogutu, B. Remote Sensing for Detection and Monitoring of Vegetation Affected by Oil Spills. *Int. J. Remote Sens.* **2018**, *39* (11), 3628–3645. <https://doi.org/10.1080/01431161.2018.1448483>.
- (95) Ozgis, M. S.; Kaduk, J. D.; Jarvis, C. H. Mapping Terrestrial Oil Spill Impact Using Machine Learning Random Forest and Landsat 8 OLI Imagery: A Case Site within the Niger Delta Region of Nigeria. *Environ. Sci. Pollut. Res.* **2019**, *26*, 3621–3635. <https://doi.org/10.1007/s11356-018-3824-y>.
- (96) van der Meijde, M.; van der Werff, H. M. A.; Jansma, P. F.; van der Meer, F. D.; Groothuis, G. J. A Spectral-Geophysical Approach for Detecting Pipeline Leakage. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11* (1), 77–82. <https://doi.org/10.1016/j.jag.2008.08.002>.
- (97) Brigot, G.; Colin-Koeniguer, E.; Plyer, A.; Janez, F. Adaptation and Evaluation of an Optical Flow Method Applied to Coregistration of Forest Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9* (7), 2923–2939. <https://doi.org/10.1109/JSTARS.2016.2578362>.
- (98) Smith, G. M.; Milton, E. J. The Use of the Empirical Line Method to Calibrate Remotely Sensed Data to Reflectance. *Int. J. Remote Sens.* **1999**, *20* (13), 2653–2662. <https://doi.org/10.1080/014311699211994>.
- (99) Stagakis, S.; Vanikiotis, T.; Sykioti, O. Estimating Forest Species Abundance through Linear

- Unmixing of CHRIS/PROBA Imagery. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 79–89. <https://doi.org/10.1016/j.isprsjprs.2016.05.013>.
- (100) Colomina, I.; Molina, P. Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. <https://doi.org/10.1016/j.isprsjprs.2014.02.013>.
- (101) Lu, B.; He, Y. Species Classification Using Unmanned Aerial Vehicle (UAV)-Acquired High Spatial Resolution Imagery in a Heterogeneous Grassland. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 73–85. <https://doi.org/10.1016/j.isprsjprs.2017.03.011>.

Abstract

Oil exploration and contamination monitoring remain limited in regions covered by vegetation. Natural seepages and oil leakages due to facility failures are often masked by the foliage, making ineffective the current technologies used for detecting crude oil and petroleum products. However, the exposure of vegetation to oil affects its health and, consequently, its optical properties in the [400:2500] nm domain. This suggests being able to detect seepages and leakages indirectly, by analyzing vegetation health through its spectral reflectance. Based on this assumption, this thesis evaluates the potential of airborne hyperspectral imagery with high spatial resolution for detecting and quantifying oil contamination in vegetated regions. To achieve this, a three-step multiscale approach was adopted. The first step aimed at developing a method for detecting and characterizing the contamination under controlled conditions, by exploiting the optical properties of *Rubus fruticosus* L. The proposed method combines 14 vegetation indices in classification and allows detecting various oil contaminants accurately, from leaf to canopy scale. Its use under natural conditions was validated on a contaminated mud pit colonized by the same species. During the second step, a method for quantifying total petroleum hydrocarbons, based on inverting the PROSPECT model, was developed. The method exploits the pigment content of leaves, estimated from their spectral signature, for predicting the level of hydrocarbon contamination in soils accurately. The last step of the approach demonstrated the robustness of the two methods using airborne imagery. They proved performing for detecting and quantifying mud pit contamination. Another method of quantification, based on multiple regression, was proposed. At the end of this thesis, the three methods proposed were validated for use both on the field, at leaf and canopy scales, and on airborne hyperspectral images with high spatial resolution. Their performance depends however on the species, the season and the level of soil contamination. A similar approach was conducted under tropical conditions, allowing the development of a method for quantifying the contamination adapted to this context. In a perspective of operational use, an important effort is still required for extending the scope of the methods to other contexts and for anticipating their use on satellite- and drone-embedded hyperspectral sensors. Finally, the contribution of active remote sensing (radar and LiDAR) should be considered in further research, in order to overcome some of the limits specific to passive optical remote sensing.

Keywords: Hyperspectral remote sensing, Vegetation optical properties, Soil contamination, Total petroleum hydrocarbons, Vegetation index, Radiative transfer model

Résumé

L'exploration pétrolière et le monitoring de la contamination demeurent très limités dans les régions colonisées par la végétation. La présence de suintements naturels et de fuites d'installations pétrolières est bien souvent masquée par le feuillage, rendant inopérantes les technologies actuelles de détection du pétrole brut et des produits pétroliers. L'exposition de la végétation à ces composés affecte toutefois son état de santé et, par conséquent, ses propriétés optiques dans le domaine [400:2500] nm. Cela suggère de pouvoir détecter les suintements et les fuites d'installations de manière indirecte, en analysant l'état de santé de la végétation au travers de sa réflectance spectrale. Basée sur cette hypothèse, la présente thèse évalue le potentiel de l'imagerie hyperspectrale aéroportée à très haute résolution spatiale pour détecter et quantifier la contamination pétrolière en région tempérée végétalisée. Pour cela, une approche multi-échelles en trois étapes a été adoptée. La première étape a eu pour objet de développer une méthode de détection et de caractérisation de la contamination en conditions contrôlées, exploitant les propriétés optiques de *Rubus fruticosus* L. La méthode proposée combine 14 indices de végétation en classification et permet de détecter divers contaminants pétroliers avec précision, depuis l'échelle de la feuille jusqu'à celle du couvert. Son utilisation en conditions naturelles a été validée sur un bourbier de production contaminé, colonisé par la même espèce. Au cours de la seconde étape, une méthode de quantification des hydrocarbures pétroliers totaux, basée sur l'inversion d'un modèle de transfert radiatif, a été développée. Cette méthode exploite le contenu en pigments des feuilles, estimé à partir de leur signature spectrale, afin de prédire précisément le taux de contamination en hydrocarbures du sol. La dernière étape de l'approche a démontré la robustesse des deux méthodes en imagerie aéroportée. Celles-ci se sont montrées très performantes pour détecter et quantifier la contamination des bourbiers. Une autre méthode de quantification, basée sur la régression multiple, a également été proposée. Au terme de cette thèse, les trois méthodes proposées ont été validées pour une utilisation sur le terrain, à l'échelle de la feuille et du couvert, ainsi qu'en imagerie hyperspectrale aéroportée à très haute résolution spatiale. Leurs performances dépendent toutefois de l'espèce, de la saison et du niveau de contamination du sol. Une approche similaire a été conduite en conditions tropicales, permettant de mettre au point une méthode de quantification de la contamination adaptée à ce contexte. En vue d'une utilisation opérationnelle, un effort important reste nécessaire pour étendre le champ d'application des méthodes à d'autres contextes et envisager leur application sur les futurs capteurs hyperspectraux embarqués sur satellite et sur drone. Enfin, l'apport de la télédétection active (radar et LiDAR) est à considérer dans les recherches futures, afin de lever certaines limites propres à l'utilisation de la télédétection optique passive.

Mots-clés : Télédétection hyperspectrale, Propriétés optiques de la végétation, Contamination du sol, Hydrocarbures pétroliers totaux, Indice de végétation, Modèle de transfert radiatif