Hyperspectral remote sensing to detect petroleum pollution in the Amazon forest

Paul Arellano[a,b,c]
Kevin Tansey[a,d]
Heiko Balzter[a,d,f]
Doreen S. Boyd[e]

[a] Centre of Landscape and Climate Research, Department of Geography, University of Leicester, University Road, Leicester, LE1 7RH, UK
[b],School of Geological Sciences & Engineering, YachayTech University, San Miguel de Urcuqui, Hacienda San Jose s/n, Imbabura, Ecuador
[c] Centre of Earth Observation (CEO), YachayTech University, San Miguel de Urcuqui, Hacienda San Jose s/n, Imbabura, Ecuador
[d] Department of Geography, University of Leicester, University Road, LE1 7RH, UK
[e] School of Geography, University of Nottingham, Nottingham, NG7 2RD, UK
[f] National Centre for Earth Observation, University of Leicester, University Road, Leicester, LE1 7RH, UK

Abstract. The global demand for fossil energy is triggering oil exploration and production projects in remote areas of the world. During the last few decades hydrocarbon production has caused pollution in the Amazon forest inflicting considerable environmental impact. Until now it is not clear how hydrocarbon pollution affects the health of the tropical forest flora. During a field campaign in polluted and pristine forest, more than 1100 leaf samples were collected and analysed for biophysical and biochemical parameters. The results revealed that tropical forests exposed to hydrocarbon pollution show reduced levels of chlorophyll content, higher levels of foliar water content and leaf structural changes. In order to map this impact over wider geographical areas, vegetation indices were applied to hyperspectral Hyperion satellite imagery. Three vegetation indices (SR, NDVI and $NDVI_{705}$) were found to be the most appropriate indices to detect the effects of petroleum pollution in the Amazon forest.

Keywords: Remote Sensing, Image Processing, Geology, Petroleum pollution, Amazon forest

1. Introduction

Global demand for energy is trigging oil and gas exploration and production across the Amazon basin, with even very remote areas leased out or under negotiation for access (Finer et al. 2008). In western Amazonia, there has been an unprecedented rise in this activity, causing environmental pollution in vast regions of forest via oil spills from pipelines networks and leakages from unlined open pits (Hurtig&San-Sebastián 2005, Bernal 2011)

Despite high international public interest in protecting Amazon rainforests, little scientific attention has focussed on the effects of oil pollution on the forest; much focus is on threats from deforestation, selective logging, hunting, fire and global and regional climate variations (Asner et al. 2004, Malhi et al. 2008, Davidson et al. 2012). The high diversity and intrinsic complex biological interactions of tropical forests and their vast expanse challenge our understanding of the impact of oil on them. Data collected *in situ* in these forests are rare, most likely due to access issues. An alternative approach to measuring and monitoring oil contamination in tropical forests at suitable spatial and temporal scales is desirable. It is suggested here that satellite imaging spectrometry, which affords the collection of hyperspectral data of the environment, could be a way forward. In order to detect vegetated landscape contamination using imaging spectrometry, environmental change as a result of contamination need to have a measurable impact upon the biochemical, and related biophysical properties (e.g., pigment concentration, leaf structural and leaf area), of the vegetation growing in that environment. Such properties measured using hyperspectral remotely sensed data may then be used as a proxy to contamination (Mutanga;Skidmore & Prins 2004).

This paper demonstrates the suitability of satellite imaging spectrometry for the detection of contamination by oil of the forest in the Ecuadorian Amazon. EO-1 (Earth-Observation 1) Hyperion imagery is analysed with supporting field data on soils and foliar properties with an overriding objective of producing a map of the spatial pattern of forest contamination by oil.

2. Methodology

2.1. Study area and sites

Three study sites within Ecuadorian Amazon rainforest were investigated. Two were located in the lowland evergreen secondary forest of Sucumbios province, in the Tarapoa region (0°11' S, 76°20' W). Due to their close proximity, both sites share soil types, weather and anthropogenic influences. Site 1 (polluted) is located by an abandoned petroleum platform where open pits have been discharging crude oil to the environment, or leaching out as the pits degrade or overflow, for the past 15 years. Site 2 (non-polluted) is some distance from Site 1 and so not directly influenced by the oil pollution evident at Site 1. Site 3 (Pristine forest-Yasuni) is situated in the highly diverse lowland evergreen primary forest of the Orellana province, in the northern section of Ecuador's Yasuni National Park (0°41' S, 76°24' W). The forest has a species richness among the highest globally (Tedersoo et al. 2010) and are situated well away from any sources of crude oil (and other anthropogenic influences).

2.2. Sampling and measurements

Fieldwork was undertaken from April to July 2012. Soils samples were collected across the study sites and several parameters related to physical properties, nutrients, metals and hydrocarbons traces were analysed in accredited laboratories following international standard methods. At leaf level, fully expanded mature leaves, with no herbivorous/pathogenic damage, were selected from each of the collected branches and analysed. Each leaf was clipped at the midpoint using cork borers to obtain a disk of known surface *(S)*; this is the optimal position from which to take chlorophyll readings (Hoel 1998). Three SPAD-502 chlorophyll meter readings were taken from each disk, at different positions, to compute a mean index value. The fresh weight ***(Fw)*** and dry weight ***(Dw)*** of each leaf disk were then calculated to measure (i) leaf water content (***Cw***) in g cm^{-2} = *(Fw-Dw)/S*. Other leaf properties computed were (ii) dry matter content (***Cm***) in g cm^{-2} = Dw/S; (iii) Specific leaf area (***SLA***) in cm^2 g^{-1} = 1/Cm; (iv) Leaf water content (***LWC***) in % = (Fw-Dw)/Fw ; (v) Leaf dry matter content (***LDMC***) in % = Dw/Fw; and (vi) Leaf thickness or leaf succulence (***Lt***) in g cm^{-2} = 1/SLA*LDMC.

2.3. Sampling and measurements

Field spectroscopy measurements of the leaf samples were taken with an ASD FieldSpec HandHeld-2 spectroradiometer (Analytical Spectral Devices Inc., Boulder, Colorado). This instrument measures in the wavelength range from 325 nm to 1075 nm at a sampling interval of 1 nm. The spectroradiometer was attached to a plant probe with an internal 4.05-W halogen light source and a leaf clip that includes a rotating head with both white and black reference panels (Arellano et al. 2016). These spectroradiometer measurements (reflectance and double-transmittance) were used to estimate chlorophyll content using PROSPECT model inversion.

2.4. Hyperion image pre-processing

USGS EO-1 Hyperion image was used in this study. Hyperion data have a spatial resolution of $30m^2$ with each pixel covering the spectral range, 400-2500 nm. Wavelength

ISBN: 978-85-17-00088-1

selection, atmospheric corrections and noise reduction were applied during the pre-processing.

2.5. Spectral vegetation indices (VI)

Several VI grouped in five broad-band, 14 narrow-band-greenness/chlorophyll, four narrow-band-other pigments and five narrow-band-water indices were computed.

INDEX TYPE	INDICES
BROAD-BAND INDICES	Simple Ratio (SR), Normalised Difference Vegetation Index (NDVI); Green Normalised Difference Vegetation Index (GNDVI); Enhanced Vegetation Index (EVI); Atmospherically Resistant Vegetation Index (ARVI)
NARROW-BAND INDICES: GREENES, CHLOROPHYLL, REP	Sum Green (SG); Pigment Specific Simple Ratio-Chl (PSSRa); Red-Edge Normalised Difference Index ($NDVI_{705}$); Modified Red-Edge Simple Ratio (mSR_{705}); Modified Red-Edge Simple Ratio (mSR_{705}); Carter Index 2 (CTR2); Lichtenthaler Index 1(LIC1) or Pigment Specific Normalised Difference – Chla (PSNDa); Optimised Soil-Adjusted Vegetation Index (OSAVI); Modified Chlorophyll Absorption Ratio Index (MCARI); Ratio of derivatives at 725 and 702 nm ($Der_{725-702}$); Red-Edge Position (REP); Vogelmann Red-Edge Index (VOG1); Chlorophyll Index (CI_{590})
NARROW-BAND INDICES: OTHER PIGMENTS	Structure Insensitive Pigment Index (SIPI); Red Green Ratio (RG); Anthocyanin Reflectance Index 1 (ARI1); Anthocyanin Reflectance Index 2 (ARI2)
NARROW BAND INDICES: WATER	Water Band Index (WBI); Normalised Difference Water Index (NDWI); Moisture Stress Index (MSI); Normalised Difference Infrared Index (NDII) and Normalised Heading Index (NHI)

3. Results

3.1. Analysis of field-derived data

The results of the soil analysis confirm high levels of hydrocarbons in the soil in Site 1. There are no evidences of petroleum pollution in Sites 2 and 3.

3.2. Analysis of foliar biophysical and biochemical parameters

Initial focus on the plotted means and ±95% confidence intervals for each foliar biochemical/biophysical variable (Figure 1. Mean and ±95% confidence interval for the foliar biophysical and biochemical parameters), and the ANOVA and associated pairwise comparisons via the Holm method (Table 1), was on how different site 1 (polluted) was from sites 2 and 3. The chlorophyll content (C_{ab}) was significantly lower at site 1 with values strongly different (99.9%) to those for the two non-polluted sites (2 and 3). No significant difference in chlorophyll content was evident between the two unpolluted sites. Leaf water content (*LWC*) and Leaf dry matter content (*LDMC*) also exhibited strongly significant differences (99.9%) between the unpolluted site 1 and sites 2 (strongly significant at 99.9%) and 3 (highly significant at 99%). Total water content (*Cw*) difference however had a slightly different pattern with differences observed between site 1 and 2 only significant at 95% level but highly significant (at 99.9%) between site 1 and site 3.

Organic matter content (*Cm*) was significantly different (95%) between Site 1 and 2 but insignificant in difference between Site 1 and 3, with a high (99%) level of significance difference being shown between the two unpolluted sites. Leaf thickness (*Lt*) was strongly significantly different (99.9%) between Site 1 and 3 but no difference was observed between Sites 1 and 2 for this foliar property. No differences in SLA were observed between any of the sites.

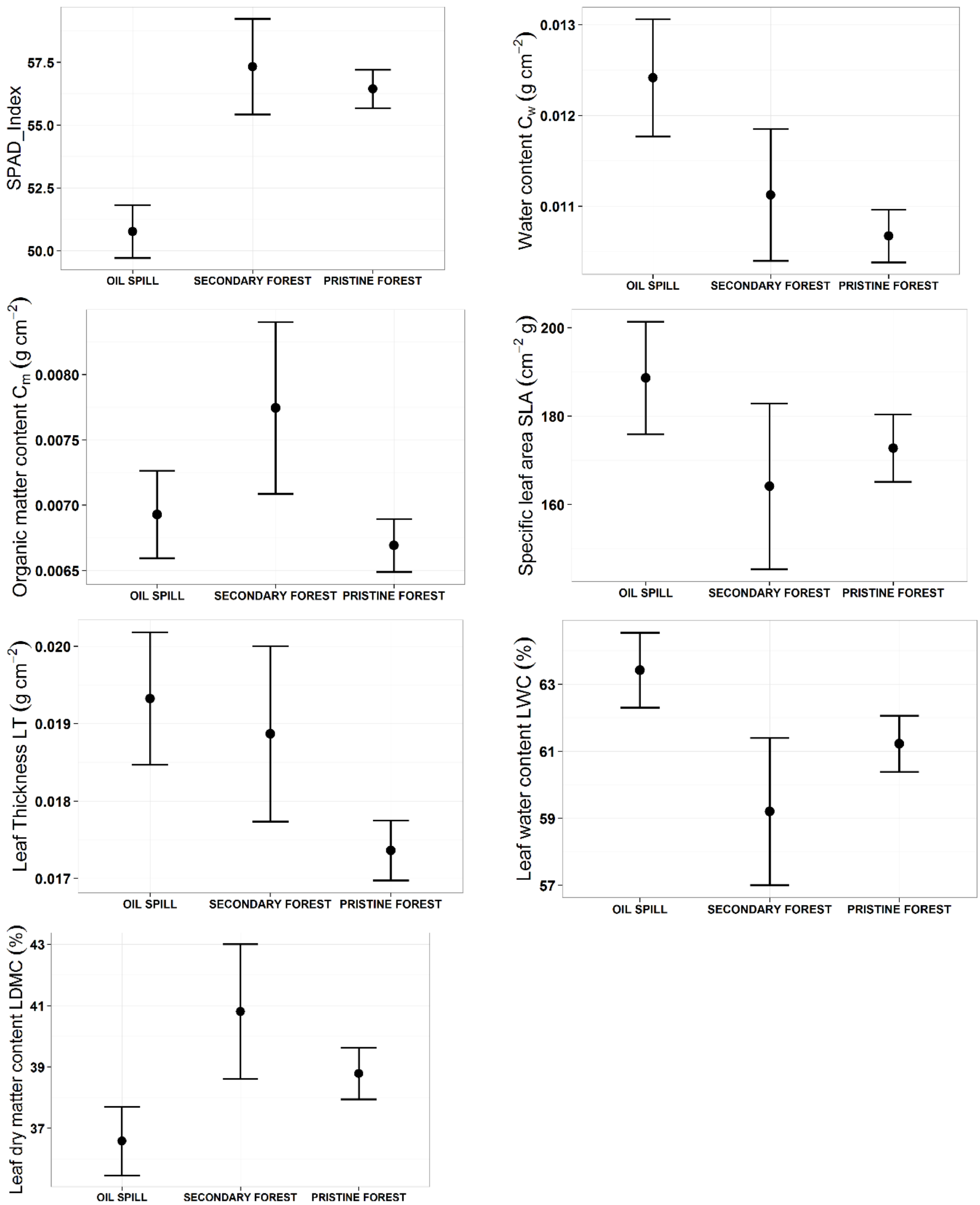

Figure 1. Mean and ±95% confidence interval for the foliar biophysical and biochemical parameters

Table 1. Pairwise comparison of p-values with holm adjustment method

	C_{ab}	*Cw*	*Cm*	*SLA*	*Lt*	*LWC*	*LDMC*
ANOVA test	2.0E-16 ***	4.2E-07 ***	1.7E-03 **	2.8E-02 *	2.2E-05 ***	1.5E-05 ***	1.5E-04 ***
Pairwise comparison – Holm adjustment method							
Oil spill (Site 1)- No Polluted (Site 2)	6.2E-10 ***	2.1E-02 *	1.5E-02 *	7.8E-02	5.0E-01	4.4E-04 ***	4.4E-04 ***

ISBN: 978-85-17-00088-1

Oil spill (Site 1)- Pristine forest (Site 3)	1.2E-14 ***	2.2E-07 ***	2.3E-01	7.8E-02	2.0E-05 ***	4.2E-03 **	4.2E-03 **
No polluted (Site 2)- Pristine forest (Site 3)	2.1E-01	3.5E-01	1.1E-03 **	4.2E-01	4.3E-02 *	5.8E-02	5.8E-02
*** Strongly significant (99.9%) ** Highly significant (99%) * Significant (95%) No significant difference							

3.3. Analysis of vegetation indices from Hyperion images

Most vegetation indices (23 of the 28) illustrated 99.9% significance difference between Site 1 (polluted) and Site 3 (pristine forest) which are the most dissimilar sites in terms of forest structure, plant species and conservation. 16 of 28 indices showed 99.9% significance differences between Site 2 (secondary non-polluted forest) and Site 3 (pristine forest) and just 11 vegetation indices registered 99.9% significance between Site 1 (polluted) and Site 2 (non-polluted forest). Of those 11 vegetation indices which were able to discriminate as strongly significant (99.9%) the difference between the two sampled secondary forests (Site 1 and Site 2), all of them corresponding to broad-band indices and narrow-band-greenness-chlorophyll-red-edge index groups. Lower and no-significance were found in indices grouped under other pigments and water indices.

3.4. Mapping vegetation stress

The eleven vegetation indices that strongly discriminated polluted and non-polluted secondary forests (strongly significant at 0.1% level of confidence) were selected as the more sensitive indices to detect the effects of petroleum pollution. Those indices are: SR, NDVI, GNDVI, SG, PSSRa, NDVI705, CTR2, LIC1, OSAVI, VOG1 and MTCI.

To ascertain the importance of each the 11 VI in the mapping of contamination a discriminant function analysis was undertaken which illustrates that three VI (the SG, NDVI and $NDVI_{705}$) explain 83% of the ability to separate between the 3 sites. Figure 2 remaps contamination based on these 3 VI only. By way of validation Figure 3 depicts those sites sampled in the field that have been correctly allocated as either contaminated or uncontaminated. This Figure also affords closer examination of the cause and effect of the hydrocarbon contamination in these forests.

ISBN: 978-85-17-00088-1

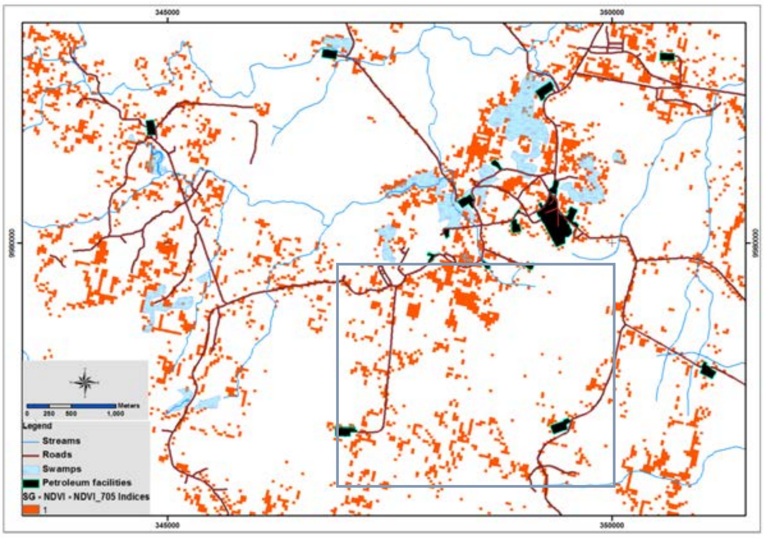

Figure 2. Areas identified as vegetation stress based on the SG, NDVI and NDVI_705 indices which together contribute to 83% of the site separability. The blue square is that highlighted in Figure 10.

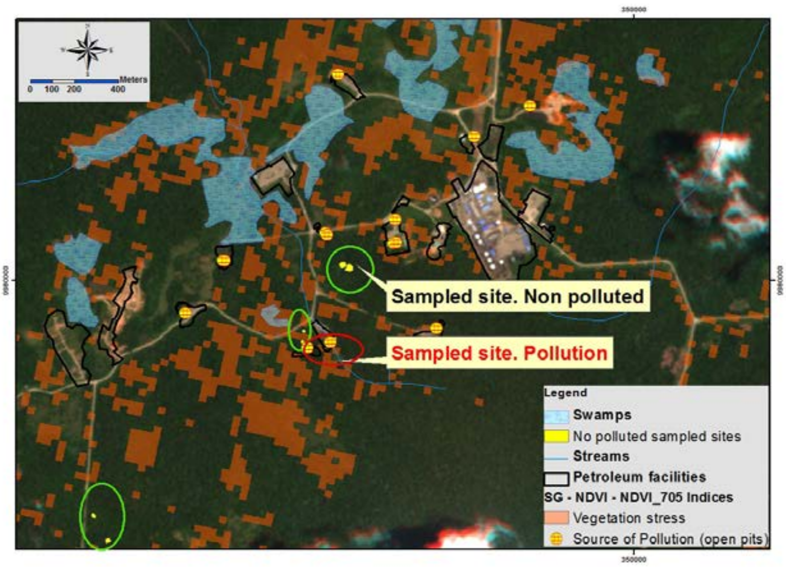

Figure 3. Areas detected as vegetation stress in petroleum productive area. Open pits identified as source of pollution and RAPID EYE images (background) have been provided by the Environmental Ministry of Ecuador (PRAS-program)

4. Discussion

4.2. Impact of petroleum contamination on leaf properties

Of the leaf biochemical and biophysical properties measured it was chlorophyll content and those associated with water content that exhibited significant differences between the polluted site and non-polluted sites. The low levels of chlorophyll content seen at site 1 indicate vegetation stress caused by a reduction of photosynthetic activity in vegetation exposed to petroleum contaminant. The C_{ab} content is responsive to a range of stresses on vegetation because of its direct role in the photosynthetic processes of light harvesting and initiation of electron transport (Zarco-Tejada et al. 2000). The higher values of water content (*Cw*) observed at the polluted site may be linked to the adaptation process of plants to close

ISBN: 978-85-17-00088-1

the stomata under stress conditions as strategy to reduce transpiration, which in turns reduce photosynthetic rate linked to the lower chlorophyll and thus total tree metabolism (Larcher 2003, Zweifel;Rigling & Dobbertin 2009). Other foliar properties related to water, those expressed on mass basis (% LWC and % LDMC) also differed and is due to the fact that as these parameters are not normalised by the leaf area, these differences can be explained by the high species diversity of the sample sites where leaves vary greatly in morphology, anatomy and physiology in response to their growing conditions (Tedersoo et al. 2010). Of these leaf variables, it is chlorophyll content that lends itself to be measured from space using a hyperspectral sensor, and since it is this that showed differences between the polluted and unpolluted sites, this suggests that by measuring this biochemical in vegetation compartments, detection of petroleum contamination across vast expanse of tropical forests is indeed possible.

4.3. Vegetation indices to detect the occurrence of petroleum pollution

As suggested by the field data it was those vegetation indices with sensitivity to photosynthetic pigments that were most useful in discriminating between the contaminated and non-contaminated sites. The Sum Green vegetation index (SG) clearly identified an increased reflectance signal in the visible spectral region of the area affected by petroleum pollution which confirms the sensitivity of Hyperion image to register reduced chlorophyll content levels in the polluted site. Also of use are the broad-band and narrow-band vegetation indices related to the traditional NDVI (SR, GNDVI, NDVI705), endorsing the conclusions of (Zhu et al. 2013).

The three indices of most use for mapping (explaining 83% of separability between the three sites), were the SG, NDVI and NDVI705, and are a mixture of both multispectral and hyperspectral vegetation indices. This particular selection of indices seems to be based on their ability to highlight lower levels of photosynthetic pigments, in particular chlorophyll (SG index) and dense vegetation with the high LAI (NDVI) characteristic of tropical forest environments. To employ these indices within a monitoring system to detect petroleum contamination is attractive, particularly given the imminent improvements in sensor technology, capability and the simplicity of using the spectra measured by these sensors. Indeed, new sensors coming on stream will offer superspectral and hyperspectral data which will improve spectral resolution at medium spatial resolution (i.e., Sentinel, EnMap, FLEX, HyspIRI, IMS-1 HYSI). In addition, high resolution sensors will improve their spectral resolution (e.g., WV 4) and the fast developed of miniaturized cameras on board of Unmanned Aerial Vehicle (UAVs). Moreover improving the temporal resolution (mainly through constellation approaches or targeting capture) mean that monitoring for new pollution impacts becomes a reality.

5. Conclusions

This study provides evidence of leaf biochemical alterations in the rainforest caused by petroleum pollution and demonstrates that these can be detected by spaceborne satellite remote sensing. The results indicate that tropical forests exposed to petroleum pollution show principally reduced levels of chlorophyll content, accompanied by higher levels of foliar water content. These alterations were detectable from space using the EO-1 Hyperion sensor by way of vegetation indices that are sensitive to detection changes of photosynthetic activity of the forest based on chlorophyll content and indices related to canopy density and vegetation vigour. This investigation has shown a potential for the use of imaging spectrometers for the identification and characterisation of hydrocarbon pollution or hydrocarbon seeps in dense tropical forests.

6. References

Arellano, P., Tansey, K., Balzter, H. & Boyd, D. 2016, "Field spectroscopy and radiative transfer modelling to assess impacts of petroleum pollution on biophysical and biochemical parameters of the Amazon rainforest.", *Journal of Environmental Earth Scences*, vol. Under consideration.

Asner, G.P., Nepstad, D., Cardinot, G. & Ray, D. 2004, "Drought stress and carbon uptake in an Amazon forest measured with spaceborne imaging spectroscopy", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 101, no. 16, pp. 6039-6044.

Bernal, A.M. 2011, "Power, powerlessness and petroleum: Indigenous environmental claims and the limits of transnational law", *New Political Science,* vol. 33, no. 2, pp. 143-167.

Davidson, E.A., De Araüjo, A.C., Artaxo, P., Balch, J.K., Brown, I.F., C. Bustamante, M.M., Coe, M.T., Defries, R.S., Keller, M., Longo, M., Munger, J.W., Schroeder, W., Soares-Filho, B.S., Souza Jr., C.M. & Wofsy, S.C. 2012, "The Amazon basin in transition", *Nature,* vol. 481, no. 7381, pp. 321-328.

Finer, M., Jenkins, C.N., Pimm, S.L., Keane, B. & Ross, C. 2008, "Oil and gas projects in the Western Amazon: Threats to wilderness, biodiversity, and indigenous peoples", *PLoS ONE,* vol. 3, no. 8.

Hoel, B.O. 1998, "Use of a hand-held chlorophyll meter in winter wheat: Evaluation of different measuring positions on the leaves", *Acta Agriculturae Scandinavica - Section B Soil and Plant Science,* vol. 48, no. 4, pp. 222-228.

Hurtig, A.K. & San-Sebastián, M. 2005, "Epidemiology vs epidemiology: The case of oil exploitation in the Amazon basin of Ecuador [5]", *International journal of epidemiology,* vol. 34, no. 5, pp. 1170-1172.

Larcher, W. 2003, *Physiological Plant Ecology,* Fourth edn, Spinger, Germany.

Malhi, Y., Roberts, J.T., Betts, R.A., Killeen, T.J., Li, W. & Nobre, C.A. 2008, "Climate change, deforestation, and the fate of the Amazon", *Science,* vol. 319, no. 5860, pp. 169-172.

Mutanga, O., Skidmore, A.K. & Prins, H.H.T. 2004, "Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features", *Remote Sensing of Environment,* vol. 89, no. 3, pp. 393-408.

Tedersoo, L., Sadam, A., Zambrano, M., Valencia, R. & Bahram, M. 2010, "Low diversity and high host preference of ectomycorrhizal fungi in Western Amazonia, a neotropical biodiversity hotspot", *ISME Journal,* vol. 4, no. 4, pp. 465-471.

Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H. & Noland, T.L. 2000, "Chlorophyll fluorescence effects on vegetation apparent reflectance: I. Leaf-level measurements and model simulation", *Remote Sensing of Environment,* vol. 74, no. 3, pp. 582-595.

Zhu, L., Zhao, X., Lai, L., Wang, J., Jiang, L., Ding, J., Liu, N., Yu, Y., Li, J., Xiao, N., Zheng, Y. & Rimmington, G.M. 2013, "Soil TPH Concentration Estimation Using Vegetation Indices in an Oil Polluted Area of Eastern China", *PLoS ONE,* vol. 8, no. 1.

Zweifel, R., Rigling, A. & Dobbertin, M. 2009, *Species-specific stomatal response of trees to drought - A link to vegetation dynamics?.*